Unbundling Quantitative Easing: Taking a Cue from Treasury Auctions*

Walker Ray
London School of Economics

Michael Droste
Harvard University

Yuriy Gorodnichenko
UC Berkeley and NBER

August 2, 2023

Abstract

We study empirically and theoretically the role of preferred habitat in understanding the economic effects of the Federal Reserve's quantitative easing (QE) programs. Using high-frequency identification and exploiting the structure of the primary market for U.S. Treasuries, we isolate demand shocks that are transmitted solely through preferred habitat channels, but otherwise mimic QE shocks. We document large "localized" yield curve effects when financial markets are disrupted. Our calibrated model, which embeds a preferred habitat model in a New Keynesian framework, can largely account for the observed financial effects of QE. We find that QE is modestly stimulative for output and inflation, but alternative policy designs can generate stronger effects.

Keywords: quantitative easing, monetary policy, market segmentation, treasury auctions

JEL Classification: E52, E43, E44

*We thank Harald Uhlig, anonymous referees, Michael Bauer, Rhys Bidder, Michael Fleming, Ed Knotek, Matteo Maggiori, Eric Swanson, Michael Weber, Milena Wittwer, and seminar participants at Berkeley, New York Fed, Cleveland Fed, San Francisco Fed, Chicago Fed, ASU, DEC, CMC, Bilkent, GWU, Texas A&M, and Northwestern for comments on an earlier version of the paper. We are grateful to Maxime Sauzet for excellent research assistance.

1 Introduction

While evaluating the first rounds of quantitative easing (QE), then-Fed chair Ben Bernanke observed, "The problem with QE is it works in practice but it doesn't work in theory." Indeed, QE was successful in reducing short- and long-term interest rates, but the mechanisms behind these effects are still not well-understood. Nevertheless, this theoretical ambiguity has not stopped the Fed from continuing to utilize QE programs, both during the onset of COVID-19, and then reversing course by implementing quantitative tightening (QT) in response to recent inflationary pressures.

Although workhorse macroeconomic models imply that Treasury demand is determined solely by economic agents' intertemporal consumption decisions, several explanations have been put forth to rationalize the workings of QE. For instance, QE could have signaled to financial markets a commitment to keep short-term interest rates low for a long time (forward guidance). Or, perhaps the Fed exploited financial market frictions (limited arbitrage and market segmentation) by purchasing securities in a particular segment. Alternatively, large-scale asset purchases by the Fed could signal a poor state of the economy, pushing interest rates down ("Delphic effect").[1] Given the paucity of QE events, it has been difficult to provide clear empirical evidence for (and assess the relative contributions of) the proposed channels. Moreover, QE policies have been implemented as responses to severe macro-financial conditions, which further confounds identification and interpretation and raises questions about the effectiveness of QT going forward.

The objective of this paper is to unbundle the effects of QE by focusing on a specific channel: market segmentation in the form of *preferred habitat*, which posits that certain investors have preferences for specific maturities. To this end, we take the following approach. First, we identify shifts in *private* demand for Treasuries that mimic QE, but are independent of all other plausible channels of QE. Next, we analyze the propagation of these demand shocks across financial markets. In particular, we assess the ability of preferred habitat theory to rationalize the observed responses, and we confirm key predictions regarding pass-through of demand shocks in and out of financial crises. Informed by our empirical analysis, we then develop a general equilibrium macroeconomic model designed to study QE policies. We find that the preferred habitat channel accounts for the bulk of the observed response to QE in financial markets. Our model suggests that the first rounds of QE had modest stimulative effects on output and inflation, but given the relative health of financial markets in the present period, QT alone is unlikely to be successful in reducing inflation to target. Finally, we explore alternative QE implementations that can help improve the design of future asset purchases.

Our analysis starts with the key insight that the mechanism through which market segmentation and preferred habitat forces operate is not the source of Treasury quantity

[1] See, Campbell et al. (2012), Martin and Milas (2012) and Bhattarai and Neely (2020) for surveys.

shocks *per se*, but rather how marginal investors absorb these shocks. We utilize the primary market for Treasuries to identify demand shifts that are independent of all QE propagation mechanisms besides preferred habitat. Although the primary market is the venue through which the Treasury issues debt (a supply-side action), the institutional structure of Treasury auctions has a number of desirable features for identifying demand-side shocks. Because all of the supply information is announced by the Treasury in advance of each auction, the release of the auction results reveal unexpected shifts in demand alone, allowing us to rule out a host of confounding factors. By utilizing intraday changes in Treasury yields around the close of Treasury auctions, we construct a novel measure of Treasury demand shocks.

We document that demand shocks are reasonably large and persistent, with effects on yields typically lasting for many weeks following the auction. Furthermore, the surprise movements in demand are driven by institutional investors such as foreign monetary authorities, investment funds, insurance companies and the like. We show that these shocks are not driven by market-wide changes in expectations about inflation, output, or other general macro-financial conditions. Therefore, variation in Treasury yields around the release of Treasury auction results can help us to isolate the effect of *idiosyncratic* purchases in specific asset segments on the level and shape of the yield curve, which is difficult to achieve by examining only QE events. Because Treasury auctions are frequent and information spans many decades, we can study state dependence in the effect of targeted purchases of assets (e.g., crisis vs. non-crisis states), which is instrumental for understanding how QE programs can work in normal times. Importantly, because Treasury auctions for specific maturities are spread over time, we can identify changes in demand for government debt of specific maturities and trace the effect of these changes on other parts of the yield curve. In this sense, we have natural experiments which can mimic targeted purchases of the Fed during QE programs.

We use our auction demand shocks to empirically evaluate the "localization hypothesis," a characteristic prediction of preferred habitat theory and a key input into our subsequent quantitative analyses. Using a simple regression specification informed by theory, we document strong evidence in favor of localized yield curve effects during financial crisis periods (i.e., the location of the demand shock in maturity space matters and the effects on the yield curve are larger for bonds of similar maturities). These results imply potentially powerful effects of QE on targeted yields in crises. However, we cannot reject the null of no localization effects during normal times, which suggests a limited use of QE as a conventional policy tool.

Building on seminal work by Vayanos and Vila (2021), we develop a general equilibrium model with risk-free and risky debt to rationalize the empirical responses of the yield curve to shocks in demand for Treasury securities and to better understand the effects of QE on financial markets and the broader macroeconomy. We calibrate the model to match

a variety of moments for yields and macroeconomic variables as well as the responses of yields to surprise movements in *private* demand during Treasury auctions in "crisis" and "non-crisis" times. When fed a shock mimicking QE1 in size and duration, our model generates movements in the yield curve remarkably close to those observed in the data. This is consistent with the view that QE works mainly via market segmentation and preferred habitat, and that the *net* effect of other channels is small. Given the disruption in financial markets, our calibration implies that QE1 stimulated output and inflation by as much as a 50-75 basis point rate cut.

Policy experiments in our model indicate that these relatively modest macroeconomic effects are sensitive to implementation details. In particular, holding securities on the balance sheet longer (and making this clear to markets on announcement) boosts the stimulative power of QE significantly. The expansionary effects of QE fall precipitously when undertaken during periods when bond markets are relatively healthy. We also show that QE may have unintended consequences: uncertainty surrounding the Fed's asset purchases may lead to excess macroeconomic volatility, thus calling for policy guidance. Finally, QE programs tilted towards risky assets (e.g., mortgage-backed securities) are more effective in stimulating the economy, particularly when these assets are more volatile. Our results indicate that QE should remain in policymakers' toolkit, but must be utilized with caution and realistic expectations. For example, while our model predicts that the Fed's ongoing QT program is disinflationary, the effects are similar to a 25-50 basis point rate hike, and thus dwarfed by the current rise in inflation.

Our paper makes three primary contributions. Theoretically, we embed a financial model of the entire term structure of interest rates for risky and safe assets within a dynamic model of the macroeconomy and thus can provide an integrated analysis of QE.[2] This is important because previous studies of QE largely focus separately on either financial variables or macroeconomic variables. For example, Krishnamurthy and Vissing-Jorgensen (2012) and Chodorow-Reich (2014) study the effects of QE announcements on financial markets, but do not quantify the macroeconomic effects of QE. Vayanos and Vila (2021), Greenwood and Vayanos (2014), Greenwood et al. (2016), King (2019a) and related work explore the financial market implications of preferred habitat theory, but are silent on any potential effects on output or inflation. On the other hand, recent developments in macroeconomic theory (e.g., Gertler and Karadi (2011), Cúrdia and Woodford (2011), Chen et al. (2012), Gertler and Karadi (2013), Sims and Wu (2020), Karadi and Nakov (2020), Carlstrom et al. (2017), Ippolito et al. (2018)) concentrate on aggregate variables, but are unable to capture the rich dynamics in bond markets which we document. Moreover, many of these theories rely on reserve requirements or moral hazard/enforcement constraints on banks as the key channel through which QE works. In contrast, we focus on the interaction of preferred habitat with limited risk-bearing

[2] Ray (2019) develops additional analytical and normative results in a version of this model.

capacity in bond markets, which we argue is a core mechanism behind QE effects. Our quantitative model merges these literatures: we utilize financial data and high-frequency identification to discipline our model, which we then use to quantify the macroeconomic effects of QE.

Empirically, we develop a novel, high-frequency measure of demand shocks for Treasuries by exploiting the institutional structure of Treasury auctions. A well-identified shock series is crucial for testing various macro-finance theories, and we hope our measure will be analyzed and extended in future work in a fashion similar to the high-frequency monetary shocks pioneered by Kuttner (2001), Bernanke and Kuttner (2005), Gürkaynak et al. (2007), and others. Our approach is a natural complement to existing empirical approaches estimating affine term structure models (Hamilton and Wu (2012), Kaminska and Zinna (2020)) or demand systems (Koijen et al. (2021)) using lower-frequency data.

Combining the insights of our model with our high-frequency demand shocks (and building on extensive research studying how QE purchases impacted the yield curve, such as D'Amico and King (2013), Li and Wei (2013), Cahill et al. (2013), King (2019b)), we provide strong empirical evidence for *state-dependent* (i.e., "crisis" vs. "non-crisis") localization effects in the spillovers across maturities and asset classes of surprise movements in Treasury demand. Put differently, we use QE-like events (rather than QE directly, in the spirit of Fieldhouse et al. (2018) and Di Maggio et al. (2020)) in order to exploit rich cross-sectional and time-series variation to obtain sharp identification and precise estimates of demand shock spillovers. This new finding confirms one of the characteristic predictions of our model (and preferred habitat theory more generally), and thus our paper shows that these mechanisms are crucial in understanding the Treasury market and channels through which QE affects yields.

Our paper is related to a broad literature studying Treasury auctions, investigating how yields move around Treasury auctions (e.g., Lou et al. (2013), Fleming and Liu (2016)) and respond to variation in demand (e.g., Cammack (1991), Beetsma et al. (2016, 2018), Forest (2018)). However, our focus is not Treasury auctions in and of themselves; rather, we exploit the institutional structure of the primary market for Treasuries to better understand the mechanisms behind QE. Relative to this early important work, we *structurally* link demand shocks identified from Treasury auctions to a general equilibrium preferred habitat model.

2 Data and Institutional Details

In this section we describe the primary sources of our data and present basic statistics. First, we describe the U.S. Treasury auctions for U.S. government notes and bonds (coupon-bearing nominal securities). Second, we describe the details of the data regarding intraday secondary-market Treasury prices.

2.1 Primary Market for Treasury Securities

The Treasury sells newly issued securities to the public on a regular basis through auctions. In recent years, 2-, 3-, 5- and 7-year notes are auctioned monthly. 10-year notes and 30-year bonds are auctioned in February, May, August and November with reopenings in the other 8 months. The frequency of auctions has changed over time. For example, 30-year bonds were not issued between 1999 and 2006 and were issued only twice a year between 1993 and 1999; and 20-year bonds were auctioned in May 2020, the first time since 1986.[3]

There are two types of bids: noncompetitive and competitive. Noncompetitive bidders agree to accept the terms settled at the auction, and are typically limited to $5 million per bidder. Competitive bidders submit the amount they would like to purchase, not exceeding 35% of the amount offered at auction, and the price (the interest rate) at which they would like to make the purchase.

Auction participants include primary dealers, other non-primary brokers and dealers, investment funds (e.g., pension, hedge, mutual), insurance companies, depository institutions, foreign and international entities (governmental and private), the Federal Reserve System Open Market Account (SOMA), and individuals. These participants are classified into three groups: primary dealers, direct bidders, and indirect bidders. Primary dealers (brokers and banks) trade on their account with the Federal Reserve Bank of New York; they are required to participate in every Treasury auction, and typically buy the largest share of auctioned debt. Direct bidders are non-primary dealers who submit bids for their own proprietary accounts. Indirect bidders submit competitive bids via a direct submitter, including foreign and international monetary authorities placing bids through the Federal Reserve Bank of New York. Additionally, the Treasury divides investors into the following classes: Investment Funds; Pension and Retirement Funds and Insurance Companies; Depository Institutions; Individuals; Dealers and Brokers; Foreign and International; Federal Reserve System; Other.[4]

Figure 1 depicts the stages of a Treasury auction.[5] First, the Treasury releases all pertinent information regarding an upcoming auction a few days prior to the auction date. An announcement includes the amount offered, additional security information (e.g., maturity, CUSIP, coupon schedule), and other information describing the rules of the auction. After the announcement, investors may submit bids up until the auction

[3] In our analysis, we exclude inflation-protected securities (TIPS) and floating rate notes because these securities have different structural arrangements than simple coupon-bearing nominal securities. We also exclude Treasury bills (zero-coupon securities with maturity one year or less) because the QE programs mainly bought long-maturity nominal U.S. government debt.

[4] Data for announcement and results of each auction since late 1979 are available from TreasuryDirect.gov. Data regarding amounts accepted and tendered by bidder type (primary dealer, direct, and indirect) are available starting in 2003. The Treasury provides information regarding allotment by investor class starting in 2000 (see Fleming (2007) for a breakdown by types and class of bidders in greater detail). Appendix F provides detailed information on all data sources.

[5] Appendix Figure B2 presents a typical announcement of an auction and its results.

closing time. For typical note and bond auctions, non-competitive bids may be submitted by 12:00 p.m., while the deadline for competitive bids is 1:00 p.m.

After the auction closes, competitive bids are accepted in ascending order (in terms of yields) until the quantity meets the amount offered minus the amount of non-competitive bids. Winning bidders receive the same yield as the highest accepted bid. Once the winning bids are determined, auction results are released immediately; beginning in the early 2000s, results are released within minutes of the close of the auction (see Garbade and Ingber (2005)). Besides the winning yield, the Treasury announces various aggregate statistics regarding the bidding, such as the total demand (bids tendered) for the security and the composition of bids and winners by investor and bid type. One particularly salient piece of information is the "bid-to-cover," the ratio of all bids received to all bids accepted. A few days after the close of an auction, the Treasury delivers the securities and charges the winning bidders for payment of the security.

Panel A of Table 1 presents summary statistics for note and bond auctions from 1995-2017, the period for which we have intraday Treasury yields (Appendix Figure B1 plots the number and size of note and bond auctions split by maturity). Since 1995, a typical offering of $20 billion generates more than $50 billion in demand. Primary dealers account for the largest source of demand (bid-to-cover ratio ≈2), but other types of bidders also account for a large fraction of auction offerings. Primary dealers purchase ≈60 percent of auctioned Treasuries, with the rest split between investment funds and foreign buyers. There is considerable variation in the offered amounts (standard deviation ≈$9 billion) as well as the level and composition of demand (standard deviation for the bid-to-cover ratio ≈0.5, and the standard deviation of bid-to-cover ratio for primary dealers is ≈0.35).

2.2 Intraday Treasury Yields

Once a Treasury security auction is complete, the security is issued to the winning bidders and the security is free to trade in the over-the-counter (OTC) secondary market. Following the announcement of an auction but before issuance, there is a forward market for newly auctioned Treasuries. The forward contracts mature on the same day as the securities are issued, and hence this market is referred to as the "when-issued" market. Our data on secondary market yields (including when-issued yields) comes from GovPX, which provides comprehensive intraday coverage of all outstanding U.S. Treasuries for the period 1995-2017. We use changes in intraday Treasury yields in order to construct market-based measures of demand surprises occurring during Treasury auctions.[6]

[6] An earlier draft of this paper used Treasury futures to construct auction demand shocks. Treasury futures provide a natural market-based measure of such shifts, but in practice movements in the secondary market around auctions are not predictable. In addition, Treasury futures markets are less liquid and cannot be not tied to a specific CUSIP-level bond. Thus, we construct demand shocks using secondary market Treasury yields; but our results are robust to using futures (see Gorodnichenko and Ray (2017)).

3 Quantifying Demand Shocks

This section describes our procedure to measure surprise movements in Treasury yields around Treasury auctions and documents properties of these surprises. Our key assumption is that in a small window around the release of Treasury auction results, shifts in Treasury yields reflect unexpected changes in market beliefs about the demand for Treasuries with a specific maturity. Indeed, the Treasury announces an offered amount well before an auction happens, thus fixing supply in advance of investor bidding. Hence, between the announcement and close of the auction, Treasury yields should move only in response to unexpected changes in demand conditions. Our high-frequency approach isolates variation only due to unexpected shifts in demand arising from a specific auction.

3.1 Shock Construction

Let $y_{t,pre}^{(m)}, y_{t,post}^{(m)}$ be the Treasury yields before and after the close of the auction on date t with maturity m. We measure the surprise movements in Treasury yields as:

$$D_t^{(m)} = y_{t,post}^{(m)} - y_{t,pre}^{(m)}. \tag{1}$$

For all auctions, $y_{t,pre}^{(m)}$ is the last yield observed 10 minutes before the close of the auction, while $y_{t,post}^{(m)}$ is the first yield observed 10 minutes following the release of the auction results. If the date t auction is a re-opening of a previously issued security, we use secondary market yields to construct our shocks. If the date t auction is instead a newly issued security, we use yields from the "when-issued" market. In our sample, auctions typically close at 1:00PM, or less frequently at 11:30AM. However, the time between the close of the auction and the release of the results is a function of how long it takes the Treasury to compile the results. The Treasury began releasing results within minutes in the early 2000s, but in the 1990s frequently took longer. We collect wire reports from Bloomberg, which gives a tight upper bound on the release time.[7]

Summary statistics of our constructed shock measures $D_t^{(m)}$ are presented in Panel B of Table 1 (time series are plotted in Appendix Figure B3). The shock means are close to zero (and statistical tests do not reject the null of zero means). Moreover, there is essentially no serial correlation in $D_t^{(m)}$ ($\rho = -0.03$). Hence, these summary statistics indicate that surprises are not systematic and do not contain predictable movements. In our sample period, the standard deviation of $D_t^{(m)}$ increases in maturity m and ranges from 1.4 basis points for 2-year maturity to 3.3 basis points for 30-year maturity. For comparison, Chodorow-Reich (2014) reports that the largest intraday movements in yields following a QE announcement occurred on March 18, 2009, when Treasury (5-year) yields

[7] Lou et al. (2013) and Fleming and Liu (2016) show there are predictable price movements in the days and hours before and after the auction, but price movements are unpredictable very near the close of the auction. Hence, the use of small intraday windows is key to identifying unanticipated demand shocks.

fell by 23 basis points. Panel C presents statistics separately for periods of a binding/non-binding ZLB and for expansion/recessions. Regardless of regime, on average our demand shocks are close to zero.

To verify that these shocks are not spurious, we also report movements in Treasury yields on non-auction days (Panel D of Table 1; for days without auctions, the same "pre" and "post" windows are used as auctions in the same period). In all cases, the variance of the shocks on auction dates is larger than on non-auction dates. This pattern further suggests that surprise auction results influence secondary-market Treasury yields.

3.2 Narrative Evidence

To provide a better understanding of what forces are behind these surprise movements, Figure 2 plots 30-year Treasury yields during two 30-year Treasury bond auctions. The dashed vertical lines denote the close of the auction and the release of results, respectively. The first auction occurred on December 9, 2010. This auction was a reopening of previously issued 30-year bonds from the month prior. Yields are relatively stable in the lead up to the close of the auction, but drop sharply and immediately following the release of the the results. The Financial Times wrote:

> "Large domestic financial institutions and foreign central banks were big buyers at an auction of 30-year US Treasury bonds on Thursday. 'Investors weren't messing around...You don't get the opportunity to buy large amounts of paper outside the auctions and 'real money' were aggressive buyers.' "

The second is from an auction of newly issued bonds on August 11, 2011. When-issued yields were relatively stable prior to auction close, but after the close and release of the auction results, yields immediately rose. The Financial Times wrote:

> "An auction of 30-year US Treasury bonds saw weak demand...bidders such as pension funds, insurers and foreign governments shied away. 'There's not too many ways you can slice this one, it was a very poorly bid auction.' "

We interpret these two examples as follows. Before the auctions closed, the market information set consists of all the supply information, both for outstanding securities and the amount on offer for the current 30-year auction. The 30-year Treasury yields reflect beliefs about the expected path of short-term interest rates, inflation expectations, and demand for long-maturity Treasury securities. After the auctions closed and the results were released, the only update to the information set is the news regarding the bidding that took place in the auction, which solely reflects demand for Treasury debt. The change in the 30-year yields is a reaction to the unexpected shift in Treasury demand revealed at the auction (of course, this shift reflects many factors, including individual investors' beliefs about interest rates and economic fundamentals).

These articles also highlight why auctions can have important elements of price discovery: when investors wish to purchase large amounts of Treasuries to meet their needs, they may prefer to use auctions rather than make large transactions on the secondary market. As a result, auctions reveal new information about demand for government debt that is not already reflected in OTC secondary market trades.

3.3 Demand Determinants

The institutional structure of the primary market and narrative evidence from the financial press suggest that $D_t^{(m)}$ captures unexpected shifts in the demand for Treasuries. But because $D_t^{(m)}$ is an equilibrium response, it is important to establish that these shifts are related to observable measures of demand. Table 2 presents formal evidence by regressing our shocks on measures of demand reported at the auction:

$$D_t^{(m)} = \alpha^{(m)} + \beta^{(m)} X_t^{(m)} + \varepsilon_t^{(m)}, \tag{2}$$

where $X_t^{(m)}$ are various estimates of the shift in demand at a given auction. Columns (1)-(6) present estimates separately for auctions of different maturities, while column (7) pools across maturities. Panel A uses changes in the bid-to-cover (the change is taken relative to the most recent note or bond auction), which is a natural measure of demand shifts, since an increase in the bid-to-cover indicates higher demand relative to the amount of Treasuries offered. Consistent with our interpretation of $D_t^{(m)}$ as a reaction to unexpected demand, Panel A shows that an increase in the bid-to-cover ratio predicts a larger intraday fall in Treasury yields following the close of the auction.[8]

Our results show that the effect of typical surprise increases in demand is economically large. For example, a one standard deviation (0.45) increase in the bid-to-cover ratio in a Treasury auction for 10-year notes leads to a $2.66 \times 0.45 \approx 1.2$ basis point decline in 10-year Treasury yields. We can back out a simple estimate for the sensitivity of yields as a function of the change in quantity demanded (in terms of dollars). A typical offering amount in a 10-year Treasury note auction is between \$20 and \$30 billion. Hence, our estimates imply that an increase in demand for 10-year Treasuries by \$10 billion decreases 10-year yields by $2.66 \times \left(\frac{10}{30}, \frac{10}{20}\right) \approx (0.89, 1.33)$ basis points.

In order to assess sensitivity of our demand shocks to changes in demand by bidder type, Panel B reports estimates of equation (2) using the change in the bid-to-cover ratio of indirect bidders, direct bidders, and primary dealers. The sensitivity of surprises $D_t^{(m)}$ to unexpected demand of indirect bidders increases with maturity. For example, a unit

[8] Appendix Figure B4 reports binscatter plots of these results. Table 2 uses changes as there is predictable low-frequency movement in the bid-to-cover ratio (Appendix Figures B7 and B8). However, our results are not sensitive to this choice. Appendix Figures B5 and B6 repeat our analysis using the bid-to-cover in levels or a residualized measure of the bid-to-cover from a univariate AR(4) model.

increase in the bid-to-cover ratio for indirect bidders decreases the yields of 2-year Treasuries by 2.24 basis points, and the yields of 30-year Treasuries by 11.3 basis points. Direct bidders exhibit the same pattern, although the coefficients are smaller. The sensitivity to changes in the bid-to-cover ratio coming from primary dealers is also smaller. When we pool across maturities, demand of direct and especially indirect bidders generates *ceteris paribus* more variation in Treasury yields than demand of primary dealers, although for all bidder types an increase in bidder demand implies a decline in intraday yields $D_t^{(m)}$.

Panel C uses additional investor allotment data from the Treasury to break down the amount accepted by types of bidders: Investment Funds, Foreign, Dealers, and remaining smaller investors classes (aggregated into a "Miscellaneous" category). Since the fractions by group add up to one, we set Dealers as the leave-out category. Our estimates suggest that as the fraction accepted for investment funds and foreign buyers increases, $D_t^{(m)}$ declines. Estimates for the Miscellaneous category are generally smaller and less robust.

