IJCB
INTERNATIONAL JOURNAL OF CENTRAL BANKING
Volume 18, Number 4 October 2022

Volume 18, Number 4 October 2022

IJCB

INTERNATIONAL JOURNAL OF CENTRAL BANKING

IJCB International Journal of Central Banking

Volume 18, Number 4 October 2022

*The contents of this journal, together with additional materials provided by article authors, are available without charge at **www.ijcb.org**.*

The *International Journal of Central Banking* is published quarterly (ISSN: 1815-4654). Online access to the publication is available free of charge at **www.ijcb.org**. Changes to hard copy subscriptions and orders should be placed by e-mail at editor@ijcb.org.

Renewals, claims, address changes, and requests for permission to reprint material from this journal should be addressed to:

International Journal of Central Banking
Printing & Fulfillment K1-120
Federal Reserve Board
Washington, DC 20551

Phone: 202-452-3425
Fax: 202-728-5886
E-mail: editor@ijcb.org

ISSN: 1815-4654

How Has Monetary and Regulatory Policy Affected Trading Relationships in the U.S. Repo Market?*

Sriya Anbil and Zeynep Senyuz
Federal Reserve Board

We analyze the effects of changes in monetary and regulatory policy on trading dynamics in the U.S. triparty repo market. Using a confidential data set of transactions, we find that the Fed's reverse repo (RRP) facility led to a 16 percent reduction in cash lending by money market mutual funds (MMFs) eligible to transact with the Fed. We show that the RRP facility increased the bargaining power of MMFs on days when their borrowers, non-U.S. dealers, increased their window-dressing activity due to Basel III capital reforms. For those dealers reliant on eligible MMF funding, window dressing became more expensive, but the average rates they paid on other days remained stable because of anchoring by the facility. We also show that the RRP facility influenced the way MMFs managed their balance sheets, and directed them towards safer investments.

JEL Codes: C32, E43, E52.

*An earlier version of this paper was circulated under the title "The Regulatory and Monetary Policy Nexus in the Repo Market." We thank Alyssa Anderson, Jeff Huther, Elizabeth Klee, Christoffer Koch, Marco Macchiavelli, Xiumin Martin, Benjamin Munyan, Heather Tookes, Emre Yoldas, and two unanimous referees for useful suggestions. We also thank participants at the 2018 IBEFA ASSA Conference, 2017 JFI Conference, 2017 EFA Meeting, 2017 Congress of the EEA, 2017 World Finance Conference, and seminar participants at the Tepper School of Business and at the Federal Reserve Board for helpful discussions. Grant Carney, Hampton Dennis, and Margaret Sabelhaus provided excellent research assistance. The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Federal Reserve, or other members of its staff. Author e-mails: sriya.l.anbil@frb.gov and zeynep.senyuz@frb.gov.

1. Introduction

Financial intermediaries rely on the repo market to finance the securities on their balance sheets. Persistent inability to fund these securities would lead to stress in money markets, as witnessed during the global financial crisis (GFC) of 2007–09. The repo market, which is a crucial source of short-term funding for many financial institutions, has been perceived as a potential source of instability since the GFC (see, for example, Adrian and Shin 2011, Copeland, Martin, and Walker 2014, and Gorton and Metrick 2012).

As part of its policy response to the GFC and the following Great Recession, the Federal Reserve (Fed) reduced the federal funds rate to its effective lower bound and conducted large-scale asset purchases (LSAPs) to provide accommodative financial conditions and promote economic recovery. The expansion of reserves due to LSAPs prompted changes in the way the Fed implements its monetary policy. In October 2008, the Fed starting paying interest on reserves, and this administered rate became its primary policy tool. In September 2013, the overnight reverse repurchase (RRP) facility was introduced to enhance rate control in an environment of abundant reserves. By offering a secured rate through the RRP facility where many counterparties, including money market mutual funds (MMFs), could lend to the Fed, the Fed effectively set a soft floor on repo rates (see Klee, Senyuz, and Yoldas 2019 for an analysis of how the RRP facility affected overnight funding rates).

During this time, the regulatory environment also evolved substantially. Prior to the GFC, broker-dealers (dealers), who are the main borrowers in the repo market, used to operate with substantial leverage, as they were not subject to strict regulatory limits. While capital requirements were much less restrictive than they are today, they were not uniform among dealers from different jurisdictions. For example, non-U.S. dealers did not have to meet a certain leverage ratio, unlike U.S. dealers. In the aftermath of the GFC, the international regulatory authority at Basel implemented a series of major financial reforms prompting banks and other financial intermediaries to reevaluate their risk-management practices. Among the new regulations, the Basel III capital reforms introduced a formal leverage ratio, requiring banks to hold Tier 1 capital equivalent to at least 3 percent of their leverage exposure calculated using their on- and off-balance-sheet assets. These requirements directly affected the incentives and trading strategies of dealers, who typically borrow cash in the repo market to finance the securities on their balance sheets.

The regulatory requirements for dealers operating in different jurisdictions created different incentives for their activity in money markets. Munyan (2015) shows that pre-Basel III, non-U.S. dealers were already reducing their repo activity on financial reporting days, as their leverage ratios were based on quarter-end snapshots of their balance sheets. Their withdrawal from the market on financial reporting days is one form of the so-called window-dressing strategy that dates back to the 1800s.[1] The difference in regional implementation of the Basel III capital reforms further incentivized foreign dealers to engage in window dressing while it did not affect the U.S. dealers, which continued to report leverage ratios based on their daily activity.

In this paper, we analyze how these changes in U.S. monetary policy implementation and Basel III capital reforms affected the activity of two major repo market players: (i) MMFs—the primary cash lenders, and (ii) dealers—the primary cash borrowers. We use a confidential data set of repo transactions at the intradaily level in the triparty market, which is a major repo segment where a third party provides custodial services for operational efficiency. Our sample covers the period from January 2013 until August 2016, during when the RRP facility was introduced and the Basel III leverage ratio implementation took place.

On the supply side of the repo market, we show that the introduction of the overnight RRP facility led to a reduction of lending by MMFs eligible to transact with the Fed. These lenders compared the

[1]Financial institutions have historically modified the composition or size of their balance sheets ahead of their quarter-end filings to report more favorable ratios to their regulators or to the public. This phenomenon has been well-documented in the literature: see, for example, Agarwal, Gay, and Ling (2014), Allen and Saunders (1992), Lakonishok et al. (1991), and Sias and Starks (1997), among others. Prior to the GFC, an important motivation for window dressing was to improve the profitability measures mainly for public reporting. However, financial intermediaries started focusing more on the capital and liquidity measures in the more stringent post-crisis regulatory environment.

return for their cash investments in the private market and the RRP offer rate, and typically invested at the facility when the offer rate was more favorable relative to other options. On quarter-ends, when non-U.S. dealers pulled back from the market to reduce the size of their balance sheets for financial reporting, the facility provided a backstop to eligible MMFs to place their excess cash. We show that cash supply in the repo market provided by eligible MMFs declined by 16 percent, on average, after the inception of the facility. On the demand side, we show that after the Basel III leverage ratio implementation, window-dressing activity by European dealers increased by about 19 percent. With this increase in window dressing, total reduction in repo borrowing on financial reporting days reached 30 percent.

The inception of the RRP facility and the intensified window-dressing activity by European dealers constituted a supply shock and a demand shock in the repo market, respectively. Putting these shocks together, we show that window dressing became more expensive for dealers reliant on eligible MMF funding in comparison with dealers that were less reliant. Further, reliant dealers were unable to reduce their borrowing as much on financial reporting days. However, on other days, dealers reliant on eligible MMF funding ended up paying, on average, similar rates to their MMF lenders in comparison with dealers that were less reliant, because of the anchoring effect of the RRP facility. Our results highlight the importance of trading relationships and how these relationships affected bargaining power dynamics in the triparty repo market.

Finally, we examine the implications of the RRP facility for balance sheet management of MMFs. Using Securities and Exchange Commission (SEC) N-MFP filings, we find evidence that the inception of the RRP facility made eligible MMFs safer, as they shifted the composition of their balance sheets towards Treasury repo and away from repo backed by other collateral, commercial paper (CP), asset-backed commercial paper (ABCP), or corporate debt. Further, we find some evidence that ineligible MMFs shortened the duration of their portfolios after the inception of the facility, likely reflecting their strategy to rely more on shorter-duration investments, as they did not have the RRP facility available as a backstop.

Our work contributes to the growing literature on repo market dynamics. Adrian and Shin (2011) and Macchiavelli and Zhou

(2019) show that dealers rely on repo for short-term funding needs and adjust the size of their balance sheets mainly through their activity in this market. Focusing on different market segments, Copeland, Martin, and Walker (2014), Gorton and Metrick (2012), Krishnamurthy, Nagel, and Orlov (2014), and Martin, Skeie, and von Thadden (2014) analyze repo market dynamics during the GFC. Regarding the effects of the RRP facility in money markets, Anderson and Kandrac (2018) argue that eligible funds were able to command higher rates during the testing phase of the facility. Using transaction-level data for all MMFs investing at the facility over a much longer horizon, we show that the RRP facility did not lead to differential pricing. Our finding is consistent with Klee, Senyuz, and Yoldas (2019), who show that the facility rate anchored rates in the repo market.

Our work also contributes to the vast literature on the importance of trading relationships in different financial market segments: Han and Nikolaou (2016) for the repo market; Afonso, Kovner, and Schoar (2014) and Cocco, Gomes, and Martins (2009) for the interbank market; Bharath et al. (2011) and Dass and Massa (2011) for bank-firm relationships; and Chernenko and Sunderam (2014) for MMF lending, among others.

Another strand of literature that our paper is related to focuses on the effects of the network structure in over-the-counter (OTC) markets on prices. Hendershott et al. (2016), Li and Schürhoff (2019), Maggio et al. (2019), Maggio, Kermani, and Song (2017), and Neklyudov (2019) all show that dealers that are less interconnected in the OTC network are unable to charge their clients more. We document the increase in bargaining power of eligible MMFs after the RRP facility was introduced, as their dependence on dealer borrowing declined, and show that strong relationships prevented some dealers from window dressing as effectively. Our results are also consistent with Babus and Hu (2017) and Gofman (2017), who show that if a network becomes less interconnected, it also becomes more stable. We provide some evidence that when MMFs became less dependent on dealers, their investments shifted towards safer investments. Finally, our results on the effects of increasing bargaining power of MMFs also provides empirical evidence for the theoretical results shown in Duffie, Gârleanu, and Pedersen (2005, 2007) in the search-and-bargaining literature.

The rest of the paper proceeds as follows. The next section provides background information on the repo market and the Fed's RRP facility. Section 3 describes the confidential data set of repo trading and the control variables for the empirical analysis. The effects of the RRP facility on MMF lending are documented in Section 4. Section 5 turns to the demand side of the repo market and analyzes the effects of Basel III capital regulations on dealer behavior. Section 6 investigates how the changes in the Fed's monetary policy framework and regulatory framework affected bargaining power in trading relationships. Section 7 examines the implications of the RRP facility for MMFs' balance sheet management. Section 8 concludes.

2. The Repo Market and the Policy Environment

2.1 Repo Market Mechanics

A repo transaction facilitates the sale and future repurchase of a security that serves as collateral between the two parties: (i) the borrower who owns a security and seeks cash and (ii) the lender who receives the security as collateral when lending the cash. The cash borrower sells securities to the cash lender with the agreement to repurchase them at the maturity date. On the maturity date, the borrower returns the cash with interest to the lender and the collateral is returned from the lender to the borrower.[2] In the event of a default, the cash lender can sell the cash borrower's collateral to recover the loan amount.

The repo market during our time period can be divided into two broad segments: the bilateral market and the triparty market, where all types of securities are used as collateral. In the bilateral market, lenders and borrowers interact directly to negotiate the terms of the trade, settle the trade, and organize all back-office support themselves. In the triparty market, lenders and borrowers use the services of a third party to act as a custodian, settle the trade, and provide all back-office support. In the triparty repo platform for our sample

[2]From the cash borrower's perspective, this transaction is called a repo, and from the cash lender's perspective, it is called a reverse repo. Fed transactions in the repo market are defined as the opposite of market convention. If executed by the Fed, a cash-out/securities-in transaction is called a "repo" and a cash-in/securities-out transaction is called a "reverse repo."

period, two clearing banks—Bank of New York Mellon (BNYM) and J.P. Morgan (JPM)—provided these back-office services.

Our analysis focuses on the triparty market for which we have daily transaction data. The overall daily triparty repo volume, which includes trades of all maturities and collateral types, was around $2 trillion during our sample period. More than $1.5 trillion of this volume involved general collateral (GC).[3] GC repo transactions are backed by securities that meet the predetermined eligibility criteria to be accepted as collateral. The cash lender does not know the specifics of the securities collateralizing the transaction prior to settlement, as opposed to a specific collateral repo where the cash lender typically tries to obtain a particular security. While other collateral types are also traded in the triparty segment, we focus on GC repo transactions in the triparty market where the collateral include U.S. Treasury securities, agency debt, and agency mortgage-backed securities (MBS). GC repo is the largest, safest, most liquid, and primarily overnight.

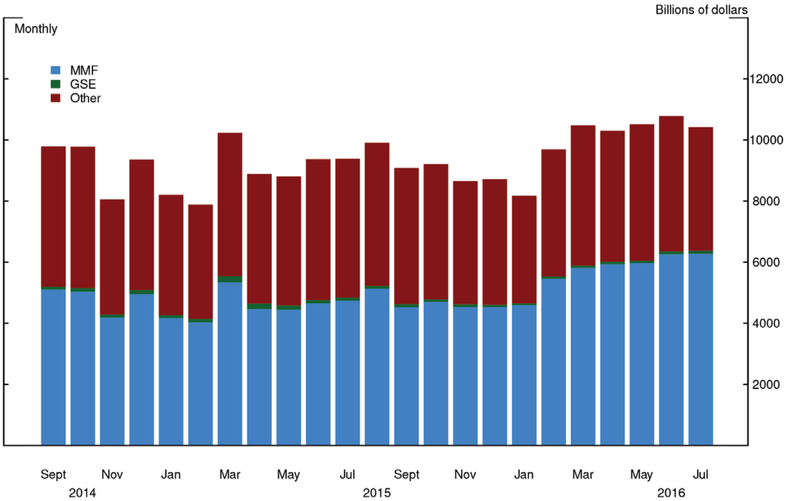
Figure 1 shows the monthly totals of overnight GC triparty transaction volumes by lender types. As shown by the gray shaded bars, MMFs account for about 60 percent of overnight cash lending. The repo market is also an important investment platform for government-sponsored enterprises (GSEs), which constitute about 5 percent of the total monthly volume. All other types of cash lenders account for the remaining 35 percent of the volume. These lenders include mutual funds, asset managers, and hedge funds.

On the demand side of the triparty market, the main cash borrowers are dealers. Figure 2 shows the monthly totals of overnight GC triparty transaction volumes by U.S. and foreign dealers (includes European, U.K., Canadian, and Japanese dealers). Borrowing activity by the two dealers groups have been following parallel trends suggesting that, on average, they have been responding to broad market factors similarly.

The triparty GC repo market also includes the overnight RRP operations by the Fed and the general collateral finance (GCF) segment, which is a blind-brokered, interdealer repo platform that provides funding for dealers that may not have sustainable access to

[3] See Copeland, Martin, and Walker (2014) for a more detailed discussion of the mechanics of the triparty repo market.

Figure 1. Triparty GC Repo Lending by Lender Type



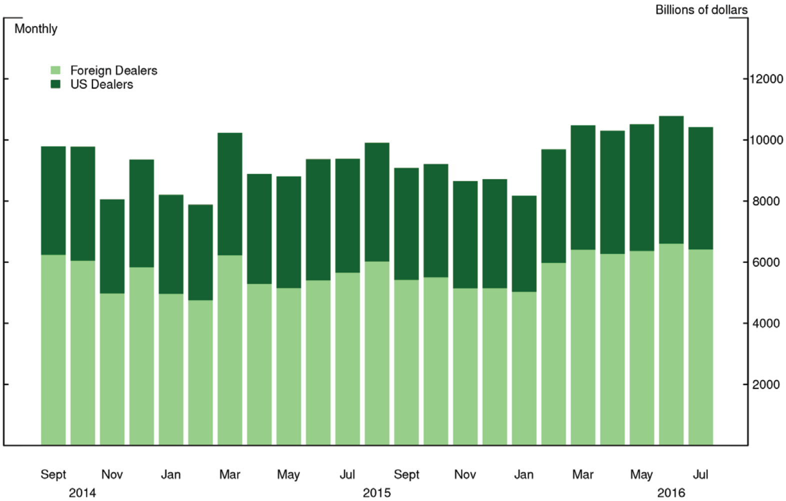
Note: This figure displays monthly totals of overnight triparty lending against GC by lender type. Data are overnight daily triparty repo transactions, obtained from FRBNY, and aggregated monthly. The sample shown here is from August 2014 to August 2016. "MMF" include money market mutual funds. "GSE" are government-sponsored enterprises such as Fannie Mae and Freddie Mac. "Other" includes asset managers, hedge funds, and mutual funds.

cash in the broader triparty market. Figure 3 shows an organizational diagram of the triparty GC repo market with the approximate shares of three segments from 2014 to 2016: (i) triparty market where BNYM and JPM serve as the custodian banks (70 percent of volume), (ii) interdealer GCF market (15 percent of volume), and (iii) transactions with the Fed via the RRP facility (15 percent of volume). In our triparty repo analysis, we exclude the interdealer GCF market as well as transactions with the Fed via the RRP facility.

2.2 Fed's RRP Facility

During the GFC and its aftermath, the Fed increased the size of its balance sheet through several liquidity facilities and LSAPs.

Figure 2. Triparty Repo Borrowing by Dealers



Note: This figure displays monthly totals of overnight triparty lending against GC by lender type. Data are overnight daily triparty repo transactions, obtained from FRBNY, and aggregated monthly. The sample shown here is from August 2014 to August 2016. Foreign dealers include European, U.K., Canadian, and Japanese dealers.

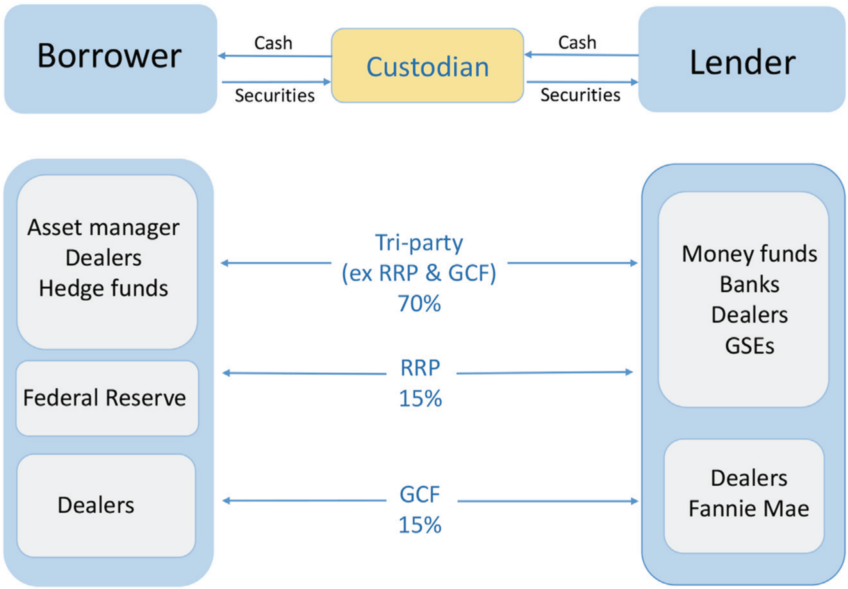
As a result, reserves in the financial system have reached unprecedented levels, resulting in a change in the Fed's monetary policy implementation.[4]

In October 2008, the Fed started paying interest on excess reserves (IOER) to banks that have accounts at the Fed, and the IOER became the primary tool of the new policy framework. However, the IOER could not set an effective floor for the federal funds rate because of the fragmented market structure. In September 2013, the Fed introduced the overnight RRP facility as a supplementary tool of its new policy framework to enhance monetary control.

The Fed has been offering overnight RRPs on a daily basis at a preannounced rate since September 2013. Through this facility, the Fed borrows cash from eligible counterparties in exchange for Treasury securities in its portfolio. These transactions take place with

[4]See Ihrig, Weinbach, and Meade (2015) for details of the Fed's monetary policy implementation framework during and after the crisis.

Figure 3. Triparty Repo Market Mapping



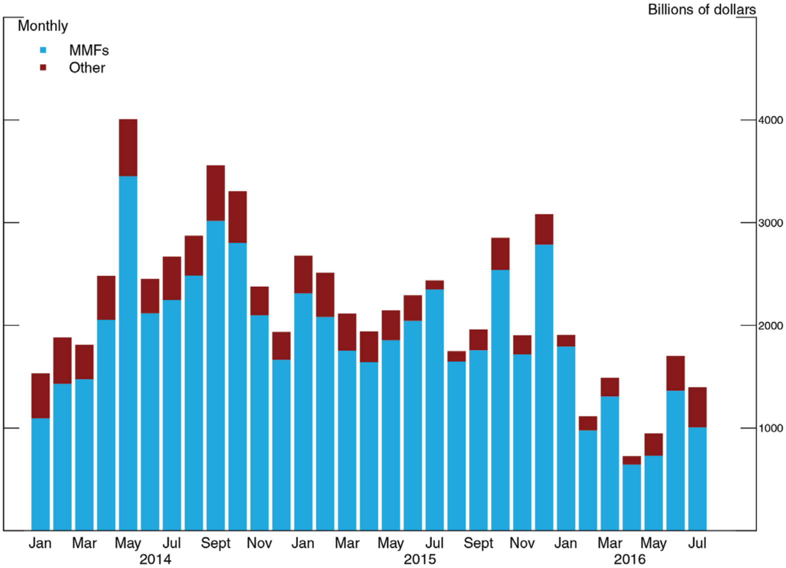
Note: The triparty GC repo platform consists of repo transactions for which BNYM and JPM are the custodian banks, interdealer transactions that take place in the GCF segment, and transactions with the Fed via the RRP. Shares of each segment shown above reflect approximate averages over the period from 2014 to 2016. These shares fluctuate on certain calendar days, such as quarter-ends. Since mid-2016, GCF volume has dropped significantly with the rise of the Fixed Income Clearing Corporation delivery-versus-payment segment. The data are from FRBNY and Bloomberg.

the agreement to repurchase the same security at a specified price at a specific time in the future. Overnight RRPs are offered to a broad set of financial institutions including nonbank cash lenders in the repo market, such as MMFs.[5]

Figure 4 shows the composition of participants coming to the facility from January 2014 to August 2016. MMFs were the primary participants, accounting for the majority of takeup at the facility.

[5]There are currently more than 150 RRP counterparties including MMFs, GSEs, primary dealers, and other banks. A list of RRP counterparties and information on eligibility requirements can be found at https://www.newyorkfed.org/markets/rrp_counterparties.html.

Figure 4. Overnight RRP Participation

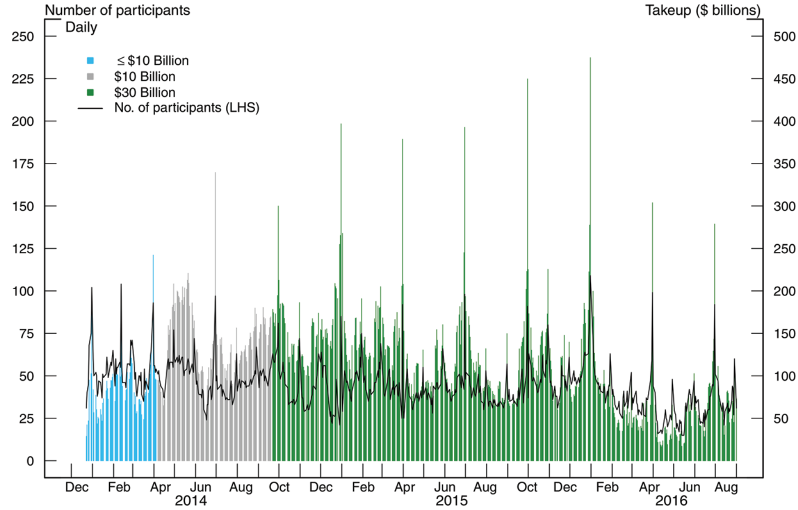


Note: This figure displays monthly totals of RRP participation by counterparty type. Data on overnight RRP takeup by counterparty type are available from FRBNY. The monthly sample shown here is from January 2014 to August 2016.

MMFs are the biggest cash lenders in the repo market and view the overnight RRP facility as a low-return and low-risk investment compared with lending to dealers.[6] The RRP facility provided a convenient alternative investment vehicle for MMFs who compared the facility's offering rate with other rates in the market and determined whether to participate in the overnight RRP operation offered each day. As MMFs would be unwilling to lend in the market at any rate below the RRP rate, the offering rate at the facility helped establish a floor for overnight funding rates, as shown by Klee, Senyuz, and Yoldas (2019). Most importantly, when dealers withdrew from overnight funding markets on quarter-ends for financial reporting purposes, the RRP facility provided a backstop for MMFs.

[6] Among other counterparties, GSEs account for most of the remaining takeup, with Fannie Mae and Freddie Mac increasing their facility usage on days ahead of their principal and interest payment dates.

Figure 5. RRP Facility Usage



Note: This figure displays daily RRP takeup from both overnight and term RRP operations. Data on overnight and term RRP takeup are available from FRBNY. The daily sample shown here is from December 23, 2013 to August 1, 2016.

Figure 5 shows the daily RRP takeup after the individual bid size increased to $3 billion on December 23, 2013. Takeup at the facility was very low in the first few months of testing, during which the facility parameters were modified frequently. Following gradual increases in the bid size over a few months, facility takeup increased. For our sample from January 2013 to August 2016, average daily volume was around $100 billion.[7] Seasonal spikes correspond to quarter-ends, when total RRP takeup hit record levels as MMFs shifted their lending to the Fed due to reduced demand for repo financing.

[7]Investment capacity at the facility proved to be an important factor affecting repo rates, especially on quarter-ends. Before the September 2014 quarter-end, the Fed introduced an overall cap of $300 billion, which led to a sharp drop in repo rates as cash lenders scrambled for alternative investments. A series of term RRPs were conducted ahead of quarter-ends to provide extra capacity to RRP counterparties until late 2015. The facility cap was lifted in December 2015 and there have been no term RRP operations over quarter-ends since then. Throughout the paper RRP refers to the sum of overnight and term operations.

Further details on the design of the overnight RRP facility can be found in Frost et al. (2015).

3. Data

3.1 Triparty Repo Transaction-Level Data Set

Our confidential data set of daily triparty repo transactions of overnight maturity backed by GC which is available from January 2, 2013, to August 1, 2016, allows us to examine the implications of the new policy environment on the repo market.[8] These data are reported by BNYM to the Federal Reserve Bank of New York (FRBNY).[9] We examine repo transactions backed by Treasury securities, agency debt, and agency MBS, and of overnight maturity.

To fully analyze MMF-dealer trading dynamics, it is necessary to identify those funds that are eligible to lend to the Fed via the RRP facility among all MMF lenders in the data set.[10] However, identification of eligible lenders is challenging, as lender names are not uniform throughout the data set, and one must identify funds at the fund rather than complex level to determine the eligible funds. In our data set, a single fund may be listed in multiple ways. For example, Money Fund A may appear as "mmf A," "moneyf A," "moneyfunda," etc. To collapse the multiple ways a money fund may appear in the data set to the appropriate fund, we use the monthly SEC Form N-MFP filings for MMFs, which provide monthly snapshots of MMF fund-level repo transactions including the volume, collateral, price, and maturity of the trade. We are able to match the various lender strings in our transaction-level data set to a single

---

[8] We end our sample at August 1, 2016 to avoid data issues of MMF fund closures from the compliance of SEC MMF Reform in mid-October 2016.

[9] During this time period, BNYM was a custodian for about 80 percent of trades in the triparty repo market. BNYM started providing daily transaction-level data after August 22, 2014. Prior to this date, BNYM used to provide a Tuesday snapshot of all outstanding open trades, for which it acts as custodian. For this period we merged Tuesday snapshots with the daily transaction data and use day fixed effects to account for the cross-sectional snapshots in our analysis. We sum all transactions for a given lender-borrower pair at the daily level.

[10] Since the list of RRP-eligible counterparties has changed over time, we use a dynamic list of eligible funds. See https://www.newyorkfed.org/markets/rrp_counterparties.html for the up-to-date list of eligible counterparties.

Table 1. Summary Statistics for Triparty Transaction-Level Data

Items	No. of Transactions
Eligible MMFs	71,322 (41%)
Ineligible MMFs	100,636 (59%)
Foreign Dealers	104,358 (61%)
U.S. Dealers	67,600 (39%)
Transactions	171,958
Treasury Securities	87,704 (51%)
Agency Securities	84,254 (49%)

Note: This table displays summary statistics about the triparty transaction-level data set. Data are obtained from FRBNY. The data are reported by Bank of New York Mellon to FRBNY, and available from January 2, 2013, to August 1, 2016.

fund by matching the volume, collateral, maturity, and price of the repo transaction on these month-end N-MFP filing days.[11]

We are able to match 288 MMFs and 25 dealers. Of the approximately 60 percent of transaction volume that are likely MMF lenders, we can identify the MMF lenders and dealers of 82 percent of these transactions (see Figure 1). Consequently, our data correspond to about 50 percent of the total volume in the triparty repo market. We identify a total of 1,101 unique MMF-dealer pairs. As reported in Table 1, eligible and ineligible MMFs account for 41 percent and 59 percent of these transactions, respectively. Foreign dealers (European, U.K., Japanese, and Canadian) participate in 61 percent of transactions (16 dealers) while the U.S. dealers (9 dealers) account for the remaining volume. Our data set consists of trades backed by Treasury (51 percent) and agency securities (49 percent).

[11] We merge repo transactions with N-MFP filings on October 31, 2014; November 30, 2014; and December 31, 2014. A fund is considered matched if we match overnight repo volumes within a 1 percent error given collateral type, maturity, and price. By using only three N-MFP filing dates, we do assume some static MMF lending in our data set. Because the string match between the N-MFP filings and the repo transaction-level data set is quite intensive, and we match 82 percent of transaction volume on these dates, we are comfortable with our methodology.

3.2 Control Variables

We construct several control variables to establish robustness of our empirical results, where we refer to dealer i on day t. First, using the quarterly Consolidated Report of Condition and Income Reports from the Federal Financial Institutions Examination Council (FFIEC), or Call Reports, we construct two balance sheet measures for each dealer borrower i at quarter q: (i) dealer's total assets ($TotalAssets_{i,q}$), (ii) dealer's short-term funding dependence ($STFD_{i,q}$), calculated as follows.[12]

$$STFD_{i,q} = \frac{(ST\ Noncore\ Funding)_{i,q} - (ST\ Investments)_{i,q}}{(LT\ Assets)_{i,q}}. \tag{1}$$

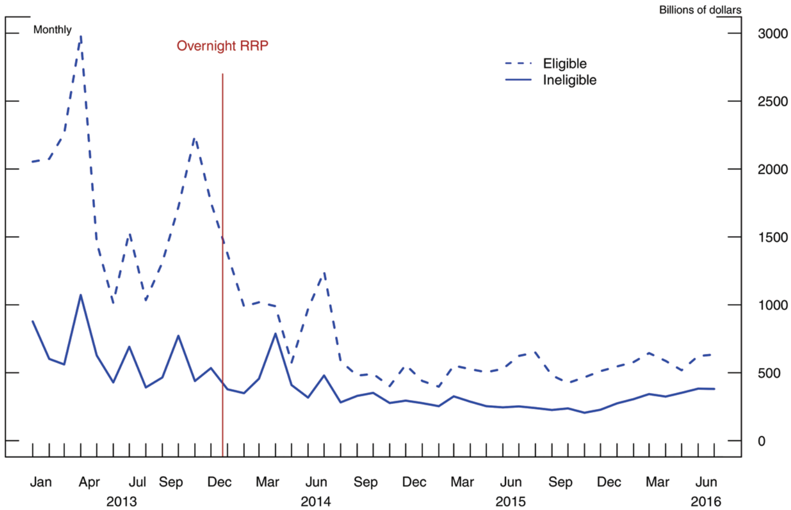
Second, we construct control variables for the lenders (MMFs) using the SEC N-MFP filings to calculate MMF complex total assets under management (AUM), Treasury repo investments, and the amount of Treasury securities held. N-MFP filings are filed monthly by each MMF fund. These control variables capture MMFs' dependence on lending in repo markets and their preference for lending in Treasury repo versus holding Treasury securities outright, which are close substitutes.

4. The RRP Facility and Cash Supply

We first analyze how the introduction of the overnight RRP facility affected cash lending in the repo market by eligible versus ineligible MMFs. When the Fed started test operations at the facility, it released a list of eligible counterparties. Those MMFs that are eligible to participate to lend to the Fed via the overnight RRP compared

[12] STFD was developed by bank supervisors as a measure of banks' short-term funding dependence. See pages 3–6 of https://www.federalreserve.gov/boarddocs/supmanual/bhcpr/UsersGuide13/0313.pdf for specific definitions of the variables. Total assets is item RCFD2170 on FFIEC 002 or FFIEC 031; or item RCON2170 on FFIEC 041 or FFIEC 051. Call Report data are available for 19 of the 25 dealers in the triparty repo market. The remaining participants do not have commercial bank operations in the United States and, therefore, do not need to report regulatory ratios to the FFIEC.

Figure 6. MMF Lending



Note: This figure displays weekly totals of MMF lending in overnight triparty repo backed by GC. Data are daily overnight triparty repo transactions backed by GC, obtained from FRBNY, and summed weekly. The weekly sample is from January 1, 2013, to August 1, 2016. We estimate total weekly lending before August 22, 2014 (when daily data are available) from Tuesday snapshots of triparty transaction data.

the offering rate at the facility with private market rates and determined whether to participate in the RRP operation offered each day. In addition to the spread between the RRP offer rate and market rates, calendar factors also affected facility usage. When dealers, the major repo borrowers, contracted their balance sheets on financial reporting days, eligible MMFs could invest their surplus cash at the RRP facility, while ineligible MMFs were forced to find other investment opportunities in the private market.

Figure 6 shows weekly lending by eligible and ineligible MMFs from January 2, 2013 to August 1, 2016. Both series are trending down during the first half of the sample, likely reflecting factors such as the low rate environment, the implementation of more conservative risk measures, and the third round of LSAPs by the Fed which reduced the supply of Treasury collateral in the market. To quantify the MMF response to the RRP, we first test for unit roots in the

MMF lending series using Elliott, Rothenberg, and Stock (1996) and Ng and Perron (2001) and tests that are powerful against persistent alternatives. We fail to reject the null of unit root due to the presence of breaks in both series.[13] After confirming that the series do not exhibit unit-root behavior once the breaks are accounted for, we regress them on their respective break dates and retrieve the residuals, which are stationary. We then estimate the following regression for weekly MMF repo lending, RL_t:

$$RL_t = \beta_0 + \beta_1 RL_{t-1} + \psi_1 RRP1 + \psi_2 RRP2 + \epsilon_t, \qquad (2)$$

where RRP1 takes the value 1 for the weeks from September 23 to December 23, 2013 and 0 otherwise, to represent the initial testing period of the facility. During this time, takeup was very low amid small bid limits which were then increased gradually. The individual bid size reached \$5 billion on December 23, 2013, and takeup increased to levels consistent with the sample average. Our second indicator variable, RRP2, marks the beginning of the period during which the facility started to be perceived as a viable investment option by market participants. We also account for autocorrelation by including the first lag of repo lending.

As shown in Table 2, we do not find a significant response of MMFs to the inception date of the facility as captured by RRP1, although we find evidence of a shift in eligible MMF in response to RRP2. Once the RRP facility became a significant investment option with increased individual caps, eligible MMF lending declined by an average of 16 percent. While we find a significant effect of RRP facility on eligible MMF lending, we do not find any effect on ineligible MMF lending. These results are unlikely to be driven by dealer demand since overall repo borrowing remained steady during our sample period.

To quantify the extent of substitution between eligible MMF investments in the private repo market and at the RRP facility, we

[13]Perron (1989) points out that conventional unit-root tests are biased towards a false unit-root null when the data are stationary and include a structural break. See Hansen (2001) for a comprehensive review of the literature on structural breaks. We find evidence of breaks in both series using the Quandt-Andrews unknown breakpoint test. We find the break dates of 2014:W28 and 2014:W17, for eligible and ineligible lending, respectively. For brevity, we do not report the unit root and the structural break test statistics.

Table 2. Effects of RRP on Repo Lending to Dealers

	Eligible MMF (1)	Ineligible MMF (2)
RRP1	0.03	–0.08
	(0.25)	(–0.8)
RRP2	–0.16***	–0.10
	(–2.32)	(–1.56)
Repo Lending (–1)	0.45***	0.30***
	(6.31)	(4.01)
Observations	180	180
R^2	0.27	0.11

Note: This table presents the results of time-series regressions for the log of weekly repo lending by MMFs to dealers. The weekly sample runs from January 2, 2013, to August 1, 2016, and is obtained from FRBNY. Dealers comprise all foreign and U.S. dealers in our sample. RRP1 takes the value 1 from September 23, 2013, to December 23, 2013 and 0 otherwise. RRP2 is equal to 0 until December 23, 2013, and 1 thereafter. ***, **, and * indicate statistical significance at the 1 percent, 5 percent, and 10 percent level, respectively. t-ratios are reported in parentheses and are calculated using robust standard errors.

estimate a time-series regression for the log of private repo lending by MMFs, shown in Equation (3), where we have the first lag of RRP facility takeup as an explanatory variable along with the first lag of the dependent variable. This specification is estimated using daily data from August 22, 2014 to August 1, 2016.[14]

$$RL_t = \beta_0 + \beta_1 RL_{t-1} + \psi_1 RRP_{t-1} Takeup + \epsilon_t \tag{3}$$

Table 3 shows the results of this estimation. We find that while overall repo lending by eligible MMFs declined by 16 percent, about 9 percent of this decline was replaced by investments in the RRP facility. On average, a one-standard-deviation increase in RRP

[14]We focus on the sample after daily data collection has started on August 22, 2014, to avoid the break in the data due to changes in data collection. Also, in this part of the sample, the RRP facility was well established as an alternative investment option for cash lenders in the repo market with sufficient counterparty limits, which helps us get a clean read on the extent of substitution between investing in the repo market and in the RRP facility.

Table 3. Effect of RRP on Repo Lending to Dealers

	Eligible MMF (1)
RRP Takeup (–1)	–0.09***
	(–3.88)
Repo Lending (–1)	0.34***
	(2.86)
Observations	477
R^2	0.0923

Note: This table presents the results of a time-series regression for the log of daily repo lending by MMFs to dealers. The daily sample runs from August 22, 2014, to August 1, 2016, and is obtained from FRBNY. Dealers comprise all foreign and U.S. dealers in our sample. RRP Takeup(–1) is the log of takeup by eligible MMFs the day prior. Repo Lending(–1) is the log of volume lent by eligible MMFs to dealers the day prior. ***, **, and * indicate statistical significance at the 1 percent, 5 percent, and 10 percent level, respectively. t-ratios are reported in parentheses and are calculated using robust standard errors

takeup is associated with a 9 percent decline in repo lending by eligible MMFs, showing strong evidence of substitution between repo lending and investing at the RRP facility.

5. Basel III Capital Regulations and Cash Demand

In this section, we turn to the demand side of the triparty repo market and analyze how Basel III affected dealer behavior on financial reporting days. Dealers in the triparty repo market have typically exhibited some form of window-dressing behavior. Adrian and Shin (2010, 2011) show that dealers adjust the size of their balance sheets mainly through short-term repo borrowing. Prior to the GFC, dealers used to operate with substantial leverage, as they were not subject to binding regulatory limits. When regulators responded to the GFC with requirements of higher-quality assets and lower leverage, dealers were prompted to reevaluate their risk-management practices and adjust their balance sheet management. As a result, dealer risk-taking has moderated since the GFC (see Adrian et al. 2013).

Basel III capital reforms formally introduced a leverage ratio, requiring banks to hold Tier 1 capital equal to 3 percent of an exposure measure which includes on-balance-sheet assets and certain off-balance-sheet items, including repo transactions. The calculation of the leverage ratio depends on banks' jurisdictions. Although the Basel III capital reforms were determined by an international body, each country could decide how they choose to implement their own version. In the United States, it is implemented as the supplementary leverage ratio (SLR), which is calculated on a daily basis. For European banks, the leverage ratio was computed as an average of the three month-end values over the quarter until October 2014 when the rule was amended to require only quarter-end reporting. U.K. dealers have been reporting their leverage ratios based on quarter-end snapshots of their balance sheet until they switched to reporting based on daily averages in January 2016. For other foreign banks, such as Canadian and Japanese banks, their leverage ratio was calculated as a quarter-end snapshot. Therefore, European and most other foreign dealers have been subject to less stringent implementation of the Basel III leverage ratio than to U.S. dealers.

The difference in regional implementation of the Basel III leverage ratio created different incentives for U.S. dealers and other dealers for financial reporting purposes. If the leverage ratio is calculated on a month- or quarter-end basis, then banks are likely to contract the size of their balance sheets on these dates and expand it on other days. While foreign banks in most jurisdictions are incentivized to engage in window dressing because of the less stringent leverage ratio implementation, U.S. dealers which report leverage ratios based on daily values do not have any reason to do so.[15] All else equal, in the absence of this difference in regional implementation of the leverage ratio, we would not expect repo trading behavior to be different between U.S. and non-U.S. dealers on reporting days.

Appendix A describes our empirical model that quantifies how much European dealers pulled back their borrowing in triparty repo

[15]Basel III also introduced two liquidity measures: the liquidity coverage ratio (LCR) and the net stable funding ratio (NSFR). The LCR required banks to hold high-quality liquid assets sufficient to meet a 30-day liquidity stress scenario, and the NSFR complemented it by promoting liquidity buffers over a longer horizon. Although LCR requirements may have affected repo activity for collateral other than Treasury securities, there have been no implications for repo backed by Treasury collateral. Moreover, all banks are required to be compliant with LCR on a daily basis.

on regulatory reporting days, specifically quarter-end days, after the implementation of Basel III. Table A.1 in Appendix A reports the estimation results from July 1, 2008, to August 1, 2016, for European and U.S. dealers in columns 1 and 2, respectively. European dealers have already been reducing their borrowing by about 12 percent on the days leading up to quarter-ends, and an additional 11 percent on quarter-end days. We find that European dealers further reduced their repo borrowing by 19 percent on the quarter-end day after Basel III was implemented, implying a $122 billion drop in European dealer borrowing. The total decline in European dealer repo borrowing on quarter-end days was, on average, 30 percent in the post-Basel III period. While these dynamics are strongly pronounced for European dealers, there is no significant pattern for U.S. dealers at or around quarter-ends, and no change in their patterns following the implementation of Basel III leverage ratio.

The increased window-dressing activity by European dealers post-Basel III introduced a demand shock into the repo market, and thus it allows us to estimate whether non-U.S. dealers had to pay a premium to keep their relationships stable with their cash lenders when they increased their window-dressing activity. As discussed above, on quarter-ends, when non-U.S. dealers pulled back from the market to reduce the size of their balance sheets for financial reporting, the facility offered a backstop to eligible MMFs to place their excess cash. In the absence of the facility, withdrawal of dealers from the market on calendar days would be inconvenient for their lenders. This dynamic could in turn affect trading relationships and induce changes in dealers' trading strategies to lock in funding on and around quarter-end days as MMF funding could potentially shift away from these dealers. In the next section, we turn to the question of whether these dynamics forced non-U.S. dealers to pay a premium to keep their trading relationships stable with their cash lenders.

6. Bargaining Power in Trading Relationships

Having shown that changes in the monetary policy framework (inception of the RRP facility) and changes in the regulatory framework (implementation of the Basel III leverage ratio) reduced overall

MMF cash lending in the triparty repo market and reduced borrowing demand from some dealers on quarter-end days, we now investigate how these shocks may have affected the bargaining power of MMFs and dealers in their trading relationships.

Trading relationships in the triparty repo market are important due to their over-the-counter nature, that is, lenders and borrowers seek each other out to negotiate prices every day. First, we show that by providing an alternative investment option, the RRP facility increased the bargaining power of eligible MMFs over their dealer borrowers with whom they have strong relationships. Second, we show that, while the cost of window dressing increased for dealers trading with eligible funds after the inception of the RRP facility, this was not the case for dealers trading with ineligible funds.

6.1 Empirical Framework

We test for changes in MMFs' bargaining power after the inception of the RRP facility in two steps. First, we examine whether the volatility of the rates in the triparty market declined for eligible MMFs in comparison with ineligible MMFs. Klee, Senyuz, and Yoldas (2019) show that the RRP facility led to stronger co-movement of overnight money market rates and reduced volatility of the repo rate. In light of these results, we look into the potential effect of the RRP facility on rate volatility to assess if the facility had anchored the rates charged by eligible MMFs in the private market. Second, we examine whether the overall decline in triparty repo lending by MMFs, shown in Table 2, was due to eligible MMFs pulling back from lending.

We then analyze the effects of Basel III implementation on bargaining power dynamics, which took place after the inception of the RRP facility. In particular, we analyze whether bargaining power shifted towards window-dressing dealers post-Basel III as their MMF lenders needed to place their cash elsewhere on financial reporting days.

We construct two measures of trading relationship strength for each lender and borrower reflecting the intensive and extensive margin of the importance of their relationship (see Maggio, Kermani, and Song 2017). To capture the importance of the borrower (lender) to the lender's (borrower's) business—the intensive margin—we define two daily variables for each borrower-lender (i, j) pair. Borrower's share of the lender's business is given by

$$(Share\ of\ Business)_{i,j,t} = \frac{(Volume)_{i,j,t}}{(Total\ Volume\ by\ i(j))_t}. \tag{4}$$

To capture the importance of the borrower (lender) to the lender's (borrower's) frequency of trading—the extensive margin—we define the *Frequency of Trading* between a borrower i and a lender j on day t, as follows:

$$\begin{aligned}(Frequency\ of\ Trading)_{i,j,t} \\ = \frac{(Number\ of\ Transactions)_{i,j,t}}{(Total\ Number\ of\ Transactions\ by\ i(j))_t}.\end{aligned} \tag{5}$$

To test for changes in MMF bargaining power after the RRP facility started, we estimate Equation (6) for the volume-weighted average repo rate, $r_{i,j,t}$, and also for its 30-day rolling standard deviation as a proxy for rate volatility, which is $\sigma(r)_{i,j,t}$.[16]

$$\begin{aligned} r_{i,j,t} = \beta_0 &+ \beta_{1i,j}\,(Rel\ Strength)_{i,j,t} \times 1(t = RRP2) \\ &+ \beta_{2i,j}\,(Rel\ Strength)_{i,j,t} + \theta_{1i,j} + \phi_t \\ &+ \delta_{1i,t-1} + \delta_{2j,t-1} + \epsilon_{i,j,t}, \end{aligned} \tag{6}$$

where *Rel Strength* is a vector capturing the relationship strength measures described above. β_1 is our coefficient of interest and shows how the reliance of dealers on MMF funding changed their rate for repo financing after the inception of the RRP facility.[17] We also include relationship fixed effects, $\theta_{1i,j}$, daily time fixed effects ϕ_t, and the borrower and lender control variables, $\delta_{1i,t-1}$ and $\delta_{2j,t-1}$, respectively, as defined in Section 3.1. Borrower (dealer) controls include

[16]Prior to August 22, 2014, given our Tuesday snapshot data we calculate the standard deviation by using the rate over the last four Tuesdays.

[17]We consider December 23, 2013 (that is, RRP2), as the start date of the facility because eligible funds did not reduce their lending until after individual bid amounts increased as of this date, as demonstrated in Table 2.

last quarter's $STFD$ and $Total\ Assets$. Lender (MMF) level controls are last month's AUM, Treasury repo outstanding, and Treasury securities held. Standard errors are clustered at the relationship level.

To test for potential changes in bargaining power after the implementation of Basel III, we estimate Equation (7) for the volume-weighted average repo rate, $r_{i,j,t}$ between borrower i and lender j on day t from January 2, 2013 to August 1, 2016.[18]

$$\begin{aligned} r_{i,j,t} = {} & \beta_0 + \beta_1 r_{i,j,t-1} + \theta_1 D_{1t} + \theta_{1i,j}\left(D_{1t} \times (Rel\ Strength)_{i,j,t}\right) \\ & + \theta_{2i,j}\,(Rel\ Strength)_{i,j,t} + \psi_{i,j} + \phi_t \\ & + \delta_{1i,t-1} + \delta_{2j,t-1} + \epsilon_{i,j,t}, \end{aligned} \tag{7}$$

where D_{1t} is a 3×1 vector including calendar-day indicators of one day prior to a quarter-end, the quarter-end, and one day after the quarter-end. Positive and significant coefficients in $\theta_{1i,j}$ would suggest that dealers that depend on their lenders pay higher rates on financial reporting days. We also include relationship fixed effects, $\psi_{i,j}$, and daily time fixed effects ϕ_t. The borrower and lender controls, as defined in Section 3.1, are reflected by $\delta_{1i,t-1}$ and $\delta_{2j,t-1}$, respectively. Standard errors are clustered at the relationship level.

6.2 Results

Table 4 presents the estimation results from Equation (6) using the volume-weighted average rate. We do not identify any significant and consistent effect of trading relationship strength on the level of the repo rate. This result is likely due to the anchoring effect of the RRP, as shown by Klee, Senyuz, and Yoldas (2019). However, if the facility anchored repo rates, we should observe a reduction in rate volatility after the facility was introduced. To explore this possibility, we estimate the same regression using the volatility of the repo rate as the dependent variable. Results are summarized in Table 5.

[18]Mid-June 2014 reflects the approximate implementation date for Basel III. Table A.1 in Appendix A shows that European dealers significantly increased their window dressing after the implementation of Basel III, not after its announcement.

Table 4. Effect of Trading Relationships on Repo Pricing

	Eligible MMF (1)	Ineligible MMF (2)	Eligible MMF (3)	Ineligible MMF (4)
Lender's Share of Business × RRP2	0.003**	–0.001		
	(1.98)	(–0.35)		
Dealer's Share of Business × RRP2	0.001	0.007		
	(0.31)	(0.41)		
Lender's Freq. × RRP2			0.001	–0.002
			(0.39)	(–0.81)
Dealer's Freq. × RRP2			–0.007***	0.025
			(–2.76)	(1.08)
Lender's Share of Business	–0.002	0.004		
	(–1.18)	(1.55)		
Dealer's Share of Business	0.000	–0.016		
	(0.07)	(–0.91)		
Lender's Freq. of Trading			–0.002	0.004
			(–1.09)	(1.62)
Dealer's Freq. of Trading			0.007**	–0.046**
			(2.40)	(–2.20)
Rate (–1)	0.235***	–0.008*	0.236***	–0.008*
	(9.10)	(–1.72)	(9.11)	(–1.72)
Controls	Yes	Yes	Yes	Yes
Observations	37,560	65,220	37,560	65,220
R^2	0.9922	0.4558	0.9922	0.4558

Note: This table presents the results of the panel regression shown in Equation (6) except where the dependent variable is the triparty rate charged between dealer i and MMF j. We include all foreign and U.S. dealers. The daily sample runs from January 2, 2013, to August 1, 2016, and is obtained from FRBNY. RRP2 is equal to 0 until December 23, 2013, and 1 thereafter. ***, **, and * indicate statistical significance at the 1 percent, 5 percent, and 10 percent level, respectively. t-ratios are reported in parentheses. Standard errors are clustered at the relationship level.

Table 5. Effect of Trading Relationships on Repo Volatility

	Eligible MMF (1)	**Ineligible MMF (2)**	**Eligible MMF (3)**	**Ineligible MMF (4)**
Lender's Share of Business × RRP2	0.005	–0.007		
	(1.35)	(–1.33)		
Dealer's Share of Business × RRP2	–0.010***	–0.014		
	(–2.96)	(–0.61)		
Lender's Freq. × RRP2			0.010*	–0.002
			(1.96)	(–0.44)
Dealer's Freq. × RRP2			–0.013***	–0.017
			(–2.90)	(–0.42)
Lender's Share of Business	–0.007*	0.003		
	(–1.95)	(0.71)		
Dealer's Share of Business	0.009***	0.003		
	(2.74)	(0.11)		
Lender's Freq. of Trading			–0.011**	0.002
			(–2.37)	(0.49)
Dealer's Freq. of Trading			0.012**	–0.059
			(2.51)	(–0.87)
Controls	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Relationship FE	Yes	Yes	Yes	Yes
Observations	37,560	65,219	37,560	65,219
R^2	0.6780	0.3806	0.6782	0.3808

Note: This table presents the results of the panel regression shown in Equation (6). The dependent variable is the rolling standard deviation of the triparty repo rate between dealer i and MMF j. We include all foreign and U.S. dealers. The daily sample runs from January 2, 2013, to August 1, 2016, and is obtained from FRBNY. RRP2 is equal to 0 until December 23, 2013, and 1 thereafter. ***, **, and * indicate statistical significance at the 1 percent, 5 percent, and 10 percent level, respectively. t-ratios are reported in parentheses. Standard errors are clustered at the relationship level.

As shown in columns 1 and 3, a one-standard-deviation increase in dealers' *Share of Business* and *Frequency of Trading* is associated with a 1 basis point decline in the volatility of repo rates received by eligible MMFs after the RRP facility started. This finding, which is consistent with Klee, Senyuz, and Yoldas (2019), reflects that the facility offering rate anchored rates that eligible funds charged when lending. Insignificant coefficients on the lender side indicate that the dependence of the lender on the borrower does not play a role in the stability of the repo rate. Further, as shown in columns 2 and 4, there is no statistically significant change to the volatility of repo rates received by ineligible MMFs from their dealer borrowers regardless of relationship strength after the RRP facility, which is consistent with the fact that these funds do not have access to the facility.[19]

Table 6 shows the estimation results from Equation (7). As shown in columns 1 and 3, a one-standard-deviation increase in dealers' *Share of Business* or *Frequency of Trading* is associated with higher rates charged for dealers on financial reporting days when borrowing from eligible MMFs. Specifically, dealers dependent on eligible MMF business pay about half a basis point more for funding on the day before the quarter-end than on other days, reflecting the cost of window dressing. The RRP facility allowed eligible MMFs to maintain their bargaining power over their dealer borrowers with whom they had strong relationships, regardless of their window-dressing activity. As shown in columns 2 and 4, there is no significant change in the rates that ineligible MMFs charged their dealer borrowers on or around quarter-end.

Table 7 presents the results when the dependent variable in Equation (7) is replaced with the log of transaction volume between dealer i and MMF j on day t. As shown in columns 1 and 3, a one-standard-deviation increase in dealer's *Share of Business* and dealer's *Frequency of Business* is associated with about 90 percent and 150 percent increase in repo volumes on quarter-ends in comparison to other days, suggesting that dealers dependent on eligible MMF business were unable to window dress as effectively as with

[19] Although some coefficients of interest appear larger in magnitude in columns 2 and 4 than those in columns 1 and 3, we are unable reject the null hypothesis that they are equal to zero.

Table 6. Effect of Trading Relationships on Repo Pricing during Quarter-ends

	Eligible MMF (1)	Ineligible MMF (2)	Eligible MMF (3)	Ineligible MMF (4)
Dealer's Share of Business × Quarter-end (–1)	0.005**	–0.002		
	(2.59)	(–0.18)		
Dealer's Share of Business × Quarter-end	0.008	–0.005		
	(1.17)	(–0.44)		
Dealer's Share of Business × Quarter-end (+1)	–0.003	0.005		
	(–0.91)	(1.30)		
Dealer's Freq. × Quarter-end (–1)			0.007***	0.004
			(2.77)	(0.32)
Dealer's Freq. × Quarter-end			0.008	–0.008
			(1.15)	(–0.65)
Dealer's Freq. × Quarter-end (+1)			–0.004	0.003
			(–0.95)	(0.75)
Dealer's Share of Business	–0.001	–0.012***		
	(–0.72)	(–2.89)		
Dealer's Freq. of Trading			–0.002	–0.027*
			(–0.81)	(–1.78)
Rate (–1)	0.254***	–0.009**	0.254***	–0.009**
	(9.25)	(–2.10)	(9.25)	(–2.11)
Controls	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Relationship FE	Yes	Yes	Yes	Yes
Observations	32,331	56,128	32,331	56,128
R^2	0.9922	0.4303	0.9922	0.4303

Note: This table presents the results of the panel regression shown in Equation (7). The dependent variable is the triparty repo rate charged by MMF j to dealer i. We include all foreign and U.S. dealers. The daily sample runs from January 2, 2013, to August 1, 2016, and is obtained from FRBNY. ***, **, and * indicate statistical significance at the 1 percent, 5 percent, and 10 percent level, respectively. t-ratios are reported in parentheses. Standard errors are clustered at the relationship level.

Table 7. Effect of Trading Relationships on Repo Volumes during Quarter-ends

	Eligible MMF (1)	Ineligible MMF (2)	Eligible MMF (3)	Ineligible MMF (4)
Dealer's Share of Business × Quarter-end (–1)	–0.272	–0.684		
	(–1.29)	(–0.82)		
Dealer's Share of Business × Quarter-end	0.897**	0.421		
	(2.31)	(1.29)		
Dealer's Share of Business × Quarter-end (+1)	0.309	–0.498		
	(0.99)	(–1.57)		
Dealer's Freq. × Quarter-end (–1)			–0.277*	–1.569**
			(–1.83)	(–2.19)
Dealer's Freq. × Quarter-end			1.520***	1.796***
			(3.53)	(5.35)
Dealer's Freq. × Quarter-end (+1)			0.897**	0.797***
			(2.26)	(2.90)
Dealer's Share of Business	3.262***	3.421***		
	(6.71)	(5.22)		
Dealer's Freq. of Trading			1.567***	1.445***
			(4.42)	(4.66)
Volume (–1)	0.489***	0.632***	0.598***	0.695***
	(16.03)	(28.80)	(21.52)	(37.71)
Controls	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Relationship FE	Yes	Yes	Yes	Yes
Observations	32,331	56,128	32,331	56,128
R^2	0.8040	0.9126	0.7693	0.9049

Note: This table presents the results of the panel regression shown in Equation (7), except the dependent variable is the log of the volume borrowed by dealer i from MMF j. We include all foreign and U.S. dealers. The daily sample runs from January 2, 2013, to August 1, 2016, and is obtained from FRBNY. ***, **, and * indicate statistical significance at the 1 percent, 5 percent, and 10 percent level, respectively. t-ratios are reported in parentheses. Standard errors are clustered at the relationship level.

other dealers. Further, results in column 4 suggest that ineligible MMFs were able to maintain some bargaining power over the dealers that are reliant on their business, but only through the extensive margin, as indicated by the positive and significant coefficient on dealer's *Frequency of Business* on quarter-ends.

In summary, we find that, even after the implementation of Basel III, which strengthened window-dressing incentives for certain dealers, eligible MMF bargaining power increased in triparty repo trading. Those dealers that were reliant on eligible MMF lending were unable to window dress as much as other dealers, and ended up paying higher rates for funding on financial reporting days.

One caveat to these results is that they do not take into account the formation and termination of relationships, and that the eligibility of MMFs to the RRP was not random. In Section B.1 of Appendix B, we show results that test the robustness of Tables 5, 6, and 7. We condition on trading relationships between dealers and MMFs that existed prior to the RRP and persisted through the end of our sample. The results are presented in Tables A.3, A.4, and A.5. Our results are consistent with those presented in Section 6.2, suggesting that trading relationship formations and terminations are not the factors driving our results. In Section B.2 of Appendix B, we explore the possibility that observable characteristics of eligible MMFs, rather than their eligibility to the RRP, might be driving our results. In Tables A.6, A.7, and A.8, we compare eligible funds to ineligible funds that could have been eligible had they submitted the paperwork to participate at the RRP. All funds in the robustness analysis had an AUM of at least $5 billion for six consecutive months, which was the most stringent condition of RRP eligibility. Our analysis shows that RRP eligibility, and not observable characteristics of eligible MMFs, drives our main results.

Finally, in Appendix C we present some additional results exploring how the network between dealers and MMFs changed after the inception of the RRP facility. We find evidence that the strength of relationships between non-U.S. dealers and ineligible MMFs weakened during the post-Basel III period and, at the same time, non-U.S. dealers became more reliant on eligible MMF funding. These results are presented in Tables A.9 and A.10. We argue that this shift in the network was likely a contributing factor for window dressing to become more expensive for non-U.S. dealers.

7. Implications for MMFs' Balance Sheet Management

Financial intermediaries, such as dealers, typically finance their portfolio of high-quality but long-term assets in the repo market. To meet their short-term funding needs, dealers rely on MMFs as their primary cash lenders in the repo market. If dealers were unable to fund their assets in the repo market on a consistent basis, resulting liquidity mismatch has the potential to disrupt financial stability.

In this section, we analyze how the inception of the RRP facility as a backstop for eligible MMFs may have affected MMFs' investment strategies. The data on MMF balance sheets come from the SEC N-MFP filings at the end of each month from January 31, 2013 through July 31, 2016 for all eligible and ineligible MMFs in our data set.

Tables 8 and 9 summarize the regression results for eligible and ineligible MMFs, respectively. Control variables include the MMFs' assets under management and standard errors are clustered at the fund level. In each table, column 1 shows the effect of changes in Treasury repo activity (which includes RRP takeup) on changes in repo activity backed by other collateral including asset-backed commercial paper (ABCP), commercial paper (CP), and corporate debt on MMF j's balance sheet; column 2 shows the effect of changes in Treasury repo activity on the market value of all other items, excluding repo, on MMFs' balance sheet; and column 3 shows the effect of changes in Treasury repo activity on the weighted-average maturity (WAM) of all other items, excluding repo, on MMFs' balance sheet.

Table 8 presents some evidence of substitution between Treasury repo and repo backed by other collateral for eligible funds after the inception of the RRP facility. A one-standard-deviation increase in the change in Treasury repo is associated with an 18 percentage point decline in repo backed by other collateral (column 1). The evidence is much stronger for substitution between Treasury repo and other items on the balance sheet. We observe an almost 40 percentage point decline in all other items, excluding repo, in response to a one-standard-deviation increase in Treasury repo (column 2). These results translate to an average of \$152 million and \$2.9 billion decline per fund in repo backed by other collateral and in other items on the balance sheet, respectively. Eligible MMFs appeared to have preferred Treasury repo over repo backed by other collateral that

Table 8. Effect of RRP Facility on Eligible MMF Holdings

	Chg. Other Repo (1)	**Chg. All Other (2)**	**Chg. WAM All Other (3)**
Chg. Tsy. Repo × RRP2	–0.183**	–0.397***	
	(–2.06)	(–4.37)	
Chg. Tsy. Repo	0.080	–0.394***	
	(1.33)	(–3.47)	
Tsy. Repo × RRP2			–0.111*
			(–1.80)
Tsy. Repo			0.089
			(1.10)
Chg. All Other			0.490***
			(2.95)
Controls	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
MMF FE	Yes	Yes	Yes
Observations	1,124	1,124	1,124
R^2	0.1618	0.5629	0.1529

Note: This table presents the results of a panel regression for eligible MMFs. The dependent variable in column 1 is the change in repo outstanding backed by all collateral other than Treasury securities. The dependent variable in column 2 is the change in the market value of all other items on MMF j's balance sheet excluding repo. The dependent variable in column 3 is the change in the weighted-average maturity (WAM) of all other items, excluding repo, on MMF j's balance sheet. The monthly sample run uses month-end snapshots from N-MFP filings from January 30, 2013 to July 31, 2016, and is obtained from the SEC. ***, **, and * indicate statistical significance at the 1 percent, 5 percent, and 10 percent level, respectively. t-ratios are reported in parentheses. Standard errors are clustered at the fund level.

might be considered more risky, and also over investing in ABCP, CP, and corporate debt (components of "all other"). Further, there is some evidence of eligible MMFs reducing their WAMs by about 11 percentage points, as indicated by the negative coefficient on the interaction term that is statistically significant at the 10 percent level.

Table 9 shows the results on changes in the balance sheets of ineligible funds after the RRP facility. As expected, since they could not access the facility, ineligible funds did not change their balance sheets

Table 9. Effect of RRP Facility on Ineligible MMF Holdings

	Chg. Other Repo (1)	Chg. All Other (2)	Chg. WAM All Other (3)
Chg. Tsy. Repo × RRP2	–0.251	–0.645*	
	(–1.15)	(–1.95)	
Chg. Tsy. Repo	0.031	–0.127	
	(0.21)	(–0.48)	
Tsy. Repo × RRP2			–0.090***
			(–3.39)
Tsy. Repo			0.197**
			(2.06)
Chg. All Other			–0.050
			(–0.76)
Controls	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
MMF FE	Yes	Yes	Yes
Observations	1,999	1,999	1,979
R^2	0.1489	0.3031	0.0415

Note: This table presents the results of a panel regression for ineligible MMFs. The dependent variable in column 1 is the change in repo outstanding backed by all collateral other than Treasury securities. The dependent variable in column 2 is the change in the market value of all other items on MMF j's balance sheet excluding repo. The dependent variable in column 3 is the change in the weighted-average maturity (WAM) of all other items, excluding repo, on MMF j's balance sheet. The monthly sample run uses month-end snapshots from N-MFP filings from January 30, 2013 to July 31, 2016, and is obtained from the SEC. ***, **, and * indicate statistical significance at the 1 percent, 5 percent, and 10 percent level, respectively. t-ratios are reported in parentheses. Standard errors are clustered at the fund level.

materially after its inception. However, as shown in column 3, ineligible MMFs shortened their WAMs by 9 percentage points on average, which corresponds to about a decline of six days to accommodate outflows. While there is also some evidence of WAM shortening for eligible funds during this period, much stronger evidence for ineligible MMFs may reflect their potential hedging behavior against risks of not having the RRP facility as a backstop. Their inability to tap the RRP facility when they needed to place their excess cash overnight may have induced them to rely on investments with shorter duration, other things equal.

Our results suggest that the inception of the RRP facility influenced the way MMFs manage their balance sheets. With the changing incentives, eligible MMFs shifted some of their investments to safer alternatives such as Treasury repo. While the eligible funds benefited from having the RRP facility as a backstop, balance sheet management of ineligible funds remained largely unchanged, except for a slight shortening of their WAMs, which may reflect a desire to hedge against increased outflows.

8. Concluding Remarks

We analyze the effects of monetary and regulatory policy on trading dynamics in the U.S. repo market. On the supply side, we find that the introduction of the RRP facility led to a 16 percent reduction in cash lending by MMFs eligible to transact with the Fed in the repo market. On the demand side, we find that after the Basel III leverage ratio implementation, window dressing by European dealers intensified notably with total reduction in their repo borrowing on quarter-ends reaching 30 percent. Putting these two shocks together, we show that bargaining power of MMFs increased after the inception of the RRP facility. For those dealers reliant on eligible MMF funding, window dressing became more expensive than for dealers less reliant on MMF funding. However, the average rates they paid remained stable given the anchoring effect of the RRP facility. Our results highlight the importance of MMF-dealer relationships and their effects on bargaining power dynamics in the triparty repo market.

We also find evidence that the changing trading dynamics in the repo market following the RRP facility influenced the way MMFs manage their balance sheets. While eligible MMFs shifted the composition of their balance sheets towards Treasury repo and away from repo backed by other collateral, ineligible funds reduced their WAMs modestly. These findings highlight the financial stability implications of the RRP facility through its effects on MMF balance sheet management.

Appendix A. Basel III Capital Regulations and Repo Borrowing

In this section, we describe our empirical framework to analyze the effects of Basel III capital regulations on triparty repo borrowing on regulatory reporting days. For this analysis, we use a separate confidential data set of daily overnight repo borrowing by each dealer in the triparty market, from July 1, 2008, to August 1, 2016, which were reported by two custodian banks—BNYM and JPM—to FRBNY.[20]

Basel III rules that required repo positions to be included in the leverage exposure calculations were announced in the United States in June/July 2013, which coincided with the timing of leverage ratio requirements being transposed to local rules in Europe. Although full compliance with the new regulations was not mandated until January 2018, banks started to adjust their strategies earlier in order to signal that they are well positioned to meet regulatory targets by the compliance deadlines. On January 1, 2015, dealers began reporting the new leverage ratios to the public, including three quarters of historical data, making 2014:Q2 the first quarter-end entering into the calculations. Our sensitivity analysis across alternative quarter-end dates also confirmed that 2014:Q2 was associated with the largest statistically significant dealer response on a quarter-end.

To measure the extent of window dressing by dealer groups from different regional jurisdictions and their response to the implementation of Basel III leverage ratio requirements, we estimate the following time-series regression for daily repo borrowing by European and U.S. dealers from July 1, 2008, to August 1, 2016.

$$\begin{aligned} RB_t = {} & \beta_0 + \beta_1 D_{1t} + \beta_2 D_{2t} + \phi_1 RB_{t-1} \\ & + \theta_1 \big(\mathit{Quarter\text{-}end} \times \mathit{Announced}\big)_t \\ & + \theta_2 \big(\mathit{Quarter\text{-}end} \times \mathit{Implemented}\big)_t + \epsilon_t, \end{aligned} \tag{8}$$

where RB_t is the log of aggregate repo borrowing by dealers on day t. *Announced* equals 1 after the details of Basel III regulations were announced on June 15, 2013 and turns 0 after the implementation takes place. *Implemented* takes the value 0 before June 15,

[20]We exclude data on interdealer GCF segment and RRP trades with the Fed.

Table A.1. Overnight Repo Borrowing by Dealers on Calendar Days and Effects of Regulations

	European Overnight Borrowing (1)	U.S. Overnight Borrowing (2)
Quarter-end × Announced	–0.021	0.030
	(–0.59)	(1.14)
Quarter-end × Implemented	–0.193***	0.066
	(–5.84)	(1.62)
Quarter-end (–5)	–0.005	0.016
	(–0.71)	(1.51)
Quarter-end (–4)	–0.033**	–0.014
	(–2.45)	(–1.03)
Quarter-end (–3)	–0.013	–0.003
	(–0.96)	(–0.35)
Quarter-end (–2)	–0.051***	–0.008
	(–5.75)	(–1.34)
Quarter-end (–1)	–0.043***	0.031***
	(–4.13)	(4.07)
Quarter-end	–0.107***	0.037
	(–4.46)	(1.46)
Quarter-end (+1)	0.155***	–0.047***
	(5.43)	(–3.23)
Quarter-end (+2)	0.022***	0.004
	(2.82)	(0.33)
Quarter-end (+3)	–0.008	–0.010
	(–0.84)	(–1.24)
Quarter-end (+4)	–0.000	–0.010
	(–0.05)	(–0.85)
Quarter-end (+5)	0.019**	0.014
	(1.96)	(1.61)
Overnight Borrowing (–1)	0.887***	0.957***
	(54.14)	(98.62)
Observations	1,056	1,056
R^2	0.977	0.981

Note: This table presents the results of time-series regressions for the log of overnight repo borrowing by European and U.S. dealers, in columns 1 and 2, respectively. *Announced* equals 1 after the details of Basel III regulations were announced on June 15, 2013 and turns 0 when the implementation of the new rules takes place. *Implemented* takes the value 0 before June 15, 2014, and 1 after this date. The daily sample runs from January 2, 2011, to August 1, 2016. ***, **, and * indicate statistical significance at the 1 percent, 5 percent, and 10 percent level, respectively. t-ratios are reported in parentheses and are calculated using robust standard errors.

2014, and 1 after this date. Since these models are estimated at the daily frequency, we switch on the indicator variables in the middle of the month to capture dynamics of the first quarter-end that follows. *Quarter-end* is an indicator variable that take the value 1 on quarter-ends, and 0 on other days. D_{1t} is a 12×1 vector that includes calendar-day indicators of five days prior to a quarter-end, the quarter-end, five days after the quarter-end, and *Month-end.* D_{2t} is a 2×1 vector that includes the *Announced* and *Implemented* indicators.

Our specification has two interaction terms to capture the change in repo borrowing with respect to our dates of interest: θ_1 measures the change in borrowing on quarter-ends after the announcement; θ_2 captures the quarter-end change after the implementation. If European dealers were incentivized to contract their balance sheet on financial reporting days because of less stringent implementation of the Basel III leverage ratio, we expect θ_1 and θ_2 to be negative and significant for European dealers and insignificant for U.S. dealers whose calculations are based on daily averages. Since repo borrowing exhibits substantial persistence, we also include its first lag in the model to account for autocorrelation. We calculate standard errors robust to heteroskedasticity and winsorize all continuous variables at the 1 percent level, although our results are robust to keeping the outliers.

Appendix B. Robustness Analysis

B.1 Trading Relationships

Since short-term funding is difficult to replace, dealers tend to trade with the same MMFs every day. Table B.1 illustrates the persistence of these relationships. There are 1,101 dealer-MMF pairs in the data from January 1, 2013 to August 1, 2016, with an average number of days in a trading relationship equal to 307. A dealer-MMF pair is considered to be in a relationship if the pair trades for at least two trading days. Conditional on dealer-MMF pairs that had a trading relationship from September 23, 2013, to August 1, 2016, we identify 623 dealer-MMF pairs with an average number of days in a relationship equal to 319 days. These statistics suggest that relationships that terminated before the inception of the RRP were just as persistent as relationships that existed well after the inception of the RRP.

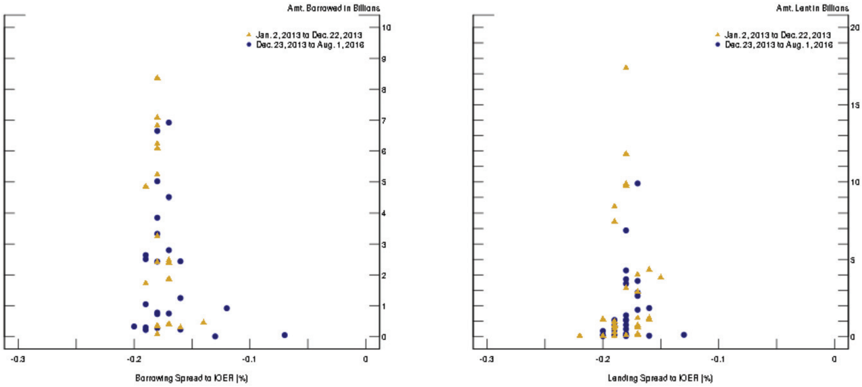
Table B.1. Summary Statistics of Trading Relationships

	Statistic
No. Dealers	23
No. MMFs	288
No. Dealer-MMF Pairs	1,101
No. Dealer-MMF Pairs, Conditional on Existing before RRP	623
Avg. No. Days in Relationship	307
Avg. No. Days in Relationship, Conditional on Existing before RRP	319
Avg. Share of MMF's Business	0.39
Avg. Share of Dealer's Business	0.39
Avg. Count of MMF's Business	0.08
Avg. Count of Dealer's Business	0.07

Figure B.1 shows the transacted volumes and the associated rates by borrowers (dealers) and lenders (MMFs), respectively, from January 2, 2013, to December 22, 2013 (triangles) and from December 23, 2013, to August 1, 2016 (dots) for overnight GC repo. Each triangle (dot) represents a borrower or lender. We observe that dealers borrow roughly the same amount each day from the same MMF lenders, with very similar transaction volumes before and after the RRP facility. Accordingly, Figure B.1 illustrates the inelasticity of demand and supply for short-term funding in the triparty repo market.

Tables B.2, B.3, and B.4 display robustness analysis for the results in Tables 5, 6, and 7, respectively, which are not conditional on existing trading relationships before the RRP facility. We reestimate Equations (6) and (7) conditioning on trading relationships that existed before the RRP facility. Table B.2 displays the regression results of Equation (6) conditioning on existing relationships. As shown in columns 1 and 3, all else equal, a one-standard-deviation increase in *Dealer's Share of Business* and *Dealer's Frequency of Trading* is associated with about a 1 basis point decline in the standard deviation of repo rates received by eligible MMFs after the inception of the RRP facility. Further, as shown in columns 2 and 4, there is no change to the standard deviation of repo rates received by ineligible MMFs from their dealer borrowers regardless of relationship strength after the RRP facility. Table B.3 displays results from

Figure B.1. Trading Dynamics in Triparty Repo



Note: This figure displays the borrowing (left) and lending (right) spread against total volume for each borrower (and lender) in the overnight GC triparty repo market across two date ranges: before the inception of the RRP between January 2 and December 22, 2013 (triangles); and after the inception of the RRP between December 23, 2013 and August 1, 2016 (circles). The borrowing spread is defined as the volume-weighted average interest rate a borrower paid minus IOER. The lending spread is defined as the volume-weighted average interest rate a lender gained minus IOER. The data are confidential transaction data and are from the Federal Reserve Bank of New York.

Equation (7) conditional on existing relationships before the RRP. Columns 1 and 3 show that a one-standard-deviation in dealers' *Share of Business* or *Frequency of Trading* is associated with higher borrowing rates from eligible MMFs on financial reporting days. As shown in columns 2 and 4, there is no change in the rates that ineligible MMFs charged their dealer borrowers that window dress with whom they have strong relationships. Finally, Table B.4 shows the results where the dependent variable in Equation (7) is replaced with the log of transaction volume between dealer i and MMF j on day t, for those dealer-MMF relationships that existed before the RRP facility. As shown in columns 1 and 3, a one-standard-deviation increase in *Dealer's Share of Business* and *Dealer's Frequency of Business* is associated with about 106 percent and 170 percent *increase* in repo volumes on quarter-ends in comparison to other days. All these results are consistent with our main results shown in Tables 5, 6, and 7.

Table B.2. Effect of Existing Trading Relationships on Repo Volatility

	Eligible MMF (1)	**Ineligible MMF (2)**	**Eligible MMF (3)**	**Ineligible MMF (4)**
Lender's Share of Business × RRP2	0.006	–0.004		
	(1.58)	(–0.84)		
Dealer's Share of Business × RRP2	–0.010***	–0.013		
	(–2.90)	(–0.53)		
Lender's Freq. × RRP2			0.011**	0.001
			(2.22)	(0.13)
Dealer's Freq. × RRP2			–0.011**	–0.028
			(–2.45)	(–0.72)
Lender's Share of Business	–0.006*	0.005		
	(–1.77)	(1.09)		
Dealer's Share of Business	0.008**	–0.001		
	(2.32)	(–0.03)		
Lender's Freq. of Trading			–0.010**	0.003
			(–2.11)	(0.72)
Dealer's Freq. of Trading			0.012**	0.004
			(2.36)	(0.16)
Controls	Yes	Yes	Yes	Yes
Observations	28,693	41,455	28,693	41,455
R^2	0.6725	0.2487	0.6730	0.2488

Note: This table presents the results of the panel regression shown in Equation (6) but conditions on trading relationships that existed prior to September 22, 2013. The dependent variable is the rolling standard deviation of the overnight GC triparty repo rate between dealer i and MMF j. We include all foreign (European, U.K., Japanese, Canadian) and domestic dealers. The daily sample runs from January 2, 2013, to August 1, 2016, and is obtained from FRBNY. RRP2 is equal to 0 until December 23, 2013, and 1 thereafter. ***, **, and * indicate statistical significance at the 1 percent, 5 percent, and 10 percent level, respectively. t-ratios are reported in parentheses. Standard errors are clustered at the dealer-MMF level.

Table B.3. Effect of Existing Trading Relationships on Repo Pricing during Quarter-ends

	Eligible MMF (1)	**Ineligible MMF (2)**	**Eligible MMF (3)**	**Ineligible MMF (4)**
Dealer's Share of Business × Quarter-end (–1)	0.005**	0.005		
	(2.50)	(0.71)		
Dealer's Share of Business × Quarter-end	0.003	–0.011		
	(0.49)	(–1.14)		
Dealer's Share of Business × Quarter-end (+1)	–0.004	0.010		
	(–0.97)	(1.56)		
Dealer's Freq. × Quarter-end (–1)			0.006***	–0.001
			(2.87)	(–0.07)
Dealer's Freq. × Quarter-end			0.003	–0.015
			(0.45)	(–1.57)
Dealer's Freq. × Quarter-end (+1)			–0.004	0.005
			(–1.06)	(0.91)
Dealer's Share of Business	–0.002	–0.014**		
	(–1.14)	(–2.49)		
Dealer's Freq. of Trading			0.000	–0.016
			(0.09)	(–0.86)
Rate (–1)	0.248***	0.030	0.248***	0.030
	(7.68)	(0.92)	(7.68)	(0.92)
Controls	Yes	Yes	Yes	Yes
Observations	23,722	33,165	23,722	33,165
R^2	0.9920	0.8267	0.9920	0.8267

Note: This table presents the results of the panel regression shown in Equation (7) but conditions on trading relationships that existed prior to September 22, 2013. The dependent variable is the overnight GC triparty repo rate charged by MMF j to dealer i. We include all foreign (European, U.K., Japanese, Canadian) and domestic dealers. The daily sample runs from January 2, 2013, to August 1, 2016, and is obtained from FRBNY. ***, **, and * indicate statistical significance at the 1 percent, 5 percent, and 10 percent level, respectively. t-ratios are reported in parentheses. Standard errors are clustered at the relationship level.

Table B.4. Effect of Existing Trading Relationships on Repo Volumes during Quarter-ends

	Eligible MMF (1)	**Ineligible MMF (2)**	**Eligible MMF (3)**	**Ineligible MMF (4)**
Dealer's Share of Business × Quarter-end (–1)	–0.387	0.412		
	(–1.66)	(0.55)		
Dealer's Share of Business × Quarter-end	1.061**	0.810		
	(2.46)	(1.27)		
Dealer's Share of Business × Quarter-end (+1)	0.276	–0.604		
	(0.78)	(–1.06)		
Dealer's Freq. × Quarter-end (–1)			–0.337**	–0.367
			(–2.14)	(–0.53)
Dealer's Freq. × Quarter-end			1.695***	2.319***
			(3.50)	(4.38)
Dealer's Freq. × Quarter-end (+1)			0.887*	1.060**
			(1.89)	(2.32)
Dealer's Share of Business	3.427***	3.953***		
	(6.18)	(4.08)		
Dealer's Freq. of Trading			1.603***	1.770***
			(4.26)	(3.96)
Volume (–1)	0.454***	0.621***	0.559***	0.698***
	(15.13)	(20.31)	(20.25)	(27.44)
Controls	Yes	Yes	Yes	Yes
Observations	23,722	33,165	23,722	33,165
R^2	0.7813	0.8947	0.7426	0.8828

Note: This table presents the results of the panel regression shown in Equation (7), except the dependent variable is the log of overnight GC borrowing by dealer i from MMF j, and conditions on trading relationships that existed prior to September 22, 2013. We include all foreign (European, U.K., Japanese, Canadian) and domestic dealers. The daily sample runs from January 2, 2013, to August 1, 2016, and is obtained from FRBNY. ***, **, and * indicate statistical significance at the 1 percent, 5 percent, and 10 percent level, respectively. t-ratios are reported in parentheses. Standard errors are clustered at the relationship level.

To sum up, trading relationships between dealers and MMFs are long term and persistent, with very stable transacted volumes. Relationship formations and terminations do not seem to affect MMF bargaining power in this market.

B.2 MMF Characteristics

We perform robustness tests of the results discussed in Section 6.2 to explore potential bias in comparison of the bargaining power of eligible versus ineligible MMFs, given that eligible MMFs could choose to become RRP counterparties. Although the process of applying to the Federal Reserve Bank of New York to become an RRP counterparty was not arduous, a major criterion included having AUM of at least $5 billion over the last consecutive six months.

Table B.5 shows the results from Equation (6). In columns 2 and 4 we now condition on funds that could have been eligible to use the RRP facility, as they are similar to eligible MMFs in that they have AUM above $5 billion. We observe no changes to the standard deviation of repo rates received by funds that met the eligibility requirements regardless of relationship strength after the inception of the facility, consistent with the results reported in Table 5, suggesting that fund characteristics were not driving the standard deviation of repo rates. Table B.6 shows the estimation results from Equation (7) where columns 2 and 4 condition on funds that met the eligibility requirements but did not become eligible. Again, no change is observed in the rates charged by MMFs, consistent with the results in Table 6, providing further evidence that observable characteristics of eligible funds were not driving the repo rates. Finally, Table B.7 shows the results where the dependent variable in Equation (7) is replaced with the log of transaction volume between dealer i and MMF j on day t. As shown in columns 2 and 4, we observe fewer changes in repo volumes on quarter-ends in comparison to other days for funds that met the eligibility requirements in comparison to eligible MMFs. These results are consistent with our main results shown in Table 7, and provide further evidence that it was the presence of the RRP facility that affected the bargaining power of eligible MMFs.

Table B.5. Effect of Trading Relationships on Repo Pricing Volatility Comparing Eligible MMFs to Funds that Met RRP Requirements

	Eligible MMF (1)	**Meets Req. (2)**	**Eligible MMF (3)**	**Meets Req. (4)**
Lender's Share of Business × RRP2	0.005	–0.017		
	(1.35)	(–1.58)		
Dealer's Share of Business × RRP2	–0.010***	–0.044		
	(–2.96)	(–0.89)		
Lender's Freq. × RRP2			0.010*	–0.007
			(1.96)	(–1.01)
Dealer's Freq. × RRP2			–0.013***	–0.106
			(–2.90)	(–1.08)
Lender's Share of Business	–0.007*	0.008		
	(–1.95)	(1.37)		
Dealer's Share of Business	0.009***	0.043		
	(2.74)	(0.79)		
Lender's Freq. of Trading			–0.011**	0.002
			(–2.37)	(0.19)
Dealer's Freq. of Trading			0.012**	0.064
			(2.51)	(0.73)
Controls	Yes	Yes	Yes	Yes
Observations	37,560	17,571	37,560	17,571
R^2	0.6780	0.1989	0.6782	0.1987

Note: This table presents the results of the panel regression shown in Equation (6). Columns 1 and 3 display the results for eligible MMFs. Columns 2 and 4 display the results for MMFs that met the eligibility requirements to approach the RRP. The dependent variable is the rolling standard deviation of the overnight GC triparty repo rate between dealer i and MMF j. We include all foreign (European, U.K., Japanese, Canadian) and domestic dealers. The daily sample runs from January 2, 2013, to August 1, 2016, and is obtained from FRBNY. RRP2 is equal to 0 until December 23, 2013, and 1 thereafter. ***, **, and * indicate statistical significance at the 1 percent, 5 percent, and 10 percent level, respectively. t-ratios are reported in parentheses. Standard errors are clustered at the relationship level.

Table B.6. Effect of Trading Relationships on Repo Pricing during Quarter-ends Comparing Eligible MMFs to Funds that Met RRP Requirements

	Eligible MMF (1)	Meets Req. (2)	Eligible MMF (3)	Meets Req. (4)
Dealer's Share of Business × Quarter-end (–1)	0.005**	0.027**		
	(2.59)	(2.47)		
Dealer's Share of Business × Quarter-end	0.008	0.012		
	(1.17)	(1.33)		
Dealer's Share of Business × Quarter-end (+1)	–0.003	0.013		
	(–0.91)	(1.24)		
Dealer's Freq. × Quarter-end (–1)			0.007***	0.006
			(2.77)	(0.14)
Dealer's Freq. × Quarter-end			0.008	0.010
			(1.15)	(1.18)
Dealer's Freq. × Quarter-end (+1)			–0.004	0.012
			(–0.95)	(1.25)
Dealer's Share of Business	–0.001	–0.012		
	(–0.72)	(–1.30)		
Dealer's Freq. of Trading			–0.002	–0.033
			(–0.81)	(–1.28)
Rate (–1)	0.254***	0.016	0.254***	0.016
	(9.25)	(0.80)	(9.25)	(0.80)
Controls	Yes	Yes	Yes	Yes
Observations	32,331	14,713	32,331	14,713
R^2	0.9922	0.7421	0.9922	0.7421

Note: This table presents the results of the panel regression shown in Equation (7). Columns 1 and 3 display the results for eligible MMFs. Columns 2 and 4 display the results for MMFs that met the eligibility requirements to approach the RRP. The dependent variable is the overnight GC triparty repo rate charged by MMF j to dealer i. We include all foreign (European, U.K., Japanese, Canadian) and domestic dealers. The daily sample runs from January 2, 2013, to August 1, 2016, and is obtained from FRBNY. ***, **, and * indicate statistical significance at the 1 percent, 5 percent, and 10 percent level, respectively. t-ratios are reported in parentheses. Standard errors are clustered at the relationship level.

Table B.7. Effect of Trading Relationships on Repo Volumes during Quarter-ends Comparing Eligible MMFs to Funds that Met RRP Requirements

	Eligible MMF (1)	Meets Req. (2)	Eligible MMF (3)	Meets Req. (4)
Dealer's Share of Business × Quarter-end (–1)	–0.272	0.876		
	(–1.29)	(0.80)		
Dealer's Share of Business × Quarter-end	0.897**	–0.156		
	(2.31)	(–0.20)		
Dealer's Share of Business × Quarter-end (+1)	0.309	–0.938		
	(0.99)	(–0.93)		
Dealer's Freq. × Quarter-end (–1)			–0.277*	0.768
			(–1.83)	(0.70)
Dealer's Freq. × Quarter-end			1.520***	2.042***
			(3.53)	(4.14)
Dealer's Freq. × Quarter-end (+1)			0.897**	1.026
			(2.26)	(1.46)
Dealer's Share of Business	3.262***	4.635***		
	(6.71)	(3.87)		
Dealer's Freq. of Trading			1.567***	2.471***
			(4.42)	(3.95)
Volume (–1)	0.489***	0.561***	0.598***	0.685***
	(16.03)	(13.51)	(21.52)	(20.20)
Controls	Yes	Yes	Yes	Yes
Observations	32,331	14,713	32,331	14,713
R^2	0.8040	0.8096	0.7693	0.7706

Note: This table presents the results of the panel regression shown in Equation (7), where the dependent variable is replaced to be the log of overnight GC volume borrowed by dealer i from MMF j. Columns 2 and 4 display the results for ineligible MMFs that met the eligibility requirements to approach the RRP. We include all foreign (European, U.K., Japanese, Canadian) and domestic dealers. The daily sample runs from January 2, 2013, to August 1, 2016, and is obtained from FRBNY. ***, **, and * indicate statistical significance at the 1 percent, 5 percent, and 10 percent level, respectively. t-ratios are reported in parentheses. Standard errors are clustered at the relationship level.