Our results indicate that movements in demand, proxied by the bid-to-cover ratio, are a key determinant of $D_t^{(m)}$. Furthermore, we observe that the demand from institutional investors is important in accounting for variation in $D_t^{(m)}$. We stress that our identifying assumption only relies on observing the changes in *yields* immediately following the close and release of auction results. Our empirical approach does not require a measure of the unanticipated movement in quantity demanded. Hence, our empirical analysis in the remainder of the paper only relies on our constructed measure $D_t^{(m)}$.

3.4 Co-movement Across Markets

We now turn to analyzing how our demand shocks for Treasuries propagate across other financial markets. We measure the impact of demand shocks on other assets by estimating univariate regressions of the form

$$y_t = \gamma + \phi D_t + u_t, \tag{3}$$

where y_t is the change in the price or yield of some asset on auction date t and D_t is our auction demand shock. We pool across all auction maturities to simplify presentation, but our results are robust to estimating equation (3) separately by maturity groups.

Where available, we use intraday changes measured in the same window as our shocks. We also examine daily changes, both due to data limitations for some series and because daily changes may pick up responses that do not occur immediately. A strong relationship between D_t and y_t indicates either that D_t and y_t have a common determinant (e.g., changes in inflation expectations alter the behavior of bids in Treasury auctions and change prices of inflation swaps) or that there is a propagation channel from D_t to y_t (e.g., higher revealed demand for Treasuries results in repricing of other debt securities).

Panel A of Table 3 reports results for debt markets. The dependent variable in the

first row is the intraday change in the *price* (co-moves negatively with the yield) of the Exchange Traded Fund (ETF) "LQD," the iShares iBoxx $ Investment Grade Corporate Bond tracking investment grade corporate bonds. The estimated coefficient $\hat{\phi}$ is interpreted as the impact in log points of a one basis point increase in D_t. We observe a strong reaction to the Treasury demand shock, accounting for more than 50 percent of variation observed in corporate bond ETF prices during the short windows around the close and release of the Treasury auction results. The second row reports the results using "HYG," the iShares iBoxx $ High Yield Corporate Bond ETF tracking high yield corporate bonds. Although the sign of the estimated coefficient is as expected, the pass-through from our demand shocks to high yield corporate bonds is weaker than that of investment grade bonds. The estimated coefficient is not statistically significant, and the magnitude is smaller than for investment grade debt by a factor of ten. Moreover, the R^2 is much lower, suggesting that our demand shocks account for only 1 percent of the observed variation in high yield corporate bond ETF prices during these intraday periods. The next two rows examine pass-through to mortgage rates, as measured by the ETFs "MBB" and "VMBS" (from iShares and Vanguard respectively; both ETFs track investment-grade mortgage-backed pass-through securities guaranteed by U.S. government agencies). Similar to our findings for investment-grade corporate debt, we find economically large and statistically significant pass-through of demand shocks to mortgage borrowing rates. We also find relatively high R^2s of 34% and 13%, respectively.

The next rows repeat the above analysis using daily measures of corporate bond *yields*, as measured by Moody's Aaa, Moody's Baa, and Bank of America's C corporate yield indices. Consistent with the intraday results, our demand shocks have a strong effect on safe (Aaa) corporate bonds. Moreover, the pass-through of our demand shocks to corporate bond yields is nearly one-to-one. However, using daily rather than intraday changes as the dependent variable leads to a decline in R^2, which underscores the benefits of using intraday data. While our demand shocks still have large effects on moderately safe debt (Baa), there appears to be much smaller transmission to highly risky corporate debt (C), consistent with our intraday findings.

Secondary market yields react not only strongly, but also persistently. Figure 3 plots the contemporaneous reaction of 10-year Treasury spot rates (top panel) and the Aaa corporate bond yields (bottom panel) to our shocks D_t, as well as the reactions up to 60 days in the future. The reaction remains strongly statistically significant over three weeks later, while the point estimate is quite stable even 2 months later. To provide a perspective on the magnitude of this persistence, Figure 3 also plots the change in yields following the QE1 announcement on March 18, 2009 (normalized such that the change on impact is equal to 1). Consistent with Wright (2012) and Greenlaw et al. (2018), yields had returned to roughly their starting point within a few months after QE1.

One may be concerned that these reactions in the bond market are driven by some

omitted factor, rather than idiosyncratic changes in institutional investors' demand for Treasuries. Our high-frequency identification goes a long way towards assuaging this concern, but it is still possible that our demand shocks reflect systematic changes in bidders' expectations of macroeconomic or financial fundamentals. This is an issue for our interpretation if two conditions hold: (i) this information was not already reflected in market prices, and instead is only revealed at the auction; (ii) the market as whole updates their beliefs about fundamentals by observing the results of the auction. If this "information factor" drives the bond market reactions we observe, we should find strong co-movement between D_t and indicators capturing beliefs about current or future states of the economy and financial markets. The "information factor" can take any number of forms, so it is difficult to conclusively rule out this channel; we focus on a battery of key indicators to assess the quantitative significance of this alternative explanation.

Panel B of Table 3 reports estimates for the response of equities to our demand shocks. Rows 1 and 2 report the results for the intraday change in ETFs tracking the S&P 500 and the Russell 2000 indices. Rows 3 and 4 are for the daily changes in these indices. Although the estimated slope is generally positive, the estimate is typically insignificant. Moreover, the quantitative importance is small, as the auction demand shocks account for a trivial share of variation in equities. Thus, there is little evidence on average for a common factor driving Treasury and equity prices during small windows around the release of auction results, or for strong propagation of our demand shocks to equity prices.

Panel C of Table 3 presents results for a wider array of assets. The dependent variable in row 1 is the intraday change in the ETF "GLD," which tracks the price of Gold Bullion. Row 2 reports results for the daily change in the S&P Total Commodity Index. For the Commodity Index we do not find a significant correlation with D_t. For Gold, while the relationship is statistically significant, the R^2 is very low. Rows 3 and 4 use the daily change in inflation expectations implied by inflation swaps at the 10- and 2-year horizons. Our demand shocks are not associated with a significant comovement in inflation expectations. Row 5 uses the daily change in the 3-month USD LIBOR-Overnight Index Swap spread (a common measure of credit risk in the banking sector). Rows 6 and 7 use daily changes in two CDS indices from Credit Market Analysis that track the automotive industry (a highly cyclical industry) and banks (a proxy for the financial sector). These three measures proxy for expectations about future output and market conditions. Our demand shocks have no tangible relationship with these measures, consistent with our interpretation that the shocks do not capture superior information of Treasury auction bidders about future recessions and the like. Finally, row 8 documents that D_t shocks are not associated with the VIX (a measure of market perceptions of future volatility). In short, these null results suggest that our demand shocks are not driven by changes in expectations regarding inflation, output, liquidity, default risk, or volatility.

As a final test, Panel D of Table 3 reports results with expected federal funds rates,

which are derived from federal funds futures contracts. We view this as a "catch-all" test of the "information factor" story: if our demand shocks reflect changes in market expectations of fundamentals, then the market should also expect a Fed response (to the extent these fundamentals matter for the macroeconomy). We find that our demand shocks are not associated with changes in expected federal funds rates: the estimated coefficients on our demand shocks are very close to zero in these specifications, and demand shocks explain virtually none of the variation in federal funds futures in our sample.[9]

3.5 Localization Hypothesis

The results of Tables 2 and 3 allow for some broad observations. First, given our high-frequency approach and the institutional structure of Treasury auctions, our constructed shocks are likely only driven by new information regarding the demand side of the market. Second, these shifts are largely driven by shifts in the demand arising from institutional investors. Third, these demand shocks from the primary market for Treasuries propagate to private borrowing rates. We do not observe the underlying sources of demand shocks, and indeed it is likely driven by a host of factors, including shifts in a given investor's idiosyncratic beliefs about the macroeconomy. However, these demand shifts are not driven by *systematic* or market-wide shifts in macroeconomic expectations (e.g., flight to quality or inflation expectations) that may move demand for Treasuries at all maturities.

We now turn to testing the key predictions of *preferred habitat* theory, first discussed in Modigliani and Sutch (1966). Preferred habitat theory posits that certain investor clienteles specialize in bonds of specific maturities; idiosyncratic positive (negative) shocks to their demand leads to decreases (increases) in bond yields. This prediction contrasts with the neutrality result of standard theories and the expectations hypothesis: holding fixed the expected path of short rates, idiosyncratic demand shocks for specific bonds should have no effect on yields. Our results in Table 2 are consistent with the basic preferred habitat theory (and are inconsistent with the neutrality result of the expectations hypothesis): increases in idiosyncratic demand (proxied by the bid-to-cover ratio at a given auction) imply decreases in yields following the close of a given auction.

However, a naïve preferred habitat view where yields for a given bond are determined *solely* by idiosyncratic demand shocks is unrealistic, as this would imply large arbitrage profits to be made by term-structure arbitrageurs. Thus, modern preferred habitat theory (Vayanos and Vila (2021)) formalizes the interaction between such arbitrageurs and preferred habitat investors. One of the characteristic predictions of modern preferred habitat theory is the *localization hypothesis*. When arbitrageur risk-bearing capacity is high, habitat demand shocks have *global* effects on the yield curve. That is, the relative response of the interest rates across the yield curve does not depend on where in maturity

[9] We plot rolling regressions for select asset prices in Appendix Figure B9. Appendix Figure B10 plots rolling regressions for h-month ahead federal funds futures h up to 12 months.

space the demand shock occurs. However, as arbitrageur risk-bearing capacity declines, the spillovers of demand shocks become more *localized*. That is, the relative response of interest rates becomes more concentrated on parts of the yield curve that are closer in maturity space to where the demand shock occurs.

Intuitively, as risk-bearing capacity falls, term-structure arbitrageurs find it more and more costly to integrate the markets for bonds of different maturities. Thus, demand shocks in one maturity segment of the yield curve will have relatively larger effects for bonds with similar maturities to those directly affected by the shift in demand. Yields for maturities which are far removed from those of the bonds directly affected by the demand shock will instead exhibit relatively small movements.

Thus, testing the localization hypothesis requires empirical measures of the conditional response of yields to demand shifts for short- and long-maturity Treasuries, holding all other risk factors constant. Our high-frequency identification strategy is designed to isolate the reactions of the yield curve to Treasury demand shocks alone, effectively ruling out any other shocks during these small windows around auctions. This insight allows us to arrive at a simple regression specification that can test the localization hypothesis using only high-frequency changes in yields around auctions $D_t^{(\tau)} = y_t^{(\tau),post} - y_t^{(\tau),pre}$:

$$D_t^{(\tau)} = \alpha^{(\tau)} + \gamma_s^{(\tau)} \mathbf{I}(m_t = short) D_t^{(\tau^*)} + \gamma_\ell^{(\tau)} \mathbf{I}(m_t = long) D_t^{(\tau^*)} + \varepsilon_t^{(\tau)}, \tag{4}$$

where $\mathbf{I}(m_t = short)$ and $\mathbf{I}(m_t = long)$ are indicator variables equal to one if there is an auction of short-maturity and long-maturity bonds on date t, respectively. We estimate equation (4) for all maturities τ, holding some baseline maturity τ^* fixed. Coefficients $\hat{\gamma}_s^{(\tau)}$ and $\hat{\gamma}_\ell^{(\tau)}$ measure the relative effect of demand shocks for short-maturity and long-maturity auctions respectively. Now we can restate the localization hypothesis formally as a function of the coefficients $\gamma_s^{(\tau)}, \gamma_\ell^{(\tau)}$ as follows: i) when risk-bearing capacity is high, $\left|\gamma_s^{(\tau)} - \gamma_\ell^{(\tau)}\right| \to 0$ for all maturities τ; ii) when risk-bearing capacity is low, then $\gamma_s^{(\tau)} > \gamma_\ell^{(\tau)}$ if $\tau < \tau^*$ and $\gamma_s^{(\tau)} < \gamma_\ell^{(\tau)}$ if $\tau > \tau^*$.

3.6 Empirical Localization Results

To test the state-dependent localization hypothesis, we estimate specification (4) separately for two subsamples: a "non-crisis" period (when risk-bearing capacity is high), and a "crisis" period (when risk-bearing capacity is low). Testing for the equality of coefficients $\hat{\gamma}_s^{(\tau)}$ and $\hat{\gamma}_\ell^{(\tau)}$ in each subsample shows whether demand shocks have more localized effects when risk-bearing capacity is low. In our baseline estimates, we take the "crisis" period to be 2008-2012. We divide the auctions into short-maturity (maturity of 5 years or less) and long-maturity (maturity of 7 years or greater). The results are robust to alternative choices.

Since we construct $D_t^{(\tau)}$ from secondary-market yields, we do not have measures for all

maturities τ at all times. Our approach is to run rolling regressions (across maturities), including yields for maturities within ± 2 years for each maturity $\tau \leq 20$ years and within ± 4 years for each maturity $\tau > 20$ year. Given that the choice of benchmark maturity τ^* is arbitrary, we set $\tau^* = 3$, with an eye towards applying our results to QE in order to focus on the differential effects of intermediate and long-maturity yields.

Panel A of Figure 4 plots the estimates of $\hat{\gamma}_s^{(\tau)}$ and $\hat{\gamma}_\ell^{(\tau)}$ for the non-crisis sample. The estimated responses follow a similar hump-shaped pattern peaking at intermediate maturities (around roughly 5-7 years), and then declining and stabilizing for longer term maturities (with a possible slight uptick for very long (20+ years) maturities). In general, we cannot reject equality of $\hat{\gamma}_s^{(\tau)}$ and $\hat{\gamma}_\ell^{(\tau)}$.

When we estimate specification (4) on the crisis sample, we observe a strikingly different pattern (Panel B of Figure 4). In response to shocks in demand for short- and long-maturity Treasuries, both $\hat{\gamma}_s^{(\tau)}$ and $\hat{\gamma}_\ell^{(\tau)}$ increase rapidly for $\tau < \tau^*$, with the $\hat{\gamma}_s^{(\tau)}$ estimates increasing somewhat more quickly than $\hat{\gamma}_\ell^{(\tau)}$. Once we move to maturities greater than τ^*, the estimates quickly diverge. Specifically, in response to shocks in demand for short-maturity Treasuries, the yields for maturities greater than τ^* fall relative to the yield response for τ^* (as shown by the estimates of $\hat{\gamma}_s^{(\tau)}$). On the other hand, in response shocks in demand for long-maturity Treasuries, yields for maturities $\tau > \tau^*$ continue to increase relative to the benchmark maturity τ^* (as shown by the estimates of $\hat{\gamma}_\ell^{(\tau)}$). For maturities of 20 to 30 years, the relative response of yields to long-maturity demand shocks ($\hat{\gamma}_\ell^{(\tau)}$) is over twice as large as the response of yields to short-maturity demand shocks ($\hat{\gamma}_s^{(\tau)}$). We can strongly (at the 0.1% level) reject the null of $\hat{\gamma}_s^{(\tau)}$ and $\hat{\gamma}_\ell^{(\tau)}$ being equal.

These results are consistent with the key predictions of preferred habitat theory: during "normal" periods when financial risk-bearing capacity is high, demand shocks for short- and long-maturity securities have relatively similar impacts on the yield curve. During periods of financial distress when risk-bearing capacity is low, the impacts are more localized: the impact of short-maturity demand shocks are largest for short maturities, while the impact of long-maturity demand shocks peaks at the long end of the term structure. These results provide support for the view that during financial crises, arbitrageurs are less willing or able to integrate bond markets.

The localization results are robust to a variety of alternative assumptions. In the interest of space, we provide summaries for a few key robustness specifications in this section. First, we explore the sensitivity of our results to using alternative measures of financial distress. Specifically, we consider two alternatives: i) an aggregate "intermediary capital ratio," a market-based measure of financial distress (low intermediary capital ratios are associated with lower risk-bearing capacity) described in He et al. (2016); ii) a narrative-based measure of financial crisis from Romer and Romer (2017) (higher values of the crisis indicator are associated with lower risk-bearing capacity). Both of these alternatives generate results similar to our baseline findings (see Appendix Figures B12 through

B16). Second, our results are robust to using different cutoffs for separating auctions into short and long maturities. For example, when we use 10 years (rather than 7 years in the baseline) as the cut-off for long-maturity auctions, the results (Appendix Figure B17) still strongly support the localization hypothesis and, if anything, the point estimates for the "non-crisis" period are even more similar than in our baseline specification. Third, as we discussed above, the choice of the benchmark maturity τ^* is arbitrary in our framework. To verify that this choice is indeed immaterial for our results, we experiment with values other than $\tau^* = 3$ and find similar results (e.g. Appendix Figure B18 reports estimates for $\tau^* = 6$). Fourth, our results are nearly identical when dropping auctions that occurred during the weeks of QE announcements (Appendix Figure B19).

Note that specification (4) does not require us to observe the underlying shock to demand *quantity*, but instead only relies on the change in yields conditional on such (unobserved) changes in demand. However, the change in the bid-to-cover ratio is a proxy for demand shifts. Although we do not have a high-frequency market-based measure of the expected bid-to-cover before and after an auction, we nevertheless considered an alternative regression specification to test the localization hypothesis:

$$D_t^{(\tau)} = \alpha^{(\tau)} + \nu_s^{(\tau)} \mathbf{I}(m_t = short)\tilde{\beta}_t + \nu_\ell^{(\tau)} \mathbf{I}(m_t = long)\tilde{\beta}_t + \varepsilon_t^{(\tau)}, \tag{5}$$

where $\tilde{\beta}_t$ is the change in the bid-to-cover ratio following an auction at date t (as in Table 2). Figure 5 reports the results (after flipping the sign of the coefficients for comparability with our baseline specification). Unlike specification (4), the estimates $\hat{\nu}_s^{(\tau)}, \hat{\nu}_\ell^{(\tau)}$ measure the *absolute* response of the yield curve to shifts in demand (proxies), as opposed to the *relative* response. Hence the formal test of the localization hypothesis is not as straightforward. With this caveat, we see clear patterns of localization in the "crisis" estimates, but little evidence of localization in the "non-crisis" estimates. In the non-crisis subsample estimates (Panel A of Figure 5), the response to both short- and long-maturity demand follows a similar hump-shaped pattern: the response is increasing from short to intermediate maturities, peaking around $\tau = 5$ to $\tau = 10$ before declining (though there is some evidence that long demand shocks have larger effects on very long-maturity yields). On the other hand, the estimates from the crisis subsample (Panel B of Figure 5) exhibit significant differences across short and long estimates: the response to short demand shocks remains hump-shaped, but peaks for short maturities before quickly declining. Conversely, the response to long demand shocks increases almost without fail as the maturity increases, peaking at very long maturities.[10]

We additionally explore state-dependence in the pass-through of Treasury demand

[10] Appendix Figures B20 and B21 repeat the estimation of specification (5), but use the level of the bid-to-cover, or the residuals of a univariate AR(4) model. Appendix Figures B22-B24 conduct additional robustness exercises using measures of the bid-to-cover only from Indirect Bidders. In all cases, the patterns observed are very similar to our results in Figure 5.

shocks to private borrowing rates. Although preferred habitat models typically only feature risk-free bonds, the logic of the localization hypothesis also suggests that the pass-through to risky assets of Treasury demand shocks should lessen when risk-bearing capacity is low. Panel A of Table 4 estimates a version of equation (4) where we regress the intraday change of various ETF prices on $D_t^{(\tau^*)}$, interacted with whether the auction took place during the crisis period. For very safe or very risky corporate borrowing (LQD and HYG in columns 1 and 2, respectively), we find little evidence of localization. Although the sign of the crisis interaction coefficient in both cases is positive, it is not statistically significant. On the other hand, mortgage borrowing rates (MBB and VMBS in columns 3 and 4, respectively) seem to exhibit substantial localization. In both cases, the interaction coefficient is positive, economically large compared to the non-crisis coefficient, and highly statistically significant. Next, we explore the pass-through to equity indices. For the S&P500 (SPY, column 5), we find little evidence for localization; while for the Russell 2000 (IWM, column 6), there is some support for localization: the coefficient is positive in non-crisis times, but the interaction coefficient suggests that during crisis times the pass-through drops to zero. However, in both cases, the R^2 is low.

Finally, we explore state- and maturity-dependent localization for safe corporate borrowing rates. We estimate a version of equation (4), but use intraday changes in various corporate bond ETFs as the dependent variable interacted with both the regime (crisis/non-crisis) and maturity (short/long) of the auction. We utilize two sets of ETFs: "BSV," "BIV," and "BLV" from Vanguard; and "CSJ," "CIU,", and "CLY" from iShares. These ETFs track investment-grade bonds with short, intermediate, and long maturities respectively.[11] The results for this exercise (Panel B of Table 4) are not directly comparable to our baseline Treasury results from Figure 4 because the ETFs are measured as (log) price changes and because we do not observe the entire term structure for corporate debt. Nevertheless, the results are consistent with our findings in Figure 4. The first row shows that demand shocks in Treasury auctions also have large effects on corporate bond prices across the term structure. The second row shows that this pass-through is similar for demand shocks originating in long-term auctions in non-crisis times (with the one exception of the coefficient for BLV). The third and fourth rows report the interaction coefficients for the "crisis" periods. The final row shows that the pass-through of demand shocks for long-maturity Treasuries occurring during crisis periods has the same kind of localization pattern as Figure 4. The coefficient becomes monotonically more negative as we move from short-maturity to long-maturity bond ETFs. In particular, the coefficient for the long-maturity ETFs is negative and statistically significant.

[11] The Vanguard ETFs were introduced in 2007, but include Treasuries in addition to corporate bonds. The iShares ETFs only include corporate bonds, but have shorter time series.

4 A New Keynesian Preferred Habitat Model

The results of Section 3 show that our shock series offer useful variation to study empirically the financial effects of Treasury demand shocks. However, our end goal is to make sense of the transmission of QE, both to asset prices and to the broader macroeconomy. Thus, in order to make more progress understanding shock transmission along these dimensions, we need to develop a theoretical framework with meaningful interactions between financial markets and the macroeconomy. This will allow us to be precise about the transmission mechanisms of demand shocks, as well as help us extrapolate from private shocks observed at auctions to large scale asset purchases (or sales). Informed by our empirical findings in Section 3, this section develops such a model.

4.1 Overview: Agents and Financial Assets

The financial block of our model builds on Vayanos and Vila (2021), extended to allow for both riskless and risky assets. Formally, we assume two sets of assets: riskless and risky zero coupon bonds with maturity $\tau \in (0, T)$. A τ-maturity riskless bond has price $P_t^{(\tau)}$ and pays \$1 at maturity with certainty in period $t + \tau$. A risky τ-maturity bond has price $\tilde{P}_t^{(\tau)}$ and instead pays $D_{t+\tau} \equiv e^{d_{t+\tau}}$ at maturity in period $t + \tau$, where D_t is a stochastic process. We will assume that both sets of assets are in zero net supply, but for completeness we denote the supply of assets by $S_t^{(\tau)}$ and $\tilde{S}_t^{(\tau)}$, respectively.

As in Vayanos and Vila (2021), financial assets are traded between two sets of investors: arbitrageurs and preferred habitat funds. Crucially, there exist "clientèle" investors who have idiosyncratic demand ("preferred habitat") for specific assets and maturities. For example, pension funds can have a preference for long-maturity assets to better match the maturity structure of pension liabilities. The other side of the market are risk-averse arbitrageurs such as hedge funds and dealers, who smooth out these demand shocks.

The macro side of our model is largely built on a standard New Keynesian framework. A representative household consumes and supplies labor to firms producing differentiated goods while facing pricing frictions. Equilibrium aggregate dynamics will depend on two familiar equations: an IS curve which determines the dynamics of the output gap x_t, and a Phillips curve which determines the dynamics of inflation π_t. The links between the financial and the macroeconomic blocks of our model are two-fold. First, the central bank sets the policy rate i_t in reaction to aggregate fluctuations. In equilibrium, i_t will be the return on a riskless bond as maturity $\tau \to 0$; therefore, aggregate fluctuations will result in changes in asset prices. Second, we depart from the usual New Keynesian setup and assume that households cannot access financial markets directly, but instead must borrow through the preferred habitat funds. Through the household Euler equation, output in equilibrium will depend on a weighted average of the returns on both short- and long-term assets, rather than only on the policy rate. Therefore, asset price movements

will lead to changes in macroeconomic conditions. Our model formalizes this intuition, and characterizes the general equilibrium solution. We stress that our aim is to develop a tractable model, and discuss the simplifying assumptions necessary for our model.