Appendix C. Triparty Repo Network Structure

We present additional results on the effects of changing monetary (RRP facility) and regulatory policy (Basel III leverage ratio) changes on the triparty repo network structure. Given the salience of these changes, we should expect changes to the network structure of the triparty repo market. To assess these changes, we estimate the following equation for the two trading relationship measures defined in Equations (4) and (5). For dealer i and MMF j on day t from January 2, 2013 to August 1, 2016, $y_{i,j,t} = (\textit{Share of Business, Frequency of Trading})_{i,j,t}$.

$$\begin{aligned} y_{i,j,t} = \beta_0 &+ \beta_1 1(i = Foreign)_i \times 1(t = RRP1)_t \\ &+ \beta_2 1(i = Foreign)_i \times 1(t = RRP2)_t \\ &+ \beta_3 1(i = Foreign)_i \times 1(t = Implemented)_t \\ &+ \theta_{1i,j} + \phi_t + \delta_{1i,t-1} + \delta_{2j,t-1} + \epsilon_{i,j,t} \end{aligned} \tag{9}$$

$1(i = \textit{Foreign})_i$ equals 1 if dealer i is a foreign dealer, and 0 otherwise; $1(t = \textit{RRP}1)_t$ equals 1 between September 23, 2013 and December 23, 2013, and 0 otherwise; $1(t = \textit{RRP}2)_t$ equals 1 between December 23, 2013 and June 15, 2014, and 0 otherwise; $1(t = \textit{Implemented})_t$ equals 1 after June 15, 2014, and 0 otherwise, labeling the implementation date of Basel III regulations. We include relationship fixed effects, $\theta_{1_{i,j}}$, daily time fixed effects ϕ_t, and the borrower and lender control variables, $\delta_{1i,t-1}$ and $\delta_{2j,t-1}$, respectively, as defined in Section 3.1. Borrower (dealer) controls include last quarter's *STFD* and *Total Assets*. Lender (MMF) level controls are last month's AUM, Treasury repo outstanding, and Treasury securities held. Standard errors are clustered at the relationship level.

Table C.1 presents the results from the estimation of Equation (9) from the perspective of MMF j. Columns 1 and 2 show the results for $\textit{Share of Business}_{i,j,t}$ and columns 3 and 4 show the results for $\textit{Frequency of Trading}_{i,j,t}$. From column 2 we see that ineligible MMFs lent 2.1 percentage points less to foreign dealers after Basel III was implemented. In the post-Basel III era, relationships between ineligible MMFs and foreign dealers weakened.

Table C.1. Effect of RRP on Relationship Strength: MMF Perspective

	Eligible MMF (1)	**Ineligible MMF (2)**	**Eligible MMF (3)**	**Ineligible MMF (4)**
Foreign × RRP1	0.006	–0.004	–0.005	0.001
	(0.63)	(–0.56)	(–0.71)	(0.11)
Foreign × RRP2	0.003	–0.002	–0.005	–0.003
	(0.23)	(–0.27)	(–0.42)	(–0.58)
Foreign × Implemented	0.009	–0.021***	–0.004	–0.010
	(0.87)	(–2.59)	(–0.59)	(–1.48)
Lender's Share of Business (–1)	0.655***	0.740***		
	(28.14)	(35.05)		
Lender's Freq.			0.649***	0.757***
			(16.76)	(30.03)
Controls	Yes	Yes	Yes	Yes
Observations	37,560	65,220	37,560	65,220
R^2	0.7998	0.8425	0.8372	0.8826

Note: *Foreign* equals 1 if dealer i is a foreign dealer, and 0 otherwise. RRP1 equals 1 between September 23, 2013 and December 23, 2013. RRP2 equals 1 between December 23, 2013 and June 15, 2014. *Implemented* equals 1 after June 15, 2014 to capture the Basel III implementation period. We include all foreign and domestic dealers. The daily sample runs from January 2, 2013, to August 1, 2016, and is obtained from FRBNY. ***, **, and * indicate statistical significance at the 1 percent, 5 percent, and 10 percent level, respectively. t-ratios are reported in parentheses. Standard errors are clustered at the relationship level.

Table C.2. Effect of RRP on Relationship Strength: Dealer Perspective

	Eligible MMF (1)	**Ineligible MMF (2)**	**Eligible MMF (3)**	**Ineligible MMF (4)**
Foreign × RRP1	0.006	0.000	0.004	0.002
	(0.67)	(0.05)	(0.87)	(1.27)
Foreign × RRP2	0.014*	0.004*	0.006	0.007**
	(1.66)	(1.76)	(1.20)	(2.54)
Foreign. × Implemented	0.023***	0.003	0.010**	0.007***
	(2.61)	(0.78)	(2.21)	(2.79)
Dealer's Share of Business (–1)	0.713***	0.680***		
	(22.03)	(14.91)		
Dealer's Freq.			0.656***	0.603***
			(21.23)	(10.70)
Controls	Yes	Yes	Yes	Yes
Observations	37,560	65,220	37,560	65,220
R^2	0.8755	0.8280	0.9222	0.8566

Note: *Foreign* equals 1 if dealer i is a foreign dealer, and 0 otherwise. RRP1 equals 1 between September 23, 2013 and December 23, 2013. RRP2 equals 1 between December 23, 2013 and June 15, 2014. *Implemented* equals 1 after June 15, 2014 to reflect the approximate implementation date of Basel III regulations. We include all foreign and domestic dealers. The daily sample runs from January 2, 2013, to August 1, 2016, and is obtained from FRBNY. ***, **, and * indicate statistical significance at the 1 percent, 5 percent, and 10 percent level, respectively. t-ratios are reported in parentheses. Standard errors are clustered at the relationship level.

Table C.2 presents the results from the estimation of Equation (9) from the perspective of dealer i. Columns 1 and 2 show the results for *Share of Business*$_{i,j,t}$ and columns 3 and 4 show the results for *Frequency of Trading*$_{i,j,t}$. From column 1, we observe some evidence of foreign dealers concentrating their borrowing from eligible MMFs. The reliance of foreign dealers on eligible MMF funding in the post-Basel III era was likely a contributing factor of why window dressing became more expensive for non-U.S. dealers.

References

Adrian, T., M. Fleming, J. Goldberg, M. Lewis, F. M. Natalucci, and J. J. Wu. 2013. "Dealer Balance Sheet Capacity and Market Liquidity during the 2013 Selloff in Fixed Income Markets." FEDS Notes, Board of Governors of the Federal Reserve System.

Adrian, T., and H. S. Shin. 2010. "Liquidity and Leverage." *Journal of Financial Intermediation* 19 (3): 418–37.

———. 2011. "Financial Intermediary Balance Sheet Management." Staff Report No. 532, Federal Reserve Bank of New York.

Afonso, G., A. Kovner, and A. Schoar. 2014. "Trading Partners in the Interbank Lending Market." Staff Report No. 620, Federal Reserve Bank of New York.

Agarwal, V., G. D. Gay, and L. Ling. 2014. "Window Dressing in Mutual Funds." *Review of Financial Studies* 27 (11): 3133–70.

Allen, L., and A. Saunders. 1992. "Bank Window Dressing: Theory and Evidence." *Journal of Banking and Finance* 16 (3): 585–623.

Anderson, A., and J. Kandrac. 2018. "Monetary Policy Implementation and Financial Vulnerability: Evidence from the Overnight Reverse Repurchase Facility." *Review of Financial Studies* 31 (9): 3643–86.

Babus, A., and T.-W. Hu. 2017. "Endogenous Intermediation in Over-the-Counter Markets." *Journal of Financial Economics* 125 (1): 200–215.

Bharath, S. T., S. Dahiya, A. Saunders, and A. Srinivasan. 2011. "Lending Relationships and Loan Contract Terms." *Review of Financial Studies* 24 (4): 1141–1203.

Chernenko, S., and A. Sunderam. 2014. "Frictions in Shadow Banking: Evidence from the Lending Behavior of Money Market Funds." *Review of Financial Studies* 27 (6): 1717–50.

Cocco, J. F., F. J. Gomes, and N. C. Martins. 2009. “Lending Relationships in the Interbank Market.” *Journal of Financial Intermediation* 18 (1): 24–48.

Copeland, A., A. Martin, and M. Walker. 2014. “Repo Runs: Evidence from the Tri-Party Repo Market.” *Journal of Finance* 69 (6): 2343–80.

Dass, N., and M. Massa. 2011. “The Impact of a Strong Bank-Firm Relationship on the Borrowing Firm.” *Review of Financial Studies* 24 (4): 1204–60.

Duffie, D., N. Gârleanu, and L. H. Pedersen. 2005. “Over-the-Counter Markets.” *Econometrica* 73 (6): 1815–47.

———. 2007. “Valuation in Over-the-Counter Markets.” *Review of Financial Studies* 20 (6): 1865–1900.

Elliott, G., T. J. Rothenberg, and J. H. Stock. 1996. “Efficient Tests for an Autoregressive Unit Root.” *Econometrica* 64 (4): 813–36.

Frost, J., L. Logan, A. Martin, P. McCabe, F. Natalucci, and J. Remache. 2015. “Overnight RRP Operations as a Monetary Policy Tool: Some Design Considerations.” Finance and Economics Discussion Series No. 2015-010, Board of Governors of the Federal Reserve System.

Gofman, M. 2017. “Efficiency and Stability of a Financial Architecture with Too-Interconnected-to-Fail Institutions.” *Journal of Financial Economics* 124 (1): 113–46.

Gorton, G., and A. Metrick. 2012. “Securitized Banking and the Run on Repo.” *Journal of Financial Economics* 104 (3): 425–51.

Han, S., and K. Nikolaou. 2016. “Trading Relationships in the OTC Market for Secured Claims: Evidence from Triparty Repos.” Finance and Economics Discussion Series No. 2016-064, Board of Governors of the Federal Reserve System.

Hansen, B. E. 2001. “The New Econometrics of Structural Change: Dating Breaks in U.S. Labour Productivity” *Journal of Economic Perspectives* 15 (4): 117–28.

Hendershott, T., D. Li, D. Livdan, and N. Schurhoff. 2016. “Relationship Trading in OTC Markets.” Paper presented at the American Finance Association 2016 Meeting, San Francisco, CA, January 3–5.

Ihrig, J. E., G. C. Weinbach, and E. E. Meade. 2015. “Rewriting Monetary Policy 101: What’s the Fed’s Preferred Post-Crisis

Approach to Raising Interest Rates?" *Journal of Economic Perspectives* 29 (4): 177–98.

Klee, E., Z. Senyuz, and E. Yoldas. 2019. "Effects of Changing Monetary and Regulatory Policy on Money Markets." *International Journal of Central Banking* 15 (4, October): 165–205.

Krishnamurthy, A., S. Nagel, and D. Orlov. 2014. "Sizing Up Repo." *Journal of Finance* 69 (6): 2381–2417.

Lakonishok, J., A. Shleifer, R. Thaler, and R. Vishny. 1991. "Window Dressing by Pension Fund Managers." *American Economic Review: Papers and Proceedings* 81 (May): 227–31.

Li, D., and N. Schürhoff. 2019. "Dealer Networks." *Journal of Finance* 74 (1): 91–144.

Macchiavelli, M., and X. A. Zhou. 2019. "Funding Liquidity and Market Liquidity: The Broker-Dealer Perspective." http://dx.doi.org/10.2139/ssrn.3311786.

Maggio, M. D., F. Franzoni, A. Kermani, and C. Sommavilla. 2019. "The Relevance of Broker Networks for Information Diffusion in the Stock Market." *Journal of Financial Economics* 134 (2): 419–46.

Maggio, M. D., A. Kermani, and Z. Song. 2017. "The Value of Trading Relations in Turbulent Times." *Journal of Financial Economics* 124 (2): 266–84.

Martin, A., D. Skeie, and E.-L. von Thadden. 2014. "Repo Runs." *Review of Financial Studies* 27 (4): 957–89.

Munyan, B. 2015. "Regulatory Arbitrage in Repo Markets." Working Paper No. 15-22, Office of Financial Research.

Neklyudov, A. 2019. "Bid-Ask Spreads and the Over-the-Counter Interdealer Markets: Core and Peripheral Dealers." *Review of Economic Dynamics* 33 (July): 57–84.

Ng, S., and P. Perron. 2001. "Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power." *Econometrica* 69 (6): 1519–54.

Perron, P. 1989. "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis." *Econometrica* 57 (6): 1361–1401.

Sias, R. W., and L. T. Starks. 1997. "Institutions and Individuals at the Turn-of-the-Year." *Journal of Finance* 52 (4): 1543–62.

What Drives Dollar Funding Stress in Distress?*

Yuewen Tang[a] and Alfred Wong[b]
[a]Hong Kong Monetary Authority
[b]Hong Kong Institute for Monetary and Financial Research

We study the forces driving dollar funding stress under adverse market conditions for Asia-Pacific economies. We find that the response of dollar funding conditions to changes in macrofinancial variables differs significantly between orderly and turbulent markets. In orderly markets, idiosyncratic dollar strength, currency volatility, and monetary policy divergence are key factors affecting the stress for the economy. Currency expectations and FX market liquidity also play an important role in determining long-term funding pressure. In turbulent markets, the effect of these variables—except idiosyncratic dollar strength and currency volatility, which retain a strong influence—diminishes or even vanishes. Instead, the creditworthiness of the government and corporate sectors, which is found to have little impact under normal market conditions, emerges as a major stress determinant, and becomes increasingly influential as adversity intensifies.

JEL Codes: C31, E44, F31, G15.

1. Introduction

Social distancing measures have been an inevitable policy response in practically all countries in combating the spread of the COVID-19 virus. In cities and countries where the spread takes hold, the

*The authors are grateful to two anonymous referees for insightful comments. They also thank Adolfo Barajas, Tom Fong, Giorgio Valente, Eric Wong, and Jiayue Zhang for invaluable comments and discussions. Any remaining errors and the views expressed are the authors' own and do not represent those of the Hong Kong Monetary Authority, Hong Kong Academy of Finance, or Hong Kong Institute for Monetary and Financial Research. Author e-mails: ytang@hkma.gov.hk and alfred_yt_wong@hkma.gov.hk.

authorities have had no alternative but to enforce lockdowns, bringing all social activities to a standstill and paralyzing the economy. As a result, the world has slipped rapidly into an economic abyss, the deepest since the Great Depression in the 1930s.[1]

In view of the sharply deteriorating economic outlook for the United States and globally, the Federal Reserve eased monetary policy aggressively and took extraordinary steps to cushion the economy and provide sufficient liquidity for U.S. financial institutions and corporations.[2] Given the unrivaled role the U.S. dollar played in facilitating international finance and trade, the availability and adequacy of dollar funding were equally important for financial institutions and corporations outside the United States to weather the global health crisis. However, the actions of the Federal Reserve did not help them initially. Shortly after the announcement on March 11, 2020 by the World Health Organization (WHO) that the outbreak had evolved into a pandemic, a severe shortage of U.S. dollars quickly developed internationally, imposing significant dollar funding stress on the global economy (Avdjiev, Egemen, and McGuire 2020).

Dollar funding conditions have eased after the Federal Reserve reinvigorated its existing swap line arrangements with major central banks and introduced new ones with the central banks of some emerging markets.[3] However, although a global economic recovery is now under way, uncertainty still looms large given that the vaccination process is likely to take a protracted period of time while the virus continues to mutate. Adding to the uncertainty is the recent

[1]The global economy registered a dismal –3.3 percent growth in 2020 against the backdrop of what is sometimes referred to as the Great Lockdown (https://blogs.imf.org/2020/06/24/reopening-from-the-great-lockdown-uneven-and-uncertain-recovery/).

[2]The Federal Open Market Committee decided to lower the target range for the federal funds rate by 1/2 percentage point to 1 to 1-1/4 percent on March 3, 2020 (https://www.federalreserve.gov/newsevents/pressreleases/monetary20200303a.htm) and by 1 full percentage point to 0 to 1/4 percent on March 15, 2020 (https://www.federalreserve.gov/newsevents/pressreleases/monetary20200315a.htm).

[3]According to the Board of Governors of the Federal Reserve System (2020), the Bank of Canada, the Bank of England, the Bank of Japan, the European Central Bank, the Federal Reserve, and the Swiss National Bank have taken coordinated actions to enhance the provision of liquidity via the standing U.S. dollar liquidity swap line arrangements. In addition, the Federal Reserve has established temporary dollar liquidity swap lines with nine additional foreign central banks.

sharp escalation of the United States–China tensions over a wide range of geopolitical issues. Could this lead to a potential reduction in U.S. dollar supply globally or to some economies? Are we already in the run-up to more significant global financial turmoil than in March 2020 when the WHO declared a pandemic? There seems to be a pressing need for policymakers to understand better the driving forces behind dollar funding stress so that they can be more prepared when the next crisis hits. This is particularly true for Asia Pacific, given the region's considerably larger share of economic activity accounted for by global supply chains and its traditionally stronger demand for dollar funding and higher funding cost sensitivity (Hong et al. 2019).

Research on the driving forces behind dollar funding stress is not new. However, much of it is focused on explaining what causes or sustains it—in particular, why the phenomenon has persisted for years after the 2007/08 global financial crisis and the European debt crisis (Baba and Packer 2009b; Ivashina, Scharfstein, and Stein 2015; Borio et al. 2016; Sushko et al. 2017; Du, Tepper, and Verdelhan 2018). No study has so far attempted to identify the forces in crisis times. Also, the spotlight of the research always follows the major currencies, especially those in Europe. However, the dollar also plays an undisputable and dominant role as the primary reserve currency, cross-border financing currency, and cross-border trade invoicing currency in Asia Pacific.[4] The importance of the potential implications has made EMEAP (2020) turn the spotlight to the region with an intention to better understand the role of the dollar in cross-border financing activities and its impact on financial stability.

This paper contributes to the literature as a first attempt to study the dynamics of these forces under extreme market situations, focusing on eight EMEAP economies. We first examine the normal market relationship between dollar funding stress and a number of potentially determining factors, leveraging on the findings of

[4]In addition, according to EMEAP (2020), around 70 percent of outstanding international debt securities and more than half of banks' cross-border claims and liabilities in the region are denominated in U.S. dollars at end-2019. EMEAP stands for the Executives' Meeting of East Asia-Pacific Central Banks, a forum launched in 1991 to strengthen the cooperative relationship among central banks in the East Asia and Pacific region (https://www.emeap.org/).

previous studies. We then estimate the distressed market relationship between them to uncover any behavioral changes.

Interestingly, we find that bilateral dollar strength, rather than the strength of the dollar globally, plays a much more important role in determining dollar funding stress for the economies in the region, a result that is in contrast with Avdjiev et al. (2019), who find that idiosyncratic currency strength plays, if any, a passive role. Currency volatility, currency expectations, and foreign exchange (FX) market liquidity, which reflect various risks pertaining to dollar strength on a bilateral basis, and monetary policy divergence between the United States and the economy concerned also contribute to the stress. However, these drivers, except idiosyncratic currency strength and currency volatility, are left on the back burner by the dollar lender in extending credit in financial turmoil, especially under the most extreme scenario. Instead, the credit risk of the government and corporate sectors, which has little impact on both the short- and long-term stress during normal market times, emerges as a primary consideration for the dollar lender in turbulent markets. This suggests that borrowers from economies with a weaker fiscal position or higher corporate debt may face greater challenges in dollar funding markets in crisis times. Overall, the results are more aligned with those of the studies that emphasize currency volatility (Liao and Zhang 2020) and credit channels (Hartley 2020; Liao 2020).

The rest of the paper is organized as follows. In the next section we explain and define what we mean by dollar funding stress. Section 3 discusses the potential factors driving dollar funding stress. In Section 4 we set out the model, its specifications, and the data used for estimation. Section 5 presents and discusses our results. Section 6 concludes the paper with a brief discussion of the policy implications.

2. What Is Dollar Funding Stress?

Cost of USD borrowing per se should not be taken to mean or reflect dollar funding stress. Otherwise, most funding markets would have registered a reduction, rather than an increase, in the stress at the peak of the outbreak in early 2020. For example, the three-month USD LIBOR, the cost of borrowing in the interbank money market, in fact fell substantially, but the fall was mainly attributed to the large reductions in the target federal funds rate by the Federal

Reserve in March 2020.[5] It is also debatable to label larger spreads of the LIBOR over its respective overnight indexed swap (OIS) rate as evidence for higher dollar funding stress. As OIS is practically risk-free, the LIBOR-OIS spread reflects mostly the premium the lender charges for lending on an uncollateralized basis.[6] Other things being equal, this premium would thus increase as the economic environment deteriorates and vice versa. True, when the premium increases, it pushes up the LIBOR, the total cost of USD borrowing. However, a higher cost resulting from a larger premium to compensate the lender for taking a higher risk is not what the stress refers to here.

Dollar funding stress refers to what is above and beyond the cost of borrowing in the LIBOR market. The interest rates that matter to the borrower are the ones at which he can borrow. The well-known problem with the LIBOR market is that it is not a market that is always accessible to all.[7] For those shut out of the interbank money market in times of financial turbulence, the LIBOR is simply irrelevant. Under these circumstances, it is the market they turn to which matters. And the most popular alternative for them is probably the cross-currency swap market, in which they could swap their domestic currency for U.S. dollars, effectively borrowing U.S. dollars by pledging another currency as collateral.

The more stressful the situation is, the more financial institutions and corporations would be forced to resort to the market for USD funding. Other things being equal, this would push up the cost of USD borrowing in the cross-currency swap market. Theoretically, the cost of borrowing should equalize across markets under covered interest parity. Any difference in the borrowing cost between the LIBOR and cross-currency swap markets cannot sustain even for minute periods of time, as traders and other market participants

[5] The LIBOR, the London interbank offered rate, remains the most popular benchmark cost of borrowing in the interbank money market despite its widely known problems (King and Lewis 2020).

[6] A LIBOR transaction is one in which a party lends to another without any collateral, but an OIS—an interest rate swap between a fixed rate and a predetermined floating daily overnight rate—is a transaction that does not involve lending or exchange of principals.

[7] This may be best exemplified by the difficulties experienced by European financial institutions during the 2007/08 global financial crisis (Baba and Packer 2009b).

would arbitrage it away. However, following the 2007/08 global financial crisis, the pressure has grown so much that the cost of borrowing U.S. dollars in the cross-currency swap market now often exceeds that in the interbank money market, making possible material violations to the parity condition.

The resulting deviation, which often occurs in favor of the U.S. dollar, is widely interpreted as an indication of how much the market is under stress in its hunger for U.S. dollars.[8] The level of stress can thus be measured by the deviation, that is, the extent to which the cost of USD borrowing in the cross-currency swap market exceeds the cost of USD borrowing in the interbank money market. For short-term dollar funding, the dollar funding stress can be measured by the FX swap basis, which is the difference between the USD LIBOR and the implied USD interest rate in the FX swap transaction. For long-term dollar funding, it can be gauged directly by the basis traded in the cross-currency basis swap (CCBS) market.[9] The more negative the FX swap or CCBS basis, the greater is the USD shortage and the higher is the dollar funding stress.

In March 2020, the three-month FX swap basis increased 100 basis points for the Japanese yen, 54 for the euro, 63 for the British pound, and 91 for the Swiss franc within one week following the WHO announcement, while the five-year CCBS basis widened 13 basis points for the Japanese yen, 4 for the euro, 5 for the British pound, and 6 for the Swiss franc. In emerging Asia, economies also took a similar beating, reaching levels unseen in recent years (Figures 1 and 2). The same week saw, for instance, the three-month FX swap basis rise 266 basis points for the Singapore dollar, 9 for the Hong Kong dollar, and 161 for the Korean won, whereas there was

[8]There are, however, competing theories arguing that when counterparty credit risk is present (which is the case after the global financial crisis, as evidenced by the LIBOR-OIS spread for most currencies), the deviation merely reflects that interest rates for unsecured lending/borrowing are no longer appropriate for pricing the secured transactions in cross-currency swap markets (Wong and Zhang 2018).

[9]In both cases, the counterparty credit risk involved in the transaction is minimal, as the USD borrowing party needs to pledge another currency as collateral. At the same time, the funding liquidity risks of the two parties are swapped, where the USD lending party takes the funding liquidity risk of USD while the USD borrowing party takes the funding liquidity risk of the other currency.

Figure 1. Three-Month FX Swap Bases of EMEAP Currencies, Jan. 2017–Mar. 2021

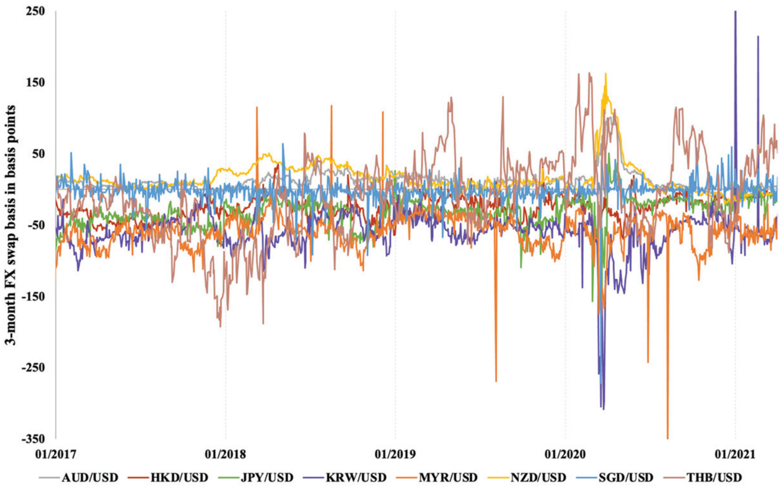
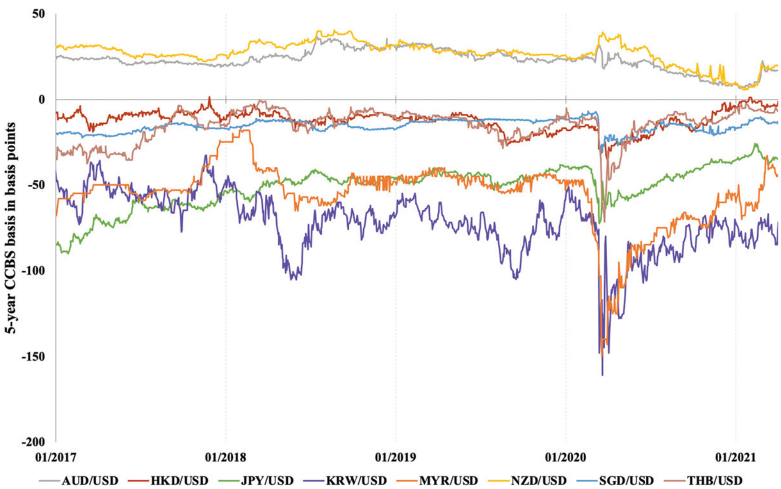


Figure 2. Five-Year CCBS Bases of EMEAP Currencies, Jan. 2017–Mar. 2021



a widening of the basis by 17, 10, and 72 basis points, respectively, for these currencies in the CCBS market.

3. Potential Driving Forces of Dollar Funding Stress

Cross-currency bases for most currencies vis-à-vis the U.S. dollar are in favor of the lender of U.S. dollars. It is common to see, for example, in CCBS transactions, that the dollar-lending party effectively borrows foreign currency at a discount to the foreign currency LIBOR and this discount is the basis. There are only a few exceptions where the deviations are in favor of the foreign currency lender, e.g., the Australian dollar and the New Zealand dollar.[10] However, apart from these rare cases, the fact that the higher implied dollar interest rates compared with the LIBOR in cross-currency swap transactions are generally taken as a sign of global shortage of U.S. dollars and, hence, also the amount of stress associated with the shortage.

Since the covered interest parity condition broke down following the onset of the global financial crisis in 2007/08, a large volume of literature has surfaced, offering a wide range of explanations and theories about the phenomenon. This study is not another attempt to add more explanations or theories to the debate or argue which explanation or theory is more credible. Instead, we wish to take into account the existing explanations and theories and adopt a more pragmatic approach to tackling the question by examining the potential driving forces empirically. Consequently, we consider the candidates for which data, especially daily data, are available.[11] We have identified the following list of variables and discuss each of them briefly below.

- *Global Dollar Strength.* Avdjiev et al. (2019) find the global strength of the U.S. dollar to be an important driving force

[10] In addition, there are a small number of other currencies, such as the British pound, whose cross-currency bases occasionally reverse themselves although they are on balance in favor of the dollar.

[11] As a result, institutional factors that may aggravate or contain the stress such as central bank swap lines are not considered, although the effect of these factors would filter through to some of the macrofinancial variables in the list below (Bahaj and Reis 2018).

behind the currency swap bases for major currencies, as the effect of a stronger dollar feeds through to the shadow price of cross-border bank leverage for non-U.S. financial institutions. This study employs the same proxy they use for global dollar strength, the broad dollar index vis-à-vis the currencies of a large group of U.S. trading partners compiled by the Federal Reserve Board.

- *Idiosyncratic Dollar Strength.* By the same token, the strength of the dollar vis-à-vis the currency concerned may arguably be a more relevant shadow price for the financial institutions in the economy concerned. Idiosyncratic dollar strength here refers to the strength of the dollar against the currency not attributable to the global strength of the dollar. It is represented by the residual obtained by regressing the bilateral exchange rate of each of the currencies vis-à-vis the U.S. dollar (indexed to share the same base with the broader dollar index) on the broad dollar index.[12]
- *Currency Volatility.* If dollar strength affects the shadow price of cross-border bank leverage, the risk to its stability would naturally be a source of concern. This price is likely to increase in times of market turbulence as hedging demand grows and the use of FX derivatives intensifies (Liao and Zhang 2020). Consequently, other things being equal, higher currency volatility may result in a higher price premium to pay for the certainty of the cost of dollar funding.
- *Currency Expectations.* Similarly, the expected appreciation (or depreciation) of the exchange rate of the U.S. dollar vis-à-vis the currency concerned is also likely to play a role. The risk reversal of the currency, which is the difference in implied volatility between the call and put currency options, is used to proxy currency expectations.[13]

[12]We have also used the bilateral exchange rate instead of the idiosyncratic dollar strength in our estimation. The results, which can be available upon request, are similar.

[13]The price of an option reflects the market expectation of the likelihood of an adverse outturn happening. A call option gives the right to buy the asset at a certain price and a put option offers the right to sell. Hence, the buyer of a call bets on the asset to rise above the strike price within a certain period, while the seller thinks it may not and accepts a payment for taking the risk. A put

- *FX Market Liquidity.* Arai et al. (2016) argue that regulatory reforms make global banks less reluctant to engage in market making and arbitrage trading in the FX swap market, while Krohn and Sushko (2020) find that deterioration in dollar funding is linked to a reduction in market liquidity in the spot FX market. It is thus reasonable to consider bid-ask spreads in both the spot and forward markets as potential candidates.
- *Financial Market Volatility.* Stock market volatility is often regarded as a signal of global banks' leverage cycle that drives cross-border fund flows and hence global liquidity conditions (Borio and Disyatat 2011; Forbes and Warnock 2012; Obstfeld 2012a, 2012b; Bruno and Shin 2015; Rey 2015). We adopt the widely used forward-looking S&P 500 volatility measure as a gauge of market risk or "fear" to capture the impact of market sentiment on dollar funding conditions (Whaley 2000; Giot 2005).
- *Interest Differential.* Arai et al. (2016) observe that corporate borrowers tend to arbitrage any interest differential by issuing bonds denominated in currencies with a lower yield to hedge their proceeds back via FX swaps, thus putting upward pressure on the higher-yielding currency in the cross-currency swap market. Du, Tepper, and Verdelhan (2018) also find that cross-currency bases tend to be positively correlated with the level of nominal interest rates and increase with interest rate shocks.
- *Monetary Policy Divergence.* A larger "interest margin" or steeper yield curve for a currency is more conducive to funding investment denominated in the currency by other currencies (with a smaller "interest margin" or a flatter yield curve) through the cross-currency swap market, causing a greater demand for the currency and pushing its basis up (Iida, Kimura, and Sudo 2018). Therefore, as the yield curve

option works exactly the other way round. However, the prices of the call and the put are not necessarily the same, as there may be heavier betting for a rise in the asset than for a fall, or the other way round. Hence, the price difference can measure how asymmetric the market is in expecting a rise and a fall in the asset. See Wong and Fong (2018) for a more detailed discussion.

is essentially a function of monetary policy, term spread differential, which reflects the monetary policy divergence between two countries, can affect funding pressure in the swap market (Borio et al. 2016).

- *Credit Spread.* Last, but not least, counterparty credit risk is often cited as a prominent reason for the emergence and persistence of cross-currency bases in crisis periods (Baba and Packer 2009a; Coffey, Hrung, and Sarkar 2009; Ivashina, Scharfstein, and Stein 2015; Wong and Zhang 2018). The credit spread of the sovereign and corporate USD bonds of the economy (over U.S. Treasury securities) is used to account for the credit risk of the non-U.S. borrowers in the region.[14]

4. Model Specification and Data

In light of the potential driving forces identified above, our baseline regression model is specified as follows:

$$\begin{aligned} y_{it} = {} & \alpha_i + \beta_1 DollarG_t + \beta_2 DollarI_{it} + \beta_3 VOL_{it} + \beta_4 RiskRev_{it} \\ & + \beta_5 BASprd_{it} + \beta_6 BASprd3M_{it} + \beta_7 VIX_t \\ & + \beta_8 IntDiff_{it} + \beta_9 TermSprd_{it} + \beta_{10} CreditSprd_{it}, \end{aligned}$$

where

y_{it}	denotes the dollar funding stress for economy i as represented by the three-month FX swap or five-year CCBS basis of currency i vis-à-vis the U.S. dollar;
α_i	is the fixed effect of currency i;
$DollarG_t$	stands for the U.S. trade-weighted broad dollar index compiled by the Federal Reserve Board;
$DollarI_{it}$	is the residual obtained by regressing the exchange rate of currency i vis-à-vis the U.S. dollar on $DollarG_t$;

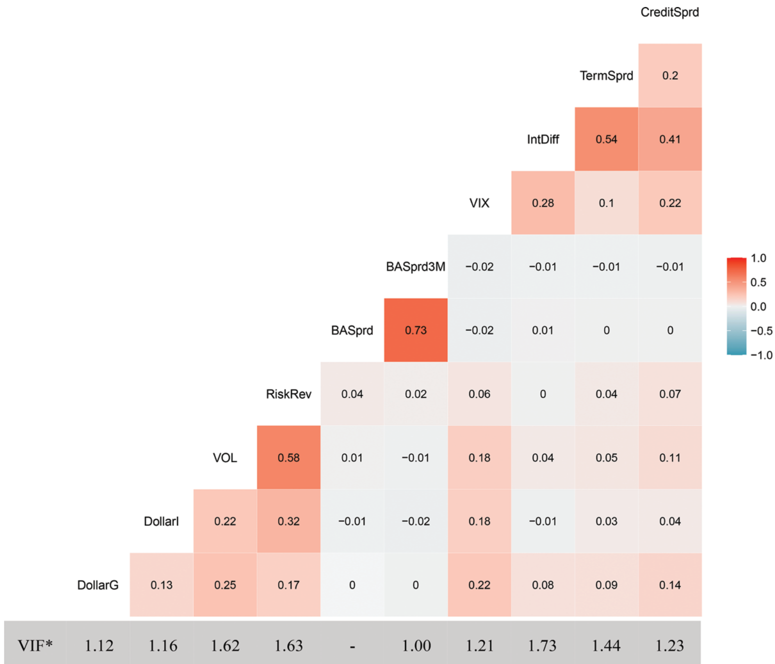
[14] We assume that the difference between the true credit risk and the credit risk as implied by credit spread, if any, is largely stable (due, for example, to stable investor preferences towards taking the credit risk of holding the bonds) such that the change in credit spread is a reasonable approximation of the change in the true credit risk premium.

VOL_{it}	represents the three-month 25-delta FX call option implied volatility of the exchange rate of currency i vis-à-vis the U.S. dollar;
$RiskRev_{it}$	denotes the three-month 25-delta FX option risk reversal of the exchange rate of currency i vis-à-vis the U.S. dollar;
$BASprd_{it}$	represents the bid-ask spread of the spot exchange rate of currency i vis-à-vis the U.S. dollar;
$BASprd3M_{it}$	denotes the bid-ask spread of the three-month forward exchange rate of currency i vis-à-vis the U.S. dollar;
VIX_t	is the 30-day forward-looking Volatility Index compiled by the Chicago Board Options Exchange;
$IntDiff_{it}$	represents the spread of the 10-year government bond yield of currency i over the 10-year U.S. Treasury bond yield;
$TermSprd_{it}$	measures the 10-year over 2-year spread differential between currency i government bond and U.S. Treasury markets; and
$CreditSprd_{it}$	represents the JP Morgan global aggregate bond credit spread index, which measures the spread of U.S.-dollar denominated sovereign and corporate bond yields of economy i over U.S. treasury yields.

All the variables take the form of their first difference. For simplicity, the three-month and five-year funding markets are chosen to represent the short- and long-term markets, respectively. All data are daily, covering the period from January 2007 to March 2021. There are 11 EMEAP economies, but data availability allows us to study only 8 of them, namely, Australia, Hong Kong, Japan, Korea, Malaysia, New Zealand, Singapore, and Thailand. The sources and statistical characteristics of the data are summarized in Appendix A.

Before estimating the model, we test the data for multicollinearity between the independent variables, which could potentially compromise the efficiency of the estimation. Figure 3 shows the correlation matrix of the independent variables. The correlation coefficient between *BASprd* and *BASprd3M* is 0.73, which indicates a very high positive correlation, though this should not be surprising (Krohn and Sushko 2020). In view of this, we need to drop one of

Figure 3. Correlation Matrix and Variance Inflation Factors of Explanatory Variables



*The variance inflation factor (VIF) is estimated after *BASprd* is removed.

them. We remove *BASprd* given the intuition that dollar funding in the cross-currency swap market is probably associated more with the liquidity in the forward market than in the spot market, but it does not really matter which one we retain given such a high correlation. The correlation coefficients of 0.58 between *VOL* and *RiskRev* and 0.54 between *IntDiff* and *TermSprd* are less than desired but acceptable in view of their variance inflation factors (VIFs).[15] The bottom

[15]The VIF is a measure of how much the variance (i.e., the standard error squared) of the estimated coefficient is inflated due to the existence of correlation among the independent variables in the regression. Specifically, the VIF for the *jth* independent variable is the reciprocal of $(1 - R_j^2)$, where R_j^2 is the R^2 value obtained by regressing the *jth* independent variable on the remaining independent variables. A VIF of one means that there is no correlation between

row in the figure shows that the VIF, a test of multicollinearity, for each remaining independent variable is way below five.

5. Empirical Results

To examine the effect of the potential candidates on dollar funding stress, we first estimate their normal market relationship in both the short- and long-term funding markets by regular linear regressions. We then move on to assess their distressed market relationship, employing the technique of quantile regression.

The estimates of the linear regressions, which are essentially least-squares based, represent the relationships between the FX swap or CCBS basis and the various driving forces under normal market conditions. Each of them is the average, or the conditional mean to be exact, relationship over a long period of time, more than 14 years in the present study. However, the relationship of interest to policymakers, e.g., those relationships described in Section 3, can possibly be buried in the data covering protracted and uneventful periods, in which non-U.S. financial institutions generally experience little dollar funding pressure. As a result, there should be little surprise that some of these estimates turn out to be insignificant or sometimes even carry the wrong sign. As we shall see, it happens to this study, just as it happens to many previous studies (e.g., Avdjiev et al. 2019; Barajas et al. 2020; Cerutti, Obstfeld, and Zhou 2021).

The distinct advantage of quantile regression is that it can enable us to evaluate the relationship between the dependent and independent variables across different quantiles of the conditional distribution of the dependent variable, which can be taken to represent different market conditions. Hence, this technique makes it possible to estimate the response of the basis to any potential factor under extreme scenarios, thereby helping uncover the relationship that is relevant and more important for policymaking. Appendix B explains why and how quantile regression can do it while regular linear regression cannot.

---

the *jth* independent variable and the remaining independent variables, since the variance of b_j is not inflated at all. As a rule of thumb, a VIF greater than five indicates a problem of multicollinearity (Craney and Surles 2007).

5.1 Relationship in Normal Markets

Tables 1 and 2 present six sets of results for the short- and long-term dollar funding stress, respectively, under normal market conditions. In each of the tables, the second column lists the expected signs for the variables for ease of reference. The next six columns present the estimates obtained from (i) simple pooled regression; (ii) simple pooled regression with a lagged term of the dependent variable for correcting the first-order serial correlation; (iii) panel regression with currency fixed effects and a lagged term of the dependent variable for correcting the first-order serial correlation; (iv) panel regression with currency random effects and a lagged term of the dependent variable for correcting the first-order serial correlation; (v) pooled regression with the Newey-West (1987) robust covariance matric estimator to obtain the heteroskedasticity and autocorrelation corrected (HAC) standard errors; and (vi) pooled regression with the Driscoll-Kraay (1998) method to correct both cross-sectional and serial correlation to obtain the spatial correlation consistent (SCC) standard errors.

When the simple pooled regression model is estimated, diagnostic tests suggest the presence of potential serial correlation and cross-sectional correlation problems in both the short-term and long-term dollar funding stress equations. We thus reestimate the pooled model, and the panel models, with a lagged term. The fixed-effects model assumes that currency-specific effects are correlated with the independent variables (i.e., a group-specific fixed quantity), while the random effects model assumes currency-specific effects are uncorrelated (i.e., a random sample from the population). However, we find that both the fixed effects and random effects are insignificant at the 90 percent confidence level, suggesting that a simple pooled model is perhaps a better choice.[16] Indeed, as can be seen, there is practically no difference in the estimates between the models with and without currency-specific effects. When more sophisticated techniques are employed to correct for heteroskedasticity, and higher-order and cross-sectional autocorrelation, the results are of similar flavor, with some variables losing their statistical significance to different extents.

[16]The pooled regression model does not consider heterogeneity across groups or time.

Table 1. Short-Term Dollar Funding Stress: Currency-Specific Effects and Pooled Panel Data Regressions, January 2007–March 2021

	Expected Sign	Pooled	Pooled	Fixed	Random	HAC	SCC
(Intercept)		0.070	0.092		0.092	0.070	0.070
		(0.222)	(0.211)		(0.211)	(0.129)	(0.042)
lag(FXSwap)			**–0.311*** ***	**–0.311*** ***	**–0.311*** ***		
			(0.006)	(0.006)	(0.006)		
DollarG	–	–0.197	0.391	0.391	0.391	–0.197	–0.197
		(0.676)	(0.643)	(0.643)	(0.643)	(1.945)	(1.633)
DollarI	–	**–3.636*** ***	**–3.599*** ***	**–3.601*** ***	**–3.599*** ***	**–3.636*** *	**–3.636*** ***
		(0.439)	(0.417)	(0.417)	(0.417)	(1.418)	(1.014)
VOL	–	**–4.340*** ***	**–5.390*** ***	**–5.389*** ***	**–5.390*** ***	–4.340	**–4.340*** **
		(0.647)	(0.615)	(0.615)	(0.615)	(2.465)	(1.613)
RiskRev	–	**5.064*** **	**3.569*** *	**3.570*** *	**3.569*** *	5.064	**5.064*** *
		(1.544)	(1.468)	(1.468)	(1.468)	(8.895)	(2.382)
BASprd3M	–	**0.065*** ***	**0.050*** ***	**0.050*** ***	**0.050*** ***	0.065	**0.065*** ***
		(0.012)	(0.011)	(0.011)	(0.011)	(0.035)	(0.009)
VIX	–	0.082	0.112	0.112	0.112	0.082	0.082
		(0.126)	(0.120)	(0.120)	(0.120)	(0.281)	(0.074)
IntDiff	+	4.409	6.064	6.072	6.064	4.409	4.409
		(4.494)	(4.271)	(4.272)	(4.271)	(9.174)	(2.444)

(*continued*)

Table 1. (Continued)

	Expected Sign	Pooled	Pooled	Fixed	Random	HAC	SCC
TermSprd	+	**13.768***	**13.351***	**13.327***	**13.351***	13.768	**13.768***
		(5.726)	(5.442)	(5.443)	(5.442)	(12.148)	(5.751)
CreditSprd	−	**0.135***	0.104	0.104	0.104	0.135	0.135
		(0.056)	(0.054)	(0.054)	(0.054)	(0.117)	(0.090)
R^2		0.006	0.103	0.103	0.103		
Adj. R^2		0.006	0.102	0.102	0.102		
Num. Obs.		28,957	28,956	28,956	28,956		
Adequacy Test		DW = 1.951; Pesaran CD Test: p-value < 0.001		F-test: p-value = 0.946	Breusch-Pagan LM Test: p-value = 0.151		

Note: The third and fourth columns present the results of the pooled regression model without and with the lagged dependent variable, respectively. A Durbin–Watson statistic of 1.951 (with a p-value of less than 0.001) indicates the presence of positive serial correlations. A Pesaran (2004) cross-sectional dependence test yielding a p-value of less than 0.001 rejects the null hypothesis that the residuals are not correlated across the currencies. The fifth and sixth columns present the results of the fixed-effects and random-effects regression models. An F-test yielding a p-value of 0.946 fails to reject the null hypothesis that the fixed effects are insignificant, while a Breusch-Pagan (1980) Lagrange multiplier test yielding a p-value of 0.151 cannot reject the null hypothesis that the variances of the random effects are zero at the 90 percent confidence level. In the seventh column the Newey-West (1987) robust covariance matrix estimator is used to correct for serial correlation to obtain the heteroskedasticity and autocorrelation corrected (HAC) standard errors. In the eighth column, the Driscoll-Kraay (1998) method is used to correct both cross-sectional and serial correlation to obtain the spatial correlation consistent (SCC) standard errors. ***, **, and * denote that the estimated coefficient is statistically significant at 0.1 percent, 1 percent, and 5 percent, respectively.

Table 2. Long-Term Dollar Funding Stress: Currency-Specific Effects and Pooled Panel Data Regressions, January 2007–March 2021

	Expected Sign	Pooled	Pooled	Fixed	Random	HAC	SCC
(Intercept)		–0.002	–0.002		–0.002	–0.002	**–0.002***
		(0.021)	(0.021)		(0.021)	(0.019)	(0.001)
lag(CCBS)			**–0.062*****	**–0.062*****	**–0.062*****		
			(0.006)	(0.006)	(0.006)		
DollarG	–	–0.031	–0.031	–0.031	–0.031	–0.031	–0.031
		(0.063)	(0.062)	(0.062)	(0.062)	(0.105)	(0.070)
DollarI	–	**–0.226*****	**–0.208*****	**–0.208*****	**–0.208*****	**–0.226***	**–0.226****
		(0.040)	(0.040)	(0.040)	(0.040)	(0.111)	(0.073)
VOL	–	**–1.240*****	**–1.216*****	**–1.216*****	**–1.216*****	**–1.240*****	**–1.240*****
		(0.060)	(0.060)	(0.060)	(0.060)	(0.277)	(0.299)
RiskRev	–	**–0.867*****	**–0.904*****	**–0.904*****	**–0.904*****	–0.867	**–0.867*****
		(0.143)	(0.143)	(0.143)	(0.143)	(0.666)	(0.143)
BASprd3M	–	–0.002	**–0.002***	**–0.002***	**–0.002***	–0.002	**–0.002***
		(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
VIX	–	**0.052*****	**0.053*****	**0.053*****	**0.053*****	**0.052***	0.052
		(0.012)	(0.012)	(0.012)	(0.012)	(0.022)	(0.029)
IntDiff	+	**–2.709*****	**–2.655*****	**–2.655*****	**–2.655*****	**–2.709*****	**–2.709***
		(0.405)	(0.404)	(0.404)	(0.404)	(0.783)	(1.107)

(*continued*)

Table 2. (Continued)

	Expected Sign	Pooled	Pooled	Fixed	Random	HAC	SCC
TermSprd	+	**1.785*****	**1.739*****	**1.739*****	**1.739*****	1.785	**1.785****
		(0.515)	(0.514)	(0.514)	(0.514)	(1.078)	(0.569)
CreditSprd	−	0.002	0.002	0.002	0.002	0.002	0.002
		(0.005)	(0.005)	(0.005)	(0.005)	(0.009)	(0.002)
R^2		0.037	0.041	0.041	0.041		
Adj. R^2		0.037	0.041	0.041	0.041		
Num. Obs.		29,111	29,110	29,110	29,110		
Adequacy Test		DW = 1.962; Pesaran CD Test: p-value < 0.001		F-test: p-value = 1.000	Breusch-Pagan LM Test: p-value = 0.046		

Note: The third and fourth columns present the results of the pooled regression model without and with the lagged dependent variable, respectively. A Durbin–Watson statistic of 1.951 (with a p-value of less than 0.001) indicates the presence of positive serial correlations. A Pesaran (2004) cross-sectional dependence test yielding a p-value of less than 0.001 rejects the null hypothesis that the residuals are not correlated across the currencies. The fifth and sixth columns present the results of the fixed-effects and random-effects regression models. An F-test yielding a p-value of 1 fails to reject the null hypothesis that the fixed effects are insignificant, while a Breusch-Pagan (1980) Lagrange multiplier test yielding a p-value of 0.046 cannot reject the null hypothesis that the variances of the random effects are zero at the 90 percent confidence level. In the seventh column the Newey-West (1987) robust covariance matrix estimator is used to correct for serial correlation to obtain the heteroskedasticity and autocorrelation corrected (HAC) standard errors. In the eighth column, the Driscoll-Kraay (1998) method is used to correct both cross-sectional and serial correlation to obtain the spatial correlation consistent (SCC) standard errors. ***, **, and * denote that the estimated coefficient is statistically significant at 0.1 percent, 1 percent, and 5 percent, respectively.

Looking at the overall picture, somewhat surprising to us is the coefficient of *DollarG*, which is found to be insignificant for both short- and long-term stress across different models and estimation techniques. On the contrary, the coefficient of *DollarI*, which measures dollar strength vis-à-vis the currency concerned that is not attributable to global dollar strength, turns out to be significant and carries a negative sign as expected in both the short- and long-term stress equations. This suggests that it is idiosyncratic, rather than global, dollar strength that causes dollar funding stress to the economy, a result that is at odds with that of Avdjiev et al. (2019). Nonetheless, it seems to make more sense than otherwise, given that dollar funding stress is largely an adversity suffered by non-U.S. financial institutions. From their perspective, when the U.S. dollar is weak against other major currencies, these borrowers are still likely to find it more difficult and costly to obtain dollar funding if their own currency is even weaker. The reason is that a larger amount of their own currency would be required as collateral for obtaining the same amount of dollar funding. We find that the risks pertaining to the stability of dollar strength vis-à-vis individual currencies are also important, as evidenced by the highly significant and negative coefficient of *VOL* in both the short- and long-term stress equations. This suggests that dollar funding stress in general tends to be associated with a volatile exchange rate. *RiskRev* is found to affect short-term stress positively and affect long-term stress negatively, though the impact is insignificant when it is estimated under HAC.

Turning to the variables other than those related to dollar strength, the effect of FX market liquidity is found to be positive on short-term dollar funding stress but insignificant under HAC. There is some negative, but very mild, impact on long-term dollar funding stress. General financial market volatility as proxied by *VIX* is also found to have little effect on short-term stress; it positively affects long-term stress, but the impact becomes much less significant under HAC and insignificant under SCC. Long-term interest differential as represented by *IntDiff* supposedly has a positive influence on dollar funding stress but is found to have no impact on short-term stress. Like Avdjiev et al. (2019), we find that it has a negative influence on long-term stress. *TermSprd*, which denotes the relative stance of monetary policy of the economy vis-à-vis the United States, is found to have a significant positive impact on both short- and long-term stress as expected. Finally, *CreditSprd* is found to have little influence on both.

For robustness and comparison, we present the results of the estimation of the short-term impact employing one-month and six-month FX swap bases and those of the long-term impact using one-year and three-year CCBS bases in Appendix C. Broadly speaking, they are highly consistent with those presented in the above. The major difference is that the magnitude of the coefficient is generally larger at the shorter end of the funding market, which is not surprising given the larger fluctuation of the data.

5.2 Relationship in Extreme Markets

While it is important to know the long-term driving forces behind what seems to be an intriguing global phenomenon, policymakers would probably find it more useful to understand the dynamics underpinning the phenomenon in times of market stress. This can be achieved with the aid of quantile regression.

Regardless of how the dependent variable responds to the independent variables under normal market conditions, it can behave quite differently in stressful times. As discussed, quantile regression can help us estimate the response of dollar funding stress to the driving forces under extreme market scenarios. In this study, the extremity of the scenarios is defined by dollar funding stress set progressively at the 25 percent, 20 percent, 10 percent, and 5 percent quantiles of its conditional distribution given that the greater the dollar funding stress, the more negative (or the less positive) is the FX swap or CCBS basis. The results of the quantile regressions for the three-month FX swap and five-year CCBS bases are presented in a progressive manner in Tables 3 and 4, respectively. As can be seen, compared with the least-squares estimates, the results of the quantile regressions apparently seem to be more clear-cut, especially when we move along the extremity scale. In the most extreme situation, no estimate which is found to be significant carries a wrong sign. These findings are highly robust across the basis spectrum, with the bases of other maturities showing consistent results (Appendix C). There are four points we wish to highlight.

Firstly, similar to what is found for the orderly market, idiosyncratic dollar strength is also an important driving force behind

Table 3. Short-Term Dollar Funding Stress: Quantile Regressions, January 2007–March 2021

	Expected Sign	Quantile			
		25%	**20%**	**10%**	**5%**
(Intercept)		**–4.067*****	**–5.599*****	**–12.001*****	**–22.113*****
		(0.070)	(0.089)	(0.211)	(0.485)
lag(FXSwap)		**–0.262*****	**–0.270*****	**–0.274*****	**–0.306****
		(0.001)	(0.002)	(0.005)	(0.005)
DollarG	–	**0.515****	0.392	**1.140***	1.140
		(0.166)	(0.222)	(0.571)	(1.346)
DollarI	–	**–0.619*****	**–0.684*****	**–1.395*****	**–3.011*****
		(0.137)	(0.110)	(0.326)	(0.786)
VOL	–	**–1.288*****	**–1.435*****	**–2.628*****	**–4.266*****
		(0.189)	(0.213)	(0.538)	(1.066)
RiskRev	–	–0.174	–0.422	–1.471	–3.754
		(0.514)	(0.639)	(1.477)	(3.232)
BASprd3M	–	**0.023*****	**0.024*****	**0.022***	0.026
		(0.001)	(0.003)	(0.010)	(0.020)
VIX	–	**0.073***	**0.089***	0.085	0.430
		(0.037)	(0.039)	(0.101)	(0.225)
IntDiff	+	**5.811*****	**6.756*****	**10.313****	**22.794****
		(1.207)	(1.518)	(3.600)	(8.509)
TermSprd	+	**–3.264***	–2.231	6.318	15.966
		(1.534)	(1.820)	(3.878)	(10.698)
CreditSprd	–	**–0.077*****	**–0.113*****	**–0.177*****	**–0.359****
		(0.016)	(0.020)	(0.052)	(0.116)
Num. Obs.		28,956	28,956	28,956	28,956

Note: ***, **, and * denote the estimated coefficient is statistically significant at 0.1 percent, 1 percent, and 5 percent, respectively.

both short- and long-term funding stress in the turbulent market with the coefficient of *DollarI* being negative and significant at all the quantiles. This suggests that as the market becomes chaotic, a stronger dollar vis-à-vis the currency concerned inflicts more stress on the borrower in the region. On the other hand, *DollarG* has a positive and negative impact on short-term and long-term dollar funding stress, respectively, as turbulence picks up initially. However, as we move towards the most extreme market, the impact vanishes in statistical significance in both cases, suggesting that global dollar

Table 4. Long-Term Dollar Funding Stress: Quantile Regressions, January 2007–March 2021

	Expected Sign	Quantile			
		25%	**20%**	**10%**	**5%**
(Intercept)		**–0.468*****	**–0.767*****	**–1.982*****	**–3.663*****
		(0.011)	(0.019)	(0.035)	(0.084)
lag(CCBS)		**–0.05*****	**–0.064*****	**–0.069*****	**–0.067****
		(0.002)	(0.004)	(0.007)	(0.020)
DollarG	–	**–0.061***	–0.056	–0.136	–0.361
		(0.027)	(0.053)	(0.086)	(0.227)
DollarI	–	**–0.060*****	**–0.061***	**–0.197*****	**–0.315****
		(0.016)	(0.029)	(0.059)	(0.112)
VOL	–	**–0.338*****	**–0.464*****	**–0.800*****	**–1.043*****
		(0.025)	(0.030)	(0.099)	(0.196)
RiskRev	–	**–0.268*****	**–0.239*****	**–0.515***	–0.992
		(0.079)	(0.072)	(0.253)	(0.522)
BASprd3M	–	0.000	0.000	–0.002	–0.002
		(0.000)	(0.001)	(0.002)	(0.002)
VIX	–	**0.017*****	**0.023****	**0.039***	0.064
		(0.005)	(0.009)	(0.017)	(0.040)
IntDiff	+	–0.145	–0.263	–1.161	–2.153
		(0.185)	(0.334)	(0.593)	(1.447)
TermSprd	+	0.106	0.148	1.278	3.384
		(0.226)	(0.440)	(0.711)	(1.849)
CreditSprd	–	**–0.005***	**–0.010****	**–0.020*****	**–0.033***
		(0.002)	(0.004)	(0.004)	(0.016)
Num. Obs.		29,110	29,110	29,110	29,110

Note: ***, **, and * denote the estimated coefficient is statistically significant at 0.1 percent, 1 percent, and 5 percent, respectively.

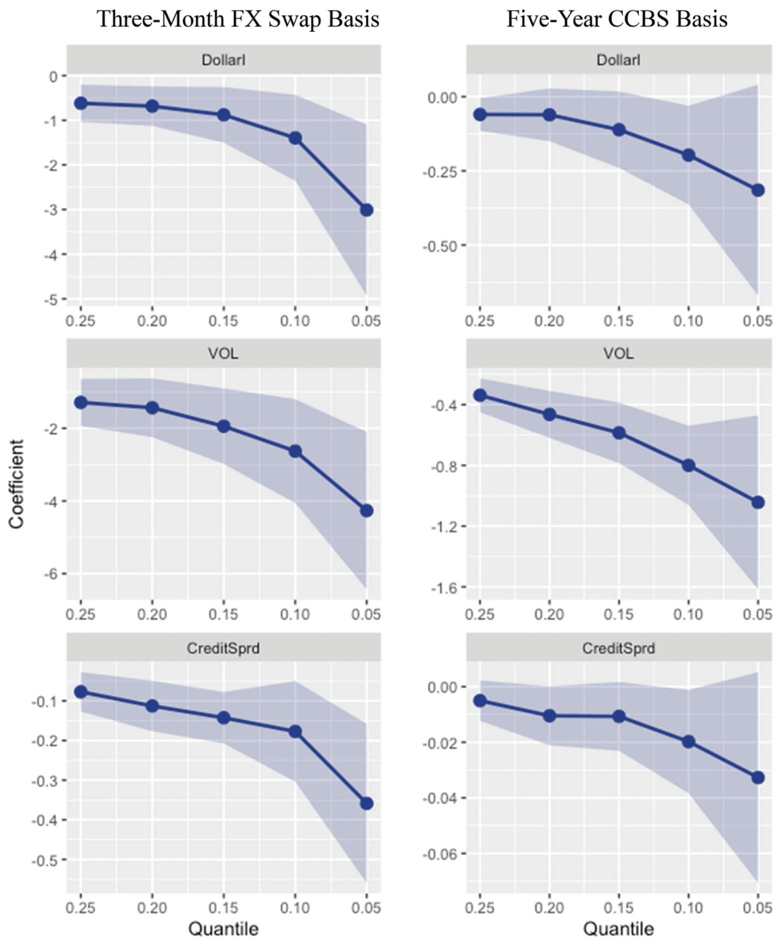
strength is irrelevant for the determination of dollar funding stress in crisis times as well as in normal markets.

Second, the volatility of dollar strength vis-à-vis individual currencies continues to play a critical role in determining both the FX swap and CCBS bases under extreme scenarios, with a consistently negative quantile estimate found for both short- and long-term stress. Exchange rate expectations play a much smaller role in stressful situations compared with normal market times. *RiskRev*, which can be interpreted as the expected dollar strength vis-à-vis the local currency, is found to have no impact on short-term stress. It maintains its impact on long-term stress initially as we move from the simple pooled regression to the quantile regressions, but the impact also disappears when we reach the lowest quantile in the estimation.

Third, the positive effect of *BASprd3M* on short-term stress found in the normal market lingers into the beginning of a turbulent market, but its statistical significance diminishes gradually towards the lowest quantile; it has no impact on long-term stress. Similarly, the coefficient of *VIX* is found to be positive amid increasing turbulence in both the short- and long-term equations but also becomes statistically insignificant at the lowest quantile. These findings suggest that FX market liquidity and general financial market volatility do not contribute to dollar funding stress, at least not under the most extreme scenario. Interestingly, *IntDiff,* which is found to have no impact on short-term stress in the orderly market, now shows up with a significant and increasing impact alongside the degree of market extremity. It has little impact on long-term stress, however, under extreme scenarios. On the contrary, *TermSprd*, which affects both the FX swap and CCBS bases significantly in the normal market, is found to have little influence over both short- and long-term funding stress, meaning that relative monetary conditions are only an important driver in the long run but not at critical moments. Most interesting to us is *CreditSprd*, a variable that is found to have no effect on short- or long-term stress at all during normal times, shows up as an important factor affecting dollar funding stress in both the short- and long-term markets as market conditions deteriorate. This means counterparty risk is an important element from the perspective of the dollar lending party, as reflected by the steeper compensation it demands from the borrower in turbulent times as compared with what would normally be required in a quiet market.

Finally, it is worth noting that the response of the variables that are found to be significant tends to intensify as turbulence gathers momentum. For example, the coefficient of *VOL* becomes more negative as we move towards a lower quantile in the estimation, reflecting that the more volatile the currency, the larger is the impact on the stress for the economy. Figure 4 provides a graphical exposition of the results of the various quantiles for *DollarI*, *VOL,* and *CreditSprd* to illustrate visually the extent to which the response exacerbates as market conditions worsen.

Figure 4. Response Sensitivity at Various Quantiles



Note: The shaded area represents the 95 percent confidence interval.

6. Concluding Remarks

Overall, the results of our estimation suggest that macrofinancial variables tend to behave quite differently in terms of how they affect dollar funding stress for the Asia-Pacific economies, as compared with what is found by previous studies. To some extent, this may be attributed to the fact that most previous studies are centered on the most advanced economies, while the economies under study here are a much more diverse group. The results also highlight the importance of differentiating the responses of the stress between normal and extreme market circumstances for policymaking and market surveillance.

Some recent studies have identified global dollar strength as probably the single most important factor that drives dollar funding stress. However, we find that it plays little role in determining both short- and long-term stress faced by EMEAP dollar borrowers in normal markets. It adds to long-term borrowing stress when turbulence begins to pick up, but the effect also dissipates as adversity deepens. On the contrary, idiosyncratic dollar strength is a major source of dollar funding pressure, regardless of whether it is in the short- or long-term market and irrespective of market conditions. In addition, uncertainty about dollar strength against individual currencies is found to be important both in orderly markets and in times of crisis. Currency expectations also play a role in the long-term market during normal market times, but their impact fades gradually when turbulence intensifies. Dollar funding stress also depends critically on monetary policy divergence over the long haul but not during crisis times.

Our findings suggest that credit risk which does not affect dollar funding stress in normal markets is an important consideration for the dollar lender in extending credit in turbulent times. This means that under stressful scenarios, economies that suffer a sharper deterioration in the credit outlook for their government, banks, and corporations (due possibly to a larger public debt or heavy borrowing) are likely to experience tighter funding conditions.

These results lend support to the policy of the Federal Reserve on establishing USD swap lines with other central banks. It was timely that the Federal Reserve extended the arrangement to nine more central banks early in the health crisis.[17] It is well known that some of the economies concerned have very volatile currencies, and most of them were expected to suffer a severe fiscal setback at the onset of

[17]On March 20, 2020, the arrangement of swaps lines was extended to nine more central banks, comprising US$60 billion each with the Monetary Authority of Singapore, Reserve Bank of Australia, Banco Central do Brasil, Danmarks Nationalbank, Bank of Korea, and Banco de Mexico, and US$30 billion each with the Reserve Bank of New Zealand, Norges Bank, and Sveriges Riksbank (https://www.federalreserve.gov/monetarypolicy/bst_liquidityswaps.htm).

the pandemic, which could trigger significant credit risk reappraisal for both the sovereign and financial institutions. In broadening the coverage of the facility further in the future, the Federal Reserve may wish to give more consideration to economies where the currencies are more likely to come under pressure and credit conditions tend to be more fragile in turbulent markets.

The findings also provide food for thought for policymakers in the region. For example, instead of monitoring global dollar strength as suggested in the literature, they should perhaps focus more on their own currency movement, volatility, and the market expectations about it. If they are concerned with potential financial contagion from their neighbors, they may also wish to keep a close eye on those who have a larger public debt or heavier corporate borrowing, which could render these economies more susceptible to a major credit risk reappraisal in times of crisis.

Appendix A. Sources and Descriptive Statistics of the Data

Table A.1. Sources of the Data

Variable	Description	Source
FXSwap	Three-month FX swap basis of foreign currency versus U.S. dollar	Bloomberg, RBNZ
CCBS	Five-year CCBS of foreign currency versus U.S. dollar	Bloomberg
DollarG	Federal Reserve Board U.S. trade-weighted broad dollar index	FRB of St. Louis
DollarI	Residual from regressing *DollarG* on bilateral exchange rate	Authors' estimation
VOL	Three-month 25-delta FX call option-implied volatility	JP Morgan database
RiskRev	Three-month 25-delta FX option risk reversal	JP Morgan database
BASprd	Bid-ask spread of spot exchange rate	Bloomberg
BASprd3M	Bid-ask spread of three-month forward exchange rate	Bloomberg
VIX	CBOE Volatility Index	Bloomberg
IntDiff	Yield spread of 10-year foreign govt. over 10-year U.S. Treasury	Bloomberg
TermSprd	10-year over 2-year spread differential (foreign govt. over U.S. Treasury)	Bloomberg
CreditSprd	JP Morgan global aggregate bond credit spread index	JP Morgan database

Table A.2. Descriptive Statistics of the Data

	Min.	Median	Mean	Max.	S.D.	Num. Obs.
FXSwap						
AUD/USD	–130.35	9.25	10.84	241.21	15.04	3,716
HKD/USD	–83.37	–16.68	–17.05	57.36	14.49	3,716
JPY/USD	–256.51	–24.06	–27.48	71.32	21.06	3,716
KRW/USD	–1,761.30	–60.93	–97.50	1,215.67	137.80	3,716
MYR/USD	–753.62	–50.41	–58.31	494.56	73.49	3,715
NZD/USD	–54.43	14.94	17.67	162.50	16.19	3,715
SGD/USD	–271.74	0.42	2.80	301.83	17.68	3,715
THB/USD	–302.16	–13.83	56.19	1,266.14	210.72	3,715
CCBS						
AUD/USD	–50.00	22.63	20.82	48.00	9.56	3,717
HKD/USD	–63.00	–9.00	–8.81	20.50	12.34	3,716
JPY/USD	–102.50	–49.00	–48.86	34.00	25.19	3,717
KRW/USD	–324.00	–76.00	–92.02	5.50	51.59	3,716
MYR/USD	–240.00	–79.00	–85.57	–3.00	41.73	3,715
NZD/USD	–5.50	26.30	24.85	52.00	11.90	3,717
SGD/USD	–69.00	–18.86	–20.78	2.50	12.40	3,716
THB/USD	–205.00	–20.00	–31.31	6.00	34.82	3,717
DollarG						
All Currencies	85.47	97.55	101.84	126.52	10.83	3,717
DollarI						
USD/AUD	–10.89	–0.77	0.06	32.79	6.44	3,717
USD/HKD	–0.55	–0.16	0.00	0.89	0.39	3,717
USD/JPY	–16.85	–0.67	0.01	23.15	9.08	3,717
USD/KRW	–18.54	–0.37	0.03	45.72	8.85	3,716
USD/MYR	–7.14	–0.96	–0.01	11.27	3.54	3,716
USD/NZD	–12.97	–0.98	0.06	41.67	8.22	3,717
USD/SGD	–7.52	–1.02	0.03	12.36	4.71	3,717
USD/THB	–9.06	–0.22	0.04	9.06	4.40	3,717
VOL						
USD/AUD	6.26	12.09	12.76	38.81	4.70	2,716
USD/HKD	0.28	0.77	0.91	5.46	0.56	3,716
USD/JPY	4.43	9.44	9.58	24.47	2.72	3,716
USD/KRW	3.54	10.70	12.39	68.31	7.65	3,716
USD/MYR	2.78	8.36	8.67	20.27	3.40	3,716
USD/NZD	6.72	12.96	13.56	36.22	4.56	3,716
USD/SGD	2.90	6.09	6.57	17.89	2.38	3,716
USD/THB	3.80	6.80	7.51	14.86	2.38	3,716

(*continued*)

Table A.2. (Continued)

	Min.	Median	Mean	Max.	S.D.	Num. Obs.
RiskRev						
USD/AUD	0.11	1.63	1.90	8.25	1.24	3,716
USD/HKD	–1.30	–0.35	–0.31	2.13	0.42	3,716
USD/JPY	–10.07	–1.23	–1.47	1.50	1.58	3,716
USD/KRW	–0.63	1.96	2.55	27.00	2.54	3,716
USD/MYR	–0.85	1.38	1.38	5.61	1.00	3,716
USD/NZD	0.14	1.62	1.89	8.00	1.20	3,716
USD/SGD	–0.68	0.88	0.89	4.47	0.68	3,716
USD/THB	0.27	1.08	1.15	4.63	0.60	3,716
VIX						
All Currencies	9.14	17.23	20.02	82.69	9.66	3,715
BASprd						
USD/AUD	–77.66	2.35	3.94	151.14	7.44	3,717
USD/HKD	0.77	0.78	1.39	32.81	1.34	3,717
USD/JPY	0.76	1.77	2.62	56.51	3.15	3,717
USD/KRW	1.01	16.95	17.26	372.65	25.77	3,716
USD/MYR	–167.61	12.34	13.42	59.91	9.81	3,677
USD/NZD	–1.53	4.31	7.02	473.35	13.34	3,717
USD/SGD	1.22	4.29	5.93	321.64	8.21	3,717
USD/THB	–647.95	8.32	15.38	448.33	29.18	3,717
BASprd3M						
USD/AUD	–72.60	2.68	4.43	168.38	8.10	3,717
USD/HKD	0.13	1.39	1.90	36.66	1.96	3,717
USD/JPY	–1.57	1.39	2.31	60.78	3.73	3,717
USD/KRW	1.59	47.26	63.16	652.27	61.87	3,716
USD/MYR	0.58	12.96	13.72	324.52	12.54	3,715
USD/NZD	–1.07	5.13	7.99	477.36	13.83	3,717
USD/SGD	0.23	5.26	7.70	215.21	10.03	3,717
USD/THB	–548.91	16.54	25.60	451.12	31.25	3,716
IntDiff						
AUD/USD	–0.86	1.14	1.01	2.77	0.93	3,717
HKD/USD	–2.05	–0.59	–0.61	0.28	0.28	3,716
JPY/USD	–3.36	–1.98	–1.92	–0.50	0.64	3,717
KRW/USD	–0.65	0.72	0.78	2.93	0.81	3,716
MYR/USD	–1.77	1.43	1.22	2.82	0.79	3,715
NZD/USD	–0.67	1.62	1.30	3.13	0.97	3,717
SGD/USD	–2.36	–0.34	–0.42	0.78	0.55	3,673
THB/USD	–1.29	0.47	0.55	2.17	0.65	3,715

(*continued*)

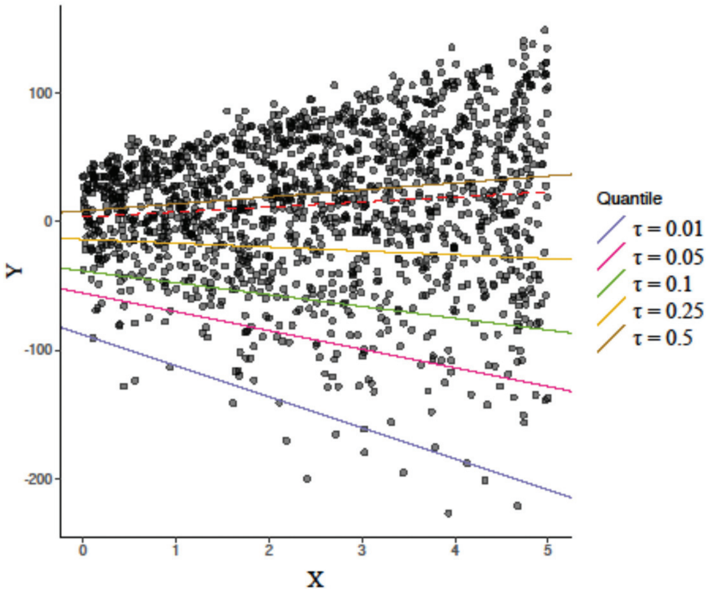
Table A.2. (Continued)

	Min.	**Median**	**Mean**	**Max.**	**S.D.**	**Num. Obs.**
TermSprd						
AUD/USD	–2.51	–0.75	–0.72	0.52	0.73	3,717
HKD/USD	–1.70	–0.40	–0.38	0.52	0.32	3,422
JPY/USD	–2.02	–0.87	–0.82	1.12	0.65	3,717
KRW/USD	–2.20	–0.92	–0.82	1.06	0.65	3,716
MYR/USD	–2.07	–0.67	–0.66	0.45	0.70	3,715
NZD/USD	–2.17	–0.59	–0.58	0.78	0.66	3,717
SGD/USD	–1.15	–0.09	–0.15	1.20	0.33	3,673
THB/USD	–2.52	–0.36	–0.45	1.15	0.74	3,715
CreditSprd						
AUD	27.96	144.79	164.21	470.00	69.20	3,716
HKD	63.75	246.37	253.45	673.31	81.69	3,716
JPY	50.94	124.59	139.45	347.59	49.28	3,716
KRW	49.71	131.03	169.99	706.27	114.14	3,716
MYR	66.00	159.70	175.21	495.97	64.73	3,716
NZD	48.06	112.91	137.25	408.43	72.67	3,716
SGD	59.32	136.94	151.04	466.11	60.64	3,716
THB	137.00	226.00	270.10	839.00	130.16	3,457

Appendix B. Quantile Regression

An ordinary least squares (OLS) regression estimates the mean response of the dependent variable to the independent variables based on the conditional mean function. Hence, this provides only a general or average view of the relationship between them. However, sometimes we are only interested in the relationship at certain points in the conditional distribution of the dependent variable, rather than at the mean. And in some cases, it is possible that a relationship does not exist at the mean at all but only at the tails of the conditional distribution. Quantile regression is an elegant technique of estimating the conditional median (or other quantiles) of the response variable. This technique is appealing due to its robustness to outliers and especially useful in the analysis of extreme events that lie in the high (or low) conditional quantiles for heavy tailed distributions.

Figure B.1. OLS and Quantile Regression Lines



Taking a similar formulation as the classical regression model, the quantile regression model for τth quantile can be written as

$$Q_\tau(y_i|x_1, x_2, \cdots, x_p) = \beta_0(\tau) + \beta_1(\tau)x_1 + \beta_2(\tau)x_2 + \cdots + \beta_p(\tau)x_p.$$

In contrast to being constants in the OLS regression, the beta coefficients are now functions with a dependency on the quantile level τ. The corresponding conditional quantile of y_i given x_p can be written as $Q_\tau(y_i|x_p)$ such that the quantile level τ is the probability of y_i equal to or less than its value estimated by the model, i.e., $Pr(y_i \leq Q_\tau(y_i|x_p)|x_p)$.

Figure B.1 presents an example of regression data for which both the mean and the variance of the response Y increase as the predictor X increases. The dashed line in the middle represents a simple OLS fit. The OLS regression models the conditional mean $E(Y|X)$ but does not capture the conditional variance $Var(Y|X)$. By fitting a series of quantile regression models for a grid of values of τ in the interval (0,1), we can describe the entire conditional distribution of the response. The solid lines in Figure B.1 show the fitted quantile regressions for the quantile levels at 1 percent, 5 percent, 10 percent, 25 percent, and 50 percent. In this particular example, the OLS regression line (the dashed line) conveys little information about the relationship between X and Y, as the fitted regression line has only a slight positive slope and does not describe the increasing dispersion of Y, while the quantile regression lines reveal interesting relationships. As can be seen, the decrease in response Y accelerates along the quantile scale as the predictor X increases, meaning that the relationship becomes more prominent as we move to the lower quantiles. This relationship, which apparently is negative, is not observable at the mean level.

Appendix C. Results of Cross-Currency Bases of Other Maturities

This appendix presents the results obtained by estimating the model using one-month and six-month FX swap bases, and one-year and three-year CCBS bases. Broadly speaking, they are highly consistent with those found for the three-month and five-year bases. As the fixed and random effects are insignificant, Table C.1 only shows the results of the simple pooled regression (with a lagged term of the dependent variable for correcting the first-order serial correlation) for comparison. As can be seen, like what we find for the three-month and five-year bases, *DollarI* and *VOL* are the only two variables that display a negative and significant coefficient across all maturities as expected. The estimates of all other variables can be significant with the right sign, insignificant, or significant with the wrong sign. As discussed in Section 5, this reflects the problem of estimating the mean relationship between the dependent and independent variables over a long period of time characterized mainly by a fairly calm market.

Tables C.2 and C.3 show the results estimated at various quantiles of the one- and six-month FX swap bases and one- and three-year CCBS bases, respectively. Similar to those for the three-month and five-year bases, the problem of ambiguity is gone as one moves towards the lowest quantile. No variable is found to be significant

Table C.1. FX Swap and CCBS Bases of Other Maturities: Pooled Regressions, January 2007–March 2021

	Expected Sign	1-Month	6-Month	1-Year	3-Year
(Intercept)		–0.018	0.002	–0.005	–0.002
		(0.598)	(0.105)	(0.024)	(0.020)
lag(FXSwap/ CCBS)		**–0.366*****	**–0.207*****	**–0.126*****	**–0.038*****
		(0.005)	(0.006)	(0.006)	(0.006)
DollarG	–	–1.226	**–0.864****	**0.263*****	0.040
		(1.815)	(0.317)	(0.072)	(0.060)
DollarI	–	**–11.812*****	**–1.108*****	**–0.243*****	**–0.267*****
		(1.163)	(0.207)	(0.046)	(0.038)
VOL	–	**–12.154*****	**–2.515*****	**–2.116*****	**–1.402*****
		(1.717)	(0.298)	(0.069)	(0.057)
RiskRev	–	**23.635*****	–1.306	**–2.394*****	**–1.078*****
		(4.133)	(0.730)	(0.165)	(0.140)
BASprd3M	–	**0.131*****	0.003	–0.002	–0.001
		(0.031)	(0.006)	(0.001)	(0.001)
VIX	–	**0.867****	0.055	**0.070*****	**0.063*****
		(0.332)	(0.059)	(0.013)	(0.011)
IntDiff	+	7.339	2.829	**–3.229*****	**–3.123*****
		(11.734)	(1.983)	(0.468)	(0.388)
TermSprd	+	3.234	0.564	**3.523*****	**2.180*****
		(14.946)	(2.278)	(0.595)	(0.498)
CreditSprd	–	0.220	0.015	0.008	–0.003
		(0.149)	(0.025)	(0.006)	(0.005)
R^2		0.142	0.051	0.104	0.054
Adj. R^2		0.141	0.050	0.104	0.054
Num. Obs.		29,069	27,603	29,124	28,918

Note: ***, **, and * denote the estimated coefficient is statistically significant at 0.1 percent, 1 percent, and 5 percent, respectively.

with the wrong sign in the most extreme scenario. Again, *DollarI* and *VOL* display consistently their undesirable impact on both short- and long-term dollar funding stress in times of market turmoil as expected, while *CreditSprd,* which is found to have no influence at all in normal markets, emerges as a key determinant in crisis times.

Table C.2. One-Month and Six-Month FX Swap Bases: Quantile Regressions, January 2007–March 2021

		Quantile				Quantile			
		25%	**20%**	**10%**	**5%**	**25%**	**20%**	**10%**	**5%**
	Expected Sign	*1-Month FX Swap Basis*				*6-Month FX Swap Basis*			
(Intercept)		**–9.699*****	**–13.687*****	**–29.930*****	**–56.621*****	**–1.975*****	**–2.717*****	**–5.655*****	**–10.080*****
		(0.181)	(0.239)	(0.552)	(1.323)	(0.034)	(0.043)	(0.091)	(0.229)
lag(FXSwap)		**–0.312*****	**–0.321*****	**–0.349*****	**–0.372*****	**–0.100*****	**–0.108*****	**–0.099*****	**–0.092*****
		(0.003)	(0.004)	(0.004)	(0.007)	(0.002)	(0.002)	(0.005)	(0.010)
DollarG	–	**2.095*****	**2.537*****	**3.204***	6.845	**–0.177****	–0.125	**–0.308***	–0.451
		(0.503)	(0.638)	(1.546)	(3.685)	(0.089)	(0.108)	(0.178)	(0.637)
DollarI	–	**–0.899****	**–1.074***	**–2.556*****	**–7.949*****	**–0.561*****	**–0.607*****	**–0.668*****	**–0.926****
		(0.331)	(0.436)	(0.675)	(2.141)	(0.058)	(0.067)	(0.125)	(0.360)
VOL	–	**–2.732*****	**–3.093*****	**–5.433*****	**–10.719*****	**–1.472*****	**–1.704*****	**–2.124*****	**–3.301*****
		(0.463)	(0.520)	(0.881)	(3.006)	(0.051)	(0.091)	(0.185)	(0.491)
RiskRev	–	1.746	0.410	2.667	5.788	**–0.474*****	**–0.871*****	**–1.458****	**–2.734****
		(1.262)	(1.685)	(2.697)	(7.953)	(0.154)	(0.187)	(0.588)	(1.160)
BASprd3M	–	**0.051*****	**0.055*****	0.040	0.076	**0.005*****	**0.002***	0.004	0.011
		(0.008)	(0.012)	(0.026)	(0.064)	(0.001)	(0.001)	(0.003)	(0.010)
VIX	–	0.144	0.140	0.181	0.339	–0.018	–0.014	0.019	0.124
		(0.085)	(0.116)	(0.283)	(0.629)	(0.016)	(0.019)	(0.020)	(0.087)
IntDiff	+	**14.317*****	**14.255*****	**23.792***	**49.455****	**6.171*****	**7.066*****	**7.519*****	**9.954****
		(3.061)	(4.240)	(9.752)	(18.718)	(0.547)	(0.676)	(1.410)	(4.167)
TermSprd	+	**–9.012***	–4.141	–7.296	8.771	**–5.798*****	**–6.004*****	**–5.057*****	–0.609
		(3.965)	(4.977)	(12.615)	(16.138)	(0.738)	(0.728)	(1.490)	(5.410)
CreditSprd	–	**–0.214*****	**–0.250*****	**–0.337****	**–0.647***	**–0.042*****	**–0.058*****	**–0.072*****	**–0.109****
		(0.043)	(0.055)	(0.122)	(0.302)	(0.006)	(0.009)	(0.016)	(0.051)
Num. Obs.		29,069	29,069	29,069	29,069	27,603	27,603	27,603	27,603

Note: ***, **, and * denote the estimated coefficient is statistically significant at 0.1 percent, 1 percent, and 5 percent, respectively.

Table C.3. One-Year and Three-Year CCBS Bases: Quantile Regressions, January 2007–March 2021

		Quantile				Quantile			
		25%	**20%**	**10%**	**5%**	**25%**	**20%**	**10%**	**5%**
	Expected Sign	*One-Year CCBS Basis*				*Three-Year CCBS Basis*			
(Intercept)		**–0.440*****	**–0.733*****	**–1.911*****	**–3.559*****	**–0.429*****	**–0.702*****	**–1.780*****	**–3.440*****
		(0.012)	(0.018)	(0.034)	(0.078)	(0.012)	(0.018)	(0.037)	(0.080)
lag(CCBS)		**–0.067*****	**–0.076*****	**–0.092*****	**–0.095*****	**–0.051*****	**–0.062*****	**–0.071*****	**–0.053*****
		(0.002)	(0.004)	(0.005)	(0.012)	(0.003)	(0.003)	(0.009)	(0.016)
DollarG	–	0.041	0.061	0.127	0.102	**–0.063***	–0.087	**–0.234***	–0.398
		(0.031)	(0.049)	(0.082)	(0.219)	(0.031)	(0.049)	(0.102)	(0.213)
DollarI	–	**–0.110*****	**–0.147*****	**–0.268*****	**–0.522*****	**–0.067*****	**–0.095****	**–0.236*****	**–0.467*****
		(0.018)	(0.032)	(0.036)	(0.129)	(0.019)	(0.031)	(0.068)	(0.080)
VOL	–	**–0.475*****	**–0.626*****	**–1.190*****	**–1.710*****	**–0.373*****	**–0.502*****	**–0.896*****	**–1.423*****
		(0.029)	(0.047)	(0.046)	(0.207)	(0.029)	(0.047)	(0.101)	(0.183)
RiskRev	–	**–0.290*****	**–0.419*****	**–0.936*****	–0.792	**–0.149***	–0.181	–0.450	–0.492
		(0.054)	(0.123)	(0.113)	(0.589)	(0.076)	(0.129)	(0.232)	(0.479)
BASprd3M	–	**–0.002*****	**–0.002****	**–0.002***	–0.003	**–0.001*****	–0.001	–0.002	0.000
		(0.000)	(0.001)	(0.001)	(0.003)	(0.000)	(0.001)	(0.002)	(0.003)
VIX	–	0.011	**0.023****	**0.051*****	0.075	0.008	0.015	0.035	0.052
		(0.006)	(0.009)	(0.013)	(0.041)	(0.005)	(0.008)	(0.020)	(0.043)
IntDiff	+	**–0.615*****	**–1.163*****	**–1.927*****	–2.089	–0.023	–0.077	–0.774	–1.660
		(0.167)	(0.259)	(0.491)	(1.353)	(0.179)	(0.290)	(0.669)	(1.449)
TermSprd	+	**0.773****	**1.323*****	**2.895*****	**3.708***	0.075	0.202	1.518	**4.196***
		(0.240)	(0.340)	(0.697)	(1.640)	(0.222)	(0.375)	(0.779)	(1.846)
CreditSprd	–	**–0.005***	**–0.009***	**–0.019*****	**–0.036***	**–0.016*****	**–0.020*****	**–0.036*****	**–0.057***
		(0.002)	(0.004)	(0.005)	(0.017)	(0.003)	(0.004)	(0.010)	(0.022)
Num. Obs.		29,124	29,124	29,124	29,124	28,918	28,918	28,918	28,918

Note: ***, **, and * denote the estimated coefficient is statistically significant at 0.1 percent, 1 percent, and 5 percent, respectively.

References

Arai, F., Y. Makabe, Y. Okawara, and T. Nagano. 2016. "Recent Trends in Cross-Currency Basis." Bank of Japan Review No. 2016-E-7.

Avdjiev, S., W. Du, D. Koch, and H. S. Shin. 2019. "The Dollar, Bank Leverage, and Deviations from Covered Interest Parity." *American Economic Review: Insights* 1 (2): 193–208.

Avdjiev, S., E. Egemen, and P. McGuire. 2020. "Dollar Funding Costs during the Covid-19 Crisis through the Lens of the FX Swap Market." BIS Bulletin No. 1.

Baba, N., and F. Packer. 2009a. "From Turmoil to Crisis: Dislocations in the FX Swap Market Before and After the Failure of Lehman Brothers." *Journal of International Money and Finance* 28 (8): 1350–74.

———. 2009b. "Interpreting Deviations from Covered Interest Parity during the Financial Market Turmoil of 2007–08." *Journal of Banking and Finance* 33 (11): 1953–62.

Bahaj, S., and R. Reis. 2018. "Central Bank Swap Lines." CEPR Discussion Paper No. 13003.

Barajas, A., A. Deghi, C. Raddatz, D. Seneviratne, P. Xie, and Y. Xu. 2020. "Global Banks' Dollar Funding: A Source of Financial Vulnerability." IMF Working Paper No. 20/113.

Board of Governors of the Federal Reserve System. 2020. "Central Bank Liquidity Swaps." https://www.federalreserve.gov/monetarypolicy/central-bank-liquidity-swaps.htm.

Borio, C., and P. Disyatat. 2011. "Global Imbalances and the Financial Crisis: Link or No Link?" BIS Working Paper No. 346.

Borio, C., R. N. McCauley, P. McGuire, and V. Sushko. 2016. "Covered Interest Parity Lost: Understanding the Cross-Currency Basis." *BIS Quarterly Review* (September): 45–64.

Breusch, T., and A. Pagan. 1980. "The Lagrange Multiplier Test and Its Applications to Model Specification in Econometrics." *Review of Economic Studies* 47 (1): 239–53.

Bruno, V., and H. S. Shin. 2015. "Capital Flows and the Risk-taking Channel of Monetary Policy." *Journal of Monetary Economics* 71 (April): 119–32.

Cerutti, E. M., M. Obstfeld, and H. Zhou. 2021. "Covered Interest Parity Deviations: Macrofinancial Determinants." *Journal of International Economics* 130 (May): Article 103447.

Coffey, N., W. Hrung, and A. Sarkar. 2009. "Capital Constraints, Counterparty Risk, and Deviations from Covered Interest Rate Parity." Staff Report No. 393, Federal Reserve Bank of New York.

Craney, T. A., and J. Surles. 2007. "Model-Dependent Variance Inflation Factor Cutoff Values." *Quality Engineering* 14 (3): 391–403.

Driscoll, J. C., and A. C. Kraay. 1998. "Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data." *Review of Economics and Statistics* 80 (4): 549–60.

Du, W., A. Tepper, and A. Verdelhan. 2018. "Deviations from Covered Interest Rate Parity." *Journal of Finance* 73 (3): 915–57.

EMEAP. 2020. "Study on US Dollar Liquidity and Funding Dynamics in the EMEAP Region." Report by EMEAP Working Group on Financial Markets, August.

Forbes, K. J., and F. E. Warnock. 2012. "Capital Flow Waves: Surges, Stops, Flight and Retrenchment." *Journal of International Economics* 88 (2): 235–51.

Giot, P. 2005. "Relationships between Implied Volatility Indexes and Stock Index Returns." *Journal of Portfolio Management* 31 (3): 92–100.

Hartley, J. S. 2020. "Covered Interest Rate Parity Deviations in External Emerging Market Sovereign Debt." *Journal of Fixed Income* 29 (4): 92–99.

Hong, G. H., A. Oeking, K. H. Kang, and C. Rhee. 2019. "What Do Deviations from Covered Interest Parity and Higher FX Hedging Costs Mean for Asia?" IMF Working Paper No. 19/169.

Iida, T., T. Kimura, and N. Sudo. 2018. "Deviations from Covered Interest Rate Parity and the Dollar Funding of Global Banks." *International Journal of Central Banking* 14 (4, September): 275–325.

Ivashina, V., D. S. Scharfstein, and J. C. Stein. 2015. "Dollar Funding and the Lending Behavior of Global Banks." *Quarterly Journal of Economics* 130 (3): 1241–81.

King, T. B., and K. F. Lewis. 2020. "Credit Risk, Liquidity, and Lies." *International Journal of Central Banking* 16 (5, October): 219–68.

Krohn, I., and V. Sushko. 2020. "FX Spot and Swap Market Liquidity Spillovers." BIS Working Paper No. 836.

Liao, G. Y. 2020. "Credit Migration and Covered Interest Rate Parity." *Journal of Financial Economics* 138 (2): 504–25.

Liao, G. Y., and T. Zhang. 2020. "The Hedging Channel of Exchange Rate Determination." International Finance Discussion Paper No. 1283, Board of Governors of the Federal Reserve System.

Newey, W. K., and K. D. West. 1987. "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica* 55 (3): 703–8.

Obstfeld, M. 2012a. "Does the Current Account Still Matter?" *American Economic Review* 102 (3): 1–23.

———. 2012b. "Financial Flows, Financial Crises, and Global Imbalances." *Journal of International Money and Finance* 31 (3): 469–80.

Pesaran, M. H. 2004. "General Diagnostic Tests for Cross-Section Dependence in Panels." *Journal of Econometrics* 69 (7): 393–414.

Rey, H. 2015. "Dilemma Not Trilemma: The Global Financial Cycle and Monetary Policy Independence." NBER Working Paper No. 21162.

Sushko, V., C. Borio, R. McCauley, and P. McGuire. 2017. "The Failure of Covered Interest Parity: FX Hedging Demand and Costly Balance Sheets." BIS Working Paper No. 590.

Whaley, R. E. 2000. "The Investor Fear Gauge." *Journal of Portfolio Management* 26 (3): 12–17.

Wong, A., and T. Fong. 2018. "Safehavenness of Currencies." *European Journal of Finance* 24 (4): 300–332.

Wong, A., and J. Zhang. 2018. "Breakdown of Covered Interest Parity: Mystery and Myth." *BIS Papers* 96 (March): 57–78.

Withering Cash: Is Sweden Ahead of the Curve or Just Special?*

Hanna Armelius,[a] Carl Andreas Claussen,[b] and André Reslow[b]

[a]Confederation of Swedish Enterprise

[b]Sveriges Riksbank

Cash in circulation has increased in most countries but has fallen dramatically in Sweden. We explore the drivers behind this development using panel data consisting of 129 countries. In line with the previous literature, we find that GDP, demography, and the interest rate are key explanatory variables. A new finding is that lower corruption is associated with lower demand for cash in developed countries. Our empirical model performs relatively well in explaining the developments in most OECD countries. However, our model cannot explain the divergent Swedish development. We argue that a unique combination of events and policy measures have led to the decline of cash in Sweden.

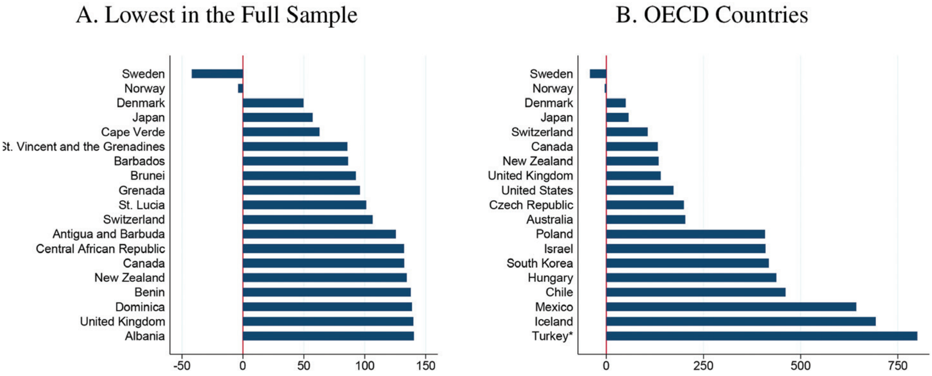
JEL Codes: E41, E42, E51.

1. Introduction

There is much in our increasingly digitized economies to suggest that the use of (physical) cash should be falling. For example, the number of online purchases is increasing; digital payments at physical points of sale are widespread; and payment applications for smartphones

*We would like to thank Niklas Amberg, Thomas Jansson, Franz Seitz, two anonymous referees, and the participants in the seminars held at the Sveriges Riksbank for comments and suggestions. Armelius conducted most of her work on this paper while employed at Sveriges Riksbank. The opinions expressed in this article are the sole responsibility of the authors and should not be interpreted as reflecting the views of Confederation of Swedish Enterprise, or Sveriges Riksbank or its Executive Board. Author e-mails: hanna.armelius@ekonomifakta.se; carlandreas.claussen@riksbank.se; andre.reslow@riksbank.se.

Figure 1. Percentage Change in Currency in Circulation between 2001 and 2018



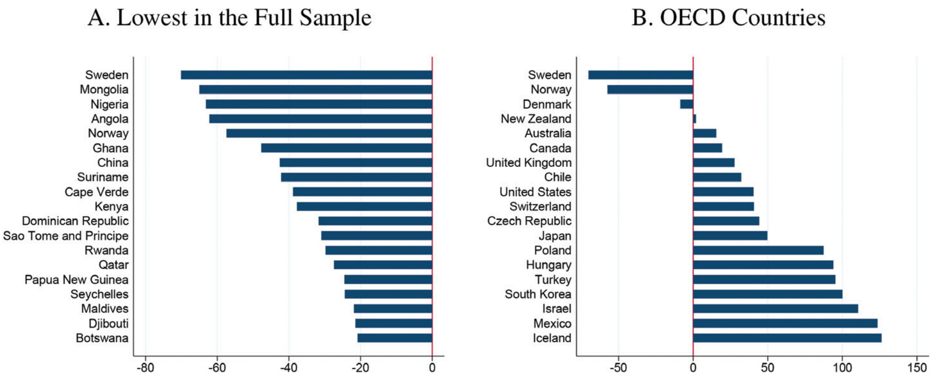
Note: In graph A we show the development for the 19 countries with the lowest increase in our sample, while graph B shows the development for all the OECD countries in our sample. *For illustrative purposes, the graph shows an increase of 800 percent for Turkey, while the actual increase was 2,864 percent.

and other mobile devices are advancing fast. This makes digital payments cheaper and more convenient, and also allows for non-cash payments in situations where cash used to be the only option.

However, the amount of cash in circulation keeps increasing (Figure 1). In many countries—for instance, the United States—the increase has been more than 170 percent since 2001. The growth in cash has even surpassed the growth of the economy in most countries (Figure 2). Sweden stands out as a notable exception since cash in circulation has fallen by almost 50 percent there. Cash as a share of GDP has fallen by even more and now stands at less than 1.5 percent. Neighboring Norway has experienced a similar, but less pronounced, development.

How can we explain the differences in the growth of cash in circulation in different countries? Are Sweden and Norway just ahead of the others, or is there something special about them? Can other countries expect a similar development in the near future? What drives the demand for cash more generally? In this paper, we address these questions. They have become even more relevant during the COVID-19 pandemic when the use of cash for transaction purposes

Figure 2. Percentage Change in the Cash-to-GDP Ratio between 2001 and 2018



Note: In graph A we show the development for the 19 countries with the lowest increase in our sample, while graph B shows the development for all the OECD countries in our sample.

has fallen in many countries, while the use of cash for store of value has increased.[1]

Understanding what drives the development of cash is important for several reasons. One is that cash payments can be more costly for society than digital payments (see, e.g., Schmiedel, Kostova, and Ruttenberg 2012).[2] Another is that cash might facilitate criminal activity. Some countries may therefore want to understand how they can reduce the use of cash. Conversely, cash may be fundamental to our monetary systems, as it enforces the uniformity of money and makes commercial bank money appear less risky (see, e.g., Armelius, Claussen, and Hendry 2020). Furthermore, cash facilitates anonymous payments and competition in the payment market; cash enhances economic resilience and seigniorage revenues; and cash makes it possible for everyone to make their daily payments. Some

[1]See, for example, Ashworth and Goodhart (2020a) and Heinonen (2020) for a global survey. However, Sweden continues to be an outlier with non-increasing cash in circulation (Sveriges Riksbank 2020).

[2]For an alternative view, see Carbo-Valverde and Rodriguez-Fernandez (2019).

also argue that cash protects against "digital dollarization."[3] Countries may therefore want to stop a potential marginalization of cash. Whatever the reasons may be, if we want to influence the development of the amount of cash in circulation, we need to understand what drives it.

We use an extensive data set consisting of 129 countries and covering the years 2001 to 2018 to explore if econometric models can explain the development of cash in general, and the Swedish divergence in particular. Panel regressions using the full sample suggest that economic development is a key explanatory variable. In general, richer countries have less cash in circulation relative to GDP. In line with the previous literature, we also find that increases in the opportunity cost of cash (the interest rate) reduce cash demand, while a higher average age of the population increases cash demand. When we limit the sample to OECD countries, higher corruption is associated with higher demand for cash.

Our main specification performs well in explaining the development in most OECD countries. However, it cannot explain the development in Sweden, where the model fit is more than twice as bad as for any other country. We therefore discuss potential explanations as to why Sweden is "unexplained" by the model. More specifically, we discuss Swedish policy measures to reduce tax evasion; an aggressive banknote and coin changeover; the introduction of a new mobile payment application; as well as a few other recent events in Sweden. These policy measures and developments appear to have affected access to, and demand for, cash. Thus, while our estimations do not indicate that all countries will soon see a reduction in cash in circulation, the Swedish experience suggests that countries that simultaneously implement a combination of reforms that make cash less attractive and electronic payments more convenient may see a significant reduction in the use of cash.

This paper contributes to the literature in the following ways. First, our study covers a large number of countries, providing results for both developed and developing countries. Second, we consider variables that are often excluded in cash demand studies, such as

[3] "Digital dollarization" is a situation in which the national currency is supplanted by a digital platform's currency rather than another developed country's currency (Brunnermeier, James, and Landau 2019).

corruption, trust, and technology adaptation. Third, we provide a thorough discussion of events and institutional settings that can help us understand the divergent development in Sweden relative to other countries. The latter is highly policy relevant, since the development in Sweden is often in the spotlight in international policy discussions.

The paper proceeds as follows. The next section provides an overview of the relevant literature. Section 3 describes the data, while Section 4 explains the empirical strategy. Section 5 presents the main estimation results, and Section 6 discusses the predictions of the model in comparison to actual developments. Section 7 discusses potential reasons why the model cannot explain the development in Sweden, and Section 8 concludes.

2. Related Literature

Theories of cash demand often start from the Baumol (1952)–Tobin (1956) inventory model and predict that cash demand will be increasing in income or spending, decreasing in the opportunity cost of holding cash, and increasing in the cost of acquiring cash. Keynes's (1937) three motives for holding cash give similar predictions and also suggest that people will hold higher cash balances when there is increased uncertainty.

The empirical literature on money demand, taking theory as a starting point, is vast. Most relevant for us are the more recent papers where researchers estimate cash demand relations.[4] A robust finding in these papers is that cash in circulation increases with GDP and falls with the interest rate, in line with what theory predicts. Evidence is mixed for the cost of acquiring cash; some find negative effects of the number of ATMs and bank branches and some find positive effects. There is scarce empirical evidence to support that increased uncertainty would increase cash balances. Furthermore, there is evidence that increased penetration of electronic payment

[4]See, for example, Amromin and Chakravorti (2009); Arango-Arango and Suárez-Ariza (2019); Ashworth and Goodhart (2020b); Assenmacher, Seitz, and Tenhofen (2019); Bech et al. (2018); Cusbert and Rohling (2013); Huynh, Schmidt-Dengler, and Stix (2014); Jobst and Stix (2017); Seitz, Fischer, and Köhler (2004); Shirai and Sugandi (2019).

alternatives reduces the demand for cash. Papers that include proxies for the informal sector tend to find positive effects, albeit not always significant. Finally, papers that include some measure of the average age of the population usually find that it has a positive effect on cash demand. We summarize all these potential explanatory factors often used in the empirical literature, and variables used to capture these factors, in Table 1.[5]

3. Data

Our variable of interest is currency in circulation (CiC), specifically the ratio between CiC and GDP.[6] This ratio is convenient since it allows us to compare countries without worrying about exchange rates, and it has a simple theoretical interpretation as the inverse of money velocity. Our data span the period 2001–18, and consist of 129 countries, out of which 19 are OECD members. We exclude countries for which we could not find key data and countries in the European Monetary Union. The sample period is mainly defined by the CiC data availability in the International Monetary Fund (IMF) database. All the countries in our final sample are listed in Table A.1 in the appendix.[7]

In addition to CiC and GDP, we collect a large number of potential explanatory variables, both standard variables from the existing literature and some new ones. Among the variables in Table 1, we have collected data on GDP per capita, the interest rate, the share of self-employed, uncertainty, and the old-age dependency ratio.[8] We

[5]Other related studies, but somewhat less relevant for our study, include empirical papers using microdata and theoretical papers that study consumer behavior and cash usage (see, e.g., Alvarez and Lippi 2009; Attanasio, Guiso, and Jappelli 2002; Bagnall et al. 2016; Wakamori and Welte 2017; Wright et al. 2017). See also Bartzsch, Rösl, and Seitz (2013) for the role of foreign demand.

[6]Currency (or cash) in circulation refers to the outstanding amount of money in the form of notes and coins issued by the central bank and/or government.

[7]We focus on those countries where we observe CiC throughout 2001–18. For Djibouti, we extrapolate using a spline function to obtain a missing value in 2001.

[8]We use the short-term interest rate from the OECD database. When the OECD interest rate data are unavailable, as they are for most countries in our full sample, we create a measure that is the mean of four different short-term interest rates (the deposit rate, the money market rate, rates on government T-bills, and the central bank policy rate) from the IMF's International Financial Statistics (IFS) database. For many countries, only a subset of the four rates is

Table 1. Explanatory Factors in the Literature

Explanatory Factor	Variables	Estimated Coefficient
Scaling Factor	GDP, GDP per capita	+
Alternative Cost	Interest Rates	−
Cost of Withdrawing Cash	Number of ATMs, Number of Bank Branches	+/−
Uncertainty	“Uncertainty Index,” Crisis Dummy	+/No Effect
Ease of Electronic Payments	Number of EFTPOS Terminals	−/No Effect
Informal Sector	Share Shadow Economy	+/No Effect
Small Business	Ratio of Self-Employed	+
Age Structure	Life Expectancy, Old-Age Dependency Ratio	+

Note: The signs refer to the factor and not the variable. As the elasticities in different studies are not directly comparable, we only refer to the signs. The listed variables represent a selected sample of commonly used variables.

have also considered variables like the number of ATMs, commercial bank branches, and debit/credit card ownership, but decided to leave them out of our final data set, for two main reasons. First, these variables are likely to be determined in tandem with cash demand and will therefore lead to simultaneity bias in the estimations. Second, when included in the estimations, we find no clear relationships, and our main coefficients are robust to the inclusion of these variables.

Some of the new variables that we consider are motivated by the fact that cash provides anonymity and leaves no electronic traces—features that can be desirable for illegal activities. We therefore include measures of corruption and organized crime. We may also notice that higher crime rates may, on the one hand, raise the cost of distributing cash, and thereby the cost of getting hold of cash, and thus increase cash holdings. On the other hand, it might induce people to hold less cash for security concerns. The anonymity provided by cash might also be desirable in oppressive regimes. Hence, we include a variable measuring human rights and a variable measuring trust in politicians. In addition, trust in politicians (and crime rates) matters for the development of cash in circulation more broadly since it influences the investment climate in general and, therefore, investments in ATMs and infrastructure for electronic payments.

People who do not trust banks to protect their integrity might prefer cash to commercial bank deposits. People might also prefer cash because they do not trust retail banks to be sufficiently safe. This hypothesis is supported by monetary theory, which suggests that people will prefer cash or other forms of central bank money over private money if institutions that facilitate trust in commercial bank money are weak (see, e.g., Armelius, Claussen, and Hendry 2020). Therefore, we include a variable measuring trust in the financial sector and a variable measuring the regulatory quality in each country.

As a country-specific measure of uncertainty, we use the World Uncertainty Index by Ahir, Bloom, and Furceri (2019). Measuring digitization and technology adaptation is not straightforward, and the data that exist are often not observable for many countries or

available, and for many countries, the rates are not observed in all the years. We therefore use the mean of the available rates. In our OECD sample, the correlation between the OECD interest rate and our mean of the IMF rates is 0.97.

an extended period of time. We collect data on Internet usage as a proxy for general attitudes towards technology adaptation. It will also capture technological possibilities and ease of making electronic payments.

All collected variables and their descriptive statistics are presented in Table 2. Given the large heterogeneity among the countries in our sample, we also consider a subsample, limited to the OECD countries in our data set. Table 2 shows that the average CiC/GDP is 7 percent in the full sample, while it is 5.57 in the OECD subsample. The standard deviation presented in the table is the overall variation. It is worth noting that for some of the variables (e.g., Self-Employed, Human Rights, Regulatory Quality, Control of Corruption) most of the variation comes from between countries such that they display less variation within countries. With 129 countries and 18 years, we have a potential maximum of 2,322 observations for each variable. However, for most variables, we do not have observations for all countries and all years, resulting in a number of missing observations.[9]

4. Empirical Strategy

We estimate the following fixed-effects reduced-form cash-demand model,

$$C_{i,t} = \alpha_i + \delta_t + \boldsymbol{\beta}\mathbf{X}_{i,t} + \varepsilon_{i,t}, \tag{1}$$

where i is a country indicator, t a year indicator, $\mathbf{X}_{i,t}$ a set of explanatory variables, and $\varepsilon_{i,t}$ is a random error with mean zero. We use the natural logarithm of the cash-to-GDP ratio (*log* CiC/GDP) as the dependent variable $C_{i,t}$. As mentioned before, this ratio is convenient since it allows us to compare countries without worrying about exchange rates. Although our main specification will be a fixed-effects model, we will also estimate the model replacing the country fixed effects, α_i, with a common constant.

[9]We treat all missing observations as "missing at random."

Table 2. Descriptive Statistics

	Full Sample			OECD		
	Obs.	**Mean**	**Std. Dev.**	**Obs.**	**Mean**	**Std. Dev.**
Panel A						
CiC/GDP	2,322	7.00	4.16	342	5.57	3.91
Interest Rate	2,206	6.18	5.49	342	4.20	5.79
log GDP per capita	2,322	1.32	1.47	342	3.32	0.76
Self-Employed	2,232	45.15	27.15	342	17.71	9.04
Age Dependency Ratio	2,286	10.93	7.03	342	20.95	6.50
Individuals Using Internet	2,227	29.86	28.30	342	68.29	23.16
Human Rights	2,286	0.40	1.46	342	1.61	1.65
Regulatory Quality	2,320	48.75	25.50	342	86.45	11.32
Control of Corruption	2,320	47.25	27.94	342	83.51	17.45
Panel B						
World Uncertainty Index	1,782	0.18	0.16	324	0.21	0.13
Confidence in Financial Sector	1,072	58.03	16.40	222	53.29	15.75
Public Trust in Politicians	1,043	3.05	1.19	209	3.82	1.33
Organized Crime, 1–7 (Best)	1,043	4.87	1.02	209	5.55	0.88

Note: CiC is collected from the IMF (IFS), Bank of England, the People's Bank of China, the Reserve Bank of India, the National Bank of Switzerland, and Singapore Department of Statistics. Interest rates are collected from the OECD and the IMF (IFS). From the World Bank we collect GDP (local currency), GDP per capita (1,000 USD), Self-Employed (% of total employment), Individuals Using Internet (% of the population; we interpolate and extrapolate to obtain nine missing observations in the OECD sample), and the Age Dependency Ratio (old as % of the working-age population). Control of Corruption (percentile rank; we extrapolate to obtain values for 2001) and Regulatory Quality (percentile rank) are collected from the Worldwide Governance Indicators (Kaufmann, Kraay, and Mastruzzi 2011). Organized Crime (1–7, with 7 being the best) and Public Trust in Politicians (1–7, with 7 being the best) are collected from the Global Competitiveness Index (World Economic Forum). Human Rights score is collected from Fariss (2019); we extrapolate to obtain values for 2018. Confidence in Financial Sector (% responding yes) is collected from the Gallup World Poll, and the Uncertainty Index is from Ahir, Bloom, and Furceri (2019).

In the absence of sharp identification, the panel data structure is essential since it allows us to utilize two sources of variation: variation across countries within each year and variation within countries across years. The year fixed effects capture any common time trend and are important since they absorb global trends and global shocks—such as the global financial crisis (2007–08). Some of our variables display less variation within countries than between countries. Hence, estimations excluding country fixed effects should be interpreted as cross-country estimates that compare cash demand factors between countries. In contrast, specifications that include country fixed effects allow for within-country interpretations since country fixed effects control for different levels and for omitted time-invariant elements (e.g., culture and religion).[10]

One concern regarding the estimation of Equation (1) is stationarity. Testing for stationarity in a panel like ours can be a bit problematic, and therefore we use a number of different tests. Using a Harris and Tzavalis (1999) unit-root test, we cannot reject the null hypothesis that the panels (countries) contain unit roots. However, we can reject (at the 0.1 level) that the panels contain unit roots when we include a time trend. Hence, the panels appear to be trend stationary. One drawback with the Harris and Tzavalis (1999) test is that it is based on the assumption that all panels have the same autoregressive parameter. To combat this limitation, we turn to alternative tests. Using the Im, Pesaran, and Shin (2003) unit-root test, we reject the null that all panels contain unit roots. However, using a Hadri (2000) Lagrange multiplier (LM) test, we also reject the null that all panels are stationary. Hence, some countries seem to be stationary, while others are not. One caveat with the country-specific unit-root tests is that they assume that both N (the number of countries) and T (the number of years) tend to infinity. In our data, $N = 129$ and $T = 18$. Therefore, while the assumption might be fine for N, it is likely less so for T. Hence, we should interpret the tests with some caution.[11]

[10] In all estimations we consider standard errors clustered at the country level to account for likely error correlation within each country (see, e.g., Abadie et al. 2017; Angrist and Pischke 2008; Cameron and Miller 2015).

[11] On the other hand, the Harris and Tzavalis (1999) test assumes that N approaches infinity while T is fixed, but has the drawback of assuming a common unit root.

Nevertheless, in order to ensure that our results are robust and not contested due to non-stationarity, we also estimate the following model,

$$\Delta C_{i,t} = \alpha \Delta C_{i,t-1} + \delta_t + \boldsymbol{\beta} \Delta \mathbf{X}_{i,t} + \varepsilon_{i,t}, \tag{2}$$

using generalized method of moments (GMM), where Δ denotes first differences.

When deciding on the final set of variables to include in $\mathbf{X}_{i,t}$, we face several trade-offs. If we were to include all of our collected variables, we would reduce the risk of omitted-variable bias. Still, at the same time, we would drastically reduce the number of observations, since many variables are observed only for some scattered years. A second concern is multicollinearity. Many of our variables are correlated, although we do not have any variables that have a very high correlation (above 0.9).[12] We have chosen to focus on variables where we have a large amount of data. In Table 2, the variables in panel A are the ones included in $\mathbf{X}_{i,t}$ in our main specification (net of CiC/GDP that serves as our dependent variable). By excluding the variables in panel B, we obtain a set of variables that will constitute a fully balanced panel for the OECD sample, and we limit the multicollinearity concerns. However, we are still interested in assessing the relationship and importance of the variables in panel B. Hence, we also estimate

$$\Delta C_{i,t} = \alpha \Delta C_{i,t-1} + \delta_t + \boldsymbol{\beta} \Delta \mathbf{X}_{i,t} + \gamma \Delta z_{i,t} + \varepsilon_{i,t}, \tag{3}$$

where $z_{i,t}$ is each of our additional explanatory variables (those not included in $\mathbf{X}_{i,t}$) added one at a time.

5. Empirical Results

We estimate Equations (1) and (2) using both the full sample of all countries and the subsample of OECD countries. The estimated coefficients are presented in Table 3. Columns 1 and 4 suppress the country fixed effects, while the full specification of Equation (1) is

[12] We have performed a variance inflation factor (VIF) test to assess the multicollinearity problem. All VIF values are below the rule-of-thumb threshold of 10, indicating that we do not have any severe multicollinearity problems.

Table 3. Cash Demand Estimation Results

	Full Sample			OECD		
	OLS (1)	**FE (2)**	**GMM (3)**	**OLS (4)**	**FE (5)**	**GMM (6)**
Interest Rate	–0.035***	–0.007**	–0.004***	–0.043***	–0.013***	–0.008***
	(0.006)	(0.003)	(0.002)	(0.009)	(0.004)	(0.002)
log GDP per capita	–0.271***	–0.155**	–0.174***	–0.115	–0.017	0.034
	(0.074)	(0.078)	(0.055)	(0.276)	(0.136)	(0.039)
Age Dependency Ratio	0.034***	0.003	0.007	0.075***	0.011	0.008
	(0.008)	(0.010)	(0.006)	(0.018)	(0.019)	(0.005)
Self-Employed	–0.004	–0.005	–0.003	0.005	–0.041**	–0.000
	(0.003)	(0.006)	(0.003)	(0.016)	(0.015)	(0.008)
Individuals Using Internet	0.003	0.000	–0.000	–0.005	0.008	–0.000
	(0.003)	(0.001)	(0.001)	(0.006)	(0.005)	(0.001)
Human Rights	–0.055*	0.033	0.008	–0.107*	0.078	0.024
	(0.032)	(0.028)	(0.018)	(0.056)	(0.052)	(0.021)
Regulatory Quality	–0.002	–0.003	–0.001	0.000	–0.007	–0.002
	(0.003)	(0.002)	(0.001)	(0.011)	(0.007)	(0.002)
Control of Corruption	–0.003	0.000	–0.001	–0.016*	–0.008*	–0.004**
	(0.003)	(0.002)	(0.001)	(0.008)	(0.004)	(0.002)
Observations	2,010	2,010	1,780	342	342	304
R^2	0.335	0.910		0.636	0.940	
R^2 Adjusted	0.326	0.903		0.607	0.931	
Year FE	√	√	√	√	√	√
Country FE		√			√	

Note: The dependent variable is the natural logarithm of the cash-to-GDP ratio. In columns 1 and 4 the country fixed effects have been suppressed and replaced by a common constant. Standard errors robust to clustering at country level are in parentheses. *, **, and *** represent the 10 percent, 5 percent, and 1 percent significance level, respectively.

presented in columns 2 and 5. Columns 3 and 6 present the GMM estimations of Equation (2).

In line with earlier studies, we find a negative and statistically significant effect of the interest rate on cash demand. Between countries, a 1 percentage point higher interest rate is associated with a 3.5 to 4.3 percent lower cash-to-GDP ratio. When adding country fixed effects, the coefficients are attenuated to around –0.01 but are still significant, such that a 1 percentage point higher interest rate is associated with around 1 percent lower cash-to-GDP ratio. In line with, for example, Bech et al. (2018), we find that richer countries have a lower cash-to-GDP ratio. The coefficient on *log* GDP per capita is negative and significant for the whole sample, but insignificant (and attenuated) for the OECD subsample. In the full sample, a 1 percent increase in GDP per capita is associated with a 0.3 percent lower cash-to-GDP ratio between countries and 0.16 percent lower in the within-country estimates.

As expected, and in line with earlier findings, we find that age matters. The coefficient is positive in all specifications and highly significant in models without country fixed effects. Countries with a 1 percentage point higher age dependency ratio will have a 3.4 percent higher cash-to-GDP ratio in the full sample and 7.5 percent higher in the OECD sample. When adding the country fixed effects, the age variable becomes smaller and insignificant. This could be because much of our variation is between countries, while the variation within countries over time is limited. This is not very surprising since demography does not change that much over time. The same pattern holds for the human rights variable. We observe a significant (negative) relationship when we exclude the country fixed effects. When adding the country fixed effects or estimating in first differences, the coefficient becomes insignificant. The fact that coefficients and significance change when adding fixed effects is not surprising. Many variables have different amounts of variation between and within countries. It is important to note that by including country fixed effects we control for and absorb time-invariant differences between the countries.

In all estimations, Internet usage and regulatory quality turn out to be insignificant and close to zero. The number of self-employed is generally insignificant and not consistently estimated. Our estimates suggest that better control of corruption reduces the amount of cash.

A one-unit increase in the control of corruption is associated with a 0.8 to 1.6 percent lower cash-to-GDP ratio in the OECD estimation. Hence, a one-standard-deviation increase of 17.45 in control of corruption would imply a decrease in the cash-to-GDP ratio of 16 to 30 percent.

The results so far have omitted the variables listed in panel B of Table 2. We are still interested in assessing their relationship and importance for cash demand. As described in Section 4, we therefore also estimate Equation (3). The results are presented in Table A.2 in the appendix. All variables in panel B turn out to be insignificant and close to zero, and adding these variables does not alter the main takeaways from Table 3. We may note that the old-age dependency ratio turns significant in the OECD sample due to increased precision.

As noted earlier, there is some scarce evidence in the previous literature that uncertainty positively affects the amount of currency in circulation. This seems to be visually supported for some countries in our data. There appears to be, for some countries, a more pronounced increase during 2007–18, following the global financial crisis, compared with 2001–07. However, looking at the estimated year effects, we do not find any evidence that the years associated with the financial crisis would significantly differ from the other years in our sample. One caveat with this approach is that the year effects assume that all countries had a homogeneous exposure to the crisis. Therefore, when assessing uncertainty, it is preferable to include variables that capture each country's heterogeneous exposure. However, as shown in Table A.2, we do not find any significant relationship between the World Uncertainty Index by Ahir, Bloom, and Furceri (2019) and the cash-to-GDP ratio.

5.1 Robustness

As a test for the robustness of our model selection, we perform an exercise using a *lasso*-model selection (Hastie, Tibshirani, and Wainwright 2015; Tibshirani 1996). We allow the lasso selection to choose from our main set of variables in $\mathbf{X}_{i,t}$ (i.e., panel A of Table 2), but force the selection of year and country fixed effects. The results from this exercise reveal that the final model selection differs from our main specification for both the full and the OECD samples.

In the full sample, the variables capturing self-employment, Internet usage, and corruption are excluded in the lasso selection. In the OECD sample, the *log* GDP per capita and the human rights variable are excluded. However, the estimation of the remaining coefficients aligns very well with the results from our main model presented in Table 3.

We further assess the robustness of our results by removing countries where foreign demand for the (physical) currency is large. We first remove the United States since the U.S. dollar is widely used outside of the United States for daily transactions and store of value. We then perform a second estimation where we also remove Switzerland, Japan, and the United Kingdom—countries whose currencies also are used abroad for daily transactions and store of value. From these tests, we conclude that our results are robust to these exclusions.

As a final robustness check we also perform estimations where we replace Self-Employed, Human Rights, Regulatory Quality, and Control of Corruption with their averages. This exercise provides some additional observations, as we average out some missing values. Note that these variables are then omitted in the fixed-effects and GMM estimations. Again, the results are robust to these alternative specifications.

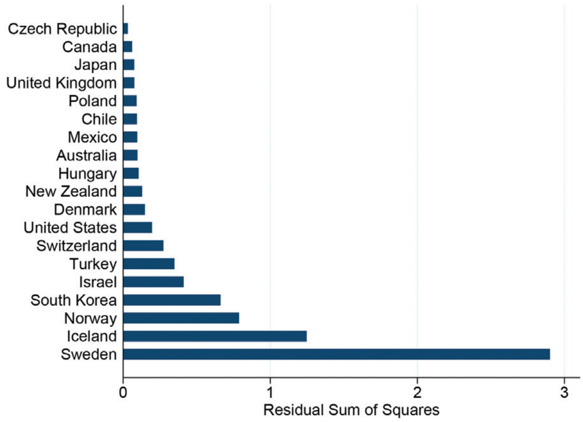
6. Can the Model Explain the Divergent Development?

In this section, we analyze if our empirical model can predict ("explain") actual outcomes. We limit the analysis to the OECD sample, and we use the estimation presented in column 5 of Table 3.[13] Based on this estimation, we calculate the residual sum of squares (RSS) for each country. We report these RSS values for each country in Figure 3, where we have ordered the countries from best to worst model fit.

The figure shows that the model has a very good fit for countries like the Czech Republic, Canada, and Japan. The countries

[13] We limit this analysis to the OECD countries since, in that sample, we have a fully balanced panel using our main specification and we believe that the OECD sample is more homogeneous.

Figure 3. Residual Sum of Squares



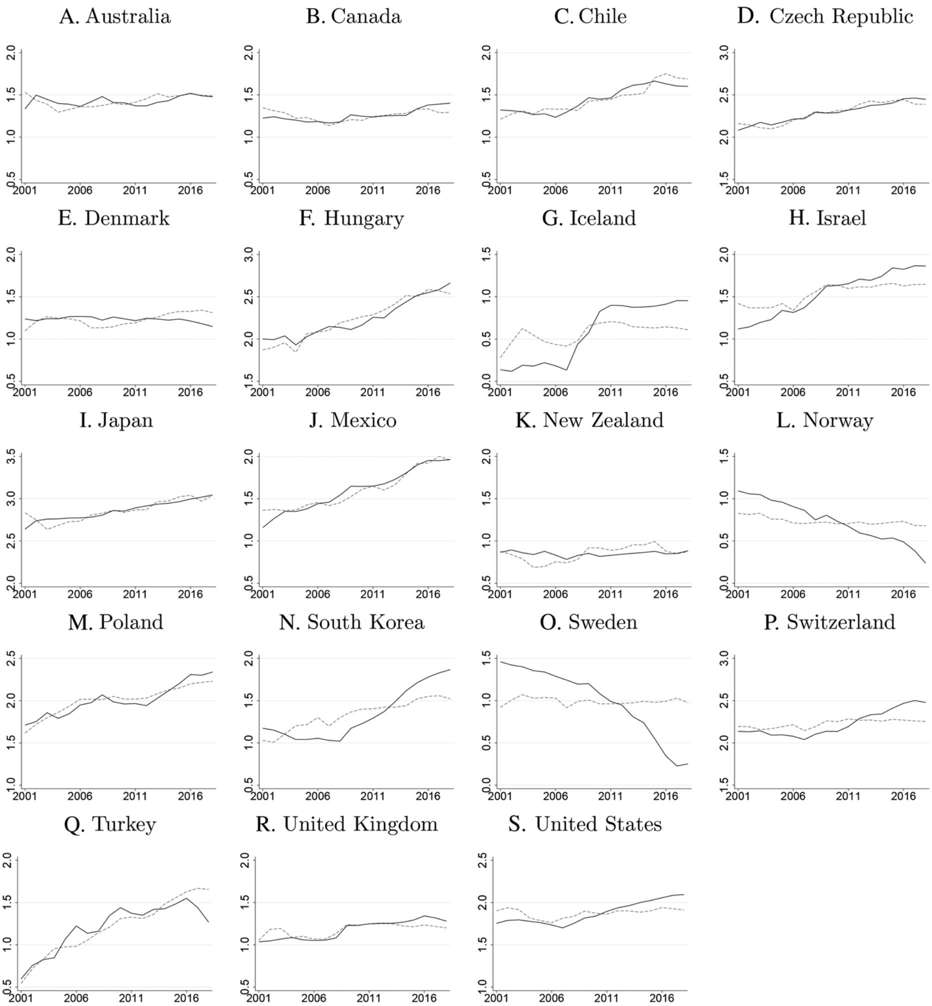
Note: The graph shows the residual sum of squares (RSS) based on the estimation in column 5 of Table 3.

that stand out as being poorly predictable are Iceland and Sweden. The low predictability for Iceland is likely due to the global financial crisis, which hit Iceland particularly hard and led to a substantial increase in the cash-to-GDP ratio. We notice that the development in Sweden has been exceptionally hard to predict; the RSS value is more than twice as large as for any other country, indicating that Sweden is, indeed, special.[14] An interesting observation is that Norway, which also stands out with a fall in CiC (Figure 1), is better explained by the model than Sweden—although relatively poorly explained compared with the rest of the sample.

In order to visualize the model's fit and explanatory power over time, we plot the fitted values (as dashed lines) and the actual values (as solid lines) for each country in Figure 4. The figure shows that the model has a good fit for most countries. It predicts an increase in cash in circulation in several countries. The increase in actual *log* CiC/GDP in Iceland after the financial crisis, which gives the high RSS value, is evident from the figure. We can also notice that South Korea, Switzerland, and the United States are countries where the

[14]It is worth noting that the RSS value for Sweden is the largest in the full sample as well.

Figure 4. *log* CiC/GDP, Actual Value and Model Prediction



Note: The figure shows the model predictions (fitted values) based on the estimation in column 5 in Table 3 as dashed gray lines, and the actual outcomes as solid black lines. The graphs show the *log* CiC-to-GDP ratio.

financial crisis might have had a substantial impact on the trend in actual *log* CiC/GDP. Looking at Sweden, we see that the model fails to capture the sharp decline in cash in circulation, since the model predicts an unchanged level. The model also fails to fully predict the decrease in Norway, although we notice that Norway is the only country where the model predicts a decline.

7. Discussion: What Is Special about Sweden?

Having explored what we can learn from cross-country data, we now discuss some Swedish policy measures and developments that may help explain why the model cannot explain the divergent development in Sweden. More specifically, we suggest that the combination of Swedish policy measures to reduce tax evasion, an aggressive banknote and coin changeover, and the introduction of a new mobile payment application could be important for the development of CiC in Sweden. While these types of events and changes are not unique to Sweden, the fact that they were all implemented within a short period could have reinforced their effects. The timing of these events is illustrated in Figure 5. We also discuss, in Section 7.4, a few other aspects that could help explain why Sweden is special.

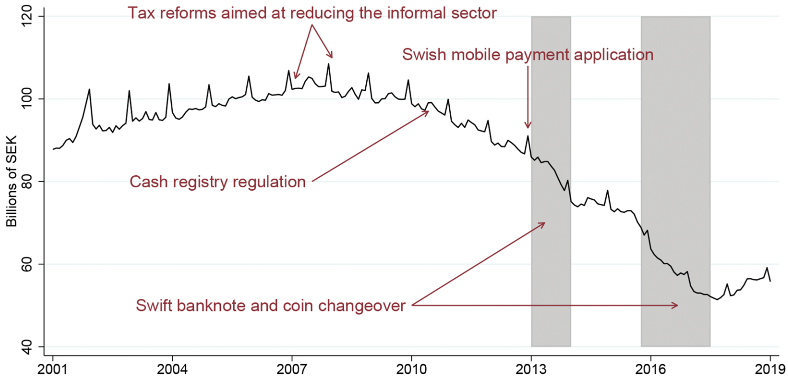
7.1 Reduced Tax Evasion and a Smaller Informal Sector

Starting in 2007, the Swedish authorities introduced measures aimed at transferring jobs from the informal to the formal sector and reducing tax evasion.

In 2007 a substantial tax deduction for the purchase of household services, such as cleaning, was introduced. In 2008 a similar tax deduction for services related to house repairs and maintenance was introduced. The objective of the schemes was to reduce undeclared work. The measures appear to have had an effect. Tillväxtanalys (2019) used microdata on Swedish firms to create a control group consisting of firms that had similar characteristics to the firms that were eligible for the deductions prior to the reform. They then ran fixed-effects regressions and found that the tax reform for household services had increased the amount of formal work in the household services sector by around 10 percent.

In 2010 it became mandatory for firms selling goods or services in return for cash to have a certified cash register and report the cash register to the Swedish Tax Agency. The provisions also involved an

Figure 5. Events that Help Explain the Decline in CiC in Sweden



Note: The graph shows monthly currency in circulation (in billions of SEK) in Sweden.

obligation to produce and offer the customer a receipt. In addition, the Tax Agency was allowed to conduct more supervision and inspection visits. The combination of the new law and the increased number of inspections made it more difficult for businesses to withhold income by receiving payments in cash. The number of fines levied by the Tax Agency when irregularities were discovered increased from 500 in 2010 to 2,900 in 2012. Swedish Tax Agency (2012) conducted a study that exploited differences in timing in the submission of the first report by different companies to the Tax Agency. The study found that reported turnover was 5 percent higher in the months following the notification of a tax register as compared with the turnover by similar companies that had not submitted a report.

These measures are not directly captured by the explanatory variables of our model. Although variables such as regulatory quality and corruption might capture some of the effects, the reforms are likely to be too narrow to be proxied by the broader measures that we observe on a country level. Here we would also like to note that although Swedish Tax Agency (2012) and Tillväxtanalys (2019) report that the measures have reduced the informal sector and tax evasion, it is hard to disentangle the measures' effect on cash demand empirically. A key reason is that we do not have sector-specific cash demand data. Moreover, as noted by Engert, Fung, and Segendorf (2019), numerous countries in the last 10 to 20 years have experienced a general trend of declining underground economies.

7.2 An Aggressive Banknote and Coin Changeover

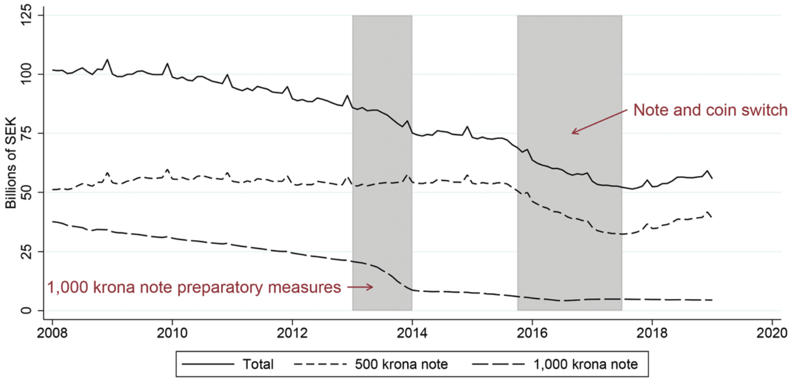
During two intervals between 2012 and 2017, the Riksbank conducted a changeover of banknotes and coins. A particular feature of this changeover was that the window for exchanging old notes for new ones was short, only nine months. Furthermore, the Riksbank applies relatively strict redemption rules. Invalid notes can only be redeemed at the Riksbank's main office in Stockholm for a fee, and only if proper documentation of their origin is presented (in order to avoid money laundering).

The changeover started in November 2012 when, as a preparatory measure before new notes would be introduced, older 50- and 1,000-krona notes without a foil strip still in circulation were declared to be invalid from year-end 2013. At the same time, Sveriges Riksbank (2012) announced that the versions of the 50- and 1,000-krona notes with foil strips, which had been introduced in 2006, would be valid only until June 2017. This meant that anyone holding the oldest version of, for example, the 1,000-krona banknote knew that they would have to do at least one more switch in the near future.

After these preparatory measures, the changeover was conducted in two steps. The first began in October 2015, when the Riksbank issued new 20-, 50-, 200-, and 1,000-krona banknotes. In September and October 2015, the Riksbank sent out information brochures to the general public and also informed the public about the banknote and coin changeover through other channels. This information included the announcement that the old versions of the respective notes would become invalid after nine months.[15] The second step was initiated in October 2016, when the Riksbank issued new 100- and 500-krona banknotes and new 1- , 2- , and 5-krona coins. The procedure was once again that the old banknotes and coins were valid for only nine months after the new ones had started to be

[15] The Riksbank's communication measures regarding the banknote and coin changeover are documented in Sveriges Riksbank (2018b).

Figure 6. The Swedish Banknote and Coin Changeover



Note: The graph shows monthly currency in circulation (in billions of SEK) for different denominations in Sweden. The "Total" series corresponds to all coins and banknote denominations available.

issued. However, this validity limit had already been announced in September 2015.[16]

Having to exchange notes is inconvenient for cash holders. In addition, large-denomination notes were already cumbersome to use since many shops did not accept them. Furthermore, by 2013 it had become harder to exchange notes at bank offices. The number of bank offices had declined, and many of the remaining ones had become cashless.

Looking at the timing of the changeover and the time series, we can see clear drops in the largest denominations during the two changeover periods. In Figure 6, we see that the preparatory period of 2013 coincides with a significant drop in circulation of the 1,000-krona banknote, while the main changeover period coincides with a large decline in circulation of the 500-krona banknote.[17]

[16] There is still SEK 5.4 billion worth of banknotes outstanding that have not been redeemed. See also Sveriges Riksbank (2018a) for a summary and evaluation of the banknote and coin changeover.

[17] During the changeover period, the 500-krona banknote made up around 70 percent of the total amount of currency in circulation.

Banknote and coin changeovers are not uncommon elsewhere, but the recent Swedish ones were aggressive in an international comparison. Internationally, old notes are often legal tender for a very long time after the introduction of new notes and sometimes even indefinitely. In the United States, for instance, all notes issued since 1861 are legal tender. In Denmark, all banknotes issued after 1945 are legal tender. The bank of Canada did not get the power to remove legal tender status of banknotes before 2018. After that, legal tender status has only been removed for banknotes that had not been produced for at least two decades. Compared with these economies, the window given in Sweden was short. We may add to this that, historically, Swedish changeover periods have been much longer than this one. Engert, Fung, and Segendorf (2019), who compare the development of cash in Sweden and Canada, assess that the relatively aggressive banknote and coin changeover is likely to have reduced the demand for larger notes in Sweden relative to Canada.

A very aggressive banknote changeover will probably not on its own reduce the demand for cash permanently, as is evidenced by, for instance, the measures taken in India in 2016. There, currency in circulation showed a large drop immediately following the announcement that some large-denomination notes would become invalid and exchangeable for new notes for only 50 days. However, a couple of years later, currency in circulation was back at the old level, and it has continued to grow with the old trend since. In the Swedish case, it is possible that the changeover had a larger effect since there were attractive digital substitutes for cash available when the changeover took place.

7.3 An Attractive Mobile Payments Application

In December 2012 a new payment application for smartphones called *Swish* was introduced in Sweden. The application offers digital real-time payments (person-to-person and person-to-business) between commercial bank accounts in different banks.[18] Its user-friendliness,

[18]The service is only available for SEK accounts in banks operating in Sweden and is therefore not available to foreign tourists and others not holding a Swedish bank account.

real-time properties, and broad reach made digital payments possible in essentially all areas where cash payments have previously been the only option. More than 80 percent of the adult Swedish population now has the app installed. Since this corresponds to the latest available estimate of the share of smartphone ownership, Swish has essentially reached full market penetration in the adult population.

The introduction and rise of Swish as an alternative to cash are not captured directly by any of the explanatory variables in our model. However, including variables like the number of Swish users as an explanatory variable would lead to spurious estimation results. Swish and CiC are likely to be just mirrors of each other since both are determined by the same exogenous variables—for instance, age of the population, regulatory quality, and technology adaptation.

Other countries have implemented similar services, but Swish differs from most of these in that it essentially covers the whole banking sector and has, in principle, universal reach. In most countries, the services appear to be more piecemeal. The fact that Swedish banks were able to develop a common solution is in line with a long tradition in Sweden. Swedish banks are used to setting up jointly owned, infrastructure-related companies that provide services for all banks while still promoting competition among them. One example is *Bankomat AB*, which operates the vast majority of ATMs in Sweden, and is jointly owned by the major banks. Another example is a common digital identification system supplied by the banks (called *BankID*) and used by all banks for online banking services, by Swish, and by public authorities. This is different from the workings of banking sectors in many other countries and is hard to measure and include in the empirical estimation.

7.4 Other Aspects that Could Explain the Fall in CiC

As noted above, Swedish banks have reduced their cash services over the studied period. Between 2011 and 2016, the number of bank offices offering cash services more than halved.[19] The number

[19] Sometimes there can be an ATM in (or close to) a cashless bank office. They do not, however, offer cash services over the counter, and in particular they offer no means of depositing cash.

of ATMs fell by 14 percent, and the number of cash service boxes (that smaller businesses use for handling their daily takings) fell by 15 percent from 2011 to 2017 (The Riksbank Committee 2018). Engert, Fung, and Segendorf (2019) notice that Sweden has fewer bank branches that handle cash per inhabitant than Canada and suggest that access to cash through banks can play a role. As noted in Section 3, we have not included any variables for bank branches accepting cash in our empirical model. This is partly due to lack of data (it would be close to impossible to gather time series for that variable for all of our countries) and partly due to econometric (simultaneity) reasons. As in the case with ATMs, the number of bank offices offering cash services is likely to be determined in tandem with cash demand.

During the 1990s and early 2000s, the Riksbank reduced the number of cash distribution centers and thus withdrew implicit subsidies for cash. By 2014, the Riksbank only had one banknote distribution center. This differs from the situation in many other countries, where the central bank often has a much more prominent role in cash distribution. Since most of the reduction in the Riksbank's cash distribution centers happened prior to our sample period, it is impossible to include in the estimations. However, it could still be important, and could also have contributed to the reduction in commercial bank offices offering cash services.

Finally, it is worth noting that the increase in cash in circulation in many countries since the financial crisis is often due to higher demand for large-denomination notes—as documented by Engert, Fung, and Segendorf (2019) and Judson (2018)—while demand for small-denomination notes has fallen. The increased demand for cash is thus likely to be at least partly for store-of-value purposes. In Sweden, there was no similar increase in demand for cash during the financial crisis, nor has demand increased during the COVID-19 pandemic (Sveriges Riksbank 2020). This could be because there is strong trust in the ability and willingness of the Swedish government to protect money held in banks in times of crisis. Sweden has experienced two systemic banking crises during the last three decades, and public authorities have proven willing and able to protect commercial bank deposits. The payment systems have been up and running without interruptions, and no reductions have been applied to the value of commercial bank deposits. In other countries, which have not experienced similar systemic banking crises, there might be weaker trust in commercial bank money and, therefore, higher demand for cash for store-of-value purposes in times of financial turmoil.

We may conclude our discussion of why the model cannot explain the divergent development in Sweden, and what is special in Sweden, as follows. Several events and policy measures that have had mutually reinforcing effects on cash demand that are not captured in our model may explain the divergence. These include measures to reduce tax evasion and the informal sector, an aggressive banknote and coin changeover, the introduction of Swish, and the withdrawal of central bank subsidies to cash distribution. These factors are, however, hard to capture in an econometric time-series model covering multiple countries.

Interestingly, Norway—a country that also has a downward trend in cash demand that is not fully explained by the model—has had similar developments. Norway had a relatively aggressive banknote changeover, introduced an attractive mobile payments application, and has seen a reduction in bank offices. The Norwegian mobile application (Vipps) is almost identical to Swish, and it was introduced around the same time (2015). It has also reached the same degree of market penetration. However, regarding the other factors, development in Norway is somewhat less clear-cut or came later than in Sweden. The reduction in bank offices was somewhat less pronounced in Norway, and during our sample period Norwegian bank offices—in contrast to Swedish ones—were legally obliged to provide cash services. The Norwegian banknote changeover was also quite restrictive but happened later in our sample period.[20] Notice also that Norway had less cash in circulation in 2001 than Sweden, but the two countries are now at approximately the same level. We leave further comparative analysis of the developments in Norway and Sweden for later work.

[20]The Norwegian changeover started in 2017. In Norway, the old notes became invalid one year after the announcement date; the 100- and 200-krone note became invalid in May 2018, the 50- and 500-krone note in October 2019, and the 1,000-krone note in November 2020.

8. Conclusions

In this paper, we have analyzed developments in the amount of cash in circulation using a data set consisting of 129 developed and developing countries. Our main specification performs well in explaining cash developments for most OECD countries. We find that economic development, demography, and the level of the interest rate are key explanatory variables. We also find that better control of corruption is negatively related to the demand for cash in developed countries.

The development in Sweden consistently stands out. It is one of few countries where cash in circulation has decreased over the past couple of decades, not only as a share of GDP but since 2008 also in nominal terms. We find that our model cannot explain the divergent development in Sweden, while it performs relatively well for neighboring Norway, where cash in circulation has also declined. We discuss some events and policy measures that could have accelerated the decline in cash usage in Sweden. These include measures to fight tax evasion and an aggressive banknote and coin changeover. The combination of these measures, which had a negative influence on the incentives to hold and accept cash, combined with the rise of an electronic peer-to-peer alternative to cash (the mobile application Swish) has probably been decisive for developments in Sweden. However, it is not possible to reach a firm conclusion regarding the effects of these measures and events, as more detailed data is lacking.

With this paper, we have shed some light on the divergent development of cash in circulation in Sweden. Our empirical results and our discussion of some recent events in Sweden suggest that the demand for cash is shaped not only by general economic conditions but also by central bank policies, such as banknote and coin changeovers, government policies targeting tax evasion and the informal sector, and the competition in and the general workings of the banking sector.

Appendix

Table A.1. Country List

Non-OECD			OECD
Albania	Egypt	Papua New Guinea	Australia
Algeria	Equatorial Guinea	Paraguay	Canada
Angola	Eswatini	Philippines	Chile
Antigua and Barbuda	Fiji	Qatar	Czech Republic
Armenia	Gabon	Romania	Denmark
Azerbaijan	Georgia	Russia	Hungary
Bangladesh	Ghana	Rwanda	Iceland
Barbados	Grenada	Samoa	Israel
Belarus	Guatemala	Sao Tome and Principe	Japan
Belize	Guinea Bissau	Senegal	Mexico
Benin	Guyana	Serbia	New Zealand
Bhutan	Haiti	Seychelles	Norway
Bolivia	Honduras	Sierra Leone	Poland
Bosnia and Herzegovina	India	Singapore	South Korea
Botswana	Indonesia	Solomon Islands	Sweden
Brazil	Jamaica	South Africa	Switzerland
Brunei	Kazakhstan	Sri Lanka	Turkey
Bulgaria	Kenya	St. Kitts and Nevis	United Kingdom
Burkina Faso	Kuwait	St. Lucia	United States
Burundi	Kyrgyzstan	St. Vincent and the Grenadines	
Cambodia	Lesotho	Sudan	
Cameroon	Macao	Suriname	
Cape Verde	Malaysia	Tajikistan	
Central African Republic	Maldives	Tanzania	
Chad	Mali	Thailand	
China	Mauritius	Togo	
Colombia	Moldova	Tonga	
Comoros	Mongolia	Trinidad and Tobago	
Congo, Dem. Rep.	Morocco	Tunisia	
Congo, Rep.	Mozambique	Uganda	
Costa Rica	Myanmar	Ukraine	
Cote d'Ivoire	Namibia	United Arab Emirates	
Croatia	Nepal	Uruguay	
Djibouti	Nicaragua	Vanuatu	
Dominica	Niger	Zambia	
Dominican Republic	Nigeria		
	North Macedonia		
	Oman		
	Pakistan		

Table A.2. Adding Panel B Variables

	Full Sample				OECD			
	GMM (1)	**GMM (2)**	**GMM (3)**	**GMM (4)**	**GMM (5)**	**GMM (6)**	**GMM (7)**	**GMM (8)**
Interest Rate	–0.003**	–0.010***	–0.013***	–0.013***	–0.007***	–0.012***	–0.008**	–0.007*
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)
log GDP per capita	–0.136**	–0.134**	–0.231***	–0.229***	0.070*	0.024	0.027	0.021
	(0.059)	(0.058)	(0.059)	(0.060)	(0.038)	(0.036)	(0.040)	(0.034)
Age Dependency Ratio	0.008	0.005	0.009	0.009	0011**	0.008*	0.012*	0.012**
	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.004)	(0.006)	(0.006)
Self-Employed	–0.001	–0.004	–0.002	–0.002	0.004	–0.007	–0.011	–0.011
	(0.003)	(0.003)	(0.003)	(0.003)	(0.006)	(0.005)	(0.009)	(0.008)
Individuals Using Internet	–0.001	0.000	0.001	0.001	–0.001	–0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Human Rights	0.004	–0.001	0.036*	0.035	0.018	0.018	0.043	0.040
	(0.021)	(0.017)	(0.022)	(0.022)	(0.021)	(0.022)	(0.031)	(0.028)
Regulatory Quality	–0.002*	–0.001	–0.002	–0.002	–0.001	0.002	–0.002	–0.002
	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.003)	(0.002)
Control of Corruption	–0.001	0.000	–0.001	–0.001	–0.004***	–0.002	–0.004**	–0.004**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
World Uncertainty Index	–0.013				0.004			
	(0.029)				(0.022)			
Confidence in Financial Sector		–0.001				–0.001		
		(0.001)				(0.001)		
Public Trust in Politicians			0.001				0.002	
			(0.012)				(0.019)	
Organized Crime, 1–7 (Best)				0.006				–0.010
				(0.009)				(0.016)
Observations	1,455	831	893	893	288	190	190	190
Year FE	√	√	√	√	√	√	√	√

Note: The dependent variable is the natural logarithm of the cash-to-GDP ratio. Standard errors robust to clustering at country level are in parentheses. *, **, and *** represent the 10 percent, 5 percent, and 1 percent significance level, respectively.

References

Abadie, A., S. Athey, G. W. Imbens, and J. Wooldridge. 2017. "When Should You Adjust Standard Errors for Clustering?" Working Paper No. 24003, National Bureau of Economic Research.

Ahir, H., N. Bloom, and D. Furceri. 2019. "The World Uncertainty Index." Working Paper No. 19-027, Stanford Institute for Economic Policy Research.

Alvarez, F., and F. Lippi. 2009. "Financial Innovation and the Transactions Demand for Cash." *Econometrica* 77 (2): 363–402.

Amromin, G., and S. Chakravorti. 2009. "Whither Loose Change? The Diminishing Demand for Small-Denomination Currency." *Journal of Money, Credit and Banking* 41 (2–3): 315–35.

Angrist, J. D., and J.-S. Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Arango-Arango, C. A., and N. F. Suárez-Ariza. 2019. "Digital Payments Adoption and the Demand for Cash: New International Evidence." Borradores de Economia 1074, Banco de la Republica de Colombia.

Armelius, H., C. A. Claussen, and S. Hendry. 2020. "Is Central Bank Currency Fundamental to the Monetary System?" Discussion Paper No. 2-2020, Bank of Canada.

Ashworth, J., and C. A. Goodhart. 2020a. "Coronavirus Panic Fuels a Surge in Cash Demand." CEPR Discussion Paper No. 14910.

———. 2020b. "The Surprising Recovery of Currency Usage." *International Journal of Central Banking* 16 (3, June): 239–77.

Assenmacher, K., F. Seitz, and J. Tenhofen. 2019. "The Demand for Swiss Banknotes: Some New Evidence." *Swiss Journal of Economics and Statistics* 155 (1): 1–22.

Attanasio, O. P., L. Guiso, and T. Jappelli. 2002. "The Demand for Money, Financial Innovation, and the Welfare Cost of Inflation: An Analysis with Household Data." *Journal of Political Economy* 110 (2): 317–51.

Bagnall, J., D. Bounie, K. P. Huynh, A. Kosse, T. Schmidt, S. Schuh, and H. Stix. 2016. "Consumer Cash Usage: A Cross-Country Comparison with Payment Diary Survey Data." *International Journal of Central Banking* 12 (4, December): 1–62.

Bartzsch, N., G. Rösl, and F. Seitz. 2013. "Estimating the Foreign Circulation of Banknotes." *Economics Letters* 119 (2): 165–67.

Baumol, W. J. 1952. "The Transactions Demand for Cash: An Inventory Theoretic Approach." *Quarterly Journal of Economics* 66 (4): 545–56.

Bech, M. L., U. Faruqui, F. Ougaard, and C. Picillo. 2018. "Payments Are a-changin' but Cash Still Rules." *BIS Quarterly Review* (March): 67–80.

Brunnermeier, M. K., H. James, and J.-P. Landau. 2019. "The Digitalization of Money." Working Paper No. 26300, National Bureau of Economic Research.

Cameron, A. C., and D. L. Miller. 2015. "A Practitioner's Guide to Cluster-Robust Inference." *Journal of Human Resources* 50 (2): 317–72.

Carbo-Valverde, S., and F. Rodriguez-Fernandez. 2019. "An International Approach to the Cost of Payment Instruments: The Case of Cash." Unpublished Memo.

Cusbert, T., and T. Rohling. 2013. "Currency Demand during the Global Financial Crisis: Evidence from Australia." Research Discussion Paper No. 2013-01, Reserve Bank of Australia.

Engert, W., B. Fung, and B. Segendorf. 2019. "A Tale of Two Countries: Cash Demand in Canada and Sweden." Working Paper No. 376, Sveriges Riksbank.

Fariss, C. J. 2019. "Yes, Human Rights Practices Are Improving Over Time." *American Political Science Review* 113 (3): 868–81.

Hadri, K. 2000. "Testing for Stationarity in Heterogeneous Panel Data." *Econometrics Journal* 3 (2): 148–61.

Harris, R. D., and E. Tzavalis. 1999. "Inference for Unit Roots in Dynamic Panels Where the Time Dimension Is Fixed." *Journal of Econometrics* 91 (2): 201–26.

Hastie, T., R. Tibshirani, and M. Wainwright. 2015. *Statistical Learning with Sparsity: The Lasso and Generalizations.* CRC Press.

Heinonen, A. 2020. "Global Banknote Developments before and after Covid-19." https://www.centralbanking.com/central-banks/currency/banknotes/7691731/global-banknote-developments-before-and-after-covid-19.

Huynh, K., P. Schmidt-Dengler, and H. Stix. 2014. "The Role of Card Acceptance in the Transaction Demand for Money." Staff Working Paper No. 14-44, Bank of Canada.

Im, K. S., M. H. Pesaran, and Y. Shin. 2003. "Testing for Unit Roots in Heterogeneous Panels." *Journal of Econometrics* 115 (1): 53–74.

Jobst, C., and H. Stix. 2017. "Doomed to Disappear? The Surprising Return of Cash Across Time and Across Countries." CEPR Discussion Paper No. 12327.

Judson, R. 2018. "Big Note, Small Note: Central Bank Digital Currency and Cash." In *Do We Need Central Bank Digital Currency?* ed. E. Gnan and D. Masciandaro, Number 2018/2 in SUERF conference proceedings, 33–45 (chapter 3). Bocconi University and BAFFI CAREFIN.

Kaufmann, D., A. Kraay, and M. Mastruzzi. 2011. "The Worldwide Governance Indicators: Methodology and Analytical Issues." *Hague Journal on the Rule of Law* 3 (2): 220–46.

Keynes, J. M. 1937. "The General Theory of Employment." *Quarterly Journal of Economics* 51 (2): 209–23.

Schmiedel, H., G. L. Kostova, and W. Ruttenberg. 2012. "The Social and Private Costs of Retail Payment Instruments: A European Perspective." ECB Occasional Paper No. 137.

Seitz, F., B. Fischer, and P. Köhler. 2004. "The Demand for Euro Area Currencies: Past, Present and Future." ECB Working Paper No. 330.

Shirai, S., and E. A. Sugandi. 2019. "What Explains the Growing Global Demand for Cash?" Working Paper No. 1006, Asian Development Bank Institute.

Sveriges Riksbank. 2012. "50-krona and 1,000-krona Banknotes without Foil Strips Will Become Invalid after 31 December 2013." Press Notice, November 8.

———. 2018a. "Banknote and Coin Changeover in Sweden: Summary and Evaluation."

———. 2018b. "Communication Measures during the Banknote & Coin Changeover."

———. 2020. "Payments in Sweden 2020." https://www.riksbank.se/en-gb/payments--cash/paymen-in-sweden/payments-in-sweden-2020/.

Swedish Tax Agency. 2012. "Taxes in Sweden 2012. An English Summary of Tax Statistical Yearbook of Sweden."

The Riksbank Committee. 2018. "Tryggad tillgång till kontanter [Secure Access to Cash] (in Swedish)." SOU No. 2018:42, Ministry of Finance.

Tibshirani, R. 1996. "Regression Shrinkage and Selection via the Lasso." *Journal of the Royal Statistical Society: Series B (Methodological)* 58 (1): 267–88.

Tillväxtanalys. 2019. "Utvärdering av RUT-avdraget—Effekter på företagens tillväxt och överlavnad [Evaluation of the RUT Tax Deduction—Effects on Firm Growth and Survival] (in Swedish)." Report No. 2019:08, Swedish Agency for Growth Policy Analysis.

Tobin, J. 1956. "The Interest-Elasticity of Transactions Demand for Cash." *Review of Economics and Statistics* 38 (3): 241–47.

Wakamori, N., and A. Welte. 2017. "Why Do Shoppers Use Cash? Evidence from Shopping Diary Data." *Journal of Money, Credit and Banking* 49 (1): 115–69.

Wright, R., E. Tekin, V. Topalli, C. McClellan, T. Dickinson, and R. Rosenfeld. 2017. "Less Cash, Less Crime: Evidence from the Electronic Benefit Transfer Program." *Journal of Law and Economics* 60 (2): 361–83.

Is Inflation Domestic or Global? Evidence from Emerging Markets*

Rudolfs Bems, Francesca Caselli, Francesco Grigoli, and Bertrand Gruss
International Monetary Fund

Following a period of disinflation during the 1990s and early 2000s, inflation in emerging markets has remained remarkably low. The volatility and persistence of inflation also fell considerably and remained low despite large swings in commodity prices, the global financial crisis, and periods of strong and sustained U.S. dollar appreciation. A key question is whether this improved inflation performance is sustainable or reflects global disinflationary forces that could prove temporary. In this paper, we use a New Keynesian Phillips-curve framework and data for 19 large emerging market economies over 2004–18 to assess the contribution of domestic and global factors to domestic inflation dynamics. We find that long-term inflation expectations, linked to domestic factors, were the main determinant of inflation. External factors played a considerably smaller role. The results suggest that although emerging markets are increasingly integrated into the global economy, policymakers still hold significant leverage in domestic inflation developments.

JEL Codes: E31, E58, F62.

*The views expressed in this paper are those of the authors and do not necessarily represent those of the IMF or IMF policy. We thank, without implicating, Oya Celasun, Gian Maria Milesi-Ferretti, Petra Geerats, Weicheng Lian, Maury Obstfeld, Rafael Portillo, Jay Shambaugh, and the participants to the International Monetary Fund–IMF Economic Review–Central Bank of Chile joint conference "Current Policy Changes Facing Emerging Markets" in Santiago, Chile in July 2019 for their comments and suggestions. Jungjin Lee and Jilun Xing provided excellent research assistance. Author e-mails: rbems@imf.org, fcaselli@imf.org, fgrigoli@imf.org, bgruss@imf.org.

1. Introduction

There is a lively debate about the so-called globalization of inflation hypothesis—that is, whether globalization is responsible for a weakening in the relationship between inflation and domestic slack and a strengthening in the relationship between inflation and global factors (International Monetary Fund 2005b; Ball 2006; Fisher 2006; Kohn 2006; Yellen 2006; and Carney 2017). The empirical evidence, which is almost entirely limited to advanced economies, is mixed: Ihrig et al. (2010) find little support for an increasing role of global factors in the inflation process, while Borio and Filardo (2007) and Auer, Borio, and Filardo (2017) argue that the role of global factors increased since the 1990s.[1] More recently Ha, Kose, and Ohnsorge (2019) turned the attention towards emerging markets. They find that global shocks contributed more to domestic inflation variation in advanced economies than in emerging market and developing economies. However, while global shocks became more important over time, domestic shocks still account for the largest share in domestic inflation variation in both country groups. Even though most of the attention on the role of external factors in inflation dynamics focused on advanced economies—owing chiefly to the underwhelming reaction of prices to the global financial crisis and the subsequent wage puzzles[2]—this is a particularly relevant issue for understanding the recent macroeconomic performance of increasingly globalized emerging markets.

Following a period of disinflation during the 1990s and early 2000s, inflation in emerging markets has been, on average, remarkably low and stable (IMF 2016, 2018; and Ha, Kose, and Ohnsorge 2019). Even in the aftermath of large commodity price swings, the global financial crisis, and sizable appreciation of the U.S. dollar, inflation in most countries was quick to stabilize, and the short-lived effects of inflationary shocks, in turn, allowed central banks to cut interest rates to fight off recessions. A confluence of economic factors, including improved domestic policy frameworks (Rogoff 2003;

[1]The discussion of whether globalization has an impact on domestic inflation applies in the short to medium term, as in the long run the rate of inflation is set by monetary policy (Ihrig et al. 2010).

[2]See, for instance, IMF (2013), Danninger (2016), and Draghi (2017).

IMF 2005a; and Vegh and Vuletin 2014) and global disinflationary forces (Carney 2017; and Auer, Levchenko, and Sauré 2019) have likely affected the recent inflation performance in emerging markets.

This paper examines the underpinnings of the recent inflation experience in emerging markets. We review the inflation performance in a sample of 19 large emerging markets over the past few decades and quantify the impact of domestic and global factors in determining inflation dynamics since the start of the post-disinflation period in the mid-2000s.[3] To do so, we rely on a hybrid variant of the New Keynesian Phillips curve that is augmented with foreign variables (similar to Borio and Filardo 2007, Ihrig et al. 2010, and Auer, Borio, and Filardo 2017 for advanced economies) and estimate the determinants of domestic core and headline inflation over 2004–18.

Our results show that long-term inflation expectations are the main factor driving inflation from target and inflation variability. Although the reduced-form nature of the analysis carries some limitations, we find evidence that inflation expectations reflect domestic developments and the impact of global factors on inflation expectations is marginal when compared with that of domestic factors. Beyond inflation expectations, we find that while some external factors, such as foreign price pressures, have a statistically significant impact on domestic inflation, they played a relatively small role in driving inflation dynamics in our sample. Our findings also reveal significant cross-country heterogeneity, and that there is still significant room for improvement in inflation performance in some emerging markets from further reductions in the level and variability of long-term inflation expectations.

Overall, our results indicate that domestic rather than global factors were the main contributor to the gains in inflation performance among emerging markets since the mid-2000s. They suggest that although these economies are increasingly interconnected with

[3]The country coverage is defined by data availability of long-term (that is, three-year-ahead and longer) forecasts for inflation and a minimum population of two million people. It includes the following countries: Argentina, Brazil, Bulgaria, Chile, China, Colombia, Hungary, India, Indonesia, Malaysia, Mexico, Peru, Philippines, Poland, Romania, Russia, South Africa, Thailand, and Turkey.

the global economy, their policymakers still have significant leverage on domestic inflation developments.[4]

The rest of the paper is organized as follows. Section 2 discusses the globalization of inflation hypothesis and presents some stylized facts about the recent inflation performance in our sample of emerging markets. Section 3 presents the empirical analysis, starting with the estimation of the Phillips curve, moving to the quantification of the contributions of domestic and global factors, and concluding with a battery of robustness tests. Section 4 reports a few concluding remarks.

2. Background

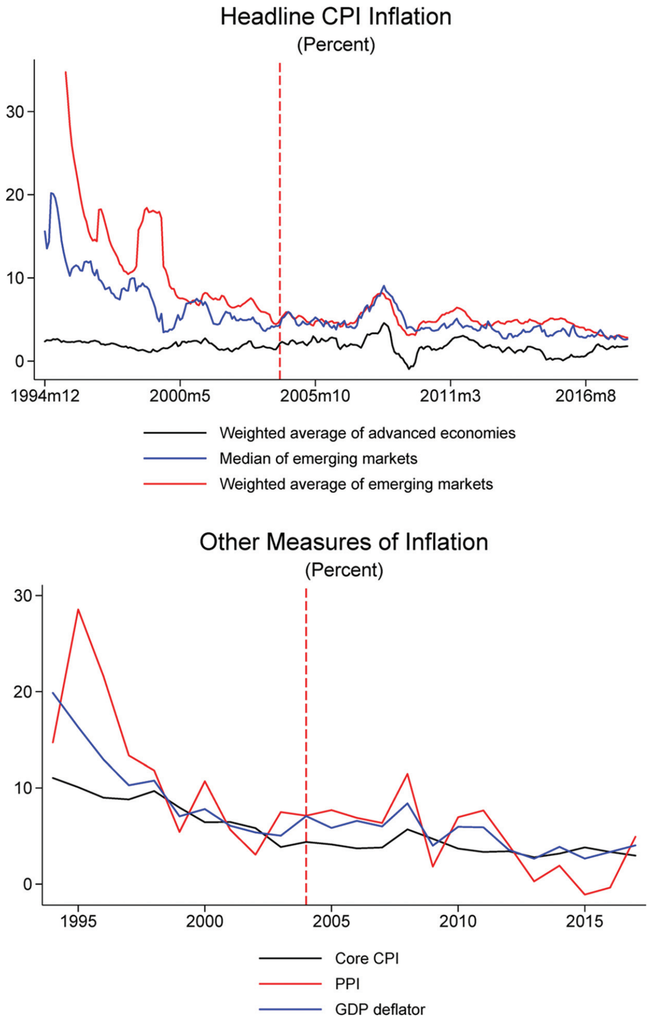
This section first reviews the inflation performance for the 19 emerging markets in the sample. It then introduces the two main arguments that could help explain such performance: the globalization of inflation hypothesis, which relates the integration of emerging markets in the global economy with the price dynamics; and the adoption of rule-based policy frameworks (such as inflation targeting and fiscal rules), which are likely to have strengthened predictability in policy decisions and increased price stability.

2.1 Inflation Performance in Emerging Markets

Following a period of disinflation during the 1990s and early 2000s, inflation in emerging markets remained relatively low and stable. The upper panel of Figure 1 shows that the weighted average of headline consumer price index (CPI) inflation (hereafter, headline inflation) for the 19 emerging markets in the sample declined

[4]A recent related paper by Jaŝová, Moessner, and Takáts (2018) examines the role of the global output gap in driving inflation in emerging markets in a similar New Keynesian Phillips-curve framework. The main difference between this study and the current paper concerns the measurement of inflation expectations. While Jaŝová, Moessner, and Takáts (2018) use short-term (end-year) inflation expectations, we focus on the role of long-term inflation expectations (three years ahead and beyond). Given the persistence inherent in the inflation process, we see inflation expectations for longer horizons as essential for capturing the link between the extent of anchoring and inflation in emerging markets.

Figure 1. Disinflation in Emerging Markets (percent)



Source: Haver Analytics; IMF, World Economic Outlook; and authors' calculations.

Note: The vertical dashed line marks the start of the post-disinflation period. The vertical axis in the upper panel is truncated at 35 percent to ease visualization. Weighted averages are constructed using weights of nominal GDP, expressed in U.S. dollar terms, for 2010–12. The lines in the lower panel denote medians across sample emerging markets of each indicator.

dramatically—by more than a 100 percentage points[5] from 1995 to 2004—and leveled off at about 5 percent thereafter, which is about 3 percentage points higher than the weighted average of advanced economies.[6] (For figures in color, see the online version of the paper at http://www.ijcb.org.) Median headline inflation, which abstracts from a few hyperinflation episodes of the 1990s, still shows a significant decline from about 20 percent to about 5 percent since 2004.

We now turn to other measures of price inflation. The inflation rate for core CPI (hereafter, core inflation), which excludes food and energy items (typically characterized by more volatile prices), also declined until the mid-2000s and remained low and stable thereafter, as shown in the lower panel of Figure 1. The inflation rate of producer prices fell drastically during the 1990s and remained at relatively low levels ever since. Finally, GDP deflators, which encompass the prices of all domestically produced final goods and services, exhibit the same pattern.

Despite this generalized decline in inflation rates across emerging markets, there is some heterogeneity. To illustrate this, Figure 2 shows the share of emerging markets in the sample with inflation rates exceeding 10 percent. In the late 1990s, about half of the countries in the sample experienced inflation rates above 10 percent. Since 2004, such share declined significantly, yet one country out of 10 emerging markets still experienced relatively high inflation rates.

Figure 3 shows that inflation volatility—defined as the standard deviation of detrended inflation—has been stable or declining in emerging markets since 2004.[7] While the volatility of core inflation toward the end of the sample became broadly comparable to the average level observed for advanced economies, the volatility of

[5]For ease of visualization, the vertical axis of the upper panel in Figure 1 is truncated at 35 percent.

[6]The 19 countries in the sample constitute 80 percent of the GDP of all emerging market and developing economies.

[7]The decline in the volatility of inflation rates is not driven by exchange rate behavior, as there is no clear evidence of a decline in the volatility of exchange rate movements since the late 1990s. See Ilzetzki, Reinhart, and Rogoff (2017) for a discussion of changes in de facto exchange rate volatility.

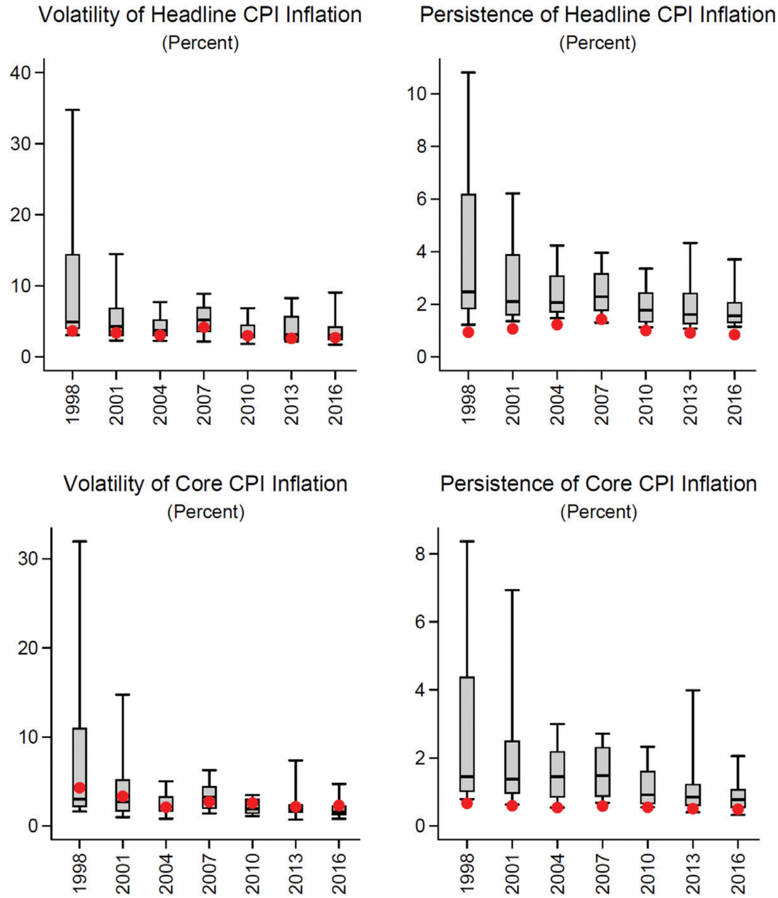
Figure 2. Share of Countries with Double-Digit Inflation (percent)

Source: Haver Analytics; and authors' calculations.
Note: The vertical dashed line marks the start of the post-disinflation period.

headline inflation remained somewhat higher. Inflation persistence—defined as the tendency for price shocks to elevate inflation above its long-term level for a prolonged period—also declined gradually during the sample period, even though it remains somewhat above the level observed in advanced economies.[8] One aspect that may explain the higher volatility for headline inflation in emerging markets is that food and other commodities, whose prices tend to be more volatile, account for a larger share of their consumption baskets than in advanced economies. Higher volatility and persistence of inflation in emerging markets than in advanced economies may also reflect higher pass-through of external shocks to local prices as a result of larger dollar import invoicing shares (Bonadio, Fischer,

[8]We calculate inflation persistence following Stock and Watson (2007, 2010). The approach consists of decomposing inflation, π_t, into a permanent component, ζ_t, and a transitory component, η_t, where $\zeta_t = \zeta_{t-1} + \epsilon_t$ and η_t and ϵ_t are independently normally distributed with time-varying variances $\sigma^2_{\eta,t}$ and $\sigma^2_{\epsilon,t}$, respectively. The measure of inflation persistence underlying the calculations in Figure 3 is the estimated standard deviation of the shock to the permanent component of inflation.

Figure 3. Inflation Dynamics



Source: Haver Analytics; and authors' calculations.
Note: The volatility is computed as the standard deviation of detrended inflation. Persistence is calculated as the standard deviation of the permanent component of inflation based on Stock and Watson (2007). The horizontal lines in each box denote the medians, the upper and the lower edges of each box show the top and bottom quartiles, the vertical lines denote the ranges between the top and bottom deciles, and the red dots denote the averages for advanced economies. The labels on the horizontal axis denote the start of the three-year windows.

and Sauré 2020; Gopinath et al. 2020) and monetary policy institutions and frameworks that are less developed and credible, and thus less effective.[9]

There is, however, substantial cross-country heterogeneity in terms of volatility and persistence of inflation among emerging markets. Either in the case of headline inflation or core inflation, the cross-country distribution for the latest observation covering 2016–18 suggests that the volatility and persistence of inflation for 10 percent of the sample are about two to three times higher than for the median country. Similarly to inflation levels, we conclude that there is some cross-country heterogeneity with respect to the improvements in inflation volatility and persistence.

2.2 The Globalization of Inflation Hypothesis

The globalization of inflation hypothesis posits that, as economies deepen their level of integration in the global markets, prices end up being driven by external factors. The discussion dates back to the oil price swings of the 1970s, but it gained renewed prominence in the context of increased exports from low-wage countries. That is, global competition could lead to downward pressure on prices. However, the evidence on the effects of globalization on inflation is mixed. In the United States, for instance, although globalization could be assimilated to a supply shock that temporarily reduced inflation, it did not in fact affect the underlying inflation process (Ball 2006).

A critical aspect of globalization is that the global supply chain became increasingly integrated, and with that the possibility of outsourcing and offshoring raised the degree of substitutability of production stages (Auer, Levchenko, and Sauré 2019). Thus, it might be economically convenient to relocate production where slack is larger to enjoy lower costs. A related argument is that the increased ability to purchase final goods from the cheapest locations led to

[9]See Mishkin (2007) for a discussion of how better monetary policy can contribute to a decline in inflation persistence. Bems et al. (2018) document that inflation expectations are on average better anchored in advanced economies than in emerging markets, and show that external shocks tend to have a more persistent effect on domestic inflation when inflation expectations are worse anchored.

greater price competition.[10] Thus, as it is easier to move production abroad, domestic prices should display a stronger sensitivity to external conditions (Borio and Filardo 2007; Auer, Levchenko, and Sauré 2019). In addition, if these factors weigh on the bargaining power of workers, the relationship between domestic slack and wage (or price) inflation would become weaker. For example, a higher share of imports from low-wage countries and competition in traded goods could make it more difficult for domestic firms to adjust prices when labor market conditions are tight and workers demand higher wages (Auer, Degen, and Fischer 2013).

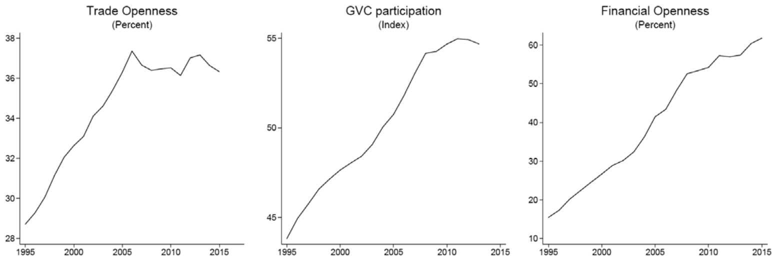
Globalization and increased market contestability can affect inflation indirectly via the components of the Phillips curve, including the domestic output gap and/or inflation expectations. If this is true, foreign factors become progressively more dominant in shaping inflation dynamics, and, in the words of Auer, Borio, and Filardo (2017), the Phillips-curve equation should take a more "globe-centric" view of the inflation process—for example, by including the foreign output gap (Borio and Filardo 2007; Ihrig et al. 2010; and Auer, Borio, and Filardo 2017).

In the past few decades the process of integration in emerging markets was remarkably intense. Emerging markets went from producing a third of global output in the 1990s to more than half. Figure 4 shows that trade openness increased steadily since 1995 for our sample of 19 countries and leveled off thereafter. The participation in global value chains (GVCs) also shows a marked increase over the past two decades, reflecting the intensification of outsourcing and offshoring of production.[11] The flip side of the increase in GVC participation is a deeper financial integration through foreign direct investment and portfolio investment. As a result, the

[10]The higher concentration of market power in some firms, however, could hamper these effects. As noted by Autor et al. (2020), market power can translate to pricing power.

[11]The GVC participation index is calculated as the sum of backward participation (imported intermediate inputs used to generate output for export) and forward participation (that is, exports of intermediate goods used as inputs for the production of exports of other countries) as a ratio of gross exports (see Aqib, Novta, and Rodrigues-Bastos 2017 for more details about the global value chain participation measure).

Figure 4. Integration of Emerging Markets into Global Markets



Source: Aqib, Novta, and Rodrigues-Bastos (2017); IMF, Balance of Payments Statistics; IMF, World Economic Outlook; and authors' calculations.
Note: Trade openness is defined as imports in percent of GDP. GVC participation is defined as the sum of backward participation (imported intermediate inputs used to generate output for export) and forward participation (exports of intermediate goods used as inputs for the production of exports of other countries) as a ratio of gross exports; financial openness is defined as the sum of foreign direct investment and portfolio equity liabilities in percent of GDP. All variables are expressed as five-year moving averages.

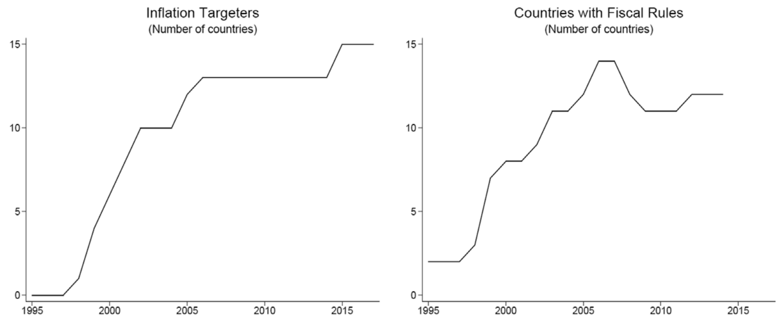
average financial openness indicator for these economies displays a significant surge.

2.3 Institutional Changes

The literature stresses how an independent central bank and sound and sustainable fiscal policy are key attributes for the credibility of monetary policy (Mishkin 2000; Mishkin and Savastano 2001) and therefore potential drivers of the extent of anchoring of inflation expectations and inflation performance. Other papers find that transparency about the objective and conduct of monetary policy is also a key determinant of inflation expectations. Finally, some studies find an association between fiscal institutions and credibility on the one hand, and inflation performance and the anchoring of inflation expectations on the other hand (Combes et al. 2017; Montes and Acar 2018), or a link between expected fiscal performance and inflation expectations (Celasun, Gelos, and Prati 2004).

The last two decades witnessed important institutional changes in emerging markets, as shown in Figure 5. Out of the 19 countries in the sample, the number of inflation targeters increased from zero

Figure 5. Adoption of Policy Frameworks



Source: IMF, Balance of Payments Statistics; national authorities; and authors' calculations.
Note: The number of countries with fiscal rules is the sum of the countries with any fiscal rule, as defined in the IMF Fiscal Rules Dataset (2016).

in 1995 to 15 in 2017. At the same time, the number of countries with some type of fiscal rule rose from 2 to 14 in 2007; by 2011, it fell to 11 as Argentina, India, and Russia suspended their fiscal rules, and rose again to 12 when Russia implemented a new fiscal rule in 2013. These institutional changes towards rule-based policymaking generally come with increased price stability and some predictability in policy decisions. If this is the case, the sensitivity of inflation to domestic factors may have increased.

3. The Role of Domestic and Global Factors: An Empirical Assessment

The empirical analysis to uncover the role of domestic and foreign factors in determining inflation consists of two stages. The first stage estimates a Phillips curve augmented with variables proxying external factors for a panel of 19 emerging markets using quarterly data from the first quarter of 2004—the start of the post-disinflation period—to the first quarter of 2018.[12] After establishing the statistical significance of the inflation determinants, the second stage

[12]The results are broadly unchanged if the start of the disinflation period is set to any quarter of 2004 or 2005.

explores the contribution of domestic and foreign factors to inflation variation, across countries and over time.

Excluding the inflationary period prior to 2004 allows us to focus on a time frame of current relevance, during which price stability was at the forefront of the monetary frameworks in emerging markets. The pre-2004 sample is, instead, characterized by the presence of several runaway inflation episodes (and the subsequent disinflation periods), which are associated with large exchange rate devaluations as a result of specific events rather than factors studied in the paper.[13]

The section concludes by presenting a set of tests to ensure the robustness of the results.

3.1 An Augmented Phillips-Curve Framework

3.1.1 Empirical Strategy

The analysis relies on a hybrid variant of a standard New Keynesian Phillips curve (Galí and Gertler 1999; Galí, Gertler, and Lopez-Salido 2001, 2003). Drawing from the literature, the specification is augmented with variables that serve as proxies for macro developments abroad (Borio and Filardo 2007; Ihrig et al. 2010; and Auer, Borio, and Filardo 2017). Formally, we estimate the following equation:

$$\pi_{i,t} = \gamma^b \pi_{i,t-1} + \gamma^f \pi^e_{i,t} + \beta Y^{gap}_{i,t} + \theta Z^*_{i,t} + \eta_i + \epsilon_{i,t} \tag{1}$$

in which π is either core inflation or headline inflation; π^e denotes three-year-ahead inflation expectations; Y^{gap} is the domestic output gap; Z^* is a vector of external variables that includes, depending on

[13]These include, among others, the military coup in Turkey in 1997, the abandonment of the Convertibility Plan in Argentina in 2001 and the currency board in Bulgaria in 1997, the elimination of subsidies in 1997 in Romania, the financial crisis in Russia in 1998, and the effects of the Asian crisis in Indonesia in 1998. As a result, the standard deviation of core (headline) inflation in the period 1997–2003—the period for which we can retrieve information on long-term inflation expectations and inflation targets—is about 10 (5) times the standard deviation of the period 2004–18. Core and headline inflation peaked at 519.1 percent and 583.5 percent during 1997–2003, respectively, compared with 28.3 percent for core inflation and 53.2 percent for headline inflation during 2004–18.

the specification, the import-weighted foreign output gap, an indicator for external price pressure in the previous period, and the lag of energy and food price inflation. Differently from Borio and Filardo (2007), Ihrig et al. (2010), and Auer, Borio, and Filardo (2017), we include the foreign output gap *and* external price pressure in the specification to capture both demand and supply shocks.[14] η_i denotes country fixed effects; ϵ is the error term; and i and t are the subindexes for the country and the time period, respectively.[15]

Inflation expectations, a key variable in the analysis, are from Consensus Economics and report the average of inflation forecasts across professional forecasters.[16] These forecasts are available biannually up to 2014 and at quarterly frequency thereafter. In the case of South Africa, the source is the Bureau for Economic Research, and data are available at quarterly frequency for the entire sample period. In all cases, inflation expectations are based on headline inflation forecasts, but it should be noted that the CPI definition may have changed over time.

Among the variables in vector Z^*, the variable capturing external price pressures is defined as the percent change in the import-weighted producer price index of countries from which country i imports, converted to local currency using the nominal effective exchange rate, and relative to the percent change in the GDP deflator:[17]

$$\Delta P^*_{i,t} = \Delta mPPI_{i,t} + \Delta neer_{i,t} - \Delta P_{i,t} \tag{2}$$

[14]Energy price inflation and food price inflation, which are constructed interacting global prices by the weight of energy and food products in the price index, are not included in the specifications for core inflation.

[15]Despite the relatively high correlation between inflation expectations and past inflation, the variance inflation factor is well below 10 for all explanatory variables, ruling out multicollinearity concerns.

[16]The use of inflation forecasts collected through surveys covering professional forecasters is standard in the literature. However, some studies documented significant differences between forecasts of households and firms and those of professional analysts (see, for instance, Mankiw, Reis, and Wolfers 2003). However, such surveys are only available for a handful of countries and their methodologies are not necessarily comparable across countries.

[17]One may argue that, when pass-through from external to domestic prices is high, the external price pressure variable would understate the impact of external prices. While this is true, the pass-through within the same quarter to a broad measure of domestic prices such as the GDP deflator is likely to be limited.

in which $P_{i,t}$ is the natural logarithm of country i's GDP deflator. The change in the import-weighted foreign producer price index is given by

$$\Delta mPPI_{i,t} = \sum_{j=1}^{J} \omega_{ij,t} \Delta PPI_{j,t}, \tag{3}$$

where $i \neq j$, $PPI_{j,t}$ is the natural logarithm of country j's producer price index. And the change in the nominal effective exchange rate is constructed as the change in the bilateral exchange rate of each trading partner vis-à-vis the U.S. dollar, weighted by their import shares (Gopinath 2015; and Carriere-Swallow et al. 2016):[18]

$$\Delta neer_{i,t} = \sum_{j=1}^{J} \omega_{ij,t} (\Delta e_{i,t} - \Delta e_{j,t}), \tag{4}$$

where $i \neq j$, $e_{i,t}$ is the natural logarithm of country i's bilateral exchange rate (expressed in local currency per U.S. dollar, so that an increase denotes a depreciation of the domestic currency); and Δ is the first difference operator.