Households: The representative household chooses a consumption bundle C_t and labor supply N_t given nominal wealth A_t in order to maximize lifetime discounted utility. The household faces a nominal wage $\mathcal{W}_t$ and aggregate price index P_t. Our point of departure from a textbook New Keynesian model is that households can only invest with the preferred habitat funds, and allocate a fixed fraction of their wealth across these funds. The household problem is given by

$$\max E_0 \int_0^\infty e^{-\rho t} \Psi_t \left(\frac{C_t^{1-\varsigma}}{1-\varsigma} - \frac{N_t^{1+\phi}}{1+\phi} \right) \mathrm{d}t \tag{6}$$

$$\text{s.t. } \mathrm{d}A_t = [\mathcal{W}_t N_t - P_t C_t]\,\mathrm{d}t + A_t \int_0^T \left(\eta(\tau) \frac{\mathrm{d}P_t^{(\tau)}}{P_t^{(\tau)}} + \tilde{\eta}(\tau) \frac{\mathrm{d}\tilde{P}_t^{(\tau)}}{\tilde{P}_t^{(\tau)}} \right) \mathrm{d}\tau + \mathrm{d}F_t\,, \tag{7}$$

and A_0 given. The parameter ϕ is the Frisch elasticity of labor supply, and the parameter ρ is the household discount factor. We introduce aggregate demand shocks through discount rate shocks Ψ_t. The term $\mathrm{d}F_t$ captures all (flow) nominal transfers to the household. Households therefore borrow or lend their wealth A_t at an "effective" borrowing rate (rather than the short-term policy rate), given by the the integral term in the household budget constraint. The weight functions satisfy $\int_0^T (\eta(\tau) + \tilde{\eta}(\tau))\,\mathrm{d}\tau = 1$. Thus, household borrowing depends on a weighted average of the returns across the term structure and asset types. Denote the drift of the effective borrowing rate by

$$\hat{r}_t \equiv E_t \int_0^T \left(\eta(\tau) \frac{\mathrm{d}P_t^{(\tau)}}{P_t^{(\tau)}} + \tilde{\eta}(\tau) \frac{\mathrm{d}\tilde{P}_t^{(\tau)}}{\tilde{P}_t^{(\tau)}} \right) \mathrm{d}\tau\,, \tag{8}$$

which is endogenous but taken as given by households.

We think of the weighting functions $\eta(\tau), \tilde{\eta}(\tau)$ as capturing a type of preferred habitat for households (for instance, due to demographics such as age) without explicitly introducing heterogeneous households. This also captures imperfect access to financial markets, as households typically do not directly buy and sell financial assets such as corporate bonds; instead, investments are done through pension funds and mutual funds, which tend to allocate their portfolio using fixed weights and rebalance sluggishly (e.g., Koijen and Yogo (2019), Koijen and Yogo (2022), Bretscher et al. (2022), Peng and Wang (2022)).

To be clear, these weighting functions abstract from many important features of household finance such as borrowing constraints and heterogeneity of households as well as financial intermediation that offers a variety of products ranging from mortgages to credit card debt (e.g., Gomes et al. (2021) and Zinman (2015) for recent surveys). This creates some distance between the model and the data. For example, QE can change access

to credit at different maturities by affecting banks' balance sheets, inducing endogenous changes in the weighting functions $\eta(\tau)$ and $\tilde{\eta}(\tau)$ (e.g., Rodnyansky and Darmouni (2017) and Di Maggio et al. (2020)). In a similar spirit, borrowing constraints and household heterogeneity call for more cross-sectional and time-series variation in $\eta(\tau)$ and $\tilde{\eta}(\tau)$ (e.g., Cui and Sterk (2021)).

Our key assumption that we can take the weighting functions as model primitives is strong. However, this assumption permits a tractable solution and keeps the macroeconomic dynamics as close as possible to a standard New Keynesian model, while still allowing fluctuations in long-term borrowing rates to affect household decisions. Enriching the model with elements necessary to capture additional financial complexities is a fruitful avenue for future research.

Firms: The production block of the model is standard. Intermediate goods are produced by monopolistically competitive firms indexed by $j \in [0,1]$. Each firm hires workers and produces using linear technology $Y_{j,t} = N_{j,t}$. Firms face Rotemberg pricing frictions when choosing prices, and must pay a cost $\Theta(\pi_{j,t}) \equiv \frac{\theta}{2}\pi_{j,t}^2 P_t Y_t$, where $\frac{\mathrm{d}P_{j,t}}{P_{j,t}} = \pi_{j,t}\,\mathrm{d}t$ is the inflation rate of firm j. Firms are owned by households and hence maximize expected real profits discounted using the household real stochastic discount factor $e^{-\rho t}Q_t$. The firm problem is therefore given by

$$\max E_0 \int_0^\infty e^{-\rho t} Q_t \frac{\Pi_{j,t}}{P_t}\,\mathrm{d}t\,, \quad \text{where } \Pi_{j,t} = P_{j,t}Y_{j,t} - \mathcal{W}_t N_{j,t} - \Theta(\pi_{j,t}). \tag{9}$$

Intermediate goods are sold to competitive final goods producers with time-varying demand elasticity ϵ_t. This implies the usual demand and price index for differentiated goods (see e.g., Gali (2015)): $Y_{j,t} = \left(\frac{P_{j,t}}{P_t}\right)^{-\epsilon_t} Y_t$ where the price index $P_t = \left(\int_0^1 P_{j,t}^{1-\epsilon_t}\right)^{\frac{1}{1-\epsilon_t}}$.

Arbitrageurs: Arbitrageurs have mean-variance preferences and allocate their wealth W_t to holdings $X_t^{(\tau)}, \tilde{X}_t^{(\tau)}$ across τ riskless and risky bonds, respectively. They solve

$$\max E_t\,\mathrm{d}W_t - \frac{a}{2} Var_t\,\mathrm{d}W_t \tag{10}$$

$$\text{s.t. } \mathrm{d}W_t = W_t i_t\,\mathrm{d}t + \int_0^T X_t^{(\tau)} \left(\frac{\mathrm{d}P_t^{(\tau)}}{P_t^{(\tau)}} - i_t\,\mathrm{d}t\right) + \tilde{X}_t^{(\tau)} \left(\frac{\mathrm{d}\tilde{P}_t^{(\tau)}}{\tilde{P}_t^{(\tau)}} - i_t\,\mathrm{d}t\right)\mathrm{d}\tau\,, \tag{11}$$

where the parameter a governs the risk-return trade-off that arbitrageurs face. This parameter can be taken literally as a risk aversion parameter, or more generally can be thought of as a proxy for factors that lead to the imperfect risk-bearing capacity of arbitrageurs. Arbitrageurs transfer all profits or losses to the household each period. Taking a as time-invariant is important for tractability; however, we might expect a to increase in periods of financial distress (e.g., He and Krishnamurthy (2013), Kyle and Xiong (2001)). We analyze how the predictions of the model depend on the risk-bearing capacity of arbitrageurs.

Preferred Habitat Funds: On the other side of the market is a continuum of habitat investors, who specialize in (riskless and risky) bonds of specific maturities. We follow Vayanos and Vila (2021) and assume that these investors respond to prices through the following demand curves:

$$Z_t^{(\tau)} = -\alpha(\tau)\log P_t^{(\tau)} - \beta_t(\tau), \quad \tilde{Z}_t^{(\tau)} = -\tilde{\alpha}(\tau)\log \tilde{P}_t^{(\tau)} - \tilde{\beta}_t(\tau). \tag{12}$$

The functions $\alpha(\tau), \tilde{\alpha}(\tau)$ are the semi-elasticities of a τ-habitat investor's demand for riskless and risky assets, respectively. The time-varying demand intercept is given by $\beta_t(\tau), \tilde{\beta}_t(\tau)$, where we assume a factor structure such that

$$\beta_t(\tau) = \sum_{k=1}^{K} \theta^k(\tau)\beta_t^k - \zeta(\tau), \quad \tilde{\beta}_t(\tau) = \sum_{k=1}^{\tilde{K}} \tilde{\theta}^k(\tau)\tilde{\beta}_t^k - \tilde{\zeta}(\tau). \tag{13}$$

The functions $\theta^k(\tau), \tilde{\theta}^k(\tau)$ govern how demand factors $\beta_t^k, \tilde{\beta}_t^k$ lead to changes in demand from τ-habitat investors. In general, there are no restrictions on the dynamics of the demand factors (up to linearizing). For our quantitative analysis, we assume that the habitat demand factors are independent from one another, but may respond to the short rate (as in King (2019a)). Hence, for each demand factor we have

$$\mathrm{d}\beta_t^k = -\left(\kappa_\beta^k \beta_t^k + \phi_{i,\beta}^k i_t\right)\mathrm{d}t + \sigma_\beta^k \,\mathrm{d}B_{\beta^k,t}\,, \tag{14}$$

where the parameters $\kappa_\beta^k, \phi_{i,\beta}^k$, and σ_β^k govern the mean-reversion, response to short rates, and volatility of each demand factor, respectively (and with analogous parameterization and dynamics for demand factors $\tilde{\beta}_t^k$).

The τ-maturity riskless and risky funds have wealth $W_t^{(\tau)}$ and $\tilde{W}_t^{(\tau)}$. In order to fund their positions in their respective asset classes, the fund receives investment from the household sector based on the weights $\eta(\tau), \tilde{\eta}(\tau)$. Any remaining wealth is invested at the short rate. The budget constraint of a τ-maturity riskless fund is therefore

$$\mathrm{d}W_t^{(\tau)} = Z_t^{(\tau)}\frac{\mathrm{d}P_t^{(\tau)}}{P_t^{(\tau)}} + \left[W_t^{(\tau)} - Z_t^{(\tau)} + \eta(\tau)A_t\right] i_t\,\mathrm{d}t - \eta(\tau)A_t\frac{\mathrm{d}P_t^{(\tau)}}{P^{(\tau)}},$$

and the risky fund faces an analogous budget constraint. Each preferred habitat fund also transfers profits or losses to the household each period.

Interest Rate and Payoff Processes: To close the model, we need to specify how the instantaneous interest rate i_t and the payoff process d_t is determined. We assume the central bank sets the nominal interest rate i_t according to the following policy rule:

$$\mathrm{d}i_t = -\kappa_i(i_t - \phi_\pi\pi_t - \phi_x x_t - i^*)\,\mathrm{d}t + \sigma_i\,\mathrm{d}B_{i,t}\,, \tag{15}$$

where the response of the policy rate to inflation and the output gap are governed by ϕ_π, ϕ_x, κ_i is a mean-reversion parameter, and i^* is the central bank's target policy rate. Note that if $\kappa_i \to \infty$, we recover a standard Taylor rule.

Finally, because risky assets are in zero net supply ($\tilde{S}_t^{(\tau)} = 0$), they can be interpreted as synthetic securities which are created by the financial sector (since arbitrageurs and preferred-habitat funds take non-zero positions in risky assets across all maturities). We assume that the payoff processes for these synthetic securities evolves according to

$$\mathrm{d}d_t = -\kappa_d \left(d_t - \psi_x x_t - \psi_\pi \pi_t - d^*\right) \mathrm{d}t + \sigma_d \, \mathrm{d}B_{d,t} \,. \tag{16}$$

This reduced-form approach to modelling payoffs of risky assets allows us to capture salient features of how private borrowing rates fluctuate over the business cycle. The process is mean-reverting (with inertia κ_d) and may depend on movements in output and inflation through parameters ψ_x, ψ_π (which are primitives of the model). Stepping outside of the model, assets besides financial institution securities influence the borrowing decisions of the household sector. In reality, this also includes risky private bonds, mortgage debt, equities, and a host of other assets. Rather than complicate the model by specifying a corporate structure for firms, we collapse all of these sources of payoff risk into a single process $D_t \equiv e^{d_t}$, which we allow to respond in a flexible manner to aggregate dynamics.[12]

4.2 Equilibrium

The household and firm problems are highly similar to a textbook New Keynesian model; detailed derivations are in Appendix C. The optimizing behavior of these agents and market clearing give rise to a Phillips curve and IS curve for inflation and the output gap π_t, x_t, where the output gap is the deviation of output from the "natural" level that would prevail in the absence of nominal rigidities and fluctuations in desired markups ($\theta = 0, \epsilon_t = \bar{\epsilon}$). The only modification to the standard (linearized) New Keynesian dynamics is that the effective borrowing rate $\hat{r}_t$ replaces the policy rate i_t. We have

$$E_t \, \mathrm{d}\pi_t = \left(\rho \pi_t - \delta x_t - z_{\pi,t}\right) \mathrm{d}t \,, \tag{17}$$

$$E_t \, \mathrm{d}x_t = \varsigma^{-1} \left(\hat{r}_t - \pi_t - \bar{r} - z_{x,t}\right) \mathrm{d}t \,, \tag{18}$$

where the model is linearized around a steady state with $\bar{\pi} = \bar{x} = 0$. The parameter δ measures the aggregate degree of price rigidity, and $\bar{r}$ is the "natural" borrowing rate.

[12] Our results are not sensitive to this precise microfoundation of the payoff process. We could have chosen a formulation whereby the payoff process D_t captures claims on a portion of firm profits not paid directly and immediately to households. Under the zero net supply assumption and with the payoff process given by equation (16), the macroeconomic dynamics of the model are identical. Although this formulation brings the interpretation of D_t closer to corporate bonds, it does not explicitly model the financing frictions leading to this distribution of profits (and our model does not feature firm investment).

The aggregate supply and demand shocks are assumed to evolve according to

$$\mathrm{d}z_{\pi,t} = -\kappa_{z_\pi} z_{\pi,t}\,\mathrm{d}t + \sigma_{z_\pi}\,\mathrm{d}B_{z_\pi,t}\,, \tag{19}$$

$$\mathrm{d}z_{x,t} = -\kappa_{z_x} z_{x,t}\,\mathrm{d}t + \sigma_{z_x}\,\mathrm{d}B_{z_x,t}\,. \tag{20}$$

The parameters $\kappa_{z_\pi}, \kappa_{z_x}$ govern the persistence in these processes, and the terms $B_{z_\pi,t}, B_{z_x,t}$ are standard independent Brownian motions, with respective volatility terms σ_{z_π}, σ_x.

We collect the state variables, jump variables, and Brownian terms into vectors $\mathbf{y}_t = \begin{bmatrix} i_t & d_t & \dots & \beta_t^k & \dots \end{bmatrix}^\top$, $\mathbf{x}_t = \begin{bmatrix} \pi_t & x_t \end{bmatrix}^\top$, and $\mathbf{B}_t = \begin{bmatrix} B_{i,t} & B_{d,t} & \dots & B_{\beta^k,t} & \dots \end{bmatrix}^\top$, respectively. The following Lemma describes the aggregate dynamics of the model, taking as given the effective borrowing rate. All proofs are in Appendix A.

Lemma 1 (Aggregate Dynamics)**.** *Suppose the effective borrowing rate drift is given by*

$$\hat{r}_t = \hat{\mathbf{A}}^\top \mathbf{y}_t + \hat{C}. \tag{21}$$

Then the rational expectations equilibrium is given by

$$\mathrm{d}\mathbf{y}_t = -\boldsymbol{\Gamma}\left(\mathbf{y}_t - \bar{\mathbf{y}}\right)\mathrm{d}t + \boldsymbol{\sigma}\,\mathrm{d}\mathbf{B}_t\,, \tag{22}$$

$$\mathbf{x}_t - \bar{\mathbf{x}} = \boldsymbol{\Omega}\left(\mathbf{y}_t - \bar{\mathbf{y}}\right), \tag{23}$$

where the matrices $\boldsymbol{\Gamma}, \boldsymbol{\Omega}$ are a function of the eigendecomposition of the linearized dynamics of the model (and therefore functions of $\hat{\mathbf{A}}$).

If $\hat{\mathbf{A}} = \mathbf{e}_i$, the vector which "selects" the policy rate $\mathbf{e}_i^\top \mathbf{y}_t = i_t$, then the effective borrowing rate responds one-for-one with the policy rate i_t and the aggregate dynamics of the model reduce to a standard New Keynesian model.[13]

Next, we turn to characterizing the behavior of asset prices. We focus on a solution to the model in which (log) asset prices are affine functions of the state variables, given by (endogenous) coefficient functions:

$$-\log P_t^{(\tau)} = \mathbf{A}(\tau)^\top \mathbf{y}_t + C(\tau),\ -\log \tilde{P}_t^{(\tau)} = \tilde{\mathbf{A}}(\tau)^\top \mathbf{y}_t + \tilde{C}(\tau). \tag{24}$$

Lemma 2 (Asset Prices)**.** *Suppose that the state variables $\mathbf{y}_t$ evolve according to equation (22). Then the affine coefficients in equation (24) are given by*

$$\mathbf{A}(\tau) = \left[\mathbf{I} - e^{-\mathbf{M}\tau}\right]\mathbf{M}^{-1}\mathbf{e}_i,\ \ \tilde{\mathbf{A}}(\tau) = \left[\mathbf{I} - e^{-\mathbf{M}\tau}\right]\mathbf{M}^{-1}\mathbf{e}_i - e^{-\mathbf{M}\tau}\mathbf{e}_d. \tag{25}$$

where $\mathbf{e}_i, \mathbf{e}_d$ are vectors such that $\mathbf{e}_i^\top \mathbf{y}_t = i_t$ and $\mathbf{e}_d^\top \mathbf{y}_t = d_t$. The matrix $\mathbf{M}$ solves the fixed

[13] In this case, the central bank can ensure determinacy when the response to inflation $\phi_\pi > 1$. However, in general $\hat{\mathbf{A}} \neq \mathbf{e}_i$ and the determinacy conditions are more complicated. In our quantitative exercises, we verify that these determinacy conditions are satisfied numerically.

point problem:

$$\begin{aligned}\mathbf{M} = \mathbf{\Gamma}^\top - a \int_0^T &\left[-\alpha(\tau)\mathbf{A}(\tau) + \mathbf{\Theta}(\tau)\right] \mathbf{A}(\tau)^\top \\ &+ \left[-\tilde{\alpha}(\tau)\tilde{\mathbf{A}}(\tau) + \tilde{\mathbf{\Theta}}(\tau)\right] \tilde{\mathbf{A}}(\tau)^\top \, \mathrm{d}\tau \, \boldsymbol{\sigma}\boldsymbol{\sigma}^\top,\end{aligned} \tag{26}$$

where $\mathbf{\Theta}(\tau)$, $\tilde{\mathbf{\Theta}}(\tau)$ *stack the habitat demand functions* $\theta^k(\tau)$, $\tilde{\theta}^k(\tau)$ *into vectors.*

The matrix $\mathbf{M}$ can be thought of as the risk-adjusted dynamics of the state. Note that when arbitrageurs are risk-neutral ($a = 0$), we have $\mathbf{M} = \mathbf{\Gamma}^\top$. However, when $a \neq 0$, $\mathbf{M}$ appears on both sides of equation (26) (through the affine coefficients $\mathbf{A}(\tau), \tilde{\mathbf{A}}(\tau)$).

With the results in Lemmas 1 and 2, we can characterize the equilibrium of the model.

Proposition 1 (Existence and Uniqueness)**.** *An affine equilibrium is one in which the state and jump variables evolve according to equations* (22) *and* (23)*, and asset prices are determined by the solution to the expressions* (25) *and* (26)*. In this case, the effective borrowing rate is given by equation* (21)*, where* $\hat{\mathbf{A}}$ *solves the fixed point problem*

$$\hat{\mathbf{A}} = \mathbf{e}_i + \left(\mathbf{\Gamma}^\top - \mathbf{M}\right) \int_0^T \left[\eta(\tau)\mathbf{A}(\tau) + \tilde{\eta}(\tau)\tilde{\mathbf{A}}(\tau)\right] \mathrm{d}\tau \,. \tag{27}$$

In a neighborhood of risk-neutrality ($a \approx 0$)*, the equilibrium exists and is (locally) unique.*

Note that the dynamics matrix of the state $\mathbf{\Gamma}$ depends on the effective borrowing rate coefficients $\hat{\mathbf{A}}$, which itself is a function of the risk-adjusted dynamics matrix $\mathbf{M}$. Thus, equilibrium is determined as a fixed point that produces asset price dynamics consistent with equilibrium dynamics of the macroeconomy and vice versa. In general, an affine equilibrium of this type may not exist, or there may be multiple solutions to this fixed point problem. However, when $a = 0$, the model reduces to a standard New Keynesian model. The result in Proposition 1 shows that this equilibrium persists and is locally unique as we depart from risk neutrality. We use this insight in our numerical continuation solution algorithm, described in Appendix D.

4.3 Localization and the Channels of Quantitative Easing

Our model makes precise the channels through which QE may have financial and macroeconomic effects. The immediate and direct effect of a QE purchase shock is a reduction in the amount of assets held by arbitrageurs through market clearing (a decline in $X_t^{(\tau)}$ or $\tilde{X}_t^{(\tau)}$ in equation (11), depending on the type of assets purchased during QE). If arbitrageurs are risk-neutral ($a = 0$), we recover a typical neutrality result: QE changes the arbitrageurs' portfolio, but not asset prices (and thus has no affect on real activity).

On the other hand, when arbitrageurs are risk-averse ($a > 0$), the change in arbitrageurs' portfolio implies changes in risk exposure, which is priced. We can broadly dis-

tinguish between three main sources of risk: *duration* or *short-rate* risk (from fluctuations in i_t); *payoff* risk (from fluctuations in d_t); and *habitat demand* risk (from fluctuations in demand factors β_t^k). By reducing arbitrageur exposure to these sources of risk, in equilibrium QE reduces the risk compensation arbitrageurs require, which manifests as a decline in expected returns. As a result, the household effective borrowing rate $\hat{r}_t$ falls, which through the intertemporal decisions of the household leads to lower savings and a jump in consumption, and from the pricing decisions of firms leads to an increase in inflation.

However, the precise mapping from a QE shock (or private demand shocks) to the repricing of risk depends qualitatively on the risk-bearing capacity of arbitrageurs. Our model helps clarify the intuition behind the empirical localization results found in Section 3. Intuitively, when arbitrageurs are nearly risk-neutral ($a \approx 0$), macroeconomic fundamentals affecting the path of the short rate (duration risk) are the dominant factor in determining of the term structure of interest rates. Hence, when arbitrageurs hold bonds, short-rate fluctuations are their main source of risk. Demand shocks that re-allocate bonds away from arbitrageurs reduce their exposure to short-rate risk, and hence decrease the compensation arbitrageurs require to hold bonds. Since all bonds are sensitive to short-rate risk, any such demand shock will push down yields of all bonds. Importantly, this mechanism is independent of the location (in maturity space) of the demand shock.

As arbitrageur risk aversion increases ($a > 0$), habitat demand shocks become more prominent as additional sources of risk. Arbitrageurs try to limit their exposure to these sources of risk, leading to less propagation from the location of the demand shock to other parts of the term structure. Arbitrageurs become less willing to integrate bond markets across maturities, and hence the response of the yield curve becomes more localized around the location (in maturity space) of a given demand shock.

We can now use the model to formalize the logic of the localization effects documented in Section 3. Take two demand factors: a "short" maturity factor β_t^s and "long" maturity factor β_t^ℓ with the same overall magnitude across maturities ($\int_0^T \theta^s(\tau)\,\mathrm{d}\tau = \int_0^T \theta^\ell(\tau)\,\mathrm{d}\tau$), but the short factor is more concentrated in bonds of short maturities relative to the long factor ($\exists \tau' : \ \theta^s(\tau) < \theta^\ell(\tau) \iff \tau > \tau'$). In this context, the localization hypothesis involves the differential responses of the entire yield curve to movements in the short and long demand factors. We formally state a version of the localization hypothesis:

$$\text{if } a \approx 0 \ : \frac{\partial y_t^{(\tau)} \Big/ \partial \beta_t^s}{\partial y_t^{(\tau^*)} \Big/ \partial \beta_t^s} \approx \frac{\partial y_t^{(\tau)} \Big/ \partial \beta_t^\ell}{\partial y_t^{(\tau^*)} \Big/ \partial \beta_t^\ell}, \tag{28}$$

$$\text{if } a \gg 0 \ : \tau > \tau^* \iff \frac{\partial y_t^{(\tau)} \Big/ \partial \beta_t^s}{\partial y_t^{(\tau^*)} \Big/ \partial \beta_t^s} < \frac{\partial y_t^{(\tau)} \Big/ \partial \beta_t^\ell}{\partial y_t^{(\tau^*)} \Big/ \partial \beta_t^\ell}, \tag{29}$$

where τ^* is an arbitrary "baseline" maturity (and unrelated to τ) and $\partial y_t^{(\tau)} \Big/ \partial \beta_t^k$ is the

response of τ-maturity yields to the demand shock $k \in \{s, \ell\}$. When arbitrageur risk-bearing capacity is high ($a \approx 0$), the relative response of the yield curve is the same for both factors. On the other hand, when risk-bearing capacity is low ($a \gg 0$), then long demand shocks have relatively larger effects on long-maturity yields and *vice versa* for short demand shocks. These expression make statements about the movements of the yield curve to short and long demand shocks, *relative* to movements in some fixed maturity τ^*.[14] Thus, the model maps to our specification (4), and the qualitative predictions of the model line up with our results. In the next section, we move from the qualitative to quantitative features of the model.