The foreign output gap is defined as

$$Y_{i,t}^{*gap} = \sum_{j=1}^{J} \omega_{ij,t} Y_{j,t}^{gap}, \tag{5}$$

where $i \neq j$, $\omega_{ij,t}$ is the share of exports from country j to country i in country i's total imports (lagged one year and measured annually), and $Y_{j,t}^{gap}$ is the Hodrick-Prescott filtered series of real GDP of country j.

We estimate the baseline specification employing median regressions to account for a few extreme observations. Alternatively, the analysis uses robust regressions, which downplay the influence of outliers, and constrained regressions that restrict the sum of the coefficients on past inflation and inflation expectations to be equal

[18] See also Auer, Chaney, and Sauré (2018) for a discussion of pass-through determinants at the firm level and Vogel (2008) for a discussion of firm-level pricing strategy.

to one. Although potential endogeneity is a limitation for the estimation techniques used, the structure of the data (with gaps in the first part of the sample because inflation expectations are available at lower frequency) prevents the use of estimators that rely on lags, such as the system generalized method of moments.

While we first present the results of the standard hybrid New Keynesian Phillips-curve estimation that controls for foreign variable, from the outset we want to rule out the possibility that inflation expectations might reflect global developments rather than domestic factors, thereby overestimating the role of the latter. To do that, we employ a two-stage approach in which we first run a regression of inflation expectations on foreign price pressure, foreign output gap, and country and time fixed effects. This step effectively purges the inflation expectations variable of external factors. Then, in a second stage, we modify the baseline specification to replace inflation expectations with the residual from the first stage, which is orthogonal to all foreign factors (and to domestic effects co-moving over time and fixed across countries).[19]

3.1.2 Estimation Results

Table 1 presents the estimation results. Overall, the explanatory variables account for 52 percent (44 percent) of the variation of core (headline) inflation. The findings suggest that price setting was, to some extent, forward looking, with a coefficient on three-year-ahead inflation expectations of 0.6 in the regressions for core inflation and ranging between 0.4 and 0.6 in the regressions for headline inflation.[20] Domestic cyclical conditions, for which the output gap serves

[19]That said, foreign shocks that have an impact on the domestic output gap but are not captured by changes in the foreign output gap and the external price pressure variable can also lead to a downward bias in the estimated contribution of global factors. On the other hand, some of the fluctuations in the exchange rate embedded in the external price pressure variable can be due to domestic factors, potentially biasing the estimated contribution of foreign factors upward. Further tests can be found in Section 3.4.

[20]Argentina does not have data for core inflation. Moreover, headline inflation statistics have been heavily criticized and Cavallo (2013) shows that inflation calculated using online prices is about three times higher than the official estimates. To deal with this, we rely on estimates from the IMF's country team. Also, to ensure that our results are not dependent on Argentina, we run the regressions of

Table 1. Hybrid Phillips-Curve Estimation, Specifications Augmented for External Factors

	Core Inflation			Headline Inflation		
	Median Regression (1)	**Robust Regression (2)**	**Constrained Regression (3)**	**Median Regression (4)**	**Robust Regression (5)**	**Constrained Regression (6)**
Inflation Expectations Three Years Ahead	0.587***	0.631***	0.566***	0.396***	0.303***	0.564***
	(0.111)	(0.077)	(0.062)	(0.134)	(0.067)	(0.088)
Lag of Core/Headline Inflation	0.494***	0.500***	0.434***	0.422***	0.481***	0.436***
	(0.037)	(0.023)	(0.062)	(0.047)	(0.028)	(0.088)
Output Gap	0.159***	0.168***	0.103	0.188**	0.182***	0.110
	(0.045)	(0.037)	(0.070)	(0.086)	(0.067)	(0.095)
Lag of External Price Pressure	0.018***	0.018***	0.032***	0.005	–0.001	0.020
	(0.004)	(0.003)	(0.011)	(0.008)	(0.007)	(0.014)
Foreign Output Gap	0.021	0.060	0.070	0.117	0.085	0.169
	(0.050)	(0.053)	(0.100)	(0.087)	(0.103)	(0.130)
Lag of Food Price Inflation				0.013***	0.018***	0.025***
				(0.004)	(0.004)	(0.006)
Lag of Energy Price Inflation				0.000	–0.001	–0.001
				(0.002)	(0.002)	(0.003)
Countries	18	18	18	19	19	19
Observations	633	633	633	668	668	668
R-squared	0.525			0.445		

Source: Authors' calculations.
Note: All specifications include country fixed effects. Constrained regressions force the sum of the coefficients on past inflation and expected inflation to be one. Median regressions report the pseudo R-squared. Robust standard errors are in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

as a proxy, also matter, but the size of the impact is small in economic terms: a 1 percentage point increase in the output gap is associated with an increase in the core headline inflation rate by 0.2 percentage point.

With respect to the external variables, the foreign output gap is not significant, even if the external price pressure variable is excluded from the specification. This is in contrast to the results of Borio and Filardo (2007) and Auer, Borio, and Filardo (2017) for advanced economies, which find that foreign slack affects domestic inflation. External price developments, on the other hand, are an important determinant of inflation, as indicated by the positive and significant coefficient on the lag of external price pressure (food price inflation) in the regressions for core (headline) inflation. The effects, however, are economically small: a 1 percentage point increase in the external price pressure variable (food price inflation) is associated with an increase of 0.02 to 0.03 (0.01 to 0.02) percentage point in the core (headline) inflation rate.

We now turn to the results of the two-stage approach in Table 2. Columns 1 and 5 report the estimations of the first-stage regressions, which aim at purging inflation expectations from all external factors and where we include time fixed effects to remove all co-movements across countries. When we only control for external price pressure and foreign output gap in column 1, the coefficient on the former turns out not statistically significant and the latter is positive and only marginally statistically significant. Switching to headline inflation, in column 5 the coefficient on external price pressure becomes negative and statistically significant, while the one on foreign output gap remains positive. Importantly, the results of the second-stage regressions in columns 2 to 4 for core inflation and 6 to 8 for headline inflation are remarkably similar to the ones obtained in the one-stage regressions. The coefficients on inflation expectations are only marginally smaller, confirming that inflation expectations are mostly driven by domestic factors.

headline inflation excluding it. While the coefficient on foreign output gap sometimes turns significant, the one on inflation expectations gets larger. All in all, the results are qualitatively comparable to the ones discussed in this section.

Table 2. Hybrid Phillips-Curve Two-Stage Estimation, Specifications Augmented for External Factors

		Core Inflation				Headline Inflation		
	First Stage for Infl. Exp. (1)	Median Regression (2)	Robust Regression (3)	Constrained Regression (4)	First Stage for Infl. Exp. (5)	Median Regression (6)	Robust Regression (7)	Constrained Regression (8)
Residual of Infl. Exp. Three Years Ahead		0.464***	0.485***	0.329***		0.384***	0.265***	0.430***
		(0.108)	(0.090)	(0.062)		(0.122)	(0.071)	(0.074)
Lag of Core/Headline Price Inflation		0.529***	0.562***	0.671***		0.459***	0.513***	0.570***
		(0.032)	(0.022)	(0.062)		(0.044)	(0.027)	(0.074)
Output Gap		0.179***	0.182***	0.076		0.139*	0.171**	0.112
		(0.033)	(0.037)	(0.078)		(0.082)	(0.067)	(0.103)
Lag of External Price Pressure	–0.001	0.018***	0.016***	0.027**	–0.058***	–0.001	–0.005	0.009
	(0.002)	(0.003)	(0.003)	(0.011)	(0.015)	(0.007)	(0.007)	(0.014)
Foreign Output Gap	0.225*	–0.008	0.041	0.044	1.083**	0.159	0.088	0.122
	(0.129)	(0.042)	(0.052)	(0.097)	(0.530)	(0.176)	(0.103)	(0.125)
Lag of Food Price Inflation					0.122***	0.015***	0.020***	0.029***
					(0.035)	(0.004)	(0.004)	(0.006)
Lag of Energy Price Inflation					–0.068***	0.001	–0.001	–0.002
					(0.025)	(0.002)	(0.002)	(0.003)
Countries	18	18	18	18	19	19	19	19
Observations	633	633	633	633	668	668	668	668
R-squared	0.845	0.518			0.797	0.444		

Source: Authors' calculations.
Note: All specifications include country fixed effects. Constrained regression force the sum of the coefficients on past inflation and expected inflation to be one. Median regressions report the pseudo R-squared. Robust standard errors are in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

3.2 Contributions to Inflation Deviations from "Target"

After establishing that both domestic and external factors play a role in determining inflation, we use the estimated panel coefficients in Table 1 to compute the country-specific contributions of the explanatory variables.[21] Following Yellen (2015), we calculate the contributions to inflation in each quarter for each regression by taking into account the persistence of the inflation process:

$$C_{i,t}^{x} = C_{i,t-1}^{x}\gamma^{b} + (\varphi^{x} x_{i,t}), \tag{6}$$

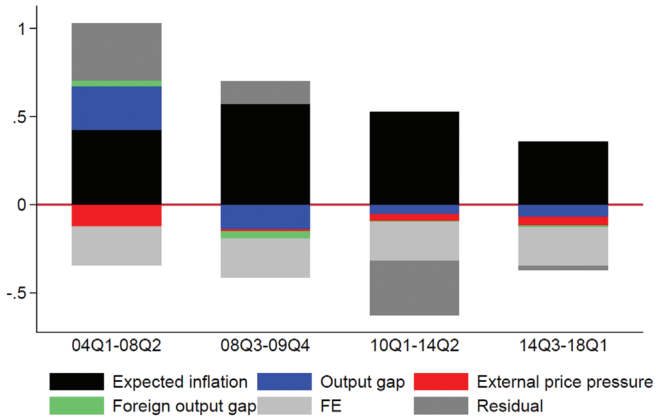
where $C_{i,t}^{x}$ is the contribution to inflation dynamics in country i at period t of each explanatory variable x in vector $X = \left[\pi^{e}, Y^{Gap}, Z^{*}, \eta_i\right]$, γ^{b} is the coefficient on past inflation which captures the persistence of the inflation process, and φ^{x} is the coefficient on variable x. In other words, a dynamic simulation of the model is run by setting the initial value of each explanatory variable to zero and using the coefficient on lagged inflation to incorporate the effects of inflation persistence that are attributable to previous movements in the explanatory variables. To evaluate what factors contributed to average deviations of inflation from the target, the contribution of inflation expectations is re-expressed in terms of deviation from either an *explicit* target (the one announced under the inflation-targeting regime) or an *implicit* one (the moving average of 10-year-ahead inflation expectations).[22]

Figure 6 shows the contribution of each factor to deviations of core inflation from target over four subperiods, which loosely correspond to the precrisis boom (from the first quarter of 2004 to the second quarter of 2008), the global financial crisis (from the third quarter of 2008 to the end of 2009), the post-crisis recovery (the start of 2010 to the second quarter of 2014), and the oil price decline and its aftermath (from the third quarter of 2014 to the first quarter of

[21] The conclusions in Section 3.2 and 3.3 hold when we use the coefficients of Table 2.

[22] Such decomposition can be performed under the assumption that the coefficients on the lag of inflation and inflation expectations sum to one. Both for median and robust regressions—in which the coefficients are unconstrained—Wald tests cannot reject the hypothesis of the sum of the coefficients being equal to one.

Figure 6. Contributions to Deviation of Core Inflation from Target, by Subperiod (percentage points)



Source: Authors' calculations.
Note: The bars represent the average contribution of each factor averaged across countries.

2018).[23] The largest contributor to deviations of core inflation from target over the four subperiods is inflation expectations. That is, inflation expectations for the sampled emerging markets, on average, exceeded the inflation target.[24] Domestic cyclical conditions played a smaller role. Upswings during the boom period led inflation to move above the target, while downturns during the global financial crisis led to lower inflation compared with the target.

Among the external factors, the largest contributor is the variable capturing external price pressures, which was, on average, deflationary during the sample period. However, the magnitude of this effect (–0.05 percentage point annually, on average, over the sample period) was considerably smaller than that of long-term inflation

[23]We report the results for core inflation in the rest of the analysis to abstract from the volatility induced by energy and food prices and focus on the underlying inflationary pressures. However, the results for headline inflation are qualitatively similar to the ones for core inflation.

[24]This could reflect the public's doubts about the central bank's commitment to the inflation target or concerns about fiscal sustainability that may imply higher inflation in the future.

expectations (0.5 percentage point). The deflationary pressure from external prices was most pronounced during the boom that preceded the global financial crisis.[25] The contribution of foreign slack is economically insignificant.

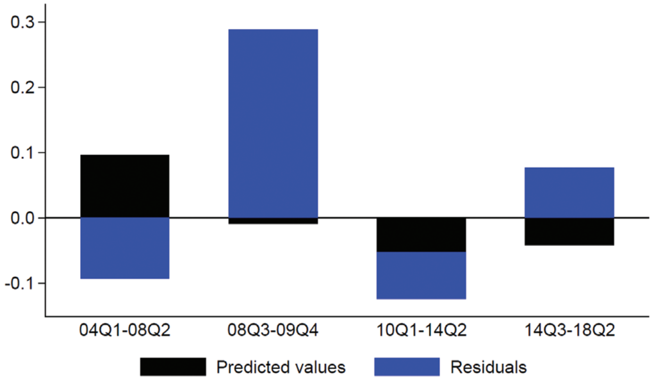
Figure 6 also shows that the overall deviation of inflation from the target declined gradually during 2004–14, by 0.7 percentage point. This trend is partly explained by output gaps (domestic and foreign), which stimulated inflation during the boom of 2004–07, and depressed it during the bust of 2008–09, and partly by the remaining residual.

Could the decrease in the average decomposition residual during 2004–14 of Figure 6 reflect a neglected common source of downward pressure on inflation? To address this question, the analysis estimates a common driver of inflation across emerging markets that cannot be explained by domestic factors. The approach is implemented in two steps. First, we include time fixed effects in a model specification as Equation (1) but without the external variables in vector Z^*. Second, we regress the common component—that is, the time fixed effects—on the cross-country averages of the domestic determinants of core inflation, and obtain the predicted values and the residuals, which can be thought as the "true" residual of the first regression.

As shown in Figure 7, the common component (the sum of the predicted values and the residuals) captures the commodity-induced inflation surge during 2008, but for other sample subperiods its contribution to inflation deviations from target is small in economic terms. Furthermore, the estimated time fixed effects correlate with domestic explanatory variables, suggesting that the risk of neglecting other external forces is reduced. Beyond these factors, the residual provides a negligible average contribution to inflation during the post-global financial crisis period. These findings corroborate

[25] Breaking up the contribution of the external price pressure variable into its subcomponents reveals that the contribution of the import-weighted nominal effective exchange rate—which in principle could also reflect domestic developments—is small, hovering around zero with the exception of the global financial crisis subperiod, when it reached 0.15 percentage point. The other two subcomponents, the import-weighted foreign PPI inflation and the percent change in the GDP deflator, present larger contributions ranging between 0 and 0.17 percentage point and –0.12 and –0.25 percentage point, respectively.

Figure 7. Common Driver of Core Inflation (percentage points)

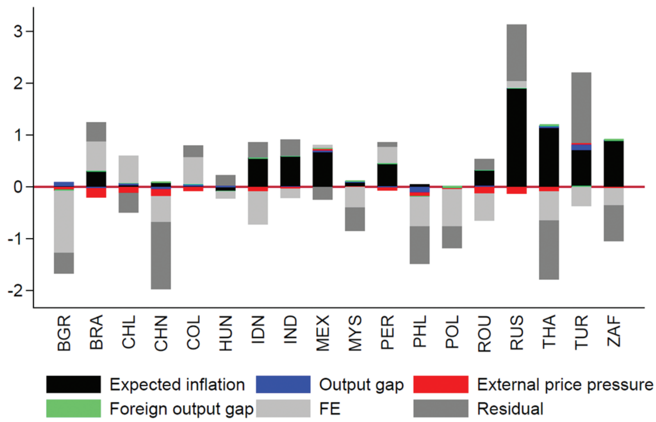


Source: Authors' calculations.
Note: Time fixed effects are based on a panel regression that excludes external variables. Residuals are from a regression of these time fixed effects on country averages of the domestic determinants of core inflation. Predicted values are displayed in terms of deviation from the mean over the sample period.

the earlier results on the comparatively limited average impact of global factors in driving inflation in emerging markets. Overall, the results of this section point to the centrality of fluctuations in long-term inflation expectations in driving inflation in emerging countries, which are interpreted to be of domestic origin.

Examining the contributions at the country level reveals that although changes in long-term inflation expectations are the main overall contributor to the deviations of actual inflation from target, there is noticeable cross-country heterogeneity. As shown in Figure 8, countries such as Chile and Poland, for example, show small contributions of inflation expectations from the target, consistent with the maturity of their monetary frameworks. On the other hand, in Russia and Thailand deviations of inflation expectations from target were large. Overall, the average inflationary impact of expectations is sizable for only half of the economies in the sample. In contrast, external price developments exerted downward pressure on domestic prices for three-fourths of the economies in the sample, even though the magnitude of this contribution is small. The impact of cyclical factors is by construction limited, when averaged over 2004–18.

Figure 8. Contributions to Deviations of Core Inflation from Target, by Country (percentage points)



Source: Authors' calculations.
Note: The bars represent the average contribution of each factor averaged across periods.

3.3 Contributions to Inflation Variation

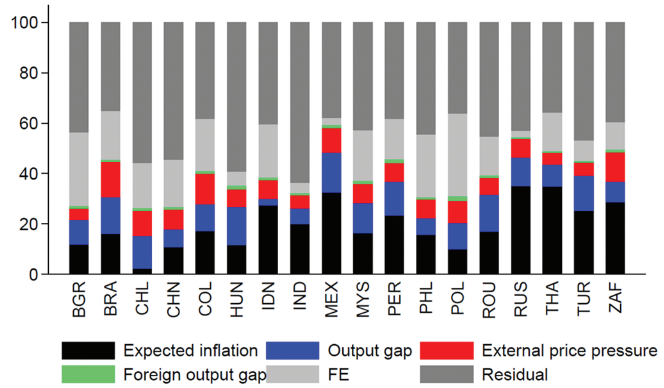
To assess what factors contributed to the variation of inflation deviations from target, we perform an alternative decomposition. In the spirit of a variance decomposition exercise, we calculate the contribution for each x variable of the N vector as

$$C_i^{var,x} = \frac{\frac{1}{T}\sum_t^T |C_{i,t}^x|}{\sum_N \frac{1}{T}\sum_t^T |C_{i,t}^n|}, \tag{7}$$

where the contribution of inflation expectations, C^{var,π^e}, is expressed in terms of deviations from the target. In words, the expression in Equation (7) calculates the ratio of the average absolute value of the contribution of each variable to the sum of the same average absolute value of the contributions of all variables.

Figure 9 presents the normalized contributions. The results confirm the importance of fluctuations in long-term inflation expectations around the inflation target. Inflation expectations are the

Figure 9. Normalized Contributions to Deviations of Core Inflation from Target, by Country (percent of total contributions)



Source: Authors' calculations.
Note: The bars represent the average of the absolute values of the country-specific contributions over the period 2004:Q1–2018:Q1, as a percent of the overall deviation of core inflation from the target.

largest contributing explanatory factor for four-fifths of the sample countries, explaining, on average, 20 percent of the variation in inflation. Similar to the evidence in Figure 8, there is substantial heterogeneity across countries, with the share attributable to inflation expectations ranging from 2 percent to 35 percent. One should note that a low average contribution for a given factor over the entire sample does not mean it does not play an important role in driving inflation dynamics over the short term. For instance, Figure 9 shows that the share of inflation variation explained by inflation expectations was sizable in Colombia despite the very small average contribution reported in Figure 8, indicating that the contribution of fluctuations of inflation expectations around the target were relatively large but tended to cancel out along the sample. With respect to the other variables, the results confirm that external price movements played a more limited role for variability in inflation rates, on average explaining 8 percent of inflation deviations, and that the contribution of the foreign output gap is negligible in all decomposition results.

To establish the relative importance of domestic and foreign factors in determining inflation dynamics, we group the contributions into two subsets S^n with $n = [1, 2]$:

$$C_i^{var,S^n} = \frac{\sum_{x \in S^n} \frac{1}{T} \sum_t^T |C_{i,t}^x|}{\sum_{x \in [S^1 \wedge S^2]} \frac{1}{T} \sum_t^T |C_{i,t}^x|}, \tag{8}$$

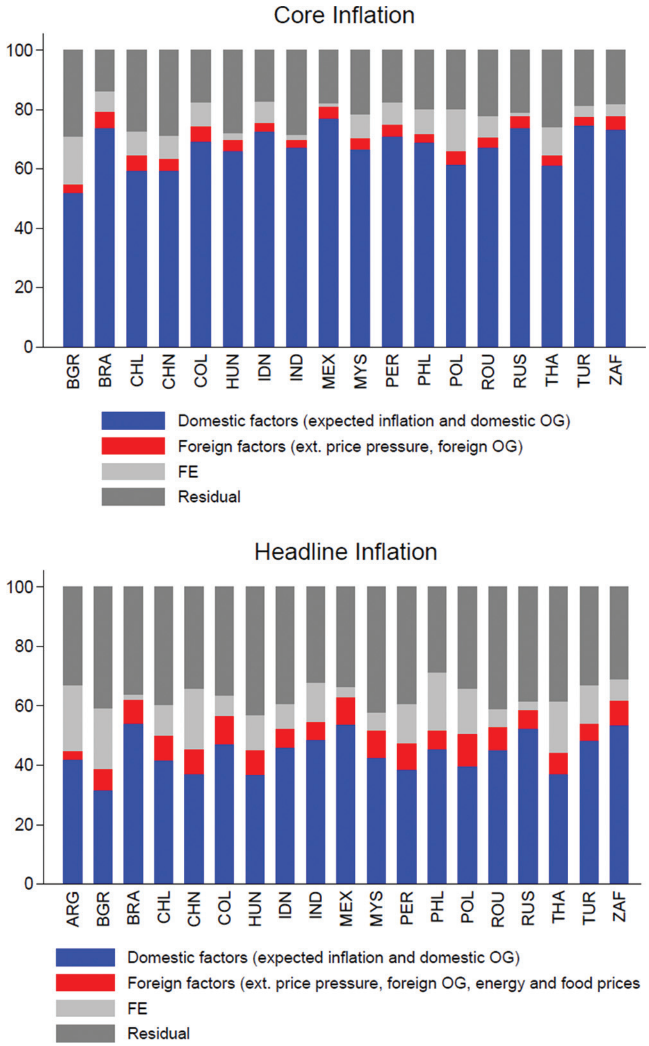
where S^1 denotes a first subset consisting of domestic factors (inflation expectations and the output gap) and S^2 a second subset consisting of foreign factors (foreign output gap, external price pressure, and commodity price inflation). The contribution of inflation expectations here is *not* expressed in terms of deviations from the target.

Applying this definition of global factors, the results shown in Figure 10 confirm that domestic contributions to inflation variation are much larger than foreign contributions, for both core inflation and headline inflation. Domestic contributions explain between 52 percent and 77 percent of core inflation dynamics and between 32 percent and 55 percent of headline inflation dynamics. The proportion of inflation dynamics explained by foreign factors is much smaller, ranging between 3 percent and 5 percent for core inflation and 3 percent and 11 percent for headline inflation.

3.4 Robustness Exercises

The analysis in this paper is subject to some limitations. First, some variables categorized as domestic (foreign) could in reality contain foreign (domestic) elements; also, the results are subject to sizable uncertainty since 45 percent of the variation in inflation remains unexplained. Second, as in many other empirical exercises involving a Phillips-curve estimation, the estimates can be affected by endogeneity arising from omitted variables. Third, three-years-ahead inflation expectations might not be representative of long-term inflation expectations. In this section, we present the results of a series of robustness tests that provide some evidence to limit the concerns about these issues.

Figure 10. Normalized Contributions of Domestic and Global Factors to Inflation Dynamics, by Country (percent of total contributions)



Source: Authors' calculations.
Note: The bars represent the average of the absolute values of the country-specific contributions (accounting for persistence of inflation) over the period 2004:Q1–2018:Q1, as a percent of the sum of all contributions.

3.4.1 Global Factors

The baseline specification in Equation (1) includes a vector of external variables, so that the coefficient on inflation expectations already abstracts from any change in external factors. Still, one concern is that the evolution of inflation expectations may be capturing global developments that are common across countries. If one were to make the extreme assumption that all the residual is due to uncaptured foreign factors, the average contribution of foreign factors to inflation variation would be 26 percent for core inflation and 44 percent for headline inflation, still less than or comparable to the average contribution of domestic factors (68 percent for core inflation and 44 percent for headline inflation).

In the alternative specifications of columns 1 and 2 of Table 3, the vector of external variables is replaced with time fixed effects as catch-all variables for foreign factors. In this case, the average contribution of foreign factors to inflation would be 11 percent for both core and headline inflation. Time fixed effects, however, do not capture idiosyncratic movements in external price pressures, given that such pressures can vary by country. Therefore, in columns 3 and 4 of Table 3, we add back the external price pressure variable to the specification that includes time fixed effects.[26] The results confirm that external price pressures remain significant despite the inclusion of time fixed effects, and that the average contribution of foreign factors to inflation variation would be 17 percent for core inflation and 14 percent for headline inflation.

Finally, drawing on Choi et al. (2018), in the regression for headline inflation, we interact energy and food price inflation with the weight of these items in CPI baskets. The results in column 5 of Table 3 show that the coefficient for food price inflation remains significant and becomes larger in magnitude, consistent with the large weight of food in the CPI baskets of the 19 sample countries, which averages 32.9 percent. The coefficient for energy inflation, however, is still insignificant, in line with its smaller weight in the CPI basket, which averages 9.6 percent. The results for other variables are virtually unchanged.

[26]The foreign output gap is not included in these specifications because it turns out to be insignificant in the baseline specifications.

Table 3. Hybrid Phillips-Curve Estimation, Alternative Specifications

	Core Inflation (1)	Headline Inflation (2)	Core Inflation (3)	Headline Inflation (4)	Headline Inflation (5)
	With Time Fixed Effects	With Time Fixed Effects	With Time Fixed Effects	With Time Fixed Effects	With Weighted Commodity Inflation
Inflation Expectations Three Years Ahead	0.832***	0.327***	0.862***	0.353***	0.354***
	(0.111)	(0.082)	(0.104)	(0.080)	(0.102)
Lag of Core/Headline Inflation	0.444***	0.488***	0.435***	0.490***	0.417***
	(0.039)	(0.036)	(0.040)	(0.033)	(0.045)
Output Gap	0.172***	0.230***	0.138***	0.225***	0.167**
	(0.049)	(0.059)	(0.041)	(0.065)	(0.081)
Lag of External Price Pressure			0.016***	0.018***	0.006
			(0.003)	(0.005)	(0.008)
Foreign Output Gap					0.158**
					(0.076)
Lag of Food Price Inflation					0.045***
					(0.013)
Lag of Energy Price Inflation					0.016
					(0.018)
Countries	18	19	18	19	19
Observations	634	669	634	669	668
R-squared	0.561	0.494	0.568	0.498	0.445

Source: Authors' calculations.
Note: The table presents median regression results, for which the pseudo R-squared is reported. All specifications include country fixed effects. Robust standard errors are in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

3.4.2 Extensions

As discussed in Section 2.2, the past few decades witnessed a trade integration process that led many emerging markets to participate more in GVCs. Deeper integration should be reflected in stronger competition from abroad, possibly affecting inflation dynamics. To capture the role of stronger trade integration that is not yet reflected in the external price pressure variable, the baseline specification is extended to include trade openness and participation in GVCs, as well as their interactions with external variables:

$$\pi_{i,t} = \gamma^b \pi_{i,t-1} + \gamma^f \pi^e_{i,t} + \beta Y^{gap}_{i,t} + \theta Z^*_{i,t} + \varphi T_{i,t} Z^*_{i,t} + \psi T_{i,t} + \eta_i + \epsilon_{i,t} \tag{9}$$

in which $T_{i,t}$ is a measure of trade openness or participation in GVCs. The results in Table 4 suggest there is no significant evidence that deeper trade integration has a significant effect on domestic inflation. As shown in columns 1, 3, and 4, if anything, the coefficients on trade openness and GVC participation are positive when they are significant, but they are relatively small, and the results are not consistent across inflation measures. The interaction term between trade openness and foreign output gap in the specification for headline inflation is significant in column 2, suggesting that movements in foreign cyclical conditions have an impact on inflation when the economy is more open, although the magnitude of the effect is small.

Since China joined the World Trade Organization in 2001, China quickly increased its share in global trade owing to relatively lower export prices and became an important trading partner for many emerging markets in the sample, possibly affecting their inflation dynamics. The analysis explores the role of price pressure from China by decomposing the external price pressure variable into its Chinese component and the non-Chinese component. The results in columns 5 and 6 indicate that external price pressure from China does not have any significant impact on core or headline inflation dynamics, while non-Chinese external price pressures remain a significant determinant in the specification for core inflation, consistent with the results of the baseline specification.

Table 4. Hybrid Phillips-Curve Estimation, Extensions

	Core Inflation (1)	Headline Inflation (2)	Core Inflation (3)	Headline Inflation (4)	Core Inflation (5)	Headline Inflation (6)
	Interaction: Trade Openness	Interaction: Trade Openness	Interaction: GVC Participation	Interaction: GVC Participation	Interaction: China's External Price Pressure	Interaction: China's External Price Pressure
Inflation Expectations Three Years Ahead	0.643***	0.406***	0.632***	0.378***	0.551***	0.399***
	(0.100)	(0.107)	(0.096)	(0.121)	(0.096)	(0.104)
Lag of Core/Headline Inflation	0.479***	0.422***	0.479***	0.427***	0.502***	0.426***
	(0.031)	(0.047)	(0.032)	(0.049)	(0.030)	(0.046)
Output Gap	0.154***	0.223***	0.173***	0.194**	0.163***	0.206***
	(0.044)	(0.073)	(0.040)	(0.085)	(0.037)	(0.079)
Lag of External Price Pressure	0.009	0.011	–0.001	0.029		
	(0.008)	(0.016)	(0.014)	(0.036)		
Foreign Output Gap	–0.047	–0.195	0.038	–0.141	0.019	0.082
	(0.106)	(0.139)	(0.160)	(0.290)	(0.040)	(0.095)
Lag of Food Price Inflation		0.014		0.020		0.012***
		(0.009)		(0.017)		(0.004)
Lag of Energy Price Inflation		–0.002		–0.008		0.000
		(0.004)		(0.008)		(0.002)
Trade Openness	0.015*	0.026				
	(0.008)	(0.020)				
Trade Openness * Lag of External Price Pressure	0.000	–0.000				
	(0.000)	(0.001)				
Trade Openness * Foreign Output Gap	0.002	0.007**				
	(0.003)	(0.003)				

(*continued*)

Table 4. (Continued)

	Core Inflation (1)	Headline Inflation (2)	Core Inflation (3)	Headline Inflation (4)	Core Inflation (5)	Headline Inflation (6)
	Interaction: Trade Openness	**Interaction: Trade Openness**	**Interaction: GVC Participation**	**Interaction: GVC Participation**	**Interaction: China's External Price Pressure**	**Interaction: China's External Price Pressure**
Trade Openness * Lag of Food Price Inflation		–0.000				
		(0.000)				
Trade Openness * Lag of Energy Price Inflation		0.000				
		(0.000)				
GVC Participation			0.060**	–0.033		
			(0.030)	(0.065)		
GVC Participation * Lag of External Price Pressure			0.000	–0.000		
			(0.000)	(0.001)		
GVC Participation * Foreign Output Gap			–0.001	0.004		
			(0.003)	(0.006)		
GVC Participation * Lag of Food Price Inflation				–0.000		
				(0.000)		
GVC Participation * Lag of Energy Price Inflation				0.000		
				(0.000)		
External Price Pressure Excl. China					0.018***	0.007
					(0.003)	(0.007)
External Price Pressure from China					–0.004	–0.002
					(0.004)	(0.009)
Countries	18	19	18	19	18	19
Observations	624	659	633	668	627	662
R-squared	0.524	0.453	0.526	0.446	0.523	0.446

Source: Authors' calculations.
Note: The table presents median regression results, for which the pseudo R-squared is reported. All specifications include country fixed effects. Robust standard errors are in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

3.4.3 Inflation Expectation Horizons

Inflation expectations in the baseline specification correspond to three-year-ahead inflation forecasts, a sufficiently long horizon to capture beliefs about inflation in the long term rather than the effect of transitory shocks and the response of monetary policy. However, to ensure that the results are not dependent on the selection of this specific horizon, we perform a series of robustness tests using inflation expectations of up to seven years ahead. The results in Table 5 for core inflation are robust to the change of the horizon for inflation expectations, with the magnitude of the coefficient decreasing only marginally as the horizon gets larger (the coefficient on expected inflation for horizons three to seven years ahead ranges from 0.56 to 0.64).[27] In the case of headline inflation, inflation expectations become insignificant for horizons of six years ahead and beyond, reflecting the higher volatility of headline inflation compared with core inflation.

4. Conclusions

Following a period of disinflation during the 1990s and early 2000s, inflation in emerging markets has remained remarkably low and stable despite large swings in commodity prices, the global financial crisis, and periods of strong and sustained U.S. dollar appreciation. A key question is whether this improved inflation performance is sustainable, or if instead it reflects a temporary constellation of global factors that put downward pressure on inflation. The literature on the role of global factors in driving domestic inflation focuses on advanced economies and presents mixed results.

This paper studies the role of domestic and global factors in driving inflation dynamics in emerging markets. We estimate a New Keynesian Phillips-curve model for core and headline inflation using data for 19 large emerging markets over 2004–18. Following recent contributions in the literature (Borio and Filardo 2007; Ihrig et al.

[27] One potential concern with the Phillips-curve specification is reverse causality from current inflation to inflation expectations, especially at shorter horizons. The decrease in estimated coefficients as the horizon lengthens is consistent with this concern. But the small magnitude of the differences suggests the effect is limited in economic terms.

Table 5. Hybrid Phillips-Curve Estimation, Varying Inflation Expectation Horizon

	Core Inflation				Headline Inflation			
	Four-Year Ahead Infl. Exp. (1)	**Five-Year Ahead Infl. Exp. (2)**	**Six-Year Ahead Infl. Exp. (3)**	**Seven-Year Ahead Infl. Exp. (4)**	**Four-Year Ahead Infl. Exp. (5)**	**Five-Year Ahead Infl. Exp. (6)**	**Six-Year Ahead Infl. Exp. (7)**	**Seven-Year Ahead Infl. Exp. (8)**
Inflation Expectations	0.637***	0.614***	0.585***	0.560***	0.397**	0.448*	0.256	–0.066
n Years Ahead	(0.125)	(0.130)	(0.131)	(0.155)	(0.158)	(0.245)	(0.262)	(0.247)
Lag of Core/Headline	0.502***	0.524***	0.548***	0.549***	0.459***	0.461***	0.502***	0.537***
Inflation	(0.037)	(0.037)	(0.037)	(0.036)	(0.047)	(0.044)	(0.048)	(0.040)
Output Gap	0.138***	0.136***	0.144***	0.168***	0.164*	0.152**	0.179**	0.172**
	(0.036)	(0.040)	(0.038)	(0.041)	(0.088)	(0.077)	(0.081)	(0.081)
Lag of External Price	0.021***	0.018***	0.020***	0.020***	0.008	0.004	0.008	0.003
Pressure	(0.003)	(0.003)	(0.004)	(0.003)	(0.009)	(0.008)	(0.009)	(0.008)
Foreign Output Gap	0.050	0.042	0.057	0.002	0.080	0.119	0.034	0.051
	(0.047)	(0.051)	(0.048)	(0.050)	(0.098)	(0.077)	(0.088)	(0.101)
Lag of Food Price					0.013***	0.013***	0.013***	0.013***
Inflation					(0.004)	(0.004)	(0.005)	(0.005)
Lag of Energy Price					0.000	0.000	–0.000	–0.000
Inflation					(0.002)	(0.002)	(0.002)	(0.002)
Countries	18	18	18	18	19	19	19	19
Observations	577	603	576	576	612	638	611	610
R-squared	0.514	0.519	0.513	0.511	0.446	0.439	0.442	0.443

Source: Authors' calculations.
Note: The table presents median regression results, for which the pseudo R-squared is reported. All specifications include country fixed effects. Robust standard errors are in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

2010; and Auer, Borio, and Filardo 2017), we augment the model with variables capturing foreign macro developments, including the import-weighted output gap and producer price inflation of trading partners.

We find that domestic factors accounted for the lion's share of inflation dynamics in emerging markets, in line with the findings of Ha, Kose, and Ohnsorge (2019). Fluctuations in long-term inflation expectations, linked to domestic developments, were the main driver of average deviations of inflation from target and inflation variability. The contribution of global variables is not always statistically significant and, in any case, substantially smaller than the one from domestic factors in economic terms. To address potential endogeneity concerns, we implement a battery of robustness tests that confirm the marginal impact of global factors compared with that of domestic factors, and that inflation expectations reflect the evolution of domestic variables rather than global developments.

Our findings have important implications for monetary policy in emerging markets. The results show that the gains in inflation performance since the mid-2000s are largely attributable to domestic factors, which could capture improved policy frameworks and gains in credibility. One implication, suggested by these findings, is that although emerging markets are increasingly integrated with the global economy, domestic policies, through their impact on inflation expectations, continue to hold significant leverage over domestic inflation outcomes.

References

Aqib, A., N. Novta, and F. Rodrigues-Bastos. 2017. "Calculating Trade in Value Added." IMF Working Paper No. 17/178.

Auer, R. A., C. Borio, and A. Filardo. 2017. "The Globalization of Inflation: the Growing Importance of Global Value Chains." BIS Working Paper No. 602.

Auer, R. A., T. Chaney, and P. Sauré. 2018. "Quality Pricing-to-Market." *Journal of International Economics* 110 (January): 87–102.

Auer, R. A., K. Degen, and A. M. Fischer. 2013. "Low-Wage Import Competition, Inflationary Pressure, and Industry Dynamics in Europe." *European Economic Review* 59 (April): 141–66.

Auer, R. A., A. A. Levchenko, and P. Sauré. 2019. "International Inflation Spillovers through Input Linkages." *Review of Economics and Statistics* 101 (3): 507–21.

Autor, D., D. Dorn, L. F. Katz, C. Patterson, and J. Van Reenen. 2020. "The Fall of the Labor Share and the Rise of Superstar Firms." *Quarterly Journal of Economics* 135 (2): 645–709.

Ball, L. M. 2006. "Has Globalization Changed Inflation?" NBER Working Paper No. 12687.

Bems, R., F. G. Caselli, F. Grigoli, B. Gruss, and W. Lian. 2018. "Expectations' Anchoring and Inflation Persistence." IMF Working Paper No. 18/120.

Bonadio, B., A. M. Fischer, and P. Sauré. 2020. "The Speed of Exchange Rate Pass-Through." *Journal of the European Economic Association* 18 (1): 506–38.

Borio, C., and A. Filardo. 2007. "Globalisation and Inflation: New Cross-Country Evidence on the Global Determinants of Domestic Inflation." BIS Working Paper No. 227.

Carney, M. 2017. "[De] Globalization and Inflation." Speech at the 2017 IMF Michel Camdessus Central Banking Lecture, International Monetary Fund, Washington, DC, September 18.

Carriere-Swallow, Y., B. Gruss, N. E. Magud, and F. Valencia. 2016. "Monetary Policy Credibility and Exchange Rate Pass-Through." IMF Working Paper No. 16/240.

Cavallo, A. 2013. "Online and Official Price Indexes: Measuring Argentina's Inflation." *Journal of Monetary Economics* 60 (2): 152–65.

Celasun, O., R. G. Gelos, and A. Prati. 2004. "Obstacles to Disinflation: What is the Role of Fiscal Expectations?" *Economic Policy* 19 (40): 442–81.

Choi, S., D. Furceri, P. Loungani, S. Mishra, and M. Poplawski-Ribeiro. 2018. "Oil Prices and Inflation Dynamics: Evidence from Advanced and Developing Economies." *Journal of International Money and Finance* 82 (April): 71–96.

Combes, J.-L., X. Debrun, A. Minea, and R. Tapsoba. 2017. "Inflation Targeting, Fiscal Rules and the Policy Mix: Cross-effects and Interactions." *Economic Journal* 128 (615): 2755–84.

Danninger, S. 2016. "What's Up with US Wage Growth and Job Mobility?" IMF Working Paper No. 16/122.

Draghi, M. 2017. "Accompanying the Economic Recovery." Speech at the ECB Forum on Central Banking, Sintra, June 27.

Fisher, R. 2006. "Coping with Globalization's Impact on Monetary Policy." Speech at 2006 Allied Social Science Associations Meeting, Boston, MA, January 6.

Galí, J., and M. Gertler. 1999. "Inflation Dynamics: A Structural Econometric Analysis." *Journal of Monetary Economics* 44 (2): 195–222.

Galí, J., M. Gertler, and J. D. Lopez-Salido. 2001. "European Inflation Dynamics." *European Economic Review* 45 (7): 1237–70.

———. 2003. "Erratum to 'European Inflation Dynamics'." *European Economic Review* 47 (4): 759–60.

Gopinath, G. 2015. "The International Price System." NBER Working Paper No. 21646.

Gopinath, G., E. Boz, C. Casas, F. J. Díez, P.-O. Gourinchas, and M. Plagborg-Møller. 2020. "Dominant Currency Paradigm." *American Economic Review* 110 (3): 677–719.

Ha, J., M. A. Kose, and F. Ohnsorge, eds. 2019. *Inflation in Emerging and Developing Economies*. World Bank Publications 30657. The World Bank.

Ihrig, J., S. B. Kamin, D. Lindner, and J. Marquez. 2010. "Some Simple Tests of the Globalization and Inflation Hypothesis." *International Finance* 13 (3): 343–75.

Ilzetzki, E., C. M. Reinhart, and K. S. Rogoff. 2017. "Exchange Arrangements Entering the 21st Century: Which Anchor Will Hold?" NBER Working Paper No. 23134.

International Monetary Fund. 2005a. "Does Inflation Targeting Work in Emerging Markets?" In *World Economic Outlook: Building Institutions*, (September), chapter 4. International Monetary Fund.

———. 2005b. "Globalization and External Imbalances." In *World Economic Outlook: Globalization and External Imbalances*, (April), chapter 3. International Monetary Fund.

———. 2013. "The Dog That Didn't Bark: Has Inflation Been Muzzled or Was It Just Sleeping?" In *World Economic Outlook: Hopes, Realities, Risks*, (April), chapter 3. International Monetary Fund.

———. 2016. "Global Disinflation in an Era of Constrained Monetary Policy." In *World Economic Outlook: Subdued Demand:*

Symptoms and Remedies, (October), chapter 3. International Monetary Fund.

———. 2018. "Challenges for Monetary Policy in Emerging Economies as Global Financial Conditions Normalize." In *World Economic Outlook: Challenges to Steady Growth*, (October), chapter 3. International Monetary Fund.

Jaŝová, M., R. Moessner, and E. Takáts. 2018. "Domestic and Global Output Gaps as Inflation Drivers: What Does the Phillips Curve Tell?" BIS Working Paper No. 748.

Kohn, D. 2006. "The Effects of Globalization and Their Implications for Monetary Policy." Speech at Federal Reserve Bank of Boston's 51st Economic Conference, Chatham, MA, June 16.

Mankiw, N. G., R. Reis, and J. Wolfers. 2004. "Disagreement about Inflation Expectations." In *NBER Macroeconomics Annual 2003,* Vol. 18, ed. M. Gertler and K. Rogoff, 209–48 (chapter 4). MIT Press.

Mishkin, F. S. 2000. "Inflation Targeting in Emerging-Market Countries." American Economic Review 90 (2): 105–9.

———. 2007. "Inflation Dynamics." *International Finance* 10 (3): 317–34.

Mishkin, F. S., and M. A. Savastano. 2001. "Monetary Policy Strategies for Latin America." *Journal of Development Economics* 66 (2): 415–44.

Montes, G. C., and T. Acar. 2018. "Fiscal Credibility and Disagreement in Expectations about Inflation: Evidence for Brazil." *Economics Bullet*in 38 (2): 826–43.

Rogoff, K. 2003. "Globalization and Global Disinflation." *Economic Review* (Federal Reserve Bank of Kansas City) 88 (4): 45–80.

Stock, J. H., and M. W. Watson. 2007. "Why Has US Inflation Become Harder to Forecast?" *Journal of Money, Credit and Banking* 39 (s1): 3–33.

———. 2010. "Modeling Inflation After the Crisis." NBER Working Paper No. 16488.

Vegh, C. A., and G. Vuletin. 2014. "Overcoming the Fear of Free Falling: Monetary Policy Graduation in Emerging Markets." In *The Role of Central Banks in Financial Stability: How Has It Changed?* ed. D. D. Evanoff, C. Holthausen, G. G. Kaufman, and M. Kremer, 105–29. World Scientific.

Vogel, J. 2008. "Spatial Competition with Heterogeneous Firms." *Journal of Political Economy* 116 (3): 423–66.

Yellen, J. 2006. "Monetary Policy in a Global Environment." Speech at the Conference "The Euro and the Dollar in a Globalized Economy," University of California at Santa Cruz, Santa Cruz, CA, May 27.

———. 2015. "Inflation Dynamics and Monetary Policy." Speech at the Philip Gamble Memorial Lecture, University of Massachusetts, Amherst, Amherst, Massachusetts, September 24.

Excess Persistence in Employment of Disadvantaged Workers*

Bruce Fallick and Pawel Krolikowski
Federal Reserve Bank of Cleveland

We examine persistence in employment-to-population ratios among less-educated individuals in excess of that implied by persistence in aggregate labor market conditions, using state-level data for the United States. Dynamic panel regressions indicate only a moderate degree of excess persistence, which dissipates within three years. We find no significant asymmetry between the excess persistence of high versus low employment rates. The cumulative effect of excess persistence in the business cycle surrounding the 2001 recession was mildly positive, while the effect in the cycle surrounding the 2008–09 recession was decidedly negative. Simulations suggest that the lasting employment benefits of temporarily running a "high-pressure" economy are small.

JEL Codes: E24, J21, J24.

1. Introduction

The relationship between current employment experience and future employment outcomes, especially for disadvantaged workers, has long interested both researchers and policymakers. Notably, during the expansion of the 2010s policymakers asked whether temporarily running a "high-pressure economy," with robust aggregate demand

---

*We thank seminar participants at the Federal Reserve Bank of Cleveland, the Midwest Macroeconomic meetings, the Southern Economic Association meetings, and the American Economic Association meetings for helpful comments. Anne Chen, Emily Dohrman, Anna Weixel, and Meifeng Yang provided excellent research assistance. The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Federal Reserve Bank of Cleveland or of any other person associated with the Federal Reserve System. Authors' contact: Federal Reserve Bank of Cleveland, 1455 E 6th St, Cleveland, OH 44114. E-mails: bruce.fallick@clev.frb.org and pawel.krolikowski@clev.frb.org.

and a tight labor market, might produce long-run benefits to workers with weak workforce attachment even after the economy as a whole returns to a more "normal" state (Stockhammer and Sturn 2012; Ball 2015; Reifschneider, Wascher, and Wilcox 2015; Yellen, 2016, 2019).[1]

We address this question by estimating excess persistence in employment among less-educated individuals using state-level data for the United States. We define excess persistence in employment as a lasting effect of past employment conditional on macroeconomic conditions, as in Okun (1973).

We find a moderate but ephemeral degree of excess persistence: For the group with the greatest excess persistence among those we examine—prime-age men with no more than a high school education—the effects of past employment rates on subsequent employment rates can be substantial early on but essentially dissipate within three years. Furthermore, we find little evidence for asymmetric effects of high or low past employment on present employment. Our estimates imply that the cumulative effect of excess persistence in the business cycle surrounding the 2001 recession was mildly positive, while the effect in the cycle surrounding the 2008–09 recession was decidedly negative. Our simulations suggest that, despite large contemporaneous benefits, the lasting employment benefits of temporarily running a "high-pressure" economy are small.

Microeconomic evidence seems to support the notion of excess persistence in employment. This evidence includes findings that macroeconomic conditions at the time a person completes his or her education and starts a career have lasting effects on relative individual earnings, that the state of the labor market earlier in one's tenure at an employer influences one's subsequent wage rate at that employer, and that a person's early employment experience may affect her later employment (see von Wachter 2020 for a recent discussion).

[1] A related literature addresses persistence in aggregate conditions themselves. This literature has found that in at least some countries, loose labor markets appear to have had adverse long-run effects (e.g., Blanchard and Summers 1986; Ball 2009).

However, such evidence on how conditions at an early point in one's labor market experience affect individual outcomes does not establish the existence of excess persistence in aggregate employment, for two reasons.

First, the effects on those who, say, initially enter the labor force during a tight labor market are measured relative to the effects on those who enter during a slack labor market. This form of comparative excess persistence at the individual level does not imply excess persistence at an aggregate level: More employment in my history may enhance my chances of being employed today at the expense of reducing the chances of a competing person (with less employment in his history) being employed today.[2]

Second, given the great heterogeneity across jobs and persons and the multiplicity of mechanisms through which employment experience may affect future employment probabilities, the dynamic effects of employment at the microeconomic level may depend on the source of the variation in employment. That is, the microeconomic evidence on the dynamic effects of more employment in general does not imply that greater employment achieved through tighter macroeconomic conditions, as opposed to other causes, will have lasting effects on overall employment rates.

In addition, the microeconomic literature has mostly concentrated on excess persistence in wage rates or earnings, which need not imply excess persistence in employment. Indeed, depending on the mechanism at work, persistence in wage rates may work against persistence in employment. For example, Schmieder and von Wachter (2010) find that lower unemployment rates during a worker's job spell, which are associated with higher wage premiums, significantly increase the probability of job loss.

That said, past labor market conditions may affect subsequent employment outcomes at the aggregate level even conditional on

[2]In the context of trade policy, Abraham and Kearney (2018, p. 8) write, "as Pierce and Schott (2016) acknowledge, their difference-in-differences identification strategy precludes an estimate of the effect of the policy change on overall U.S. employment. This is because the estimated effects are all about relative job losses and there is not an obvious way to translate their findings into an estimate of overall absolute job losses." Similarly, see Gautier et al. (2018) for an example of how microeconomic welfare evaluation of job search assistance may differ from aggregate evaluation.

subsequent macroeconomic conditions because they affect employment experience. In particular, experience provides human and market capital that enhance future employability, such as “soft” skills (Almlund et al. 2011) and job contacts that facilitate employment after job loss (Cingano and Rosolia 2012; Glitz 2017) or improve match quality (Dustmann et al. 2017).

Unfortunately, previous research directly addressing the question of excess persistence in aggregate employment in the United States is thin. We follow the general approach of Fleischman and Gallin (2001) and Fleischman, Gallin, and Smith (2018), who estimate a dynamic model to extract the persistence of the employment-to-population ratio (e/p) in excess of that implied by the persistence of the macroeconomic conditions themselves, as measured by overall labor market tightness. Their evidence is consistent with our results. They also do not find a large degree of persistence in cohort-level e/p in response to fluctuations in macroeconomic conditions. They use variation among synthetic birth cohorts over time to identify possible excess employment persistence in the national data, as opposed to variation among states over time that we exploit.

Hotchkiss and Moore (2018) use state-level variation to compare individual outcomes in recessions following expansions of varying intensities. They find that, for some demographic groups, a person is likely to experience better outcomes during a period of high unemployment if that period was preceded by a tighter labor market. Yagan (2019) and Hershbein and Stuart (2020) find a large amount of persistence in local e/p ratios following recessions. However, Hershbein and Stuart (2020) find that this relationship is likely driven by persistent declines in overall labor demand, rather than the result of excess persistence in employment that is our interest here. Using employer survey data from the 1990s, Holzer, Raphael, and Stoll (2006) find that the relative demand for disadvantaged workers rose and racial discrimination likely declined during that expansion. Unfortunately, their data cover only the period 1992 to 2001 and so cannot separate the contemporaneous implications of cyclical conditions from their longer-term effects.

To measure the excess persistence in aggregate employment, we estimate a dynamic panel model in the detrended employment-to-population ratio (e/p) of disadvantaged workers, while controlling for aggregate labor market conditions using the unemployment rate (UR) gap among all workers.[3] We use variation among states over time for identification in these regressions.

The validity of the policy implications from our reduced-form model requires two assumptions. The first is that the variations in e/p under consideration be driven only by variations in overall labor market conditions as represented by the UR gap. This means that the phenomena (including economic policies) that drive the UR gap have no direct effect on the cyclical component of the e/p of the disadvantaged group, or that any direct effect is highly correlated with the UR gap (see Section 3.2.1). The second is that the degree of excess persistence identified by the state panel regressions is applicable to the aggregate level. We argue in Section 3.2.2 that this assumption is appropriate in our application.

The paper proceeds as follows. Section 2 describes our data. Section 3 describes the dynamic panel model. Section 4 presents our baseline estimates. Section 5 presents robustness exercises, including various detrending methods for the e/p of disadvantaged workers and instrumenting for the UR gap. Section 6 considers various definitions of the disadvantaged group. Section 7 investigates whether the degree of excess persistence in employment differs between high and low employment rates. Section 8 simulates the implications of our estimates of excess persistence for employment over the business cycle, and their implications for temporarily running a tight labor market. Section 9 concludes.

2. Data and Definitions

2.1 Baseline Sample

We focus our analysis on individuals with no more than a high school education, for four reasons. First, this education group has seen its relative earnings (Acemoglu and Autor 2011) and employment (Juhn 1992; Council of Economic Advisers 2017) decline markedly since the 1970s, which has made it a frequent focus of concern. Second, the mechanisms mentioned above for possible excess persistence at the

[3] We do not address the possibility of persistence generated by *long-term* unemployment, as in Song and von Wachter (2014) and Kallenberg and von Wachter (2017).

aggregate level would seem to be more important for this population, whose lower employment rates in general mean that they may benefit less from households and neighborhoods that provide human and market capital independent of an individual's own employment history (Conley and Topa 2002). Third, the employment of these populations tends to be more procyclical, so any change in overall labor market conditions can be expected to have a larger effect on their employment (Devereux 2002; Hoynes, Miller, and Schaller 2012; Aaronson et al. 2019), making any degree of excess persistence more important for this group. Fourth, Blacks and Hispanics are more likely to be less educated (Stoops 2004) and if these groups face discrimination in the labor market, higher levels of employment among the less educated mean greater direct exposure of employers to this group, which may reduce discrimination (Boisjoly et al. 2006; Miller 2017).

While these factors apply equally to women and to men, in this paper we focus on men because of practical difficulties in detrending the employment rates of women (see Section 2.3).

We further concentrate on prime-age men, ages 25 to 54, in order to abstract from most education and ordinary retirement decisions. Also for practical reasons we mostly examine all races together. However, we explore excess persistence among alternative education, race, and age groups in Section 6, and find qualitatively similar results as for our baseline group.

2.2 Data

Our analysis uses U.S. annual data for the e/p and URs at the state level. We include only the 50 states, omitting Washington DC and territories. We calculate the e/p for particular demographic groups for 1978–2018 from individual data in the basic monthly Current Population Survey (CPS).[4] We use published data on state URs. We measure labor market tightness by the UR gap, the difference between the overall UR in the state and an estimate of the state's

[4]CPS state-level data (National Bureau of Economic Research 2019) are also available for 1976 and 1977, but due to confidentiality restrictions some states are not identified in those years.

trend UR. For our baseline specification we use estimates of state-level trend URs from Fallick and Tasci (2020) (henceforth FT).

2.3 Detrending e/p

There are secular trends in the e/p of all groups of workers that we study. We isolate the cyclical component of e/p of each group in each state using the method recommended by Hamilton (2018). This method derives the trend of a variable as the predicted value from a regression of that variable at date $t+h$ on the d most recent values as of date t. We set the horizon parameter, h, at five years, and d at four, but our results are not sensitive to other reasonable choices. Except where noted, all of the results reported below use these detrended e/p.[5] In order to minimize end-of-sample bias, when estimating the trend we augment the e/p series on both ends with univariate forecasts (Kaiser and Maravall 1999; Stock and Watson 1999a).[6]

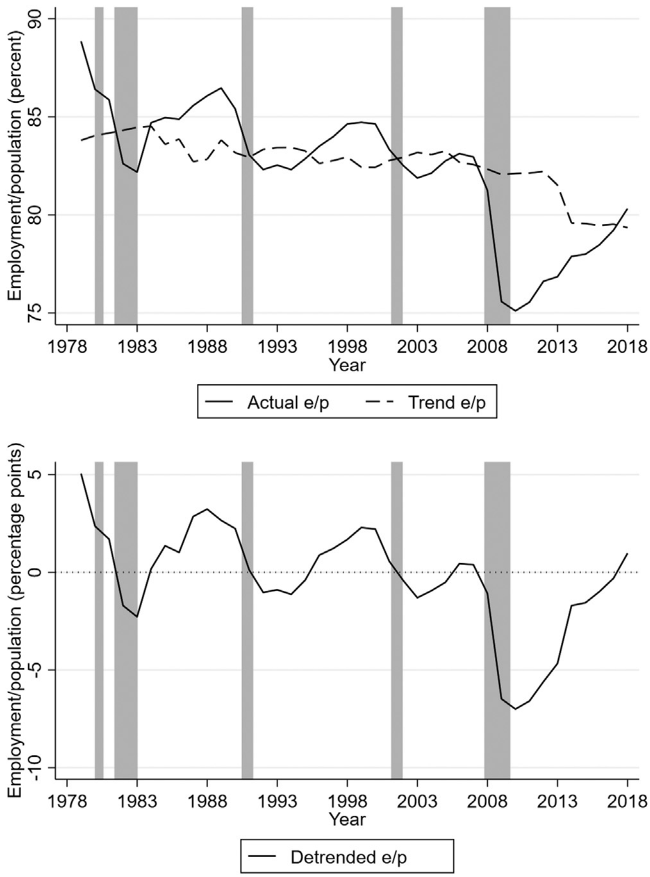
In addition to the advantages proposed by Hamilton (2018), this detrending method is backward looking. Alternative detrending methods that use subsequent data are not suitable for our purposes, as they may include the effects of excess persistence in the estimates of trend, thereby understating the amount of excess persistence in the data. We investigate alternative detrending methods (including no detrending) in Section 5.1.

The upper panel of Figure 1 shows the actual e/p and trend e/p for prime-age men with no more than a high school education

[5]Removing the trend in the e/p allows us to concentrate on persistence stemming from cyclical fluctuations. Notice that this detrended e/p will move fairly closely with (the negative of) the unemployment-population ratio in each state. This is because $e/p = L/p - u/p$, in which L denotes the size of the labor force and u denotes the number of unemployed workers, and removing the trend in the e/p primarily removes the secular movements in the labor force participation rate. However, movements in the participation rate caused by cyclical fluctuations ought to remain in the detrended e/p.

[6]We use second-order autoregressive models for this purpose, similar to Clark and Kozicki (2005) and Mise, Kim, and Newbold (2005), to extend the e/p series 10 years backward and forward from the beginning and end of our sample period, respectively.

Figure 1. Actual e/p and Trend e/p of Disadvantaged Group, Aggregated



Note: State-level actual and trend e/p for prime-age men with no more than a high school education aggregated to the national level. Trend e/p is calculated separately for each state using the method in Hamilton (2018). The dotted horizontal line in the lower panel denotes zero.

(the disadvantaged group in our baseline results), aggregated from the state to the national level for ease of display.[7] The lower panel shows the detrended e/p. Unfortunately, we were unable to estimate

[7]We aggregate by weighting within year by the number of observations in our CPS data for the baseline sample for each state in that year.

Table 1. Summary Statistics for e/p of Baseline Sample, State-Level Data

	Mean	Std. Dev.	Min.	Max.
Actual e/p$_{s,t}$ (%)	82.4	5.1	62.6	94.4
Detrended e/p$_{s,t}$ (pp)	–0.3	3.4	–14.0	9.7

Note: Summary statistics for prime-age men with no more than a high school education for the years 1978 to 2018. Mean and standard deviation are weighted by the population of the state. "Actual e/p$_{st}$" is the e/p of prime-age men with no more than a high school education in state s at time t. "Detrended e/p$_{st}$" is "Actual e/p$_{st}$" less the estimated trend for each state and is measured in percentage points (pp). Trend e/p is calculated using the method in Hamilton (2018). The detrended e/p will be the dependent variable in our baseline specification (Section 3).

reasonable trends for the like group of women, so we confine our attention to men.[8]

Table 1 provides summary statistics for the state-level (actual and detrended) e/p used in the regression analysis for our baseline group. Not surprisingly, there is more variation in the state-level data than is evident in the aggregate data in Figure 1.[9]

2.4 Disjoint Samples

Measuring employment status in the CPS is subject to measurement error from at least two sources. The first is sampling error, which makes any particular sample imperfectly representative of the population. The second is misreporting, due to misunderstanding, proxy responses, etc. (Poterba and Summers 1986; Elsby, Hobijn, and Şahin 2015). Sampling error, in particular, is positively correlated

[8]We suspect that the substantial changes in the trajectory of women's labor force participation in the 1970s and 1990s (e.g., Aaronson et al. 2014), and thus e/p, pose difficulties for univariate methods of estimating trends without longer time series than we have available for prime-age women with no more than a high school education.

[9]This trend differs from a suitably lagged moving average because the coefficients (effectively weights) on the various lags for each state are estimated, not imposed. The estimated coefficients are neither uniform across lags nor uniform across states, nor do they sum to 1 within any state (indeed, in every state the coefficients on the lags sum to less than 1). All in all, the estimated trends are smoother than moving averages.

over time due to the repeated sampling of individuals in the monthly CPS (Tiller 1992), which would bias up our estimates of excess employment persistence.

To avoid this bias, we use “disjoint” samples from one year to the next. That is, we calculate the e/p in state s in a given year for use on the left-hand side (LHS) of our regression equation (Equation (1) in Section 3 below) from a sample of individuals who are distinct from those used to calculate the e/p in previous years for use on the right-hand side (RHS) of this equation.[10] Since the disjoint samples still provide an unbiased estimate of the population e/p in each year, we obtain consistent estimates of excess employment persistence.

Such disjoint samples could be constructed in a number of ways. For simplicity and to balance the sample sizes used for the LHS and RHS measures, we choose to calculate the LHS e/p from a sample that includes only observations in the CPS that are in rotation groups 1 to 4, and the RHS e/p from a sample that includes only observations in rotation groups 5 to 8. These samples are disjoint because an individual in rotation groups 5 to 8 in year $t-1$ (or $t-2$) cannot be in rotation groups 1 to 4 in year t. There are other schemes that would provide slightly larger samples, but they would involve more complicated interactions between rotation group and calendar year. Summary statistics similar to those in Figure 1 and Table 1 for the full sample, but for our disjoint samples, are provided in Appendix A.

3. Dynamic Panel Methodology

3.1 Estimating Equation

Equation (1) is our baseline estimating equation, in which e/p is the detrended employment-to-population ratio, DA denotes the

[10] Indeed, as expected, estimation with the full sample suggests a larger amount of excess persistence than with the disjoint samples, although our conclusions in Section 8, in which we use simulations to assess the magnitude of our estimates, are not materially affected. We recognize that the smaller estimates from the disjoint samples could be due to attenuation bias from the smaller sizes of the disjoint samples. However, experimentation with random subsamples of the full sample that mimic the size of our disjoint samples indicates that attenuation bias is not a serious concern in this case.

disadvantaged group, s denotes state, t denotes year, the α are state fixed effects, the γ are year fixed effects, $Ugap$ is the UR gap, β and δ are coefficients, and ϵ is an error term:

$$(e/p)_{s,t}^{DA} = \alpha_s + \gamma_t + \beta_1 (e/p)_{s,t-1}^{DA} + \beta_2 (e/p)_{s,t-2}^{DA} + \delta_0 Ugap_{s,t} + \delta_1 Ugap_{s,t-1} + \delta_2 Ugap_{s,t-2} + \epsilon_{s,t}. \quad (1)$$

As described in Section 2.4, the e/p's on the left-hand and right-hand sides of Equation (1) are derived from disjoint samples.

This specification for $(e/p)_{s,t}^{DA}$ accounts for factors that differ across states but are constant over time (α_s) and for aggregate factors that change over time but are constant across states (γ_t). Thus, the estimation uses variation that is left over after removing within-state and within-year variation.

The coefficients β on the lagged detrended e/p terms capture persistence in the e/p *in excess of* that implied by the persistence in the overall UR gap.[11] Our approach for obtaining estimates of β_1 and β_2 in Equation (1) is equivalent to the following two-step procedure. First, regress $(e/p)_{s,t}^{DA}$ on state and year fixed effects and the UR gap and its two lags and obtain residuals, $\xi_{s,t}$. In this first step, we use two lags because the annual UR gap obtained with the FT approach is well approximated by an AR(2).[12] Second, obtain the two β coefficients from regressing $\xi_{s,t}$ on two own lags, $\xi_{s,t-1}$ and $\xi_{s,t-2}$. This second step captures the excess persistence in e/p after controlling for the effects of aggregate labor market conditions, as measured by the overall UR gap. In Appendix B.1 we show that these β parameters are a function of both individual effects on a person's employment, such as human capital accumulation and depreciation, as well as cross-individual effects, such as employment networks and competition.[13]

Of course, we are concerned about the endogeneity of the overall UR gap with respect to the detrended e/p for the disadvantaged

[11]This estimation strategy depends, of course, on there being sufficient cyclical variation in the UR gap to identify the relationships. If there were complete hysteresis in UR, for example, the strategy would fail.

[12]We investigated various lag lengths and, based on both formal tests and to avoid overfitting, settled on this lag structure.

[13]Our dynamic panel approach bears some resemblance to Blanchard and Katz (1992), which we discuss in Appendix B.2.

group, if for no other reason than that the disadvantaged group make up a sizable proportion of the labor force. In Section 5.2 we use several instruments for the UR gap and show that they do not change our conclusions.

A regression of squared residuals on the inverse of the number of observations, as suggested by Solon, Haider, and Wooldridge (2015), indicates significant heteroskedasticity in our data. Therefore we weight the regressions by the number of observations in the disadvantaged group in each state in each year.

It is well known that estimating dynamic panels with fixed effects may lead to biased estimates if the panel is short. Arellano (2003) argues that if the number of periods is at least 10, then this bias is likely small. Nickell (1981) shows that with reasonably long panels, the bias is around order $-(1+\beta)/T$, in which T is the length of the panel. As our data effectively span 38 years, if there is excess persistence ($\beta > 0$), the downward bias in the coefficient is likely small. Note, however, that Hershbein and Stuart (2020) argue otherwise.

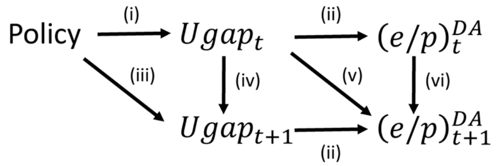
We are also not concerned that the size of our cross-section (N) induces bias. Monte Carlo simulations by Nerlove (1967), in which $N = 25$, suggest that the approximate formula for bias in Nickell (1981) is more or less exact when β is not too large. We have 50 cross-section observations in our baseline sample and our estimates of β are well below unity.

Although our data for the e/p go back to 1978, our estimates for trend UR from the FT model begin only in 1979. Between this constraint and the lag structure in Equation (1), the sample period for our baseline regressions is 1981 to 2018 (38 years), which, with 50 states, yields a total of 1,900 observations.

3.2 Identifying Aggregate Excess Persistence with Equation (1)

Below in Section 8 we will use the estimates from the reduced-form Equation (1) to simulate the implications of excess persistence for employment of the disadvantaged group over the business cycle. The validity of these inferences depends upon two assumptions. The first is that the variation in e/p under consideration is driven by variation in overall labor market conditions as represented by the UR gap, and that the degree of excess persistence in e/p is invariant to

Figure 2. Effects of Policy on the UR Gap and e/p



Note: We call channel (vi) excess persistence in e/p. We assume that variation in the UR gap is the only source of variation in detrended e/p that is relevant to generating excess persistence. See Section 3.2.1 for a discussion.

the source of that variation in the UR gap. The second is that the degree of excess persistence identified by the state panel regressions is applicable to the aggregate level. We discuss each of these in turn.

3.2.1 The UR Gap as a Sufficient Statistic for Labor Market Conditions

Figure 2 describes the causal relationship between economic policy actions, overall labor market conditions, and the (detrended) e/p of the disadvantaged group. A given policy action affects contemporaneous overall labor market conditions as represented by the UR gap (channel i), which, in turn, affects the e/p ratio of the disadvantaged group (channel ii). The e/p in year t further affects the e/p in year $t+1$ (channel vi) conditional on the contemporaneous and lagged influence of the UR gap (channels ii and v). It is this channel (vi) that we call excess persistence in e/p and estimate with the coefficients β in Equation (1). (For ease of exposition, the diagram includes only one lag of UR gap and of e/p, although our empirical model includes two lags. We also omit the s subscript.)

The appropriateness of β for the simulations in Section 8 depends importantly on the assumption that variation in overall labor market conditions represented by the UR gap is the only source of variation in detrended e/p that is relevant to generating excess persistence. This means that the phenomena (including economic policies) that drive the UR gap have no direct effect on the cyclical component of the e/p (i.e., the detrended e/p) of the disadvantaged group, or, more precisely, that any direct effect is highly correlated with the

UR gap. This assumption would seem to be approximately appropriate for monetary policy actions as contemplated by Yellen (2016, 2019). However, our data doubtless include variations in the UR gap driven by other phenomena. We depend upon these other phenomena being sufficiently acyclical that their persistent effects on e/p are captured in the estimated trend.

An alternative would be to identify particular types of policy shocks and estimate excess persistence in the response of e/p to those shocks. If the degree of excess persistence varies by the type of policy, contrary to our assumption that the policies we contemplate operate on e/p only through overall labor market tightness, that would have the value of allowing differentiated policy simulations. Of course, the validity of such estimates depends upon proper identification of the policy shocks, which is often controversial (Nakamura and Steinsson 2018, p. 61). In addition, the number of incidents of any particular policy intervention in the sample period are not large, which may pose a challenge for estimation.

Our approach does not rely on identification of particular shocks, but if the degree of excess persistence does, in fact, vary by the type of policy, then our estimates are an average across different true coefficients. In that case, they may not be valid for any particular policy intervention, but instead provide a general sense of magnitudes.

3.2.2 Aggregate Inference from State-Level Variation

Although we are ultimately interested in excess employment persistence at the national level, we use state-level variation in our estimation to improve identification. We then assume that the estimates from our state-level panel regressions can be used to simulate excess persistence at the national level in Section 8 below.

However, Beraja, Hurst, and Ospina (2019), among others, note several impediments to extrapolating state-level coefficients to the national level.[14] There are at least three ways to address these impediments. First, instead of applying the coefficients from the state-level equations directly, one could use them to discipline structural models. Second, one could obtain aggregate estimates

[14]Other examples include Nakamura and Steinsson (2014), Charles, Hurst, and Schwartz (2019), and Adao, Arkolakis, and Esposito (2020).

by embedding the state-level equation within an explicitly spatial model, as in Adao, Arkolakis, and Esposito (2020). Third, one could concentrate on situations in which the impediments are, as a practical matter, minor. We adopt this third solution. For this solution to be valid, our application needs to satisfy three conditions.

The first condition is that the sources of change, and mechanisms through which those changes operate, be the same at the state and national level. As noted above, our model assumes that changes in overall labor market conditions as represented by the UR gap are the only contemporaneous drivers of changes in detrended e/p, so the proximate source of change is the same at the state and national levels. Furthermore, the mechanisms posited to produce excess persistence—accumulation of human capital and market capital—operate at the individual level, and so are the same in national as in state-level data.

The second condition is that there be sufficient variation across states over time in the right-hand side variables to identify the coefficients.[15] In our case, it is well recognized that both the size and timing of business cycle changes in unemployment and employment rates vary substantially across states (e.g., Owyang, Piger, and Wall 2005). The quantity of state-specific variation in these variables need not be large relative to the common variation to obtain reliable estimates. But the assumption that the phenomena that drive the movements in the UR gap have no direct effect on detrended e/p is important here. As emphasized by Nakamura and Steinsson (2014) and Chodorow-Reich (2019), if nationally uniform policies that move the unemployment rate also directly affected the detrended e/p, these would be absorbed by the year indicators in Equation (1) and our β coefficients would not capture that effect.

The third condition is that the change in one state not affect the outcome in another state in ways that are not accounted for in the econometric model. In our application, the main concern in this regard is interstate migration in response to differences across states in overall labor market conditions. In work not shown, we found that over our sample period the interstate migration of our

[15] Static cross-sectional variation will be captured by the state fixed effects, while movements over time that are common to all states will be captured by the time fixed effects.

Table 2. Baseline Estimates

	Coefficient (Std. Err.)
$(e/p)_{s,t-1}$	0.25*** (0.02)
$(e/p)_{s,t-2}$	0.14*** (0.03)
$Ugap_{s,t}$	–1.24*** (0.09)
$Ugap_{s,t-1}$	0.36*** (0.12)
$Ugap_{s,t-2}$	0.42*** (0.11)
Observations	1,900
Within R-squared	0.29

Note: The degree of excess persistence among prime-age men with no more than a high school education is moderate. These are the estimated coefficients from Equation (1). The dependent variable is the detrended e/p of disadvantaged workers, $(e/p)_{s,t}$. $Ugap_{s,t}$ is the UR gap in state s at time t. Weighted by number of observations of the disadvantaged group. Driscoll-Kraay standard errors in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1. See Section 3 for the specification and Section 4 for a discussion of the results.

baseline demographic group did not respond in an economically or statistically significant degree to such differences—findings that are consistent with those of Bound and Holzer (2000) and Notowidigdo (2020).[16]

If these conditions are met, then the coefficient β estimated from state-level variation in UR gaps is a valid estimate of the degree of excess persistence at the national level.

4. Baseline Estimates

We find a moderate degree of excess persistence in employment among disadvantaged men. The estimates are shown in Table 2.[17] The coefficients on the lagged detrended e/p are significantly positive, indicating some excess persistence. However, these coefficients,

[16]Details are available from the authors.

[17]Throughout the paper, we show Driscoll-Kraay standard errors, with a lag length of 3, to allow for both spatial and temporal dependence in our state-panel regressions (Driscoll and Kraay 1998). As an alternative, we have also estimated standard errors clustered on year and state. There was no consistent pattern across the coefficients of which method yielded larger estimates of the standard errors, and our conclusions are not sensitive to this choice.

as well as the results in Section 8, indicate that within three years the effect of the lagged e/p has virtually no effect on the current e/p.

Note that the coefficients on $Ugap_{s,t}$ and its lags are of opposite signs in Table 2. This is similar to the results in Fleischman and Gallin (2001) and Fleischman, Gallin, and Smith (2018), in which the coefficients on the GDP gap and the lagged GDP gap have opposite signs. One interpretation of this is that changes in the UR gap, in addition to the level, have a short-run effect on the detrended e/p of the disadvantaged group, as is common in models of wage growth (Blanchard and Galí 2010). As we will see in Section 8.2, this property leads to the e/p "overshooting" its trend in some simulations.

5. Robustness Exercises

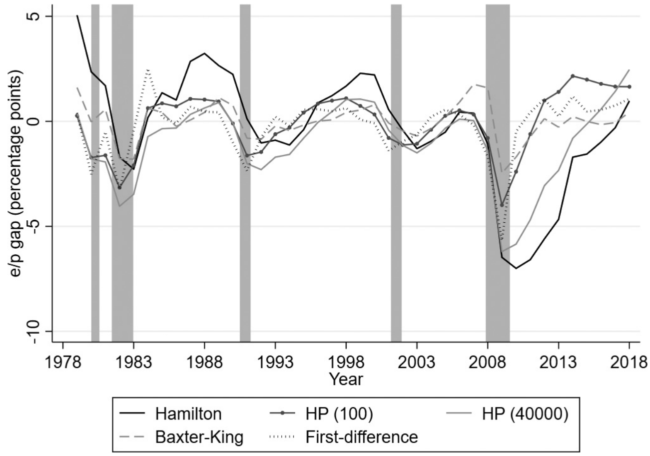
5.1 Detrending Methods for e/p

In this section we show that our results are robust to various methods of addressing trends in the e/p of disadvantaged workers.

As described above, our baseline estimates use the detrended e/p of the disadvantaged group in Equation (1) to concentrate on persistence stemming from cyclical fluctuations instead of lower-frequency (structural) phenomena. The choice of detrending method may therefore be important for identifying these cyclical movements, and in turn for the estimates of excess persistence. We try several methods for detrending state e/p, including one-sided filters, two-sided filters, and simple parametric time trends, as well as no attempt to account for trends. We find that these alternatives all imply similar or smaller estimates of excess persistence than our baseline approach.

Beginning with the filtering methods, in addition to our baseline approach that uses Hamilton (2018), we use the one-sided Hodrick-Prescott (HP) filter (Stock and Watson 1999b) with two smoothing parameters (100 and 40,000); first differencing; and a Baxter-King band-pass filter with a period of two to eight years and three-year smoothing. These filters amplify various frequencies of a series. The cyclical component from the Hamilton filter with a five-year horizon parameter recovers the spectral density function of white noise, as shown in Appendix C.1. The remaining filters remove lower-frequency components of the time series and pass through higher-frequency components, to a greater or lesser extent.

Figure 3. Estimates of State Detrended e/p, Aggregated



Note: State-level detrended e/p estimated by five different approaches, aggregated to the national level. See Section 5.1 for details.

The cyclical component of each filter gives a qualitatively similar account of the national detrended e/p since 1978, as shown in Figure 3. In particular, the deviations of e/p relative to trend were largest during the 2008–09 recession, less severe during the 1980s recession, and smallest during the 1991 and 2001 recessions. All the filters suggest that e/p was above trend in 2018. The detrended e/p from the Hamilton filter, reproduced from the lower panel of Figure 1, is the most procyclical out of our chosen filters. For example, during the 2008–09 recession, the Hamilton filter suggests that e/p fell over 6 percentage points relative to trend. (The cyclical components of the disjoint samples are similar to the full sample and shown in Appendix C.2.)

Our estimates of Equation (1) using the various detrending methods are shown in Table 3. Column 1 repeats our baseline specification with the detrended e/p using the approach in Hamilton (2018). Columns 2 through 5 present the results using the Baxter-King filter, the first-difference approach, and HP filter with 100 and 40,000 smoothing parameters, respectively. Column 6 uses the Hamilton method to detrend both the e/p ratio and the UR, to examine

Table 3. Excess Persistence Using Different Estimates of e/p Trends

	Baseline (Ham.) (1)	BK (2)	First Diff. (3)	HP (100) (4)	HP (40,000) (5)	Ham. for e/p and UR (6)
$(e/p)_{s,t-1}$	0.25***	–0.08***	0.27***	0.07***	0.21***	0.22***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
$(e/p)_{s,t-2}$	0.14***	–0.04**	0.18***	0.04**	0.15***	0.12***
	(0.03)	(0.02)	(0.01)	(0.02)	(0.02)	(0.03)
$Ugap_{s,t}$	–1.24***	–0.53***	–0.04	–0.94***	–1.22***	–1.10***
	(0.09)	(0.04)	(0.1)	(0.05)	(0.09)	(0.17)
$Ugap_{s,t-1}$	0.36***	0.29***	–1.25***	0.56***	0.56***	–0.10
	(0.12)	(0.07)	(0.15)	(0.07)	(0.10)	(0.19)
$Ugap_{s,t-2}$	0.42***	0.04	0.76***	0.41***	0.39***	0.45***
	(0.11)	(0.06)	(0.09)	(0.07)	(0.10)	(0.12)
Obs.	1,900	1,900	1,900	1,900	1,900	1,900
Within R2	0.29	0.06	0.32	0.19	0.25	0.30

Note: Using different estimates of state trend e/p implies similar or smaller estimates of excess persistence to our baseline approach (column 1). These are the estimated coefficients from Equation (1), in which we use different methods to detrend e/p. The disadvantaged group is prime-age men with no more than a high school education. The dependent variable is the detrended e/p of disadvantaged workers. $Ugap_{s,t}$ is the UR gap in state s at time t. In columns 1 through 5 we estimate the UR gap using the model developed in Fallick and Tasci (2020) (see Section 2). In column 6 we use the method described in Hamilton (2018) to detrend the UR. Weighted by number of observations of the disadvantaged group. Driscoll-Kraay standard errors in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$. See Section 5.1 for details.

whether using different methods for estimating trends for these two quantities in the baseline specification is driving our results.

Using different estimates of state trend e/p implies similar or smaller estimates of excess persistence to our baseline approach. The first-difference and HP (40,000) filters imply a similar amount of excess persistence to our baseline approach. Other filters imply smaller estimates of excess persistence. Using the Hamilton (2018) method for detrending both e/p and UR—instead of the Hamilton method for e/p and the FT method for UR—yields similar results to our baseline approach, as shown in column 6. The detrended e/p from the Baxter-King and HP (100) filters are less procyclical than the other methods, consistent with the aggregated results in Figure 3.

Despite the robustness to various filtering methods, one may be concerned that if excess persistence is sufficiently long-lived, then any of these methods may attribute some of the excess persistence to the trend. This may be of particular concern if some secular movements have their origins in cyclical phenomena, as may be the case, for example, for the number of persons receiving disability payments (Aaronson et al. 2014). We therefore also estimate Equation (1) with no detrending at all, which should provide an upper bound on the degree of excess persistence, and with simple parametric time trends, which should be less susceptible to mistaking persistent cyclical movements for structural trends.

We present these results in Table 4. Column 1 repeats our baseline specification (with the detrended e/p). Columns 2 through 4 instead use actual e/p on both sides of Equation (1). Column 2 makes no attempt to account for trends in e/p. Column 3 includes linear time trends in e/p and column 4 includes quadratic time trends.

As expected, the coefficients on the lagged e/p in column 2 are larger than in the baseline. However, the differences are small. The inclusion of the parametric time trends results in smaller estimates of excess persistence than in the baseline.

5.2 Instrumenting for the UR Gap

The overall UR gap may be endogenous with respect to the detrended e/p for the disadvantaged group, if for no other reason

Table 4. No Time Trend and State-Specific Time Trends

	Baseline (1)	No Trends (2)	Linear Trends (3)	Quadratic Trends (4)
$(e/p)_{s,t-1}$	0.25***	0.27***	0.14***	0.10***
	(0.02)	(0.02)	(0.03)	(0.03)
$(e/p)_{s,t-2}$	0.14***	0.19***	0.08***	0.05*
	(0.03)	(0.02)	(0.03)	(0.03)
$Ugap_{s,t}$	–1.24***	–1.29***	–1.36***	–1.42***
	(0.09)	(0.11)	(0.09)	(0.09)
$Ugap_{s,t-1}$	0.36***	0.59***	0.43***	0.42***
	(0.12)	(0.11)	(0.08)	(0.08)
$Ugap_{s,t-2}$	0.42***	0.20**	0.021	–0.12
	(0.11)	(0.09)	(0.08)	(0.09)
Observations	1,900	1,900	1,900	1,900
R-squared	0.29	0.32	0.28	0.27

Note: Without detrending, estimates of excess persistence are slightly larger than in our baseline and using state-specific time trends reduces the estimates of excess persistence. These are the estimated coefficients from Equation (1), in which we include no detrending, as well as linear and quadratic state-specific time trends. The dependent variable is the actual employment-to-population ratio of disadvantaged workers. The disadvantaged group is prime-age men with no more than a high school education. $Ugap_{s,t}$ is the UR gap in state s at time t, in which we estimate the trend UR gap using the model developed in Fallick and Tasci (2020) (see Section 2). Column 1 produces the baseline OLS regression from Table 2. Column 2 includes no detrending of actual e/p. Columns 3 and 4 include linear and quadratic trends, respectively. Weighted by number of observations of the disadvantaged group. Driscoll-Kraay standard errors in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$. See Section 5.1 for details.

than that the disadvantaged group in a particular state make up a sizable proportion of the labor force.

To address this potential endogeneity, we use three approaches to instrument for state UR gaps as estimated by the FT model. First, we instrument with state GDP (U.S. Bureau of Economic Analysis 2019) gaps because state GDP gaps are not mechanically related to the detrended e/p for the disadvantaged group. Second, we use the "leave-out" mean of the UR gap in the state's region because the UR gaps of other states in a state's region reflect similar demand conditions but should be approximately exogenous to the e/p in the state in question. Third, we use the detrended insured unemployment rate (IUR), defined as the number of individuals receiving UI benefits over all covered employment. The IUR is correlated with a state's UR but should be exogenous to a state's e/p because the IUR reflects the level of labor demand, benefits being designed to be paid only to individuals who lose a job through no fault of their own (e.g., laid off or position abolished), and not to individuals who quit or are fired for cause (U.S. Department of Labor 2018).

To obtain state GDP gaps we detrend state GDP using the procedure suggested by Hamilton (2018), with a five-year horizon parameter. To give this filter a running start ahead of our sample period, we estimate the filter from 1970 onward. However, data on state GDP are available beginning only with 1977. In order to minimize end-of-sample bias, we augment the GDP series on both ends, as we did for e/p in Section 2.3. Column 1 of Table 5 repeats our baseline ordinary least squared (OLS) regression. Column 2 shows the estimates using this instrument.

To obtain the "leave-out" mean of the UR gap in a state's region (the "regional Ugap"), we use the eight clusters of the 48 contiguous states identified by Crone (2005) as having similar business cycles. To adjust the baseline for comparison to this instrument, column 3 of Table 5 presents the estimated OLS coefficients from Equation (1) when using only the 48 contiguous states. Column 4 presents the results when instrumenting the UR gap with the regional UR gap, in which all UR gaps are estimated with the FT approach (as in the baseline).

To detrend the IURs we use an HP filter with a smoothing parameter of 1,600.[18] State-level insured unemployment data only start in 1986, so for comparison, in column 5 we rerun the baseline OLS regression for that shorter date range. Column 6 presents the results using the detrended IUR as an instrument.

[18]As with GDP, we augment the insured unemployment data on both ends using second-order regressive models to reduce endpoint bias.

Table 5. Instrumenting for UR Gaps with GDP Gaps, Regional UR Gaps, and IUR

	OLS Baseline (1)	**IV GDP (2)**	**OLS Excluding AK and HI (3)**	**IV Regional Ugap (4)**	**OLS 1986–2018 (5)**	**IV IUR 1986–2018 (6)**
$(e/p)_{s,t-1}$	0.25***	0.17***	0.25***	0.14***	0.26***	0.16***
	(0.02)	(0.05)	(0.02)	(0.03)	(0.02)	(0.03)
$(e/p)_{s,t-2}$	0.14***	0.14***	0.14***	0.08**	0.14***	0.06
	(0.03)	(0.02)	(0.03)	(0.04)	(0.03)	(0.04)
$Ugap_{s,t}$	–1.24***	–1.85***	–1.22***	–1.66***	–1.33***	–1.83***
	(0.10)	(0.8)	(0.10)	(0.32)	(0.10)	(0.29)
$Ugap_{s,t-1}$	0.36***	0.22	0.34***	–0.10	0.26*	0.10
	(0.12)	(1.4)	(0.11)	(0.52)	(0.14)	(0.50)
$Ugap_{s,t-2}$	0.42***	0.69	0.40***	0.53*	0.53***	0.22
	(0.11)	(0.75)	(0.10)	(0.32)	(0.13)	(0.35)
FS F-stat		942		164		60
Obs.	1,900	1,900	1,824	1,824	1,650	1,650
Within R2	0.29	0.23	0.28	0.18	0.28	0.17

Note: For each instrument, the estimated excess employment persistence is no greater than in the OLS baseline. These are the estimated coefficients from Equation (1) with instruments for the overall UR in state s at time t and its lags. The dependent variable is the detrended e/p of disadvantaged workers, $(e/p)_{s,t}$. The disadvantaged group is prime-age men with no more than a high school education. $Ugap_{s,t}$ is the UR gap in state s at time t, in which the trend is estimated using the Fallick-Tasci approach (Section 2). Column 1 reproduces the baseline OLS regression from Table 2. In column 2 we instrument for the UR gap with state GDP gaps. In column 3 we restrict the sample for the OLS regression to the 48 contiguous states. In column 4 we instrument for the UR gap with the average UR of the other states in a state's region as defined by Crone (2005). In column 5 we restrict the sample for the OLS regression to years when the IUR is available. Column 6 instruments for the UR gap with the IUR gap. Weighted by number of observations of the disadvantaged group. "FS" stands for "First Stage." Driscoll-Kraay standard errors in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$. See Section 5.2 for details.

In each case, the estimated excess employment persistence when we instrument for the UR gap is no greater than in the OLS baseline.[19]

6. Different Definitions of the Disadvantaged Population

The definition of the disadvantaged group is necessarily somewhat arbitrary. We have so far defined the disadvantaged group as prime-age men with no more than a high school education because, among men, this group has seen substantial deterioration in relative earnings and employment in recent decades, has generally lower employment rates than other education groups, has more procyclical employment rates, and its members are more likely to be Black or Hispanic. However, these characterizations are all the more apt for persons with less than a high school education, and to Blacks and Hispanics themselves. In addition, younger persons have had less opportunity for previous accumulation of human and market capital, and so may have more to gain from a bout of employment and more to lose by missing out on employment, while older persons may exhibit more excess persistence in employment because of age discrimination in hiring (Neumark, Burn, and Button 2019).

Table 6 explores these possibilities by varying the definition of the disadvantaged population. Column 1 repeats our baseline specification, which treats prime-age (25 to 54) men with no more than a high school diploma as the disadvantaged group. Column 2 narrows the baseline sample to prime-age men with less than a high school education. Column 3 narrows the baseline sample to Black men and Hispanic men ages 25 to 54 with no more than a high school education.[20] Column 4 narrows the baseline sample to men ages 18 to

[19] In addition to the estimated coefficients, Table 5 reports the first-stage F statistic to show that our instruments are all strongly correlated with the detrended e/p of the disadvantaged group. We report the statistic proposed by Cragg and Donald (1993). Our test statistics are above conventional critical values presented in Stock and Yogo (2005).

[20] We also tried samples of Black men and Hispanic men separately. Unfortunately, the samples in the CPS data were too small to allow reasonable estimation. For example, for Black men with no more than high school the smallest state averages just 6 observations, and limiting the sample to only states with even 75 observations eliminates half of the states.