4.4 Model Calibration

To study the quantitative properties of the model, we need to pick function forms and assign values to parameters. First, we must choose functional forms regarding the effective borrowing weights and the habitat elasticity and demand functions. We assume similar exponential functions as in Vayanos and Vila (2021):

$$\text{habitat elasticity function:} \quad \alpha(\tau) = \alpha_0 \exp(-\alpha_1 \tau), \tag{30}$$

$$\text{habitat demand functions:} \quad \theta^k(\tau) = \theta_0 \tau (\theta_1^k)^2 \exp\left(-\theta_1^k \tau\right), \tag{31}$$

$$\text{effective borrowing weights:} \quad \eta(\tau) = \eta_0 \tau \eta_1^2 \exp(-\eta_1 \tau), \tag{32}$$

with analogous functional forms for $\tilde{\alpha}(\tau), \tilde{\theta}^k(\tau), \tilde{\eta}(\tau)$ (and $\tilde{\eta}_0 = 1 - \eta_0$ so that the weights sum to one). Equation (30) implies that the habitat investor elasticity with respect to (log) price is declining with maturity, and is single-peaked with respect to yields. Similarly, equations (31) and (32) imply that the demand factor and borrowing weights are single-peaked functions. In addition to improving numerical properties of the model, these exponential functional forms allow some flexibility in capturing key modeling features, such as demand shocks targeted in specific areas of maturity space, without significantly increasing the dimensionality of the problem. The parameters α_0, θ_0 govern the overall size of the habitat elasticity and demand functions, while $\alpha_1, \theta_1^k, \eta_1$ govern the shape as a function of maturity: a lower value of these parameters imply that more of the weight of these functions lies at longer maturities.

Next, to map the model to our empirical estimates in Section 3, we assume there are three demand factors, corresponding to short- and long-maturity Treasury auctions, and a risky asset. We denote these demand factors respectively by $\beta_t^s, \beta_t^\ell, \tilde{\beta}_t$. In order to impose more discipline on the model, we assume that the habitat demand shocks are identical, except for the shape of the demand factor functions $\theta^s(\tau)$, $\theta^\ell(\tau)$, and $\tilde{\theta}(\tau)$. Specifically, we set $\theta_1^s = 0.5$ and $\theta_1^\ell = 0.2$ to match our regression analysis in Section 3. These choices

[14] Appendix E discusses sufficient conditions and derives further localization predictions of the model.

imply that the short factor is concentrated in short maturities less than 5 years (and the mean of $\theta^s(\tau)$ is 4 years), while the long factor has more weight in intermediate and long maturities above 7 years (and the mean of $\theta^\ell(\tau)$ is 10 years). We set the parameter $\alpha_1 = 0.1$ such that the habitat elasticity with respect to yields ($\equiv \tau \cdot \alpha(\tau)$) is maximized for the 10-year maturity, since this is often taken as a key benchmark yield by investors. We further assume that the habitat demand and elasticity functions are identical for Treasuries and risky assets: $\alpha(\tau) = \tilde{\alpha}(\tau)$ and $\theta^s(\tau) = \tilde{\theta}(\tau)$. Finally, we assume that the dynamics parameters $\kappa_\beta, \phi_{i,\beta}, \sigma_\beta$ for each demand factor in equation (14) are identical.

An important input into our model is the effective borrowing rate weights $\eta(\tau), \tilde{\eta}(\tau)$. If household borrowing responds only to movements in the instantaneous short rate, then the macroeconomic dynamics of the model are identical to a textbook New Keynesian model. Thus, the effective borrowing weights can be thought of broadly as capturing the sensitivity of the macroeconomy to borrowing rates besides the federal funds rate. For our baseline calibration, we assume that the effective borrowing rate only depends on risky rate ($\eta_0 = 0$ and hence $\eta(\tau) = 0$ and only $\tilde{\eta}(\tau)$ is non-zero). Further, we set $\eta_1 = 2$, such that $\tilde{\eta}(\tau)$ peaks at 0.5 years and has an average value of 1 year. Our baseline choice puts a heavy weight on short and intermediate maturities because trading volume in debt markets is concentrated at shorter-term maturities. For instance, estimates from the Securities Industry and Financial Markets Association (SIFMA) show that over the last two decades, trading volume for Treasury securities with maturity of 3 years or less made up about 45% of total trading volume on average (and this ratio was relatively stable, with a minimum and maximum annual value of 41% and 52% respectively). Even within coupon-bearing securities (excluding T-bills), trading volume for Treasury notes with maturity of 3 years or less made up 33% of the trading volume in coupon-bearing Treasury notes and bonds. In comparison, trading volume for Treasury bonds with a maturity of greater than 11 years made up only 8% of the trading volume in coupon-bearing Treasury notes and bonds. Trading volume for corporate debt markets across maturities is less readily available. However, comparing trading volume in corporate bond markets and commercial paper markets (from SIFMA and the Federal Reserve, respectively) further suggests an outsized role of short-maturity rates. For instance, for 2020-2021 the trading volume in commercial paper markets was on average 4 times that of trading volume in corporate debt markets (for securities with maturity of at least one year). In short, our calibration implies that household borrowing is a function of risky borrowing rates (rather than the policy rate), largely concentrated at short and intermediate maturities. We explore the sensitivity of our results to alternative calibrations below.

Finally, in order to reduce the number of parameters needed to be calibrated, we simplify and assume that the central bank only responds to inflation ($\phi_x = 0$ in Equation (15)), and the payoff process only responds to the output gap ($\psi_\pi = 0$ in Equation (16)). We explore sensitivity to these assumptions below.

We jointly estimate the remaining parameters through a moment-matching exercise.[15] We target volatility and cross-correlations of Treasury and corporate yields $y_t^{(\tau)}, \tilde{y}_t^{(\tau)}$ as well as inflation and the output gap π_t, x_t.[16] Specifically, we target the following set of moments: i) the volatility of 1-year Treasury yields, 1-year corporate-Treasury spreads, the output gap, and inflation (in levels, 1-year changes, and 1-month changes); ii) the correlation of inflation and 1-year Treasury yields, the correlation of output gap and 1-year corporate-Treasury spreads, and the correlation of inflation and the output gap (in levels); iii) the volatility of the entire term structure of Treasury yields (1-year changes and 1-month changes); iv) the correlation of the entire term structure of Treasury yields and 1-year Treasury yields (1-year changes and 1-month changes); v) the localization regression coefficients from Figure 4. Besides the localization regression coefficients, the data are monthly. Inflation is defined as the year-over-year log change in the Personal Consumption Expenditure price index. The output gap is defined as the cyclical component of Industrial Production using an HP filter. Zero-coupon Treasury yields are from Gürkaynak et al. (2007). Zero-coupon corporate bond yields are from the Treasury's High Quality Market (HQM) Corporate Bond Yield Curve data. We utilize the HQM data because it is a long-running measure of the term structure of zero-coupon private borrowing rates which are important for macroeconomic activity; however, these rates reflect relatively safe corporate rates (rated A or above). Besides the localization coefficients, moments are computed starting in 1986 (when 30-year yields are available in Gürkaynak et al. (2007)).

By targeting the localization regression coefficients from Figure 4, our model is informed by our empirical localization results, which are well-suited to identify the preferred habitat parameters of our model. We allow only the habitat risk-adjusted parameters ($a \cdot \alpha_0$, $a \cdot \sigma_\beta \cdot \theta_0$, and $a \cdot \phi_{i,\beta}$) to be different in "crisis", which we target to match the "crisis" localization regression coefficients from Figure 4. All other parameters are assumed to be the same across "crisis" and "non-crisis" periods. Given that the number of observations in the "non-crisis" period is vastly larger than the number of observations in the "crisis" period, we first estimate all of the parameters from the "non-crisis" sample.

The fit of the model is summarized in Figure 6 (moments which depend on maturity) and Panel A of Table 5 (additional moments). Panel B of Table 5 reports our baseline

[15] Collecting unknown parameters into a vector $\boldsymbol{\rho}$, we estimate the model by choosing $\hat{\boldsymbol{\rho}}$ to minimize the weighted sum of squares: $L(\boldsymbol{\rho}) = \sum_{n=1}^{N} w_n(\hat{m}_n - m_n(\boldsymbol{\rho}))^2$, where $\{\hat{m}_n\}_{n=1}^N$ are empirical moments and $\{m_n(\boldsymbol{\rho})\}_{n=1}^N$ are model-implied counterparts as a function of $\hat{\boldsymbol{\rho}}$. The terms $\{w_n\}_{n=1}^N$ are weights placed on each target moment. These weights are set to $1/\mathcal{N}_T$ for moments that are a function of maturity, where $\mathcal{N}_T$ is the number of maturities, and 1 otherwise, so that our estimates are not overly influenced by maturity-based moments. Appendix D derives the model-implied moments.

[16] The habitat elasticity functions and demand functions enter the solution multiplicatively with risk aversion and demand shock volatility, so they are only identified up to a scaling factor. If shocks to quantity demand for Treasuries and risky assets can be measured directly across all periods (including outside auction times), one can separately identify these parameters. Because such measures are not available, the literature (e.g., Vayanos and Vila (2021)) typically chooses some normalization, such as $\sigma_\beta = 1$. In our auction data, the bid-to-cover ratio captures this information partially; this information is not necessary for our calibration strategy, but we utilize this information in our QE exercises.

parameter values. Our parameterized model not only picks up the qualitative features of the data but is also successful at matching the data quantitatively. Besides matching a wide variety of volatility and correlation moments across borrowing rates and aggregate variables, the model localization coefficients are close to the data localization coefficients in "crisis" and "non-crisis" periods (Panels A through D of Figure 6).[17] Nonetheless, given the challenges associated with estimating DSGE models, one should treat this parameterization with caution, and we run extensive sensitivity analysis in Section 4.8.

4.5 Response of the Yield Curve to QE1

To assess the contribution of preferred habitat theory to the observed reaction of yields to quantitative easing, we feed a "QE1" shock into our model, compute predicted responses of yields at different maturities, and compare predictions to actual changes in yields. We focus on QE1 because it was arguably the "cleanest" QE shock: there are a clear set of policy announcement events, and the observed response to these events are unlikely to be plagued by anticipation issues relative to later rounds of QE. But in order to accurately capture the relevant features of QE1, we do not treat QE as precisely equivalent to a habitat demand shock in our model. Besides the obvious fact that QE1 was significantly larger than typical private demand shocks, QE1 also occurred unexpectedly, and markets had different expectations about the dynamic properties QE1 compared to typical private demand shocks. Moreover, QE1 simultaneously purchased Treasuries and mortgage-backed securities (MBS), and the profile of maturities was different than in auctions. Thus, in our model QE1 is not a simple rescaling of the habitat demand shocks during auctions. We discuss in detail how we mimic the actual QE1 shock in the model.

Persistence: We assume that this shock was completely unexpected ("MIT shock"), and that afterwards markets expected purchases to be unwound slowly and deterministically:

$$\mathrm{d}\beta_t^{QE} = -\kappa_{QE}\beta_t^{QE}\,\mathrm{d}t\,. \tag{33}$$

We set the inertia parameter $\kappa_{QE} = 0.2$. This magnitude roughly implies that markets expected the Fed to unwind its purchases somewhat faster than holding to maturity (more precisely, the half-life of the purchases is roughly 3.5 years). Ex-post, the MBS holdings roughly followed this process until the reintroduction of MBS purchases during QE3. On

[17] We also examine localization across asset classes. Appendix Figure B26 plots various measures of risky borrowing rate localization in the model. We find evidence in the model of increased maturity localization in crisis periods: the relative effects of our short and long riskless demand shocks $\beta_t^{(s)}, \beta_t^{(\ell)}$ have larger differential effects on risky rates across maturities in the model calibrated to the crisis period compared to the non-crisis period (consistent with Panel B of Table 4). Additionally, we find that compared to the response of the riskless yield curve, the pass-through of riskless demand shocks to risky borrowing rates decreases in the crisis calibration (consistent with Panel A of Table 4). However, unlike our baseline Treasury localization results, the mapping from our empirical results to the model is not as precise, and so these results should be taken as qualitative.

the other hand, the Treasuries purchased as part of the Fed's QE programs were held on the balance sheet for a very long time, and holdings remained elevated well beyond that implied by our parameterization. However, this does not imply that markets expected this ex-ante.[18] We explore the sensitivity of this assumption below.

Composition: QE1 involved purchasing a total of roughly $1.25 trillion mortgage-backed securities (MBS) and $300 billion Treasuries, concentrated on intermediate and long-term maturity purchases (over 5 years).[19] Given this focus, we assume the same functional form as in equation (31) for the QE demand function $\theta^{QE}(\tau)$ and set $\theta_1^{QE} = 0.35$, such that the model-implied QE1 purchases are of relatively long-term maturity but in between the maturities of the preferred habitat "short" and "long" factor described above (hence the mean of $\theta^{QE}(\tau)$ is roughly 6 years). These purchases are split up such that roughly 80% of QE1 purchases are risky assets in the model, while the remaining 20% are of safe bonds, in order to match the fraction of actual MBS and Treasury purchases during QE1.

Size: Matching the overall magnitude of QE1 requires a few additional steps. As discussed above, our calibration strategy does not separately identify the size of preferred habitat demand shifts and arbitrageur risk aversion. This allows us to compute the response of yields to a "unit" demand shock $\partial y_t^{(\tau)} \Big/ \partial \beta_t^k$ for short- or long-term auctions. However, our moment-matching exercise thus far provides no information about the dollar magnitude of a "unit" demand shock in the model, since we only can estimate the product $a \cdot \sigma_\beta \cdot \theta_0$. Our model is meant to capture all the volatility of demand shocks, not just those which occur during auctions. But by utilizing additional information from Treasury auctions, we can pin down the dollar magnitude of demand shocks.

The results of our alternative localization specification in Figure 5 measures the response of the yield curve to a unit change in the bid-to-cover, across short-term and long-term auctions and in normal vs. crisis times. During the QE1 period, a unit change in the bid-to-cover ratio corresponded to about $30 billion for long-term auctions, while short-term auctions were 50% larger. Thus we estimate this quantity of long-term auctions (which we denote by $\Delta\beta_t^{(auc)}$) by targeting the results of our alternative localization regression; results are in Appendix Figure B25. Finally, since QE1 was roughly $1.5 trillion, we have that $\Delta\beta_t^{(QE)} \approx 50 \times \Delta\beta_t^{(auc)}$. To be clear, this does not imply the yield curve response to QE1 is hard-wired to be the same as a (scaled-up) auction demand shock. Both the maturity composition, asset composition, and the stochastic properties of the QE1 shock

[18] The actual purchases of MBS and Treasuries during QE1 were planned to take place over the 6 months following the March 2009 announcement. The FOMC statements during this time do not make any reference to selling these securities off, but state that the FOMC will "carefully monitor the size and composition of the Federal Reserve's balance sheet in light of evolving financial and economic developments." Eventually, the FOMC made clear their policy of Treasury reinvestment. For instance, in Chairman Bernanke's July 2010 report to Congress, he stated that "the proceeds from maturing Treasury securities are being reinvested in new issues of Treasury securities with similar maturities."

[19] In our analysis, we exclude purchases of agency debt (which accounts for a small fraction of the total purchases during QE1) given the complexities associated with the health of Freddie Mac and Fannie Mae.

in the model differ from the estimated demand factors. Thus, the yield response to the QE shock vs. the demand factors will not be proportional: $\partial y_t^{(\tau)} \Big/ \partial \beta_t^{(QE)} \not\propto \partial y_t^{(\tau)} \Big/ \partial \beta_t^k$.

Panel A of Figure 7 plots the cumulative change observed in the data and predicted in the model. We use the cumulative change in the yield curve reported in Krishnamurthy and Vissing-Jorgensen (2011) following the 5 major QE1 event dates as the empirical response of the yield curve to QE1. The remarkable consistency between the responses suggests that the actual market reaction to QE1 announcements aligns with the predictions of our model in response to observed shifts in private demand for Treasuries. This finding implies that the *net* effect of other channels of QE (e.g., inflation expectations, forward guidance, signaling) could be smaller than thought before.

4.6 Macroeconomic Effects of Quantitative Easing

While there is extensive research documenting the responses of financial markets to rounds of QE, little is known about how QE affected the broader economy because QE events are so infrequent. We cannot shed more light on this using regression analysis or similar tools, but we can use our calibrated model to quantify the macroeconomic effects of QE.

Panel B of Figure 7 shows that our baseline calibration implies that in terms of macroeconomic effects, on impact QE1 increased output by roughly 0.73 percentage points, and inflation by 0.36 percentage points. These effects then monotonically fall towards zero over the next few years. The cumulative effects are $\int_0^\infty E_0 [x_t] \, dt \approx 0.56$ percentage points and $\int_0^\infty E_0 [\pi_t] \, dt \approx 0.64$ percentage points. For comparison, these stimulative effects of QE1 are comparable to a rate cut of roughly 50-75 b.p. in the model, undertaken when financial markets are not distressed (and note that conventional monetary policy rate cuts have a faster degree of mean reversion in our model).[20]

4.7 Comparing Quantitative Easing and Tightening

To rein in inflation, central banks are beginning the implementation of quantitative tightening (shrinking their balance sheets through asset sales or a reduction in reinvestment of maturing assets). In our linearized model, QE and QT are mirror images of one another. Differences in the macroeconomic impacts of QE and QT can arise only from differences in the financial environment or in the design of the program. These differences are salient in

[20] In our model, there is little difference between (1) the Fed buying assets (and thus reducing assets available to the private market) and (2) the Treasury issuing less debt (and thus reducing assets available to the private market). Thus, while comparing our estimated effect of QE1 and actual data, one should keep in mind that the Treasury increased issuance of debt by approximately $2.7 trillion between December 2008 and October 2010 (the duration of QE1) relative to the pre-crisis trend. The average maturity of debt issued over this period was 6.4 years. If we assume that this shock to the supply of government debt is as persistent as QE1 (i.e., debt is held to maturity), then QE1 was roughly offsetting the increased supply of Treasuries and the combined macroeconomic effect of changes in government debt due to fiscal and monetary policies was a wash.

the context of the Fed's QT policy, which began in June 2022. First, the Fed is conducting QT in a relatively passive, back-loaded fashion: QT began with a monthly reduction of up to $30 billion in Treasuries and $17.5 billion in MBS, with both caps doubling after one quarter. The increasing pace of QT (which was announced in advance) differs from how the Fed conducted QE1. Second, the Fed's QT has focused more on a reduction of Treasury holdings than MBS, again in contrast with QE1. Third, the maturity composition of QT is shorter than QE1. Fourth, although the Fed has not specified the size of QT, this program is likely to be larger than the purchases undertaken during QE1. Finally, risk-bearing capacity of financial markets is higher now than during QE1.

While there is substantial uncertainty associated with how the Fed will conduct QT going forward, we attempt to parsimoniously capture these salient differences within the model. We assume that the QT sales shock β_t^{QT} is composed of 65% Treasuries and 35% MBS, mimicking the composition of the Fed's QT balance sheet reduction caps. In line with the maturity composition of the Fed's balance sheet, we also assume that $\theta_1^{QT} = 0.5$ so that the maturity of the assets sold during QT is shorter than the maturity of assets purchased during QE1 (and the mean of $\theta^{QT}(\tau)$ is 4 years; Appendix Figure B27 compares the change in the Fed's Treasury holdings by maturity following QE1 and QT). Capturing the passive, back-loaded approach of QT while maintaining the mean-reverting properties necessary in the model requires some additional changes. We posit that β_t^{QT} is a mixture of an "active" component ($\beta_t^{QT,A}$) and a "passive" component ($\beta_t^{QT,P}$) such that $\beta_t^{QT} = \lambda_{QT}\beta_t^{QT,A} + (1 - \lambda_{QT})\beta_t^{QT,P}$ where

$$\mathrm{d}\beta_t^{QT,A} = -\kappa_{QT,A}\beta_t^{QT,A}\,\mathrm{d}t\,, \tag{34}$$

$$\mathrm{d}\beta_t^{QT,P} = -\left(\gamma_{QT,A,P}\beta_t^{QT,A} + \kappa_{QT,P}\beta_t^{QT,P}\right)\mathrm{d}t\,. \tag{35}$$

Thus, QT sales initially are driven by the "active" component $\beta_t^{QT,A}$, which reverts back to steady state just as our QE1 shock (and we choose the same inertia $\kappa_{QT,A} = \kappa_{QE}$). However, the "passive" factor $\beta_t^{QT,P}$ then kicks in, capturing the slow increase in the magnitude of QT sales following the announcement of QT (taken into account by market participants). We choose $\gamma_{QT,A,P} = 1.75$, $\kappa_{QT,P} = 2.25$, and $\lambda_{QT} = 0.25$ so that the QT shock doubles in magnitude after one quarter, peaks after 1 year, and then reverts to half of the peak after roughly 4-5 years. The overall magnitude of the initial QT shock is chosen to be double that of QE1: $\Delta\beta_t^{QT,A} = 2 \times \Delta\beta_t^{QE}$. Thus, while QT sales are smaller than QE1 purchases at the beginning of the program, QT sales eventually surpass QE1 in magnitude, remaining elevated so that the cumulative size of QT is larger than QE1.[21]

[21] Appendix Figure B28 compares the dynamics of QE1 and QT shocks in the model. The Fed has been very clear about the initial timing of the Treasury and MBS caps (e.g., see the March 2022 FOMC statement). Going forward, the FOMC has stated that they intend to "maintain securities holdings in amounts needed to implement monetary policy efficiently and effectively in its ample reserves regime," although the FOMC has been reticent regarding what this means quantitatively. But if we take the size

Finally, we choose the risk-adjusted parameters to be equal to 75% of our baseline QE1 calibration. On the one hand, despite increased volatility during COVID-19, many broad measures of financial distress (e.g., the Chicago Fed National Financial Conditions Index) remain significantly below the peaks observed during the Great Recession. On the other hand, some measures of intermediary health like He et al. (2016) which are more closely related to the risk-bearing capacity of arbitrageurs show signs of deterioration in the last year (see Appendix Figure B12). Thus, we split the difference for our baseline QT experiment. However, we view this as a lower bound on risk-bearing capacity (and hence the macroeconomic reactions to QT should be viewed as upper bounds).

Panel C of Figure 7 plots the impulse response of inflation and the output gap to our QT shock. Despite the fact that QT is larger in magnitude than QE1, the macroeconomic effects on impact are somewhat smaller. For instance, the initial response of output and inflation is a decline of between 0.25 and 0.3 percentage points. However, the persistent passive component also generates a longer-lasting effect of QT compared to QE1, and implies that the response of the output gap to QT is hump-shaped rather than monotonic.

The design of QT is in part responsible for the relatively muted response of output and inflation. However, the key difference between QE1 and QT is the risk-bearing capacity of financial markets. As discussed above, when risk-bearing capacity increases ($a \to 0$), our model collapses to a textbook New Keynesian model and recovers the standard neutrality results regarding asset purchases or sales.

In summary, our model suggests that QT will not induce substantial downward pressure on inflation. The differences in the model-implied effects of QE and QT motivate our subsequent sensitivity analysis for the design of these programs.

4.8 Alternative Designs of Quantitative Easing

We now explore how variations in structural parameters and in the implementation of QE influence the ability to move output and inflation. Because in our model output and inflation co-move strongly in response to QE shocks, we focus on the sensitivity analysis for output responses (Figure 8) and relegate results for inflation to the appendix (Appendix Figure B29). For all sensitivity analyses, we study how the macroeconomic effects vary in our baseline calibration (solid line), and also compare with counterfactual "risky-only" and "Treasury-only" QE policies (dashed and dotted lines, respectively), which are identical to our QE1 calibration but only purchase risky or safe assets.

Risk Appetite: As we discussed above, the power of QE to affect yields at target maturities is high in crisis times and weak in non-crisis times. Consistent with this

of the balance sheet pre-COVID as this target level, then the overall size of QT relative to GDP would be roughly twice the size of QE1 (see also Ennis and Kirk (2022)). Hence, these calibration choices represent our best translation into a mean-reverting shock in the model. With the linearity of the model, responses scale one-for-one with the size of shocks. All other sensitivity analyses are conducted in Section 4.8.

insight, Panel A of Figure 8 shows that the transmission is highly sensitive to the risk aversion of arbitrageurs. Given that our baseline calibration for "crisis" corresponds to the Great Recession, one has to have a financial crisis of truly unprecedented proportions to materially increase the power of QE-based tools. As a result, the balance of risks appears to be somewhat one-sided: it is difficult to raise the power of QE beyond what was achieved during the Great Recession but it is relatively easy to reduce the power as soon as financial panics calm down. This logic suggests that QE is less effective as a "conventional" tool to the extent that "conventional" entails well-functioning financial markets (and was reflected in our analysis of QT). This finding is true regardless of the design of QE purchases: the counterfactual "risky-only" and "Treasury-only" QE policies have the same pattern as a function of risk aversion.

Risky Asset Uncertainty: We next discuss how the transmission of QE depends on the riskiness of risky assets. Panel B of Figure 8 plots the long-run effect of QE on output as we increase σ_d from the baseline calibration. We find that QE policy has larger macroeconomic effects as risky assets become riskier. The intuition for this can be seen by examining the counterfactual "risky-only" and "Treasury-only" QE policies. Unlike the previous sensitivity exercise, these policies act differently as a function of risky asset uncertainty. As uncertainty increases, the effect of risky asset purchases on output increases. However, the impact of Treasury purchases decreases. Intuitively, when risky asset payoff uncertainty is very low, Treasuries and risky assets are good substitutes (in the limit of no payoff risk, these two assets become perfect substitutes). In this case, "risky-only" and "Treasury-only" QE policies have similar effects on household borrowing rates, and therefore macroeconomic transmission is similar because the payoff risk extraction channel discussed in Section 4.3 disappears. However, as risky asset uncertainty increases, Treasuries and risky assets become less substitutable. Now, "risky-only" QE policies are highly effective at moving household borrowing rates (and therefore have larger macroeconomic effects), while "Treasury-only" QE policies have smaller effects.