Table 6. Different Definitions of Disadvantaged

	Baseline (25–54, ≤ HS) (1)	**Less Educ. (25–54, < HS) (2)**	**Blacks & Hispanics (25–54, ≤ HS) (3)**	**Younger (18–34, ≤ HS) (4)**	**Older (45–64, ≤ HS) (5)**
$(e/p)_{s,t-1}$	0.25***	0.17***	0.026	0.21***	0.19***
	(0.02)	(0.02)	(0.04)	(0.03)	(0.02)
$(e/p)_{s,t-2}$	0.14***	0.091***	0.034	0.13***	0.096***
	(0.03)	(0.03)	(0.05)	(0.03)	(0.02)
$Ugap_{s,t}$	–1.24***	–1.15***	–0.91***	–1.76***	–0.56***
	(0.10)	(0.21)	(0.20)	(0.17)	(0.12)
$Ugap_{s,t-1}$	0.36***	0.37	0.27	0.54**	–0.064
	(0.12)	(0.28)	(0.46)	(0.25)	(0.18)
$Ugap_{s,t-2}$	0.42***	0.17	–0.21	0.42**	0.25**
	(0.11)	(0.23)	(0.32)	(0.18)	(0.12)
Obs.	1,900	1,900	1,900	1,900	1,900
Within R2	0.29	0.086	0.02	0.28	0.10

Note: Different definitions of disadvantaged do not suggest greater excess persistence in employment. These are the estimated coefficients from Equation (1) in which the dependent variable is the detrended e/p of various definitions of disadvantaged workers, $(e/p)_{s,t}$. $Ugap_{s,t}$ is the UR gap in state s at time t, in which we estimate the trend UR gap using the model developed in Fallick and Tasci (2020) (see Section 2). Weighted by number of observations of the disadvantaged group. Driscoll-Kraay standard errors in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$. See Section 6 for details.

34 with no more than a high school education. Column 5 narrows the baseline sample to older men ages 45 to 64 with no more than a high school education.

In all cases, the coefficients indicate less excess persistence than in the baseline. We take these results with a grain of salt, because the smaller sizes of the samples in the CPS data may lead to noisier measures of the lagged e/p and therefore to attenuation of those coefficients.[21]

7. Asymmetries

High and low detrended e/p may have asymmetric effects on future employment outcomes. For example, skills may be slower to deteriorate through non-use than they are to accrue through use, while the formation of networks may display the opposite pattern. To allow for such asymmetry, we split the lagged detrended e/p term into two components: one for the e/p above its trend (positive detrended e/p) and one for the e/p below its trend (negative detrended e/p).

Column 1 of Table 7 repeats the baseline specification. Column 2 introduces asymmetric linear terms. The estimates do not indicate significant asymmetry. Although the point estimates for the second lag of e/p do show more persistence in the positive direction, F-tests (not shown) cannot reject that the coefficients on the positive and negative e/p are equal at conventional significance levels. In column 3 we add quadratic terms in each asymmetric detrended e/p to allow for the possibility that extremely high employment or extremely low employment has a larger marginal effect than smaller deviations from trend. Here, too, one cannot reject symmetry.

8. Simulations

In this section we provide simulations to help interpret the magnitude of our baseline estimates of employment persistence from

[21] Attenuation bias is a potential concern with the baseline definition as well, of course. However, the sample sizes for that group are large: The smallest state averages 775 observations. In contrast, for example, for the less-than-high-school group the smallest state averages 139.

Table 7. Asymmetry

	Baseline (1)	**Linear Asymmetry (2)**	**Quadratic Asymmetry (3)**
$(e/p)_{s,t-1}$	0.25***		
	(0.03)		
$(e/p)_{s,t-2}$	0.14***		
	(0.02)		
$(e/p\ positive)_{s,t-1}$		0.25***	0.21*
		(0.05)	(0.12)
$(e/p\ positive\ squared)_{s,t-1}$			0.01
			(0.02)
$(e/p\ negative)_{s,t-1}$		0.25***	0.21***
		(0.03)	(0.06)
$(e/p\ negative\ squared)_{s,t-1}$			–0.01
			(0.01)
$(e/p\ positive)_{s,t-2}$		0.19***	0.25**
		(0.04)	(0.10)
$(e/p\ positive\ squared)_{s,t-2}$			–0.01
			(0.01)
$(e/p\ negative)_{s,t-2}$		0.11**	0.14**
		(0.04)	(0.06)
$(e/p\ negative\ squared)_{s,t-2}$			0.00
			(0.01)
$Ugap_{s,t}$	–1.24***	–1.25***	–1.24***
	(0.09)	(0.09)	(0.09)
$Ugap_{s,t-1}$	0.36***	0.36**	0.37**
	(0.12)	(0.12)	(0.11)
$Ugap_{s,t-2}$	0.42***	0.41***	0.41***
	(0.11)	(0.11)	(0.11)
Obs.	1,900	1,900	1,900
Within R2	0.28	0.284	0.285

Note: Estimates do not indicate significant asymmetry. These are the estimated coefficients from versions of Equation (1) in which we split the lagged detrended e/p term into two components: above and below trend. The dependent variable is the detrended e/p ratio of disadvantaged workers, $(e/p)_{s,t}$. The disadvantaged group is prime-age men with no more than a high school education. $Ugap_{s,t}$ is the UR gap in state s at time t, in which we estimate the trend UR gap using the model developed in Fallick and Tasci (2020) (see Section 2). Weighted by number of observations of the disadvantaged group. Driscoll-Kraay standard errors in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1. See Section 7 for details.

Section 4 and their implications for policymakers. We focus on our baseline demographic group—prime-age men with no more than a high school education—because this group exhibited the largest amount of excess persistence in employment of the groups that we examined (see Section 6).

8.1 Historical Simulations

As noted above, policymakers have been interested in the possibility that the employment benefits of a high-pressure economy for disadvantaged groups may persist even after overall labor market conditions have normalized. This idea implies that there are mechanisms, such as the accumulation of human capital and network capital, that are distinct from those that generate the response of the e/p of the disadvantaged to overall labor demand, and so may persist after that overall demand has normalized. In motivating Equation (1), we argued similarly that the individual-level mechanisms that would generate excess employment persistence, and are reflected in coefficients β, are distinct from those that generate the direct relation between detrended e/p and the UR gap, which are reflected in coefficients δ (see Section 3 and Appendix B.1).[22] If this distinction is valid, then by setting the β coefficients to zero while leaving the δ coefficients at their estimated values, we can obtain a counterfactual e/p of the disadvantaged group that would obtain in the absence of the individual-level mechanisms underlying excess persistence. In this section we use such a counterfactual to quantify the cumulative effect of excess persistence since the mid-1990s.

First, we simulate e/p using the estimated coefficients from Equation (1) in Table 2. Second, we simulate e/p setting the coefficients on the lagged e/p terms to zero while leaving the coefficients on the UR gap and its lags as they are in Table 2. The cumulative difference between these two simulations is a measure of the contribution of excess persistence to the e/p of the disadvantaged group over this

[22] Put another way, if Equation (1) was the outcome of a microfounded model that included the relevant mechanisms, we are assuming that the structural parameters contributing to the coefficients on lagged e/p would be distinct from those contributing to the coefficients on the UR gap.

period, as explained above.[23] Comparing the two simulations, we find that the cumulative effect of excess persistence in the business cycle surrounding the 2001 recession was mildly positive, while the effect in the cycle surrounding the 2008–09 recession was decidedly negative.

The details for both simulations are as follows. We simulate the e/p of the disadvantaged group in each state for each year between 1987 and 2018. To do so, we set the UR gap to its observed value in each state and year 1985 to 2018. We set the detrended e/p of the disadvantaged group to zero in the two years (1985 and 1986) before the simulation commences.[24] In both simulations we set the δ coefficients and state and calendar-year effects to their estimated values. As just noted, in the first simulation we set the β coefficients to their estimated values, but in the second simulation we set the βs to zero. For ease of presentation, we then aggregate the two simulated e/p's from the state to the national level, and concentrate on the period beginning in 1996, a year in which the national UR was near the Congressional Budget Office's (CBO's) estimate of the natural rate of unemployment.

Before presenting our results, we note that the e/p simulated using our estimated coefficients follows a similar trajectory to the actual e/p over the 1996 to 2018 period, as shown in Figure 4.[25] This result suggests that our dynamic panel model (Equation (1)) accounts well for the cyclical variations in the actual e/p.

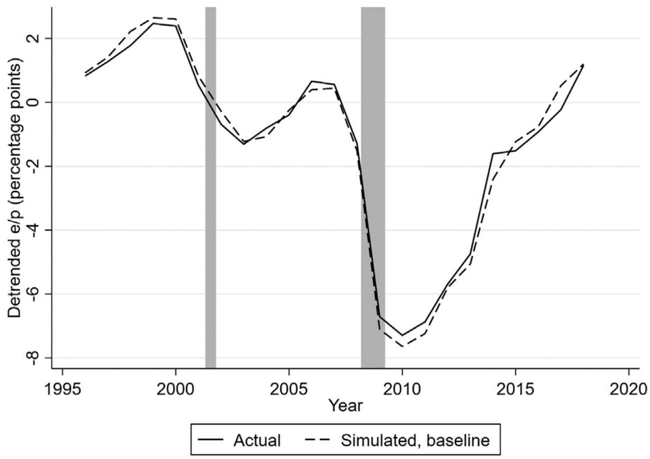
Figure 5 shows the estimated contribution of excess persistence to the e/p of the disadvantaged group. During the tight labor market

[23]An alternative would be to reestimate the equation imposing the restriction that the coefficients on the lagged e/p be zero, and use the coefficients on the *Ugap* terms from that regression in the counterfactual simulation. In that case, however, the coefficients on the *Ugap* terms would reflect excess persistence in the e/p to the extent that it is correlated with the persistence in *Ugap*. In this case the difference would not measure the contribution of excess persistence if some exists.

[24]The outcomes of interest are not sensitive to this choice of the initial e/p. Nor are they sensitive to the choice of starting year.

[25]To obtain the detrended actual e/p, we use the full sample of the CPS as opposed to the disjoint samples we used for estimating the amount of employment persistence. We take this approach because, while the disjoint samples minimize correlated measurement error for purposes of estimation (Section 2.4), the full sample provides the best estimate of the e/p for any given year.

Figure 4. Detrended e/p of Disadvantaged Group: Actual and Historical Simulation

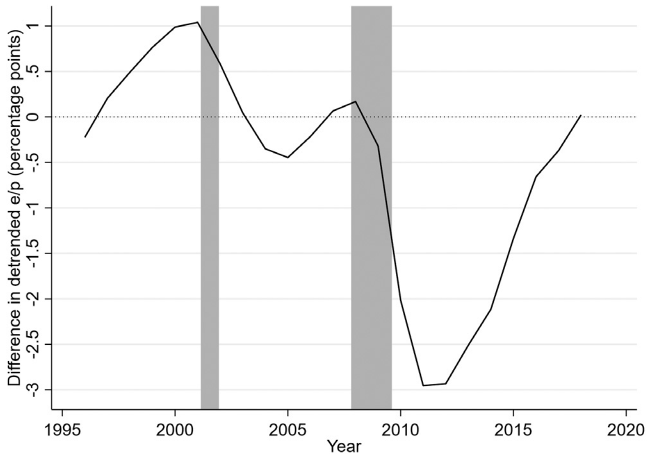


Note: The detrended e/p simulated using our estimated coefficients follows a similar trajectory to the detrended actual e/p over the 1996 to 2018 period. The figure shows the detrended actual e/p over the 1996 to 2018 period along with the one-step-ahead dynamic simulation of Equation (1) using our estimated baseline coefficients in Table 2. The simulation is done at the state level and then aggregated to the national level. For the UR gap we use the actual UR less an estimate using the model developed in Fallick and Tasci (2020) (see Section 2). See Section 8.1 for details.

toward the end of the 1990s expansion and before the 2001 recession, excess persistence served to buoy the e/p of the disadvantaged group by up to 1 percentage point at its height. Following the recession, however, excess persistence pulled in the opposite direction, weighing on the e/p of this group, with the contribution turning negative by 2004. Cumulatively, the former benefit outweighed the latter cost.

The situation is, unfortunately, quite different in the subsequent business cycle. The labor market was not as tight toward the end of the 2000s expansion as it was in the previous cycle, so the contribution of excess persistence barely moved into positive territory. The severity of the 2008–09 recession, however, meant that excess persistence weighed on the e/p of this group by almost 3 percentage points in 2011 and 2012, and only in 2018, when the national UR was

Figure 5. Contribution of Excess Persistence in Historical Simulation



Note: On net, excess persistence benefits disadvantaged workers during the business cycle around the 2001 recession but harms them during the cycle around the 2008–09 recession. This figure plots the difference between the simulated e/p from Equation (1) using all of the coefficients from Table 2 and the simulated e/p using the estimated coefficients for the *Ugap* terms but setting the coefficients on the lagged e/p terms to zero. This difference captures the contribution of excess persistence to the e/p of disadvantaged workers. The horizontal dotted line denotes zero.

0.7 percentage point below the CBO's natural rate of 4.6 percent, did excess persistence stop pushing down the e/p of disadvantaged workers.[26] Cumulatively, the costs of excess persistence during and after the 2008–09 recession far outweighed the benefits during the late stage of the previous expansion.

8.2 Policy Simulations

In this section we address policymakers' interest in the possible lasting employment benefits of a "high-pressure economy" (Ball 2015; Yellen 2016) for disadvantaged groups by providing a sense of the

[26] Replacing the baseline equation with an asymmetric specification from Section 7 does not qualitatively alter these conclusions.

likely magnitude of such benefits. These magnitudes are of particular interest if one is concerned that a high-pressure economy may increase the risk of subsequent recession, either because of high inflation and ensuing policy response (Lacker 2017; Bostic 2018) or other business cycle dynamics (Beaudry, Galizia, and Portier 2015, 2016; Feldstein 2018; Kiley 2018; Jackson and Tebaldi 2019). If this is the case, then the potential benefits of a high-pressure economy must be traded off against the possible costs.

We represent a high-pressure economy by the late stage of the 1990s expansion (we discuss this choice below). We convey a sense of the potential benefits from excess persistence by simulating the detrended e/p of the disadvantaged group from a UR that rises gradually from its low in 2000 to the natural rate in 2005 ("no-recession scenario"). We convey a sense of the potential costs from excess persistence by simulating the detrended e/p of the disadvantaged group from a UR that imitates the 2001 recession: rising to 6 percent in 2003 before falling back to the natural rate in 2005 ("recession scenario"). Not surprisingly given the modest degree of excess persistence implied by our estimated coefficients, we find that neither the lasting benefits nor the lasting costs are large, although the paths of employment are quite different in the two simulations.

In contrast to the historical simulations in Section 8.1, here we simulate directly at the national level in order to more easily specify historical paths for the UR.[27]

We chose the late stage of the 1990s expansion to represent a high-pressure labor market because in 2000 the national UR was as far below the CBO's estimate of its natural rate as occurred during the span of our data. Broad indexes of labor market conditions (KC LMCI 2021, for example) also suggest that the late stage of the 1990s expansion was the tightest labor market in that span.

To abstract from changes over time that are not due to the assumed paths for overall labor market conditions, we set the trend UR in every year of the simulations equal to the CBO's estimate of the long-run natural rate for 2005 (5.0 percent), and set all of the

[27]Because our empirical model assumes that the β and δ coefficients in Table 2 are constant across states, it would make little difference if we performed the simulations at the state level and aggregated to the national level.

year effects to the estimated year effect for 2005.[28] As in the historical simulations in Section 8.1, we begin the simulation in 1987 and set the detrended e/p of the disadvantaged group to zero in the two years before the simulations commence.

We show the hypothesized paths for the UR in the upper panel of Figure 6. In the no-recession scenario (gray line), we set the UR to the actual UR from the beginning of the simulation through the year 2000; from 2001 to 2005 we set the UR to rise linearly to the natural rate; and from 2005 on we hold the UR steady at the natural rate.

The recession scenario (black line) differs from the no-recession scenario only in the assumed path for the UR between 2000 and 2005 (the dashed lines in the upper panel of Figure 6 denote these two years). In this scenario, we set the UR to its actual value from 2001 (the year the recession commenced) through 2003 (the year in which the UR peaked during that cycle). We then set the UR to decline at a constant rate to trend in 2005. Thus this scenario includes something very much like the 2001 recession.

We show the simulated paths of detrended e/p in these two scenarios in the lower panel of Figure 6. For ease of exposition, we show the simulated detrended e/p as the deviations from the "steady-state" detrended e/p implied by our baseline coefficients.[29]

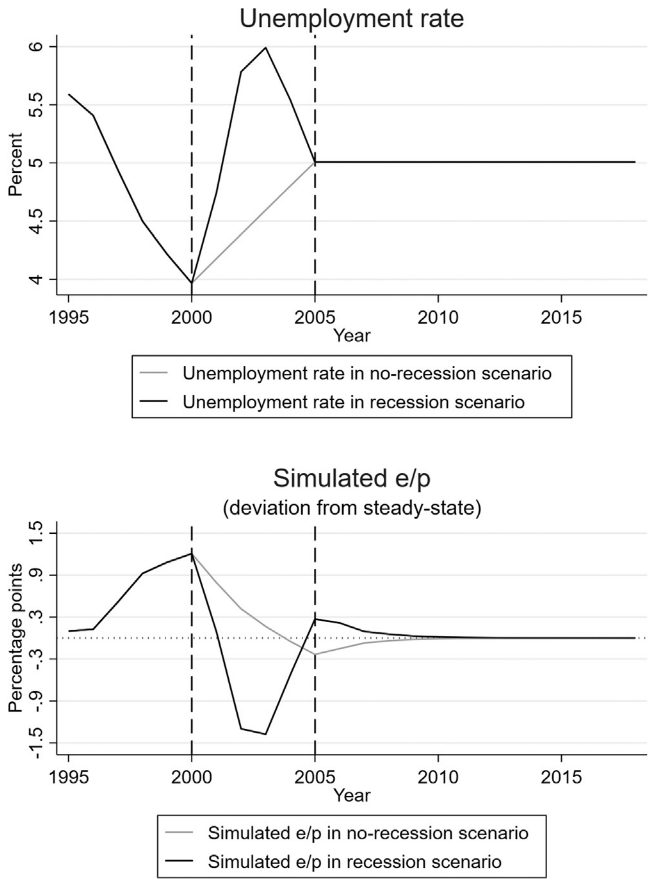
In the no-recession scenario, the lasting benefit from the tight labor market in 2000 is small. Naturally, the e/p rises above the steady state (positive in the graph) into the tight labor market of the late 1990s. The detrended e/p then falls toward the steady state as the UR reverts to trend. By 2005, when the UR returns to trend, the detrended e/p has essentially returned to its steady-state level, despite some small overshooting (see Section 4).

In the recession scenario, the lasting cost of the 2001 recession are also small. The detrended e/p falls during 2001 to 2003 as the UR rises, then rises as the UR falls. By 2005, when the UR has returned

[28] In 2005 the national UR was quite close to the CBO's estimate of the natural rate.

[29] We define a steady-state e/p as the solution for $(e/p)_t^{DA}$ in Equation (1) when $(e/p)_t^{DA} = (e/p)_{t-1}^{DA} = (e/p)_{t-2}^{DA}$, $Ugap_t = Ugap_{t-1} = Ugap_{t-2} = 0$, and $\gamma_t = \gamma_{2005}$. There are no s subscripts because these policy simulations are performed at the aggregate level.

Figure 6. No-Recession and Recession Scenarios in Policy Simulation



Note: The lasting employment benefits of temporarily running a “high-pressure” economy are small. The top panel shows the trajectory of the assumed UR for two scenarios: a “no-recession” scenario and a “recession” scenario. The lower panel shows the deviations of e/p from steady state in these two scenarios. After 2005, when the UR returns to trend, the lasting employment benefits and costs are small in the two scenarios. The time between the two vertical lines denotes the period over which the aggregate UR is assumed to be different between the two scenarios. The horizontal dotted line in the lower panel denotes zero.

to neutral, the detrended e/p has returned to its steady-state level except for a small amount of overshooting.[30]

In short, while the contemporaneous benefit for the e/p of disadvantaged workers of a high-pressure economy, and the contemporaneous cost should it be followed by a recession, are clear, neither has a significant lasting effect on the e/p of this group.[31]

These simulations used the coefficients estimated from our baseline sample. As noted in Section 6, this baseline group exhibits the largest amount of excess persistence among the groups we examined. At the other end of the spectrum, we found the least amount of excess persistence in the sample of Black or Hispanic men, ages 25 to 54, with no more than a high school education. Figure 7 repeats the simulation exercises using the coefficients estimated for this latter group, with the upper panel showing the historical simulations and the lower panel the policy simulations. Unsurprisingly, the contribution of excess persistence for this group in the historical exercise is minimal, and the policy simulations exhibit little difference between the two scenarios.

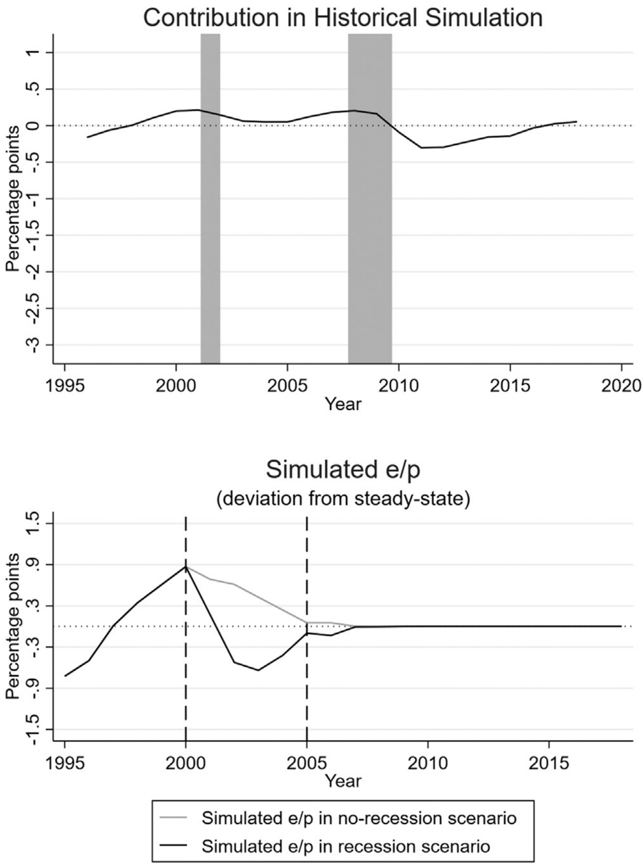
9. Conclusion

In this paper, we estimate a dynamic model on a panel of state-level data to quantify the persistence in the e/p of disadvantaged workers beyond that implied by the persistence of aggregate labor market conditions, which we call excess persistence in employment. We find that the e/p of less-educated prime-age males exhibits a moderate degree of excess persistence, which dissipates within three years. This finding is robust to a number of variations in sample and specification. Most notably, we find no indication of policy-relevant amounts of excess persistence for several definitions of disadvantaged populations that vary education levels, race, and age. In addition, we find no substantial asymmetry in the excess persistence of high

[30] Our historical simulation (Figure 5) suggests that the lasting cost of the 2008–09 recession would be larger. But even there the employment effects of excess persistence faded in just one year after UR returned to its natural rate in 2017.

[31] Replacing the baseline equation with an asymmetric specification from Section 7 does not qualitatively alter these conclusions.

Figure 7. Simulations for Group with Least Excess Persistence



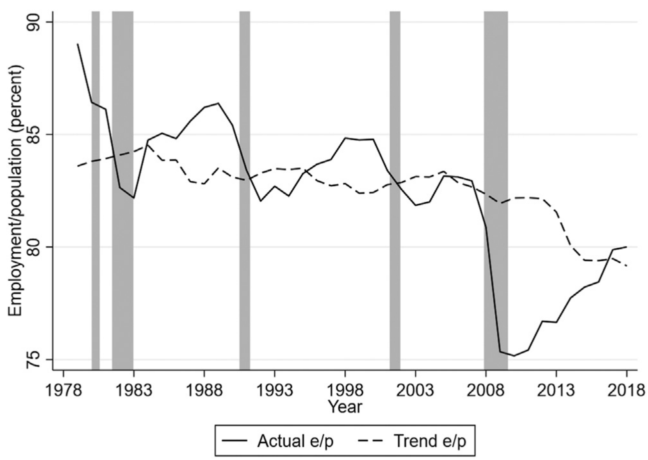
Note: For the demographic group with the smallest estimated amount of excess persistence (prime-age Black or Hispanic men with no more than a high school education), the contribution of that persistence in either historical or policy simulations is small. The time between the two vertical lines denotes the period over which the aggregate UR is assumed to be different between the two scenarios. The horizontal dotted line each panel denotes zero.

versus low employment rates. The cumulative effect of excess persistence in the business cycle surrounding the 2001 recession was mildly positive for our baseline group, while the effect in the cycle surrounding the 2008–09 recession was decidedly negative. Our simulations suggest that, despite large contemporaneous benefits, the lasting benefits to the employment rates of disadvantaged workers of temporarily running a "high-pressure" economy are small.

Appendix A. Summary Statistics for Disjoint Samples

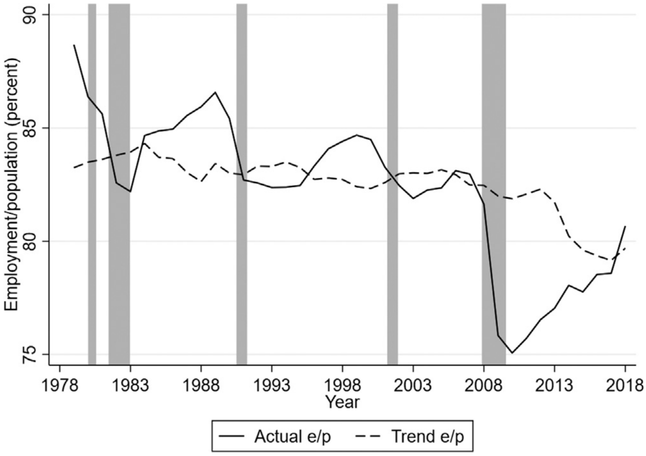
Figure 1 in the main text shows the full-sample estimates of the e/p and the trend e/p for our baseline sample of disadvantaged workers, in which both series have been aggregated from the state to the national level. These full-sample estimates provide the best estimates of the actual e/p and trend e/p in the economy. However, as described in Section 2.4, for the estimation we use disjoint samples. Figures A.1 and A.2 replicate Figure 1 for those disjoint samples. Tables A.1 and A.2 do the same for Table 1. On average the disjoint samples look much like the full sample, although there is more variability across states and years.

Figure A.1. Actual e/p and Trend e/p of Disadvantaged Group, Aggregated, LHS Sample



Note: State-level actual and trend e/p for prime-age men with no more than a high school education, aggregated to the national level. The trend e/p is calculated separately for each state using the method in Hamilton (2018). We use disjoint samples of the LHS and RHS e/p in the main analysis (Section 2.4). This presents the e/p ratios for the sample used in the LHS e/p. See Figure 1 for the e/p ratios using the full sample.

Figure A.2. Actual e/p and Trend e/p of Disadvantaged Group, Aggregated, RHS Sample



Note: State-level actual and trend e/p for prime-age men with no more than a high school education, aggregated to the national level. The trend e/p is calculated separately for each state using the method in Hamilton (2018). We use disjoint samples of the LHS and RHS e/p in the main analysis (Section 2.4). This presents the e/p ratios for the sample used in the RHS e/p. See Figure 1 for the e/p ratios using the full sample.

Table A.1. Summary Statistics for Baseline Group, State-Level Data (LHS sample)

	Mean	Std. Dev.	Min.	Max.
$e/p_{s,t}$, Actual (%)	82.4	5.2	61.4	95.6
$e/p_{s,t}$, Detrended (pp)	–0.3	3.6	–14.7	10.6

Note: Summary statistics for baseline samples for the years 1978 to 2018. "e/p_{st}, Actual" is the e/p of prime-age men with no more than a high school education in state s at time t. "e/p_{st}, Detrended" is "e/p_{st}, Actual" less the estimated trend for each state and is measured in percentage points (pp). The trend e/p is calculated using the method in Hamilton (2018). We use disjoint samples for the LHS and RHS e/p in the main analysis (Section 2.4). This table presents the summary statistics for the sample used in the LHS e/p. See Table 1 for the summary statistics using the full sample.

Table A.2. Summary Statistics for Baseline Group, State-Level Data (RHS Sample)

	Mean	Std. Dev.	Min.	Max.
e/p$_{s,t}$, Actual (%)	82.3	5.2	59.2	95.6
e/p$_{s,t}$, Detrended (pp)	–0.3	3.6	–16.5	10.3

Note: Summary statistics for baseline samples for the years 1978 to 2018. "e/p$_{st}$, Actual" is the e/p of prime-age men with no more than a high school education in state s at time t. "e/p$_{st}$, Detrended" is "e/p$_{st}$, Actual" less the estimated trend for each state and is measured in percentage points (pp). The trend e/p is calculated using the method in Hamilton (2018). We use disjoint samples for the LHS and RHS e/p in the main analysis (Section 2.4). This presents the summary statistics for the sample used in the RHS e/p. See Table 1 for the summary statistics using the full sample.

Appendix B. Our Specification and Blanchard and Katz (1992)

B.1 Motivating Our Baseline Equation

Our estimating Equation (1) can be thought of as the aggregated version of an individual-level equation. Ignoring some lags and the state subscripts for ease of exposition, the individual-level equation is

$$(e/p)_{i,t} = \alpha_i + \gamma_t + \phi(e/p)_{i,t-1} + \lambda \sum_{j \neq i}(e/p)_{j,t-1} + \delta Ugap_t,$$

in which ϕ represents sources of persistence such as human capital accumulation and depreciation, and λ represents cross-individual effects of the sort we discussed in Section 1.

Summing across i,

$$\begin{aligned}\sum_i (e/p)_{i,t} &= \sum_i \alpha_i + N\gamma_t + \phi \sum_i (e/p)_{i,t-1} \\ &\quad + \lambda \sum_i \sum_{j \neq i}(e/p)_{j,t-1} + N\delta Ugap_t \\ &= \sum_i \alpha_i + N\gamma_t + \phi \sum_i (e/p)_{i,t-1} \\ &\quad + \lambda \sum_i \left[\sum_j (e/p)_{j,t-1} - (e/p)_{i,t-1}\right] + N\delta Ugap_t.\end{aligned}$$

Denote $(e/p)_t$ as the mean of $(e/p)_{i,t}$ across i to obtain

$$N(e/p)_t = \sum_i \alpha_i + N\gamma_t + N\phi(e/p)_{t-1} + \lambda \sum_i [N(e/p)_{t-1} - (e/p)_{i,t-1}] + N\delta U gap_t,$$

and divide through by N to get

$$(e/p)_t = \frac{1}{N}\sum_i \alpha_i + \gamma_t + \phi(e/p)_{t-1} + \lambda[N(e/p)_{t-1} - (e/p)_{t-1}] + \delta U gap_t$$

or

$$\begin{aligned}(e/p)_t &= \frac{1}{N}\sum_i \alpha_i + \gamma_t + [\phi + \lambda(N-1)](e/p)_{t-1} + \delta U gap_t \\ &= \alpha + \gamma_t + \beta(e/p)_{t-1} + \delta U gap_t,\end{aligned} \tag{2}$$

in which $\beta = \phi + \lambda(N-1)$. Equation (2) is our estimating equation, in which β is the object of primary interest.

B.2 Relation to Blanchard and Katz (1992)

Our analysis is similar to Blanchard and Katz (1992)—henceforth B&K—and Dao, Furceri, and Loungani (2017). Both those studies and ours use state-level labor market data in a VAR-type framework. In particular, B&K estimate a VAR in three state-level variables: the change in employment, the employment-to-labor-force rate (that is, one minus the UR), and the labor force participation rate (LFPR), and identify innovations in employment with shocks to labor demand.[32]

However, we and B&K address different questions. B&K are interested in how a state's labor market adjusts to unexpected

[32] B&K use defense spending and predicted growth rates of employment using state industry shares and national growth rates as two observable and plausibly exogenous demand shocks.

changes in labor demand that cause its overall labor market conditions (employment in particular) to differ from that of other states, while our study focuses on the persistence of employment among disadvantaged workers in excess of that implied by overall labor market conditions.

This difference between the research questions leads to three important distinctions between our setup and B&K's: First, given their focus on aggregate adjustment mechanisms, B&K estimate equations in the change in employment. Because we do not emphasize adjustment, the change in employment does not enter into our system. Rather, the dependent variable in Equation (1) is the level of the detrended e/p.[33] Second, we examine employment (e/p) of the disadvantaged group, rather than employment of the overall population. This allows us to examine the persistence of employment in this group in excess of the persistence in overall labor market conditions.[34] Third, since our focus is on the possible lasting effects of past employment of the disadvantaged group on their current employment *conditional* on overall labor market conditions, we take overall labor market conditions as given. We neither model them in a separate equation nor attempt to identify unexpected changes in those conditions.

Appendix C. Robustness Appendix

C.1 Spectral Analysis of e/p Filters

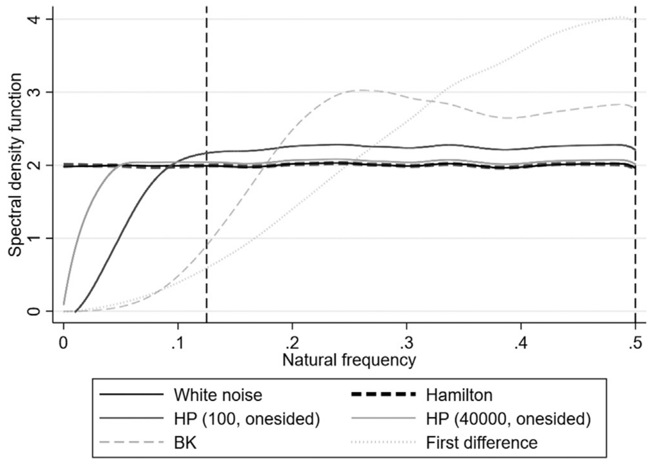
The filters in Section 5.1 amplify various frequencies. Because our qualitative results about excess persistence of e/p are robust to the choice of filter (Table 3), they are not driven by the particular frequencies passed by the filter we chose to feature.

To assess the gain of each filter, we first simulate standard normal white noise for 100,000 periods and compute its spectral density function and then pass this white noise through each of our filters

[33]We do not separately address the LFPR, which is consistent with B&K (footnote 35).

[34]Mechanically, focusing on the e/p of all workers in the B&K setup would mean we would be interested in the coefficient of the lagged e/p on the RHS, but would also include the employment-to-labor-force rate and the LFPR, which imply e/p.

Figure C.1. Spectral Density Function of Different Filters



Note: The filters we use pass through different frequencies of a standard normal white-noise series. The Hamilton approach recovers the spectral density function of the original series, and we use this approach as our baseline for detrending the state e/p of disadvantaged workers. The vertical dashed lines represent the 1/8 and 1/2 frequencies (eight- and two-year periodicities, respectively). See Appendix C.1 for more details.

(in turn) and compute the spectral density of the resulting time series.[35] The results are plotted in Figure C.1. The horizontal solid black line depicts the spectral density function of the original series, which is white noise and represents all the frequencies equally. The vertical dashed lines represent the 1/8 and 1/2 frequencies (eight- and two-year periodicities, respectively). The other lines represent the estimated gain from the various filters. We are reassured that our estimation method recovers the gains accurately for those cases in which the gains are well known (e.g., the first-difference and Baxter-King filters).

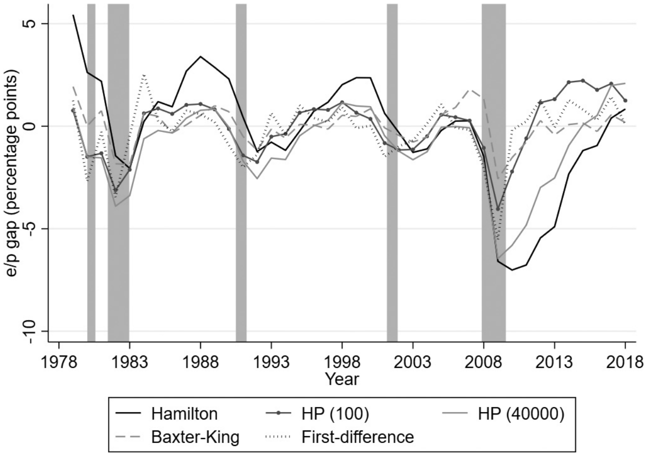
[35]Computing the spectral density function requires one to take a stand on a particular approach and some parameter values. Our approach was to compute the sample cumulative spectral-distribution function, and then compute the sample probability density function (pdf) using small differences. To smooth through the simulation error, we present the lowest-smoothed pdf using a small bandwidth (0.1).

The filters we use pass through different frequencies. The cyclical component from the Hamilton filter with a five-year horizon parameter recovers the spectral density function of the original series. As such, this filter does not amplify or mute any frequencies from the original series. The remaining filters remove lower-frequency components of the time series and pass through higher-frequency components, to a greater or lesser extent. The first-difference filter, as is well known, mostly passes through very high frequencies: it downplays frequencies below 0.25 (periodicities below $1/0.25 = 4$ periods) relative to higher frequencies, with the most weight on a natural frequency of 1/2. The remaining filters have qualitatively similar gain functions. Per design, the Baxter-King filter with a period of two to eight years and a three-year filter window puts less weight on frequencies below 0.18 (periodicities $1/0.18 = 5.6$ periods) and more weight on higher frequencies. The gain function of the HP filter depends on the smoothing parameter, with a higher smoothing parameter placing more weight on lower frequencies. As the HP smoothing parameter rises, the HP-filtered series approaches that of an ideal filter, whose gain would rise from 0 to its maximum level over an infinitesimally small frequency window at 0 frequency.

C.2 e/p Trends for Disjoint Samples

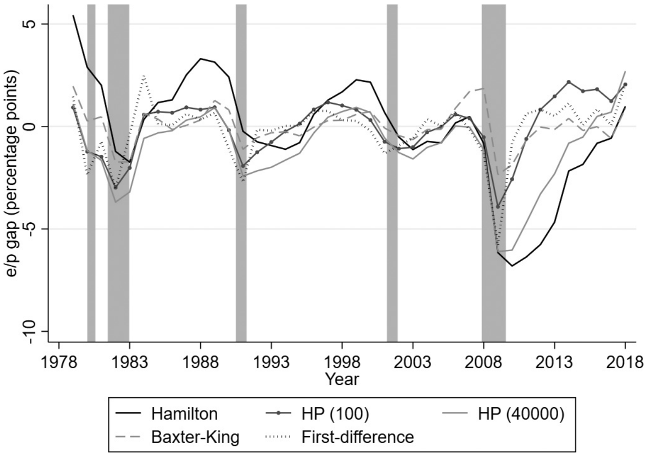
Figure 3 in the main text shows the full-sample estimates of the cyclical component of e/p using various filters, aggregated from the state to the national level. However, as described in Section 2.4, for the estimation we use disjoint samples. Figures C.2 and C.3 replicate Figure 3 for those disjoint samples. The disjoint samples look much like the full sample.

Figure C.2. Estimates of State Detrended e/p, Aggregated, LHS Sample



Note: Trends estimated by five different approaches. Aggregated to the national level. See Appendix C.2 for details.

Figure C.3. Estimates of State Detrended e/p, Aggregated, RHS Sample



Note: Trends estimated by five different approaches. Aggregated to the national level. See Appendix C.2 for details.

References

Aaronson, S., T. Cajner, B. Fallick, F. Galbis-Reig, C. Smith, and W. Wascher. 2014. "Labor Force Participation: Recent Developments and Future Prospects." *Brookings Papers on Economic Activity* (2): 197–275. doi:10.1353/eca.2014.0015.

Aaronson, S. R., M. C. Daly, W. Wascher, and D. W. Wilcox. 2019. "Okun Revisited: Who Benefits Most from a Strong Economy?" *Brookings Papers on Economic Activity* (1): 333–404.

Abraham, K. G., and M. S. Kearney. 2018. "Explaining the Decline in the U.S. Employment-to-Population Ratio: A Review of the Evidence." NBER Working Paper No. 24333. doi:10.3386/w24333.

Acemoglu, D., and D. Autor. 2011. "Skills, Tasks and Technologies: Implications for Employment and Earnings." In *Handbook of Labor Economics*, Vol. 4B, ed. D. Card and O. Ashenfelter, 1043–1171 (chapter 2). Elsevier. doi:10.1016/S0169-7218(11)02410-5.

Adao, R., C. Arkolakis, and F. Esposito. 2020. "General Equilibrium Effects in Space: Theory and Measurement." Manuscript.

Almlund, M., A. L. Duckworth, J. Heckman, and T. Kautz. 2011. "Personality Psychology and Economics." In *Handbook of the Economics of Education,* Vol. 4, ed. E. A. Hanushek, S. Machin, and L. Woessmann, 1–181 (chapter 1). Elsevier. doi:10.3386/w16822.

Arellano, M. 2003. *Panel Data Econometrics.* Oxford: Oxford University Press. doi:10.1093/0199245282.001.0001.

Ball, L. M. 2009. "Hysteresis in Unemployment: Old and New Evidence." In *Understanding Inflation and the Implications for Monetary Policy: A Phillips Curve Retrospective,* ed. J. Fuhrer et al. Federal Reserve Bank of Boston. doi:10.7551/mitpress/9780262013635.003.0015.

———. 2015. "Monetary Policy for a High-Pressure Economy." Center on Budget and Policy Priorities, March 30. https://www.cbpp.org/sites/default/files/atoms/files/3_30_15fe_ball.pdf.

Beaudry, P., D. Galizia, and F. Portier. 2015. "Reviving the Limit Cycle View of Macroeconomic Fluctuations." NBER Working Paper No. 21241. doi:10.3386/w21241.

———. 2016. "Putting the Cycle Back into Business Cycle Analysis." NBER Working Paper No. 22825. doi:10.3386/w22825.

Beraja, M., E. Hurst, and J. Ospina. 2019. "The Aggregate Implications of Regional Business Cycles." *Econometrica* 87 (6): 1789–1833. doi:10.3982/ECTA14243.

Blanchard, O., and J. Galí. 2010. "Labor Markets and Monetary Policy: A New Keynesian Model with Unemployment." *American Economic Journal: Macroeconomics* 2 (2): 1–30. doi:10.1257/mac.2.2.1.

Blanchard, O. J., and L. F. Katz. 1992. "Regional Evolutions." *Brookings Papers on Economic Activity* (1): 1–75. doi:10.2307/2534556.

Blanchard, O. J., and L. H. Summers. 1986. "Hysteresis and the European Unemployment Problem." In *NBER Macroeconomics Annual*, Vol. 1, ed. S. Fischer, 15–90. MIT Press. doi:10.1086/654013.

Boisjoly, J., G. J. Duncan, M. Kremer, D. M. Levy, and J. Eccles. 2006. "Empathy or Antipathy? The Impact of Diversity." *American Economic Review* 96 (5): 1890–1905. doi:10.1257/aer.96.5.1890.

Bostic, R. 2018. "Considerations on the Path from Extraordinary to Neutral." Speech, Global Interdependence Center, Madrid, Spain, November 15.

Bound, J., and H. J. Holzer. 2000. "Demand Shifts, Population Adjustments, and Labor Market Outcomes during the 1980s." *Journal of Labor Economics* 18 (1): 20–54. doi:10.1086/209949.

Charles, K. K., E. Hurst, and M. Schwartz. 2019. "The Transformation of Manufacturing and the Decline in U.S. Employment." *NBER Macroeconomics Annual 2018*, Vol. 33, ed. M. Eichenbaum and J. A. Parker, 307–72 (chapter 5). Chicago: University of Chicago Press.

Chodorow-Reich, G. 2019. "Geographic Cross-Section Fiscal Spending Multipliers: What Have We Learned?" *American Economic Journal: Economic Policy* 11 (2): 1–34. doi:10.1257/pol.20160465.

Cingano, F., and A. Rosolia. 2012. "People I Know: Job Search and Social Networks." *Journal of Labor Economics* 30 (2): 291–332. doi:10.1086/663357.

Clark, T., and S. Kozicki. 2005. "Estimating Equilibrium Real Interest Rates in Real Time." *North American Journal of Economics and Finance* 16 (3): 395–413. doi:10.1016/j.najef.2005.04.002.

Conley, T. G., and G. Topa. 2002. "Socio-Economic Distance and Spatial Patterns in Unemployment." *Journal of Applied Econometrics* 17 (4): 303–27. doi:10.1002/jae.670.

Council of Economic Advisers. 2017. *Economic Report of the President.* Washington, DC: Government Printing Office. https://www.govinfo.gov/content/pkg/ERP-2017/pdf/ERP-2017.pdf.

Cragg, J. G., and S. G. Donald. 1993. "Testing Identifiability and Specification in Instrumental Variable Models." *Econometric Theory* 9 (2): 222–40. http://www.jstor.org/stable/3532477.

Crone, T. M. 2005. "An Alternative Definition of Economic Regions in the United States Based on Similarities in State Business Cycles." *Review of Economics and Statistics* 87 (4): 617–26. doi:10.1162/003465305775098224.

Dao, M., D. Furceri, and P. Loungani. 2017. "Regional Labor Market Adjustments in the United States: Trend and Cycle." *Review of Economics and Statistics* 99 (2): 243–57. doi:10.1162/REST_a_00642.

Devereux, P. J. 2002. "Occupational Upgrading and the Business Cycle." *LABOUR* 16 (3): 423–52. doi:10.1111/1467-9914.00202.

Driscoll, J., and A. Kraay. 1998. "Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data." *Review of Economics and Statistics* 80 (4): 549–60. doi:10.1162/003465398557825.

Dustmann, C., A. Glitz, U. Schönberg, and H. Brücker. 2017. "Factors Determining Callbacks to Job Applications by the Unemployed: An Audit Study." *RSF: The Russell Sage Foundation Journal of the Social Sciences* 3 (3): 168–201. doi:10.7758/rsf.2017.3.3.08.

Elsby, M. W. L., B. Hobijn, and A. Şahin. 2015. "On the Importance of the Participation Margin for Labor Market Fluctuations." *Journal of Monetary Economics* 72 (May): 64–82. doi:10.1016/j.jmoneco.2015.01.004.

Fallick, B., and M. Tasci. 2020. "Estimating the Trend Unemployment Rate in the Fourth Federal Reserve District." Working Paper No. 20-19, Federal Reserve Bank of Cleveland. doi:10.26509/frbc-wp-202019.

Feldstein, M. 2018. "Normalizing Monetary Policy." *Cato Journal* 38 (2, Spring/Summer): 415–22. https://www.cato.org/cato-journal/spring/summer-2018/normalizing-monetary-policy.

Fleischman, C., and J. Gallin. 2001. "Employment Persistence." Finance and Economics Discussion Series No. 2001-25, Board of Governors of the Federal Reserve System.

Fleischman, C. A., J. Gallin, and C. L. Smith. 2018. "Labor Market Persistence: A Cohort-level Analysis." Manuscript.

Gautier, P., P. Muller, B. van der Klaauw, M. Rosholm, and M. Svarer. 2018. "Estimating Equilibrium Effects of Job Search Assistance." *Journal of Labor Economics* 36 (4): 1073–1125. doi:10.1086/697513.

Glitz, A. 2017. "Coworker Networks in the Labour Market." *Labour Economics* 44 (January): 218–30. doi:10.1016/j.labeco.2016.12.006.

Hamilton, J. D. 2018. "Why You Should Never Use the Hodrick-Prescott Filter." *Review of Economics and Statistics* 100 (5): 831–43. doi:10.1162/rest_a_00706.

Hershbein, B., and B. A. Stuart. 2020. "Recessions and Local Labor Market Hysteresis." Working Paper No. 20-325, Upjohn Institute. doi:10.17848/wp20-325.

Holzer, H. J., S. Raphael, and M. A. Stoll. 2006. "Employers in the Boom: How Did Hiring of Less-Skilled Workers Change During the 1990s?" *Review of Economics and Statistics* 88 (1): 283–99. doi:10.1162/rest.88.2.283.

Hotchkiss, J. L., and R. E. Moore. 2018. "Some Like It Hot: Assessing Longer-Term Labor Market Benefits from a High-Pressure Economy." Working Paper No. 2018-1, Federal Reserve Bank of Atlanta.

Hoynes, H., D. L. Miller, and J. Schaller. 2012. "Who Suffers During Recessions?" *Journal of Economic Perspectives* 26 (3): 27–48. doi:10.1257/jep.26.3.27.

Jackson, M., and P. Tebaldi. 2019. "A Forest Fire Theory of the Duration of a Boom and the Size of a Subsequent Bust." doi:10.2139/ssrn.2263501.

Juhn, C. 1992. "Decline of Male Labor Market Participation: The Role of Declining Market Opportunities." *Quarterly Journal of Economics* 107 (1): 79–121. doi:10.2307/2118324.

Kaiser, R., and A. Maravall. 1999. "Estimation of the Business Cycle: A Modified Hodrick-Prescott Filter." *Spanish Economic Review* 1 (2): 175–206. doi:10.1007/s101080050008.

Kallenberg, A. L., and T. von Wachter. 2017. "The U.S. Labor Market During and After the Great Recession: Continuities and Transformations." *RSF: The Russell Sage Foundation Journal of the Social Sciences* 3 (3): 1–19. doi:10.7758/rsf.2017.3.3.01.

KC LMCI. 2021. "Labor Market Conditions Indicators." https://www.kansascityfed.org/data-and-trends/labor-market-conditions-indicators/. Accessed: April 8, 2021.

Kiley, M. T. 2018. "Unemployment Risk." Finance and Economics Discussion Series No. 2018-067, Board of Governors of the Federal Reserve System. doi:10.17016/FEDS.2018.067.

Lacker, J. M. 2017. "Economic Outlook." Speech, Maryland Bankers Association, Baltimore, MD, January 6.

Miller, C. 2017. "The Persistent Effect of Temporary Affirmative Action." *American Economic Journal: Applied Economics* 9 (3): 152–90. doi:10.1257/app.20160121.

Mise, E., T.-H. Kim, and P. Newbold. 2005. "On Suboptimality of the Hodrick–Prescott Filter at Time Series Endpoints." *Journal of Macroeconomics* 27 (1): 53–67. doi:10.1016/j.jmacro.2003.09.003.

Nakamura, E., and J. Steinsson. 2014. "Fiscal Stimulus in a Monetary Union: Evidence from US Regions." *American Economic Review* 104 (3): 753–92. doi:10.1257/aer.104.3.753.

———. 2018. "Identification in Macroeconomics." *Journal of Economic Perspectives* 32 (3): 59–86. doi:10.1257/jep.32.3.59.

National Bureau of Economic Research. 2019. "Current Population Survey (CPS) Basic Monthly Data." http://data.nber.org/data/cps_basic.html. Accessed: June 28, 2019.

Nerlove, M. 1967. "Experimental Evidence on the Estimation of Dynamic Economic Relations from a Times Series of Cross-Sections." *Economic Studies Quarterly* 18 (3): 42–74. doi:10.11398/economics1950.18.3_42.

Neumark, D., I. Burn, and P. Button. 2019. "Is It Harder for Older Workers to Find Jobs? New and Improved Evidence from a Field Experiment." *Journal of Political Economy* 127 (2): 922–70. doi:10.1086/701029.

Nickell, S. 1981. "Biases in Dynamic Models with Fixed Effects." *Econometrica* 49 (6): 1417–26. doi:10.2307/1911408.

Notowidigdo, M. J. 2020. “The Incidence of Local Labor Demand Shocks.” *Journal of Labor Economics* 38 (3): 687–725. doi:10.1086/706048.

Okun, A. M. 1973. “Upward Mobility in a High-Pressure Economy.” *Brookings Papers on Economic Activity* (1): 207–61. doi:10.2307/2534087.

Owyang, M. T., J. Piger, and H. J. Wall. 2005. “Business Cycle Phases in U.S. States.” *Review of Economics and Statistics* 87 (4): 604–16. doi:10.1162/003465305775098198.

Pierce, J. R., and P. K. Schott. 2016. “The Surprisingly Swift Decline of US Manufacturing Employment.” *American Economic Review* 106 (7): 1632–62. doi:10.1257/aer.20131578.

Poterba, J. M., and L. H. Summers. 1986. “Reporting Errors and Labor Market Dynamics.” *Econometrica* 54 (6): 1319–38. doi:10.2307/1914301.

Reifschneider, D., W. Wascher, and D. Wilcox. 2015. “Aggregate Supply in the United States: Recent Developments and Implications for the Conduct of Monetary Policy.” *IMF Economic Review* 63 (1): 71–109. doi:10.1057/imfer.2015.1.

Schmieder, J. F., and T. von Wachter. 2010. “Does Wage Persistence Matter for Employment Fluctuations? Evidence from Displaced Workers.” *American Economic Journal: Applied Economics* 2 (3): 1–21. doi:10.1257/app.2.3.1.

Solon, G., S. J. Haider, and J. M. Wooldridge. 2015. “What Are We Weighting For?” *Journal of Human Resources* 2 (50): 301–16. doi:10.3368/jhr.50.2.301.

Song, J., and T. von Wachter. 2014. “Long-Term Nonemployment and Job Displacement.” Proceedings from the Federal Reserve Bank of Kansas City’s Economic Symposium, Jackson Hole, WY (August).

Stock, J., and M. Watson. 1999a. “Business Cycle Fluctuations in U.S. Macroeconomic Time Series.” In *Handbook of Macroeconomics,* Vol. 1, ed. J. B. Taylor and M. Woodford, 1–64 (chapter 1). Elsevier. doi:10.1016/s1574-0048(99)01004-6.

———. 1999b. “Forecasting Inflation.” *Journal of Monetary Economics* 44 (2): 293–335. doi:https://doi.org/10.1016/S0304-3932(99)00027-6.

Stock, J., and M. Yogo. 2005. "Testing for Weak Instruments in Linear IV Regression." In *Identification and Inference for Econometric Models,* ed. D. W. K. Andrews, 80–108. New York: Cambridge University Press. doi:10.1017/CBO9780511614491.006.

Stockhammer, E., and S. Sturn. 2012. "The Impact of Monetary Policy on Unemployment Hysteresis." *Applied Economics* 44 (21): 2743–56. doi:10.1080/00036846.2011.566199.

Stoops, N. 2004. "Educational Attainment in the United States: 2003." https://www.census.gov/content/dam/Census/library/publications/2004/demo/p20-550.pdf. Accessed: June 28, 2019.

Tiller, R. B. 1992. "Time Series Modeling of Sample Survey Data from the Current Population Survey." *Journal of Official Statistics* 8 (2): 149–66.

U.S. Bureau of Economic Analysis. 2019. "GDP by State." https://www.bea.gov/data/gdp/gdp-state. Accessed: August 14, 2019.

U.S. Department of Labor. 2018. "State Unemployment Insurance Benefits." https://oui.doleta.gov/unemploy/uifactsheet.asp. Accessed: June 28, 2019.

von Wachter, T. 2020. "The Persistent Effects of Initial Labor Market Conditions for Young Adults and Their Sources." *Journal of Economic Perspectives* 34 (4): 168–94. doi:10.1257/jep.34.4.168.

Yagan, D. 2019. "Employment Hysteresis from the Great Recession." *Journal of Political Economy* 127 (5): 2505–58. doi:10.1086/701809.

Yellen, J. L. 2016. "Macroeconomic Research After the Crisis." Speech, Federal Reserve Bank of Boston, Boston, MA, October 14.

———. 2019. Remarks delivered at the Hutchins Center Conference on Inflation, The Brookings Institution, Washington, DC, October 3.

Monetary Policy, Inflation Target, and the Great Moderation: An Empirical Investigation*

Qazi Haque
The University of Adelaide
Centre for Applied Macroeconomic Analysis

This paper estimates a New Keynesian model with trend inflation and contrasts Taylor rules featuring fixed versus time-varying inflation target. The estimation is conducted over the Great Inflation and the Great Moderation periods, while allowing for indeterminacy. Time-varying inflation target empirically fits better and active monetary policy prevails in both periods, thereby ruling out sunspots as an explanation of the Great Inflation episode.

JEL Codes: C11, C52, C62, E31, E32, E52.

1. Introduction

The post–World War II U.S. economy includes two particular eras: the Great Inflation and the Great Moderation. The former era is represented by highly volatile inflation and output growth, while there has been a marked decline in macroeconomic volatility in the latter

*I would like to thank the Editor, Boragan Aruoba, and two anonymous referees for their suggestions which improved the quality of the paper. I also thank Guido Ascari, Marco Bassetto, Efrem Castelnuovo, Firmin Doko Tchatoka, Mardi Dungey, Yunjong Eo, Ippei Fujiwara, Nicolas Groshenny, Yasuo Hirose, Punnoose Jacob, Mariano Kulish, Leandro Magnusson, Sophocles Mavroeidis, James Morley, Adrian Pagan, Bruce Preston, Barbara Rossi, Mark Weder, Benjamin Wong, Jacob Wong, and seminar participants at The University of Adelaide, Modelling Macroeconomic Shocks Workshop (University of Tasmania), Continuing Education in Macroeconometrics Workshop (The University of Sydney), Banque de France, RBA, and RBNZ. This work was supported with supercomputing resources provided by the Phoenix HPC service at The University of Adelaide. I acknowledge generous support from the Australian Research Council, under the grants DP170100697 and DP200101498, as well as the University of Oxford for their hospitality. Author contact: School of Economics, The University of Adelaide, 10 Pulteney St., Adelaide SA 5005, Australia; e-mail: qazi.haque@adelaide.edu.au.

period (McConnell and Perez-Quiros 2000; Blanchard and Simon 2001; and Stock and Watson 2003). But what has led to the transition from the Great Inflation to the Great Moderation era? The two main hypotheses put forth by the empirical literature are either "good luck" or "good policy." The "good luck" interpretation—a decline in the variance of exogenous shocks hitting the economy—has been supported by a number of authors, including Stock and Watson (2003), Primiceri (2005), Sims and Zha (2006), Smets and Wouters (2007), and Justiniano and Primiceri (2008). Within the "good policy framework, the monetary policy literature has offered at least two competing explanations regarding this shift to macroeconomic stability—a stronger policy response to inflation (Clarida, Galí, and Gertler 2000; Lubik and Schorfheide 2004) and an enhanced stability of the Federal Reserve's inflation target (Ireland 2007; Cogley, Primiceri and Sargent 2010).

The empirical plausibility of a link between monetary policy and macroeconomic instability in the 1970s was established by Clarida, Galí, and Gertler (2000) and further advocated by Lubik and Schorfheide (2004), who argue that U.S. monetary policy in the 1970s failed to respond sufficiently strongly to inflation, thereby generating indeterminacy.[1] Consequently, self-fulfilling inflation expectations are regarded as the driver of the high-inflation episode in the 1970s. According to this view, the switch from a passive to an active response to inflation has brought about a stable and determinate environment since the early 1980s.[2] In a conceptually related study, Boivin and Giannoni (2006) find that this switch has also been instrumental in reducing observed output and inflation volatility. Moreover, Benati and Surico (2008) show that by responding more strongly to inflation, monetary policy has contributed to the

[1]Here *indeterminacy* refers to the multiplicity of rational expectations equilibria, such that there are infinite number of paths toward a unique steady state, and that the economy can be unexpectedly volatile due to self-fulfilling beliefs or sunspot shocks. In contrast, an equilibrium that is locally isolated and uniquely determined by preferences and technologies is called *determinate*. See Farmer (1999) for a formal definition.

[2]A policy response to inflation is called *active* if it satisfies the Taylor principle—an aspect of the Taylor rule that describes how, for each 1 percent increase in inflation, the central bank should raise the nominal interest rate by more than 1 percentage point to ensure determinacy. Otherwise, it is labeled as *passive*.

decline in persistence and predictability of inflation around the time of the Volcker disinflation.

While existing studies focusing on indeterminacy only consider a fixed inflation target, a large part of the literature finds that time-varying target is empirically important in capturing the low-frequency movements of the inflation rate.[3] For example, Cogley and Sbordone (2008) structurally decompose inflation dynamics into a time-varying long-run component (i.e., trend inflation) and a short-run one (i.e., the inflation gap, given by the difference between inflation and trend inflation). Their main finding is that time-varying trend inflation captures the low-frequency variation in inflation dynamics, while the short-run inflation gap fits well into a purely forward-looking equation—the New Keynesian Phillips curve (NKPC)—without the need of any ad hoc intrinsic inertia. Moreover, Cogley, Primiceri, and Sargent (2010) argue that the decline in the variability of the Federal Reserve's inflation target is the single most important factor behind the reduction in inflation gap volatility and persistence during the Great Moderation. These findings square well with Milton Friedman's dictum that "Inflation is always and everywhere a monetary phenomenon." (Friedman 1968, p. 39). In fact, Ireland (2007) argues that Friedman's "always and everywhere" dictum strongly suggests that persistent movements in inflation, as observed in the data, could not have taken place without ongoing shifts in the Federal Reserve's inflation target. However, the Federal Reserve only explicitly revealed its inflation target in 2012 and did not have an explicit target until then. Hence, one must rely on a statistical or an econometric model to extract information about the Federal Reserve's inflation target from data on observed variables.

Empirical investigations conducted so far have either looked at the plausibility of a switch from indeterminacy to determinacy through the lens of a model featuring fixed target, or allowed for time-varying inflation target, while restricting the model to determinacy alone.[4] Unfortunately, the assumption of a fixed versus time-varying inflation target is not innocuous for both (in)determinacy

[3]See Cogley and Sargent (2005a); Kozicki and Tinsley (2005, 2009); Ireland (2007); Cogley and Sbordone (2008); Cogley, Primiceri, and Sargent (2010); Justiniano, Primiceri, and Tambalotti (2013), to name just a few.

[4]One exception is Coibion and Gorodnichenko (2011), which we discuss below.

and the role of monetary policy in the Great Moderation. For instance, the parameter estimate of the Taylor rule's response to the inflation gap depends on whether the Federal Reserve responds to deviations from a fixed target or time-varying target. This feature then affects the probability of being in a determinate or indeterminate regime. This paper evaluates the competing "good policy" views on the U.S. economy's shift from the Great Inflation to the Great Moderation by estimating a New Keynesian model with positive trend inflation, while also allowing for both indeterminacy and time-varying inflation target.[5] Notwithstanding, the paper does distinguish between trend inflation and time-varying inflation target, as in Aruoba and Schorfheide (2011). On one hand, trend inflation (a term coined by Ascari 2004) stands for a strictly positive level of steady-state inflation around which to approximate firms' first-order conditions in the derivation of the NKPC. Allowing for positive trend inflation is important, as it affects the determinacy properties of the model. Ascari and Ropele (2007, 2009) show that trend inflation makes price-setting firms more forward looking, which flattens the NKPC and widens the indeterminacy region. A fixed inflation target is simply equal to trend inflation in the model. On the other hand, following Sargent (1999), Cogley and Sargent (2005b), Primiceri (2006), and Sargent, Williams and Zha (2006), time-varying inflation target can be interpreted as the short-term goal pursued by the Federal Reserve owing to its changing beliefs about the inflation-output trade-off. Another interpretation is that the Federal Reserve opportunistically transformed supply shocks into persistent inflation changes in order to limit output losses in the 1970s, when shocks were mainly adverse. In contrast, the Federal Reserve acted in order to bring down inflation in the 1980s and 1990s, when shocks were mainly favorable (see Ireland 2007 and references cited therein). Along these lines, time-varying inflation target is assumed to follow a persistent exogenous process as in Ireland (2007) and Cogley, Primiceri, and Sargent (2010), but one whose unconditional mean

[5] Ascari, Bonomolo, and Lopes (2019) also allow for temporarily unstable paths in a simple New Keynesian model with fixed zero inflation target, while this paper requires all solutions to be stable, in line with previous contributions in the literature.

is equal to the steady state or trend inflation. Hence, we formalize neither the decisional process by the Federal Reserve to vary its target over time nor the learning process which possibly induced the evolution of such target. Nonetheless, the paper provides a quantitative assessment of the relevance of inflation target shocks hitting the low-frequency component of inflation, particularly for (in)determinacy.[6]

The estimation is conducted over two different periods covering the Great Inflation (1960:Q1–1979:Q2) and the Great Moderation (1984:Q1–2008:Q2). The paper finds that when considering the model with a fixed inflation target, indeterminacy cannot be ruled out before 1979 while determinacy prevails after 1984, which is in line with the existing empirical literature (Lubik and Schorfheide 2004; Hirose, Kurozumi, and Van Zandweghe 2020). Yet, this outcome differs when allowing for a time-varying inflation target. This time the posterior density favors determinacy for both the pre-1979 and post-1984 subsamples. This result suggests that monetary policy, even during the pre-Volcker period, was likely to be sufficiently active to ensure determinacy. Using the Bayes factor to compare the two specifications, the paper then reports evidence in favor of time variation in the inflation target process. What is driving the determinacy result? First of all, the inflation gap that enters the Taylor rule when the target is drifting over time is less volatile than the inflation gap with a fixed target. For a given historical path of the nominal interest rate, the response of the nominal rate to the inflation gap turns out to be higher in case of a time-varying target, which leads to determinacy. Moreover, as Cogley, Primiceri, and Sargent (2010) discuss, inflation target shocks induce persistent responses in the inflation gap, which helps to capture the highly persistent inflation dynamics in the 1970s. As such, the model does not require the richer endogenous inflation dynamics that arise under indeterminacy to explain the Great Inflation episode. Therefore, unlike the literature's preponderant view, this finding works against self-fulfilling inflation expectations (i.e., sunspots) as an explanation of the Great Inflation episode.

[6] For models in which inflation target evolves partly or fully endogenously, see Ireland (2007), Zanetti (2014), and Eo and Lie (2020).

The structure of the paper is as follows. Section 2 presents a brief overview of some closely related papers to provide further background and motivation. Section 3 sketches the model and its solution. Section 4 presents the econometric strategy, while Section 5 documents the estimation results. Robustness checks are performed in Section 6. Finally, Section 7 concludes.

2. Related Literature

Closely related to this paper are studies by Castelnuovo (2010), Cogley, Primiceri, and Sargent (2010), Aruoba and Schorfheide (2011), Coibion and Gorodnichenko (2011), Ettmeier and Kriwoluzky (2020), and Hirose, Kurozumi, and Van Zandweghe (2020), among others. Both Castelnuovo (2010) and Cogley, Primiceri, and Sargent (2010) estimate a New Keynesian model with time-varying inflation target using standard Bayesian Markov chain Monte Carlo (MCMC) techniques, while restricting the parameter space to determinacy, and perform counterfactual simulations to assess the drivers of the Great Moderation. This paper, on the other hand, estimates the model over the entire stable region of the parameter space using sequential Monte Carlo (SMC) techniques; that is, simultaneously estimating the model over both determinacy and indeterminacy regions. The paper also compares the fit of fixed versus time-varying target and shows that the latter specification fits better.

Coibion and Gorodnichenko (2011) use a single-equation approach to estimate a Taylor rule with time-varying coefficients using real-time data and extract a measure of time-varying trend inflation. A time series for the probability of determinacy is then constructed by feeding the empirical estimates of the Taylor rule into a New Keynesian model with firm-specific labor and trend inflation. This series indicates that the probability of determinacy was essentially zero in the second half of the 1970s. In contrast, this paper treats (in)determinacy as a property of a rational expectations system that requires a full information estimation approach, such that the parameter estimates of the Taylor rule account for the endogeneity of its targeted variables. Moreoever, Coibion and Gorodnichenko (2011) do not estimate the shock processes and so the effect on indeterminacy cannot be quantified as completely as in a fully specified and estimated dynamic stochastic general equilibrium (DSGE) model, as they point out.

Aruoba and Schorfheide (2011) develop and estimate a two-sector model comprising search-based monetary frictions and a standard New Keynesian economy with price rigidities. They study the steady-state welfare implications of the estimated model and suggest that distortions created by monetary frictions may be of similar magnitude as the distortions created by price stickiness in standard New Keynesian models. Although the focus of their paper is quite different, some of the features of their estimated model are quite similar to this study. In particular, both this paper and that of Aruoba and Schorfheide (2011) log-linearize the model around a non-zero steady-state inflation or trend inflation and assume an exogenous time-varying inflation target in the monetary policy rule.[7]

Hirose, Kurozumi, and Van Zandweghe (2020) estimate a New Keynesian model with firm-specific labor and a fixed inflation target (equal to trend inflation) using the same SMC methodology as in this paper. They find that the pre-Volcker period is characterized by indeterminacy, while better systematic monetary policy as well as changes in the level of trend inflation resulted in a switch to determinacy after 1982.[8] In contrast, this paper estimates a similar model with homogenous labor while also allowing for time variation in the inflation target process. The paper documents that a time-varying inflation target empirically fits better than a constant target and determinacy prevails in both sample periods.

In a recent paper, Ettmeier and Kriwoluzky (2020) also use the same SMC methodology and estimate a New Keynesian model with monetary and fiscal policy interactions. By estimating the model over the entire parameter space, Ettmeier and Kriwoluzky (2020) find that the pre-Volcker macroeconomic dynamics were driven by both a passive monetary/passive fiscal (indeterminate) regime and an active fiscal/passive monetary (determinate) regime. They show

[7] Aruoba and Schorfheide (2011) assume a random walk for the inflation target process, while this paper assumes a highly persistent but stationary process as in Cogley, Primiceri, and Sargent (2010).

[8] Arias et al. (2020) corroborate these findings by revisiting the relation between the sytematic component of monetary policy, trend inflation, and determinacy within a medium-scale DSGE model.

that due to fiscal dominance arising from the active fiscal/passive monetary regime, fiscal policy actions (in particular government spending) were critical in the inflation build-up in the 1970s. In contrast, this paper abstracts from fiscal policy actions and its interactions with monetary policy and instead focuses on evaluating monetary policy rules featuring fixed versus time-varying inflation target, while also analyzing its implications for (in)determinacy, through the lens of a generalized New Keynesian (GNK) model with non-zero steady-state inflation and no ad hoc backward-looking price indexation.[9] Ettmeier and Kriwoluzky's (2020) model, which is based on Bhattarai, Lee, and Park (2016), features indexation of non-reoptimized prices to both lagged inflation and steady-state inflation, thereby completely mitigating the effects of non-zero steady-state inflation on model dynamics (see Coibion and Gorodnichenko 2011 and Ascari and Sbordone 2014). However, Hirose, Kurozumi, and Van Zandweghe (2020) show that a GNK model with no price indexation fits the post-war U.S. data substantially better, which is also in line with Cogley and Sbordone's (2008) argument. An interesting avenue for future research would be to incorporate non-zero trend inflation without any backward-looking indexation into a model with monetary-fiscal interactions and empirically reexamine the drivers of the Great Inflation and subsequent Great Moderation.[10]

The finding that the pre-Volcker period could possibly be characterized by a unique equilibrium coincides with those of Orphanides (2004), Bilbiie and Straub (2013), and Haque, Groshenny, and Weder (2021). Orphanides (2004) finds an active response to expected inflation in a Taylor-type rule estimated for the pre-1979 period using real-time data, as opposed to ex post revised data, thereby claiming that self-fulfilling inflation expectations could not have been a source of macroeconomic instability during the Great Inflation. Bilbiie and Straub (2013) show that limited asset market participation results in an inverted IS curve and inverted aggregate demand logic; that

[9] Following Ascari and Sbordone (2014), we use the term GNK to refer to the New Keynesian model log-linearized around a positive inflation rate in the steady state.

[10] Ascari, Florio, and Gobbi (2018) study the long-run Taylor principle in a model with positive trend inflation and Markov-switching monetary and fiscal policies and find an important role for trend inflation.

is, interest rate increases become expansionary. Accordingly, they document passive monetary policy during the pre-Volcker period as being consistent with equilibrium determinacy. Haque, Groshenny, and Weder (2021) document that commodity price shocks, in an environment characterized by a high degree of real wage rigidities, generated a trade-off for the Federal Reserve in terms of stabilizing inflation and the output gap during the 1970s. Faced with this trade-off, they find that the Federal Reserve responded aggressively to inflation and negligibly to the output gap in the pre-Volcker period, such that its conduct did not lead to indeterminacy.

3. Model

The estimation is based on a version of Ascari and Sbordone's (2014) generalized New Keynesian (GNK) model. The model economy consists of an intertemporal Euler equation, obtained from the household's optimal choice of consumption and bond holdings, a discrete-time staggered price-setting model of Calvo (1983) that features a positive steady-state trend inflation, and a Taylor rule that characterizes monetary policy. As discussed earlier, allowing for a positive steady-state inflation is important for the following reasons: (i) positive trend inflation makes price-setting firms more forward looking, which flattens the NKPC and makes the inflation rate less sensitive to current economic conditions; (ii) it alters the determinacy properties of the model; and (iii) trend inflation generates richer endogenous persistence of inflation and output, even in the determinacy case. Unlike Ascari and Sbordone (2014), this paper assumes stochastic trend growth modeled as the technology level following a unit-root process; replaces their labor supply disturbance with a discount factor shock, which is a stand-in for a demand-type shock; and introduces (external) habit formation in consumption to generate output persistence. In light of the result of Cogley and Sbordone (2008) regarding the lack of empirical support for intrinsic inertia in the GNK Phillips curve (GNKPC), the model is estimated in the absence of rule-of-thumb price setting. Finally, the Taylor rule features responses to the inflation gap, the output gap, and output growth and also allows for interest rate smoothing.

3.1 The Log-Linearized Model

The log-linearized equilibrium conditions are given by the following equations.[11]

$$\begin{aligned}\widehat{y}_t = &\left(\frac{h}{g+h}\right)[\widehat{y}_{t-1} - \widehat{g}_t] + \left(\frac{g}{g+h}\right)[E_t\widehat{y}_{t+1} + E_t\widehat{g}_{t+1}]\\ &- \left(\frac{g-h}{g+h}\right)[\widehat{r}_t - E_t\widehat{\pi}_{t+1}]\\ &+ \left(\frac{g-h}{g+h}\right)\left[\widehat{d}_t - E_t\widehat{d}_{t+1}\right], \qquad (1)\end{aligned}$$

$$\begin{aligned}\widehat{\pi}_t = &\kappa E_t\widehat{\pi}_{t+1} + \vartheta\left[\varphi\widehat{s}_t + (1+\varphi)\widehat{y}_t\right]\\ &+ \chi\left(\frac{h}{g-h}\right)[\widehat{y}_t - \widehat{y}_{t-1} + \widehat{g}_t] - \varpi E_t\widehat{\Psi}_{t+1} + \varpi\widehat{d}_t, \qquad (2)\end{aligned}$$

$$\begin{aligned}\widehat{\Psi}_t = &(1-\xi\beta\pi^{\varepsilon})\left[\varphi\widehat{s}_t + (1+\varphi)\widehat{y}_t + \widehat{d}_t\right]\\ &+ \xi\beta\pi^{\varepsilon}\left[E_t\widehat{\Psi}_{t+1} + \varepsilon E_t\widehat{\pi}_{t+1}\right], \qquad (3)\end{aligned}$$

$$\widehat{s}_t = \varepsilon\xi\pi^{\varepsilon-1}\left(\frac{\pi-1}{1-\xi\pi^{\varepsilon-1}}\right)\widehat{\pi}_t + \xi\pi^{\varepsilon}\widehat{s}_{t-1}, \qquad (4)$$

$$\begin{aligned}\widehat{r}_t = &\rho_r\widehat{r}_{t-1} + (1-\rho_r)\{\psi_\pi(\widehat{\pi}_t - \widehat{\pi}^*_t)\\ &+ \psi_x\widehat{x}_t + \psi_{\Delta y}(\widehat{y}_t - \widehat{y}_{t-1} + \widehat{g}_t)\} + \epsilon_{r,t}, \qquad (5)\end{aligned}$$

$$\widehat{x}_t = \widehat{y}_t - \widehat{y}^n_t, \qquad (6)$$

$$\widehat{y}^n_t = \frac{h}{g(1+\varphi) - h\varphi}\left(\widehat{y}^n_{t-1} - \widehat{g}_t\right), \qquad (7)$$

where $\kappa \equiv \beta\left[1 + \varepsilon(\pi-1)(1-\xi\pi^{\varepsilon-1})\right]$, $\vartheta \equiv (1-\xi\pi^{\varepsilon-1})(1-\xi\beta\pi^{\varepsilon})/\xi\pi^{\varepsilon-1}$, $\chi \equiv (1-\xi\pi^{\varepsilon-1})(1-\xi\beta\pi^{\varepsilon-1})/\xi\pi^{\varepsilon-1}$, and $\varpi \equiv \beta(1-\pi)(1-\xi\pi^{\varepsilon-1})$. Hatted variables denote log-deviations from the steady

[11] A full description of the model is relegated to the online appendices, available at http://www.ijcb.org.

state. Here y_t and y_t^n stand for output and natural level of output, respectively; x_t is the output gap; r_t denotes the nominal interest rate; π_t denotes the inflation rate; π_t^* represents the Federal Reserve's time-varying inflation target; s_t denotes the resource cost due to relative price dispersion; Ψ_t is an endogenous auxiliary variable that appears in the Phillips curve in expectations, and thus drives inflation in response to expected changes in future demand and price dispersion; and E_t represents the expectations operator. Equation (1) is the dynamic IS curve, reflecting a Euler equation, where $h \in [0,1]$ represents the degree of habit persistence and g stands for the steady-state gross rate of technological progress, which is also equal to the steady-state balanced growth rate. Equations (2), (3), and (4) represent the GNK Phillips curve, where $\beta \in (0,1)$ is the subjective discount factor, $\xi \in (0,1)$ is the fraction of firms whose prices remain unchanged from previous period, π is the steady-state gross inflation rate or trend inflation, $\varepsilon > 1$ is the price elasticity of demand, and $\varphi \geq 0$ is the inverse elasticity of labor supply. Equation (4) is a recursive log-linearized expression for the price dispersion measure under the Calvo pricing mechanism. A few things should be noted here. First, the supply side of the GNK model includes three dynamic equations—(2), (3), and (4)—rather than one as in the simple NK model approximated around a zero steady-state inflation. Setting $\pi = 1$ in these three equations yields the standard NKPC. Second, s_t is a backward-looking variable, so its inclusion adds inertia to the adjustment of inflation. Therefore, the dynamics of the GNK model is richer, as discussed in Ascari and Sbordone (2014). Finally, to close the model, Equation (5) represents monetary policy actions—that is, a Taylor-type rule in which $\psi_\pi, \psi_x, \psi_{\Delta y}, \rho_r$ are chosen by the central bank, and echo its responsiveness to the inflation gap, output gap, output growth, and degree of inertia in interest rate setting, respectively. The term $\epsilon_{r,t}$ is an exogenous transitory monetary policy shock, whose standard deviation is given by σ_r. Equation (6) is the definition of the output gap, while the law of motion for the natural level of output is given by Equation (7).

The remaining fundamental disturbances involve a preference shock d_t, a shock to the growth rate of technology g_t, and an inflation target shock π_t^*. Each of these three shocks follows an $AR(1)$ process:

$$\log d_t = (1-\rho_d)\log d + \rho_d \log d_{t-1} + \epsilon_{d,t} \qquad 0 < \rho_d < 1,$$
$$\log g_t = (1-\rho_g)\log g + \rho_g \log g_{t-1} + \epsilon_{g,t} \qquad 0 < \rho_g < 1,$$

and

$$\log \pi_t^* = (1-\rho_{\pi^*})\log \pi + \rho_{\pi^*} \log \pi_{t-1}^* + \epsilon_{\pi^*,t} \qquad 0 < \rho_{\pi^*} < 1, \quad (8)$$

where the standard deviations of the innovations $\epsilon_{d,t}$, $\epsilon_{g,t}$, and $\epsilon_{\pi^*,t}$ are denoted by σ_d , σ_g, and σ_{π^*}, respectively.

Under a fixed inflation target, that is with no inflation target shock ($\sigma_{\pi^*} = 0$), Equation (8) drops out and the policy rules boils down to

$$\widehat{r}_t = \rho_r \widehat{r}_{t-1} + (1-\rho_r)\left\{\psi_\pi \widehat{\pi}_t + \psi_x \widehat{x}_t + \psi_{\Delta y}\left(\widehat{y}_t - \widehat{y}_{t-1} + \widehat{g}_t\right)\right\} + \epsilon_{r,t}, \quad (9)$$

in which the central bank's target becomes equal to a constant steady-state or trend inflation π. In other words, the model with a time-varying inflation target nests the one with a fixed target in the absence of inflation target shocks.

3.2 Rational Expectations Solution under Indeterminacy

To solve the model, the paper applies the method proposed by Lubik and Schorfheide (2003). The linear rational expectations (LRE) system can be compactly written as

$$A_0(\theta)\varrho_t = A_1(\theta)\varrho_{t-1} + B(\theta)\epsilon_t + C(\theta)\eta_t, \quad (10)$$

where ϱ_t, ϵ_t, and η_t denote the vector of endogenous variables, fundamental shocks, and one-step-ahead expectation errors, respectively, and $A_0(\theta)$, $A_1(\theta)$, $B(\theta)$, and $C(\theta)$ are appropriately defined coefficient matrices. From a methodological perspective, the solution of Lubik and Schorfheide (2003) follows from that of Sims (2002). However, it has the added advantage of being explicit in dealing with expectation errors, since it makes the solution suitable for solving and estimating models featuring multiple equilibria. In particular, under indeterminacy, η_t becomes a linear function of the fundamental shocks ϵ_t and the purely extrinsic sunspot disturbances ζ_t. The full set of solutions to the LRE model entails

$$\varrho_t = \Phi(\theta)\varrho_{t-1} + \Phi_\epsilon(\theta, \widetilde{M})\epsilon_t + \Phi_\zeta(\theta)\zeta_t, \quad (11)$$

where $\Phi(\theta)$, $\Phi_\epsilon(\theta, \widetilde{M})$, and $\Phi_\zeta(\theta)$[12] are the coefficient matrices.[13] The sunspot shock satisfies $\zeta_t \sim i.i.d.\ \mathsf{N}(0, \sigma_\zeta^2)$. Accordingly, indeterminacy can manifest itself in one of two different ways: (i) purely extrinsic non-fundamental disturbances can affect model dynamics through endogenous expectation errors; and (ii) propagation of fundamental shocks cannot be uniquely pinned down, and the multiplicity of equilibria affecting this propagation mechanism is captured by the arbitrary matrix $\widetilde{M}$.

Following the methodology proposed by Lubik and Schorfheide (2004), $\widetilde{M}$ is replaced with $M^*(\theta) + M$, and the prior mean for M is set equal to zero. The solution selects $M^*(\theta)$ by using a least squares criterion to minimize the distance between the impact response of the endogenous variables to fundamental shocks ($\partial \varrho_t / \partial \epsilon_t'$) at the boundary between the determinacy and indeterminacy regions.[14] Finding an analytical solution to the boundary in this model is infeasible, and hence, following Justiniano and Primiceri (2008) and Hirose (2020), the paper resorts to a numerical procedure to find the boundary by perturbing the parameter ψ_π in the monetary policy rule.

3.3 Equilibrium Determinacy and Trend Inflation

Before moving onto the empirical investigation, this subsection revisits how the determinacy properties of the model are altered by trend inflation. Ascari and Ropele (2009) and Ascari and Sbordone (2014) show that trend inflation makes price-setting firms more forward looking, thereby flattening the NKPC and widening the indeterminacy region. Figure 1 documents how trend inflation alters the determinacy region. Since analytical solution is infeasible unless one assumes indivisible labor, the determinacy results shown here are numerical.[15]

---

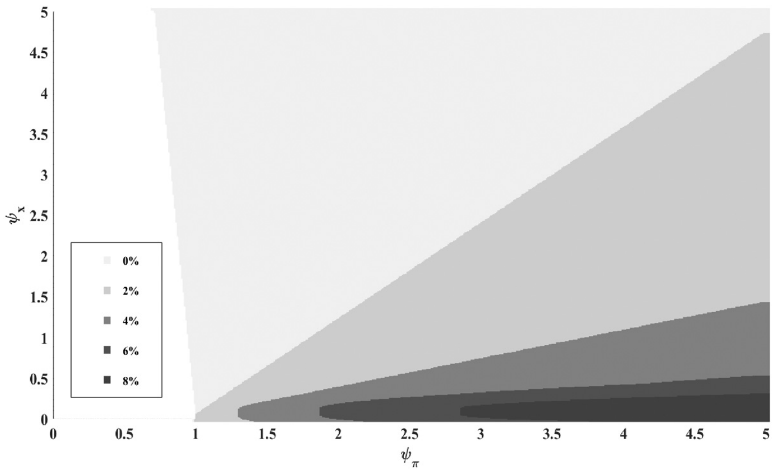
[12] Lubik and Schorfheide (2003) express this term as $\Phi_\zeta(\theta, M_\zeta)$, where M_ζ is an arbitrary matrix. For identification purpose, the paper imposes their normalization such that $M_\zeta = I$.

[13] Under determinacy, the solution boils down to $\varrho_t = \Phi^D(\theta)\varrho_{t-1} + \Phi_\epsilon^D(\theta)\epsilon_t$.

[14] This methodology has been used in previous studies, such as Benati and Surico (2009), Hirose (2007, 2008, 2013, 2020), and Doko Tchatoka et al. (2017).

[15] The parameter values and the policy rule used in the numerical computation are similar to Ascari and Sbordone (2014). In particular, $\beta = 0.99$, $\varepsilon = 11$,

Figure 1. Determinacy Region and Trend Inflation



Note: Shaded area corresponds to determinacy regions for different levels of trend inflation.

The determinacy region shrinks with trend inflation, as documented by Ascari and Ropele (2009) and Ascari and Sbordone (2014).[16] In other words, a stronger response to the inflation gap together with a weaker response to the output gap is required to generate determinacy at higher levels of trend inflation. Therefore, monetary policy should respond more to the inflation gap and less to the output gap, in order to stabilize inflation expectations. Moreover, in case of positive trend inflation, Coibion and Gorodnichenko (2011) show that both interest rate smoothing and stronger response to output growth, as opposed to the output gap, are stabilizing and therefore widen the determinacy region. Finally, it is important to note that allowing for an exogenous stochastic time-varying inflation target, as done in this paper, does not alter the (in)determinacy regions of the parameter space.

$\xi = 0.75$, $h = 0$ implying no habit formation in consumption, and $g = 1.005$ such that the steady-state growth rate of real per capita GDP is 2 percent per year. The policy rule is a simple Taylor rule of the form $r_t = \psi_\pi \pi_t + \psi_x x_t$.

[16]The figure is similar to Figure 4 in Ascari and Ropele (2009) and Figure 11 in Ascari and Sbordone (2014).

4. Econometric Strategy

4.1 Bayesian Estimation with Sequential Monte Carlo

The paper uses Bayesian techniques for estimating the parameters of the model and tests for indeterminacy using posterior model probabilities. It employs the sequential Monte Carlo (SMC) algorithm proposed by Herbst and Schorfheide (2014, 2015), which is particularly suitable for irregular and non-elliptical posterior distributions.[17]

First, priors are described by a density function of the form

$$p(\theta_S|S), \tag{12}$$

where $S \in \{D, I\}$; D and I stand for determinacy and indeterminacy, respectively; θ_S represents the parameters of the model S; and $p(.)$ stands for the probability density function. Next, the likelihood function $p(X_T|\theta_S, S)$ describes the density of the observed data, where X_T are the observations through to period T. Following Bayes' theorem, the posterior density is constructed as a combination of the prior density and the likelihood function:

$$p(\theta_S|X_T, S) = \frac{p(X_T|\theta_S, S)p(\theta_S|S)}{p(X_T|S)}, \tag{13}$$

where $p(X_T|S)$ is the marginal data density conditional on the model, which is given by

$$p(X_T|S) = \int_{\theta_S} p(X_T|\theta_S, S)p(\theta_S|S)d\theta_S. \tag{14}$$

A difficulty in the methodology of Lubik and Schorfheide (2003) is that the likelihood function of the model is possibly discontinuous at the boundary between the determinacy and indeterminacy region. As noted before, Lubik and Schorfheide (2004) propose to select $M^*(\theta)$ such that the impulse responses of the endogenous variables to fundamental shocks are continuous at the boundary. To test for indeterminacy, they estimate the model twice—first under

[17]See Hirose, Kurozumi, and Van Zandweghe (2020), who were the first to apply Bayesian estimation using the SMC algorithm to test for indeterminacy following Lubik and Schorfheide's (2003, 2004) methodology.

determinacy, and then under indeterminacy—and then compare the fit of the model under the two specifications. However, an importance sampling algorithm like SMC can use a single chain to explore the entire parameter space. Hence, to take full advantage of the algorithm, the paper estimates the model simultaneously over both determinate and indeterminate parameter space.[18] The likelihood function is then given by

$$\begin{aligned} p(X_T|\theta_S, S) = 1\{\theta_S \in \Theta^D\}p^D(X_T|\theta_D, D) \\ + 1\{\theta_S \in \Theta^I\}p^I(X_T|\theta_I, I), \end{aligned} \tag{15}$$

where Θ^D, Θ^I are the determinacy and indeterminacy regions of the parameter space; $1\{\theta_S \in \Theta^S\}$ is the indicator function, which equals 1 if $\theta_S \in \Theta^S$ and zero otherwise; and $p^D(X_T|\theta_D, D)$ and $p^I(X_T|\theta_I, I)$ are the likelihood functions under determinacy and indeterminacy, respectively. Following Herbst and Schorfheide (2014, 2015), the paper builds a particle approximation of the posterior distribution through tempering the likelihood. A sequence of tempered posteriors is defined as

$$\Pi_n(\theta_S) = \frac{[p(X_T|\theta_S, S)]^{\phi_n} p(\theta_S|S)}{\int_{\theta_S} [p(X_T|\theta_S, S)]^{\phi_n} p(\theta_S|S) d\theta_S}, \tag{16}$$

where ϕ_n is the tempering schedule, which slowly increases from zero to one.

The algorithm generates weighted draws from the sequence of posteriors $\{\Pi_n(\theta_S)\}_{n=1}^{N_\phi}$, where N_ϕ is the number of stages. At any stage, the posterior distribution is represented by a swarm of particles $\left\{\theta_n^i, W_n^i\right\}_{i=1}^{N}$, where W_n^i is the weight associated with θ_n^i and N denotes the number of particles. The algorithm has three main steps. First, in the *correction* step, the particles are reweighted to reflect the density in iteration n. Next, in the *selection* step, any

[18] Ettmeier and Kriwoluzky (2020) and Hirose, Kurozumi, and Van Zandweghe (2020) also use SMC to estimate their model over the entire parameter space. For an alternative approach that allows estimation over the entire parameter space, while using standard packages like Dynare and standard estimation algorithms, see Bianchi and Nicolò (2017).

particle degeneracy is eliminated by resampling the particles. Finally, in the *mutation* step, the particles are propagated forward using a Markov transition kernel in order to adapt to the current bridge density.

In the first stage, i.e., when $n = 1$, ϕ_1 is zero. Hence, the prior density serves as an efficient proposal density for $\Pi_1(\theta_S)$; that is, the algorithm is initialized by drawing the initial particles from the prior. Likewise, the idea is that the density of $\Pi_n(\theta_S)$ is a good proposal density for $\Pi_{n+1}(\theta_S)$.

Number of Particles, Number of Stages, Tempering Schedule. The tempering schedule is a sequence that slowly increases from zero to one, and is determined by $\phi_n = \left(\frac{n-1}{N_\phi - 1}\right)^\tau$, where τ controls the shape of the schedule. The tuning parameters N, N_ϕ, and τ are fixed ex ante. The estimation uses $N = 10,000$ particles and $N_\phi = 200$ stages. The parameter that controls the tempering schedule, denoted by τ, is set at 2 following Herbst and Schorfheide (2015).

Resampling. Resampling is necessary to avoid particle degeneracy. A rule-of-thumb measure of this degeneracy, proposed by Liu and Chen (1998), is given by the reciprocal of the uncentered variance of the particles, and is called the effective sample size (ESS). The estimation employs systematic resampling whenever $ESS_n < \frac{N}{2}$.

Mutation. Finally, one step of a single-block random-walk Metropolis-Hastings (RWMH) algorithm is used to propagate the particles forward.

The SMC algorithm has several practical advantages. First, it allows for estimation over the entire parameter space. Lubik and Schorfheide (2004) show that the shape of the likelihood function may be different under indeterminacy. This then makes MCMC-based inference complicated because it is less suited to approximating the posterior when the latter is not well shaped or has multiple modes. In order to deal with this issue, Lubik and Schorfheide (2004) estimate the model over determinacy and indeterminacy separately. However, SMC methods are more appropriate when the posterior distribution displays irregular patterns, as also pointed out by Ascari, Bonomolo, and Lopes (2019) in a similar context.

Second, the algorithm does not require one to find the mode of the posterior distribution.[19] Computing the posterior mode when allowing for indeterminacy can be computationally cumbersome in practice because of the irregular shape of the likelihood function. The SMC algorithm is an "importance sampling algorithm"; that is, instead of attempting to sample directly from the posterior, the algorithm draws from a different tractable distribution, commonly referred to as an importance distribution. The reweighting of a particle from the importance distribution gives the particle the status of an actual draw from the posterior distribution. Here, the initial particles are drawn from the prior; that is, the prior serves as the initial proposal density for this tractable distribution. In subsequent steps, the density in the current stage of the algorithm, i.e., $\Pi_n(\theta_S)$, serves as a proposal density for the next stage.

Finally, an additional advantage on the computational front is parallelization. The particle mutation phase is ideally suited for parallelization because the propagation steps are independent across particles and do not require any communication across processors. For models estimated under indeterminacy along the lines of Lubik and Schorfheide (2004), the evaluation of the likelihood function is computationally very costly because it requires running a model solution procedure that bridges the gap between the impact response of the variables to fundamental shocks at the boundary between determinacy and indeterminacy (by picking $M^*(\theta)$). Whenever analytical solution to the boundary is not available, this requires numerically tracing the boundary for every draw at every stage. Thus, gains from parallelization can be quite large.

4.2 Data

The paper uses three U.S. quarterly time series: per capita real GDP growth rate $100\Delta \log Y_t$, quarterly growth rate of the GDP deflator $100\Delta \log P_t$, and the federal funds rate $100 \log R_t$. The model is estimated over two sample periods. The first sample, 1960:Q1–1979:Q2,

[19]Standard methods like Metropolis-Hastings algorithm constructs a Gaussian approximation around the posterior mode and uses a scaled version of the asymptotic covariance matrix (taken to be the inverse of the Hessian computed at the mode) as the covariance matrix for the proposal distribution.

corresponds to the Great Inflation period. The second one, 1984:Q1–2008:Q2, corresponds to the Great Moderation period, which is characterized by dramatically milder macroeconomic volatilities. The measurement equations relating the relevant elements of ϱ_t to the three observables are given by

$$\begin{bmatrix} 100\Delta\log Y_t \\ 100\Delta\log P_t \\ 100\log R_t \end{bmatrix} = \begin{bmatrix} g^* \\ \pi^* \\ r^* \end{bmatrix} + \begin{bmatrix} \widehat{y}_t - \widehat{y}_{t-1} + \widehat{g}_t \\ \widehat{\pi}_t \\ \widehat{r}_t \end{bmatrix}, \tag{17}$$

where $g^* = 100(g-1)$, $\pi^* = 100(\pi-1)$, and $r^* = 100(r-1)$.

4.3 Calibration and Prior Distributions

Some parameters are fixed before the estimation. The elasticity of substitution among intermediate goods and the inverse of the labor-supply elasticity are fixed at $\varepsilon = 11$ and $\varphi = 1$, respectively. The former value corresponds to a steady-state markup of 10 percent, which is consistent with the estimate of Basu and Fernald (1997).[20] The latter value is a standard one in the macroeconomic literature.[21] The remaining parameters are estimated.[22] Table 1 summarizes the specification of the prior distributions. The prior for the inflation coefficient ψ_π follows a gamma distribution centered at 1.10 with a standard deviation of 0.50, while the response coefficient to the output gap ψ_x and output growth $\psi_{\Delta y}$ are both centered at 0.125 with standard deviation 0.10. The paper uses beta distribution with mean 0.50 for the smoothing coefficient ρ_r, the Calvo probability ξ, and habit persistence in consumption h, and 0.70 for the persistence of the discount factor shock ρ_d. The autoregressive parameter of the total factor productivity (TFP) shock ρ_g follows a beta distribution

[20]More recent estimates by Edmond, Midrigan, and Xu (2018) suggest that aggregate markups could be as high as 25 percent. Hence, to check the robustness of the results, the elasticity of substitution among intermediate goods is alternatively set at $\varepsilon = 5$, corresponding to a steady-state markup of 25 percent. The online appendices show that the results remain robust.

[21]See, for instance, Hirose, Kurozumi, and Van Zandweghe (2020), who also set $\varphi = 1$.

[22]For the subjective discount factor β, the steady-state condition $\beta = \frac{\pi g}{r}$ is used in estimation.

Table 1. Prior and Posterior Distributions

Name	Density	Prior Mean (Std. Dev.)	1960:Q1–1979:Q2 Posterior Mean [90% Interval]	1984:Q1–2008:Q2 Posterior Mean [90% Interval]
ψ_π	Gamma	1.10 (0.50)	2.29 [1.42,2.89]	4.04 [3.08,4.95]
ψ_x	Gamma	0.125 (0.10)	0.12 [0.00,0.25]	0.13 [0.00,0.26]
$\psi_{\Delta y}$	Gamma	0.125 (0.10)	0.16 [0.02,0.28]	0.38 [0.08,0.64]
ρ_r	Beta	0.50 (0.20)	0.41 [0.23,0.65]	0.71 [0.62,0.80]
π^*	Normal	0.98 (0.50)	1.18 [0.86,1.51]	0.69 [0.52,0.85]
r^*	Gamma	1.60 (0.25)	1.49 [1.18,1.79]	1.46 [1.19,1.69]
g^*	Normal	0.50 (0.10)	0.54 [0.38,0.68]	0.51 [0.40,0.62]
h	Beta	0.50 (0.10)	0.39 [0.31,0.51]	0.40 [0.31,0.50]
ξ	Beta	0.50 (0.10)	0.39 [0.27,0.57]	0.49 [0.36,0.61]
ρ_d	Beta	0.70 (0.10)	0.79 [0.67,0.88]	0.92 [0.89,0.95]
ρ_g	Beta	0.40 (0.10)	0.17 [0.11,0.26]	0.17 [0.11,0.24]
$\rho_{\pi*}$	Beta	0.95 (0.025)	0.96 [0.93,0.99]	0.95 [0.91,0.98]
σ_r	Inv-Gamma	0.60 (0.20)	0.39 [0.27,0.48]	0.21 [0.16,0.26]
σ_d	Inv-Gamma	0.60 (0.20)	0.71 [0.39,0.95]	1.69 [1.20,2.18]
σ_g	Inv-Gamma	0.60 (0.20)	1.28 [1.07,1.55]	0.71 [0.59,0.83]
$\sigma_{\pi*}$	Uniform	0.075 (0.0433)	0.08 [0.04,0.13]	0.04 [0.03,0.06]
σ_ζ	Inv-Gamma	0.60 (0.20)	0.57 [0.24,0.90]	0.57 [0.25,0.93]
$M_{r,\zeta}$	Normal	0.00 (1.00)	0.02 [–1.71,1.62]	0.02 [–1.64,1.65]
$M_{d,\zeta}$	Normal	0.00 (1.00)	–0.08 [–1.66,1.65]	0.00 [–1.63,1.63]
$M_{g,\zeta}$	Normal	0.00 (1.00)	–0.01 [–1.70,1.65]	0.06 [–1.54,1.61]
$M_{\pi*,\zeta}$	Normal	0.00 (1.00)	0.00 [–1.63,1.65]	0.01 [–1.62,1.71]

Note: The prior probability of determinacy is 0.498. The inverse gamma distributions are of the form $p(\sigma|\nu,\varsigma) \propto \sigma^{-\nu-1}e^{-\nu\varsigma^2/2\sigma^2}$, where $\nu = 4$ and $\varsigma = 0.45$.

Table 2. Determinacy versus Indeterminacy

Sample	Inflation Target	Log-Data Density	Probability of Determinacy
1960:Q1–1970:Q2	Fixed	–152.08	0.20
	Time Varying	–144.29	1.0
1984:Q1–2008:Q2	Fixed	–32.58	1.0
	Time Varying	–27.58	1.0

Note: The table shows the SMC-based approximations of log marginal data densities and the posterior probabilities of determinacy.

centered at 0.40, since this process already includes a unit root, while that of the inflation target shock ρ_{π^*} is assumed to be highly persistent and is centered at 0.95. The priors for the quarterly net steady-state rates of output growth, inflation, and nominal interest rate, denoted by g^*, π^*, and r^*, respectively, are distributed roughly around their average values over the entire sample period.

For the shocks, the prior distributions for all but one follow an inverse-gamma distribution with mean 0.60 and standard deviation 0.20. The exception is the standard deviation of the innovation to the inflation target shock σ_{π^*}, which is an important parameter in the analysis, as it governs the rate at which π_t^* drifts. Following Cogley, Primiceri, and Sargent (2010), the paper adopts a weakly informative uniform prior on (0,0.15) for this parameter.

Finally, in line with Lubik and Schorfheide (2004), the coefficients M follow standard normal distributions. Hence, the prior is centered around the baseline solution of Lubik and Schorfheide (2004).

Importantly, the choice of the priors leads to a prior predictive probability of determinacy of about 50 percent, which is quite even and suggests no prior bias toward either determinacy or indeterminacy.

5. Estimation Results

5.1 Model Comparison

Table 2 collects the results for the empirical performance of the model with fixed and time-varying inflation targets. To assess the quality of the models' fit to the data, log marginal data densities and posterior model probabilities are reported. The posterior probability of determinacy is calculated as the fraction of the draws in the final stage of the SMC algorithm that generate determinate equilibrium. The SMC algorithm delivers a numerical approximation of the marginal data density as a byproduct in the *correction* step, which is given by

$$p^{SMC}(X_T|S) = \prod_{n=1}^{N_\phi} \left(\frac{1}{N} \sum_{i=1}^{N} \widetilde{w}_n^i W_{n-1}^i \right),$$

where $\widetilde{w}_n^i$ is the incremental weight defined by

$$\widetilde{w}_n^i = [p(X|\theta_{n-1}^i, S)]^{\phi_n - \phi_{n-1}},$$

and W_n^i are the normalized weights. Herbst and Schorfheide (2014, 2015) show that the particle weights converge under suitable regularity conditions as follows:

$$\begin{aligned}
&\frac{1}{N} \sum_{i=1}^{N} \widetilde{w}_n^i W_{n-1}^i \\
&\Longrightarrow \int [p(X|\theta_s, S]^{\phi_n - \phi_{n-1}} \frac{[p(X|\theta_s, S]^{\phi_{n-1}} p(\theta_S|S)}{\int [p(X|\theta_s, S]^{\phi_{n-1}} p(\theta_S|S) d\theta_S} d\theta_S \\
&= \frac{\int [p(X|\theta_s, S]^{\phi_n} p(\theta_S|S) d\theta_S}{\int [p(X|\theta_s, S]^{\phi_{n-1}} p(\theta_S|S) d\theta_S}.
\end{aligned}$$

Table 2 shows that in case of a fixed inflation target, indeterminacy cannot be ruled out in the pre-Volcker period, while determinacy unambiguously prevails after 1984. Nevertheless, the fact that the posterior probability of determinacy in the pre-Volcker period is around 20 percent in the case of a fixed inflation target is a priori unexpected, given the empirical findings of Lubik and Schorfheide (2004) and Hirose, Kurozumi, and Van Zandweghe (2020), who show the pre-Volcker period to be explicitly characterized by indeterminacy. However, this is a result of estimating a GNK model with trend inflation and homogenous labor while using the GDP deflator to measure inflation. In fact, upon further investigation, the paper finds that when using CPI to measure inflation instead of the GDP deflator (as in Lubik and Schorfheide 2004), or assuming firm-specific labor instead of homogenous labor (as in Hirose, Kurozumi, and Van Zandweghe 2020), strong evidence for indeterminacy reemerges, and the results are documented in Section 6 of the paper.

In contrast, when allowing for time variation in the inflation target pursued by the Federal Reserve in the pre-Volcker period, the entire mass of the posterior distribution falls in the determinacy region of the parameter space, and this finding remains robust to various perturbations of the baseline model.[23] Phrased alternatively, it suggests that monetary policy did not result in sunspot fluctuations during the Great Inflation period, given time variation in the inflation target.

In terms of posterior odds ratio, the marginal likelihood points toward the empirical superiority of the specification featuring time variation in the inflation target in both subsamples. In particular, the Bayes factor or KR ratio involving fixed versus time-varying target is about 16 in the pre-Volcker period, and points toward "very strong" evidence in favor of the model in which the Federal Reserve follows a time-varying inflation target.[24,25]

The finding that allowing for time-varying inflation target leads to determinacy in the Great Inflation era might be surprising, given that the literature has established the pre-Volcker period as characterized by indeterminacy. What is driving this result? On one hand, Lubik and Schorfheide (2004) and Fujiwara and Hirose (2012) argue that a model under indeterminacy can generate richer persistent inflation dynamics compared with determinacy because fewer autoregressive roots are suppressed. On the other hand, as documented by Cogley, Primiceri, and Sargent (2010), inflation target

[23]The post-1984 period remains explicitly characterized by determinacy.

[24]We report the Bayes Factor or KR ratio as suggested in Kass and Raftery (1995), calculated as 2(log-data density H1 − log-data density H0), where the null hypothesis (H0) is always the less-preferred model (while the alternative hypothesis, H1, is the preferred one). Hence, we weight evidence against the null hypothesis.

[25]According to Kass and Raftery (1995), values of KR below 2 are "not worth more than a bare mention," between 2 and 6 suggest "positive" evidence in favor of one of the two models, between 6 and 10 suggest "strong" evidence, and larger than 10 suggest "very strong" evidence.

shocks induce persistent responses in the inflation gap, which help to capture the highly persistent inflation dynamics in the 1970s. According to the posterior estimates, inflation target was loosely anchored during the pre-Volcker period, as evident from its higher innovation variance. As such, the model no longer requires the richer inflation dynamics that arise under indeterminacy in order to explain the Great Inflation episode, and therefore we find the pre-Volcker period to be characterized by determinacy. Moreover, the inflation gap that enters the Taylor rule when the target is drifting over time is less volatile than the inflation gap with a fixed target. For a given historical path of the nominal interest rate, then the response of the nominal rate to the inflation gap turns out to be higher in case of a time-varying target, which leads to determinacy.

5.2 Parameter Estimates and the Federal Reserve's Inflation Target

Table 1 reports the posterior means and the 90 percent highest posterior density intervals based on 10,000 particles from the final stage of the SMC algorithm under time-varying inflation target (the specification that fits better).[26] As seen in the table, the Taylor rule's response to the inflation gap is strongly active in the pre-1979 period. In fact, the point estimate is above 2, which shows why the posterior favors determinacy under a time-varying target. Moving across the sample, the policy responses to the inflation gap and output growth and inertia in interest rate setting all increased, while trend inflation fell considerably.

The estimated response to the inflation gap is in line with the results of Fernández-Villaverde and Rubio-Ramírez (2008), who estimate a DSGE model with time-varying structural parameters. Fernández-Villaverde and Rubio-Ramírez (2008) find that the response to inflation was slightly above 1 during the 1950s, 1960s, and early 1970s, and then dramatically increased in the mid-1970s and especially after Volcker's appointment as the Federal Reserve Chairman, with the average response to inflation being roughly

[26]The online appendices report parameter estimates under a fixed target.

around 4 during the 1980s and 1990s.[27] They also find substantial variation in the estimated inflation target, with the target rising in the late 1960s and 1970s and falling after the Volcker disinflation, which is similar in pattern to what this paper finds, as discussed later. Fernández-Villaverde and Rubio-Ramírez (2010) and Fernández-Villaverde, Guerrón-Quintana, and Rubio-Ramírez (2010) also find evidence of changes in monetary policy through the lens of estimated non-linear DSGE models with both time-varying parameters and stochastic volatilities. Both these papers find that the response to inflation started above 1 and increased during the 1960s, before collapsing below 1 and therefore violating the Taylor principle during the Burns-Miller Chairmanship in the 1970s. Thereafter, the response increased strongly with the arrival of Volcker. However, Fernández-Villaverde and Rubio-Ramírez (2010) and Fernández-Villaverde, Guerrón-Quintana, and Rubio-Ramírez (2010) find that the response to inflation was again below 1 during most of Greenspan's tenure, which is at odds with the findings of this paper and others in the literature that monetary policy was strongly active during the Great Moderation (Clarida, Galí, and Gertler 2000; Lubik and Schorfheide 2004; Boivin and Giannoni 2006). Nevertheless, Fernández-Villaverde, Guerrón-Quintana, and Rubio-Ramírez (2010) find that their estimates still guarantee local equilibrium determinacy even though the response to inflation temporarily violates the Taylor principle, as the agents still expect the Taylor principle to be satisfied on average. Hence, Fernández-Villaverde, Guerrón-Quintana, and Rubio-Ramírez (2010) suggest that equilibrium was determinate even during the turbulent 1970s, while they blame the instability of the Great Inflation on bad shocks or bad luck. One possible explanation for the difference between the estimated response to inflation in the current paper and Fernández-Villaverde and Rubio-Ramírez (2010) and Fernández-Villaverde, Guerrón-Quintana, and Rubio-Ramírez (2010), apart from time-varying parameters and stochastic volatilities, is the absence of a time-varying inflation target in the latter studies.[28]

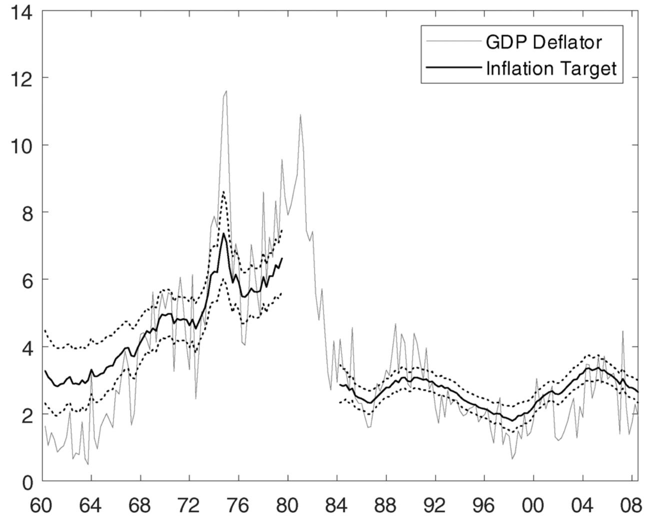
[27] See Figure 2.2 and 2.3 in Fernández-Villaverde and Rubio-Ramírez (2008).

[28] Fernández-Villaverde, Guerrón-Quintana, and Rubio-Ramírez (2010) point out the difficulties of estimating a time-varying inflation target in the context of their non-linear DSGE model solved with perturbation methods.

In terms of the policy shocks, the innovation variance of the two shocks, $\epsilon_{\pi^*,t}$ and $\epsilon_{r,t}$, both declined, which is in line with the findings of Cogley, Primiceri, and Sargent (2010). According to the posterior mean estimates, the innovation variance fell from 0.08 to 0.04 for the inflation target shock, and from 0.39 to 0.21 for the policy rate shock. However, unlike Cogley, Primiceri, and Sargent (2010), who find a moderate increase in the responsiveness to the inflation gap, this paper finds quite a substantial increase across the two periods. This suggests that both the systematic response to the inflation gap and a better anchoring of the inflation target may have played a key role in the decline of inflation volatility and persistence during the Great Moderation period.

Turning to the deep parameters, the degree of habit formation remains stable, roughly around 0.40 for both subsamples. The posterior mean for the degree of price stickiness ξ turns out to be 0.39 in the pre-Volcker period and increases to 0.49 in the post-84 period, which are smaller than the estimates reported in Smets and Wouters (2007), Fernández-Villaverde and Rubio-Ramírez (2008), Fernández-Villaverde, Guerrón-Quintana, and Rubio-Ramírez (2010), and Justiniano, Primiceri, and Tambalotti (2010). For example, allowing for time variation in the estimated Calvo parameter, Fernández-Villaverde and Rubio-Ramírez (2008) find an estimate of around 0.8 in the 1950s. Thereafter, the estimate drops somewhat in the 1960s and 1970s, before rising steadily again since the late 1970s and reaching above 0.8 during the 1990s. However, the microeconomic evidence on the average duration of prices suggests different degrees of price stickiness. For example, Bils and Klenow (2004) find that firms update prices every four to five months, roughly corresponding to $\xi = 0.40$, while Nakamura and Steinsson (2008) find longer duration ranging between 8 and 11 months on average, which roughly corresponds to $\xi = 0.70$. In any case, it turns out that the estimated degree of price stickiness has increased in the second period, which is in line with the findings of Smets and Wouters (2007) and Fernández-Villaverde and Rubio-Ramírez (2008). As they point out, this finding is consistent with the idea that low and stable inflation may reduce the cost of not adjusting prices and therefore lengthen the average price duration, thereby leading to a flatter Phillips curve. Nevertheless, the main result documented in this paper, regarding the role of time-varying inflation target in driving equilibrium determinacy

Figure 2. Inflation and the Federal Reserve's Inflation Target



Note: The solid dark line, labeled "Inflation Target," plots the mean of the target estimates based on the posterior draws of the parameters and the dotted lines show the 5th and 95th percentiles.

during the Great Inflation, does not depend on the estimated low value of the Calvo parameter, as shown in the robustness section.

Among the non-policy shocks, there is an increase in the persistence and volatility of the discount factor shock, a finding shared with Hirose, Kurozumi, and Van Zandweghe (2020). Finally, there is a decline in the volatility of technology shocks, which is in line with Lubik and Schorfheide (2004), Leduc and Sill (2007), and Smets and Wouters (2007).

Before moving on to study the drivers of the Great Moderation, Figure 2 plots the model-implied evolution of the Federal Reserve's inflation target on top of the actual GDP deflator inflation rate. Here, the paper employs the Kalman smoother to obtain ex post estimates of π_t^* based on the observations that are included in the construction of the likelihood function.[29] The inflation target was

[29]The solid dark line, labeled "Inflation Target," shows the mean of the target estimates based on the posterior draws of the parameters, and the dotted lines show the 5th and 95th percentiles.

low—even if higher than realized inflation—at the beginning of the 1960s. Thereafter, the inflation target began rising in the mid-1960s and jumped higher in the aftermath of the 1973 oil crisis. The upward trend in the inflation target in the 1970s may be interpreted as "a systematic tendency for Federal Reserve policy to translate the short-run price pressures set off by adverse supply shocks into more persistent movements in the inflation rate itself—part of an effort by policymakers to avoid at least some of the contractionary impact those shocks would otherwise have had on the real economy." (Ireland 2007, p. 1853). Subsequently, it dropped remarkably during the Volcker-disinflation period and somewhat settled around 2.5 percent since the mid-1980s. As in Leigh (2008, pp. 2022–23), the time-varying implicit inflation target for the post-1984 subsample can be divided into separate chunks: (i) "the opportunistic approach to disinflation"—a period covering from the mid-1980s to mid-1990s—during which, according to Orphanides and Wilcox (2002), the Federal Reserve did not take deliberate anti-inflation action, but rather waited for external circumstances to deliver the desired reduction in inflation; (ii) "the low-inflation equilibrium" in the late 1990s; and (iii) "the deflation scare" in the early 2000s, which led to a lowering of the federal funds rate, while the inflation target rose above actual inflation.[30] Overall, visual inspection suggests that the estimated target is similar to those previously reported by Ireland (2007), Leigh (2008), Cogley, Primiceri, and Sargent (2010), Aruoba and Schorfheide (2011), and Castelnuovo, Greco, and Raggi (2014), among others.

6. Robustness Analysis

The paper conducts robustness checks along the following dimensions: (i) alternative measure of inflation as observable in the estimation, (ii) firm-specific labor, (iii) estimating the NK model of Lubik and Schorfheide (2004) while allowing for a time-varying inflation

[30]The early 2000s was a period of low interest rates and, as noted by Eggertsson and Woodford (2003), keeping interest rates low for an extended period of time is equivalent to a rise in the inflation target. For alternative interpretation of monetary policy during the 2000s, see Groshenny (2013), Belongia and Ireland (2016), and Doko Tchatoka et al. (2017).

Table 3. Determinacy versus Indeterminacy (robustness checks)

	Constant Target	Time-Varying Target
	Log-Data Density (Probability of Determinacy)	**Log-Data Density (Probability of Determinacy)**
Sample: 1960:Q1–1979:Q2		
CPI	–152.32 (0.12)	–144.23 (0.95)
Firm-Specific Labor	–145.31 (0)	–144.07 (0.97)
Lubik and Schorfheide (2004)	–359.59 (0)	–357.92 (0.97)
Calibrate $\rho_{\pi*}$	—	–143.95 (1)
Calibrate ξ	–153.49 (0.19)	–145.87 (0.95)
Calibrate π^*	–149.29 (0)	–147.11 (1)
Sample: 1984:Q1–2008:Q2		
CPI	–89.98 (1)	–92.20 (1)
Firm-Specific Labor	–31.62 (1)	–34.35 (1)
Lubik and Schorfheide (2004)	–238.63 (0.97)	–237.38 (0.99)
Calibrate $\rho_{\pi*}$	—	–28.08 (1)
Calibrate ξ	–30.88 (0.99)	–31.16 (1)

Note: The table shows the SMC-based approximations of log marginal data densities and the posterior probabilities of determinacy for various robustness checks. "CPI" refers to the estimations with CPI inflation data; "Firm-Specific Labor" is the GNK model of Hirose, Kurozumi, and Van Zandweghe (2020); "Calibrate $\rho_{\pi*}$" refers to the estimations where $\rho_{\pi*} = 0.995$; "Calibrate ξ" refers to the estimations where $\xi = 0.75$; "Calibrate π^*" refers to the estimations where $\pi^* = 2$.

target, (iv) calibrating the persistence of the inflation target process, (v) calibrating the degree of price stickiness to a higher level, and (vi) setting the steady-state or trend inflation to a higher value for the Great Inflation sample. Table 3 summarizes the log-data densities and posterior model probabilities.[31] The top half of the table shows results for the Great Inflation period, while the bottom half shows results for the Great Moderation period.

6.1 Alternative Measure of Inflation

The baseline models are estimated using GDP deflator as a measure of inflation. To check the robustness of the results, the paper reestimates the models using CPI to measure inflation, as in Lubik and Schorfheide (2004). The posterior mass lies almost entirely in the indeterminacy region in the pre-Volcker period under a fixed inflation target, with around 90 percent of the draws from the posterior distribution generating indeterminacy. Nevertheless, time-varying inflation target continues to fit better in the pre-Volcker period, and as a result determinacy prevails as before. One difference with respect to the baseline results is that the model with a fixed target fits better in the post-1984 period, implying a larger role played by the decline in the variability of the inflation target in driving the reduction in inflation volatility.

6.2 Firm-Specific Labor

The analysis so far has relied on a GNK model with homogenous labor, following Ascari and Ropele (2009) and Ascari and Sbordone (2014). However, Kurozumi and Van Zandweghe (2017) show that a similar model with firm-specific labor leads to a distinct representation of inflation dynamics, which makes it more susceptible to indeterminacy induced by higher trend inflation. The only difference is that the household now supplies a set of labor services $N_{i,t}$, each of which is specific to intermediate-good firm $i \in [0, 1]$, instead of supplying a homogenous labor service N_t. This leads to a distinct representation of the supply side of the model. Following Hirose, Kurozumi, and Van Zandweghe (2020),

[31] The online appendix reports the parameter estimates.

the log-linearized GNKPC for the model with firm-specific labor is given by[32]

$$\widehat{\pi}_t = \kappa_f E_t \widehat{\pi}_{t+1} + (1+\varphi)\vartheta_f \widehat{y}_t + \vartheta_f \left(\frac{h}{g-h}\right)[\widehat{y}_t - \widehat{y}_{t-1} + \widehat{g}_t] + \widehat{\Psi}_t, \tag{18}$$

$$\widehat{\Psi}_t = \gamma_\psi E_t \widehat{\Psi}_{t+1} + \kappa_\psi \left(E_t \widehat{y}_{t+1} - \widehat{y}_t + E_t \widehat{g}_{t+1} + \varepsilon E_t \widehat{\pi}_{t+1} - \widehat{r}_t\right), \tag{19}$$

where $\kappa_f = \beta\xi\pi^{\varepsilon(1+\varphi)}/\xi\pi^{\varepsilon-1}$, $\vartheta_f = (1-\xi\pi^{\varepsilon-1})(1-\xi\beta\pi^{\varepsilon(1+\varphi)})/\xi\pi^{\varepsilon-1}(1+\varepsilon\varphi)$, $\gamma_\psi = \beta\xi\pi^{\varepsilon-1}$, and $\kappa_\psi = \gamma_\psi\left(\pi^{1+\varepsilon\varphi}-1\right)\left(1-\xi\pi^{\varepsilon-1}\right)/\xi\pi^{\varepsilon-1}(1+\varepsilon\varphi)$.

A few key points are particularly worth noting. First, as shown by Kurozumi and Van Zandweghe (2017), the slope of the GNKPC in the model with firm-specific labor (given by $(1+\varphi)\vartheta_f$) is less than that of the model with homogenous labor (given by $(1+\varphi)\vartheta$), as long as the elasticity of labor supply is finite, i.e., $\varphi > 0$. This reflects strategic complementarity in price setting incorporated by firm-specific labor. Therefore, inflation is less sensitive to output and so monetary policy is less capable of stabilizing inflation in the model with firm-specific labor. Second, Kurozumi and Van Zandweghe (2017) show that the long-run inflation elasticity of output implied by the GNKPC is highly sensitive to trend inflation in the model with firm-specific labor relative to the model with homogenous labor.[33] Higher trend inflation lowers this elasticity and makes the long-run version of the Taylor principle more restrictive for the Taylor rule's coefficients on inflation and output. Therefore, a model with firm-specific labor in the presence of trend inflation is meant to work against the results documented in this paper. Third, in the model with homogenous labor, the GNKPC depends on price distortion (s_t) as long as the trend inflation rate is non-zero (i.e., $\pi \neq 1$) and the elasticity of labor supply is finite (i.e., $\varphi > 0$). Therefore, the persistence of price distortion, as seen in its law of motion (4), generates endogenously persistent inflation dynamics in

[32]The Euler equation (Equation (1)), the specification of the Taylor rule (Equation (5) or Equation (9)), the definition of the output gap (Equation (6)), the expression for the natural level of output (Equation (7)), and the shock processes (Equation (8)) remain the same as in the model with homogenous labor.

[33]See Figure 2 of Kurozumi and Van Zandweghe (2017).

the model with homogenous labor, as stressed by Kurozumi and Van Zandweghe (2017). Finally, in case of infinite elasticity of labor supply (i.e., $\varphi = 0$), the GNKPC coincides between the models with firm-specific and homogenous labor, which implies that in that case the two models are equivalent.

Along these lines, the paper estimates a GNK model with positive trend inflation and firm-specific labor following Hirose, Kurozumi, and Van Zandweghe (2020).[34] In order to establish a valid comparison, this paper uses the exact same set of priors as they do. However, to achieve identification between the inflation target process and the policy rate shock, the paper assumes that the latter follows a transitory i.i.d. process while the former is a highly persistent AR(1) process following the literature.[35]

In line with Hirose, Kurozumi, and Van Zandweghe (2020), the pre-Volcker period is unambiguously characterized by indeterminacy, while the post-1984 period is characterized by determinacy, under the assumption of a fixed inflation target equal to trend inflation. However, when allowing for a time-varying inflation target, determinacy prevails in both sample periods, as before. In terms of the empirical fit, the model with a time-varying inflation target fits marginally better than one with a fixed target in the pre-Volcker period.[36] Given that the model with firm-specific labor is a priori expected to work against the baseline results, this set of findings somewhat mitigates, yet does not overturn, the key result. Despite the model being more prone to indeterminacy, the hypothesis that the inflation target has been drifting and as a consequence determinacy might have prevailed even in the pre-Volcker period is a possibility that cannot be empirically ruled out.

[34]This paper abstracts from price indexation given Hirose, Kurozumi, and Van Zandweghe's (2020) finding that the model with no indexation fits better.

[35]As in the baseline estimation, the elasticity of substitution among intermediate goods and the inverse of the labor-supply elasticity are fixed at $\varepsilon = 11$ and $\varphi = 1$, respectively

[36]In the post-1984 period, the model with a fixed inflation target fits better. Nonetheless, Table 2 shows that a model with homogenous labor and time-varying target fits even better.

6.3 Lubik and Schorfheide (2004)

To bridge the gap with key studies in the literature, the paper also estimates an NK model log-linearized around a zero inflation steady state.[37] To be transparent, the paper estimates the specification of the NK model as in Lubik and Schorfheide (2004), using the exact same set of priors, observables, and sample period as they do. In particular, the observables used in the estimation are HP-filtered output, annualized percentage change of CPI, and the average federal funds rate.[38] In line with Lubik and Schorfheide (2004), the paper considers the following sample periods: a pre-Volcker sample from 1960:Q1 to 1979:Q2 and a post-1982 sample from 1982:Q4 to 1997:Q4 that excludes the Volcker disinflation period. The findings read as follows.

First, in case of a fixed (zero) inflation target, the pre-Volcker period is explicitly characterized by indeterminacy, while determinacy prevails after 1982, basically replicating the findings of Lubik and Schorfheide (2004). The log-data densities are very similar to those reported in Lubik and Schorfheide (2004),[39] though this paper uses a different algorithm to estimate the DSGE framework over the entire region of the parameter space (Lubik and Schorfheide 2004 use standard MCMC techniques and they split the estimation separately over determinacy and indeterminacy regions). Second, when allowing for a drifting inflation target, determinacy prevails in the pre-Volcker period, which is in line with the benchmark results. Moreover, the model with a time-varying inflation target (determinacy) fits better than the one with a fixed target (indeterminacy). Again, these results raise the possibility that the Federal Reserve

[37] Hirose, Kurozumi, and Van Zandweghe (2020) find that replacing the standard NKPC with a GNKPC alters the estimated cofficients in the Taylor rule, in particular for the policy response to inflation.

[38] Note that, as in Lubik and Schorfheide (2004), the paper also estimates π^*. However, since the model of Lubik and Schorfheide (2004) is log-linearized around a zero-inflation steady state, the estimated π^* only appears in the measurement equation. As such, π^* only pertains to demeaning the inflation data used in the estimation and does not otherwise feed into the model dynamics through steady-state inflation (which is zero in the model), unlike the baseline model log-linearized around a non-zero steady-state inflation.

[39] See Table 2 on page 205 of their paper.

pursued a time-varying inflation target and possibly did not generate indeterminacy in the pre-Volcker period.

6.4 Calibrate ρ_{π^*}

Looking at the posterior distributions of the persistence of the inflation target process (ρ_{π^*}) in Table 1, the posteriors look quite similar to the prior. Hence, it seems that the data might not be sufficiently informative to pin down this parameter. As a result, the paper now calibrates ρ_{π^*}, while estimating the remaining parameters in the model as before.[40] Following Cogley, Primiceri, and Sargent (2010), ρ_{π^*} is set to 0.995. Alternatively, one may follow Ireland (2007) by assuming that the inflation target process has a unit root. Instead, the paper follows Cogley, Primiceri, and Sargent's (2010) calibration, as they show that a unit-root inflation target process counterfactually implies low inflation-gap predictability, which is at odds with the VAR evidence in their paper. A time-varying target continues to fit better than a constant target and determinacy prevails in both periods, as in the baseline estimations.

6.5 Calibrate ξ

The posterior distributions of ξ in Table 1 suggest that the estimated degree of price stickiness is relatively low. To ensure that the (in)determinacy results are not driven by a low degree of price stickiness, the paper calibrates ξ while estimating the remaining parameters of the model.[41] In particular, the degree of price stickiness is set to 0.75, which is a typical value used in calibration studies and the value used in Ascari and Sbordone (2014). A higher degree of price rigidity makes it increasingly difficult to eliminate indeterminacy. This is because when firms reset prices in the Calvo model, the weight placed on future profits depends on how likely it is for

[40]In the online appendix, I conduct identification analysis of the remaining parameters by first simulating data from the model with a time-varying target under determinacy and then estimating the model with the simulated data over the entire stable region of the parameter space (i.e., over both determinacy and indeterminacy). The results show that the estimation is able to recover the true parameter values for both the structural parameters and shocks, suggesting that the model parameters are relatively well identified.

[41]ρ_{π^*} is also set to 0.995 as above.

a firm not to alter its price by that period. Hence, greater price stickiness will increase the sensitivity of reset prices to expectations of future macroeconomic variables. As a result, a higher degree of price stickiness will widen the indeterminacy region for a given level of trend inflation. In fact, setting ξ to 0.75 implies a prior predictive probability of determinacy of about 27 percent, such that a priori it is more likely for indeterminacy to prevail.[42] As in the baseline analysis, the estimation finds that a time-varying target continues to fit better and determinacy prevails in the pre-Volcker period, despite the estimation being biased toward indeterminacy. In the post-1984 period, the fit of the model with a fixed target versus a time-varying target are quite similar and both favor determinacy.

6.6 Calibrate π^*

The paper conducts one final check. Recall that in the analysis so far, trend inflation (or steady-state inflation) and the time-varying inflation target are distinct features. There are two counteracting effects at work here. On one hand, the time-varying inflation target captures some of the low-frequency movements of inflation, so that there is less of a need for the richer dynamics characterized by the reduced form under indeterminacy. On the other hand, the presence of positive trend inflation widens the indeterminacy region of the parameter space. The paper finds that inflation target drifts higher during the Great Inflation period, making indeterminacy less likely, but trend inflation remains constant, so that indeterminacy region remains unaffected. However, this is not the case in Coibion and Gorodnichenko (2011), for example, where trend inflation increases during the Great Inflation period and expands the indeterminacy region. To address this issue, the paper estimates the GNK model with firm-specific labor in the pre-Volcker period while calibrating the steady-state inflation to a higher level and allowing for a time-varying inflation target.[43] In particular, trend inflation is set to 8 percent

[42]Recall that the prior predictive probability of determinacy in the baseline analysis is around 50 percent, such that, following the literature on testing for indeterminacy, the baseline estimations remain a priori unbiased.

[43]Firm-specific labor is assumed in order to maintain continuity with Coibion and Gorodnichenko (2011). Moreover, as discussed above, the long-run inflation elasticity of output implied by the GNKPC is more sensitive to trend inflation in

(annual level), which roughly corresponds to the highest estimate of Coibion and Gorodnichenko's (2011) time-varying trend inflation measure in the pre-Volcker period. The estimation continues to favor time-varying inflation target and determinacy prevails with a posterior probability of determinacy of 100 percent, while Coibion and Gorodnichenko (2011) find this probability to be zero with such a high level of trend inflation (see Figure 4 in their paper).[44]

7. Conclusion

This paper estimates a New Keynesian model with positive trend inflation while allowing for indeterminacy and time variation in the inflation target pursued by the Federal Reserve. The paper finds that inflation target has been drifting over time and, as a consequence, determinacy cannot be ruled out in the pre-Volcker period. The intuition for this result can be understood as follows. First, the inflation gap that enters the Taylor rule when the target is drifting over time is less volatile than the inflation gap with a fixed target. For a given historical path of the nominal interest rate, then the response of the nominal rate to the inflation gap turns out to be higher in the case of a time-varying target, which leads to determinacy. Second, inflation target shocks induce persistent responses in the inflation gap, as shown by Cogley, Primiceri, and Sargent (2010), which helps to capture the highly persistent inflation dynamics of the 1970s. As a result, the estimated model does not need to resort to the richer dynamics that arise under indeterminacy to explain the Great Inflation episode. One implication of this finding is that self-fulfilling inflation expectations, otherwise known as "sunspots," are not required to explain the high inflation outturns during this episode.

---

the model with firm-specific labor, thereby requiring a stronger response to inflation to guarantee determinacy for a given level of trend inflation relative to the model with homogenous labor. Nevertheless, we have also estimated the model with homogenous labor while calibrating trend inflation to 8 percent (annual level) and the results remain robust.

[44]The paper also estimates π^* with the priors centered around the average value for each sample period (instead of the average over the entire sample as in the baseline estimation) for both the homogenous and firm-specific labor model and the results remain robust.

The paper makes these arguments by assuming that trend inflation is positive but constant while the Federal Reserve pursues an exogenous time-varying inflation target. This choice helps to keep the analysis simple yet related and relevant to existing research. However, one could depart instead by log-linearizing the equilibrium conditions around a steady state characterized by drifting trend inflation, which would then result in a New Keynesian Phillips curve with drifting coefficients as in Cogley and Sbordone (2008). DSGE models with time-varying coefficients have been estimated by Fernández-Villaverde and Rubio-Ramírez (2008, 2010) and Fernández-Villaverde, Guerrón-Quintana, and Rubio-Ramírez (2010). I plan to pursue these lines of research in the future.

References

Arias, J. E., G. Ascari, N. Branzoli, and E. Castelnuovo. 2020. "Positive Trend Inflation and Determinacy in a Medium-Sized New Keynesian Model." *International Journal of Central Banking* 16 (3, June): 51–94.

Aruoba, B., and F. Schorfheide. 2011. "Sticky Prices versus Monetary Frictions: An Estimation of Policy Trade-offs." *American Economic Journal: Macroeconomics* 3 (1): 60–90.

Ascari, G. 2004. "Staggered Prices and Trend Inflation: Some Nuisances." *Review of Economic Dynamics* 7 (3): 642–67.

Ascari, G., P. Bonomolo, and H. F. Lopes. 2019. "Walk on the Wild Side: Temporarily Unstable Paths and Multiplicative Sunspots." *American Economic Review* 109 (5): 1805–42.

Ascari, G., A. Florio, and A. Gobbi. 2018. "High Trend Inflation and Passive Monetary Detours." *Economics Letters* 172 (November): 138–42.

Ascari, G., and T. Ropele. 2007. "Optimal Monetary Policy under Low Trend Inflation." *Journal of Monetary Economics* 54 (8): 2568–83.

———. 2009. "Trend Inflation, Taylor Principle, and Indeterminacy." *Journal of Money, Credit and Banking* 41 (8): 1557–84.

Ascari, G., and A. M. Sbordone. 2014. "The Macroeconomics of Trend Inflation." *Journal of Economic Literature* 52 (3): 679–739.

Basu, S., and J. G. Fernald. 1997. "Returns to Scale in U.S. Production: Estimates and Implications." *Journal of Political Economy* 105 (2): 249–83.

Belongia, M. T., and P. N. Ireland. 2016. "The Evolution of US Monetary Policy: 2000–2007." *Journal of Economic Dynamics and Control* 73 (December): 78–93.

Benati, L., and P. Surico. 2008. "Evolving US Monetary Policy and the Decline of Inflation Predictability." *Journal of the European Economic Association* 6 (2–3): 634–46.

———. 2009. "VAR Analysis and the Great Moderation." *American Economic Review* 99 (4): 1636–52.

Bhattarai, S., J. Won Lee, and W. Yong Park. 2016. "Policy Regimes, Policy Shifts, and US Business Cycles." *Review of Economics and Statistics* 98 (5): 968–83.

Bianchi, F., and G. Nicolò. 2017. "A Generalized Approach to Indeterminacy in Linear Rational Expectations Models." Mimeo, National Bureau of Economic Research.

Bilbiie, F. O., and R. Straub. 2013. "Asset Market Participation, Monetary Policy Rules, and the Great Inflation." *Review of Economics and Statistics* 95 (2): 377–92.

Bils, M. J., and P. J. Klenow. 2004. "Some Evidence on the Importance of Sticky Prices." *Journal of Political Economy* 112 (5): 947–85.

Blanchard, O. J., and J. Simon. 2001. "The Long and Large Decline in US Output Volatility." *Brookings Papers on Economic Activity* 2001 (1): 135–64.

Boivin, J., and M. P. Giannoni. 2006. "Has Monetary Policy Become More Effective?" *Review of Economics and Statistics* 88 (3): 445–62.

Calvo, G. A. 1983. "Staggered Prices in a Utility-Maximizing Framework." *Journal of Monetary Economics* 12 (3): 383–98.

Castelnuovo, E. 2010. "Trend Inflation and Macroeconomic Volatilities in the Post-WWII US Economy." *North American Journal of Economics and Finance* 21 (1): 19–33.

Castelnuovo, E., L. Greco, and D. Raggi. 2014. "Policy Rules, Regime Switches, and Trend Inflation: An Empirical Investigation for the United States." *Macroeconomic Dynamics* 18 (4): 920–42.

Clarida, R. H., J. Galí, and M. Gertler. 2000. "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory." *Quarterly Journal of Economics* 115 (1): 147–80.

Cogley, T., G. E. Primiceri, and T. J. Sargent. 2010. "Inflation-Gap Persistence in the US." *American Economic Journal: Macroeconomics* 2 (1): 43–69.

Cogley, T., and T. J. Sargent. 2005a. "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII US." *Review of Economic Dynamics* 8 (2): 262–302.

———. 2005b. "The Conquest of US Inflation: Learning and Robustness to Model Uncertainty." *Review of Economic Dynamics* 8 (2): 528–63.

Cogley, T., and A. M. Sbordone. 2008. "Trend Inflation, Indexation, and Inflation Persistence in the New Keynesian Phillips Curve." *American Economic Review* 98 (5): 2101–26.

Coibion, O., and Y. Gorodnichenko. 2011. "Monetary Policy, Trend Inflation, and the Great Moderation: An Alternative Interpretation." *American Economic Review* 101 (1): 341–70.

Doko Tchatoka, F., N. Groshenny, Q. Haque, and M. Weder. 2017. "Monetary Policy and Indeterminacy after the 2001 Slump." *Journal of Economic Dynamics and Control* 82 (September): 83–95.

Edmond, C., V. Midrigan, and D. Y. Xu. 2018. "How Costly Are Markups?" NBER Working Paper No. 24800.

Eggertsson, G. B., and M. Woodford. 2003. "The Zero Bound on Interest Rates and Optimal Monetary Policy." *Brookings Papers on Economic Activity* 2003 (1): 139–233.

Eo, Y., and D. Lie. 2020. "The Role of the Inflation Target Adjustment in Stabilization Policy." *Journal of Money, Credit and Banking* 52 (8): 2007–52.

Ettmeier, S., and A. Kriwoluzky. 2020. "Active, or Passive? Revisiting the Role of Fiscal Policy in the Great Inflation." DIW Discussion Paper No. 1872.

Farmer, R. E. A. 1999. *Macroeconomics of Self-fulfilling Prophecies.* Cambridge, MA: MIT Press.

Fernández-Villaverde, J., P. Guerrón-Quintana, and J. F. Rubio-Ramírez. 2010. "Fortune or Virtue: Time-Variant Volatilities Versus Parameter Drifting in U.S. Data." NBER Working Paper No. 15928.

Fernández-Villaverde, J., and J. F. Rubio-Ramírez. 2008. “How Structural Are Structural Parameters?” In *NBER Macroeconomics Annual 2007,* Vol. 22, ed. D. Acemoglu, K. Rogoff, and M. Woodford, 83–137 (chapter 2). Chicago: University of Chicago Press.

———. 2010. “Macroeconomics and Volatility: Data, Models, and Estimation.” NBER Working Paper No. 16618.

Friedman, M. 1968. “Inflation: Causes and Consequences.” In *Dollars and Deficits: Living With America’s Economic Problems*, 21–71. Englewood Cliffs, NJ: Prentice-Hall.

Fujiwara, I., and Y. Hirose. 2014. “Indeterminacy and Forecastability.” *Journal of Money, Credit and Banking* 46 (1): 243–51.

Groshenny, N. 2013. “Monetary Policy, Inflation and Unemployment: In Defense of the Federal Reserve.” *Macroeconomic Dynamics* 17 (6): 1311–29.

Haque, Q., N. Groshenny, and M. Weder. 2021. “Do We Really Know that U.S. Monetary Policy Was Destabilizing in the 1970s?” *European Economic Review* 131 (January): Article 103615.

Herbst, E., and F. Schorfheide. 2014. “Sequential Monte Carlo Sampling for DSGE Models.” *Journal of Applied Econometrics* 29 (7): 1073–98.

———. 2015. *Bayesian Estimation of DSGE Models*. Princeton, NJ: Princeton University Press.

Hirose, Y. 2007. “Sunspot Fluctuations under Zero Nominal Interest Rates.” *Economics Letters* 97 (1): 39–45.

———. 2008. “Equilibrium Indeterminacy and Asset Price Fluctuation in Japan: A Bayesian Investigation.” *Journal of Money, Credit and Banking* 40 (5): 967–99.

———. 2013. “Monetary Policy and Sunspot Fluctuations in the United States and the Euro Area.” *Macroeconomic Dynamics* 17 (1): 1–28.

———. 2020. “An Estimated DSGE Model with a Deflation Steady State.” *Macroeconomic Dynamics* 24 (5): 1151–85.

Hirose, Y., T. Kurozumi, and W. Van Zandweghe. 2020. “Monetary Policy and Macroeconomic Stability Revisited.” *Review of Economic Dynamics* 37 (July): 255–74.

Ireland, P. N. 2007. “Changes in the Federal Reserve’s Inflation Target: Causes and Consequences.” *Journal of Money, Credit and Banking* 39 (8): 1851–82.

Justiniano, A., and G. E. Primiceri. 2008. "The Time-Varying Volatility of Macroeconomic Fluctuations." *American Economic Review* 98 (3): 604–41.

Justiniano, A., G. E. Primiceri, and A. Tambalotti. 2010. "Investment Shocks and Business Cycles." *Journal of Monetary Economics* 57 (2): 132–45.

———. 2013. "Is There a Trade-Off between Inflation and Output Stabilization?" *American Economic Journal: Macroeconomics* 5 (2): 1–31.

Kass, R. E., and A. E. Raftery. 1995. "Bayes Factors." *Journal of the American Statistical Association* 90 (430): 773–95.

Kozicki, S., and P. A. Tinsley. 2005. "Permanent and Transitory Policy Shocks in an Empirical Macro Model with Asymmetric Information." *Journal of Economic Dynamics and Control* 29 (11): 1985–2015.

———. 2009. "Perhaps the 1970s FOMC Did What It Said It Did." *Journal of Monetary Economics* 56 (6): 842–55.

Kurozumi, T., and W. Van Zandweghe. 2017. "Trend Inflation and Equilibrium Stability: Firm-Specific Versus Homogeneous Labor." *Macroeconomic Dynamics* 21 (4): 947–81.

Leduc, S., and K. Sill. 2007. "Monetary Policy, Oil Shocks, and TFP: Accounting for the Decline in US Volatility." *Review of Economic Dynamics* 10 (4): 595–614.

Leigh, D. 2008. "Estimating the Federal Reserve's Implicit Inflation Target: A State Space Approach." *Journal of Economic Dynamics and Control* 32 (6): 2013–30.

Liu, J. S., and R. Chen. 1998. "Sequential Monte Carlo Methods for Dynamic Systems." *Journal of the American Statistical Association* 93 (443): 1032–44.

Lubik, T. A., and F. Schorfheide. 2003. "Computing Sunspot Equilibria in Linear Rational Expectations Models." *Journal of Economic Dynamics and Control* 28 (2): 273–85.

———. 2004. "Testing for Indeterminacy: An Application to US Monetary Policy." *American Economic Review* 94 (1): 190–217.

McConnell, M. M., and G. Perez-Quiros. 2000. "Output Fluctuations in the United States: What Has Changed Since the Early 1980's?" *American Economic Review* 90 (5): 1464–76.

Nakamura, E., and J. Steinsson. 2008. "Five Facts about Prices: A Reevaluation of Menu Cost Models." *Quarterly Journal of Economics* 123 (4): 1415—64.

Orphanides, A. 2004. "Monetary Policy Rules, Macroeconomic Stability, and Inflation: A View from the Trenches." *Journal of Money, Credit and Banking* 36 (2): 151–75.

Orphanides, A., and D. W. Wilcox. 2002. "The Opportunistic Approach to Disinflation." *International Finance* 5 (1): 47–71.

Primiceri, G. E. 2005. "Time Varying Structural Vector Autoregressions and Monetary Policy." *Review of Economic Studies* 72 (3): 821–52.

———. 2006. "Why Inflation Rose and Fell: Policy-makers' Beliefs and US Postwar Stabilization Policy." *Quarterly Journal of Economics* 121 (3): 867–901.

Sargent, T. J. 1999. *The Conquest of American Inflation.* Princeton, NJ: Princeton University Press.

Sargent, T. J., N. Williams, and T. Zha. 2006. "Shocks and Government Beliefs: The Rise and Fall of American Inflation." *American Economic Review* 96 (4): 1193–1224.

Sims, C. A. 2002. "Solving Linear Rational Expectations Models." *Computational Economics* 20 (1): 1–20.

Sims, C. A., and T. Zha. 2006. "Were There Regime Switches in US Monetary Policy?" *American Economic Review* 96 (1): 54–81.

Smets, F., and R. Wouters. 2007. "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach." *American Economic Review* 97 (3): 586–606.

Stock, J. H., and M. W. Watson. 2003. "Has the Business Cycle Changed and Why?" In *NBER Macroeconomics Annual 2002,* Vol. 17, ed. M. Gertler and K. Rogoff, 159–218 (chapter 4). MIT Press.

Zanetti, F. 2014. "Labour Market and Monetary Policy Reforms in the UK: A Structural Interpretation of the Implications." In *The Causes and Consequences of the Long UK Expansion: 1992 to 2007,* ed. J. S. Chadha, A. Crystal, J. Pearlman, P. Smith, and S. Wright, 81–110. Cambridge University Press.

Macroprudential Policy, Monetary Policy, and Euro Zone Bank Risk*

Elien Meuleman and Rudi Vander Vennet
Ghent University

We investigate the impact of macroprudential policy on the risk and return profile of euro zone banks between 2008 and 2018, conditioning on the stance of monetary policy. Using local projections, we find that a tightening in macroprudential policy increases financial stability by curbing credit growth and increasing the resilience of the banks. With respect to the policy mix, we show that tight macroprudential and monetary policies reinforce each other. But even when monetary policy is accommodating, macroprudential policy is found to be effective in deterring excessive bank risk-taking. However, we also document adverse consequences for bank franchise values.

JEL Codes: C23, E52, E61, G21, G28.

1. Introduction

Macroprudential policy is in vogue. The Great Financial Crisis amply demonstrated that microprudential regulation was insufficient to maintain the stability of banks. As a result, macroprudential policy has gained prominence in tackling the systemic risk of the banking industry. Whereas the objective of microprudential policy is to limit bank idiosyncratic risk, macroprudential policy attempts to improve financial stability from a systemic perspective (see, e.g., Crockett 2000, Borio 2003, and Caruana 2010). In the euro area,

*We gratefully acknowledge helpful comments from Martien Lamers, Olivier De Jonghe, Hans Degryse, Koen Schoors, Selien De Schryder, Gert Peersman, Ralph De Haas, an anonymous referee, and seminar participants at Ghent University. Author contact: Department of Economics, Ghent University, Sint-Pietersplein 5, 90090 Ghent, Belgium. Author e-mails: Elien.Meuleman@Ugent.be and Rudi.VanderVennet@Ugent.be.

various macroprudential tools, both bank based and borrower based, have been introduced in a period characterized by the active use of monetary policy tools by the European Central Bank (ECB). And while monetary policy and macroprudential policy have their own objectives, i.e., price stability and financial stability, there are several channels through which one policy can influence the objective of the other. This naturally raises an important policy question: is the transmission of macroprudential policy different conditional on the stance of monetary policy? The interactions can enhance or reduce the effectiveness of each policy in achieving its objective and may therefore suggest the need for coordination (Smets 2014).

We empirically analyze this question for the euro-area banking system. In particular, we examine whether or not the effectiveness of macroprudential policy is influenced by the stance of ECB monetary policy. To do this, we perform an in-depth investigation of the transmission of macroprudential policy shocks to the banks in the euro area. Our empirical analysis proceeds in different stages. First, we investigate the macroprudential transmission channels by assessing the impact of macroprudential policy on a broad set of bank risk and return profile variables that capture the resilience of the banking system. Second, as different types of macroprudential measures are expected to produce different effects depending on a bank's business model, we investigate whether the transmission of macroprudential policy is heterogeneous across different bank business models. Ultimately, we interact the macroprudential policy shock with our measure of the monetary policy stance to understand how macroprudential policies transmit to the banks' risk and return profile. We focus on the behavior of euro-area banks from 2008 to 2018, which is the period characterized by different stages of conventional and unconventional monetary policy by the ECB and which coincides with the introduction of various types of macroprudential policy in euro-area countries. Throughout the empirical analysis we use the local projections framework of Jordà (2005), which allows us to visually assess how the banks' risk and return profile is affected by macroprudential and monetary policy shocks, their interaction, whether or not these responses differ across banks, and the persistence of these effects over time.

We aim to contribute to the literature in different ways. First, we use granular bank-level data and incorporate a wide range of bank risk and return profile variables constructed with both accounting and market data, which distinguishes us from papers that use a limited set of bank variables. Second, for the construction of a macroprudential index we make use of a new database collected by experts at the ECB and national banks. This MacroPrudential Policies Evaluation Database (MaPPED) contains information on almost 2,000 macroprudential actions taken in 28 member states of the European Union. The database differs from other databases (for example, Lim et al. 2011 and Cerutti, Claessens, and Laeven 2017, among others) since it not only indicates the activation of a certain policy tool, but it also tracks the tool over time by including, for example, changes in the level or the scope of the tool. Also, where other databases have a rather limited tool coverage, this database contains information on 53 different types of policy tools. The database ensures the comparability across measures and across countries, which is one of the major drawbacks when using other existing databases (Budnik and Kleibl 2018). We assess the impact of macroprudential policy on a set of bank risk and return profile measures using a novel identification strategy that only recently has been used in economics to assess the effectiveness fiscal policy (Jordà and Taylor 2016) and macroprudential policy (Richter, Schularick, and Shim 2018; Alam et al. 2019). More specifically, we use the inverse propensity score weighting methodology as an identification strategy to re-randomize the sample of the treatment and the control group which allows us to mitigate endogeneity concerns. Third, to construct the monetary policy stance, we estimate a structural vector autoregression (SVAR) to extract an exogenous monetary policy shock. This monetary policy shock is identified by assuming that its variance increases on days on which there is a monetary policy announcement. This "identification-through-heteroskedasticity" approach yields monetary policy shocks that account for the prevailing macroeconomic and financial markets conditions, which determine the behavior of banks and the market assessment of their risk and return profile. Fourth, we add to the extant literature by exploring the interaction between monetary policy and macroprudential policy. Evidence concerning these interactions is rather limited and mainly comes from theoretical (DSGE) modeling rather than empirical analysis. In this paper we complement the literature with an in-depth empirical analysis of how different macroprudential policies affect the banking system and how they interact with monetary policy in the euro area.

Our main findings can be summarized as follows. Considered in isolation, we confirm that macroprudential policy is effective in restraining bank risk, as intended by the macroprudential authorities. Tightening macroprudential measures are typically associated with less lending and lower bank asset risk, and these features translate into lower overall bank risk, both accounting based and market based. However, the downside is that the announcement of macroprudential tools is accompanied by lower bank profitability over the projection horizon, which indicates that imposing constraints on banks causes lower current and future bank profitability. When considering the banks' business model, we find that for both lending and profitability the effects are more pronounced for retail banks than for their non-retail counterparts. This is not unexpected since banks with a retail profile are most active in traditional lending, which is the focus of macroprudential measures targeting credit growth. The negative consequences on profitability are also more pronounced for retail-oriented banks, which may affect their future viability.

Ultimately, we assess whether the effectiveness of macroprudential policy with respect to bank risk and return profiles is different conditioning on the monetary policy stance. We find that macroprudential policy and monetary policy push credit growth in the same direction, i.e., they reinforce each other. In other words, the effectiveness of macroprudential policy with respect to bank credit growth is stronger when monetary policy is also in a tight phase. Conversely, when macroprudential policy is tight but monetary conditions are accommodating, loan growth increases, suggesting that the transmission of macroprudential policy to credit growth is affected by the presence of loose monetary policy. Interestingly, while accommodating monetary policy may entail incentives for banks to take more risk, our results indicate that macroprudential measures were sufficiently strong to deter banks from excessive risk-taking. In other words, macroprudential policy succeeds in maintaining bank stability also in periods of monetary accommodation. Yet, there is an important downside: we observe a marked deterioration of the banks' market-to-book value as a reflection of the investors' conviction that low-for-long interest rates ultimately compress bank interest margins and put their profitability and franchise value under stress. Our conclusion is that the combination of restrictive macroprudential policies and prolonged monetary accommodation may turn out to be detrimental for bank health and, ultimately, financial stability.

Our main findings are corroborated when we estimate the monetary policy stance with a Taylor rule or when we use the "identification-through-external-instruments" approach. When we consider the impact of specific macroprudential policy tools, we find that credit growth measures, such as loan-to-value ratios, have an immediate and stronger negative impact on loan growth than liquidity regulation or measures aimed at the resilience of banks, such as capital regulation. However, we also find evidence for risk-shifting behavior by banks confronted with targeted credit measures: banks increase the riskiness of the loan portfolio in response to credit constraints. In trying to comply with the rules, these banks may engage in riskier activities by, e.g., shifting to more risky corporate lending or securities.

The paper proceeds in the following way. In Section 2 we review the extant literature, analyze the transmission channels of macroprudential policy, and develop our hypotheses. Section 3 describes the empirical setup we use to assess the effectiveness of macroprudential policy, both unconditional and conditional, on the stance of monetary policy. Section 4 presents the data and the selection of the sample. In Section 5 we analyze the empirical results followed by several robustness checks in Section 6. Section 7 concludes.

2. The Transmission of Macroprudential Policy

Monetary and macroprudential policies are intended to modify banks' behavior by constraining credit supply and demand. Hence, both policies may affect banks through similar transmission channels. The question thus arises how they may influence each other's effectiveness in reaching their respective objectives. The interaction between both policies can either strengthen or weaken the effectiveness of each policy in achieving its goal. In this paper we assess the impact of macroprudential policy and investigate whether or not the transmission is different conditional on the stance of monetary policy.

Macroprudential policy actions are intended to affect the balance sheet of financial institutions and to enhance financial stability. For example, banks may respond to a tightening in capital requirements by issuing more equity, by increasing retained earnings, by deleveraging, or by de-risking. All of these strategies should increase the loss-absorbing capacity of the banks and create an extra buffer in the case of unexpected losses. Liquidity-based tools force banks to hold more liquid assets or increase long-term funding, which increases the resilience of banks to unforeseen liquidity shocks. Banks can also react to tighter liquidity regulations by decreasing their lending portfolio, which also affects their resilience to adverse conditions. Borrower-based tools such as loan-to-value ratios or debt-to-income ratios affect the lending capacity of banks and should reduce the probability of default of the borrowers, which improves the stability of the bank. Macroprudential tools such as limits on certain exposures or higher risk weights on specific asset classes affect the loan supply and make banks less sensitive to shocks in, e.g., real estate markets. All these transmission channels decrease the banks' risk profile, which should limit the occurrence of systemic crises.

Existing empirical work shows that macroprudential policy is capable of smoothing the financial cycle. Lim et al. (2011) evaluate the effectiveness of different macroprudential instruments on credit growth, systemic liquidity, leverage, and capital flows. They use International Monetary Fund (IMF) survey data containing information on macroprudential instruments used in 49 countries during a 10-year period from 2000 to 2010. They find that many of the instruments used are effective in reducing procyclicality. Shim et al. (2013) investigate the impact of macroprudential tools on housing credit and housing prices using a database for policy actions covering 60 economies worldwide from 1990 to 2012. The authors find evidence that mainly the debt-service-to-income requirements and housing-related taxes can be used as tools to restrain housing credit growth. In contrast, supply-side credit policies such as risk weights and provisioning requirements had no significant impact on housing credit. Cerutti, Claessens, and Laeven (2017) use an

IMF survey, Global Macroprudential Policy Instruments (GMPI), to investigate the impact of 18 different policy instruments on credit growth in 119 countries over the period 2000 to 2013. They find that the policy tools are effective in reducing credit growth, yet the effects are more pronounced in emerging economies. Akinci and Olmstead-Rumsey (2018) use a combination of IMF survey data, Bank for International Settlement (BIS) data, and information received from national central banks and financial authorities to analyze the influence of macro policies on credit growth and housing prices. Using a dynamic panel setting, they find that tightenings in macroprudential tools are associated with lower credit growth and housing prices. Igan and Kang (2011) make use of a regional database to examine the effect of loan-to-value and debt-to-income limits on house price dynamics, residential real estate market activity, and household leverage in Korea. They find evidence that loan-to-value and debt-to-income tools are indeed associated with both a decline in house prices and a drop in the number of transactions. Dell'Ariccia et al. (2016) find that, for a large cross-country data set covering 170 countries over the period 1970–2010, macroprudential tools are effective in reducing the emergence of credit booms and the costs associated with credit busts, in contrast to monetary and fiscal policies. Meuleman and Vander Vennet (2020) investigate whether macroprudential policy is able to support financial stability by tackling the interconnectedness of banks for a sample of euro zone banks between 2000 and 2017. They find that liquidity and capital regulation is able to address the systemic linkage of banks, while credit growth tools and exposure limits have more impact on the individual risk of banks. In general, most empirical studies conclude that macroprudential policy tools achieve their stated objective, although some tools appear to be more effective than others.

Evidence on the interactions between monetary policy and macroprudential policy is still scarce and mainly comes from theoretical (DSGE) modeling rather than empirical analysis and focuses on whether the macroprudential and monetary policymakers should cooperate or not (see, for example, Angelini, Neri, and Panetta 2014, Gelain and Ilbas 2017, and Paoli and Paustian 2017). Most papers find that after a financial shock, when policies cooperate, both types of policy should work in the same direction, i.e., they complement each other. This paper empirically adds to this discussion as we investigate whether or not the effectiveness of macroprudential policy is affected by the stance of monetary policy in the euro area. Rubio and Carrasco-Gallego (2014) analyze the interactions between a macroprudential loan-to-value rule and a monetary policy Taylor rule in a DSGE model with housing and collateral constraints. They find that the actions of both policies unambiguously improve the stability of the system. Martinez-Miera and Repullo (2019) find that both tight macroprudential policy, in the form of binding capital requirements, and tight monetary policy individually reduce risk-taking; however, when the two policies are interacted, bank risk-taking increases as the transmission of monetary policy to the loan rates is affected by the presence of binding capital regulation. With respect to empirical evidence, Aiyar, Calomiris, and Wieladek (2016) find that tightening monetary policy and increasing banks' capital requirements both have negative effects on bank credit supply, and that there is no interaction between changes in monetary policy and changes in capital requirements. On the other hand, Tressel and Zhang (2016) use an interaction term between the monetary policy stance and an LTV indicator and find that LTV constraints tend to be more effective in containing credit growth and house price appreciation when monetary policy is loose. Gambacorta and Murcia (2019) use granular credit registry data of five Latin American countries and find that macroprudential policy and monetary policy reinforce each other by pushing in the same direction. David et al. (2019) confirm these results as they find benefits of synchronization between macroprudential and monetary policies using a panel data setting for a sample of 37 emerging and advanced economies.

3. Methodology

The overarching research question of this paper is to investigate the effectiveness of macroprudential policy conditional on monetary policy. To tackle this question, our empirical investigation proceeds in two stages. We first focus on the standalone effect of macroprudential policy on the bank risk and return profile variables, we identify macroprudential actions based on the MaPPED database, and we explain how we use the inverse propensity score approach to analyze the impact of macroprudential policy on bank risk and return profiles. We also check potential heterogeneous effects of these macroprudential measures across bank business models (in Subsection 3.1). Second, we identify the monetary policy stance based on an identification-through-heteroskedasticity approach in order to investigate the impact of a macroprudential shock across different monetary policy regimes (in Subsection 3.2).

3.1 Macroprudential Policy and the Bank's Risk and Return Profile

As a first step in the analysis, we need information on the macroprudential actions that have been initiated in the euro zone. We use the granular information available in the MacroPrudential Policies Evaluation Database (MaPPED), which has been collected by experts at the ECB and the national central banks (Budnik and Kleibl 2018). MaPPED provides details on 1,925 macroprudential (or similar) policy actions between 1995 and 2018 in the 28 member states of the European Union. The tools are subdivided into 11 categories: capital buffers, lending standards, maturity mismatch tools, limits on credit growth, exposure limits, liquidity rules, loan loss provisions, minimum capital requirements and risk weights, leverage ratio, and other measures (this category contains mainly crisis-related measures and resolution tools). The MaPPED survey is designed in such a way that respondents can only choose from a closed list of policy tools, in contrast to open-text questionnaires as in Lim et al. (2011) or the GMPI. These features ensure that the comparability across measures and across countries is maintained, which is one of the major drawbacks when using other existing databases (Budnik and Kleibl 2018).

MaPPED tracks every measure over time, indicating not only the activation date but also changes in the scope or the level of the measure over time, as well as the deactivation of the measure. We use the announcement date of each tool to analyze how banks react to the macroprudential policy changes using impulse response functions (IRFs) over a horizon of eight quarters.[1] Each policy action is

[1]We use the announcement date rather than the enforcement date, as we hypothesize that market participants and banks immediately respond to changes

classified as a loosening action, a tightening action, or an action with an ambiguous impact. We construct an overall indicator of macroprudential policy based on this MaPPED database. First, individual policy instruments are each coded as 1 in the quarter they are announced and 0 otherwise. An activation and a change in the scope or level of a tool are all coded as 1. Measures with an ambiguous impact are conservatively coded as 0. An overall macroprudential policy indicator is the sum of the scores on all 11 individual policies.

In terms of establishing the effect of macroprudential policy actions on the banks' risk and return profile, the main challenge is tackling the endogeneity issue. Reverse causality can be a problem in our context because macroprudential policy actions are more likely to be tightened during periods of high credit growth and increasing bank risk. Therefore, estimations that do not address the issue may be subject to a measurement error. We employ an inverse propensity score weighted (IPW) estimator specifically designed for our purposes. Propensity score methods have been originally used in biostatistics and medicine (see, for example, Rosenbaum and Rubin 1983 or Austin 2009, among others). More recently, they have been applied in economics to assess the effectiveness fiscal policy (Jordà and Taylor 2016) and macroprudential policy (Richter, Schularick, and Shim 2018; Alam et al. 2019). The IPW estimator alleviates endogeneity issues by penalizing those observations that are likely to be affected by reverse causality. More specifically, an IPW estimator gives more weight to those observations that are difficult to predict based on a set of macrovariables that are used by regulators to initiate macroprudential policy tools, and less weight to those macroprudential actions that are easy to predict based on the macrovariables. The methodology is particularly well suited to analyze macroprudential policy since the macroprudential regulator indeed uses indicators (for example, the credit-to-GDP gap for the initiation of the countercyclical buffer or housing credit/prices for credit growth measures) to initiate macroprudential policy.[2] In

in the macroprudential policy stance in the quarter of announcement, even before the tool is in force.

[2]The IPW methodology comes close to the propensity score matching technique as used in Forbes et al. (2015). We believe, however, that using the IPW technique results in more reliable results than when we use the propensity score

practice, we first specify a logit model at the country level to estimate the probability that a certain macroprudential policy tool is activated. Let $D_{j,t}$ be a tightening dummy in country j that takes on a value of 1 when a macroprudential policy action is announced in a certain quarter (or when multiple actions are announced) and 0 otherwise:

$$log\left(\frac{D_{j,t}=1|Z_{j,t-1}}{D_{j,t}=0|Z_{j,t-1}}\right)=\alpha_j+\lambda_{year}+\beta Z_{j,t-1}+\varepsilon_{j,t}. \qquad (1)$$

$Z_{j,t-1}$ is a vector of macroeconomic controls at the country level j lagged one quarter. We include the country's total bank loan growth, the change in housing prices, the growth in household debt to GDP, GDP growth, the VSTOXX, and the ECB policy rate. We also include country and year fixed effects. We refer to the probability of a tightening as the propensity score, and its estimate from Equation (1) is denoted by $\widehat{p_{j,t}}$.

In a second stage, we fit the probabilities for the logit model at the country level using regression weights given by the inverse of $\widehat{p_{j,t}}$. Weighting by the inverse of the propensity score puts more weight on those observations that were difficult to predict and thereby re-randomizes the treatment. In our application, this implies putting more weight on macroprudential tightenings that were considered as a surprise based on observed data, and putting less weight on those tightenings that could be predicted. We convert the country probabilities to the bank-level setting by assigning each bank situated in a specific country the same probabilities. With the fitted probabilities we can now estimate the cumulative responses of a shock in the macroprudential index on the change in the bank risk and

matching technique because we would lose a lot of observations, as we would only match each treated observation with one matched control observation. The matching technique does not take into account other control observations, and the control group is shrunk down to the same size as the treatment group. In contrast to the propensity score matching technique, the IPW matching occurs in both directions: from control to treated and from treated to control. That is, each observation is given weight of the inverse of the probability of the treatment they actually got so we do not lose observations. Intuitively, treatment cases that resemble the controls are interesting and given more weight, and control cases that look like they should have got the treatment also receive a higher weight.

return profile measures between 2008 and 2018 with the following local projections model using weighted least squares (WLS) as in Richter, Schularick, and Shim (2018) and Alam et al. (2019):

$$\Delta y_{i,j,t+h} = \alpha_i^h + \gamma_t^h + \beta^h D_{j,t} + \sum_{k=0}^{K} \Theta_k^h Bank_{i,t} + \sum_{l=0}^{L} \Gamma_l^h Macro_{j,t} + \varepsilon_{i,t+h}. \tag{2}$$

$\Delta y_{i,j,t+h}$ denotes the percentage change in the risk and return profile variables for bank i in country j between time t and $t+h$. $D_{j,t}$ corresponds to the tightening dummy in country j at time t. α_i^h and γ_t^h denote the bank fixed effects and the time fixed effects, respectively. The coefficient of interest is β^h, which captures the impact of a macroprudential change at time t on the bank risk and return profile variables at horizon h.[3] We expect this coefficient to be negative for the banks' risk variables since macroprudential policy tools are aimed at increasing bank stability. The variable *Bank* represents a vector of bank business model characteristics. *Macro* corresponds to the macroprudential policy indicators, which we also use in the propensity score model. We include the country's loan growth, the change in housing prices, the growth in household debt to GDP,

[3] We argue that the variables on the right-hand side are predetermined and serve as a benchmark so that $\Delta y_{i,j,t+h}$ can be seen as the deviation in Y from the expectation at time $t+h$ based on the information available at time $t-1$. If this would not be the case, the deviation in Y can also be due to (endogenous) changes in the covariates, which we want to avoid. This approach only allows us to determine the direct impact of a shock to $\Delta y_{i,j,t+h}$ rather than indirect effects through other variables. More specifically, the IRFs thus only capture the impact of a macroprudential shock at time 0, assuming all else equal over each horizon of the IRF. The monetary policy stance can however vary after the macroprudential policy tightening was announced. If the monetary policy stance changes along the impulse response estimation horizon, then the model would capture the monetary policy environment imprecisely, which can affect the results. In an attempt to address this concern, we perform a robustness check where we also consider the future stance of monetary policy by averaging the monetary policy shocks over the projection horizon of eight quarters. We find that the local projections yield the same main conclusions. If anything, the impact on net loan growth and the MES are even more pronounced when using the forward-looking monetary policy stance.

and GDP growth. The weights that are used in this weighted least squares estimation are defined by $w_{j,t} = \frac{D_{j,t}}{\widehat{p_{j,t}}} + \frac{1-D_{j,t}}{1-\widehat{p_{j,t}}}$, where we truncate $w_{j,t}$ at 10 to avoid extreme weights. These weights are consistently used in all model specifications. For all the impulse responses in the analysis, we use a horizon of eight quarters.

Different types of macroprudential measures are expected to produce different effects depending on a bank's business model. Therefore we allow for heterogeneous impulse responses across bank types. Several papers have attempted to classify banks into business models based on various statistical approaches, typically yielding between four and seven business model types (see Kok, Móré, and Petrescu 2016, Farnè and Vouldis 2017, and Roengpitya et al. 2017). However, the differences between the business models are often qualitative in nature. Therefore we opt for a parsimonious subdivision of the banks based on a limited number of observable bank balance sheet indicators. To do this, we perform a factor analysis on the bank characteristics $Bank_{i,t}$ (the loan ratio (LTA), the ratio of customer deposits to total liabilities (DEP), the ratio of total equity to total assets (CAP), the share of non-interest income in total income (DIV), and bank size (SIZE)). If there is common variance, this will be reflected by factors associated with eigenvalues above 0. The higher the eigenvalue, the more the factor is able to explain common variance. This implies that factors with low eigenvalues are less likely to reflect the broad common strategies that we relate to bank business models. Table 1 presents the results of the factor analysis.

The first factor, which explains 63 percent of all variation, is associated with a retail-based strategy. Therefore we label this factor as *RETAIL,* as it is a vector that captures the *retailness* of a bank.[4] It positively relates to the loan, deposit, and capital ratios, but is negatively related to size and income diversification. The higher the factor score, the more retail oriented the bank is. The subdivision in retail versus non-retail banks has intuitive appeal for our research question since many macroprudential measures are targeted to a specific type of bank (e.g., countercyclical capital buffers or lending restrictions in the form of LTV caps are designed to primarily

[4]We acknowledge that the labeling of factors is always somewhat subjective. In this paper, the choice for the label follows Mergaerts and Vander Vennet (2016).

Table 1. Results of Factor Analysis on a Number of Bank Business Model Characteristics

Factor	Eigenvalue	Proportion	Cumulative
Factor 1	1.84	0.63	0.63
Factor 2	0.94	0.33	0.96
Factor 3	0.09	0.03	0.99
Factor 4	0.03	0.01	1.00
Factor 5	0.00	0.00	1.00
	Correlation with Characteristics		**Communality**
SIZE	–0.87		81%
LTA	0.39		8%
DEP	0.68		46%
DIV	–0.29		4%
ETA	0.61		35%

Note: The factor analysis is conducted using the iterated principal factor method. The upper panel displays the eigenvalues of the common factors. The lower panel reports correlations of the predicted factors with the observed variables and the communality associated with each variable. A higher communality indicates that the variable is better explained by the common factors.

affect retail banks). We use the first factor obtained through the factor analysis in interaction with the macroprudential policy shock to assess heterogeneous effects across banks. Concretely, we estimate the following local projections model:

$$\Delta y_{i,t+h} = \alpha_i^h + \gamma_t^h + \beta^h D_{j,t} + \chi^h BBM_{i,t} + \pi^h D_{j,t} \times BBM_{i,t} + \sum_{k=0}^{K} \Theta_k^h Bank_{i,t} + \sum_{l=0}^{L} \Gamma_l^h Macro_{j,t} + \varepsilon_{i,t+h}. \quad (3)$$

$\Delta y_{i,t+h}$ denotes the percentage change in the risk and return profile variables for bank i between time t and $t+h$. $D_{j,t}$ corresponds to the macroprudential shock. α_i^h are the bank fixed effects. When we estimate the impact of macroprudential policy, we also include time fixed effects, γ_t^h, and estimate the model with weighted least squares, again using the weights as defined by the logit model in Equation (1), i.e., $w_{j,t} = \frac{D_{j,t}}{\widehat{p_{j,t}}} + \frac{1-D_{j,t}}{1-\widehat{p_{j,t}}}$. $BBM_{i,t}$ stands for the first factor of the factor analysis which distinguishes between retail and non-retail banks. The differential impact between retail and non-retail banks can then be calculated as the partial derivative of the bank risk and return profile variables with respect to the shock.[5]

3.2 Interactions between Monetary and Macroprudential Policy

Ultimately, we want to investigate the impact of macroprudential shocks conditional on the stance of monetary policy. Macroprudential policy is implemented by national authorities, while monetary policy is determined at the ECB level. Hence, national macroprudential policies have to take the stance of monetary policy as given. The important policy issue is whether or not the effectiveness of macroprudential policy depends on monetary policy conditions. For the identification of the monetary policy stance in the euro zone in the post-2008 period, we cannot use the policy rate because of the zero lower bound constraint. Similarly, the ECB balance sheet cannot be used because some important monetary policy measures did not affect the balance sheet (e.g., OMT was pre-announced by the Draghi "whatever it takes" speech in July 2012, operationally implemented in September 2012, but subsequently never activated). And finally, different conventional and unconventional policy measures were announced simultaneously (e.g., in January 2015, PSPP was announced jointly with a decrease in the deposit facility rate and strengthened forward guidance) and were often largely anticipated.

[5]In an estimation setup with interaction terms, the full effect is measured as the partial derivative of the bank risk and return profile variables with respect to the shock, which boils down to the sum of the standalone effect and the coefficient on the interaction term times the business model factor. The impulse responses are constructed as follows:

$$\frac{\partial \Delta y_{i,t+h}}{\partial D_{j,t}} = \hat{\beta}^h + \hat{\pi}^h BBM_{i,t}, \tag{4}$$

where $BBM_{i,t}$ corresponds with the *RETAIL* factor obtained through the factor analysis. From Equation (4) it is clear that we have impulse response functions that vary at the bank level. We therefore calculate the average impulse response corresponding to the 25 percent highest *RETAIL* factor scores (retail banks), and the average impulse response corresponding to the 25 percent lowest *RETAIL* factor scores (non-retail banks).

Based on the survey of econometric approaches used to identify monetary policy shocks in Rossi (2019), we opt for the SVAR because this approach allows us to incorporate a broad set of financial market indicators that should be linked to the decisions that banks make in terms of lending behavior, loan pricing, and the riskiness of their loan portfolio. These strategic choices should be reflected in the accounting-based and the market-based variables we use to capture the banks' risk and return profile (loan growth, loan risk, interest margin) as well as in their perceived profit potential (market-to-book value).

We estimate a time series of exogenous monetary policy shocks by modeling a set of relevant financial market variables in a structural VAR (SVAR) model at daily frequency as in Wright (2012) and Lamers et al. (2019):

$$Y_t = A_1 Y_{t-1} + \cdots + A_p Y_{t-p} + R\nu_t, \tag{5}$$

where Y_t is an N-dimensional vector of endogenous variables ($t = 1, \ldots, T$), ν_t an N-dimensional vector of orthogonal structural innovations with mean zero, and $A_1, \ldots, A_p$ and R are $N \times N$ time-invariant parameter matrices. The reduced-form residuals corresponding to this structural model are given by the relationship $\varepsilon_t = R\nu_t$.

To estimate the SVAR we use a set of variables that capture the pass-through of monetary policy to the financial sector. Following Rogers, Scotti, and Wright (2014), we select those variables that are expected to respond most to a monetary policy shock. More specifically, we include the 10-year German government bond yield, the 5-year forward inflation expectation based on inflation swap rates, an EU market index, the 5-year Spanish CDS spread, and the VSTOXX index.[6] Data are obtained through Thomson Reuters'

[6]The rationale for using the Spanish five-year CDS spread is that Spain is the prototypical euro-area periphery country which was hit by the banking crisis, a real estate crisis, and the sovereign crisis and it was not rescued with loans from the EFSF/ESM (compared with, e.g., Portugal, Ireland, or Greece). However, as a robustness check we also experimented with other sovereign stress indicators: the five-year CDS spreads of Italy and France, an index of European five-year sovereign CDS spreads, and an index based on the five-year CDS spreads of Spain, Portugal, Italy, and Ireland. Our findings do not appear to be driven by the choice of the sovereign stress indicator.

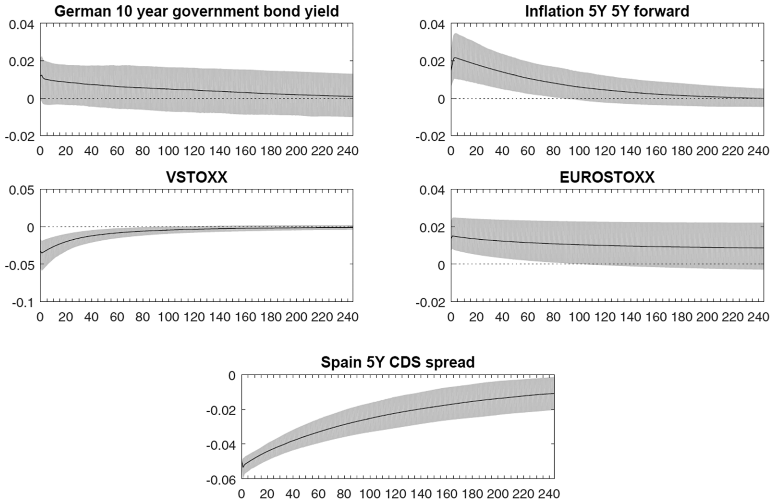
Datastream. The identification of policy shocks is based on the identification-through-heteroskedasticity strategy first proposed by Rigobon (2004), which assumes that the structural monetary policy shock is more volatile on monetary policy announcement days. The main idea is that there are days on which the volatility of the monetary policy shock is especially high, i.e., on days when there is a ECB announcement. Based on the differences in the volatility of the shocks during the two regimes, the structural VAR can uniquely be identified. In essence, we only assume that there is some kind of heteroskedastic pattern in the monetary policy shock while all other shocks are homoskedastic:

$$Var\left(\nu_t\right) = \Omega_t = \begin{cases} \Omega^{(0)} = (\omega_1, \omega_2, \ldots, \omega_N) & \text{if } \textit{no announcement} \\ \Omega^{(1)} = (\omega_1^*, \omega_2, \ldots, \omega_N) & \text{if } \textit{announcement.} \end{cases} \tag{6}$$

It can be shown that, as long as the covariance matrix of the reduced-form errors V_t changes on announcement days, these assumptions suffice to uniquely identify the first column of R and the structural monetary policy shock apart from their scale and sign. The model can be estimated following the iterative estimation procedure outlined in Lanne and Lütkepohl (2008).[7] We normalize the monetary policy shock by fixing the response on impact of one of the included variables to a unit monetary policy shock. We define a unit expansionary monetary policy shock as a shock that decreases five-year Spanish CDS spread by 5 percent. The set of days with monetary policy announcements is determined prior to the estimation of the SVAR model. This identification-through-heteroskedasticity approach is widely used in the literature to identify monetary policy shocks—for example, Caporale, Cipollini, and Demetriades (2005), Gilchrist and Zakrajŝek (2013), Rogers, Scotti, and Wright (2014), and Arai (2017). We estimate a VAR of order 2 over a sample period from October 1, 2008 to December 31, 2018, i.e., the period during which the ECB implemented various types of conventional and unconventional monetary policy. The impulse responses are shown in Figure 1.

[7]For details on this estimation procedure we refer to Lamers et al. (2019).

Figure 1. Impulse Response Function of the Variables to a Unit Monetary Policy Shock



Note: Gray areas represent 68 percent confidence intervals that are obtained through a stationary bootstrap with expected block length 10 for non-announcement days. Announcement-day residuals are bootstrapped separately. The horizontal axis represents the horizon of the impulse response function in working days, i.e., the IRFs are plotted for a horizon of 240 days.

We find that an expansionary monetary policy shock increases long-term inflation expectations and the value of the broad stock market index, while decreasing market-wide implied volatility ($VSTOXX$). Although the negative contemporaneous impact on the five-year Spanish CDS is a consequence of our identification strategy, the effect remains significantly negative across the whole horizon. We do not observe a significant impact on the yield of the long-term safe asset, possibly due to a flight-to-safety effect in which monetary easing lowers the demand for safe assets, such as German bunds, by decreasing the risk of stressed sovereign bonds (see also Rogers, Scotti, and Wright 2014 and Altavilla, Giannone, and Lenza 2016).

To capture the stance of monetary policy, we could simply take the cumulative sum of the structural monetary policy shock over time. We would however ignore monetary policy shocks that occurred in the past which may still have an impact on the financial variables in the present time. In addition, as the average of the structural shock is zero by construction, the cumulative sum of the structural shocks will mechanically converge to zero at the end of the sample period. To avoid this, we perform a historical decomposition on the data as in Peersman and Smets (2003). Historical decompositions capture the accumulated effects of a structural shock on the VAR variables during a number of periods.[8]

We compute the contribution of the monetary policy shock to changes in the Spanish CDS spread in Figure 2.[9] We multiply the series with –1 so that we can interpret the monetary policy stance as accommodating when the series is positive, which means that monetary policy decreased the Spanish five-year CDS spread. More specifically, a sequence of positive monetary policy shocks indicates that monetary policy is becoming more expansionary and therefore the cumulative series reflects the monetary policy stance with respect to the prevailing economic environment and expectations of financial markets. As a consequence, a drop in the series can reflect a tightening of monetary policy but also the lack of monetary action

[8]Intuitively, a historical decomposition converts the time series Y_t as described in Equation (5) in different components. In particular, the time series Y_t are linear functions of the history of structural shocks and an exogenous component which captures the initial conditions of the time series and the steady state. We can write the time series Y_t as follows:

$$Y_t = \underbrace{\sum_{j=0}^{t-(p+1)} A^j HC\varepsilon_{t-j}}_{\text{Contribution from shocks}} + \underbrace{A^{t-p}Y_p + \sum_{j=0}^{t-(p+1)} A^j H\mu}_{\text{Initial conditions and steady state}} ,$$

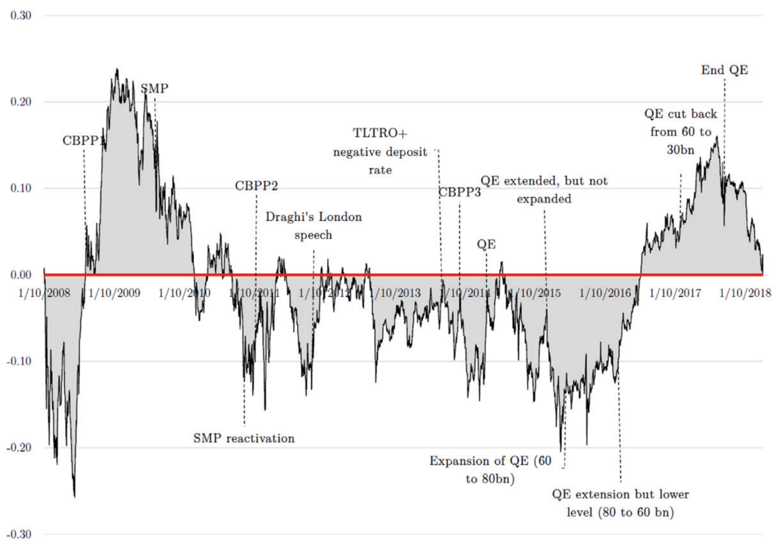
where

$$Y_t = \begin{bmatrix} y_t \\ . \\ . \\ . \\ y_{t-p+1} \end{bmatrix}, A_t = \begin{bmatrix} A_{1,t} & \dots & A_{p,t} \\ I_{N(p-1)} & \dots & 0_{N(p-1)\times N} \end{bmatrix}, H = \begin{bmatrix} I_N \\ 0_{N(p-1)\times N} \end{bmatrix}.$$

The historical decomposition provides an interpretation of historical fluctuations in the time series in terms of the identified structural shocks, in this case the monetary policy shock.