Unwinding: One practical question for policymakers is how long central banks should hold assets accumulated during QE rounds (e.g., Sims and Wu (2020) and Karadi and Nakov (2020)). To this end, we vary κ_{QE} in equation (33), which governs the unwinding process. Panel C of Figure 8 shows that the effects on output nearly double when comparing very fast and very slow unwinding (the x-axis reports results from $\kappa_{QE} \in (0.05, 1.0)$, implying a half-life of about 14 years to 3 quarters, respectively). Intuitively, if the Fed makes large purchases of assets but then quickly sells those securities, we should not observe large and long-term macroeconomic effects. Rather, it is the cumulative size of QE over time that matters. Hence, policymakers should be clear about how long the central bank expects to hold the securities on its balance sheet. Providing a type of "forward guidance" regarding the expected path of purchases can potentially increase the immediate effectiveness of these policies.

Asset Maturity: The maturity of the assets purchased during QE is one of the key decisions policymakers must make. The x-axis of Panel D of Figure 8 is the maturity of the assets purchased (where a higher value implies shorter maturity). The baseline QE policy has larger macroeconomic effects as the maturity of the purchased assets increases. Long-maturity assets are riskier in terms of duration risk, and hence purchasing these assets amplifies the duration risk extraction channel. However, we again see a dichotomy between the counterfactual "Treasury-only" and "risky-only" QE policies. This is because the payoff risk extraction channel is still active when the duration of the risky asset purchases is short. Hence, while Treasury purchases become ineffective as the maturity becomes very short-term, risky asset purchases still have sizable macroeconomic effects.

Effective Borrowing Weights: Panels E and F explore the robustness of our assumptions regarding the effective borrowing weight function $\eta(\tau)$, which as discussed above is an important input into our model. Panel E of Figure 8 plots the long-run macroeconomic effects of QE as the effective borrowing rate weights become more concentrated at short-term rates (larger values of the x-axis) or at long-term rates (smaller values of the x-axis). As expected, if the macroeconomy is more sensitive to longer-term borrowing rates, QE has a larger impact on output. Because the short end of the term structure is pinned down by the Fed's policy rate, QE will have larger effects on intermediate Treasury yields. Moreover, given the localization results discussed above, longer maturity purchases will have larger effects on long-term rates when the risk-bearing capacity of financial markets is relatively low.

Panel F varies the weights placed on safe and risky borrowing rates. The baseline effective borrowing rate depends entirely on risky rates (corresponding to the x-axis value of 0). We compare with alternative weights placed on safe borrowing rates, up until the effective rate is only a function of safe rates (x-axis value of 1). As expected, when household borrowing is more sensitive to safe rates, the counterfactual "Treasury-only" QE program becomes more powerful. However, the magnitude of the differences are small.

Interaction with Conventional Policy: Next we turn to how the macroeconomic effects of QE interact with conventional policy via the standard Taylor rule. Panels G, H, and F vary the Taylor coefficients on inflation ϕ_π, the output gap ϕ_x, and inertia κ_i (respectively). In general, if markets expect the Fed to react to the macroeconomy more quickly and aggressively with the policy rate (corresponding to larger values on the x-axis), then the macroeconomic effects of QE are attenuated. However, for a wide range of coefficient values, the macroeconomic effects of QE are quantitatively similar.

QE Uncertainty: Finally, we assumed that QE1 was completely unanticipated by markets. This may be an accurate representation of the first round of quantitative easing, but over a decade later it is likely that markets now consider the possibility of future QE shocks. To model the recurrent nature of QE, we modify equation (33) and instead

assume that QE shocks are described by

$$\mathrm{d}\beta_t^{QE} = -\kappa_{QE}\beta_t^{QE}\,\mathrm{d}t + \sigma_{QE}\,\mathrm{d}B_{QE,t}\,. \tag{36}$$

Hence, whenever $\sigma_{QE} > 0$, QE shocks themselves may lead to additional macroeconomic volatility. Further, risk-averse arbitrageurs must hedge against QE risk, in addition to fundamental sources of risk in the economy.

Panel J of Figure 8 explores the change in long-run macroeconomic volatility as a function of the volatility of QE shocks. $\sigma_{QE} = 0$ is our baseline estimate (marked by the vertical dotted line); the x-axis is the volatility of QE shocks relative to the volatility of habitat demand shocks $\frac{\sigma_{QE}}{\sigma_\beta}$; the y-axis is the the long-run (unconditional) variance of output $Var[x_t]$. The results provide an important note of caution for central bankers: increased uncertainty regarding QE leads to increased macroeconomic volatility. In other words, although policymakers may desire the added flexibility of discretionary QE tools, the downside is that this increases the uncertainty surrounding these policy tools. By communicating clearly the expected path of QE purchases, policymakers should be able to reduce market uncertainty and prevent volatility spillovers from QE into the real economy. This provides additional support for the use of QE rules or forward guidance regarding asset purchases in the spirit of Greenwood et al. (2016).[22]

5 Concluding Remarks

Quantitative easing programs during the Great Recession were a massive policy experiment. The redeployment of QE in response to the COVID-19 crisis and the current reversal through quantitative tightening is a testament to policymakers' belief in their effectiveness. But the precise channels through which QE policies work are still not clear, and so this paper seeks to unbundle QE by isolating one specific channel. Our focus is "preferred habitat," which hypothesizes that there are investor clienteles with specific preferences for bonds within a given maturity segment. Utilizing Treasury auctions, we identify shifts in *private* demand for Treasuries that mimic QE, but are independent of all other plausible channels of QE. We develop and confirm a key "localization" test of preferred habitat theory: when risk-bearing capacity is low, purchase or sale shocks of specific bonds in a given maturity segment have larger effects on bonds within this range. We then develop a general equilibrium preferred habitat model, using our demand shock results to discipline the model calibration. Our quantitative results are consistent with the view that QE programs worked largely through preferred habitat forces, and that QE is a useful (but modest) policy tool for macroeconomic stabilization during crises. However, the effects of QE weaken if financial markets are well-functioning, the holding

[22] Additional robustness exercises are reported in Appendix Figures B30 and B31.

period of assets purchases is short, or the program focuses on safer, shorter-duration assets. Further, uncertainty about future QE/QT rounds can lead to excess macroeconomic volatility. Given the design of quantitative tightening as well as the prevailing financial environment, our model predicts rather small effects on output and inflation.

Our results suggest several promising directions for future work. First, the demand shocks we construct are a potentially important input to evaluate a host of quantity-based asset pricing theories. Second, while we have focused on preferred habitat in the market for Treasuries and other debt securities, this channel may be relevant for other arenas, too. For example, we document suggestive evidence supporting state-dependence of demand shock spillovers to mortgage-backed securities, and possibly to equities. In a similar spirit, Greenwood et al. (2020) and Gourinchas et al. (2022) explore preferred habitat effects in currency markets and international bond markets. Finally, preferred habitat is one of many financial frictions relevant for economic agents. It will be interesting to explore how preferred habitat interacts with debt vs. equity structure of firms, liquidity constraints, and banking frictions. Apart from clarifying the distributional aspects of QE-like policies, this extension could also shed more light on how preferred habitat can affect investment.

Appendix A Proofs

Proof of Lemma 1. Collect all state and jump variables in a vector $\mathbf{Y}_t = \begin{bmatrix}\mathbf{y}_t^\top & \mathbf{x}_t^\top\end{bmatrix}^\top$. The interest rate process (15), risky asset payoff process (16), and habitat demand factor processes (14) are all affine functions of $\mathbf{Y}_t$. Moreover, the (linearized) Phillips curve (17) and IS equation (18) are also affine functions of $\mathbf{Y}_t$, since from equation (21) $\hat{r}_t = \hat{\mathbf{A}}^\top \mathbf{y}_t + \hat{C}$ is affine in the state variables. Aggregate dynamics can thus be written

$$\mathrm{d}\mathbf{Y}_t = -\boldsymbol{\Upsilon}\left(\mathbf{Y}_t - \bar{\mathbf{Y}}\right)\mathrm{d}t + \mathbf{S}\,\mathrm{d}\mathbf{B}_t\,. \tag{A1}$$

Note that $\boldsymbol{\Upsilon}$ depends on $\hat{\mathbf{A}}$, but which we currently take as given. Then the rational expectations equilibrium is found immediately from Buiter (1984). Partition the eigenvalues and eigenvectors as follows:

$$\boldsymbol{\Upsilon} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^{-1}, \quad \boldsymbol{\Lambda} = \begin{bmatrix}\boldsymbol{\Lambda}_1 & \mathbf{0}\\ \mathbf{0} & \boldsymbol{\Lambda}_2\end{bmatrix}, \quad \mathbf{Q} = \begin{bmatrix}\mathbf{Q}_{11} & \mathbf{Q}_{12}\\ \mathbf{Q}_{21} & \mathbf{Q}_{22}\end{bmatrix},$$

where the partitions correspond to the state $\mathbf{y}_t$ and jump $\mathbf{x}_t$ variables. If the number of "stable" eigenvalues (non-negative real parts) equals the number of state variables, then the rational expectations equilibrium dynamics are given by (22) and (23), where

$$\boldsymbol{\Gamma} = \mathbf{Q}_{11}\boldsymbol{\Lambda}_1\mathbf{Q}_{11}^{-1}, \quad \boldsymbol{\Omega} = \mathbf{Q}_{21}\mathbf{Q}_{11}^{-1}. \tag{A2}$$

□

Proof of Lemma 2. Since asset prices are given by (24) and the state evolves according to (22), Ito's Lemma implies that $\frac{\mathrm{d}P_t^{(\tau)}}{P_t^{(\tau)}} = \mu_t^{(\tau)}\,\mathrm{d}t + \boldsymbol{\sigma}^{(\tau)}\,\mathrm{d}\mathbf{B}_t$, with $\boldsymbol{\sigma}^{(\tau)} = -\mathbf{A}(\tau)^\top\boldsymbol{\sigma}$ and

$$\mu_t^{(\tau)} = \mathbf{A}'(\tau)^\top\mathbf{y}_t + C'(\tau) + \mathbf{A}(\tau)^\top\boldsymbol{\Gamma}\left(\mathbf{y}_t - \bar{\mathbf{y}}\right) + \frac{1}{2}\mathbf{A}(\tau)^\top\boldsymbol{\Sigma}\mathbf{A}(\tau), \tag{A3}$$

where $\boldsymbol{\Sigma} \equiv \boldsymbol{\sigma}\boldsymbol{\sigma}^\top$ and with analogous expressions for $\tilde{P}_t^{(\tau)}$. Differentiating the arbitrageur budget constraint (10) with respect to holdings $X_t^{(\tau)}$ gives the optimality conditions

$$\mu_t^{(\tau)} - i_t = a\left[\int_0^T X_t^{(\tau)}\mathbf{A}(\tau) + \tilde{X}_t^{(\tau)}\tilde{\mathbf{A}}(\tau)\,\mathrm{d}\tau\right]^\top\boldsymbol{\Sigma}\mathbf{A}(\tau),$$

again with analogous conditions with respect to $\tilde{X}_t^{(\tau)}$.

Substituting equation (24) into the habitat demand equations (13), we can write $Z_t^{(\tau)} = \left[\alpha(\tau)\mathbf{A}(\tau) - \boldsymbol{\Theta}(\tau)\right]^\top\mathbf{y}_t$ and $\tilde{Z}_t^{(\tau)} = \left[\tilde{\alpha}(\tau)\tilde{\mathbf{A}}(\tau) - \tilde{\boldsymbol{\Theta}}(\tau)\right]^\top\mathbf{y}_t$, where

$$\boldsymbol{\Theta}(\tau) = \begin{bmatrix}\ldots & \theta^k(\tau) & \ldots\end{bmatrix}^\top, \quad \tilde{\boldsymbol{\Theta}}(\tau) = \begin{bmatrix}\ldots & \tilde{\theta}^k(\tau) & \ldots\end{bmatrix}^\top. \tag{A4}$$

Then substitute market clearing conditions $X_t^{(\tau)} = -Z_t^{(\tau)}, \tilde{X}_t^{(\tau)} = -\tilde{Z}_t^{(\tau)}$ into the optimality conditions and collect terms that are linear in the state $\mathbf{y}_t$ to get:

$$\mathbf{A}'(\tau) + \mathbf{M}\mathbf{A}(\tau) - \mathbf{e}_i = \mathbf{0}, \quad \tilde{\mathbf{A}}'(\tau) + \mathbf{M}\tilde{\mathbf{A}}(\tau) - \mathbf{e}_i = \mathbf{0}, \tag{A5}$$

where $\mathbf{M}$ is given by equation (26). Taking $\mathbf{M}$ as given, this is a linear system of differential equations. To derive initial conditions, note that at maturity, the riskless bonds pay \$1 while the risky asset pays D_t, so the $\tau = 0$ prices are given by $P_t^{(0)} = 1, \tilde{P}_t^{(0)} = D_t$. Hence, we have $\mathbf{A}(0) = \mathbf{0}$, $\tilde{\mathbf{A}}(0) = -\mathbf{e}_d$. Then assuming $\mathbf{M}$ is diagonalizable and invertible, the solution is given by equation (25). □

Proof of Proposition 1. In an affine equilibrium where asset prices are given by equation (24), we have that $\hat{r}_t = \int_0^T\left(\eta(\tau)\mu_t^{(\tau)} + \tilde{\eta}(\tau)\tilde{\mu}_t^{(\tau)}\right)\mathrm{d}\tau$. Substituting equations (A3) and (A5) into this expression and collecting terms which are linear in the state $\mathbf{y}_t$ gives equation (27). Equilibrium is the solution of the fixed point problem implicitly defined by equations (26) and (27). Rewrite these conditions in the following function:

$$f(\hat{\mathbf{A}}; \mathbf{M}; a) = \begin{bmatrix}\mathbf{e}_i + \left(\boldsymbol{\Gamma}(\hat{\mathbf{A}})^\top - \mathbf{M}\right)\boldsymbol{\nu}(\mathbf{M}) - \hat{\mathbf{A}} \\ \mathrm{vec}\left\{\boldsymbol{\Gamma}(\hat{\mathbf{A}})^\top - a\cdot\boldsymbol{\Lambda}(\mathbf{M}) - \mathbf{M}\right\}\end{bmatrix}, \tag{A6}$$

where $\boldsymbol{\Lambda}(\mathbf{M})$ and $\boldsymbol{\nu}(\mathbf{M})$ are the integral terms from equations (26) and (27). In both cases, dependence on $\mathbf{M}$ comes through the affine coefficients $\mathbf{A}(\tau), \tilde{\mathbf{A}}(\tau)$. We have also

made explicit the dependence of $\mathbf{\Gamma}$ on $\hat{\mathbf{A}}$, which can be seen in the proof of Lemma 1. If $J \equiv \dim \mathbf{y}_t$, then $\dim \mathbf{M} = J \times J$ and $\dim \hat{\mathbf{A}} = J$ and the function $f : \mathbb{R}^{J(J+1)+1} \to \mathbb{R}^{J(J+1)}$. For any value of a, equilibrium is defined by $f(\hat{\mathbf{A}}; \mathbf{M}; a) = \mathbf{0}$.

We now analyze the solution in a neighborhood around $a = 0$. For $a = 0$, clearly $\hat{\mathbf{A}} = \mathbf{e}_i$ and $\mathbf{M} = \mathbf{\Gamma}(\mathbf{e}_i)^\top$. The partial derivatives evaluated at this point are given by:

$$\frac{\partial f}{\partial \hat{\mathbf{A}}_j} = \begin{bmatrix} \left[\partial \mathbf{\Gamma}^\top \Big/ \partial \hat{\mathbf{A}}_j\right] \boldsymbol{\nu}(\mathbf{M}) - \mathbf{e}_j \\ \operatorname{vec}\left[\partial \mathbf{\Gamma}^\top \Big/ \partial \hat{\mathbf{A}}_j\right] \end{bmatrix}, \qquad \frac{\partial f}{\partial \mathbf{M}_{kl}} = \begin{bmatrix} \mathbf{e}_k \mathbf{e}_l^\top \boldsymbol{\nu}(\mathbf{M}) \\ -\operatorname{vec} \mathbf{e}_k \mathbf{e}_l^\top \end{bmatrix},$$

where $\mathbf{e}_j, \mathbf{e}_k, \mathbf{e}_l$ are standard normal basis vectors. The matrix $\partial \mathbf{\Gamma} \Big/ \partial \hat{\mathbf{A}}_j$ is the derivative of the state dynamics matrix $\mathbf{\Gamma}$ with respect to the j-element of $\hat{\mathbf{A}}$; because this depends on derivatives of the eigendecomposion defined in the proof of Lemma 1, even in the case of $a = 0$ this is a complicated expression. Nevertheless, from this we can show that the Jacobian of f with respect to $\hat{\mathbf{A}}, \mathbf{M}$ evaluated at the $a = 0$ solution has full rank. In fact, writing this Jacobian in block form, we have

$$Df \equiv \begin{bmatrix} \mathbf{D}_{11} & \mathbf{D}_{12} \\ \mathbf{D}_{21} & -\mathbf{I}_{J^2} \end{bmatrix}, \tag{A7}$$

and $\mathbf{D}_{12} = \begin{bmatrix} \mathbf{I}_J \cdot \nu_1 & \dots & \mathbf{I}_J \cdot \nu_J \end{bmatrix}$, where ν_j is the j-element of $\boldsymbol{\nu}(\mathbf{M})$. Because the elementary row operations which transform $\mathbf{D}_{12}$ into the zero matrix simultaneously transform $\mathbf{D}_{11}$ into $-\mathbf{I}_J$, $\det Df = 1$ and the result follows from the implicit function theorem. □

References

Beetsma, R., Giuliodori, M., de Jong, F., and Widijanto, D. (2016). Price effects of sovereign debt auctions in the euro-zone: The role of the crisis. *Journal of Financial Intermediation*, 25(C):30–53.

Beetsma, R., Giuliodori, M., Hanson, J., and de Jong, F. (2018). Bid-to-cover and yield changes around public debt auctions in the euro area. *Journal of Banking & Finance*, 87(C):118–134.

Bernanke, B. S. and Kuttner, K. N. (2005). What Explains the Stock Market's Reaction to Federal Reserve Policy? *Journal of Finance*, 60(3):1221–1257.

Bhattarai, S. and Neely, C. J. (2020). An analysis of the literature on international unconventional monetary policy. *FRB St. Louis Working Paper*, (2016-21C).

Bretscher, L., Schmid, L., Sen, I., and Sharma, V. (2022). Institutional Corporate Bond Pricing. Swiss Finance Institute Research Paper 21-07, Swiss Finance Institute.

Buiter, W. H. (1984). Saddlepoint Problems in Continuous Time Rational Expectations Models: A General Method and Some Macroeconomic Examples. *Econometrica*, 52(3):665–680.

Cahill, M. E., D'Amico, S., Li, C., and Sears, J. S. (2013). Duration risk versus local

supply channel in Treasury yields: evidence from the Federal Reserve's asset purchase announcements. FEDS 2013-35, Board of Governors.

Cammack, E. B. (1991). Evidence on Bidding Strategies and the Information in Treasury Bill Auctions. *Journal of Political Economy*, 99(1):100–130.

Campbell, J. R., Evans, C. L., Fisher, J. D., and Justiniano, A. (2012). Macroeconomic Effects of Federal Reserve Forward Guidance. *Brookings Papers on Economic Activity*, 43(1):1–80.

Carlstrom, C. T., Fuerst, T. S., and Paustian, M. (2017). Targeting Long Rates in a Model with Segmented Markets. *American Economic Journal: Macroeconomics*, 9(1):205–242.

Chen, H., Cúrdia, V., and Ferrero, A. (2012). The Macroeconomic Effects of Large-scale Asset Purchase Programmes. *Economic Journal*, 122(564):289–315.

Chodorow-Reich, G. (2014). Effects of Unconventional Monetary Policy on Financial Institutions. *Brookings Papers on Economic Activity*, 45(1):155–227.

Cui, W. and Sterk, V. (2021). Quantitative easing with heterogeneous agents. *Journal of Monetary Economics*, 123(C):68–90.

Cúrdia, V. and Woodford, M. (2011). The central-bank balance sheet as an instrument of monetarypolicy. *Journal of Monetary Economics*, 58(1):54–79.

D'Amico, S. and King, T. B. (2013). Flow and stock effects of large-scale treasury purchases: Evidence on the importance of local supply. *Journal of Financial Economics*, 108(2):425–448.

Di Maggio, M., Kermani, A., and Palmer, C. J. (2020). How Quantitative Easing Works: Evidence on the Refinancing Channel. *Review of Economic Studies*, 87(3):1498–1528.

Ennis, H. and Kirk, K. (2022). Projecting the Evolution of the Fed's Balance Sheet. 22-15, Federal Reserve Bank of Richmond Economic Brief.

Fieldhouse, A. J., Mertens, K., and Ravn, M. O. (2018). The Macroeconomic Effects of Government Asset Purchases: Evidence from Postwar U.S. Housing Credit Policy. *Quarterly Journal of Economics*, 133(3):1503–1560.

Fleming, M. J. (2007). Who buys treasury securities at auction? *Federal Reserve Bank of New York Current Issues in Economics and Finance*, 13(1).

Fleming, M. J. and Liu, W. (2016). Intraday pricing and liquidity effects of us treasury auctions.

Forest, J. J. (2018). The Effect of Treasury Auction Results on Interest Rates: The 1990s Experience. Technical report, University of Massachusetts - Amherst.

Gali, J. (2015). *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework and its Applications*. Princeton University Press.

Garbade, K. D. and Ingber, J. F. (2005). The Treasury auction process: objectives, structure, and recent acquisitions. *Current Issues in Economics and Finance*, 11(Feb).

Gertler, M. and Karadi, P. (2011). A model of unconventional monetary policy. *Journal of Monetary Economics*, 58(1):17 – 34. Carnegie-Rochester Conference Series on Public Policy: The Future of Central Banking April 16-17, 2010.

Gertler, M. and Karadi, P. (2013). QE 1 vs. 2 vs. 3. . . : A Framework for Analyzing Large-Scale Asset Purchases as a Monetary Policy Tool. *International Journal of Central Banking*, 9(1):5–53.

Gomes, F., Haliassos, M., and Ramadorai, T. (2021). Household Finance. *Journal of Economic Literature*, 59(3):919–1000.

Gorodnichenko, Y. and Ray, W. (2017). The Effects of Quantitative Easing: Taking a Cue from Treasury Auctions. NBER Working Paper 24122.

Gourinchas, P.-O., Ray, W., and Vayanos, D. (2022). A Preferred-Habitat Model of Term Premia, Exchange Rates, and Monetary Policy Spillovers. NBER WP 29875.

Greenlaw, D., Hamilton, J. D., Harris, E., and West, K. D. (2018). A Skeptical View of the Impact of the Feds Balance Sheet. NBER Working Paper 24687.

Greenwood, R., Hanson, S., Stein, J., and Sunderam, A. (2020). A Quantity Driven Theory of Risk Premiums and Exchange Rates. NBER Working paper 27615.

Greenwood, R., Hanson, S. G., and Vayanos, D. (2016). Forward Guidance in the Yield Curve: Short Rates versus Bond Supply. In Albagli, E., Saravia, D., and Woodford, M., editors, *Monetary Policy through Asset Markets: Lessons from Unconventional Measures and Implications for an Integrated World*. Central Bank of Chile.

Greenwood, R. and Vayanos, D. (2014). Bond supply and excess bond returns. *Review of Financial Studies*, 27(3):663–713.

Gürkaynak, R. S., Sack, B., and Wright, J. H. (2007). The U.S. Treasury yield curve: 1961 to the present. *Journal of Monetary Economics*, 54(8):2291 – 2304.

Hamilton, J. D. and Wu, J. C. (2012). The Effectiveness of Alternative Monetary Policy Tools in a Zero Lower Bound Environment. *Journal of Money, Credit and Banking*, 44:3–46.

He, Z., Kelly, B., and Manela, A. (2016). Intermediary asset pricing: New evidence from many asset classes. Working Paper 21920.

He, Z. and Krishnamurthy, A. (2013). Intermediary Asset Pricing. *American Economic Review*, 103(2):732–770.

Ippolito, F., Ozdagli, A. K., and Perez-Orive, A. (2018). The transmission of monetary policy through bank lending: The floating rate channel. *Journal of Monetary Economics*, 95(C):49–71.

Kaminska, I. and Zinna, G. (2020). Official demand for u.s. debt: Implications for u.s. real rates. *Journal of Money, Credit and Banking*, 52(2-3):323–364.

Karadi, P. and Nakov, A. (2020). Effectiveness and addictiveness of quantitative easing. *Journal of Monetary Economics*.

King, T. (2019a). Duration Effects in Macro-Finance Models of the Term Structure. Technical report.

King, T. B. (2019b). Expectation and duration at the effective lower bound. *Journal of Financial Economics*, 134(3):736–760.

Koijen, R. S., Koulischer, F., Nguyen, B., and Yogo, M. (2021). Inspecting the mechanism of quantitative easing in the euro area. *Journal of Financial Economics*, 140(1):1–20.

Koijen, R. S. J. and Yogo, M. (2019). A Demand System Approach to Asset Pricing. *Journal of Political Economy*, 127(4):1475–1515.

Koijen, R. S. J. and Yogo, M. (2022). Understanding the Ownership Structure of Corporate Bonds. NBER Working Paper 29679.

Krishnamurthy, A. and Vissing-Jorgensen, A. (2011). The effects of quantitative easing on interest rates: Channels and implications for policy. *Brookings Papers on Economic Activity*, (2):215–265.

Krishnamurthy, A. and Vissing-Jorgensen, A. (2012). The aggregate demand for Treasury debt. *Journal of Political Economy*, 120(2):233–267.

Kuttner, K. N. (2001). Monetary policy surprises and interest rates: Evidence from the Fed funds futures market. *Journal of Monetary Economics*, 47(3):523–544.

Kyle, A. S. and Xiong, W. (2001). Contagion as a Wealth Effect. *Journal of Finance*, 56(4):1401–1440.

Li, C. and Wei, M. (2013). Term Structure Modeling with Supply Factors and the Federal Reserve's Large-Scale Asset Purchase Progarms. *International Journal of Central Banking*, 9(1):3–39.

Lou, D., Yan, H., and Zhang, J. (2013). Anticipated and Repeated Shocks in Liquid Markets. *Review of Financial Studies*, 26(8):1891–1912.

Martin, C. and Milas, C. (2012). Quantitative easing: a sceptical survey. *Oxford Review of Economic Policy*, 28(4):750–764.

Modigliani, F. and Sutch, R. (1966). Innovations in interest rate policy. *The American Economic Review*, 56(1/2):178–197.

Peng, C. and Wang, C. (2022). Factor Demand and Factor Returns. Technical report.

Ray, W. (2019). Monetary policy and the limits to arbitrage: Insights from a new keynesian preferred habitat model. Working paper.

Rodnyansky, A. and Darmouni, O. M. (2017). The Effects of Quantitative Easing on Bank Lending Behavior. *Review of Financial Studies*, 30(11):3858–3887.

Romer, C. D. and Romer, D. H. (2017). New evidence on the aftermath of financial crises in advanced countries. *American Economic Review*, 107(10):3072–3118.

Sims, E. and Wu, J. C. (2020). Evaluating central banks tool kit: Past, present, and future. *Journal of Monetary Economics*.

Vayanos, D. and Vila, J. (2021). A Preferred-Habitat Model of the Term Structure of Interest Rates. *Econometrica*, 89(1):77–112.

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.

Zinman, J. (2015). Household Debt: Facts, Puzzles, Theories, and Policies. *Annual Review of Economics*, 7(1):251–276.

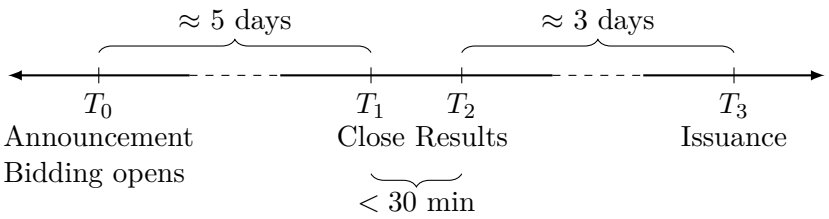

Figure 1: Auction Timing

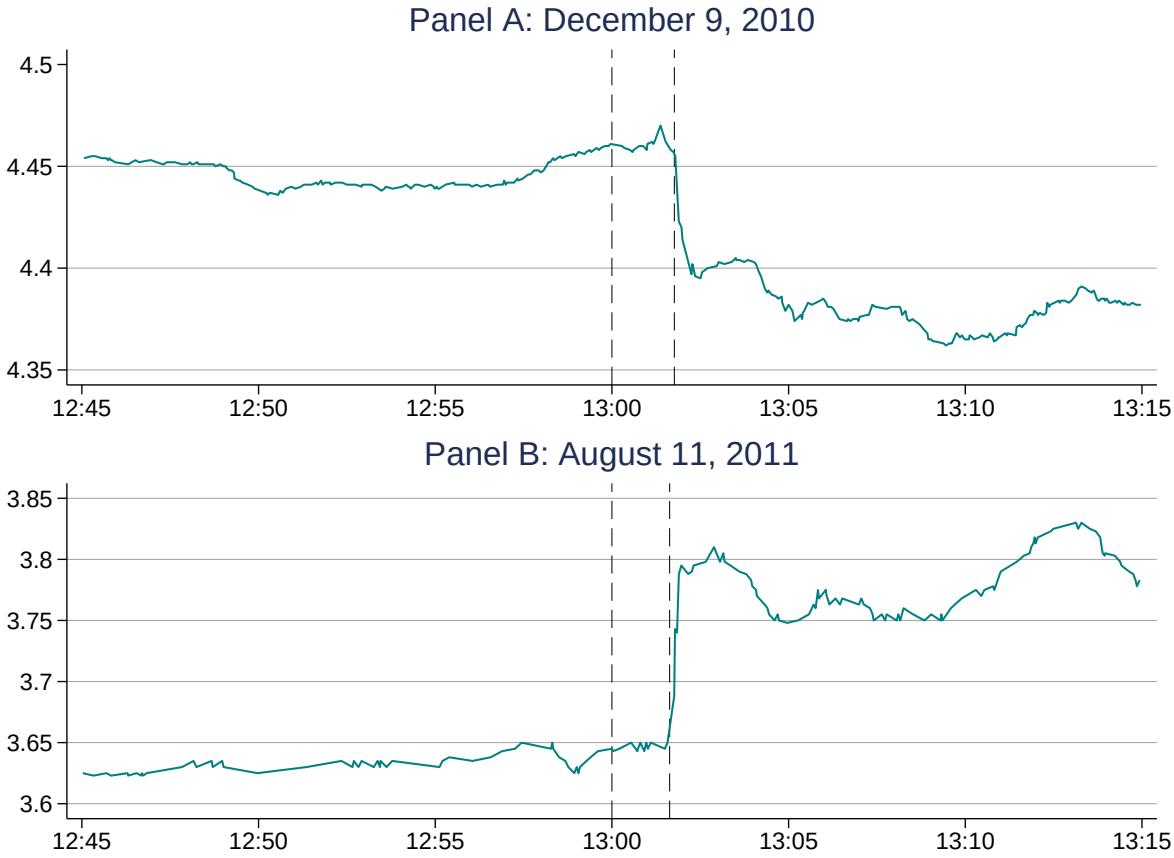

Figure 2: Intraday Treasury Yield Movements

Notes: Intraday movements in 30-year Treasury yields around the announcement of auction results for two selected auction dates. Dashed vertical lines denote the close of an auction and release of results. In both auctions considered, auctions closed at 1:00pm and results were released a few minutes thereafter.

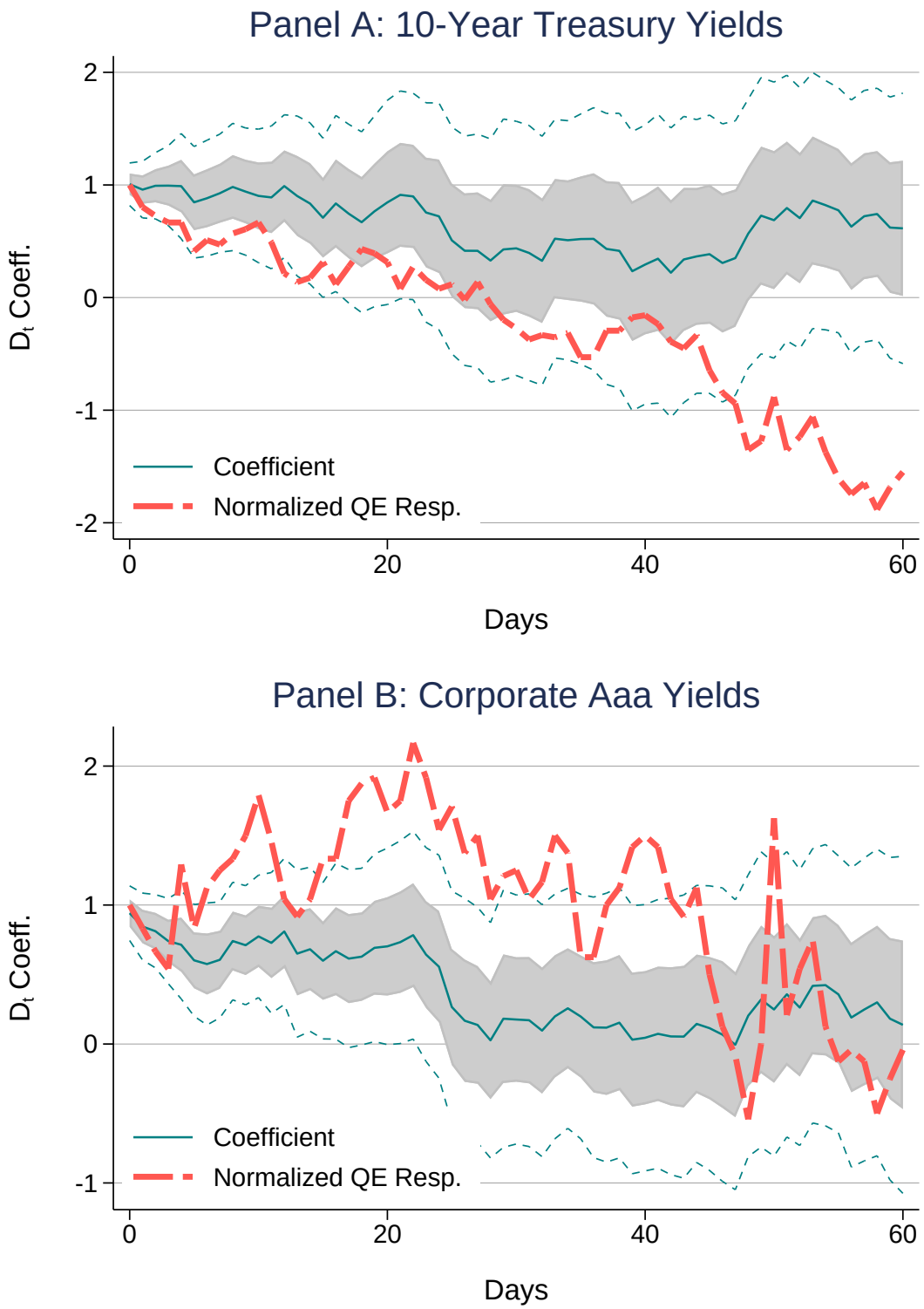

Figure 3: Long-Difference Response to Demand Shocks

Notes: Responses of 10-year Treasury spot rates (Panel A) and Moody's Aaa yields (Panel B) to demand shocks D_t (pooled across maturities). We compute "long-difference" regressions: on an auction date t, the dependent variable is $y_{t+h} - y_{t-1}$, the change h days after the auction relative to the day before the auction. The solid line plots the coefficients from regressions for $h = 0, \ldots, 60$; the shaded region and dotted lines correspond to one- and two-standard error (Newey-West, 9 lags) confidence bands. The dashed line compares the long-horizon effects of the March 2009 QE1 FOMC announcement (normalized so that the impact on announcement is one).

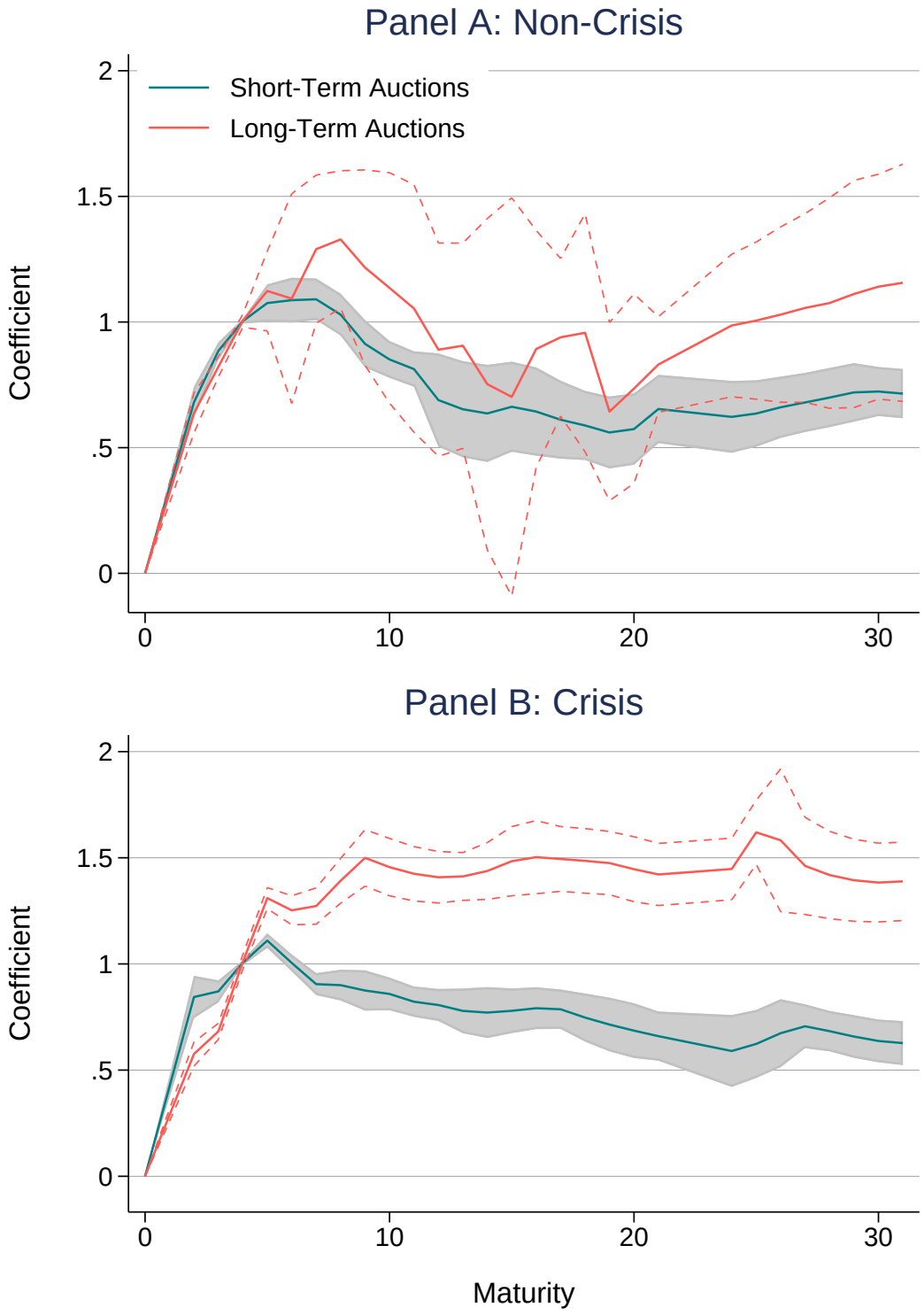

Figure 4: Localization Regression Results

Notes: Plots of the regression coefficients from regression equation (4). The top panel compares estimates from short- and long-maturity auctions during non-crisis periods; the bottom panel compares these estimates during crisis periods. 2 standard error (Newey-West, 9 lags) confidence intervals are included. Appendix Figure B11 reports p-values for the equality of estimates.

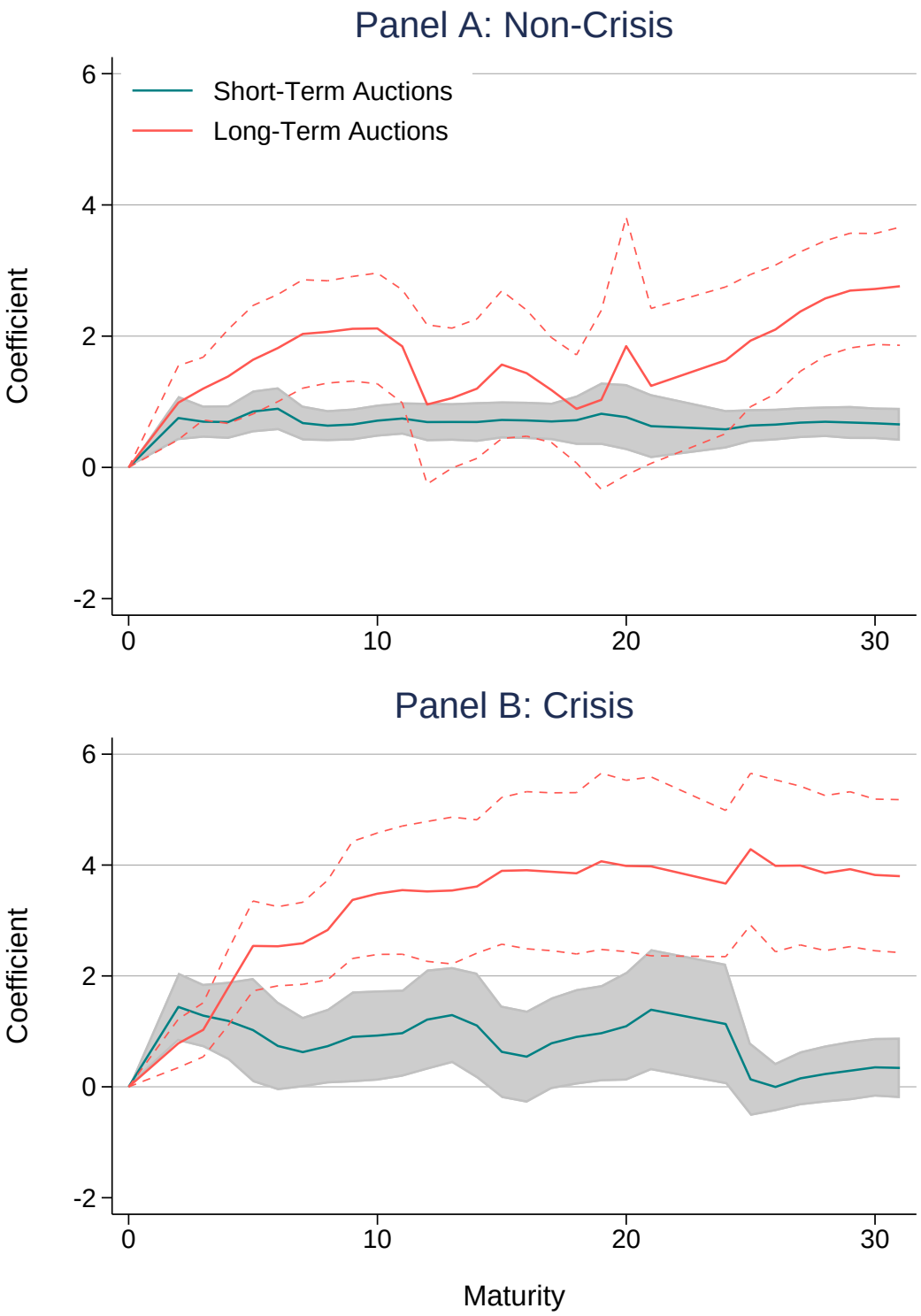

Figure 5: Alternative Localization: Bid-to-Cover

Notes: Estimates of the alternative localization regression (5), using the change in the bid-to-cover ratio as a proxy of structural demand shocks. We further flip the sign of the coefficients in order to make the estimates more comparable to our baseline specification. The top panel compares estimates from short- and long-maturity auctions during non-crisis periods; the bottom panel compares these estimates during crisis periods. 2 standard error (Newey-West, 9 lags) confidence intervals are included.

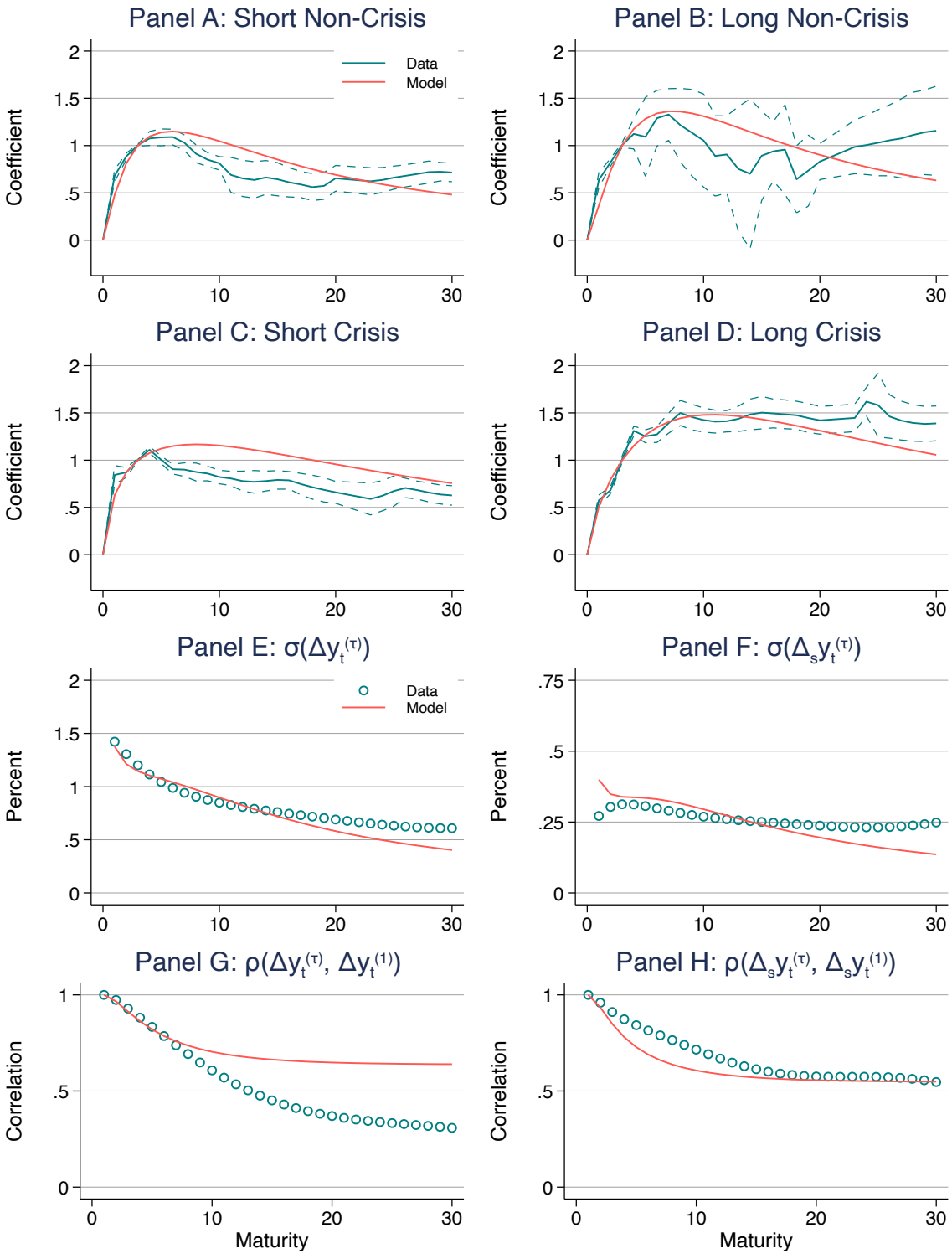

Figure 6: Model Fit

Notes: Panels A through D plot the model-implied and empirical coefficients from regression equation (4); Panels A and B plot the coefficients from the model calibrated to non-crisis periods, while Panels C and D plot the coefficients from the model calibrated to crisis periods. Panels E through H compare additional targeted volatility and correlation moments in the model (solid lines) with the data (scatter points) as a function of maturity.