[9]The monetary policy stance obtained through the historical decomposition is not altered by the variable that is chosen to perform the decomposition.

Figure 2. Cumulative Monetary Policy Shock Series



Note: The monetary policy shock is estimated using an identification-through-heteroskedasticity approach first proposed by Rigobon (2004). A sequence of positive monetary policy shocks indicates that monetary policy is becoming more expansionary; a drop in the series can reflect a tightening of monetary policy but also the lack of monetary actions.

or even that there were expansionary announcements that failed to live up to financial market expectations.

The figure shows that the shocks are able to capture important monetary policy announcements, as well as the anticipation of some measures. In October 2008, the financial crisis hit the economy and monetary policy was perceived to be insufficient given market conditions. Additional monetary policy actions were introduced in the course of 2009, which reverted the monetary policy stance to expansionary. The one-year LTRO/CBPP1 announcement in May 2009 and the SMP announcement in May 2010 are among the largest expansionary daily shocks and can therefore be considered surprises to financial markets. In the following years, the monetary policy stance is perceived by financial markets as somewhat volatile, with periods of restrictive monetary regimes followed by expansionary shocks in the monetary policy stance, caused by events such as

ECB President Mario Draghi's London speech in July 2012. The OMT announcement in September 2012 appears to have been largely anticipated following this speech in which he alluded to the implementation of additional unconventional monetary policy measures. The quantitative easing (QE) period which started in 2015 is sometimes perceived as a period of restrictive monetary policy, probably because of economic uncertainty stemming from the economic and political environment (e.g., Brexit). From 2017 onwards, the sustained monetary easing is considered by financial markets as effectively expansionary. An interesting example of the potential divide between policy intentions and market perception is described by Rostagno et al. (2019) in their account of the first 20 years of ECB monetary policy. In December 2015 the Governing Council decided to lower the deposit facility rate by 10 basis points. However, they conclude that the markets expected a larger reduction in the deposit facility rate, hence despite the intention of the ECB to be accommodating, the policy actions did not meet the expectations of financial markets (Rostagno et al. 2019). This resulted in a tightening of the monetary policy stance, as is also captured in our Figure 2, illustrating that our indicator of the monetary policy stance succeeds in identifying divergences between intended policy outcomes and actual market perceptions. This is an important value-added of the identification approach since stock market perceptions determine our market-based measures of bank systemic risk (MES) and long-term profit potential (market-to-book).

We estimate the following model combining the estimated monetary policy stance and the macroprudential index:[10]

$$\begin{aligned}\Delta y_{i,t+h} = {} & \alpha_i^h + \gamma_t^h + \beta^h D_{j,t} + \delta^h D_{j,t} \times Cum\ MP_t \\ & + \sum_{k=0}^{K} \Theta_k^h Bank_{i,t} + \sum_{l=0}^{L} \Gamma_l^h Macro_{j,t} + \varepsilon_{i,t+h}. \end{aligned} \tag{7}$$

[10]We acknowledge that macroprudential policy and monetary policy do not move independently of each other. In the local projections setup we use, we are not able to take potential regime changes into account. The impulse response functions (IRFs) show the cumulative evolution in the bank risk and return profile variables after a shock at time 0 conditional on the policy stance at time 0.

$\Delta y_{i,t+h}$ denotes the percentage change in the risk and return variables for bank i between time t and $t+h$. $D_{j,t}$ corresponds to the macroprudential tightening dummy in country j at time t. $CumMP_t$ is the cumulative monetary policy stance. α_i^h denote bank fixed effects. When we estimate the impact of macroprudential policy, we also include time fixed effects, γ_t^h, and estimate the model with weighted least squares with weights defined by the IPW model in the first step, i.e., $w_{j,t} = \frac{D_{j,t}}{\widehat{p_{j,t}}} + \frac{1-D_{j,t}}{1-\widehat{p_{j,t}}}$. The differential effect of a macroprudential shock across different monetary policy regimes is then calculated using the partial derivative of the bank risk and return variables with respect to the macroprudential index.[11]

4. Bank Risk and Return Profile

To conduct our analysis, we require accounting and market data for a sample of euro zone banks. We obtain quarterly balance sheet and income statement data from SNL Financial, which is available as of 2008:Q1. We exclude financial holding companies that are not engaged in banking activity (e.g., asset management companies, online brokers, or insurance companies). We exclude domestic subsidiary banks, but include foreign subsidiaries that satisfy the remaining criteria. Furthermore, we filter out banks that have a loans-to-assets ratio and a deposits-to-liabilities ratio lower than 20 percent. We use the accounting data to construct a set of bank business model variables to capture the asset, liability, and income structure of the banks as in Mergaerts and Vander Venner (2016). We measure a bank's asset structure by defining variables that capture the composition of earning assets (the loan ratio, LTA). We use the ratio of customer deposits to total liabilities (DEP) and an unweighted capital ratio, i.e., the ratio of total equity to total

[11] More specifically, the impulse responses are constructed as follows:

$$\frac{\partial \Delta y_{i,t+h}}{\partial D_{j,t}} = \hat{\beta}^h + \hat{\delta}^h CumMP_t.$$

We calculate both the average impulse response of a macroprudential policy shock on the bank risk and return profile variables when monetary policy is in an accommodating phase ($CumMP_t$ is larger than 0) and the average impulse response of a macroprudential policy shock on the bank risk and return profile variables when monetary policy is tight ($CumMP_t$ is lower than 0).

assets (CAP), to capture banks' funding and capital structure. As an indicator for the banks' income structure, we use the share of non-interest income in total income (DIV) as a proxy for revenue diversification. We also include bank size (SIZE), measured by total assets, as a control variable. Note that all variables have been winsorized at the 1 percent level. When quarterly data is lacking, we linearly interpolate data points that are reported at a half-yearly frequency to a quarterly frequency.[12] Income data reported at a quarterly frequency contains more variation than yearly data because of seasonality that is present in the data. To make sure the impulse responses are not influenced by this feature, we calculate the income variables (such as the net interest margin, or NIM, and the DIV) using a rolling window of the four previous quarters. Market data are obtained from Datastream.

To capture all dimensions of the bank's risk and return profile, we construct eight bank variables, of which six are based on accounting data and two on market data. First, we calculate the bank's loan growth since the most common intermediate objective of macroprudential policy is bank credit growth. Second, we use loan loss provisions as a forward-looking measure of loan quality which is a reflection of a bank's assessment of the quality of its loans. Third, we measure individual bank distress probability using the Z-score, or rather its natural logarithm as the variable itself is strongly positively skewed. This variable is defined in the following way:

$$Z\text{-}score_{i,t} = \frac{\frac{total\,equity_{i,t}}{total\,assets_{i,t}} + E_{i,t}(ROA)}{\sigma_{i,t}(ROA)} = \frac{CAP_{i,t} + E_{i,t}(ROA)}{\sigma_{i,t}(ROA)}. \tag{8}$$

We construct $E_{i,t}(ROA)$ and $\sigma_{i,t}(ROA)$ over a rolling window with three observations of ROA over the period t–2 to t. This procedure reduces the number of available observations slightly and removes banks with less than three consecutive observations. The Z-score should be interpreted as a distance-to-default measure, i.e., as the number of standard deviations ROA can diverge from its

[12]The general conclusions hold when we use the data that are not linearly interpolated.

mean before the bank defaults. A higher Z-score indicates a safer bank. Fourth, we calculate the change in the bank's leverage ratio measured by total assets divided by total equity. Fifth, we investigate the impact of policy on the change in the ratio of risk-weighted assets to total assets which provides an (rough) indication of the riskiness of the loan portfolio of the bank. Sixth, we include a measure for bank profitability in the analysis, measured by the NIM. In addition to bank balance sheet characteristics, we also investigate the impact of macroprudential policies on two measures constructed using market data. First, we include a measure for bank systemic risk. A commonly used approach is to model systemic risk as the contribution of a bank to systemwide stress. One of the most frequently used measures for systemic risk is the marginal expected shortfall (MES) by Acharya et al. (2017) calculated as the expected loss of a bank's stock price conditional on a large shock to the financial system.[13] Second, to capture the stock market's assessment concerning the franchise value of the bank, we include the market-to-book ratio. Figure 3 displays the evolution of the bank risk and return profile variables over time. The graphs demonstrate the positive evolution of euro zone bank risk during the sample period (lower loan loss provisions, lower leverage (i.e., higher capital ratios), and a higher Z-score). Most variables show the distress of the banks during the banking crisis and again during the sovereign crisis in the euro area.

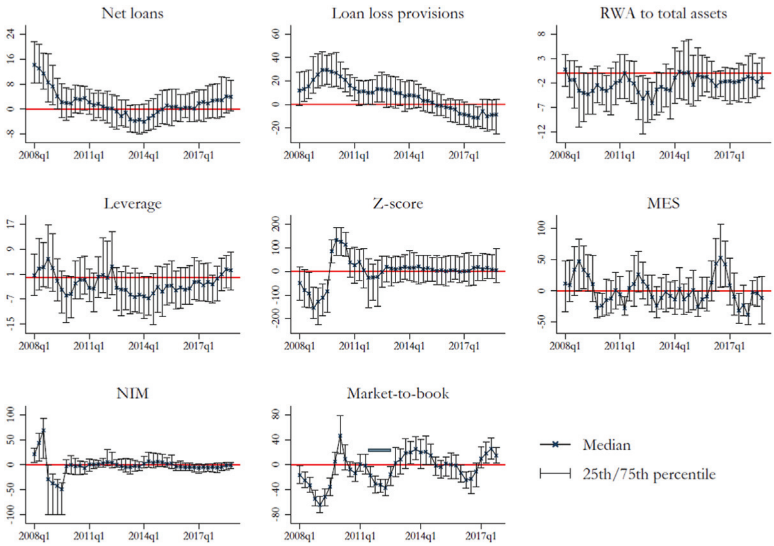
The macro control variables described in Section 3 (Methodology) that are used both in the first-stage logit regressions and in the local projections are retrieved from the ECB Statistical Data Warehouse (SDW). We include the changes in bank credit to non-financial corporations to capture domestic credit growth in each country. Second, to control for developments in the real estate market, we

[13] The MES measures a bank's expected equity loss when the market falls below a certain threshold over a given horizon and can be written as

$$MES_{i,t} = E_{t-1}(r_{i,t} | r_{m,t} < C).$$

In line with Acharya, Engle, and Richardson (2012), the threshold C that defines a crisis is set at a –2 percent loss in the relevant market index over a one-day period. As the market index we use the MSCI Europe. To estimate the different components of the MES we follow the procedure as described in Idier, Lamé, and Mésonnier (2014) and Brownlees and Engle (2017).

Figure 3. Evolution of the Growth of the Bank Risk and Return Profile Indicators for European Banks



Note: The black dashed line indicates the median. The dark area represents the 25th and 75th percentiles.

include the year-on-year change in the country-level residential property price index. Third, we include country-level GDP growth to account for economic activity. Fourth, we include the ratio of household debt to GDP in the model since policymakers use this measure as an indicator to initiate borrower-related macroprudential tools, such as loan-to-value ratios. To control for the level of volatility on the stock markets we include the VSTOXX, which is retrieved from Datastream. Last, we control for monetary policy and include the ECB MRO rate, which is also retrieved from Datastream.

The application of the sample selection criteria results in a data set of accounting measures, depending on the risk or return profile variable that is used, for around 140 banks for a total of around 3,400 bank-quarter observations at time t=0. The data set using the market-based measures results in a sample of 63 and 64 euro zone banks and around 2,200 bank-quarter observations at time t=0 when using the MES or the market-to-book variable, respectively.

The descriptive statistics are given in Table 2.

Table 2. Descriptive Statistics of the Dependent and Independent Variables

	Obs.	No. Banks	Mean	St. Dev.	Min.	Max.
Bank Risk Variables						
Net Loans Change	3,426	142	0.29%	3.99%	–10.68%	21.72%
Leverage Change	3,425	142	–0.06%	8.82%	–27.18%	28.83%
Z-score Change	3,377	140	–1.86%	55.06%	–276.06%	153.36%
Loan Loss Reserves Change	2,173	102	2.06%	10.54%	–29.51%	58.20%
RWA-to-Assets Change	3,307	136	–0.67%	4.23%	–15.32%	13.92%
NIM Change	3,426	140	2.73%	19.38%	–80.24%	95.22%
MES Change	2,208	63	0.26%	39.82%	–180.10%	156.70%
Market-to-Book Change	2,257	64	–4.44%	21.52%	–84.16%	51.95%
Bank Controls						
SIZE	3,424		17.48	1.68	13.11	21.35
DEP	3,424		0.55	0.16	0.18	0.89
LTA	3,424		0.62	0.14	0.18	0.91
CAP	3,424		0.07	0.04	0.02	0.27
DIV	3,424		0.38	0.19	0.00	1.00
Macro Controls						
Loan Growth	3,424		–1.04%	5.42%	–18.24%	31.49%
GDP Growth	3,424		1.84%	3.34%	–10.03%	19.99%
House Price Growth	3,424		0.60%	5.51%	–16.10%	33.33%
Debt Growth	3,424		–0.79%	4.15%	–15.09%	20.97%
VSTOXX	3,424		23.55	7.72	12.17	48.65
Policy Rate (MRO)	3,424		0.77%	1.06%	0.00%	4.25%

5. Empirical Results

In this section we assess the impact of macroprudential and monetary policy on euro zone banks' risk and return profile using accounting- and market-based measures of the banks' risk and return profile. Subsection 5.1 reports and discusses the results of the inverse propensity score regressions on bank risk and return profile measures in a local projections framework. We also check whether certain bank business models react more strongly to changes in the macroprudential policy stance. Subsection 5.2 investigates the impact of macroprudential policy across different monetary policy regimes.

5.1 The Impact of Macroprudential Policy on the Bank's Risk and Return Profile

In order to investigate the impact of macroprudential policy on the bank risk and return profile variables, we apply the inverse propensity score procedure. We start by performing the first-stage logit regression shown in Equation (1). We run logit classification models for the tightening dummy $D_{i,t}$ since we want to account for macroeconomic variables that are supposed to be associated with the initiation of macroprudential policy actions. Hence, we include in this regression the year-on-year percentage change in GDP growth, the country-specific housing price index, the yearly growth rate of bank credit, and household debt as well as country fixed effects and year dummies. Table 3 presents the results of the first stage.

Table 3 indicates that macroprudential tools are especially initiated after an increase in loan growth during the previous year. Also the VSTOXX appears to be a significant predictor for a tightening in the overall macroprudential policy stance. We report the AUC statistic, which indicates the area under the receiver operating curve. The statistic measures the predictive ability of a model to correctly sort observations into "tightening" and "no tightening." The AUC takes on the value of 1 for perfect classification ability and 0.5 for an uninformed random classification. The AUC of the full model is 0.758, which indicates that the first stage

Table 3. First-Stage Logit Regression to Predict a Tightening in Macroprudential Policy

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Annual Loan Growth, t–1	4.835***						3.499*
	(1.535)						(2.120)
Annual GDP Growth, t–1		–0.010					–5.792
		(1.625)					(3.832)
Annual House Price Growth, t–1			0.953				2.283
			(1.450)				(2.416)
Annual Household Debt Growth, t–1				2.420***			0.084
				(0.814)			(3.053)
VSTOXX, t–1					0.022		0.045**
					(0.016)		(0.020)
Policy Rate, t–1						0.002	0.013
						(0.218)	(0.455)
N	1,205	1,583	903	1,423	1,482	1,596	789
R^2	0.153	0.125	0.133	0.116	0.116	0.124	0.141
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
AUC	0.777	0.754	0.753	0.741	0.744	0.754	0.758

Note: The model is estimated over the sample period 2000:Q1–2018:Q4 covering 19 euro zone countries.

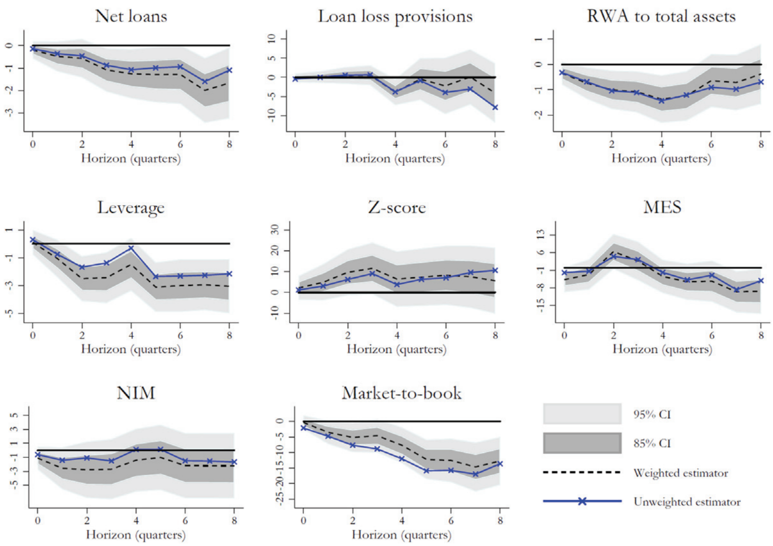
is informative in predicting a tightening in the macroprudential stance.[14]

Having estimated the first-stage logit model, we use the fitted probabilities from this model in the local projections setting as in Jordà (2005) estimated using weighted least squares as in Richter, Schularick, and Shim (2018) and Alam et al. (2019). As discussed in Section 3.1, the weights are defined by $w_{i,j,t} = \frac{D_{j,t}}{\widehat{p_{j,t}}} + \frac{1-D_{j,t}}{1-\widehat{p_{j,t}}}$, where we truncate $w_{i,j,t}$ at 10. Figure 4 shows the impulse response functions of a tightening shock in macroprudential policy on the bank risk and return profile variables. As explained in Section 4, we construct six bank risk and return profile variables based on accounting data and two variables based on market data.

Figure 4 presents the impact on net loans (top left panel), four accounting-based risk and return profile variables (LLP ratio, RWA density, leverage, and Z-score), one market-based systemic risk indicator (MES), and—in the bottom row—the two profitability variables, the banks' net interest margin (NIM) and the market-to-book ratio (for the subsample of listed banks). The first important result is that bank loan growth decreases following a macroprudential policy tightening, as intended by policymakers. Credit growth decreases by around 2 percentage points after four to six quarters following a tightening in macroprudential policy. This downward effect on bank lending is consistent with several papers focusing on aggregated indicators of bank lending, such as Lim et al. (2011), Kuttner and Shim

[14] We acknowledge that the first model specification, only including the lagged loan growth, has an even higher AUC statistic, indicating that this model specification is even better in predicting a macroprudential tightening in a certain country. In addition, specification 7 has fewer observations which could have an impact on the second-stage results. We performed several robustness checks on the first-stage regressions (different variables, different lag lengths, and different time periods), and we find that the results are not sensitive to the specification that is chosen in the first stage. As a second robustness check we investigate whether the first-stage results are different when we also include the lagged cumulative macroprudential index as a covariate in the first-stage regression. We find that the higher the macroprudential policy stance, the lower the probability of a new macroprudential initation (mainly for liquidity and capital regulation). The results in the second stage remain unaltered, except for the market-to-book ratio, which decreases less following a macroprudential shock. This effect is mainly driven by capital and liquidity regulation for which the initiation depends on the stance of macroprudential policy in that specific country and which now receives a lower weight.

Figure 4. Baseline Results of a Tightening Shock in Macroprudential Policy on a Set of Bank Risk and Return Profile Variables



Note: The response is expressed in percentage-point changes. To estimate the responses we estimate a local projections model with weighted least squares, where the weights are defined by an inverse propensity score model, as described in Section 3.1. The black dashed line represents the coefficients of this weighted estimator. The blue solid (crossed) line indicates the coefficients of the unweighted estimator. The index that is used is the overall macroprudential index covering all policy tools. The dark area represents the 95 percent confidence intervals. The lighter area indicates the 85 percent confidence intervals. The horizon is measured in quarters.

(2016), Cerutti, Claessens, and Laeven (2017), Akinci and Olmstead-Rumsey (2018), and Poghosyan (2019), among others. The estimated impact of the results is in line with the existing literature, where the impact varies between 0.3 percentage point the following quarter (Akinci and Olmstead-Rumsey 2018) to 2.2 percentage points after four quarters (Cerutti, Claessens, and Laeven 2017) for the overall macroprudential index. In terms of bank risk and return profile, the evidence in Figure 4 points to decreasing bank risk. We observe no significant change in the loan loss provision ratio, indicating that banks do not increase the riskiness of their loan portfolio. In addition, the contraction of lending is accompanied by a similar decline of the RWA density ratio, suggesting the absence of risk-shifting behavior: banks do not compensate the decline in the loan type(s) targeted by the macroprudential actions by investing in other riskier types of loans or by shifting exposures to riskier securities. It has to be noted that the post-2008 period is also characterized by the gradual implementation of Basel III and the compliance with, e.g., the LCR and the NSFR may induce banks to decrease their portfolio of risky long-term assets and shift to safer asset classes such as sovereign bonds, which would imply a decrease of the RWA density. This effect is documented by Banerjee and Mio (2018), who show that banks increase the share of high-quality liquid assets while they reduce intrafinancial loans as a response to liquidity regulation. Next to lower loan growth, we observe a decline in the leverage ratio, indicating that banks opt for deleveraging and holding more capital, which improves their risk profile. The leverage ratio decreases by around 3 percentage points after two years. Again, it has to be noted that this behavior may be driven by adherence to strengthened capital regulation in the Basel framework, which was implemented during the sample period. Combined, the improved bank risk profile metrics translate into a higher Z-score, indicating that the distance to default increases and hence bank resilience improves. Finally, we consider the MES as the market-based indicator of how stock market investors perceive the evolution of bank risk. Since the MES captures the probability of systemic stress for listed banks, the results indicate that market participants acknowledge the improved risk profile since the MES decreases significantly after four quarters. This finding corroborates the evidence in Meuleman and Vander Vennet (2020), who report that announcements of various macroprudential policy tools exert a downward effect on the MES of European banks. The conclusion from Figure 4 is that macroprudential policy in general is able to improve the risk profile of euro-area banks, and hence that it is effective in supporting financial stability.

However, the positive effect of macroprudential policy on the bank's risk profile comes with a downside: current and longer-term bank profitability experience stress. We observe a negative effect on the NIM following a macroprudential shock. This result is significant in the short term and fades to marginally significant over the projection horizon, but it is apparent that the majority of the banks experience downward pressure on their margins. This is not unexpected, since restrictions on lending or tightened liquidity rules typically result in lower interest income. Moreover, King (2013) shows that the introduction of liquidity rules such as the NSFR reduces bank net interest margins by requiring banks to use stable funding sources, which have a higher funding cost. Additional insight comes from the way stock market investors assess the impact of macroprudential actions on the long-term profitability of the banks concerned. The market-to-book ratio exhibits a significant decline over the entire impact horizon, indicating that stock markets view macroprudential regulation as negative for bank market valuations. This result is in line with Richter, Schularick, and Shim (2018), who find that stock market prices are negatively affected by the introduction of loan-to-value ratios in 56 economies.

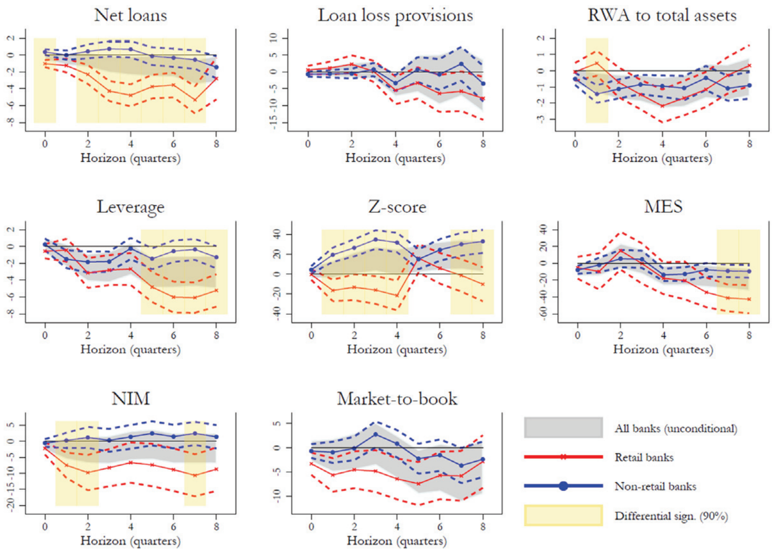
The overall conclusion from Figure 4 is that while macroprudential regulation improves the risk profile of euro-area banks, as intended, the constraints imposed by the new rules affect bank profits negatively, which may ultimately have an impact on the stability of the banking sector. We acknowledge that the results may potentially be influenced by cross-border banking flows that could lead to leakage effects and regulatory arbitrage (as found in Aiyar, Calomiris, and Wieladek 2014 and Reinhardt and Sowerbutts 2015). However, we have several reasons to believe that this bias will be rather small. In particular, we investigate the impact of a domestic macroprudential shock on a sample of domestic groups and foreign subsidiaries. First, foreign subsidiaries need to comply with regulation, which means that the impact of a macroprudential shock will be visible at the foreign subsidiary level, regardless of regulatory arbitrage or leakage effects. If there are indeed leakage effects, this can undermine the effectiveness of the macroprudential measure to curb credit growth at the country level. The incentives for regulatory arbitrage are stronger for institution-based measures, as they target the bank rather than the borrowers. This calls for an automatic and compulsory reciprocity agreement for institution-based measures. There is less incentive for regulatory arbitrage with respect to borrower-based regulation, as the regulation is linked to the borrower. Second, for domestic groups at the consolidated level, the impact of macroprudential measures may be less visible, as there can be a shift of activities to foreign subsidiaries. We perform a robustness check whereby we include the domestic subsidiary rather than the domestic group because the impact will be directly measurable at the domestic subsidiary level. We find, however, that the difference in the results is negligible. In addition, we acknowledge that the sample of macroprudential tools also contains bank-specific capital tools such as G-SII, O-SII, and systemic risk buffers which only affect large banks in the sample, while there is no impact for smaller banks. The impact on the response variables may thus be affected by these selective macroprudential policy tools. We therefore perform a robustness check which excludes the bank-specific capital buffers from the full sample of tools. The main results, in the first stage and in the second stage, remain unaltered by the exclusion of the tools. If any difference, the impact on the market-to-book value is even somewhat larger when not taking into account the buffers. This might imply that the impact of these capital buffers on the profitability of banks is less severe compared with other macroprudential policy tools.[15]

In a next stage, we analyze how macroprudential policy is transmitted across different bank business models. We hypothesize that different types of macroprudential measures will affect different types of banks in a heterogeneous way. When, e.g., the macroprudential authority undertakes actions to limit certain exposures, only banks with such exposures will need to take remedial action. We examine this hypothesis by interacting the macroprudential index with the *RETAIL* factor we obtained after running a factor analysis on a set of bank business model variables (see Section 3.1). This *RETAIL* factor captures the *retailness* of the banks since it positively loads on the loan, deposit, and capital ratios, but is negatively related to size and income diversification. Figure 5 shows the results of the local projections setting where we interact the macroprudential shock with the *RETAIL* factor from the factor analysis.

The impulse responses show that the impact of macroprudential tightenings is more pronounced for retail-oriented banks than for their non-retail counterparts. For example, credit growth decreases

[15]The results of these tests are available upon request.

Figure 5. Results of a Tightening Shock in Macroprudential Policy on a Set of Bank Risk and Return Variables Conditioning on the Bank Business Model



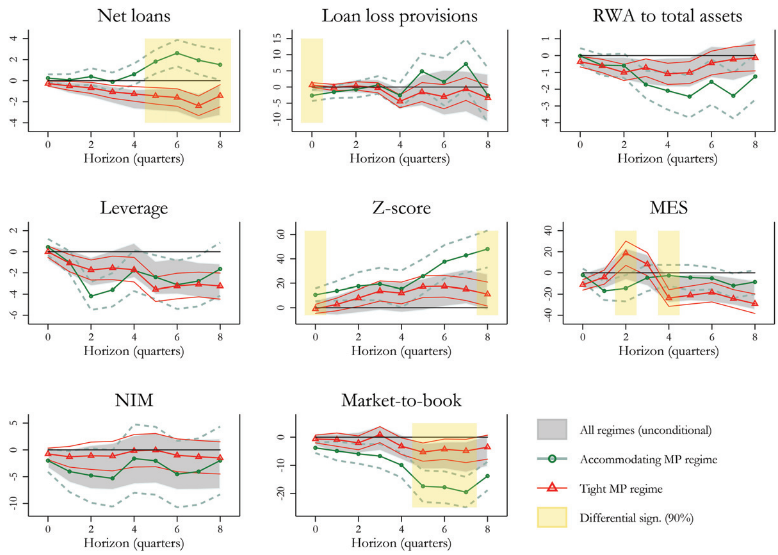
Note: The response is expressed in percentage-point changes. We obtain an indicator of the bank business model by performing a factor analysis on a set of bank characteristics. The first factor, which explains 63 percent of the variation, is related to the retail orientation of the bank, so we interact this factor with the macroprudential policy shock. The blue (circled) impulse responses indicate the response of banks that are classified by the factor analysis as being non-retail banks, i.e., the factor score is smaller than −0.5 (lowest 25 percent factor scores). The red (crossed) line impulse responses indicate the response of retail-oriented banks to a macroprudential policy shock, i.e., a factor score larger than 0.5 (highest 25 percent factor scores). The yellow bars indicate the differential significance level between retail and non-retail impulse responses at the 90 percent significance level. The index that is used is the overall macroprudential index covering all policy tools. For the unconditional impulse responses, we show the 95 percent confidence intervals. For the impulse responses of retail and non-retail banks, we use 68 percent confidence intervals. The horizon is measured in quarters.

by around 3 percentage points after four quarters for a retail-oriented bank, while for a non-retail bank the impact is limited to less than 1 percentage point. The decrease in profitability, measured by both the NIM and the market-to-book that was found in Figure 4, is mainly attributable to retail banks which seemingly suffer more from macroprudential policy actions. The negative impact on the profitability in turn negatively influences the Z-score of retail banks. In contrast, the Z-score of non-retail banks slightly increases following a macroprudential policy shock. In summary, retail-oriented banks are more sensitive to macroprudential policy shocks than other banks, probably because they have more difficulties absorbing the shock since they are more dependent on mortgage loans, which are frequently targeted by macroprudential regulators. In contrast, non-retail banks have a more diversified asset and revenue portfolio, which makes them less sensitive to changes in prudential regulations. The results are in line with the findings of Altunbas, Binici, and Gambacorta (2018), who find that smaller banks react more strongly to macroprudential changes. Meuleman and Vander Vennet (2020) find that the individual risk component, which is a subcomponent of the MES capturing idiosyncratic bank risk, decreases more strongly for retail-oriented banks. An implication of these results is that retail banks should diversify their asset portfolio in order to make them less sensitive to changes in macroprudential policy.

5.2 The Interactions between Macroprudential Policy and Monetary Policy

The crucial research question for policymakers is whether or not the transmission of macroprudential policy varies across different states of monetary policy. To check whether or not this is the case, we first interact the macroprudential shock with the stance of monetary policy, constructed as described in Section 3.2. Figure 6 shows the impulse responses of the local projections of a tightening in macroprudential policy and its effectiveness across monetary policy regimes. The red lines correspond to the response of a tightening in macroprudential policy when monetary policy is restrictive, i.e., when the monetary policy stance is below 0. The green lines indicate the impulse responses of a macroprudential tightening on the bank risk and return profile variables when monetary policy is considered to be loose by market participants, i.e., when the monetary policy stance is positive. (For figures in color, see the online version of the paper at http://www.ijcb.org.)

Figure 6. Impact of a Tightening in the Macroprudential Index across Different Monetary Policy Regimes for a Sample of Euro Zone Banks between 2008:Q4 and 2018:Q4 on a Set of Bank Risk and Return Profile Variables



Note: The response is expressed in percentage-point changes. The monetary policy shock is constructed based on the identification-through-heteroskedasticity methodology in line with Rigobon (2004) and as described in Section 3.2. The monetary policy stance is calculated as the historical contribution of the monetary policy shock to changes in the five-year Spanish CDS spread. The red (triangle) solid line indicates the response of the bank risk profile variables to a tightening in macroprudential policy conditional on the monetary policy stance being tight (i.e., the stance is below 0). The green (circle) solid line indicates the response of the bank risk profile variables to a tightening in macroprudential policy conditional on the monetary policy stance being loose (i.e., the stance is above 0). The yellow bars indicate the differential significance level between accommodating and tight monetary policy regime impulse responses at the 90 percent significance level. The gray area denotes the unconditional response to a macroprudential tightening. The confidence bounds for the unconditional impulse responses represent the 95 percent confidence intervals. For the conditional impulse responses, we use 68 percent confidence intervals. The horizon is measured in quarters.

Figure 6 shows the estimation results over a projection horizon of eight quarters. We first interpret the impulse responses for the situation in which tightening macroprudential measures are announced in a period characterized by a restrictive monetary policy stance (depicted by the solid red line). In this case, the two policies reinforce each other in lowering credit growth, as intended by both the monetary and the macroprudential authorities. In terms of bank risk, the behavior of the bank risk profile variables is consistent with improved bank stability. We not only find that net lending decreases but there is also no evidence of risk-shifting behavior, since the RWA density decreases and the LLP ratio remains constant or even decreases slightly. Simultaneously, the leverage ratio decreases significantly (banks become better capitalized) and the increasing Z-score, as an overall measure of bank health, signals improving bank resilience. For policymakers, this is the desired outcome of their actions since macroprudential policy and restrictive monetary policy work in the same direction. These findings are in line with David et al. (2019) and Gambacorta and Murcia (2019), who also document that macroprudential and monetary policy push in the same direction, i.e., they restrain credit growth. This result is also confirmed by Popoyan, Napoletano, and Roventini (2017), who find that monetary policy and macroprudential regulation are complementary in increasing the resilience of the banking sector. However, improved stability comes at a price, since we find evidence of pressure on current and future profitability. On average, the impulse response for the NIM is not significant, but the market-to-book ratio declines significantly over the projection horizon, indicating that macroprudential measures combined with restrictive monetary policy effectively impose constraints on banks. These negative consequences on the banks' franchise value are (almost) identical to the effects exhibited in Figure 4 (the standalone effect of tightening macroprudential actions), and from Figure 5 we know that these negative effects on bank profitability are particularly pronounced for retail banks. Hence, when macroprudential policy and monetary policy operate jointly in a restrictive regime, the risk profile of euro zone banks improves, but at a cost of lower anticipated profitability.

An interesting case is when there is a potential trade-off between monetary and macroprudential policy. This is the prevailing environment in the post-2008 era, since it is characterized by the simultaneous introduction of restrictive macroprudential measures following the financial and sovereign crises in Europe as well as unprecedented conventional and unconventional monetary policy by the central bank. However, as our SVAR in Figure 2 demonstrates, monetary policy actions intended by the ECB as stimulating were not always perceived as such by financial markets. Hence, our impulse responses to bank risk and return variables should capture those cases in which macroprudential measures were introduced in periods in which the monetary actions of the ECB are considered by markets as unambiguously accommodating. In Figure 6, the solid green line captures the impact on the banks' risk and return profile of macroprudential measures in periods of perceived monetary stimulus. The top left panel shows that bank loan growth does not slow down initially and even increases significantly after four quarters, suggesting that the transmission of macroprudential policy is affected by the presence of loose monetary policy. For the central bank, this is the most desired outcome since its actions are geared towards stimulating lending to the real economy. This result confirms the common finding that ECB monetary policy succeeded in decreasing loan rates and increasing bank lending (see Rostagno et al. 2019). The main concern of policymakers is that more lending may be accompanied by increased risk-taking by banks, by engaging in lending to riskier borrowers or shifting towards riskier securities (Heider, Saidi, and Schepens 2019). Our results are not compatible with this risk-taking channel. Our impulse responses show that loan loss provisions do not increase and the RWA density even decreases significantly, suggesting the absence of risk-shifting behavior. At the same time, the capital adequacy of the banks increases significantly (lower leverage) and the same observation holds for the Z-score. Our market-based risk indicator (MES) never increases over the projection horizon. The conclusion is that accommodating monetary policy may entail incentives for banks to take more risk, but in the period under investigation, our results indicate that macroprudential measures were sufficiently strong to deter banks from excessive risk-taking. This conclusion is consistent with the findings in Albertazzi et al. (2020), who examine the pricing behavior of euro-area banks and conclude that any additional risk taken in the post-2014 period was not inadequately priced. Similar evidence is reported for the rebalancing of bank securities portfolios. Albertazzi et al. (2020) report that, since the start of the APP, banks' bond portfolios have shifted through an active rebalancing out of the safest categories of securities into other investment-grade bonds. However, they argue that over the same period, this effect was more than offset by positive rating migration caused by improved macroeconomic conditions. Moreover, they show that banks' portfolio rebalancing has not translated into a loading up of domestic sovereign debt securities, not even in those economies where such securities offer higher yields. Our findings are also corroborated by Soenen and Vander Vennet (2022), who investigate the impact of ECB monetary policy on bank CDS spreads and conclude that over the post-2008 period, accommodating monetary policy by the ECB is associated with a beneficial impact on the market-perceived default risk of European banks.

Macroprudential measures announced in an environment of accommodating monetary conditions are associated with higher loan growth, but do not induce excessive risk-taking by banks. The consequences of this policy mix on bank profitability are, however, less benign. We observe in Figure 6 that the impact on the NIM is negative, although statistically not significant. According to ECB (2020), banks have increased their loan volumes in an effort to protect their interest margin, but such compensation is finite. More importantly, we observe a marked deterioration of the banks' market-to-book value as a reflection of the investors' conviction that low-for-long interest rates ultimately compress bank interest margins and put their profitability and franchise value under stress. Altavilla, Boucinha, and Peydró (2018) argue that the ECB's APP and negative deposit facility rates have a close to zero net effect on banks' ROA since positive effects (capital gains on securities and better credit quality) compensate any decline in the banks' net interest margins. However, while capital gains and lower loan losses are temporary, the gradual decline of net interest margins is a persistent phenomenon as long as monetary policy remains ultraloose. The decline in market-to-book ratios in the regime of accommodating monetary policy is significantly more pronounced than under a restrictive stance. Our results are consistent with Borio, Gambacorta, and Hofmann (2017) and Claessens, Coleman, and Donnelly (2018), who examine the impact of low policy rates on bank interest margins and conclude that long periods of low rates indeed compress bank margins. Hence, from the bottom panel of Table 6, we conclude that the combination of restrictive macroprudential policies and prolonged monetary accommodation may turn out to be detrimental for bank health and, ultimately, financial stability.

6. Extensions and Robustness Checks

In this section we perform several extensions and robustness checks to validate the results on the interaction between monetary policy and macroprudential policy. More specifically, we construct a more granular macroprudential policy index where we subdivide the index into different subindices based on their macroprudential objective in Subsection 6.1, an alternative (conventional) monetary policy stance using country-specific Taylor rules in Subsection 6.2, and an alternative unconventional monetary policy stance using the identification through external instruments approach in Subsection 6.3.

6.1 A More Granular Macroprudential Index Based on the Macroprudential Objective

In the baseline regression results we use the aggregate macroprudential policy index, including all policy actions. As different macroprudential policy tools have different objectives, it may be insightful to investigate whether different kinds of macroprudential policy tools have different effects on the bank risk and return variables and whether the effectiveness of the different tools varies over monetary policy regimes. For this setup, we regroup the tools in three types of categories according to their objective and we distinguish (i) credit growth restrictions which incorporate loan-to-value ratios, loan-to-income ratios, debt-serve-to-income ratios, maturity and amortization restrictions, and risk weights on mortgage loans and commercial loans, (ii) liquidity regulations covering liquidity coverage ratios, net stable funding ratios, loan-to-deposit ratios, and other liquidity requirements, and (iii) measures that affect the resilience of the banking sector such as minimum capital requirements (mainly the regulations under the CRR/CRD framework), capital buffers (systemic risk buffers, countercyclical buffers, capital conservation buffers), taxes on financial institutions, and loan loss provisioning rules.

We first apply the inverse propensity score procedure on the three disaggregated macroprudential indices to estimate the response of bank risk variables to a macroprudential shock. We start by performing the first-stage logit regression described in Equation (1). Table 4 presents the results of the first stage.

Table 4. First-Stage Logit Regression to Predict a Tightening in Macroprudential Policy

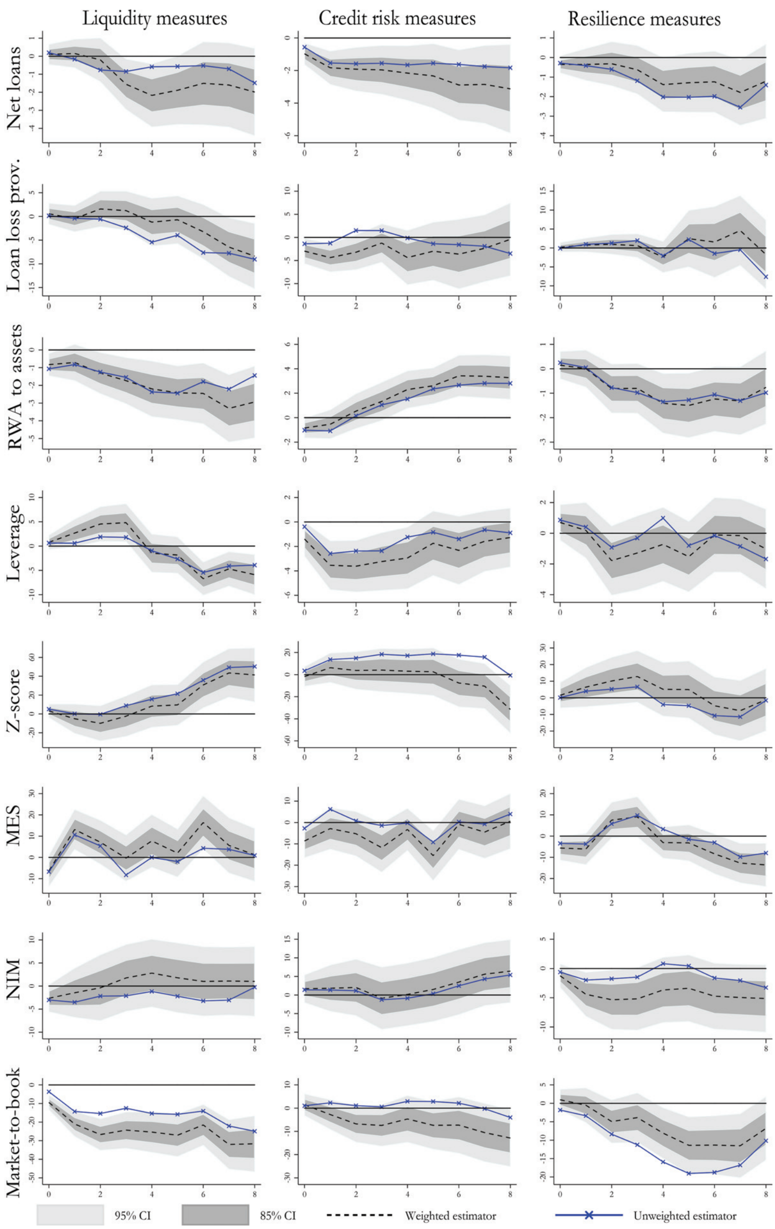
	All Policy Tools (1)	Credit Growth Tools (2)	Liquidity Tools (3)	Resilience Tools (4)
Annual Loan Growth, t–1	3.499*	8.526*	2.406	3.180
	(2.120)	(4.824)	(4.192)	(3.007)
Annual GDP Growth, t–1	–5.792	–7.882	0.618	–5.310
	(3.832)	(8.449)	(6.552)	(4.850)
Annual House Price Growth, t–1	2.283	8.475	1.003	0.883
	(2.416)	(5.467)	(4.846)	(3.047)
Annual Household Debt Growth, t–1	0.084	–4.966	4.918	–1.689
	(3.053)	(6.672)	(5.717)	(4.048)
VSTOXX, t–1	0.045**	–0.011	0.110***	0.075***
	(0.020)	(0.045)	(0.039)	(0.026)
Policy Rate, t–1	0.013	–0.668	–1.582	1.243*
	(0.455)	(1.364)	(1.376)	(0.642)
N	789	360	560	671
R^2	0.141	0.131	0.163	0.172
Country Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
AUC	0.758	0.763	0.794	0.792

Note: We perform the first-stage regressions on subindices of the macroprudential index which are constructed based on their objective. The model is estimated over the sample period 2000:Q1–2018:Q4 covering 19 euro zone countries.

From Table 4 we can see that macroprudential tools are initiated after an increase in the loan growth during the previous year. The effect is most pronounced for the credit growth tools, as expected. The $VSTOXX$ appears to be a predictor for both the liquidity tools and the resilience measures.

Figure 7 shows the impulse responses for a shock in the different macroprudential subindices. When we first focus on loan growth, we see that all three macroprudential policies decrease credit growth. While the effects of liquidity measures and resilience measures only become visible after several quarters, credit risk measures impact the bank risk variables immediately. From the MaPPED database we know that the time period between the announcement of credit growth tools and the actual enforcement is on average 2.5 months while the phase-in period is on average 3.3 months, 7.6 months, and 4.8 months for liquidity regulation, minimum capital requirements, and capital buffers, respectively. Banks thus need to adjust faster to credit risk measures than to other measures. The impact of credit risk measures is also somewhat higher than that of resilience and liquidity measures: the initiation of credit risk measures, such as loan-to-value ratios, decreases bank loan growth by 3.12 percentage points after two years, while liquidity and resilience measures decrease lending of banks by, respectively, 1.99 and 1.21 percentage points, on average. This result is in line with Kuttner and Shim (2016), who find that targeted credit policies such as debt-service-to-income requirements and housing-related taxes can be used as tools to restrain housing credit growth. In contrast, supply-side credit policies such as risk weights and provisioning requirements had no significant impact on housing credit. The estimated impact of credit growth measures on credit growth is in line with the existing literature where the estimates range between 2 to 6 percentage points per year (Zhang and Zoli 2016) and 4 to 7 percentage points per year (Kuttner and Shim 2016). With respect to loan loss provisions, we find that the implementation of liquidity regulation significantly decreases the loan loss provisions in the longer run. Turning to the risk profile of the loan portfolio, as measured by the risk-weighted assets to total assets, we observe that both liquidity measures and resilience measures induce banks to hold a safer asset portfolio. However, for the credit risk measures we see the opposite as banks increase the amount of risky assets in the total asset portfolio.

Figure 7. Robustness Check: Impact of a Tightening Shock in Different Microprudential Policy Tools on a Set of Bank Risk and Return Profile Variables



(continued)

Figure 7. (Continued)

Note: The response is expressed in percentage-point changes. To estimate the responses we estimate a local projections model with weighted least squares, where the weights are defined by an inverse propensity score model, as described in Section 3.1. The black dashed line represents the coefficients of this weighted estimator. The blue solid (crossed) line indicates the coefficients of the unweighted estimator. In this robustness check we subdivide the index into different subindices based on their macroprudential objective. We distinguish between liquidity measures (liquidity coverage ratios, net stable funding ratios, loan-to-deposit ratios, and other liquidity requirements), credit growth measures (loan-to-value ratios, loan-to-income ratios, debt-service-to-income ratios, maturity and amortization restrictions, and risk weights on mortgage loans and commercial loans), and resilience measures (regulations under the CRR/CRD framework, capital buffers, taxes on financial institutions, and loan loss provisioning rules). The dark area represents the 85 percent confidence intervals. The lighter area indicates the 90 percent confidence intervals. The horizon is measured in quarters.

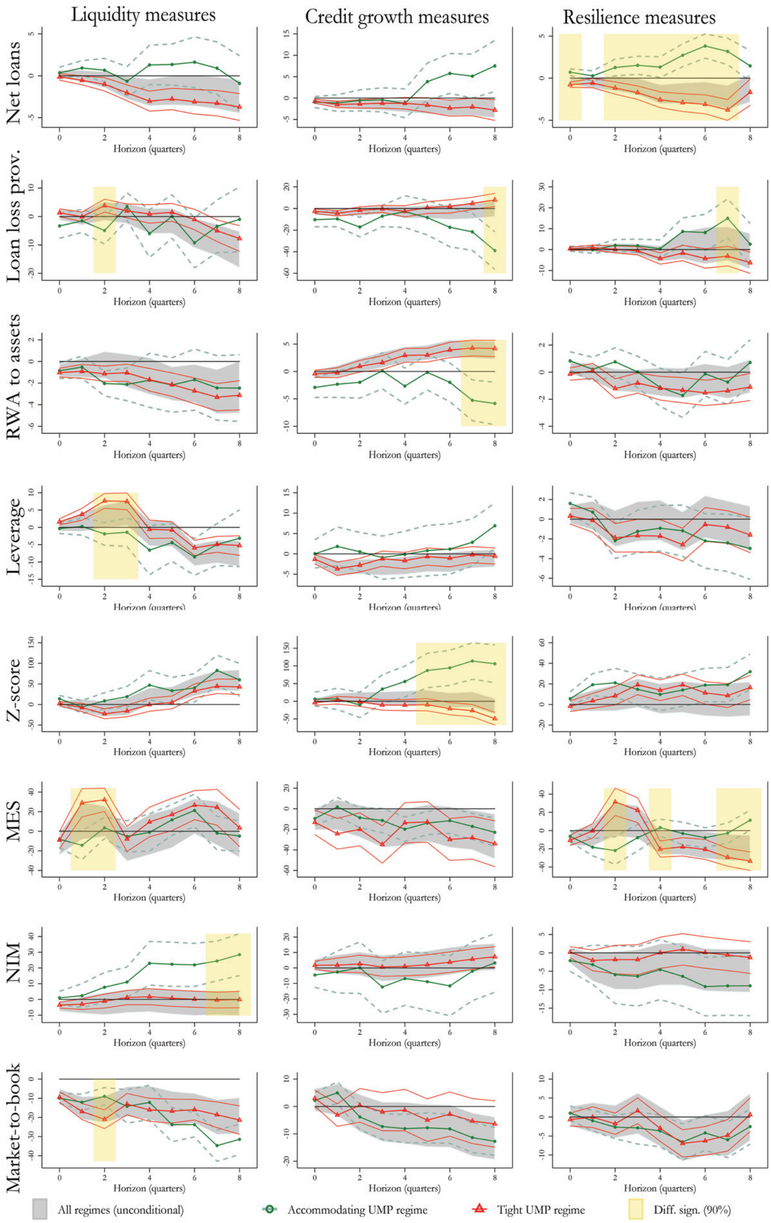
These results are compatible with a risk-shifting explanation. Since lending-oriented tools force banks to lower their exposures to certain types of counterparties or to disinvest certain types of loans or securities, the banks may shift the asset composition towards exposures that make them more interconnected to the financial system. As a typical example, restrictions on mortgage lending, e.g., in the form of loan-to-value caps or higher capital weights, may induce a shift to corporate lending or securities, which exposes these banks to business cycle shocks. This finding is in line with Acharya et al. (2018), who find that banks increase their holdings of risky securities and corporate credit in response to the introduction of loan-to-value or loan-to-income limits in Ireland. Cizel et al. (2016) also show that mainly quantity restrictions, such as exposure limits, are more prone to cause strong substitution effects. In terms of policy this calls for a careful calibration of macroprudential measures in order to avoid the unintended consequences of risk-shifting behavior by the affected banks. Auer and Ongena (2019) also find evidence of a risk-shifting channel following macroprudential tightenings as banks shift their lending to more commercial lending and to smaller and riskier firms using a loan-level data set of credit granting in Switzerland.

When we focus on the leverage ratio, we find that liquidity regulation decreases the leverage ratio after one year, which is a result of the decrease in loan growth. The leverage ratio for credit risk measures decreases more rapidly, since credit growth reacts immediately after the announcement of these tools. Surprisingly, for the resilience measures, which consist mainly of minimum capital requirements and capital buffers, the effect on the leverage ratio is limited. This may be due to the fact that the announcement of capital buffers comes on top of already enforced capital regulation (Basel III). Since most banks hold capital buffers in excess of the regulatory minimum, the announcement of additional capital buffers may not impose additional constraints. Another explanation is that capital regulation mainly targets the weighted capital ratio, while the leverage ratio is an unweighted measure for bank capitalization. In response to capital-related measures, banks react with a decrease in the risk weights of the assets, rather than with a deleveraging. This is also found by Cappelletti et al. (2019), who find that banks react to O-SII capital buffers by adjusting the risk-weighted assets rather than by reducing credit supply. With respect to the Z-score, we only find that liquidity regulation widens the distance to default after one year. For the MES, the results are less clear. The MES is a quite volatile measure; however, for the resilience measures we can see that the MES has a tendency to decrease after one year, indicating that these measures are indeed able to increase financial stability. Finally, we investigate the impact of the different macroprudential measures on the profitability indicators. For the NIM, we find that the impact is rather limited and mainly insignificant for all three measures. The negative impact on the NIM is most pronounced for the resilience measures. In contrast, the market-to-book value decreases considerably following macroprudential regulation, and this effect is visible for all three macroprudential tools. The effect is most pronounced for the liquidity tools: on average, the introduction of liquidity regulation decreases the market-to-book value with around 30 percentage points, and the effect is quite persistent.

In a next stage, we again interact the macroprudential policy tools with the monetary policy stance as calculated in Section 3.2. Figure 8 shows the results.

From the impulse responses we can see that the effects over the monetary regimes are similar in most of the cases. However, several results stand out. First, with respect to loan growth, liquidity

Figure 8. Robustness Check: Impact of a Tightening Shock in Different Macroprudential Policy Tools Conditional on the Stance of Monetary Policy on a Set of Bank Risk and Return Profile Variables



(continued)

Figure 8. (Continued)

Note: The response is expressed in percentage-point changes. To estimate the responses we estimate a local projections model with weighted least squares, where the weights are defined by an inverse propensity score model, as described in Section 3.1. We subdivide the index into different subindices based on their macroprudential objective. We distinguish between liquidity measures (liquidity coverage ratios, net stable funding ratios, loan-to-deposit ratios, and other liquidity requirements), credit growth measures (loan-to-value ratios, loan-to-income ratios, debt-service-to-income ratios, maturity and amortization restrictions, and risk weights on mortgage loans and commercial loans), and resilience measures (regulations under the CRR/CRD framework, capital buffers, taxes on financial institutions, and loan loss provisioning rules). The red (triangle) solid line indicates the response of the bank risk profile variables to a tightening in macroprudential policy conditional on the monetary policy stance being tight (i.e., the stance is below 0). The green (circle) solid line indicates the response of the bank risk profile variables to a tightening in macroprudential policy conditional on the monetary policy stance being loose (i.e., the stance is above 0). The yellow bars indicate the differential significance level between accommodating and tight monetary policy regime impulse responses at the 90 percent significance level. The gray area denotes the unconditional response to a macroprudential tightening. The confidence bounds for the unconditional impulse responses represent the 95 percent confidence intervals. For the conditional impulse responses we use 68 percent confidence intervals. The horizon is measured in quarters.

tools and resilience measures appear more effective during periods of tight monetary policy. Second, when we consider risk-weighted assets to total assets, we observe that banks reduce the riskiness of the asset portfolio both in loose and tight monetary regimes, following liquidity and resilience measures. However, the risk-shifting behavior, where banks shift to riskier assets in response to credit growth measures, is only present when monetary policy is tight. This is in line with the results of Becker and Ivashina (2014), who find that banks substitute loans with bonds when both lending standards and monetary policy are tight. Parallel to this result, also the Z-score falls following the credit growth measures, but only when monetary policy is tight. This result indicates that retail banks may become more vulnerable to business cycle shocks. A final noteworthy result is that the resilience measures appear more effective in decreasing the MES when monetary policy is tight.

6.2 An Alternative Measure for the Monetary Policy Stance: Taylor Rule

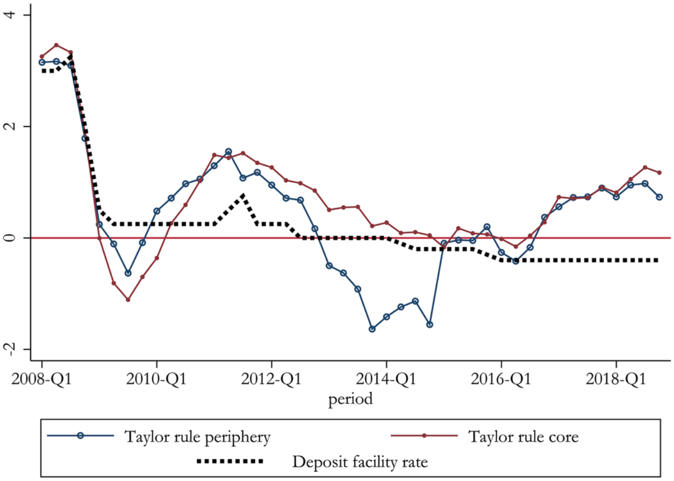
Using a VAR approach to estimate monetary policy shocks allows us to capture current and anticipated monetary policy changes. An alternative measure to capture the monetary stance is the use of a Taylor rule which indicates the optimal policy rate given deviations in inflation and output compared with their target levels. However, as we are interested in the monetary policy stance in the euro zone in the post-2008 period, we cannot use the policy rate because of the zero lower bound constraint. Therefore, we estimate a Taylor rule on the deposit facility rate which is not limited by the zero lower bound. To construct counterfactual interest rate path, we use the specification proposed by Clarida, Galí, and Gertler (1998), in which the target interest rate responds to deviations in inflation and output from their targets. We also incorporate an interest rate smoothing mechanism, in order to model the partial adjustment undertaken by central banks. We estimate the following model:

$$\begin{aligned} i_{i,t} = \rho i_{i,t-1} + (1-\rho)\alpha + \beta(\pi_{i,t} - \pi*_{i,t}) \\ + \lambda(y_{i,t} - y*_{i,t}) + \varepsilon_{i,t}, \end{aligned} \tag{9}$$

with $i_{i,t}$ the deposit facility rate, $\pi_{i,t} - \pi*_{i,t}$ the difference between the inflation rate and the target inflation in country i at time t, and $y_{i,t} - y*_{i,t}$ the output gap of country i at time t. We estimate a country-specific Taylor rule to account for different macroeconomic conditions in different euro zone member countries. As Nechio (2011) points out, a single policy rate is suboptimal, as the economic circumstances differ between countries, especially between core and peripheral countries.[16] The inflation rate corresponds to the OECD's annual growth rate of the Consumer Price Index. The output series corresponds to the Eurostat's Quarterly National Accounts' GDP data, in millions of 2010 euro, seasonally

[16]The peripheral countries in this exercise are Italy, Spain, Ireland, Greece, and Cyprus. The core countries represent all other euro zone countries.

Figure 9. Taylor Rule Estimated for All Euro Zone Countries Based on Country-Specific Macroeconomic Information

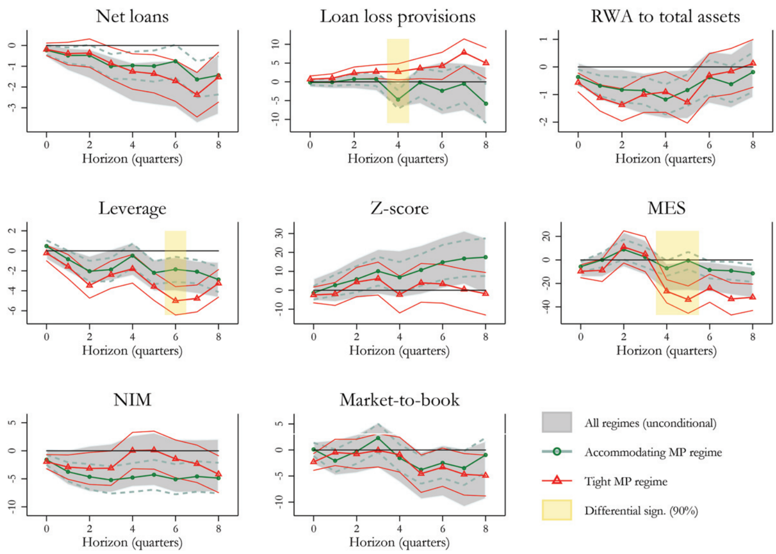


Note: In this graph, we show the average Taylor rule for the peripheral and core countries (blue circled and red crossed line, respectively) and the deposit facility rate (black dashed line).

adjusted. The output gap was obtained using a Hodrick-Prescott filter on the logarithm-transformed output series, multiplied by 100. After estimating Equation (9) we forecast the counterfactual interest rate paths. The results are presented in Figure 9, along with the observed deposit facility rate for the post-2008 period. We show the average Taylor rule for periphery and core countries.

We again interact the macroprudential policy shock with the monetary policy stance, as estimated by the Taylor rule. Figure 10 shows the impulse response functions. Looking at the impulse responses of credit growth, we again find that macroprudential policy appears to be somewhat more effective during periods of tight conventional monetary policy. In addition, the impact on the MES is also somewhat more negative during tight monetary stances. The negative impact of macroprudential policy on the bank profitability measures is more notable during times of

Figure 10. Robustness Check: Impact of a Tightening in the Macroprudential Index across Different Monetary Policy Regimes for a Sample of Euro Zone Banks between 2008:Q1 and 2018:Q4 on a Set of Bank Risk and Return Profile Variables



Note: The response is expressed in percentage-point changes. To construct the monetary policy regimes, we estimate a Taylor rule at the country level to account for macroeconomic differences across countries. The red (triangle) solid line indicates the response of the bank risk profile variables to a tightening in macroprudential policy conditional on the monetary policy stance being tight (i.e., the stance is below 0). The green (circle) solid line indicates the response of the bank risk profile variables to a tightening in macroprudential policy conditional on the monetary policy stance being loose (i.e., the stance is above 0). The yellow bars indicate the differential significance level between accommodating and tight monetary policy regime impulse responses at the 90 percent significance level. The gray area denotes the unconditional response to a macroprudential tightening. The confidence bounds for the unconditional impulse responses represent the 95 percent confidence intervals. For the conditional impulse responses we use 68 percent confidence intervals. The horizon is measured in quarters.

loose monetary policy: the NIM decreases more during periods of loose monetary policy than it does during periods of tight monetary policy. This result is comparable to the case where we interact the macroprudential policy shock with the VAR-based monetary policy stance; however, the differential effects are less significant.

6.3 An Alternative Measure for the Monetary Policy Stance: Identification Using External Instruments

An alternative way to identify monetary policy shocks is to use external sources of information, that is, external instruments. These external instruments can be thought of as noisy observations of, in this case, the monetary policy shock, but they are not correlated with other shocks. The instruments are thus not necessarily identical to the true monetary policy shock, as they might contain some measurement error, but, as long as they are uncorrelated with the other shocks in the system, they can be used to identify monetary policy (Rossi 2019). However, they need to be exogenous. We again start with a simple structural VAR model. The VAR model can be summarized as follows:

$$Y_t = \Pi(L)Y_{t-1} + \mu_t. \tag{10}$$

The reduced-form shocks and the structural shocks are linked to one another by some matrix B:

$$\mu_t = B\varepsilon_t. \tag{11}$$

We use the identification strategy of Stock and Watson (2012), Mertens and Ravn (2013), and Gertler and Karadi (2015). If we find an instrument Z for the shock of interest, we can identify the first column of B, and thus the impulse response functions of the system, without imposing zero (or other) restrictions. The following conditions need to be satisfied:

$$E(\varepsilon_{mp,t}, Z_t) = \phi \quad (relevant\,instrument) \tag{12}$$

$$E(\varepsilon_{other,t}, Z_t) = 0 \qquad (exogeneity). \tag{13}$$

We denote the structural monetary policy shock as $\varepsilon_{mp,t}$ and all other shocks as $\varepsilon_{other,t}$. The instrumental variable captures the exogenous component of the monetary policy shock. For more details and implementation, we refer to Stock and Watson (2012), Mertens and Ravn (2013), and Gertler and Karadi (2015). The VAR is estimated with the same five variables as the ones used in Section 3.2: the 5-year Spanish CDS spread, the 10-year German government bond yield, the 5-year forward inflation expectation based on inflation swap rates, an EU market index, and the VSTOXX index. The model is estimated from 2008:Q4 until 2018:Q4 at a daily frequency. In this case, we assume that monetary policy shocks affect the five-year Spanish CDS spread.

As an instrument for unconventional monetary policy, we use the monetary policy surprises as constructed by Altavilla et al. (2019). In this paper, the authors construct the "Euro Area Monetary Policy Event-Study Database (EA-MPD)."[17] The database contains tick data on a number of asset prices over relevant ECB policy windows that capture two different steps in the communication of the ECB. First, at 13:45 Central European Time (CET) a brief press release that only contains the decision on policy rates is published, while announcements of non-standard measures are mainly made as of 14:30 CET during a press conference and a Q&A session during which the ECB president reads a prepared text, the Introductory Statement (IS), on the rationale behind the decision. The database contains the change in a number of asset prices covering both the press release and press conference windows.[18] In particular, the EA-MPD provides information on the full OIS yield curve, ranging from one week to 20 years maturity; German, French, Italian, and Spanish government bond yields; the Eurostoxx; and several exchange rates. As we want to capture exogenous changes in unconventional monetary policy, we use the change in the Spanish 10-year government bond rate around ECB press conferences. This variable is presumably highly correlated with the five-year Spanish CDS spread, which makes it an optimal external instrument. We complement the database with three additional important central bank events which are non-meeting days: May 10, 2010 (Securities Market Programme (SMP)), August 8, 2011 (Reactivation of SMP), and

[17]The database is updated periodically and is available at https://www.ecb.europa.eu/pub/pdf/annex/Dataset_EA-MPD.xlsx.

[18]More specifically, the press conference window is described as the change in the median quote from the window 14:15–14:25 before the press conference to the median quote in the window 15:40–15:50 after it.

July 26, 2012 (Draghi's London speech: OMT).[19] The explanatory power of the instrument can then be examined by regressing the reduced-form VAR residuals of the monetary policy equation on a constant and the external instrument. The first-stage F-statistic of the instrument turns out to be 66.8, which is highly above the Stock, Wright, and Yogo (2002) threshold of an F-statistic of 10 for having possible weak instrument problems. We are therefore confident about our choice of an accurate instrument.

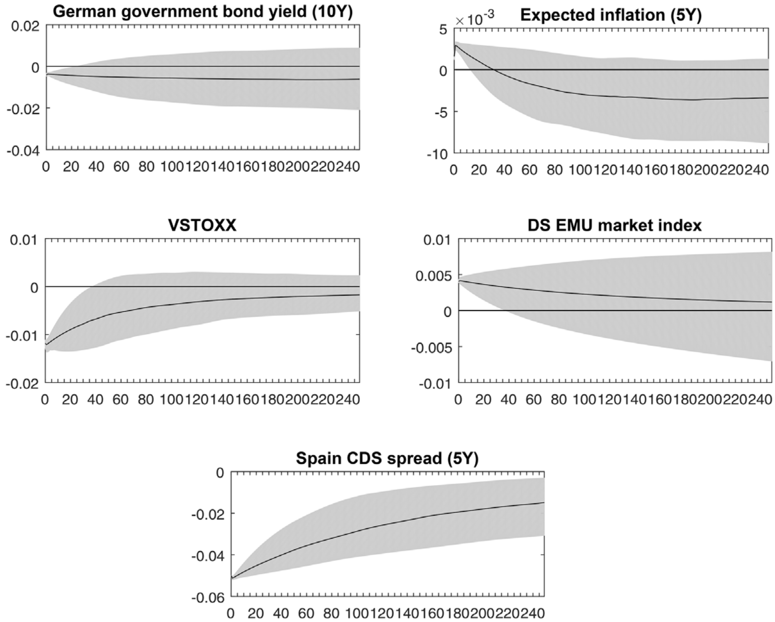
The monetary policy shock we finally extract from the VAR has a correlation of 82.6 percent with the monetary policy shock identified when using the "identification-through-heteroskedasticity" approach of Rigobon (2004). We again define a unit expansionary monetary policy shock as a shock that decreases the Spanish five-year CDS spread by 5 percent upon impact. The impulse responses are shown in Figure 11.

In line with the impulse responses obtained through identification based on heteroskedasticity of the structural shocks (Rigobon 2004), we find that an expansionary monetary policy shock increases long-term inflation expectations at impact as well as the value of the market index, while market-wide implied volatility ($VSTOXX$) decreases. In contrast to the identification-through-heteroskedasticity impulse responses, we find that the German 10-year government bond yield slightly decreases at time 0, which indicates that policy actions that affect the long-run safe yields, such as QE, receive a higher weight when estimating monetary policy through the "identification-through-external-instruments" approach.

After having estimated the alternative monetary policy stance, we transform the series to a quarterly frequency by taking the average of the series over the corresponding quarter. To convert the monetary policy shock to a monetary policy *stance,* we again calculate the contribution of the monetary policy shock to changes in the Spanish five-year CDS spread. We then interact this monetary policy stance with the macroprudential shock in order to evaluate the effectiveness of macroprudential policy across different monetary policy regimes. The results are given in Figure 12.

[19]For these days we use the daily change in the Spanish 10-year government bond yield.

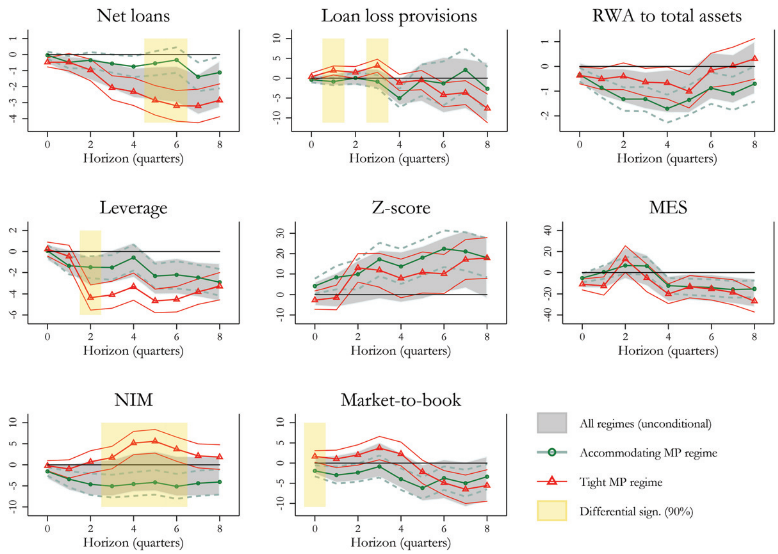
Figure 11. Impulse Response Function of the Variables to a Unit Monetary Policy Shock Which Decreases the Spanish CDS Spread by 5 Percent



Note: The monetary policy shock is obtained using the identification-through-external-instruments approach (Stock and Watson 2012; Mertens and Ravn 2013; Gertler and Karadi 2015). Gray areas represent 95 percent confidence intervals that are obtained through a wild bootstrap procedure. Because both the first- and second-stage regressions are included in the bootstrapping procedure, we avoid a potential "generated regressor" problem. The horizontal axis represents the horizon of the impulse response function in working days, i.e., the IRFs are plotted for a horizon of 240 days.

We again find that macroprudential policy appears to be more effective during periods of tight conventional monetary policy. The effect is most pronounced for bank loan growth. In addition, the impact on the MES is also negative during tight monetary stances. The negative impact of macroprudential policy on the bank profitability measures is more notable during times of loose monetary

Figure 12. Robustness Check: Impact of a Tightening in the Macroprudential Index across Different Monetary Policy Regimes for a Sample of Euro Zone Banks between 2008:Q1 and 2018:Q4 on a Set of Bank Risk and Return Profile Variables



Note: The response is expressed in percentage-point changes. The monetary policy shock is obtained using the identification-through-external-instruments approach (Stock and Watson 2012; Mertens and Ravn 2013; Gertler and Karadi 2015). The monetary policy stance is calculated as the historical contribution of the monetary policy shock to changes in the five-year Spanish CDS spread. The red (triangle) solid line indicates the response of the bank risk profile variables to a tightening in macroprudential policy conditional on the monetary policy stance being tight (i.e., the stance is below 0). The green (circle) solid line indicates the response of the bank risk profile variables to a tightening in macroprudential policy conditional on the monetary policy stance being loose (i.e., the stance is above 0). The yellow bars indicate the differential significance level between accommodating and tight monetary policy regime impulse responses at the 90 percent significance level. The gray area denotes the unconditional response to a macroprudential tightening. The confidence bounds for the unconditional impulse responses represent the 95 percent confidence intervals. For the conditional impulse responses we use 68 percent confidence intervals. The horizon is measured in quarters.

policy: the market-to-book value decreases more during periods of loose monetary policy than it does during periods of tight monetary policy. These results are comparable to those we find when we interact the macroprudential policy shock with the VAR-based monetary policy stance. Hence, our findings are robust to alternative identifications of the monetary policy stance.

7. Conclusion

Macroprudential policy is in vogue. Since the global financial crisis, macroprudential policies have gained prominence worldwide as a tool to maintain financial stability. In the euro area, the institutional framework has been adapted through the implementation of the Banking Union and the designation of macroprudential authorities in the member states. Borrower-related measures, such as LTV caps, and lender-related instruments, such as countercyclical capital buffers, have been introduced in several countries in order to deal with financial risks in the banking sector. Since both monetary policy and macroprudential policy may affect risk behavior by banks, it is important to establish whether or not the effectiveness of macroprudential policy varies across different monetary policy stances.

We tackle this important policy question empirically by analyzing the impact of macroprudential policy on the risk and return profile of euro-area banks and by examining the interaction between monetary and macroprudential policy over the 2008–18 period. Our sample consists of 140/64 euro-area banks for which we consider a coherent set of accounting-based (140 banks) and market-based (64 banks) indicators of the banks' risk and profit profile. For the identification of macroprudential policy, we apply an inverse propensity score weighting estimation in order to avoid endogeneity issues. The monetary policy stance is captured by a structural VAR in order to account for current and anticipated macroeconomic and financial market conditions. We use the local projections approach to assess the impact of macroprudential policy, and their interaction, on bank risk and return profiles over a two-year impact horizon.

The main findings can be summarized as follows. First, considered in isolation, we confirm that macroprudential policy is effective in restraining bank risk, as intended by the macroprudential authorities. Tightening macroprudential measures are typically associated with less lending and lower bank asset risk and these features translate into lower overall bank risk, both accounting based (Z-score increases) and market based (MES decreases). However, the downside is that the announcement of macroprudential tools is accompanied by lower bank profitability over the projection horizon, leading to a significant decrease in the market-to-book ratio, reflecting the market perception that imposing constraints on banks causes stressed current and future bank profitability. When considering the banks' business model, we find that for both lending and profitability the effects are more pronounced for retail banks than for their non-retail counterparts. This is not unexpected, since the banks with a retail profile are most active in traditional lending, which is the focus of macroprudential measures targeting credit growth. Nevertheless, the negative consequences for the net interest margin and the market-to-book ratio are also more pronounced for retail-oriented banks, which may affect their future viability. This conclusion indicates that regulatory authorities should mind the business model of banks when imposing constraints.

Finally, we assess whether the effectiveness of macroprudential measures varies conditional on the stance of monetary policy. We find that when tightening macroprudential measures are announced in a period characterized by a restrictive monetary policy stance, the two policies reinforce each other in lowering credit growth, as intended by both the monetary and the macroprudential authorities. Moreover, in terms of bank risk, the behavior of the bank risk profile variables is consistent with improved bank stability. From a policy perspective, the most interesting case is when there is a potential trade-off between monetary and macroprudential policy, because the prevailing environment in the post-2008 era is characterized by the simultaneous introduction of restrictive macroprudential measures following the financial and sovereign crises in Europe as well as unprecedented conventional and unconventional monetary policies by the central bank. In this case, we document that loan growth increases, suggesting that the transmission of macroprudential policy to credit growth is affected by the presence of loose monetary policy. For the central bank, this is the intended outcome since its actions are geared towards stimulating lending to the real economy. Interestingly, while accommodating monetary policy may entail incentives for banks to take more risk, our results indicate that macroprudential measures were sufficiently strong to deter banks from excessive risk-taking. In other words, macroprudential policy succeeds in maintaining bank stability also in periods of monetary accommodation. Yet, there is an important downside: we observe a marked deterioration of the banks' market-to-book value as a reflection of the investors' conviction that low-for-long interest rates ultimately compress bank interest margins and put their profitability and franchise value under stress. Our conclusion is that the combination of restrictive macroprudential policies and prolonged monetary accommodation may turn out to be detrimental for bank health and, ultimately, financial stability.

Our main findings are corroborated when we estimate the monetary policy stance with a Taylor rule or when we use the "identification-through-external-instruments" approach. When we consider the impact of specific macroprudential policy tools, we find that credit growth measures, such as loan-to-value ratios, have an immediate and stronger negative impact on loan growth than liquidity regulation or measures aimed at the resilience of banks, such as capital regulation. However, we also find evidence for risk-shifting behavior by banks confronted with targeted credit measures: banks increase the riskiness of the loan portfolio in response to credit constraints. In trying to comply with the rules, these banks may engage in riskier activities by, e.g., shifting to more risky corporate lending or securities.

References

Acharya, V., R. Engle, and M. Richardson. 2012. "Capital Shortfall: A New Approach to Ranking and Regulating Systemic Risks." *American Economic Review* 102 (3): 59–64.

Acharya, V. V., K. Bergant, M. Crosignani, T. Eisert, and F. McCann. 2018. "The Anatomy of the Transmission of Macroprudential Policies: Evidence from Ireland." Working Paper.

Acharya, V. V., L. H. Pedersen, T. Philippon, and M. Richardson. 2017. "Measuring Systemic Risk." *Review of Financial Studies* 30 (1): 2–47.

Aiyar, S., C. W. Calomiris, and T. Wieladek. 2014. "Does Macroprudential Regulation Leak? Evidence from a UK Policy Experiment." *Journal of Money, Credit and Banking* 46 (s1): 181–214.

———. 2016. "How Does Credit Supply Respond to Monetary Policy and Bank Minimum Capital Requirements?" *European Economic Review* 82 (April): 142–65.

Akinci, O., and J. Olmstead-Rumsey. 2018. "How Effective Are Macroprudential Policies? An Empirical Investigation." *Journal of Financial Intermediation* 33 (January): 33–57.

Alam, Z., A. Alter, J. Eiseman, R. Gelos, H. Kang, M. Narita, E. Nier, and N. Wang. 2019. "Digging Deeper—Evidence on the Effects of Macroprudential Policies from a New Database." IMF Working Paper No. 2019/066.

Albertazzi, U., F. Barbiero, D. Marques-Ibanez, A. Popov, C. Rodriguez D'acri, and T. Vlassopoulos. 2020. "Monetary Policy and Bank Stability: The Analytical Toolbox Reviewed." ECB Working Paper No. 2377.

Altavilla, C., M. Boucinha, and J. L. Peydró. 2018. "Monetary Policy and Bank Profitability in a Low Interest Rate Environment." *Economic Policy* 33 (96): 531–86.

Altavilla, C., L. Brugnolini, R. S. Gürkaynak, R. Motto, and G. Ragusa. 2019. "Measuring Euro Area Monetary Policy." *Journal of Monetary Economics* 108 (December): 162–79.

Altavilla, C., D. Giannone, and M. Lenza. 2016. "The Financial and Macroeconomic Effects of the OMT Announcements." *International Journal of Central Banking* 12 (3): 29–57.

Altunbas, Y., M. Binici, and L. Gambacorta. 2018. "Macroprudential Policy and Bank Risk." *Journal of International Money and Finance* 81 (March): 203–20.

Angelini, P., S. Neri, and F. Panetta. 2014. "The Interaction between Capital Requirements and Monetary Policy." *Journal of Money, Credit and Banking* 46 (6): 1073–1112.

Arai, N. 2017. "The Effects of Monetary Policy Announcements at the Zero Lower Bound." *International Journal of Central Banking* 13 (2): 159–96.

Auer, R., and S. R. G. Ongena. 2019. "The Countercyclical Capital Buffer and the Composition of Bank Lending." CEPR Discussion Paper No. 13942.

Austin, P.C. 2009. "Type I Error Rates, Coverage of Confidence Intervals, and Variance Estimation in Propensity-Score Matched Analyses." *International Journal of Biostatistics* 5 (1): Article 13.

Banerjee, R. N., and H. Mio. 2018. "The Impact of Liquidity Regulation on Banks." *Journal of Financial Intermediation* 35 (Part B, July): 30–44.

Becker, B., and V. Ivashina. 2014. "Cyclicality of Credit Supply: Firm Level Evidence." *Journal of Monetary Economics* 62 (1): 76–93.

Borio, C., L. Gambacorta, and B. Hofmann. 2017. "The Influence of Monetary Policy on Bank Profitability." *International Finance* 20 (1): 48–63.

Borio, C. E. V. 2003. "Towards a Macroprudential Framework for Financial Supervision and Regulation?" *CESifo Economic Studies* 49 (2): 181–215.

Brownlees, C., and R. F. Engle. 2017. "SRISK: A Conditional Capital Shortfall Measure of Systemic Risk." *Review of Financial Studies* 30 (1): 48–79.

Budnik, K., and J. Kleibl. 2018. "Macroprudential Regulation in the European Union in 1995-2014: Introducing a New Data Set on Policy Actions of a Macroprudential Nature." ECB Working Paper No. 2123.

Caporale, G. M., A. Cipollini, and P. O. Demetriades. 2005. "Monetary Policy and the Exchange Rate during the Asian Crisis: Identification through Heteroscedasticity." *Journal of International Money and Finance* 24 (1): 39–53.

Cappelletti, G., A. P. Marques, P. Varasso, Z. Budrys, and J. Peeters. 2019. "Impact of Higher Capital Buffers on Banks' Lending and Risk-taking: Evidence from the Euro Area Experiments." ECB Working Paper No. 2292.

Caruana, J. 2010. "Macroprudential Policy: Working Towards a New Consensus." Remarks at the high-level meeting on "The Emerging Framework for Financial Regulation and Monetary Policy," jointly organized by the BIS's Financial Stability Institute and the IMF Institute, Washington, DC, April 23.

Cerutti, E., S. Claessens, and L. Laeven. 2017. "The Use and Effectiveness of Macroprudential Policies: New Evidence." *Journal of Financial Stability* 28 (February): 203–24.

Cizel, J., J. Frost, A. Houben, and P. Wierts. 2016. "Effective Macroprudential Policy: Cross-Sector Substitution from Price and Quantity Measures." IMF Working Paper No. 16/94.

Claessens, S., N. Coleman, and M. Donnelly. 2018. "'Low-For-Long' Interest Rates and Banks' Interest Margins and Profitability: Cross-country Evidence." *Journal of Financial Intermediation* 35 (Part A, July): 1–16.

Clarida, R., J. Galí, and M. Gertler. 1998. "Monetary Policy Rules in Practice: Some International Evidence." *European Economic Review* 42 (6): 1033–67.

Crockett, A. 2000. "Marrying the Micro- and Macro-prudential Dimensions of Financial Stability." Speech before the Eleventh International Conference of Banking Supervisors, Basel, Switzerland, September 21.

David, J., G. Revelo, Y. Lucotte, and F. Pradines-Jobet. 2019. "Macroprudential and Monetary Policies: The Need to Dance the Tango in Harmony." LEO Working Papers/DR LEO 2691, Orleans Economics Laboratory / Laboratoire d'Economie d'Orleans (LEO), University of Orleans.

Dell'Ariccia, G., D. Igan, L. Laeven, and H. Tong. 2016. "Credit Booms and Macrofinancial Stability." *Economic Policy* 31 (86): 299–355.

European Central Bank. 2020. *Financial Stability Review* (November).

Farnè, M., and A. Vouldis. 2017. "Business Models of the Banks in the Euro Area." ECB Working Paper No. 2070.

Forbes, K., M. Fratzscher, R. Straub, A. Chari, K. Dominguez, J. Frankel, M. Garcia, et al. 2015. "Capital-flow Management Measures: What Are They Good For?" *Journal of International Economics* 96 (S1): S76–S97.

Gambacorta, L., and A. Murcia. 2019. "The Impact of Macroprudential Policies and Their Interaction with Monetary Policy: An Empirical Analysis Using Credit Registry Data." IFC Bulletin No. 49.

Gelain, P., and P. Ilbas. 2017. "Monetary and Macroprudential Policies in an Estimated Model with Financial Intermediation." *Journal of Economic Dynamics and Control* 78 (May): 164–89.

Gertler, M., and P. Karadi. 2015. “Monetary Policy Surprises, Credit Costs, and Economic Activity.” *American Economic Journal: Macroeconomics* 7 (1): 44–76.

Gilchrist, S., and E. Zakrajšek. 2013. “The Impact of the Federal Reserve’s Large-scale Asset Purchase Programs on Corporate Credit Risk.” *Journal of Money, Credit and Banking* 45 (s2): 29–57.

Heider, F., F. Saidi, and G. Schepens. 2019. “Life Below Zero: Bank Lending Under Negative Policy Rates.” *Review of Financial Studies* 32 (10): 3728–61.

Idier, J., G. Lamé, and J. S. Mésonnier. 2014. “How Useful Is the Marginal Expected Shortfall for the Measurement of Systemic Exposure? A Practical Assessment.” *Journal of Banking and Finance* 47 (1): 134–46.

Igan, D., and H. Kang. 2011. “Do Loan-to-Value and Debt-to-Income Limits Work? Evidence from Korea.” IMF Working Paper No. 11/297.

Jordà, O. 2005. “Estimation and Inference of Impulse Responses by Local Projections.” *American Economic Review* 95 (1): 161–82.

Jordà, O., and A. M. Taylor. 2016. “The Time for Austerity: Estimating the Average Treatment Effect of Fiscal Policy.” *Economic Journal* 126 (590): 219–55.

King, M. R. 2013. “The Basel III Net Stable Funding Ratio and Bank Net Interest Margins.” *Journal of Banking and Finance* 37 (11): 4144–56.

Kok, C., C. Móré, and M. Petrescu. 2016. “Recent Trends in Euro Area Banks’ Business Models and Implications for Banking Sector Stability.” *Financial Stability Review* (May): 1–13.

Kuttner, K. N. and I. Shim. 2016. “Can Non-interest Rate Policies Stabilize Housing Markets? Evidence from a Panel of 57 Economies.” *Journal of Financial Stability* 26 (October): 31–44.

Lamers, M., F. Mergaerts, E. Meuleman, and R. Vander Vennet. 2019. “The Tradeoff between Monetary Policy and Bank Stability.” *International Journal of Central Banking* 15 (2): 1–42.

Lanne, M., and H. Lütkepohl. 2008. “Identifying Monetary Policy Shocks via Changes in Volatility.” *Journal of Money, Credit and Banking* 40 (6): 1131–49.

Lim, C., F. Columba, A. Costa, P. Kongsamut, A. Otani, M. Saiyid, T. Wezel, and X. Wu. 2011. “Macroprudential Policy: What

Instruments and How to Use Them? Lessons from Country Experiences." IMF Working Paper No. 11/238.

Martinez-Miera, D., and R. Repullo. 2019. "Monetary Policy, Macroprudential Policy, and Financial Stability." ECB Working Paper No. 2297.

Mergaerts, F., and R. Vander Vennet. 2016. "Business Models and Bank Performance: A Long-term Perspective." *Journal of Financial Stability* 22 (February): 57–75.

Mertens, K., and M. O. Ravn. 2013. "The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States." *American Economic Review* 103 (4): 1212–47.

Meuleman, E., and R. Vander Vennet. 2020. "Macroprudential Policy and Bank Systemic Risk." *Journal of Financial Stability* 47 (April): Article 100724.

Nechio, F. 2011. "Monetary Policy When One Size Does Not Fit All." FRBSF Economic Letter No. 2011-18.

Paoli, B. D., and M. Paustian. 2017. "Coordinating Monetary and Macroprudential Policies." *Journal of Money, Credit and Banking* 49 (2-3): 319–49.

Peersman, G., and F. Smets. 2003. "The Monetary Transmission Mechanism in the Euro Area: Evidence from VAR Analysis." In *Monetary Policy Transmission in the Euro Area: A Study by the Eurosystem Monetary Transmission Network,* ed. I. Angeloni, A. Kashyap, and B. Mojon, 36–55. Cambridge University Press.

Poghosyan, T. 2019. "How Effective is Macroprudential Policy? Evidence from Lending Restriction Measures in EU Countries." IMF Working Paper No. 19/45.

Popoyan, L., M. Napoletano, and A. Roventini. 2017. "Taming Macroeconomic Instability: Monetary and Macro-prudential Policy Interactions in an Agent-based Model." *Journal of Economic Behavior and Organization* 134 (February): 117–40.

Reinhardt, D., and R. Sowerbutts. 2015. "Regulatory Arbitrage in Action: Evidence from Banking Flows and Macroprudential Policy." Staff Working Paper No. 546, Bank of England.

Richter, B., M. Schularick, and I. Shim. 2018. "The Macroeconomic Effects of Macroprudential Policy." BIS Working Paper No. 740.

Rigobon, R. 2004. "The Impact of Monetary Policy on Asset Prices." *Journal of Monetary Economics* 51 (8): 1553–75.

Roengpitya, R., N. Tarashev, K. Tsatsaronis, and A. Villegas. 2017. “Bank Business Models: Popularity and Performance.” BIS Working Paper No. 682.

Rogers, J. H., C. Scotti, and J. H. Wright. 2014. “Evaluating Asset-market Effects of Unconventional Monetary Policy: A Multi-country Review.” *Economic Policy* 29 (80): 749–99.

Rosenbaum, P. R., and D. B. Rubin. 1983. “The Central Role of the Propensity Score in Observational Studies for Causal Effects.” *Biometrika* 70 (1): 41–55.

Rossi, B. 2019. “Identifying and Estimating the Effects of Unconventional Monetary Policy: How to Do It and What Have We Learned?” CEPR Discussion Paper No. 14064.

Rostagno, M., C. Altavilla, G. Carboni, W. Lemke, R. Motto, A. S. Guilhem, and J. Yiangou. 2019. “A Tale of Two Decades: The ECB’s Monetary Policy at 20.” ECB Working Paper No. 2346.

Rubio, M., and J. A. Carrasco-Gallego. 2014. “Macroprudential and Monetary Policies: Implications for Financial Stability and Welfare.” *Journal of Banking and Finance* 49 (December): 326–36.

Shim, I., B. Bogdanova, J. Shek, and A. Subelyte. 2013. “Database for Policy Actions on Housing Markets.” *BIS Quarterly Review* (September): 83–95.

Smets, F. 2014. “Financial Stability and Monetary Policy: How Closely Interlinked?” *International Journal of Central Banking* 10 (2): 263–300.

Soenen, N., and R. Vander Vennet. 2022. “ECB Monetary Policy and Bank Default Risk.” *Journal of International Money and Finance* 122 (April): Article 102571.

Stock, J., and M. W. Watson. 2012. “Disentangling the Channels of the 2007-09 Recession: Comments and Discussion.” *Brookings Papers on Economic Activity* (Spring): 81–156.

Stock, J. H., J. H. Wright, and M. Yogo. 2002. “A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments.” *Journal of Business and Economic Statistics* 20 (4): 518–29.

Tressel, T., and Y. S. Zhang. 2016. “Effectiveness and Channels of Macroprudential Instruments: Lessons from the Euro Area.” IMF Working Paper No. 16/04.

Wright, J. H. 2012. "What Does Monetary Policy Do to Long-term Interest Rates at the Zero Lower Bound?" *Economic Journal* 122 (564): F447–F466.