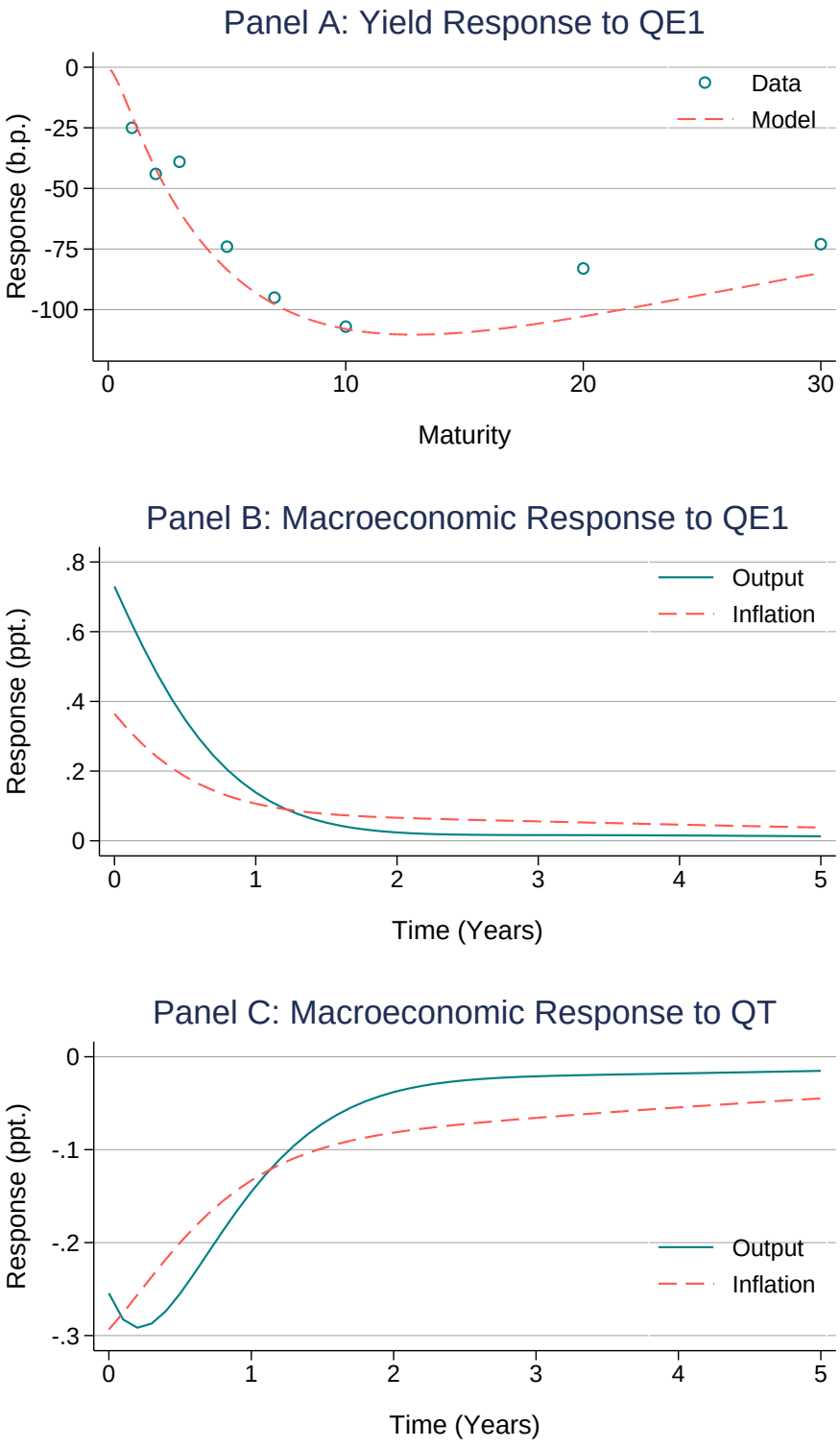

Figure 7: Model-Implied Response to QE1 and QT

Notes: Panel A compares the observed (cumulative) yield reactions and the model-implied change in yields following a QE1 shock (in basis points). Panels B and C plot the model-implied IRFs of output and inflation following QE1 and QT, respectively (in percentage points).

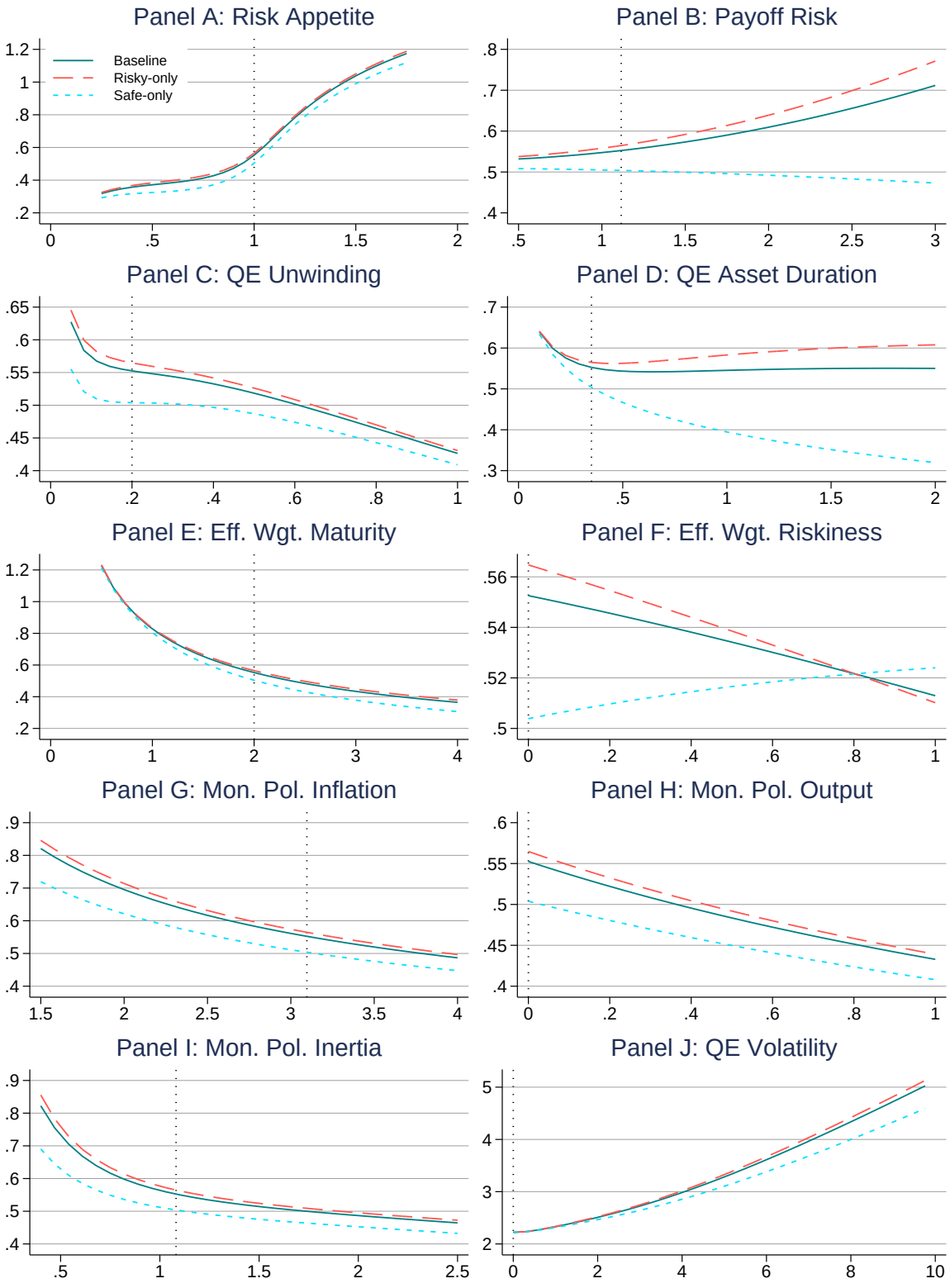

Figure 8: QE Sensitivity Analysis: Output

Notes: Response of output to a QE shock as a function of various model parameters. Panel A: arbitrageur risk aversion (a). Panel B: risky payoff uncertainty (σ_d). Panel C: QE mean reversion (κ_{QE}). Panel D: QE maturity duration (θ_1^{QE}). Panel E: effective borrowing maturity weights (η_1). Panel F: effective borrowing risky/safe weights (η_0). Panel G: inflation Taylor coefficient (ϕ_π). Panel H: output gap Taylor coefficient (ϕ_x). Panel I: inertia Taylor coefficient (κ_i). Panel J: QE volatility (σ_{QE}). In each panel, the vertical dashed line corresponds to the baseline QE calibration. The dashed orange line shows results for a hypothetical QE policy which purchases risky assets only; similarly, the dotted blue line is for Treasury-only purchases. Panel J reports the long-run volatility $Var[x_t]$, while all other panels report the cumulative effect $\int_0^\infty E_0[x_t]\,\mathrm{d}t$.

Table 1: Auction and Shock Summary Statistics

	Mean	Median	Std. Dev.	Min	Max	Obs.
Panel A: Auction Summary Statistics						
Offering Amount (billions)	22.48	22.00	9.02	5.00	44.00	1047
Total Tendered (billions)	61.98	57.91	29.91	11.35	160.96	1047
Term (Years)	8.23	5.00	8.68	2.00	30.25	1047
High Yield	3.01	2.68	1.83	0.22	7.79	1047
Bid-to-Cover	2.60	2.58	0.45	1.22	4.07	1047
Bid-to-Cover by type$^\sharp$						
Direct Bidders	0.22	0.21	0.16	0.00	0.80	827
Indirect Bidders	0.54	0.53	0.17	0.03	1.05	827
Primary Dealers	1.92	1.87	0.34	0.98	3.12	827
Fraction Accepted by type†						
Depository Institutions	0.59	0.15	2.09	0.00	31.79	896
Individuals	0.98	0.23	2.13	0.00	18.93	896
Dealers	53.44	53.07	16.22	12.08	97.73	896
Pensions	0.11	0.01	0.74	0.00	20.91	896
Investment Funds	25.81	23.05	16.95	0.23	73.91	896
Foreign	19.22	18.10	8.74	0.34	61.01	896
Other	0.25	0.03	0.91	0.00	13.46	896
Panel B: Shock Summary Statistics						
D_t	0.03	0.00	2.12	-8.30	17.80	1047
$D_t^{(2Y)}$	-0.06	0.00	1.39	-5.20	4.40	256
$D_t^{(3Y)}$	0.15	0.10	1.32	-5.30	5.00	134
$D_t^{(5Y)}$	0.20	0.30	1.94	-5.55	8.50	233
$D_t^{(7Y)}$	-0.28	0.05	2.10	-8.30	5.50	106
$D_t^{(10Y)}$	-0.04	-0.05	2.55	-6.80	12.20	186
$D_t^{(30Y)}$	0.09	-0.15	3.28	-7.40	17.80	132
Panel C: Shocks Across Regimes						
Non-ZLB	0.08	0.10	2.02	-7.40	12.20	563
ZLB	-0.03	0.00	2.23	-8.30	17.80	484
Expansion	0.03	0.00	2.03	-8.30	17.80	963
Recession	-0.07	-0.05	2.95	-7.10	9.50	84
Panel D: Non-Auction Shocks						
$\tilde{D}_t^{(2Y)}$	-0.07	0.00	1.32	-19.20	11.00	1225
$\tilde{D}_t^{(5Y)}$	-0.01	0.00	1.13	-17.40	12.90	1924
$\tilde{D}_t^{(10Y)}$	-0.01	0.00	1.03	-13.30	8.90	2151
$\tilde{D}_t^{(30Y)}$	0.01	0.00	0.92	-8.50	5.60	1914

Notes: Panel A presents summary statistics for Treasury note and bond auctions from 1995-2017 for which we have intraday data. † indicates that the moments are computed for 2000 onward, the period for which these data are available. $\sharp$ indicates that the moments are computed from 2003 onward, the period for which these data are available. Panel B summarizes demand shocks $D_t = y_{t,post} - y_{t,pre}$, the intraday change in Treasury yields before and after the close of an auction (in basis points). For newly issued securities, the yields are from when-issued trades. For reopened securities, the yields are from secondary-market trades. Panel B also reports statistics for shocks $D_t^{(m)} = y_{t,post}^{(m)} - y_{t,pre}^{(m)}$ separately for major maturities $m = 2, 3, 5, 7, 10, 30$ years. Panel C reports statistics separately for binding/non-binding ZLB periods and recessions/expansions. Panel D reports synthetic shocks on non-auction dates $\tilde{D}_t^{(m)} = y_{t,post}^{(m)} - y_{t,pre}^{(m)}$ using the same intraday windows around 1PM on non-auction dates. The yields are from secondary-market trades for on-the-run securities for maturities $m = 2, 5, 10, 30$ years.

Table 2: Demand Shocks and Measures of Demand

	(1) $D_t^{(2Y)}$	(2) $D_t^{(3Y)}$	(3) $D_t^{(5Y)}$	(4) $D_t^{(7Y)}$	(5) $D_t^{(10Y)}$	(6) $D_t^{(30Y)}$	(7) Pool D_t
Panel A: Total Bid-to-Cover Ratio							
Bid-to-Cover	-0.74***	-1.51***	-1.25***	-5.67***	-2.66***	-4.22***	-1.64***
	(0.18)	(0.27)	(0.23)	(1.22)	(0.45)	(1.23)	(0.17)
Obs.	255	134	233	106	186	132	1046
R^2	0.08	0.23	0.10	0.28	0.17	0.21	0.12
Panel B: Bid-to-Cover by Bidder Type							
Indirect	-2.24***	-3.83***	-4.22***	-8.34***	-5.59***	-11.30***	-5.10***
	(0.42)	(0.92)	(0.75)	(1.89)	(0.97)	(2.30)	(0.57)
Direct	-0.41	-0.08	-1.14	-7.40***	-1.03	-3.29	-1.63**
	(0.74)	(1.00)	(1.19)	(2.06)	(1.03)	(3.24)	(0.73)
Primary	-0.56**	-0.89**	-0.76**	-1.28	-1.74***	-1.32	-0.86***
	(0.23)	(0.41)	(0.35)	(1.82)	(0.54)	(1.38)	(0.24)
Obs.	163	120	172	106	150	115	826
R^2	0.20	0.25	0.17	0.34	0.28	0.37	0.23
Panel C: Fraction Accepted by Bidder Type							
Invest. Fund	-3.03***	-3.34***	-3.93***	-11.76***	-3.67**	-12.96***	-5.37***
	(0.90)	(1.13)	(1.20)	(2.32)	(1.53)	(3.09)	(0.73)
Foreign	-2.13**	-5.80***	-4.21***	-15.23***	-5.35***	-16.07***	-5.80***
	(0.98)	(2.02)	(1.43)	(4.37)	(1.42)	(3.92)	(0.79)
Misc.	-1.81	-3.62	-3.14	-13.91***	-4.37	-4.97	-3.96**
	(2.29)	(3.74)	(2.20)	(3.50)	(3.72)	(10.13)	(1.61)
Obs.	200	121	185	106	164	119	895
R^2	0.08	0.14	0.07	0.30	0.11	0.28	0.13

Notes: Regressions of demand shocks $D_t^{(m)}$ on the change in bid-to-cover ratio and fractions accepted, total and broken up by bidder type. Columns (1)-(6) restrict the sample to include only auctions of maturities $m = 2, 3, 5, 7, 10, 30$ years; column (7) pools across all auctions. Changes in the bid-to-cover ratio and fraction accepted are computed relative to the most recent note or bond auction. Newey-West (9 lags) standard errors in parentheses. ***, **, * denote statistical significance at 1, 5 and 10 percent.

Table 3: Asset Price Reactions to Demand Shocks

Asset type	Estimate	Std. Err.	Obs.	R^2	Sample
	(1)	(2)	(3)	(4)	(5)
Panel A: Corporate and Private Debt					
LQD	-3.93***	(0.15)	830	0.59	2002-2017
HYG	-0.34	(0.30)	678	0.01	2007-2017
MBB	-1.42***	(0.18)	662	0.34	2007-2017
VMBS	-1.46***	(0.33)	371	0.13	2009-2017
Corp. Aaa†	0.94***	(0.10)	1040	0.14	1995-2017
Corp. Baa†	0.96***	(0.10)	1040	0.15	1995-2017
Corp. C†	0.23	(0.39)	973	0.00	1997-2017
Panel B: Equities					
SPY	-0.23	(0.52)	1033	0.00	1995-2017
IWM	0.18	(0.58)	876	0.00	2000-2017
SP500†	3.61	(2.74)	974	0.00	1995-2016
Russell 2000†	6.26*	(3.25)	974	0.01	1995-2016
Panel C: Swaps, Commodities, and Spreads					
GLD	-1.16***	(0.36)	775	0.02	2004-2017
GSCI†	-1.20	(2.74)	974	0.00	1995-2016
10Y Infl. Swap†	0.02	(0.08)	724	0.00	2004-2016
2Y Infl. Swap†	0.04	(0.17)	724	0.00	2004-2016
LIBOR-OIS†	0.03	(0.04)	737	0.00	2003-2016
Auto CDS†	-0.60	(2.65)	733	0.00	2004-2016
Bank CDS†	-0.23	(0.16)	733	0.01	2004-2016
VIX†	-3.82	(3.37)	1040	0.00	1995-2017
Panel D: Federal Funds Futures					
1-Month Ahead†	0.03	(0.02)	1040	0.00	1995-2017
2-Month/1-Month Slope†	0.00	(0.01)	1040	0.00	1995-2017

Notes: Regressions of asset price changes on demand shocks D_t (pooled across maturities). † denotes daily series; otherwise, intraday changes are measured over the same window as auction demand shocks. Intraday changes are from ETFs that track securities or indices: LQD (investment grade corporate debt); HYG (high yield corporate debt); MBB and VMBS (mortgage indices); SPY (S&P 500); IWM (Russell 2000); GLD (gold bullion). Daily series: Aaa, Baa, C (Moody's & BoA corporate debt yield indices); S&P 500, Russell 2000 (equity indices); GSCI (S&P Total Commodity Index); 10- and 2-year inflation swaps; Auto and Bank credit default swap indices; LIBOR-OIS (3-month USD LIBOR-Overnight Index Swap spread); VIX (implied volatility index); Federal funds futures (h-month ahead continuous contracts). Newey-West (9 lags) standard errors in parentheses. All regressions are estimated with OLS. We report very similar instrumental variable estimates (using changes in the bid-to-cover ratio as instruments) in Appendix Table B1. ***, **, * denote statistical significance at 1, 5 and 10 percent.

Table 4: Localization Regression Results, Private Borrowing

	(1) LQD	(2) HYG	(3) MBB	(4) VMBS	(5) SPY	(6) IWM
Panel A: Crisis Localization						
$D_t^{(\tau^*)}$	-4.98***	-0.18	-3.41***	-2.83***	-0.03	2.85***
	(0.58)	(0.56)	(0.41)	(0.42)	(1.22)	(1.02)
Crisis $\times$ $D_t^{(\tau^*)}$	0.30	0.19	1.80***	4.51***	0.97	-2.64*
	(0.72)	(0.69)	(0.43)	(1.58)	(1.56)	(1.42)
Obs.	794	661	647	376	941	811
R^2	0.40	0.00	0.35	0.15	0.01	0.01
Sample	2002-2017	2007-2017	2007-2017	2009-2017	1995-2017	2001-2017
Panel B: Crisis and Maturity Localization						
$D_t^{(\tau^*)}$	-1.54***	-5.20***	-8.74***	-0.84***	-2.98***	-4.32***
	(0.18)	(0.37)	(0.93)	(0.17)	(0.25)	(1.24)
Long $\times$ $D_t^{(\tau^*)}$	-0.22	-0.88	-4.45*	0.23	-0.58	-4.37
	(0.38)	(1.46)	(2.31)	(0.28)	(0.62)	(3.37)
Crisis $\times$ $D_t^{(\tau^*)}$	0.19	0.95	1.72	-0.15	0.79*	-2.23
	(0.25)	(0.58)	(1.36)	(0.30)	(0.46)	(2.85)
Long $\times$ Crisis $\times$ $D_t^{(\tau^*)}$	0.04	-2.72*	-5.73*	-0.11	-0.54	-21.22**
	(0.51)	(1.61)	(3.15)	(0.52)	(0.81)	(10.82)
Obs.	664	653	569	647	642	401
R^2	0.36	0.62	0.38	0.14	0.28	0.19
Sample	2007-2017	2007-2017	2007-2017	2007-2017	2007-2017	2010-2017

Notes: Regressions of intraday changes in ETFs (proxying for various private borrowing rates) on the intraday change in intermediate yields $D_t^{(\tau^*)}$ (as in Figure 4). Panel A includes interactions with whether the auction took place in a "non-crisis" or "crisis" period, while Panel B additionally includes interactions with the maturity of the auction (using the same definitions as in Figure 4). The ETFs used in Panel A are the same as in Table 3. In Panel B, the Vanguard ETFs "BSV," "BIV," and "BLV" track investment-grade bonds with short, intermediate, and long maturities (respectively). The BlackRock iShares ETFs "CSJ," "CIU,", and "CLY" track investment-grade corporate bonds with short, intermediate, and long maturities (respectively). Newey-West (9 lags) standard errors in parentheses. ***, **, * denote statistical significance at 1, 5 and 10 percent.

Table 5: Model Calibration and Targeted Moments

Panel A: Matched Moments			
Target Moment	Data		Model
$\sigma(y_t^{(1)})$	1.993		1.978
$\sigma(\tilde{y}_t^{(1)} - y_t^{(1)})$	0.542		0.549
$\sigma(\pi_t)$	0.973		0.882
$\sigma(x_t)$	1.992		1.996
$\sigma(\Delta y_t^{(1)})$	1.423		1.376
$\sigma(\Delta(\tilde{y}_t^{(1)} - y_t^{(1)}))$	0.615		0.617
$\sigma(\Delta\pi_t)$	0.900		0.935
$\sigma(\Delta x_t)$	2.190		2.137
$\sigma(\Delta_s y_t^{(1)})$	0.272		0.399
$\sigma(\Delta_s(\tilde{y}_t^{(1)} - y_t^{(1)}))$	0.258		0.218
$\sigma(\Delta_s\pi_t)$	0.232		0.362
$\sigma(\Delta_s x_t)$	0.499		0.745
$\rho(\tilde{y}_t^{(1)} - y_t^{(1)}, x_t)$	-0.171		-0.158
$\rho(y_t^{(1)}, \pi_t)$	0.549		0.615
$\rho(\pi_t, x_t)$	0.268		0.271
Panel B: Calibrated Parameters			
Parameter	Value	(Crisis)	Description
σ_i	2.567		Monetary Policy Vol.
κ_i	1.082		Monetary Policy Inertia
σ_d	1.116		Risky Payoff Vol.
κ_d	0.879		Risky Payoff Inertia
$\sigma_{z,\pi}$	2.039		Cost-Push Shock Vol.
$\kappa_{z,\pi}$	0.801		Cost-Push Shock Inertia
$\sigma_{z,x}$	1.749		Agg. Demand Shock Vol.
$\kappa_{z,x}$	0.253		Agg. Demand Shock Inertia
ϕ_π	3.096		Inflation Taylor Coeff.
ψ_x	0.393		Risky Payoff Output Coeff.
δ	0.705		Nominal Rigidity
ρ	0.04		Discount Factor
ς^{-1}	1.00		Intertemporal Elasticity
κ_β	1.367		Habitat Demand Inertia
$a \cdot \alpha_0$	0.008	(0.018)	Habitat Elasticity Size
$a \cdot \sigma_\beta \cdot \theta_0$	2.509	(5.123)	Habitat Demand Size
$a \cdot \phi_{i,\beta}$	0.491	(4.620)	Habitat Demand Short Rate Response
θ_1^s	0.50		Short Treasury Factor Maturity Wgt.
θ_1^ℓ	0.20		Long Treasury Factor Maturity Wgt.
$\tilde{\theta}_1$	0.50		Risky Factor Maturity Wgt.
α_1	0.10		Habitat Elasticity Maturity Wgt.
Panel C: QE/QT Parameters			
θ_1^{QE}	0.35		QE1 Maturity Wgt.
κ_{QE}	0.20		QE1 Inertia
θ_1^{QT}	0.50		QT Maturity Wgt.
$\kappa_{QT,A}$	0.20		QT Inertia, Active Comp.
$\kappa_{QT,P}$	2.25		QT Inertia, Passive Comp.
$\gamma_{QT,A,P}$	1.75		QT Passive Comp. Response

Notes: Panel A presents the targeted moments which are not a function of maturity which we use to calibrate the model. The targeted moments are volatility and correlations across short-term rates and macroeconomic variables. Δ denotes a 1-year change, while Δ_s denote a 1-month change. $\sigma(\cdot)$ denotes standard deviation (reported in percentage points), while $\rho(\cdot,\cdot)$ denotes correlation. Panel B presents calibrated parameters used in our quantitative exercises. Panel C reports the parameterization of our QE and QT shocks.