Zhang, L., and E. Zoli. 2016. "Leaning against the Wind: Macroprudential Policy in Asia." *Journal of Asian Economics* 42 (February): 33–52.

What Do Online Listings Tell Us about the Housing Market?*

Michele Loberto,[a] Andrea Luciani,[a] and Marco Pangallo[b]
[a]Banca d'Italia
[b]Sant'Anna School of Advanced Studies

Since the Great Recession, central banks and macroprudential authorities have been devoting much more attention to the housing market. To properly assess trends and risks, policymakers need detailed, timely, and granular information on demand, supply, and transactions. This information is hardly provided by traditional survey or administrative data. In this paper, we argue that data coming from housing sales advertisements (ads) websites can be used to overcome some existing deficiencies. Using a large data set of ads in Italy, we provide the first comprehensive analysis of the problems and potential of these data. We show how machine learning tools can correct a specific bias of online listings, namely the proliferation of duplicate ads that refer to the same housing unit, increasing the representativeness and reliability of these data. We then show how the timeliness, granularity, and online nature of these data make it possible to monitor in real-time housing demand, supply, and prices.

JEL Codes: C44, C81, C31, R21, R31.

*The views expressed in this paper are those of the authors and do not reflect the views of Banca d'Italia. We are extremely grateful to Immobiliare.it for providing the data and for their assistance. We thank all the seminar and workshop participants at Banca d'Italia, Bank of England, Danmarks Nationalbank, Eurostat, and JRC Centre for Advanced Studies (JRC-CAS), for useful comments and discussions. Special thanks go to Attilio Mattiocco, Juri Marcucci, and Diana Nicoletti for their assistance and feedback in all the phases of this project. An earlier version of this paper circulated under the title "The Potential of Big Housing Data: An Application to the Italian Real-Estate Market." Marco Pangallo performed most of his work on this project while he was affiliated with the Mathematical Institute and Institute for New Economic Thinking at the University of Oxford. Marco Pangallo acknowledges support from the James S. McDonnell Foundation 21st Century Science Initiative in Understanding Dynamic and Multiscale Systems — Postdoctoral Fellowship Award. All mistakes are our own.
Author e-mails: michele.loberto@bancaditalia.it, andrea.luciani@bancaditalia.it, marco.pangallo@santannapisa.it.

1. Introduction

Following the Great Recession of 2007–09, policymakers have been following housing market trends more closely. Housing is the main component of household wealth and is one of the main drivers of private consumption (Mian, Rao, and Sufi 2013). Moreover, housing markets are highly relevant for financial stability, as housing price bubbles have historically been dramatically damaging events (Jordà, Schularick, and Taylor 2015). Therefore, assessing the trends and risks in the housing markets is crucial for central banks.

Choosing the right policy mix hinges on the availability of detailed, granular, and timely information about housing demand and supply. For instance, policymakers may want to choose different policies depending on whether excessive housing price growth is due to exuberance on the demand side or to low supply. Additionally, this information should be available at the country level but also at a more granular level. Indeed, housing bubbles may occur in local markets or even in specific market segments (Landvoigt, Piazzesi, and Schneider 2015), even if no signs of imbalances are detected at the country level. Furthermore, to take timely action, information on demand and supply should ideally be available in real time.

Collecting detailed, granular, and timely data on the housing market has been traditionally challenging. Microdata on home sales are available to researchers only in a few countries and with a significant temporal lag. Moreover, they may show limitations in the spatial and temporal dimension, or in their informational content.[1] Most importantly, extracting information on demand and supply from home sales may require strong identifying assumptions, as transactions represent "equilibrium" points.

In this paper, we show how online data from marketplace websites (such as Zillow) can fill some gaps, providing valuable

[1]For example, the United Kingdom is one of the few countries where microdata are available. However, those data contain inadequate information about the physical characteristics of homes. Many papers on the U.S. housing market have used data from Multiple Listing Services (MLS), which are pools of real estate brokers sharing information about properties to make the matching between buyers and sellers more efficient. These data report details on the price and physical characteristics of homes. Yet, because there exist many different MLSs, studies usually focused on a limited geographical area (for example, see Han and Strange 2015). In Italy, administrative microdata on housing transactions are not available for research because of privacy concerns.

information to support policy choices. Our aim is twofold. First, we investigate the measurement issues and show how to improve the representativeness and reliability of online listings data. Second, we show how these data can provide detailed, timely, and granular information on housing demand, supply, and prices, which would be harder to get from traditional sources.

Our analysis is based on a large database containing all housing sales advertisements (ads) published on Immobiliare.it, the most popular online portal for real estate services in Italy. Similar analyses could be performed using data from similar websites, such as Zillow or Trulia in the United States or Zoopla in the United Kingdom. From these sources, we can retrieve real-time and detailed information about listed dwellings, including physical characteristics, location, time on market, and asking prices. Compared with traditional listing data collected by real estate professional associations, online listing data also allow for real-time monitoring of buyers' search behavior (Piazzesi, Schneider, and Stroebel 2020), as we discuss below.

Despite the wealth of information on the housing market that these data provide, data generation could be biased in several ways. As with all non-survey or non-universal administrative data, online listings data may lead to non-representative results or feature measurement error. Additionally, online listings posted on marketplace websites may have a peculiar issue: There could be two or more *duplicate* ads referring to the same housing unit. This is a common problem in our data set, but we think that it is not just a website-specific or Italian-specific issue. For instance, Kolbe et al. (2021) report the same issue for ads on ImmobilienScout24, the largest real estate platform in Germany. To identify duplicates, we propose a procedure using machine learning algorithms, as duplicate identification based on geographical coordinates or heuristic rules is not sufficiently precise. We show that the duplicate bias is not a serious issue for a few applications, such as monitoring housing market trends at the country level. However, we also demonstrate that it is a serious problem when granularity and high frequency matter for identification. As a consequence, the results of regressions that use duplicate ads instead of unique listings may be greatly biased.[2]

[2]For instance, we show that the odds of reducing the asking price if the property on sale does not attract enough interest is reduced tenfold when considering

In terms of applications of online listings data, we first show that the number of views to the ads' webpages is a good proxy of housing demand. Indeed, when individuals visit the webpage of an ad, they convey information about the characteristics, location, and price of the home they are searching for. By aggregating all this information, we can understand which area households are searching more intensely and what they are looking for.[3] At the micro (dwelling) level, high online interest predicts lower time on market and lower probability that a downward revision of the asking price occurs. By aggregating the number of page views, we can build a measure of market tightness. We show that this indicator is a good predictor of housing prices, as suggested by the recent literature (Carillo, de Wit, and Larson 2015; Wu and Brynjolfsson 2015; van Dijk and Francke 2018).[4]

We also show that online listings are an effective tool for monitoring the number of homes for sale (so-called market inventory). Although housing supply is usually defined as the total stock of homes (Glaeser and Gyourko 2018), policymakers should focus on market inventory as a measure of short-medium run housing supply. Home sales variation is mainly driven by changes in the number of listings, and households take market conditions into account before deciding whether to put their home up for sale (Ngai and Sheedy 2020). Moreover, also the composition of market inventory changes over time with market conditions. We show that the quality of listed existing homes improves with better market conditions, as measured by housing price growth.

Finally, we discuss under which conditions listing prices can be used to nowcast and forecast sale prices (Anenberg and Laufer 2017). We stress that a good estimate of the average discount to the initial asking price is needed. When this discount is constant, asking

duplicate ads instead of unique listings. This is because brokers are likely to post a new ad when revising the price, and if this is not taken into account price revisions look excessively rare.

[3]This proxy of housing demand is complementary to web searches, which have been used recently by Piazzesi, Schneider, and Stroebel (2020).

[4]We already investigated the possibility of using webpages' views as a proxy of housing demand in a previous publication directed at a different audience (Pangallo and Loberto 2018). Here, we adopt a different econometric approach, and the sample is twice as long, highlighting the robustness of our findings.

prices are a good proxy of transaction prices. However, since discount changes with market conditions, asking prices may be a poor predictor of sale prices. In this case, auxiliary information is needed to improve the forecast (Anenberg and Laufer 2017; Lyons 2019).

This paper is organized as follows. Section 2 illustrates the main institutional details and trends of the Italian housing market. Section 3 describe the Immobiliare.it ads data set and discuss the main issues with online ads. In Section 4 we show how online listing data can be used to measure demand, supply, and housing prices. Section 5 concludes.

2. The Italian Housing Market

In this section, we describe the main trends and institutions of the Italian housing market.

The 2011 sovereign debt crisis had a strong impact on the housing market. From 2011 to 2013, home purchases and sales fell by about 30 percent and only resumed growth in 2014 (Figure B.1 in Appendix B). Housing prices experienced a more moderate but more persistent decline (Figure B.1, panel B): Between 2011 and 2018, they fell cumulatively by about 20 percent. The average time on market surged from seven to nine months between 2010 and 2015, but returned to pre-crisis levels since mid-2016 (Figure B.1, panel C). The average discount obtained by buyers relative to sellers' asking prices has followed a similar pattern, varying between 10 and 15 percent. Trends in home sales and prices diverged across geographic areas. In 2016–18, which is the period we primarily focus on in this paper, home prices were still declining in most cities. However, they had returned to growth in many large cities (Figure B.1, panel D).

In Italy, about half of all households' home purchases are financed through a mortgage loan. The relative amount of the mortgage is generally not very large: The average loan-to-value is about 60 percent. Transaction costs associated with purchasing a home depend on several factors. Costs include transaction taxes, notary fees, brokerage fees, and mortgage-related costs. Estimating the impact of transaction costs on the value of a purchased home is difficult.[5]

[5] Some costs are not proportional to the value of the home. Other costs are partially tax deductible. Moreover, many of these costs are lower if the new owner

Considering a home to be occupied by the owner and worth 100,000 euros, transaction costs can be up to 13 percent. In other cases, transaction costs can be up to 20 percent (e.g., dwellings purchased for investment purposes).

Most importantly for the focus of this paper, about half of total home sales are intermediated by real estate brokers. Real estate brokers are essential in cities and metropolitan areas. By contrast, in suburbs and rural areas most transactions do not involve an intermediary. Moreover, in Italy open listings agreements are possible, in the sense that two or more real estate agents are entitled to sell the same dwelling.

List prices are not legally binding, and the seller can always refuse to sell to a potential buyer. In general, the buyer and the seller negotiate the final price and other contractual arrangements. When a broker is involved in a sale, the seller cannot simultaneously negotiate with multiple buyers, which rules out bidding wars. Usually, the final price is below the listing price. Indeed, during 2016–18, the average discount compared with the initial asking price was about 12 percent, and the final price was equal to or higher than the initial asking price only in about 5 percent of transactions (Italian Housing Market Survey).[6]

3. Data

We analyze a data set of home listings published on Immobiliare.it, the largest online portal for real estate services in Italy. This data set covers the whole country. However, since small towns and villages may have representativeness issues, we only consider listings in the 109 main cities that are capitals of the NUTS-3 Italian regions. About 18 million people live in these cities, and the number of home sales is about one-third of all transactions in Italy.

---

bought the home as a primary residence. The total cost depends on home value, buyer income, and the reason for the purchase.

[6]The sale price may be higher than the asking price for various reasons other than bidding wars. For instance, the buyer may have particular requirements for finalizing the sale or taking possession of the home. Alternatively, the transaction includes additional amenities compared with the initial offer (e.g., a garage).

Immobiliare.it provides us with weekly snapshots of all ads visible on their website every Monday, from January 4, 2016 until December 31, 2018. For 2015 only quarterly snapshots are available.[7]

For each ad, we have detailed information about the physical characteristics and exact location of the dwelling (see Appendix A for the complete list of variables). We keep track of all variations concerning asking prices and number of times that the webpage of the ad has been visited (*clicks*). We also know the date when the ad was created and the date when it was removed. Unfortunately, we do not know if a property was sold or withdrawn from the market.

The data set counts 1,402,798 ads. Since we observe ads at a weekly frequency, the total number of records is almost 28 million. Most ads remain unchanged between two weekly snapshots, as the average turnover is about 5 percent. About 92 percent of the ads are posted by real estate agents; the remaining ads are posted by households or construction firms.

We divide the territory of each city into local housing markets using the partition developed by OMI, a branch of the Italian Tax Office. The elements of this partition are contiguous areas of the city that satisfy strict requirements in terms of homogeneity of housing prices, urban characteristics, socioeconomic characteristics, and the endowment of services and urban infrastructures. This partition is periodically revised to satisfy these criteria and better approximate local housing markets. The latest revision dates back to 2014. Thus, unlike census tracts, these zones can be considered as “local housing markets.” For each of these zones, OMI estimates the minimum and maximum housing price per square meter on a six-month basis. Table B.3 in Appendix B reports some descriptive statistics about these local housing markets.

Finally, we use information coming from the Italian Housing Market Survey, a quarterly survey covering a large sample of real estate agents. A detailed discussion about all data sources can be found in Appendix A.

---

[7] Data are available for the following days in 2015: January 5, April 25, September 7, December 28.

3.1 Duplicate Listings

The use of new, unconventional data sources is becoming increasingly common. However, using these data requires identifying potential biases that could make the data unrepresentative of the phenomenon under analysis.

The main concern with several housing marketplace websites—such as Immobiliare.it, Craigslist, Zoopla, ImmobilienScout24, Idealista, and many others—is the difficulty to strictly control the content of the ads published by the users. These websites are market platforms that allow home sellers and brokers to advertise the sale of a home in exchange for a fee. Rigorous checks on ads published by users are costly or even unfeasible. Consequently, before using online listings for economic analysis, it is necessary to assess their reliability.

A key issue is that multiple ads can be associated with the same dwelling. That may be due to various reasons. First of all, under open listing agreements, each broker could publish a different ad. Additionally, a broker could post multiple ads for the same home. In particular, the broker may delete the old ad and create a new one to refresh the time on market of the listing.[8] Furthermore, when a mandate to sell expires, the home seller may sign a listing agreement with a new agent that publishes a new ad.[9]

We are concerned with duplicates for several reasons. First, duplicate ads may provide a biased representation of housing supply, especially at granular levels. Second, the presence of duplicates may not be random but associated with the physical characteristics of the home, the urgency of the owner to sell soon, or difficulties in finding a buyer. Third, the disappearance of a duplicate ad does not necessarily correspond to a sale or a withdrawal. Likewise, new ads do not necessarily correspond to new properties entering the market.

[8] Indeed, many potential buyers search on the website from the most recently published ad to the oldest. Moreover, posting a new ad provides greater visibility to the listing because potential buyers receive notifications about new listings through the email-alert service.

[9] If the old agent does not immediately delete the ad, and the new agent posts a new one, two ads for the same dwelling exist simultaneously. Even if this does not happen, and the two ads are not simultaneously visible on the website, we still need to know that these ads refer to the same dwelling.

We identify duplicate ads using machine learning tools.[10] We depart from the original data set of ads and build a new data set of listings. In the latter, the unit of observation is a home instead of an ad. We use machine learning tools because there is no exact matching between characteristics of the homes reported in two duplicate ads. In our experimentation, using pre-specified heuristic rules (such as, consider apartments whose price difference is smaller than 5 percent) to identify duplicates was not particularly successful. Instead, machine learning algorithms autonomously learn the best criteria that identify duplicates provided the training sample is sufficiently large. Moreover, these algorithms can effectively exploit the partial similarity between dwellings' characteristics, which is crucial because different brokers can provide partially different information about the same feature. The primary input for our algorithm is location. However, other variables play a significant role (e.g., asking price, size, amenities).

After identifying duplicates, we combine them as if they were a single ad. Our final data set includes about 940,000 homes, which we will also refer to as "listings."[11] Tables B.1 and B.2 in Appendix B report descriptive statistics about the sample. In Appendix C, we provide all details about the cleaning procedure.

Once we get rid of duplicates, listing data are much more consistent with official statistics than the original ads (Table 1). The average time-on-market measure on listings data is consistent with the results of the Italian Housing Market Survey and is about two

[10] In general, it is not possible to identify duplicate ads by the address. Both in urban and in rural areas, addresses—as generally reported in the ads—may not uniquely identify homes. For example, for condo apartments in cities, sellers usually report the address of the building, and multiple apartments from the same building could be simultaneously on sale. In rural areas, non-unique addresses are also common. Georeferencing the ads may help in rural settings, where houses are more spread out. However, it is less useful in urban settings with a high concentration of homes.

[11] Duplicates are associated with a small share of listings (about 20 percent). Open listing agreements with many agents seem to be a primary source of duplicate ads. We also observe that the duplicate ads of a property appear over time: new ads are created while old ads are deleted, giving rise to a considerable number of delistings and new listings. Finally, the share of duplicates over total ads increases with city size, and there is significant variability across cities. More details about the distribution of duplicates can be found in Appendix C.2.

Table 1. Sales and Time on Market (months), for Ads, Listings (homes), and Official Data

Year	Sales			Time on Market		
	Ads	Listings	Italian Tax Office	Ads	Listings	Survey
2016	335,181	207,120	178,690	5.1	6.7	7.5
2017	312,584	187,443	186,657	4.9	6.7	6.3
2018	321,840	189,505	197,506	4.4	6.3	6.6

months longer than the average duration of ads.[12] Also, the number of delistings is much lower than the number of removed ads and broadly in line with the number of home sales, once considering that (i) delistings include withdrawals; (ii) not all the homes sold have been listed online. The correlation between the number of delistings and home sales at the city level is 0.96 (Figure 1, panel A). The fit is also excellent for housing prices. The correlation between average asking and sale prices of apartments at the local housing market level is 0.93 (Figure 1, panel B).[13]

3.2 Assessing the Distortions Generated by Duplicates

The presence of duplicates does not introduce significant distortions when estimating the trend of prices and delistings at the country level (Figure 2). However, the measurement error could be more significant at more granular levels.

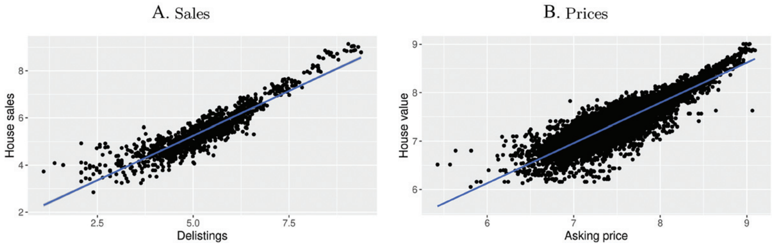
To quantify the measurement error for average asking prices and delistings, we estimate the following ordinary least squares (OLS) regressions:

$$Y_{it} = \alpha + \beta X_{it} + \varepsilon_{it}, \tag{1}$$

[12] We find a significant deviation only for 2016, when listings underestimate time on market. That is plausible because some of the homes listed in 2016 may have been initially listed in 2015. Since we only observe quarterly snapshots for 2015, we may not reconstruct the complete history of dwellings delisted in 2016.

[13] Furthermore, we compute the ratio between the listing prices and actual home values per square meter for each local housing market. On average, during 2016–18, the discount on asking prices was about 12 percent, a value consistent with evidence provided by the Italian Housing Market Survey.

Figure 1. Home Sales, Delistings, and Prices



Note: Official data on home sales and prices are provided by the Italian Tax Office. Home sales and delistings are at city level and data are quarterly. Prices are at the local housing market level and data are half-yearly. All variables are in logs.

Figure 2. Housing Prices and Sales in Italy

Note: Asking prices (panel A) are measured in euros per s.m. Delistings (panel B) were converted in index numbers, where 100 is the average number of estimated delistings (or removed ads) between 2016 and 2018.

where Y_{it} is the value of a statistic computed on the final data set of listings and referring to geographical area i during quarter t. Y can be the average asking price or the number of delistings. For both variables, we consider levels and year-on-year growth rates at a quarterly frequency. X_{it} is the same statistic as calculated on the original data set of ads. The geographical area can be a city or a local housing market because the measurement error can have a different magnitude depending on spatial granularity.

Table 2 reports the results. When considering the levels of asking prices and delistings, the distortion due to duplicate ads seems negligible. There is much heterogeneity between cities and local markets, both in terms of prices and number of delistings. Both data sets can

Table 2. Measurement Error Due to Duplicate Ads

	Levels			Y-o-y Growth Rates		
	$\boldsymbol{\alpha}$	$\boldsymbol{\beta}$	$\boldsymbol{R^2}$	$\boldsymbol{\alpha}$	$\boldsymbol{\beta}$	$\boldsymbol{R^2}$
City Level: Asking Prices	25.374	0.981	0.999	–0.295	0.916	0.917
City Level: Delistings	92.878	0.469	0.984	–3.786	0.863	0.816
Local Market Level: Prices	30.329	0.990	0.994	–0.332	0.817	0.742
Local Market Level: Delistings	2.981	0.513	0.932	–1.660	0.785	0.639

account very well for this spatial heterogeneity, and this explains the almost perfect correlation observed in the left panel of Table 2.

However, the regressions with quarterly year-on-year growth rates show that the measurement error is higher for delistings and is always significantly larger for local housing markets. Figures B.2 and B.3 reveal that the measurement error for asking prices occurs in markets with low number of listings. Delistings are harder to measure because their number is generally low, even in the largest local markets.[14] This prevents the use of the original data for most analyses where granularity and high frequency matter for identification.[15]

Moreover, the presence of multiple ads related to the same dwelling is not random. Indeed, home sellers or brokers post multiple ads to attract more attention. In Appendix D we show that using the original data set of ads implies an oversampling of homes that are relatively expensive and less attractive given their location and characteristics. This implies that, by using ads, we overestimate average listing prices. Moreover, lower attractiveness is associated with higher time on market and propensity to revise the asking price downward. Therefore, using the original data would imply severe distortions when analyzing the microstructure of the housing market (see footnote 23 for a concrete example).

[14]Table B.3 in Appendix B shows that the annual median number of delistings across local markets is 28. As local markets must be homogenous areas, their size is necessarily small.

[15]For example, Anenberg and Kung (2014) assess the impact of foreclosures in small neighborhoods by exploiting the timing of listings' entry and exit into the market. The presence of duplicate listings would seriously impair the representativeness of listing data for similar analyses.

Summing up, the measurement error implied by keeping duplicate listings in the sample is sizable at the granular level, particularly when we look at dynamics. However, it is possible to use the original data set without incurring in significant measurement error in several cases. For example, it is possible to use ads to monitor housing market trends at the country level or for sufficiently large areas. Unfortunately, the presence of duplicates is a substantial disadvantage that prevents the full exploitation of these data.

4. Measuring Demand, Supply, and Prices

This section discusses the potential of online listing data and illustrates their complementarity with traditional statistical sources.

Based on online listings, we can build timely indicators on market inventories (homes on sale), liquidity, and asking prices. By exploiting the richness of details about home characteristics and location, we can detect any diverging pattern across market segments or geographical areas. Yet, similar high-frequency data can be retrieved from some traditional providers, such as MLS or real estate broker associations. We argue that the most significant potential of these data is in the information generated by users as they browse the site, which provides insight into the search activity of potential home buyers, i.e., housing demand. Therefore, compared with traditional sources, online listings allow monitoring both sides of the housing market.

4.1 Demand

Online activity leaves digital traces of human behavior. When individuals visit an ad's webpage, they convey information about the characteristics, location, and price interval of the home they are searching for. By aggregating all this information, we can understand which area households are searching more intensely and what they are looking for. We can observe housing demand.

In our data set, we know how often website users visited the webpage of an ad during each week (clicks). Clicks are complementary to information about online housing demand that has been used in other studies (see, e.g., Piazzesi, Schneider, and Stroebel 2020), namely web searches, i.e., queries where the user specifies the location, characteristics, and price range of the home she is looking for. In principle, web searches and clicks do not convey the same information. People may search for homes with a bundle of characteristics that cannot be found in the market. In this case, we cannot observe clicks that map to those preferences. So, we do not observe the actual preferences of potential buyers. However, it is plausible that potential buyers would somewhat adapt their preferences to the composition of supply. Thus, we think that there is no loss of generality in using clicks instead of web searches for a large class of applications. Moreover, clicks are easier to be used than web searches. Home listing websites usually allow "map search," letting users specify a polygon on the map to look for a home. Extracting and aggregating this type of information about buyers' preferences is hard (Rae 2015; Piazzesi, Schneider, and Stroebel 2020). However, this problem does not arise when considering visits to webpages. Finally, clicks are available for each listing. They can be used to proxy the interest of potential buyers for each home.

To show that online interest is a proxy of housing demand, we proceed as follows. We test whether online interest for a dwelling is correlated with the time it has been on the market and with price revisions. If the webpage of a listed dwelling gets many views, it is plausible that many households are searching for that type of home. Therefore, our first hypothesis is that high online interest is associated with a shorter time on market. Moreover, it is plausible that the price interval is a key searching criterion set by all potential buyers. Suppose many households search in a given price interval for a dwelling with a particular bundle of characteristics, ceteris paribus. Then, it is less plausible to observe downward revisions of the asking price for these dwellings. Therefore, our second hypothesis is that higher online interest implies a lower propensity to revise the asking price.

We build the variable $ONLINT$ to quantify the relative interest in a particular dwelling compared with the other dwellings in the same local housing market.[16] $ONLINT$ is the average daily number of clicks on the home in the first three weeks since its initial listing, divided by the average daily number of clicks in its local

[16]We cannot use the variable $CLICKS$ because homes are listed at different times and for different periods.

housing market during the same period. Thus, when $ONLINT > 1$ it means that the home received more online interest than the average home in the same local housing market, and when $ONLINT < 1$ the reverse is true.

We consider the number of clicks in the first three weeks, as it strikes a compromise between two different problems. If we look at number of clicks over a period that is very long, say two months, online interest may be endogenous. For instance, downward price revisions that occur after a month could likely trigger a change in online interest. By contrast, a period that is too short, say a week, leads to more noise, as we observe ads only once per week.[17] Three weeks is a period that is sufficiently long to mitigate measurement error, while short enough to make it unlikely that price revisions occurred.

We restrict our sample to dwellings that have been initially listed between January 2016 and June 2018 because the observation period for any price revisions or delisting ends in December 2018. We also drop listings with duplicate ads to avoid the bias identified in Appendix D.

To test the relation between time on market and online interest, we estimate the following Weibull regression model:

$$\log(TOM_i) = \beta ONLINT_i + \delta \mathbf{X}_i + \sigma\eta_i, \tag{2}$$

where η_i are i.i.d. random variables following an extreme value distribution. TOM is the time on market—measured as the number of days between the delisting and the first listing.[18] The vector $\mathbf{X}$ includes the physical characteristics of the dwelling. We control for the relative asking price per square meter because relatively more expensive homes are less viewed.[19] We add year-quarter dummies

[17]We observe ads every Monday, but an ad could have been posted on any day of the previous week. Thus, the number of days on which online interest is measured may differ between ads.

[18]Unfortunately, we do not observe if a home has been withdrawn from the market or sold. Then, our variable TOM may be a poor proxy of the time on market. We believe that this is not the case because our measure of the time on market is consistent with survey estimates on average.

[19]The relative asking price is defined similarly to $ONLINT$ and is the ratio of the initial listing price per square meter to the average price in the local housing market during the first three weeks since initial listing.

Table 3. Online Interest

	Dependent Variable		
	***TOM* (AFT) (1)**	***PRICEREV* (LOGIT) (2)**	***PRICE* (OLS) (3)**
ONLINT	–0.069***	–0.080***	
	(0.002)	(0.005)	
$DEMAND_{t-1}$			0.038***
			(0.009)
$AVPRICE_{t-1}$			0.193***
			(0.069)
Log(*Scale*)	–0.098***		
	(0.001)		
Fixed Effects		Local Mkt.	Local Mkt.
Temporal Dummies	Year-Quarter	Year-Quarter	Year-Quarter
Observations	324,906	313,777	427,165
R^2	—	—	0.78

referring to the period of first listing of a home to control for common time-varying unobservables.

In column 1 of Table 3, we report the results. The coefficient associated with *ONLINT* is statistically significant, and its sign confirms our hypothesis. A one-standard-deviation increase in online interest in the early stage of the listing period implies a $e^{-0.069} = 0.93$ times shorter time on market.[20] Notice that the same factor would shrink to 0.70 if online interest were measured over the whole lifetime of the listings. However, in this case, the claim of exogeneity would be hard to support. As we show below, lower online interest implies a greater propensity to revise the asking price downward. Price revision affects time on market (de Wit and van der Klaauw 2013), and likely the online interest of potential buyers.

[20]The results of the Weibull regression can be alternatively interpreted in terms of a proportional hazard model. The hazard ratio associated with a one-standard-deviation increase in online interest is computed as $e^{-\left(\frac{-0.069}{0.907}\right)} = 1.08$, where 0.907 is the scale parameter.

To test whether online interest predicts the occurrence of price revisions, we introduce a binary variable $PRICEREF$. This variable takes value one if the asking price of the dwelling is revised downward and zero if it is not revised or revised upward.[21] Then, we run the following logistic regression:[22]

$$\log\left(\frac{p_{ijt}}{1-p_{ijt}}\right) = \beta ONLINT_{ijt} + \delta \mathbf{X}_{ijt} + \varepsilon_{ijt}, \tag{3}$$

where $p \equiv Prob(PRICEREF = 1)$ and, as in the previous regression, we control for the relative asking price per square meter and the physical characteristics of the dwelling. We also add local housing market and year-quarter fixed effects. We estimate that a one-standard-deviation increase in the relative number of clicks is associated with a 7 percent reduction in the odds of a downward price revision (Table 3, column 2).[23]

Finally, we test if online interest predicts aggregate housing market dynamics. We build an indicator of housing demand in each local housing market. We expect that aggregate online interest is correlated with housing prices. In particular, we hypothesize that stronger demand is associated with higher growth in housing prices, as suggested in the recent literature (Carrillo, de Wit, and Larson 2015; Wu and Brynjolfsson 2015; van Dijk and Francke 2018).

We construct the quarterly variable $DEMAND$, defined as the average daily number of clicks per listing in a local housing market. To deal with the potential endogeneity of this measure of demand to prices, we choose the following econometric strategy. We investigate whether the entry price of a new listing is positively affected by

[21] We consider only the case of downward price revisions for two reasons. First, the number of upward revisions is relatively small. Second, a price increase can be motivated by changing terms of trade or some unobserved change in dwelling quality.

[22] Pangallo and Loberto (2018) show that the relation between prices and online interest also works the other way around. We find that a 1 percent higher price is associated with a 0.66 percent lower number of clicks. We also show that this elasticity has a causal interpretation.

[23] If we did not run the deduplication procedure, running the same logistic regression on the ads data set would yield a 0.7 percent reduction in the odds of a downward price revision instead of a 7 percent reduction. This difference is explained by the fact that brokers are likely to post a new ad when revising the price, as it would attract more attention.

the intensity of search activity in the local housing market in previous months. Suppose online searches are a proxy for actual visits to homes for sale. In that case, real estate agents observe an increase in the market's tightness. Consequently, they may likely suggest higher listing prices to new sellers. We consider the entry prices of new listings; otherwise, average search activity in period $t-1$ would be correlated with prices in period t because of dwellings listed in both periods. This would be problematic, especially in smaller local markets.

We run the following OLS regression:

$$\begin{aligned}\log(P_{ijt}) &= \alpha_j + \zeta_t + \beta_1 \log(DEMAND_{j,t-1}) \\ &\quad + \beta_2 \log(\bar{P}_{j,t-1}) + \delta \mathbf{X}_i + \varepsilon_{ijt},\end{aligned} \tag{4}$$

where P_{ijt} is entry price of new listing i, located in local market j in quarter t. We control for past average asking prices, $\bar{P}$, and dwellings' characteristics. α_j control for local housing market unobservables; ζ_t is a set of year-quarter dummies. The results reported in column 3 of Table 3 confirm that online interest is a good leading indicator of prices. The elasticity of the entry prices of new listings with respect to past average search activity is about 4 percent.

In sum, webpage clicks are a valuable tool for measuring housing demand in real time and understanding buyers' preferences. Moreover, differently from buyers' web searches, clicks are easy to handle. They allow building a measure of demand for a specific home, not only for a neighborhood or a typology of dwellings.[24]

4.2 Supply

Housing supply is usually defined as the total number of dwellings, without considering whether they are on sale or not (Glaeser and Gyourko 2018). Consequently, housing supply increases because new homes are built, and it is downwardly rigid because of the durable nature of dwellings.

In the short or medium run, this definition is not necessarily the most suitable. Indeed, the number of homes potentially available for

[24] In an earlier version of this paper, we showed that the variable $DEMAND$ is a good predictor to forecast the trends of average asking price and liquidity of a local housing market. The results are available upon request.

sale changes over time, at least for two reasons. First, homeowners' decision to move into a new home can depend on macroeconomic developments and housing market conditions (Anenberg and Bayer 2020; Ngai and Sheedy 2020). Second, new homes may enter the housing market because of worsening conditions in the rental market. Owners of vacant homes always have the option to search for either a buyer or a tenant (Krainer 2001; Head, Lloyd-Ellis, and Sun 2014; Liberati and Loberto 2019).

Since the number and type of homes that are for sale may not correlate with the total number of homes, in some cases it is more reasonable to look at listings as a measure of housing supply (see Mian, Sufi, and Trebbi 2015; Piazzesi, Schneider, and Stroebel 2020).[25] For example, Ngai and Sheedy (2020) show that home sales variation is mainly driven by listings instead of a change in matching efficiency in the housing market. Here, we show that the housing supply composition is not time invariant and may change over the housing market cycle.

To show that the average quality of the homes offered for sale changes with the real estate cycle, we consider four variables that measure the average quality of listings in each city at a half-yearly frequency. We define $FLOORAREA$ as the logarithm of the average floor area of listings (measured in square meters); $BATH$ is the share of listings with at least two bathrooms; $GARDEN$ is the share of listings having a private garden; $TERRACE$ is the share of listings having a terrace. To measure the timing of the housing market cycle in each city, we use the logarithm of a hedonic asking price index ($HEDON$). We consider this variable because hedonic price indices are by construction not affected by the physical characteristics of dwellings. Therefore, they are uncorrelated with changes in average home size and quality.[26] Finally, we consider only existing dwellings. In this way, we can show that the home supply composition changes with housing prices and does not depend on the characteristics of newly built houses.

[25]It is fair to say that this distinction is the same that arises in labor market statistics, in which only people that are already working or searching actively for a job are considered inside the labor supply.

[26]Otherwise, an increase in the home average size is associated with a decrease in average asking prices. Indeed, larger homes are ceteris paribus priced at a lower price per square meter.

Table 4. Quality of Listed Dwellings and House Prices (half-yearly data)

	Dependent Variable			
	FLOORAREA **(1)**	***BATH*** **(2)**	***GARDEN*** **(3)**	***TERRACE*** **(4)**
HEDON	0.114**	0.143***	10.491***	8.946**
	(0.050)	(0.031)	(3.001)	(4.096)
Fixed Effects	City	City	City	City
Temporal Dummies	Year-Semester	Year-Semester	Year-Semester	Year-Semester
Observations	546	546	546	546
R^2	0.153	0.057	0.036	0.087

Note: Results of a panel fixed-effect estimation, using the *within* transformation. *HEDON* is the logarithm of a hedonic city-level house asking price index.

We estimate the following model for city i and half-year t:

$$Y_{i,t} = \alpha_i + \zeta_t + \beta HEDON_{i,t} + \varepsilon_{i,t}. \tag{5}$$

The dependent variable Y is one among $FLOORAREA$, $BATH$, $GARDEN$, and $TERRACE$. We add city fixed effects and time dummies. Table 4 reports the results of a panel fixed-effect estimation, using the *within* transformation. Housing supply in cities with stronger housing price dynamics is characterized by a larger average floor area and a higher number of bathrooms of listed homes. We also find an increase in the share of listings with a private garden or a terrace. Results would be qualitatively similar when using the housing prices series estimated by the Italian Tax Office (see Table B.4 in Appendix B).[27] Therefore, housing price increases are associated with a better quality of housing supply.

[27]The limited temporal dimension of our data set prevents a comprehensive analysis of potential non-stationarity in the data. However, we believe that introducing city-level fixed effects eases those concerns. Table B.5 in Appendix B reports consisting evidence.

4.3 Listing Prices

Another potential strength of listing data is the observation of sellers' asking prices. Anenberg and Laufer (2017) show that listing prices can be used to predict a standard housing price index over a short-term horizon. Indeed, listing prices are observed in real time, while sale prices are usually available with a significant lag. However, the determination of the listing price is ultimately a seller's decision possibly made in conjunction with a listing broker. Therefore, it is reasonable to question the ability of listing prices to track sale prices.

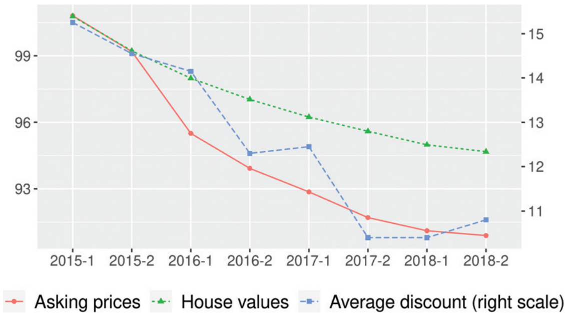
In Italy, the sale price is almost always below the list price. In this case, the asking price dynamics is a good proxy for sale price variations only if the average discount to asking prices obtained by buyers in the bargaining process were stable.[28] Since the outside option of both buyers and sellers in the bargaining process is affected by the general market conditions, the average discount on asking prices changes over time. In Italy, the discount increases during market downturns and decreases during market recoveries. Consequently, a decrease in the discount implies that sale prices decrease less or increase more than asking prices. Potentially, in some periods, sale prices may increase while asking prices decline.

We illustrate this issue in Figure 3. Between the first semester of 2015 and the second semester of 2018, average home values declined by about 6.0 percent, while asking prices diminished by 9.8 percent. That is consistent with the observation that the average discount on asking prices decreased cumulatively by 4.4 percentage points over the same period.

In sum, using asking prices to predict sale prices may require auxiliary variables to improve the fit. For example, Anenberg and Laufer (2017) show that including variables correlated with the discount—such as time on market—improves the forecasting performance of listing prices. Similarly, Lyons (2019) shows that an index based on

[28] We can express the relation between asking prices, P_t^a, and sale prices, P_t, as $P_t^a = (1 - d_t)\,P_t$, where d_t is the average discount. The dynamic relation between asking and sale prices is therefore given by $P_t^a - P_{t-1}^a = P_t - P_{t-1} - (P_t d_t - P_{t-1} d_{t-1})$.

Figure 3. Prices (index 2015S1=100) and Average Discount (percentage points)



list prices that accounts for time on market can track sale prices very well.[29]

5. Conclusion

Big data are becoming ubiquitous in business and academia and increasingly in institutions. There are many reasons for their success. Big data aim to cover the universe of entities under consideration (without the need for sampling). They provide a lot of information that can be integrated by textual analysis and image processing. If coming from online sources, they are frequently available (on a much shorter timescale than administrative data). They rely on observations rather than surveys.

There are disadvantages too. Big data may well fail to provide universal coverage (and so lead to non-representative results). They are less structured and controlled (there might be hidden factors influencing the data-generation process). They could have other sorts of measurement errors.

This study provides a concrete example of the strengths and weaknesses of big data for institutional applications. We analyze a large data set of housing ads published on the leading online portal

[29]Lyons (2019) shows that the spread between the asking and the transaction prices can be decomposed in four components corresponding to distinct market processes that take place between the time of listing and when the transaction takes place.

for real estate services in Italy. We provide a comprehensive analysis of the strengths and weaknesses of these data to study housing markets. The main issue is the existence of a substantial share of duplicate ads, leading to a misrepresentation of the volume and composition of the housing supply. However, once this issue is fixed, the potential of these data is enormous, particularly in analyzing housing demand. For example, using these data Guglielminetti et al. (2021) show how the COVID-19 pandemic has influenced the demand for housing heterogeneously across different market segments in Italy.

Although our analysis is specific to the data set we use, we think our insights could be employed more generally as economists increasingly rely on online listings websites. For example, duplicates are likely to affect all listings websites that have no incentive to control the proliferation of duplicates, e.g., because they profit from the number of ads rather than from data quality. For home listings websites, this problem is exacerbated by open mandate agreements. In all countries where these agreements are possible, duplicate ads could arise from different agencies. We find it unlikely that website administrators could correct this bias. Yet, this paper shows that machine learning techniques can correct this distortion and make online listings a powerful tool for the real-time analysis of housing markets.

Appendix A. Data Sources

Listings. The source data which we obtained from Immobiliare.it are contained in weekly files. Starting from these snapshots, we construct six data sets. The main data set is the one with unique ads. Three data sets track the weekly change of asking prices, visits, and uses of the form to contact the agency that is shown on each ad (we do not use information on contacts in this paper; in Pangallo and Loberto 2018 we show that it provides equivalent information to the number of visits). The last two data sets contain information about real estate agents and the list of hash codes of the pictures associated to each ad (we will not use these data in this paper). The information available for each ad is reported in Table A.1.

Housing Prices. Twice per year, OMI (a branch of the Italian Tax Office) disseminates estimates of minimum and maximum

Table A.1. Information Contained in the Database Provided by Immobiliare.it

Type of Data	Variables
Numerical	Price, floor area, *rooms, bathrooms*
Categorical	Property type, furniture, kitchen type, heating type, *maintenance status, balcony, terrace, floor,* air conditioning, energy class, *basement, utility room*
Related to the Building	*Elevator, type of garden, garage, porter,* building category
Contractual	Foreclosure auction, contract type
Related to the Seller	Publisher type (private citizen or real estate agency), agency name and address
Visual	Hash codes of the pictures, pictures count
Geographical	Longitude, latitude, address
Related to the Ad	Visits, contacts
Temporal	Ad posted, ad removed, ad modified
Textual	Description

Note: For a complete description of the meaning of the variables, see Loberto, Luciani, and Pangallo (2018). Italics indicates that if variables are missing, we perform semantic analysis on the textual description of the ads to recover missing information.

home values in euros per square meter, P_l and P_h, at a very granular level. Home values are available for all OMI microzones—which are uniform socioeconomic areas roughly corresponding to neighborhoods—in Italian cities. P_l and P_h are estimated based on a limited sample of home sales and valuations by real estate experts. Further information is available at https://www.agenziaentrate.gov.it/wps/content/Nsilib/Nsi/Schede/FabbricatiTerreni/omi.

We define the average home value in neighborhood (OMI microzone) j as $\bar{P}_j = \frac{P_{lj}+P_{hj}}{2}$. The average home values at city level are estimated as a simple average of the $\bar{P}_j$. For further aggregation above the city level, we compute weighted averages of the cities' average home values. As weights, we use the stock of homes measured in the 2011 census. OMI estimates are not designed for statistical purposes, and the index we construct must not be considered as equivalent to a quality-adjusted price index.

In Italy, quality-adjusted (hedonic) housing price indices are disseminated by Istat, but their reference area is not consistent with our listing data, apart from three city-level indices that refer to the main Italian cities: Rome, Milan, and Turin.

Home Sales. Quarterly data about the volume of home sales in each city are disseminated by OMI.

Italian Housing Market Survey. The Italian Housing Market Survey is a quarterly survey that has been conducted by Banca d'Italia, OMI, and Tecnoborsa since 2009. It covers a sample of real estate agents and reports their opinions regarding the current and expected course of home sales, price trends, time on market, and terms of trade. See https://www.bancaditalia.it/pubblicazioni/sondaggio-abitazioni/ for further information.

Census Data. We retrieve detailed information on socioeconomic characteristics and stock of buildings in OMI microzones from the 2011 census. Istat census tracts are much smaller than OMI microzones (quantitatively, there are approximately 400,000 Istat census tracts over the Italian territory, as compared with 27,000 OMI microzones) and do not necessarily coincide with them. We perform spatial matching of the polygons representing the tracts and the microzones and impute the Istat variables to the OMI microzones according to the overlap percentage of the polygons. For example, if an Istat census tract comprises 2,000 housing units and it straddles two OMI microzones, such that there is a 50 percent overlap for both, we impute 1,000 housing units to each of the two OMI microzones.

Appendix B. Additional Tables and Figures

Table B.1. Descriptive Statistics: Physical Characteristics and Location

Number of Observations	**936,126**
Surface (sm)	
Minimum	30
25th	70.00
Median	93.00
75th	126.00
Maximum	600
Mean	108.68
Std. Dev.	64.15
Type of Property	
Multi-family Residential Dwelling	847,008
Single-Family Home	89,118
Floor Level	
Ground Floor	122,670
Floor Level: 1–3	521,223
Floor Level: 4–	168,812
Multi-level	70,058
NA	53,363
Rooms	
Number of Rooms: 1	29,417
Number of Rooms: 2	194,115
Number of Rooms: 3	295,953
Number of Rooms: 4	240,358
Number of Rooms: 5 or More	147,063
NA	29,220
Bathrooms	
Number of Bathrooms: 1	548,843
Number of Bathrooms: 2	307,287
Number of Bathrooms: 3 or More	63,116
NA	16,880
Terrace	
Terrace: No	631,324
Terrace: Yes	304,802
Balcony	
Balcony: No	346,474
Balcony: Yes	589,652

(*continued*)

Table B.1. (Continued)

Number of Observations	**936,126**
Maintenance Status	
To Be Renovated	119,236
Good Conditions	349,691
Very Good Conditions	338,544
New-Built	85,188
NA	43,467
Kitchen Type	
Cooking Corner	165,086
Small Kitchen	121,955
Large Kitchen	558,580
NA	90,505
Utility Room	
Utility Room: No	664,806
Utility Room: Yes	271,320
Basement	
Basement: No	585,508
Basement: Yes	350,618
Garage	
No Parking Slot/Private Garage	598,023
Parking Slot	66,348
Private Garage	271,755
Garden	
Without Garden	582,787
Shared Garden	195,773
Private Garden	157,566
Janitor	
Janitor: No	861,184
Janitor: Yes	74,942
Elevator	
Elevator: No	423,983
Elevator: Yes	512,143
Air Conditioning	
Air Conditioning: No	204,779
Air Conditioning: Yes	221,551
NA	509,796
Heating	
Centralized Heating System	282,466
Autonomous Heating System	545,506
NA	108,154

(*continued*)

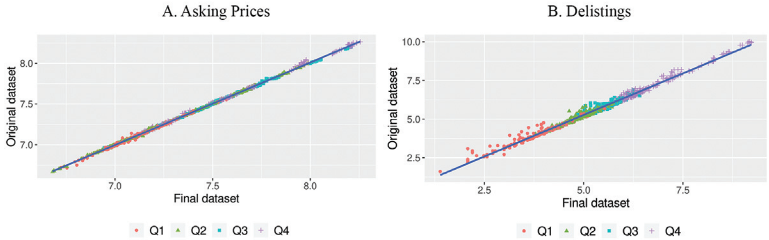
Table B.1. (Continued)

Number of Observations	**936,126**
Energy Efficiency	
Energy Efficiency: High	50,984
Energy Efficiency: Intermediate	108,744
Energy Efficiency: Low	476,659
NA	299,739
NUTS-1	
Northwest (ITC)	301,455
Northeast (ITH)	163,613
Central (ITI)	306,086
South and Insular (ITF-G)	164,972

Figure B.1. Main Trends in the Italian Housing Market

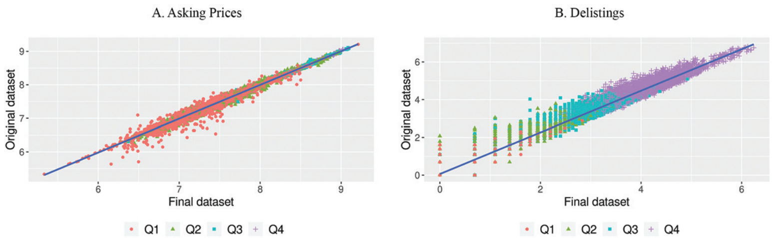
Note: Panel A: home sales, annual data from OMI (a branch of the Italian Tax Office), index 2011 = 100. Panel B: housing prices, annual data from Istat (Institute of Statistics), index 2011 = 100. Panel C: time on market (months) and average discount on the asking price obtained by the buyer (percentage points), quarterly data from the Italian Housing Market Survey. Panel D: housing prices (year-on-year percentage changes), annual data from OMI for 1,174 municipalities with a population of at least 10,000 individuals. This representation shows both a boxplot and raw data (points)—the horizontal position of a point within a year does not carry any meaning and is just needed for graphical representation. In panels A and B we report home sales and prices at country and NUTS-1 level. The other panels report quantities at country level.

Figure B.2. Comparison between the Original and the Final Data Set at City Level



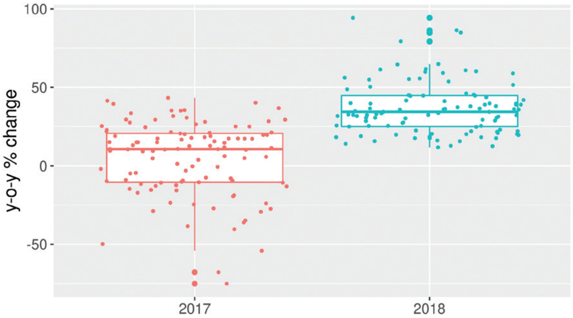
Note: Quarterly data between 2017:Q1 and 2018:Q4. Cities are ranked according to the number of listings in the final data set. The different colors of the dots represent the quartile to which the city belongs (for figures in color, see the online version of the paper at http://www.ijcb.org).

Figure B.3. Comparison between the Original and the Final Data Set at Local Housing Market Level



Note: Quarterly data between 2017:Q1 and 2018:Q4. Local markets are ranked according to the number of listings in the final data set. The different colors of the dots represent the quartile to which the local market belongs (for figures in color, see the online version of the paper at http://www.ijcb.org).

Figure B.4. Online Attention



Note: Average daily number of clicks per ad at city level.

Table B.2. Descriptive Statistics: Asking Prices and Time on Market

	Percentiles					Mean	Std. Dev.
	5	25	50	75	95		
Full Sample							
Price	63,000	122,500	190,000	305,000	690,000	201,531	257,228
Price per s.m.	834	1,389	2,049	3,004	5,000	2,378	1,377
Time on Market	20	68	147	274	680	217	235
Years							
2016							
Price	67,000	128,000	195,762	320,000	728,571	273,148	269,200
Price per s.m.	891	1,471	2,131	3,100	5,122	2,461	1,392
Time on Market	27	90	192	376	872	286	286
2017							
Price	63,549	123,850	190,000	312,077	712,593	266,542	264,460
Price per s.m.	835	1,389	2,045	3,000	5,015	2,379	1,384
Time on Market	25	89	182	336	768	259	262
2018							
Price	60,000	120,000	187,500	305,009	692,000	260,542	260,779
Price per s.m.	797	1,333	1,989	2,951	4,985	2,321	1,378
Time on Market	19	62	145	265	656	212	239
NUTS-1 Regions							
Northwest							
Price	55,000	105,000	169,000	290,540	741,000	254,214	283,708
Price per s.m.	811	1,365	1,988	2,915	5,227	2,354	1,443
Time on Market	20	69	155	293	711	228	242
Northeast							
Price	68,000	115,434	170,000	270,000	550,000	223,849	183,426
Price per s.m.	831	1,271	1,743	2,394	3,800	1,951	959
Time on Market	19	62	140	281	759	226	260
Central							
Price	91,000	166,400	246,000	373,835	800,000	322,894	287,372
Price per s.m.	1,114	1,986	2,773	3,716	5,604	2,989	1,427
Time on Market	20	68	146	272	664	215	232
South and Insular							
Price	53,000	101,732	155,000	240,000	490,000	200,444	175,545
Price per s.m.	708	1,083	1,490	2,067	3,710	1,737	1,006
Time on Market	22	69	140	242	559	193	198

Table B.3. Descriptive Statistics for Local Housing Markets

	Percentiles					Mean	Std. Dev.
	5	**25**	**50**	**75**	**95**		
Population	56.0	1,666.0	4,652.5	10.801.5	49,503.9	8,182.6	10,440.2
Households	25.0	724.0	1,954.5	4,672.5	23,467.4	3,590.4	4,809.3
Housing Units	39.0	933.5	2,499.5	5,577.0	26,856.2	4,235.9	5,392.8
Share of Owner-Occupied (perc.)	30.4	63.0	70.0	75.6	86.7	68.1	11.1
Average Asking Price	697.4	1,282.6	1,717.0	2,380.2	5,817.2	1,987.9	1,041.5
Delistings	0.2	6.8	28.0	83.8	494.9	66.5	109.2
Delistings/Housing Units (perc.)	0.0	0.8	1.4	1.9	4.8	1.5	2.3

Note: Data on the number of residents (populations), households, housing units, and owner-occupied homes are from the 2011 Census. Average asking prices are computed over the period 2016–18. For delistings, we show the average annual number during the period 2016–18.

Table B.4. Quality of Listed Dwellings and House Prices (half-yearly data)

	Dependent Variable			
	FLOORAREA **(1)**	***BATH*** **(2)**	***GARDEN*** **(3)**	***TERRACE*** **(4)**
PRICE	0.073* (0.041)	0.068*** (0.025)	10.986*** (2.525)	5.460 (3.388)
Temporal Dummies	Year-Semester	Year-Semester	Year-Semester	Year-Semester
Observations	534	534	534	534
R^2	0.168	0.043	0.052	0.083

Note: Results of a panel fixed-effect estimation, using the *within* transformation. *PRICE* is the logarithm of the housing prices as estimated by the Italian Tax Office.

Table B.5. Stationarity of Variables on Housing Supply (half-yearly data)

	$\Delta_1 HEDON_t$ (1)	$\Delta_1 PRICE_t$ (2)	$\Delta_1 FLOORAREA_t$ (3)	$\Delta_1 BATH_t$ (4)	$\Delta_1 GARDEN_t$ (5)	$\Delta_1 TERRACE_t$ (6)
$HEDON_{t-1}$	–0.222*** (0.025)					
$PRICE_{t-1}$		–0.543*** (0.045)				
$FLOORAREA_{t-1}$			–0.614*** (0.049)			
$BATH_{t-1}$				–0.775*** (0.049)		
$GARDEN_{t-1}$					–0.688*** (0.053)	
$TERRACE_{t-1}$						–0.592*** (0.052)
Fixed Effects	City	City	City	City	City	City
Observations	455	443	455	455	455	455
R^2	0.180	0.291	0.303	0.411	0.316	0.266

Note: Results of a panel fixed-effects estimation, using the *within* transformation.

Appendix C. Construction of the Housing Units Data Set

Considering the initial data set of ads, during 2016–18 the number of home sales was about 60 percent of the number of delistings (Table C.1), with significant volatility across different cities.[30] Although this statistic is broadly consistent with studies on the U.S. housing market, alternative evidence from the United Kingdom suggests that this estimate is too low.[31] Since our data set is mostly representative of home sales brokered by real estate agents—the largest share of all transactions—the assumption that each ad is associated with a different dwelling would imply that the share of sales over delistings could be well below 60 percent. Moreover, the average time on market computed on listings—as the number of months between the initial listing and the delisting—is about two months lower than the estimates provided by real estate agents in the Italian Housing Market Survey (Table C.1).

Table C.1. Number of Delistings, House Sales, and Time on Market (months)

Year	Delistings	Sales	Time on Market	
			Listings	Survey
2016	335,181	178,690	5.1	7.5
2017	312,584	186,657	4.9	6.3
2018	321,840	197,506	4.4	6.6

Note: Data on sales and time on market come from the Immobiliare.it data set and from the OMI and Italian Housing Market Survey (see Appendix A).

Given these issues, we follow the procedure to clean the original data set described in the next section.

[30] This statistic ranges between 40 percent in Florence and 70 percent in Naples.

[31] According to Anenberg and Laufer (2017) and Carrillo and Williams (2019), about half of the delistings in the United States result in withdraws. In a sample of listings from the United Kingdom, Merlo and Ortalo-Magne (2004) find that withdraws are about 25 percent of the delistings.

C.1 Deduplication at a Glance

We adopt standard methodologies for data deduplication (see Naumann and Herschel 2010; Christen 2012), which we adapt to tackle the specifics of our data set better. The deduplication process consists of three steps.

Data Preparation. To identify if two ads refer to the same dwelling, we have to compare the locations and characteristics of the homes described in the ads. This operation is complicated because the geographical coordinates or the address may not be precise enough. Moreover, some information is not accurate, but based on the best judgment of the home seller/broker.[32]

Thus, we cannot look for perfect matching between home characteristics and have to build partial similarity measures. Moreover, we use the textual description of the home provided in the ad to impute missing data and to extract information useful to identify the duplicates.

Classification. For each pair of ads, we have to decide if the ads refer to the same housing unit. To do so, we compare the characteristics of the dwellings described in the ads and based on some rules, we classify them as *duplicates* or *not duplicates*. To identify these rules, we use a machine learning algorithm, the C5.0 classification tree proposed by Quinlan (1993). The algorithm outputs a probability that the two ads are duplicates. If this probability is larger than 0.5, we consider the two ads as referring to the same housing unit.

Clusterization. The output of the previous step is a list of pairs of duplicate ads. Since multiple pairs can refer to the same dwelling, we have to create clusters of all ads referring to the same home. To do so, we use methods from graph theory and consider a cluster of ads as referring to the same housing unit if an internal similarity condition for the cluster is satisfied. Finally, for each variable, we aggregate information coming from different ads by computing the average or the most common feature observed across ads.

[32]The seller/broker of the home can identify the location on a map or provide the address as an input. The fact that two different tools are available—and the user's lack of precision—gives rise to the possibility that the same dwelling has slightly different geolocation in different ads. That is not an issue in rural areas, but in urban areas with a high concentration of housing units.

Below, we fully describe the algorithm we implemented to remove the duplicate ads. In Loberto, Luciani, and Pangallo (2018) we also show the pseudo-codes of the procedure.

C.1.1 Data Preparation

The textual description of the home provided in the ad performs a dual role. First, by using semantic analysis, information extracted from the textual description allows imputing missing data. That is important because the best way to identify duplicates is to retrieve as much information as possible from the ads. Second, we use the textual description as a further variable to identify if two ads refer to the same dwelling.

There exist standard algorithms in natural language processing that accomplish this task by considering the multiplicity of the words, such as bag-of-words (Harris 1954). However, we cannot use these algorithms here. Indeed, two different real estate agents can describe the same dwelling using different words or sentences, and this makes standard measures of distance among texts useless. For this reason, we resort to the paragraph vector (or *doc2vec*) algorithm proposed by Le and Mikolov (2014), an algorithm based on neural networks that allow representing a document by an N-dimensional vector taking into account both the order and the semantic of the words. In this way, we can measure the "distance" between two descriptions by computing the associated vectors' cosine distance.

We also convert the class of some variables to alleviate the issue of misreporting dwellings' characteristics. Indeed, two different agents can report information partially different but not completely at odds regarding the characteristics of the same housing unit. For example, consider the case of maintenance status. One real estate agent can report that the dwelling must be completely renovated, while the other agent writes that only a partial renovation is necessary. However, it is not plausible that the second agent says that the housing unit is new. As maintenance status takes only four possible ordered categories, we convert the categorical variable to an integer variable that takes value from one to four (a greater value means a better maintenance status). In this way, when we compare two dwellings, we take the absolute difference between the two variables, and we can easily allow for partial matching. We do this operation for several

Table C.2. Variable Transformations for the Deduplication Algorithm (classification tree)

Variable	Original Levels	Transformation
Garage	Missing, Single, Double	Integer: Missing = 0, Single = 1, Double = 2
Garden	Missing, Shared, Private	Integer: Missing = 0, Shared = 1, Private = 2
Maintenance Status	To renovate, Good, Excellent, New	Integer: To renovate = 0, Good = 1, Excellent = 2, New = 3
Kitchen Type	Kitchenette, Small eat-in kitchen, Large eat-in kitchen	Integer: Kitchenette = 0, Small eat-in kitchen = 1, Large eat-in kitchen = 2
Energy Class	A+, A, B, C, D, E, F, G	Integer: A+ = 0, A = 0, B = 1, C = 2, D = 3, E = 4, F = 5, G = 6
Address	Text of the address	Vector of words in the address (removing prepositions and articles)

other ordered categorical variables other than maintenance status: energy class, garage, type of garden, and kitchen type. We report the details in Table C.2.

C.1.2 Classification

We identify duplicate ads based on pairwise comparisons, meaning that we compare each ad with all other ads that are potential duplicates.

First of all, for each ad, we identify its potential duplicates to reduce the computational complexity of the pairwise approach. We define as potential duplicates those ads that refer to dwellings closer than 400 meters to each other and with a difference in asking price lower than 25 percent in absolute value.[33] In this way, we end up

[33]We compute the difference in asking price by dividing the absolute difference between the two asking prices with the lowest of the two. This condition can be quite restrictive when considering dwellings with low asking prices. Then, we

with a long list of pairs of ads, and we have to decide which pairs are duplicates.

We classify each pair of ads as duplicates (TRUE) or distinct housing units (FALSE) based on a supervised classification tree. The algorithm adopted here is the C5.0 classification tree proposed by Quinlan (1993) (http://www.rulequest.com/see5-info.html). This algorithm handles autonomously missing data, is faster than similar algorithms, and allows for boosting.

For each pair of ads, we provide as an input to the algorithm a vector of predictors (covariates in the jargon of machine learning). Based on this information, the classification tree returns the probability that the two ads are duplicates. We consider a pair of ads as duplicates if the estimated probability is higher than 0.5.

Among the predictors, we consider the following variables: floor area, price, floor, energy class, garage, garden type, air conditioning, heating type, maintenance status, kitchen type, number of bathrooms, number of rooms, janitor, utility room, location, elevator, balcony, and terrace. For continuous variables, such as price and floor area, we use both the percentage and the absolute difference; for geolocation, we take the distance in meters between the two dwellings' geographical coordinates. For binary variables, such as elevator or basement, the predictor is a dummy variable that takes value equal to one if both ads share the same characteristic. For discrete ordered multinomial variables (such as maintenance status), we consider different degrees of similarity instead, taking the absolute difference between the two variables.

We also use the distance between the textual description of the two ads as a predictor. For this variable, we consider two different measures, depending on whether the same agency posted the ads. In the first case, we use the Levenshtein distance. Otherwise, we compute the cosine similarities between the vectors produced using the paragraph vector algorithm.

We implement two different C5.0 models, depending on whether the same agency posted the ads. This choice is motivated by the observation that when an agency posts two ads for the same dwelling,

---

consider as potential duplicates also those ads with absolute difference lower than 50,000 euros.

its characteristics are almost equal. On the contrary, when the ads are posted by different agencies (or by a private user), sometimes you can tell they refer to the same dwelling only by the pictures on the website. Then, duplicate ads are less similar if posted by different agencies than if created by the same agency. Consequently, a unique model for both cases could lead to an excess of ads considered as duplicates among those published by the same agency.

C5.0 is a supervised method that requires an initial training sample of pairs of ads of which we know with certainty whether they are duplicates or not. We construct two different training samples, one for each model, by manually checking the ads on the website, comparing the pictures. The training sample for the ads of different agencies is made of 8,296 pairs of ads; among them, 3,711 are duplicates (true positive, TP). The training sample for the ads of the same agency includes 9,844 observations, and 1,850 are duplicates. These samples are constructed by iterating the following steps: (i) estimation of the model based on the initial training sample; (ii) out-of-sample validation of the models; (iii) using the results of the out-of-sample exercise to increase the training sample. We repeat this three-step approach several times until we reach a sufficiently low misclassification error.

To assess the performance of the two models, we randomly split each training sample into two different subsamples. We use the first sample (90 percent of the observations) to estimate the models. The second one (10 percent of the observations) is used for the out-of-sample assessment of the classification performance. We repeat the operation 1,000 times, and we evaluate the performance based on average results. Since the number of true negatives (ads that are not duplicates) is much larger than the number of true positives, using the standard accuracy rate can be misleading about the models' actual performance. For this reason, we consider measures of classification performance that do not rely on the number of true negatives, namely, precision, recall, and F-measure.[34]

[34]The precision rate is the ratio between the number of true positives and the sum of true and false positives. Thus, it measures how accurate a classifier is in classifying true matches. The recall rate is the ratio of true positives over the sum of true positives and false negatives; it measures the proportion of true matches that have been classified correctly. As there is a trade-off between precision and

Table C.3. Assessment of C5.0 Models

	Observations	Duplicates	Precision	Recall	F-measure
Different Agency	8,296	3,711	0.930	0.924	0.927
Same Agency	9,844	1,850	0.952	0.946	0.949

Note: Precision = TP/(TP+FP). Recall = TP/(TP+FN). F-measure = 2*(Precision*Recall)/(Precision+Recall). TP = true positive; FP = false positive; FN = false negative.

We show the results in Table C.3. As expected, the model for ads of the same agency is more precise than the one for ads of different agencies. As we said before, ads posted from the same agency and related to the same dwelling are almost the same. Therefore, it is easier to identify them. However, as the F-measure is equal to .927, also the C5.0 model for ads of different agencies has a quite good classification performance. We should remark that the variables used in the two models are not the same and have been selected to maximize the F-measure.[35] We report the set of variables for each model in Table C.4.

C.1.3 Clusterization

Once we have identified the pairs of ads that are duplicates, we need a procedure to cluster all ads that are considered related to the same housing unit and to aggregate the information in the ads. Here, we follow a standard procedure in the computer science literature (Naumann and Herschel 2010; Christen 2012).

recall, we also consider a third additional measure, the F-measure, that calculates the harmonic mean between precision and recall.

[35]We started for both models with only five predictors: the percentage difference between prices, the absolute difference between prices, the percentage difference between floor areas, the absolute difference between floor areas, and the difference between floors. Then we added each candidate predictor one-by-one, updating the initial model only if the variable provided an improvement of the F-measure (computed on the out-of-sample observations in a Monte Carlo experiment with 1,000 draws). We repeated the operation iteratively as long as there was no performance improvement from adding a new predictor.

Table C.4. Variables for the Classification Trees

Variable	Model 1	Model 2	Description of the Variable
price_abs	Yes	Yes	Absolute difference between asking prices
price_per	Yes	Yes	Percentage difference between asking prices
floorarea_abs	Yes	Yes	Absolute difference between floor area
floorarea_per	Yes	Yes	Percentage difference between floor area
floor	Yes	Yes	Absolute difference between floor level (integer)
distance	Yes	Yes	Absolute distance in meters between households
address	Yes	Yes	Indicator function: 1 if the two addresses have at least one common word
isnew	Yes	Yes	Indicator function: 1 if at least one of the ads refers to a new house
balcony	Yes	No	Indicator function: 1 if the feature balcony is the same
distdays1	Yes	Yes	Number of days between the dates the ads have been added
status	Yes	Yes	Absolute difference (integer) between categories
elevator	Yes	No	Indicator function: 1 if the feature elevator is the same
energy_class	Yes	No	Absolute difference (integer) between categories
isdetached	Yes	No	Indicator function: 1 if at least one of the ads refers to a detached or semi-detached house
bathrooms	Yes	No	Absolute difference between number of bathrooms (integer)
heating_type	Yes	No	Indicator function: 1 if the feature heating type is the same
distcontent1	Yes	No	Cosine distance of vectors (Paragraph vectors) representing textual descriptions
distcontent2	No	Yes	Levenshtein distance between textual descriptions
rooms	Yes	No	Absolute difference between number of rooms (integer)
garage	Yes	Yes	Absolute difference (integer) between categories
garden	Yes	No	Absolute difference (integer) between categories
utility_room	Yes	No	Indicator function: 1 if the feature utility room is the same
janitor	Yes	No	Indicator function: 1 if the feature janitor is the same.

(*continued*)

Table C.4. (Continued)

Variable	Model 1	Model 2	Description of the Variable
basement	Yes	No	Indicator function: 1 if the feature basement is the same
pricemq_abs	No	Yes	Absolute difference in the asking price per square meter
pricemq_min	Yes	Yes	Minimum of the two asking prices per square meter
pricemq_max	No	Yes	Maximum of the two asking prices per square meter

Table C.5. Example of Clusters

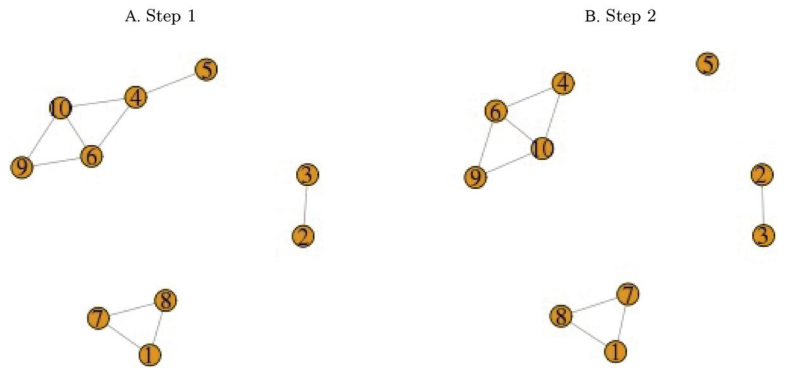
Id.x	1	1	2	4	4	4	6	6	7	9
Id.y	7	8	3	6	10	5	9	10	8	10
Prob.	0.92	0.81	0.73	0.98	1.00	0.52	0.87	0.70	0.93	0.86

Let us suppose that we have only three ads: A, B, and C. It is possible that the pairs (A,B) and (B,C) are considered as duplicates, but (A,C) is not. A simple solution is to assume transitivity: this means that since A is a duplicate of B and B is a duplicate of C, we assume that C is a duplicate of A, and all these ads are considered related to the same dwelling. However, this approach can bring several issues: let us suppose that the probability of being duplicates for the pair (A,B) is 0.95 and the probability for the pair (B,C) is 0.51. The assumption of transitivity in this case may not be reliable.

Here, we abstract from the assumption of transitivity. We decide whether a cluster of ads refers to the same housing unit based on a measure of internal similarity of the cluster. In order to illustrate our approach, we consider a simple example. Assume we have 10 ads. We compute for each of the 45 possible pairs the probability that they are duplicates, and we remove all pairs with a probability smaller than 0.5. The remaining pairs are in Table C.5.

Starting from the results of the pairwise classification step in Table C.5, we represent the information as a graph, in order to form

Figure C.1. Clustering of the Ads



clusters. The output of this step is in Figure C.1. The identifiers of the ads (here assumed to be integers between 1 and 10) are the nodes of the graph. Two nodes are connected if the probability that they are duplicates is greater than 0.5.

The tuples of ads (2,3) and (1,7,8) are considered to refer to two distinct dwellings, as in each tuple ads are all pairwise duplicates. The troubles come with the tuple (4,5,6,9,10). Here, differently than before, it is not true that each ad is a duplicate of all the others. In particular, this sub-graph only has 6 edges, while in order for it to be a fully connected graph, we would need 10 edges. More generally, an indirect graph is said to be fully connected if the number of edges is equal to $\frac{N(N-1)}{2}$, where N is the number of the nodes of the graph (in our case the number of ads).

The tuples (2,3) and (1,7,8) are fully connected, while the tuple (4,5,6,9,10) is not. We consider a cluster as representing a single housing unit if it is a group of ads with a sufficiently high internal similarity, i.e., the number of edges is at least a fraction 5/6 of the maximum number of edges in the cluster. At each step, we verify for each cluster if this condition is verified or not. If it is not satisfied, we remove the weakest edge, which we define as the one with the lowest duplicate probability among those in the cluster.

For the tuple (4,5,6,9,10), the condition is not satisfied. In this case, we delete the weakest link, represented by the edge between nodes 4 and 5 because the associated probability is 0.52. The new set of clusters is in Figure 1B, where node 5 now refers to a distinct housing unit. If we look at the new tuple (4,6,9,10), we see five edges out of six possible edges. Since our internal similarity condition is satisfied, we consider this last tuple as a distinct dwelling.

Summing up our example, we started with 10 ads, and we ended up with only four housing units.

Once we have created the clusters of ads identifying different dwellings, we must combine the information contained in multiple ads related to the same dwelling. As a general rule, for each characteristic, we take the one with the highest absolute frequency. We deviate from this rule in the case of latitude and longitude (we compute the mean across the coordinates of all ads) and when we compute the dates of entry and exit of the dwelling into the housing market (for the entry we take the date of creation of the first ad associated with the dwelling; for the exit we consider the date of removal from the database of the last ad).[36]

C.1.4 Implementation

This approach becomes computationally unfeasible once the number of ads rises. Indeed, the number of pairwise comparisons increases exponentially. Thus, the procedure described in the previous section will be applied using an iterative approach ("time machine approach").

We process the ads progressively as soon as they are published on the website. At the first iteration of the process, we run the deduplication procedure on all ads published before the first week. Once we apply the deduplication procedure, we end up with a new data set. Each row corresponds to a unique dwelling.

At the second iteration, we take as an input the data sets of ads and housing units of the first week. We check for duplicates only among the new ads added during the second week or the ads posted before but for which the price or other characteristics have been updated during the second week. We look for duplicates for all these ads among new or updated ads and the data set of housing

[36] We make a further exception to the general rule for asking prices. In this case, we take the most frequent observation among ads that have not been removed.

units from the first week. The ads that are updated are preliminarily removed from the data set of dwellings (that must be updated accordingly).

Whether the ads are duplicates is still based on the pairwise comparison, but now we can have pairs with two ads or pairs with one ad and one housing unit. Once we compute the probability that they are duplicates for each pair, we cluster the results, as explained in Section C.1.3. We impose the additional condition that in each cluster there can be at most one housing unit already identified in the previous week. This additional condition is necessary to avoid that clusters of ads that have been considered as referring to different dwellings in the past processing can be considered now as duplicates because there are new ads that are potential duplicates of both of them.

C.1.5 Additional Controls

After the deduplication procedure, we make additional controls on the data set to address potential errors. First of all, we keep only the dwellings that have been on the market for at least two weeks. Then, we drop from the data set those dwellings for which the price is not sufficiently consistent with the characteristics of the housing units. In this way, we can also identify foreclosure listings that were not previously identified because the ads did not report the foreclosure status.

Our approach consists of running a hedonic regression, estimating the ratio between actual and predicted price for each dwelling and eliminating the housing units with a ratio between asking and predicted price lower than 0.5 or higher than 1.5.[37]

C.2 Final Data Set and Representativeness

The number of homes—or "true" listings—is only 67 percent of the number of ads (about 940,000 housing units). Looking at the

[37]We keep a relatively broad range because the hedonic regression is limited to a small set of housing unit characteristics, those less affected by missing data issues. In this step, we impute missing characteristics for each housing unit using the approach proposed by Honaker, King, and Blackwell (https://gking.harvard.edu/amelia).

distribution of homes per number of associated ads, we find that duplicates are concentrated over a small share of homes: about 77 percent of dwellings have one associated ad, 13 percent have two duplicate ads, and 10 percent have more than two duplicates.

Open listing agreements with many agents seem to be the main source of duplicate ads. To see that, consider that only 15 percent of homes were listed with more than one agency, but these homes account for 35 percent of ads.

Considering a single daily snapshot, the number of listed homes is 87 percent of the number of ads on average. Thus, by taking a snapshot of the data on any specific day, we expect that only 13 percent of the ads are duplicates. These figures are consistent with those concerning the full sample because duplicate ads for the same listed home grow over time: new ads are created while old ads are deleted, and that gives rise to a huge number of delistings and new listings. We find confirming evidence when we consider only homes with multiple corresponding ads. Every week, for 90 percent of them, at most two duplicate ads are on average visible. This figure can be compared with the share of homes with two ads among those with multiple ads in the full sample, which is $10/(10 + 13) = 57\%$.

Finally, the share of duplicates over total ads increases with city size, and there is significant variability across cities. For example, the ratio between the number of ads and housing units is equal to 1.4 for Naples and 1.8 for Rome. Therefore, an additional implication of duplicates is that they can make the comparability across cities difficult.

To validate the quality of our deduplication procedure, we compare information coming from the final data set with other well-established statistical sources.

First, we compute the number of delistings and home sales in each city (obtained from OMI) at a quarterly frequency, and we find that these two variables are strongly correlated (Figure 1): their correlation coefficient is 0.94. Now, a delisting is an effective exit of a home from the market. Table 1 compares the absolute number of delistings and home sales. Compared with Table C.1, the numbers seem more plausible once we take into account that not all homes sold during these years have been listed on Immobiliare.it.

We find a strong correlation with official data when we consider prices (the correlation is 0.82; Figure 1). Our results are even stronger, because we have official estimates from OMI for each local housing market, so we can compare listing prices and average home values per square meter at a finer granularity. The non-linearity observed for very high home values is probably because OMI estimates refer to the average value of all homes in the local market. In contrast, the most expensive and prestigious homes are likely to be less liquid and, therefore, less represented among listed homes.

Moreover, we compute the ratio between listing prices and actual home values per square meter for each local housing market. On average, we find that the discount on asking prices was about 12 percent during 2016–18, a value consistent with the evidence provided by the Italian Housing Market Survey.

Finally, we look at time on the market. After our deduplication procedure, listings provide an estimate of the time on market overall consistent with the Italian Housing Market Survey (see Table 1). We find a significant deviation only for 2016 when listings underestimate time on market. That is plausible because that is the first year for which we have weekly data. Some of the homes listed in 2016 may have been initially listed in 2015. However, for 2015 we only observe quarterly snapshots, and we may not be able to reconstruct the full history of these listings due to difficulties in identifying duplicates.

Overall, information coming from our final data set of listings is consistent with official statistical sources. We consider this as evidence of the efficacy of our deduplication procedure.

Appendix D. Duplicate Ads and Systematic Bias

The presence of multiple ads related to the same dwelling is not random. In particular, we focus on two hypotheses. First, duplicate ads are more likely among those homes for which potential buyers show little interest, i.e., demand for these homes is relatively small. Intuitively, home sellers would choose to increase search intensity—through open listing agreements with multiple agencies or more generally by posting numerous ads—to compensate for the scarcity of buyers potentially interested in their homes. Second, the presence of duplicates is correlated with the listing price. It is reasonable that homes whose listing price is too high compared with similar nearby homes may have multiple associated ads because the seller increases their odds of finding a buyer.

To test for these hypotheses, we estimate the following linear probability model:[38]

$$DUPL_{ijt} = \alpha_{jt} + \beta CLICKS_{ijt} + \gamma PRICE_{ijt} + \delta \mathbf{X}_i + \varepsilon_{ijt}, \quad (6)$$

where $DUPL_{ijt}$ is an indicator variable equal to one if more than one ad is associated with home i in week t; the index j refers to the home's local housing market. $CLICKS_{ijt}$ is average daily number of visits to the webpages (*clicks*) related to dwelling i during week t.[39] Intuitively, the most-searched homes are likely to be those for which the owner or the broker receives more calls or emails from potential buyers. $PRICE_{ijt}$ is the listing price per square meter of dwelling i during the week t. We control for spatial and temporal heterogeneity at the local housing market level through the set of dummies α_{jt}. Finally, $\mathbf{X}$ is a vector of dwellings' physical characteristics: floor area (square meters), type of property (apartment, detached dwelling), floor level, number of bathrooms, maintenance status, presence of a balcony or a terrace, garage, and elevator.[40]

Since duplicates are identified through machine learning tools, any inefficiency in this first step could invalidate our analysis. However, we believe that this is not an issue for the following two reasons. First, we estimate the classification trees over a large sample of couples of ads for which we know for sure whether they are duplicates. Standard measures of performance for classification tasks used in the machine learning literature suggest that our approach is very effective (see Section C.1.2 and Table C.3). Second, duplicates' identification relies on the similarity between physical characteristics or listing prices and geographical proximity. Since the visits to the webpages and the relative (to the neighborhood) listing prices do not affect the identification of duplicates, any results of our analysis are not a consequence of the deduplication procedure.

---

[38] We use a linear probability model instead of a logit model because of computational convenience.

[39] When multiple ads refer to dwelling i, $CLICKS$ is computed in two steps. First, we compute the average daily number of clicks for each ad. Second, we compute the mean of the daily number of clicks across all ads.

[40] Given the inclusion of time-varying fixed effects and physical characteristics, there is no need to control for the housing price level in the local market to identify overpriced listings. In our context, we only need to estimate if a listing has an asking price higher than those of properties with similar characteristics.

Table D.1. Determinants of Duplicates

	Multiple Ads (1)	**New Duplicate (2)**	**New Duplicate (3)**	**New Duplicate (4)**
Listing Price t	0.0198***			
	(0.0011)			
Clicks t	–0.1221***			0.3003***
	(0.0010)			(0.0006)
Clicks t–1		–0.0015***	–0.0027***	–0.1744***
		(0.0003)	(0.0003)	(0.0007)
Clicks t–2				–0.0463***
				(0.0007)
Clicks t–3				–0.0262***
				(0.0007)
Clicks t–4				–0.0227***
				(0.0006)
Listing Price t–1		0.0024***		0.0112***
		(0.0004)		(0.0005)
Listing Price t–4			0.0013***	
			(0.0002)	
Observations	16,042,720	15,450,398	14,374,903	13,452,978
Adjusted R^2	0.0036	0.0004	0.0004	0.0178

Note: Coefficients and standard errors reported in the table have been multiplied by 100.

Column 1 in Table D.1 reports the results. The coefficients associated with $CLICKS$ and $PRICE$ are statistically significant, and their sign confirms our initial hypotheses. The estimated coefficient for $CLICKS$ is negative (–0.12), and the coefficient for $PRICE$ is positive (0.02). The presence of multiple ads is associated with lower interest by potential buyers and a relatively higher listing price. Although we cannot claim any causal relation based on model (6), the evidence is consistent with the hypothesis that the home seller increases his effort to find a buyer to compensate for a high asking price or unattractive characteristics of the home.

To identify a causal effect of demand and listing prices on the propensity to post multiple ads, we create a new indicator variable called $NEWDUPL$. This variable is equal to one if the number of ads associated with a home already on the market increases during week t. Then, we estimate the following linear probability model:

$$NEWDUPL_{ijt} = \alpha_{jt} + \beta CLICKS_{ijt-1} + \gamma PRICE_{ijt-1} + \delta \mathbf{X}_i + \zeta z_{it} + \varepsilon_{ijt}. \quad (7)$$

Compared with (6), we take as regressors the one-week lag for both demand and listing price. This model allows us to test if the home seller's or the broker's propensity to increase advertising during the week t by posting a new ad is affected by asking price and buyers' demand during the previous week.[41] We also control for the number of days dwelling i has been listed up to week t (z_{it}).

Column 2 in Table D.1 shows that our previous results are qualitatively confirmed. The propensity to post a new ad for a previously listed home decreases when online interest for that home goes up; this propensity is also increasing in the listing price. Notice that these coefficients are statistically significant, although we include many controls, and the phenomenon we are considering is not very frequent at a weekly frequency. In particular, the unconditional probability that during week t a new ad is posted for a previously listed home is 0.9 percent.

In this regression, clicks can be considered as exogenous because potential buyers cannot know the sellers' strategies a week before. Moreover, since we control for the listing price and dwellings characteristics, we deduce that the lower online attention is determined not only by an excessively high price asked by the seller but also by a genuine mismatch with potential buyers' preferences. Unfortunately, we cannot resort to this argument to claim that the lagged value of the listing price is exogenous.[42] However, in column 3, we show that replacing the one-week lagged listing price with the four-weeks lag, we still find a positive and significant effect on the propensity to post a new ad.

Finally, after showing that the listings that receive little online attention are those with the highest probability of having multiple

[41] Controlling for higher-order lags (up to $t-4$) would not affect our results.

[42] Indeed, home sellers/brokers set both the listing price and the advertising strategy, and when changing the listing price, they may have already decided to post a new ad.

ads, we want to evaluate the effectiveness of this advertising strategy. To do that, we estimate the following extension of model (7):

$$NEWDUPL_{ijt} = \alpha_{jt} + \sum_{i=0}^{4} \beta_i CLICKS_{ijt-i} + \gamma PRICE_{ijt-1} + \delta \mathbf{X}_i + \zeta z_{it} + \varepsilon_{ijt}, \tag{8}$$

where we add as regressors the contemporaneous value of the variable $CLICKS$ and all its lags up to four weeks. The results are reported in column 4 in Table D.1. We find that during the four weeks before the seller posts a new ad, his home gets a relatively poor online interest ($\beta_i < 0$ for $i = -1, -2, -3, -4$). Clicks are low especially in the previous week (β_{-1}). Then, following the publication of the new ad, a spike in clicks occurs ($\beta_0 > 0$). These results, which must be interpreted as correlations, are consistent with the hypothesis that potential buyers may believe that this is a new listing. Homebuyers may not easily recognize that the new ad refers to a previously listed home, and this is especially true when a new broker posts the ad.

Finally, homes with multiple ads show further systematic differences compared with other dwellings. We estimate the OLS regression of time on market over a dummy taking value one if a home had multiple ads, and we find that those with many ads stay longer on the market (see Table D.2). We also estimated a linear probability model where the dependent variable is an indicator variable taking value one if the home seller revised downward the initial asking price, and zero otherwise. As expected, it is more plausible to observe a price change for homes with multiple ads (Table D.2). These results are consistent with previous evidence: by using the ads, we underestimate the time on the market, and dwellings with multiple ads are overpriced (therefore more subject to price reductions).

The main conclusion is that using the original data set of ads implies an oversampling of relatively expensive homes—given their location and characteristics—and less attractive homes. Moreover, lower attractiveness is associated with higher time on market and propensity to revise downward the asking price. Therefore, using the original data would imply severe distortions when analyzing the microstructure of the housing market.

Table D.2. Duplicates, Time on Market, and Price Changes

	Time on Market (1)	Price Change (2)
Multiple Ads	125.30580***	0.17805***
	(0.74404)	(0.00150)
Fixed Effects	OMI Microzone	OMI Microzone
Temporal Effects	Quarter	Quarter
Observations	512,246	512,246
Adjusted R^2	0.06827	0.09316

References

Anenberg, E., and P. Bayer. 2020. "Endogenous Sources of Volatility in Housing Markets: The Joint Buyer–Seller Problem." *International Economic Review* 61 (3): 1195–1228.

Anenberg, E., and E. Kung. 2014. "Estimates of the Size and Source of Price Declines Due to Nearby Foreclosures." *American Economic Review* 104 (8): 2527–51.

Anenberg, E., and S. Laufer. 2017. "A More Timely House Price Index." *Review of Economics and Statistics* 99 (4): 722–34.

Carrillo, P. E., E. R. de Wit, and W. Larson. 2015. "Can Tightness in the Housing Market Help Predict Subsequent Home Price Appreciation? Evidence from the United States and the Netherlands." *Real Estate Economics* 43 (3): 609–51.

Carrillo, P. E., and B. Williams. 2019. "The Repeat Time-on-the-Market Index." *Journal of Urban Economics* 112 (July): 33–49.

Christen, P. 2012. *Data Matching: Concepts and Techniques for Record Linkage, Entity Resolution, and Duplicate Detection.* Springer Publishing Company, Inc.

de Wit, E. R, and B. van der Klaauw. 2013. "Asymmetric Information and List-Price Reductions in the Housing Market." *Regional Science and Urban Economics* 43 (3): 507–20.

Glaeser, E., and J. Gyourko. 2018. "The Economic Implications of Housing Supply." *Journal of Economic Perspectives* 32 (1): 3–30.

Guglielminetti, E., M. Loberto, G. Zevi, and R. Zizza. 2021. "Living on My Own: The Impact of Covid-19 Pandemic on Housing Preferences." Questioni di Economia e Finanza (Occasional Papers) No. 627, Bank of Italy.

Han, L., and W. C. Strange. 2015. "The Microstructure of Housing Markets: Search, Bargaining, and Brokerage." In *Handbook of Regional and Urban Economics,* Vol. 5, ed. G. Duranton, J. V. Henderson, and W. C. Strange, 813–86 (chapter 13). North-Holland.

Harris, Z. S. 1954. "Distributional Structure." *Word* 10 (2–3): 146–62.

Head, A., H. Lloyd-Ellis, and H. Sun. 2014. "Search, Liquidity, and the Dynamics of House Prices and Construction." *American Economic Review* 104 (4): 1172–1210.

Jordà, O., M. Schularick, and A. M. Taylor. 2015. "Leveraged Bubbles." *Journal of Monetary Economics* 76 (Supplement): S1–S20.

Kolbe, J., R. Schulz, M. Wersing, and A. Werwatz. 2021. "Real Estate Listings and Their Usefulness for Hedonic Regressions." *Empirical Economics* 61 (6): 3239–69.

Krainer, J. 2001. "A Theory of Liquidity in Residential Real Estate Markets." *Journal of Urban Economics* 49 (1): 32–53.

Landvoigt, T., M. Piazzesi, and M. Schneider. 2015. "The Housing Market(s) of San Diego." *American Economic Review* 105 (4): 1371–1407.

Le, Q., and T. Mikolov. 2014. "Distributed Representations of Sentences and Documents." In *Proceedings of the 31st International Conference on Machine Learning (ICML-14)*, Vol. 32, ed. E. P. Xing and T. Jebara, 1188–96.

Liberati, D., and M. Loberto. 2019. "Taxation and Housing Markets with Search Frictions." *Journal of Housing Economics* 46 (December): Article 101632.

Loberto, M., A. Luciani, and M. Pangallo. 2018. "The Potential of Big Housing Data: An Application to the Italian Real-Estate Market." Temi di discussione (Economic Working Papers), No. 1171, Bank of Italy.

Lyons, R. C. 2019. "Can List Prices Accurately Capture Housing Price Trends? Insights from Extreme Markets Conditions." *Finance Research Letters* 30 (September): 228–32.

Merlo, A., and F. Ortalo-Magne. 2004. “Bargaining over Residential Real Estate: Evidence from England.” *Journal of Urban Economics* 56 (2): 192–216.

Mian, A., K. Rao, and A. Sufi. 2013. “Household Balance Sheets, Consumption, and the Economic Slump.” *Quarterly Journal of Economics* 128 (4): 1687–1726.

Mian, A., A. Sufi, and F. Trebbi. 2015. “Foreclosures, House Prices, and the Real Economy.” *Journal of Finance* 70 (6): 2587–2634.

Naumann, F., and M. Herschel. 2010. *An Introduction to Duplicate Detection*. Morgan and Claypool Publishers.

Ngai, L. R., and K. D. Sheedy. 2020. “The Decision to Move House and Aggregate Housing-Market Dynamics.” *Journal of the European Economic Association* 18 (5): 2487–2531.

Pangallo, M., and M. Loberto. 2018. “Home Is Where the Ad Is: Online Interest Proxies Housing Demand.” *EPJ Data Science* 7 (1): Article 47.

Piazzesi, M., M. Schneider, and J. Stroebel. 2020. “Segmented Housing Search.” *American Economic Review* 110 (3): 720–59.

Quinlan, J. R. 1993. *C4.5: Programs for Machine Learning*. San Francisco, CA: Morgan Kaufmann Publishers Inc.

Rae, A. 2015. “Online Housing Search and the Geography of Submarkets.” *Housing Studies* 30 (3): 453–72.

van Dijk, D. W., and M. K. Francke. 2018. “Internet Search Behavior, Liquidity and Prices in the Housing Market.” *Real Estate Economics* 46 (2): 368–403.

Wu, L., and E. Brynjolfsson. 2015. “The Future of Prediction: How Google Searches Foreshadow Housing Prices and Sales.” In *Economic Analysis of the Digital Economy*, ed. A. Goldfarb, S. E. Greenstein, and C. E. Tucker, 89–118 (chapter 3). Chicago: University of Chicago Press.