Online Appendix for "Unbundling Quantitative Easing: Taking a Cue from Treasury Auctions"

by Walker Ray, Michael Droste, and Yuriy Gorodnichenko

Appendix B Additional Figures and Tables

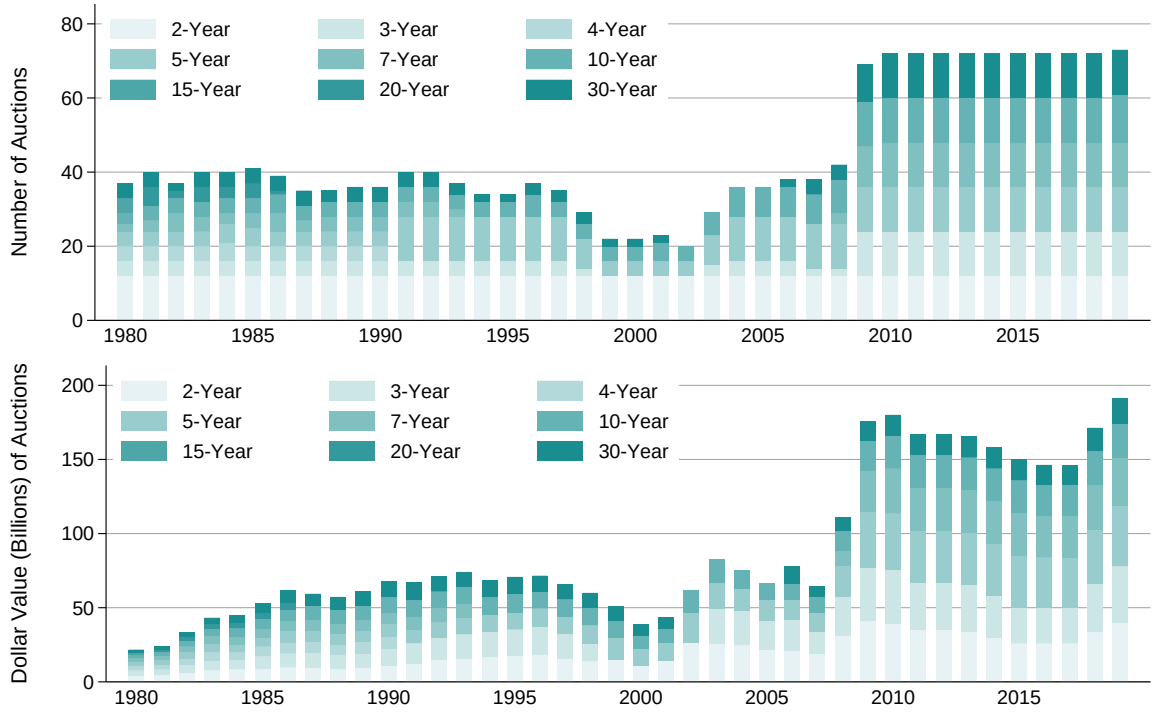

Figure B1: Number and Size of Auctions per Year

Notes: The number (top panel) and dollar value (bottom panel) of note and bond Treasury auctions per year by term length. The number of auctions temporarily fell in the late 1990s and early 2000s (during which time the Treasury ceased issuing 7- and 30-year securities). After the Great Recession, the number and size of auctions increased sharply. Since 2010, the Treasury has held auctions every month for maturities $m = 2, 3, 5, 7, 10, 30$ years.

August 03, 2011 202-504-3550

TREASURY OFFERING ANNOUNCEMENT [1]

Term and Type of Security	30-Year Bond
Offering Amount	$16,000,000,000
Currently Outstanding	$0
CUSIP Number	912810QS0
Auction Date	August 11, 2011
Original Issue Date	August 15, 2011
Issue Date	August 15, 2011
Maturity Date	August 15, 2041
Dated Date	August 15, 2011
Series	Bonds of August 2041
Yield	Determined at Auction
Interest Rate	Determined at Auction
Interest Payment Dates	February 15 and August 15
Accrued Interest from 08/15/2011 to 08/15/2011	None
Premium or Discount	Determined at Auction
Minimum Amount Required for STRIPS	$100
Corpus CUSIP Number	912803DT7
Additional TINT(s) Due Date(s) and	August 15, 2041
CUSIP Number(s)	912834KP2
Maximum Award	$5,600,000,000
Maximum Recognized Bid at a Single Yield	$5,600,000,000
NLP Reporting Threshold	$5,600,000,000
NLP Exclusion Amount	$0
Minimum Bid Amount and Multiples	$100
Competitive Bid Yield Increments [2]	0.001%
Maximum Noncompetitive Award	$5,000,000
Eligible for Holding in Treasury Direct Systems	Yes
Eligible for Holding in Legacy Treasury Direct	No
Estimated Amount of Maturing Coupon Securities Held by the Public	$24,430,000,000
Maturing Date	August 15, 2011
SOMA Holdings Maturing	$2,205,000,000
SOMA Amounts Included in Offering Amount	No
FIMA Amounts Included in Offering Amount [3]	Yes
Noncompetitive Closing Time	12:00 Noon ET
Competitive Closing Time	1:00 p.m. ET

(a) Announcement

August 11, 2011 202-504-3550

TREASURY AUCTION RESULTS

Term and Type of Security	30-Year Bond
CUSIP Number	912810QS0
Series	Bonds of August 2041
Interest Rate	3-3/4%
High Yield[1]	3.750%
Allotted at High	41.74%
Price	100.000000
Accrued Interest per $1,000	None
Median Yield[2]	3.629%
Low Yield[3]	3.537%
Issue Date	August 15, 2011
Maturity Date	August 15, 2041
Original Issue Date	August 15, 2011
Dated Date	August 15, 2011

	Tendered	**Accepted**
Competitive	$33,305,800,000	$15,985,160,000
Noncompetitive	$14,855,600	$14,855,600
FIMA (Noncompetitive)	$0	$0
Subtotal[4]	**$33,320,655,600**	**$16,000,015,600**[5]
SOMA	$489,928,400	$489,928,400
Total	**$33,810,584,000**	**$16,489,944,000**

	Tendered	**Accepted**
Primary Dealer[6]	$23,734,000,000	$10,921,532,000
Direct Bidder[7]	$6,567,000,000	$3,119,654,000
Indirect Bidder[8]	$3,004,800,000	$1,943,974,000
Total Competitive	**$33,305,800,000**	**$15,985,160,000**

(b) Results

Figure B2: Example of Auction Press Releases

Notes: Press releases from the Treasury for an auction of 30-year bonds. Source: TreasuryDirect.

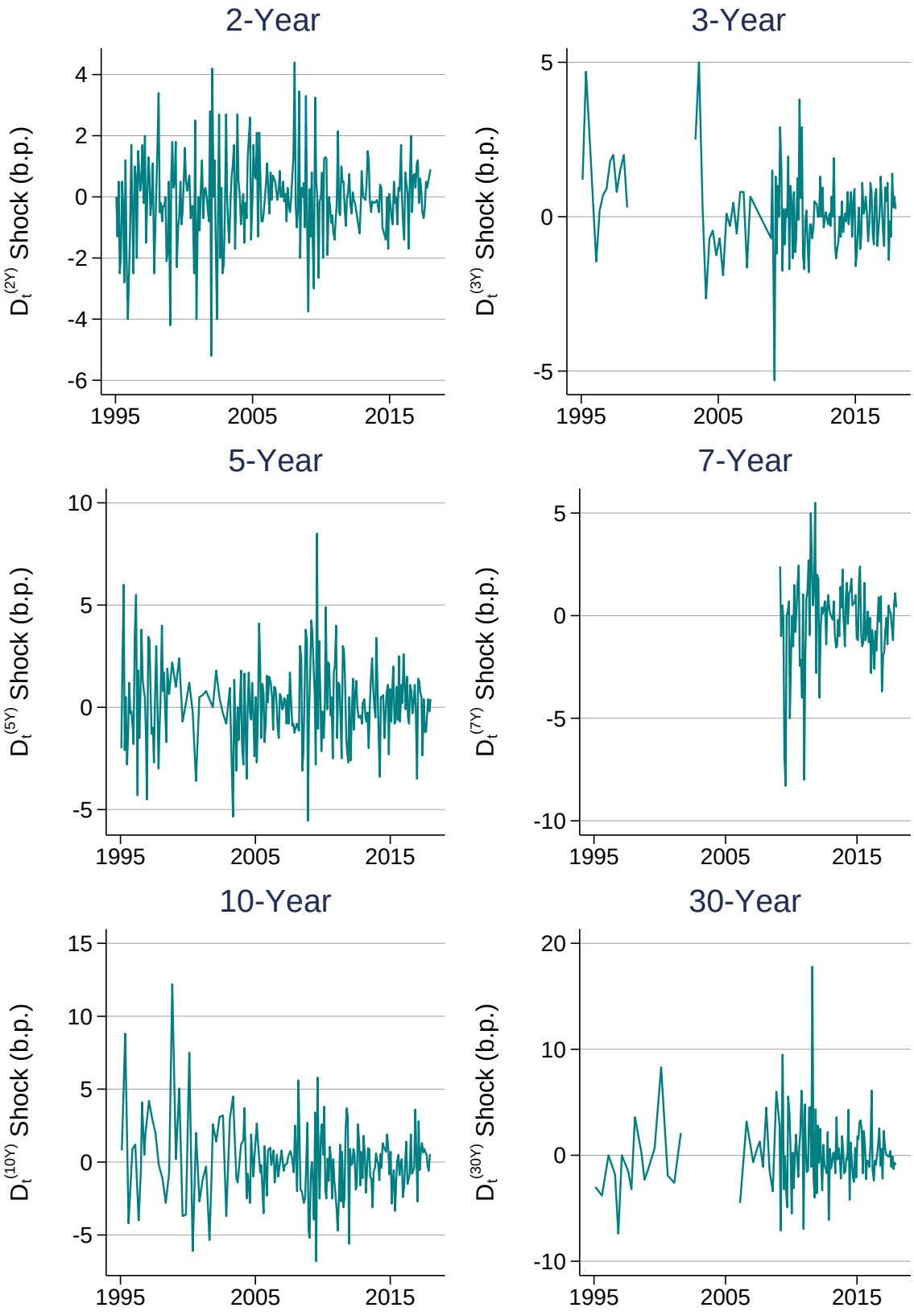

Figure B3: Time Series of Surprise Movements in Treasury Yields

Notes: Plots of the times series of intraday yield shocks following the close of Treasury auctions. Shocks $D_t^{(m)}$ are plotted separately for auctions of maturities $m = 2, 3, 5, 7, 10, 30$ years (in basis points). Gaps in the time series of 3-year, 7-year, and 30-year shocks correspond to temporary halts of newly issued securities of these maturities; see Figure B1.

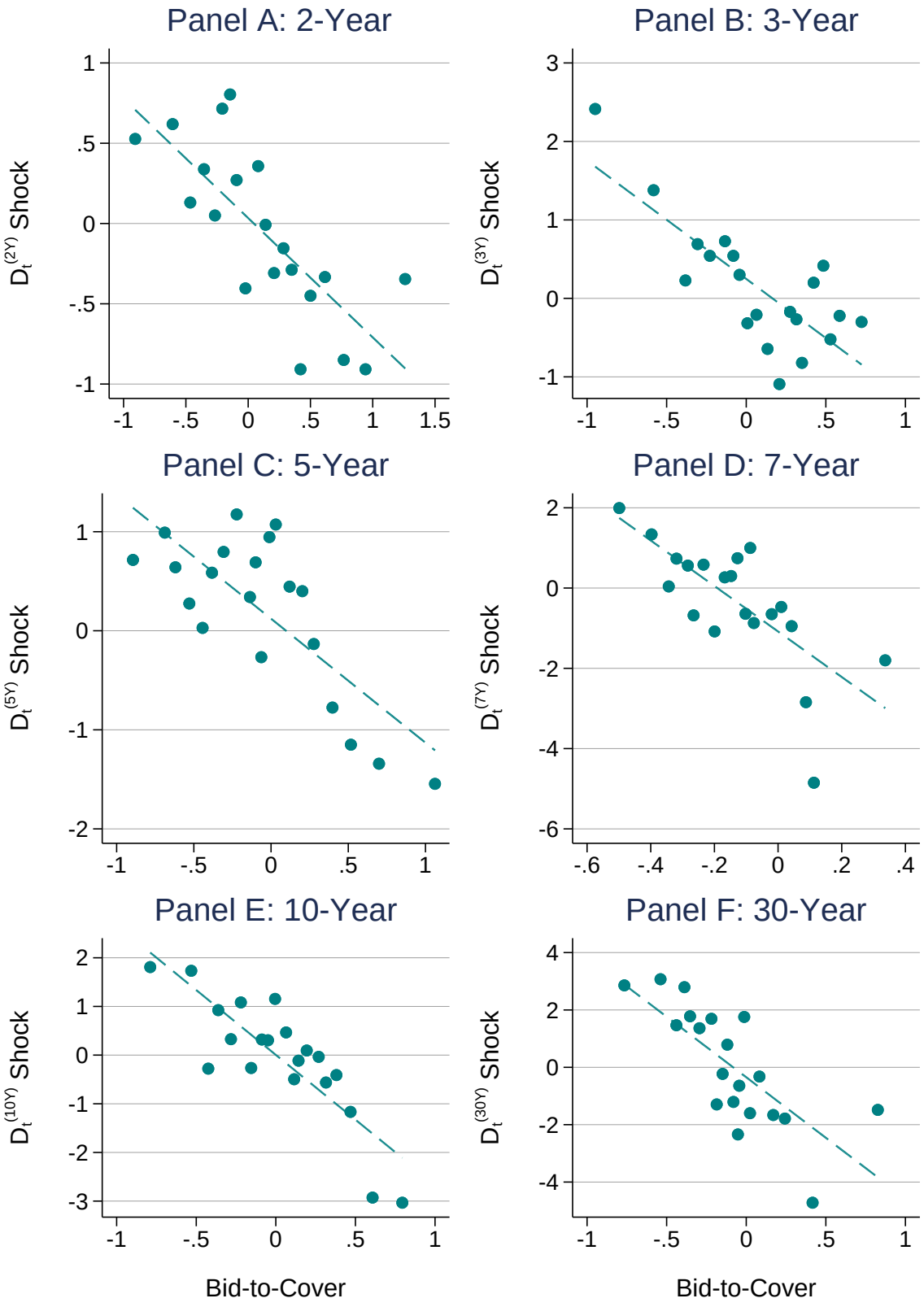

Figure B4: Demand Shocks and the Bid-to-Cover

Notes: Binned scatter plots examining the relationship between demand shocks and the change in the bid-to-cover ratio across auctions. The bid-to-cover is the ratio of the dollar value of bids received to those accepted at a given auction. Changes in the bid-to-cover ratio are computed relative to the most recent note or bond auction.

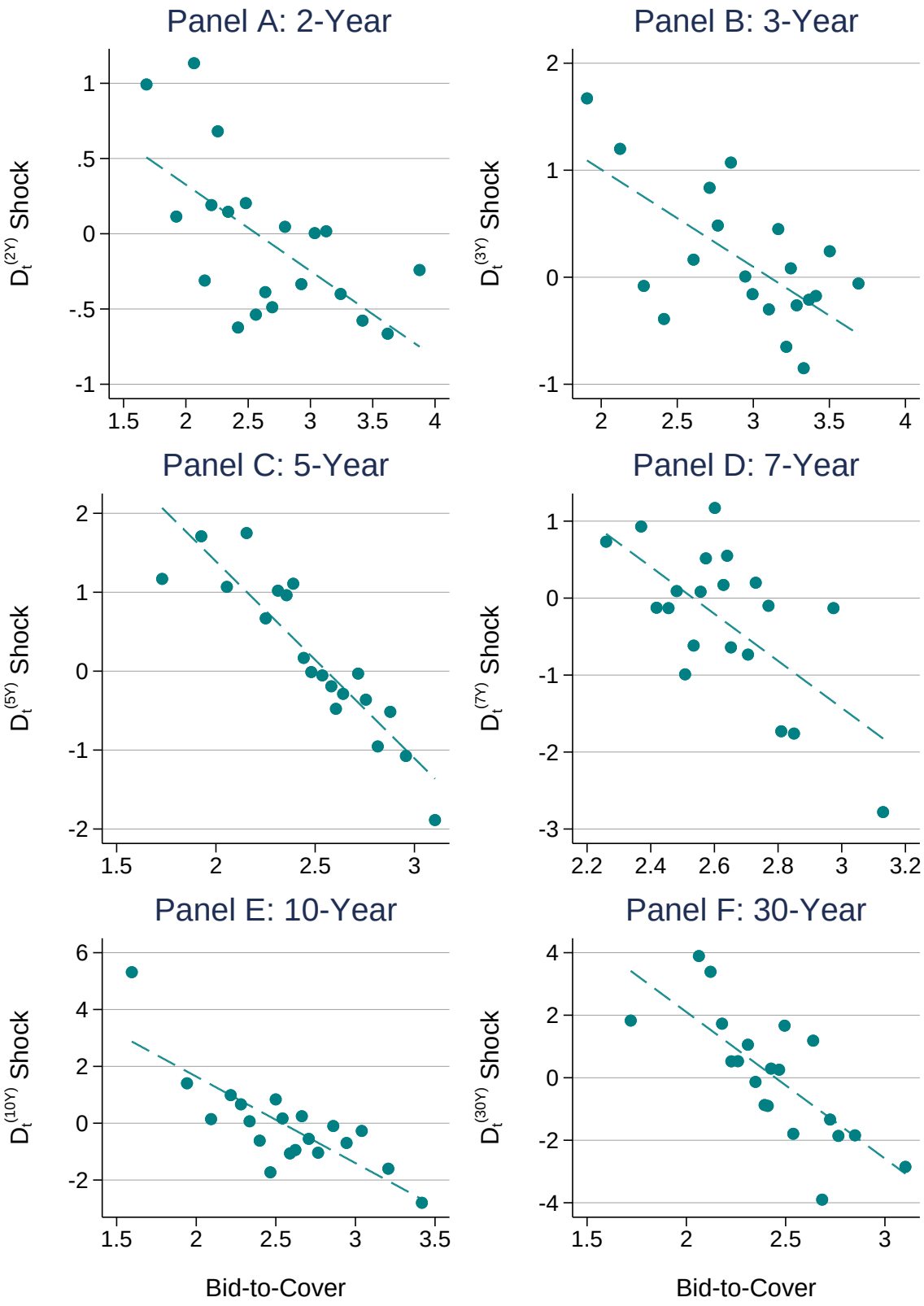

Figure B5: Demand Shocks and the Bid-to-Cover (Levels)

Notes: Binned scatter plots examining the relationship between demand shocks and the bid-to-cover ratio across auctions (in levels).

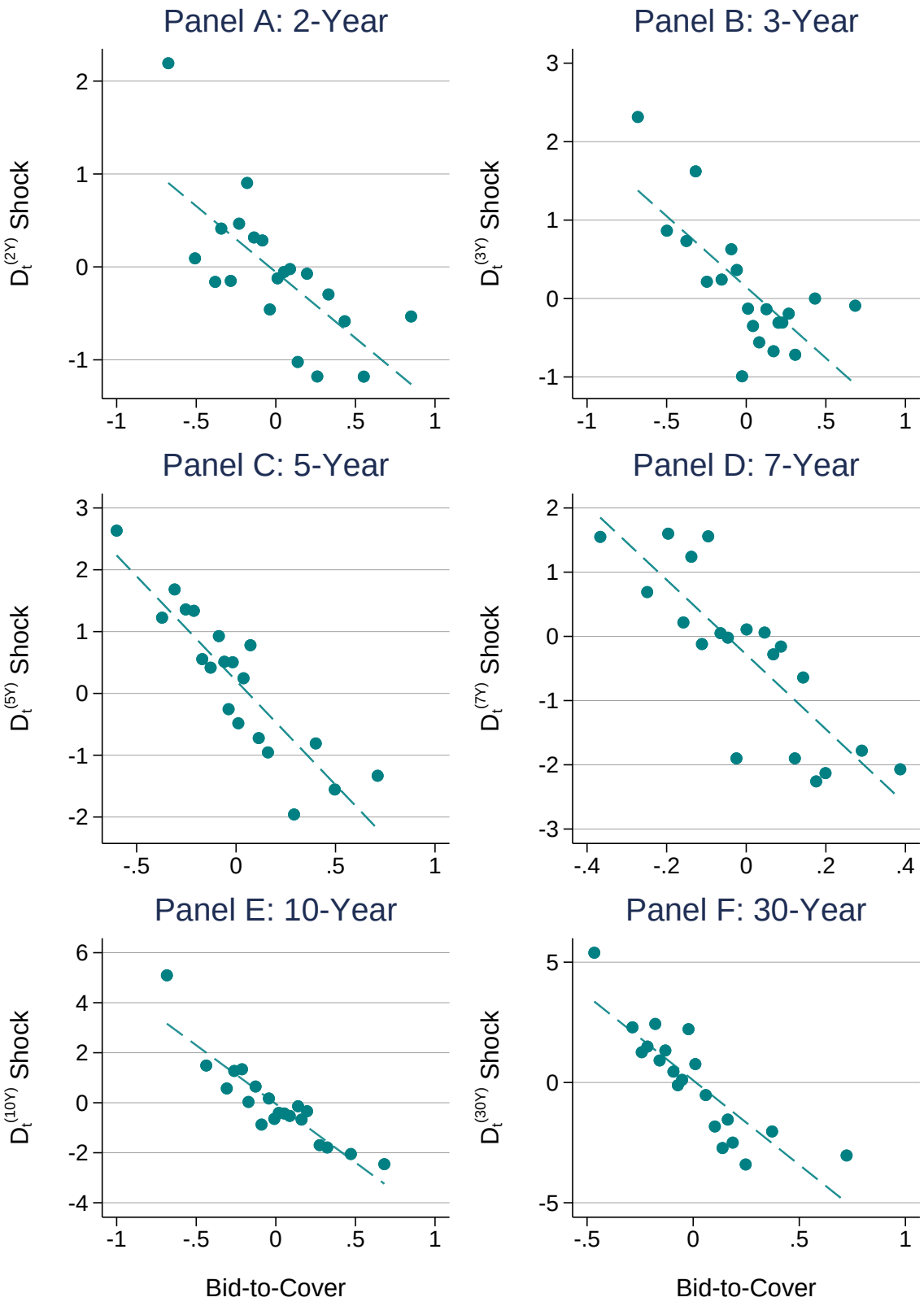

Figure B6: Demand Shocks and the Bid-to-Cover (Residualized)

Notes: Binned scatter plots examining the relationship between demand shocks and the residualized bid-to-cover ratio across auctions (residuals from an AR(4) model).

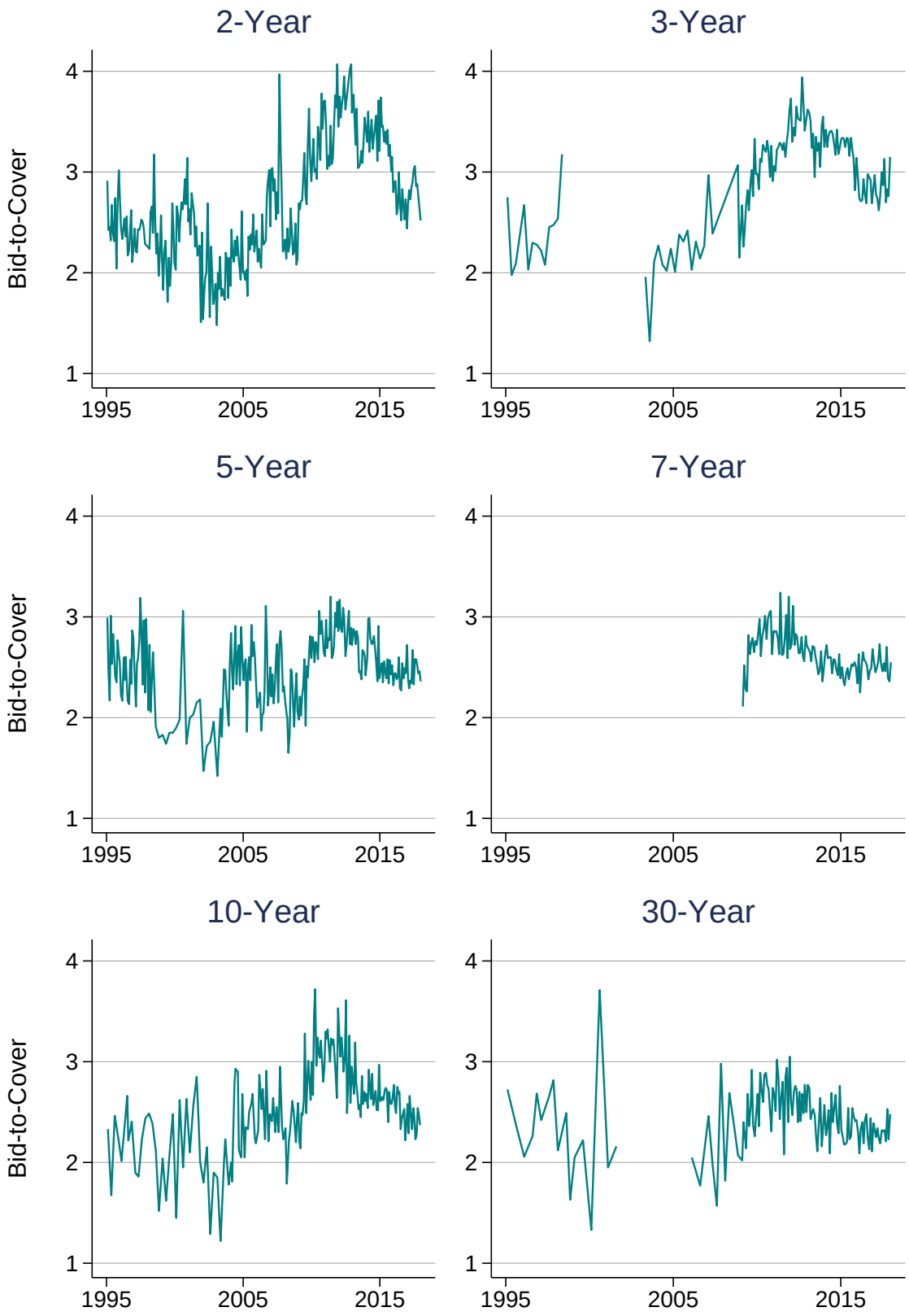

Figure B7: Time Series of Bid-to-Cover

Notes: Plots of the times series of the bid-to-cover by maturity. Gaps in the time series of 3-year, 7-year, and 30-year shocks correspond to temporary halts of newly issued securities of these maturities; see Figure B1.

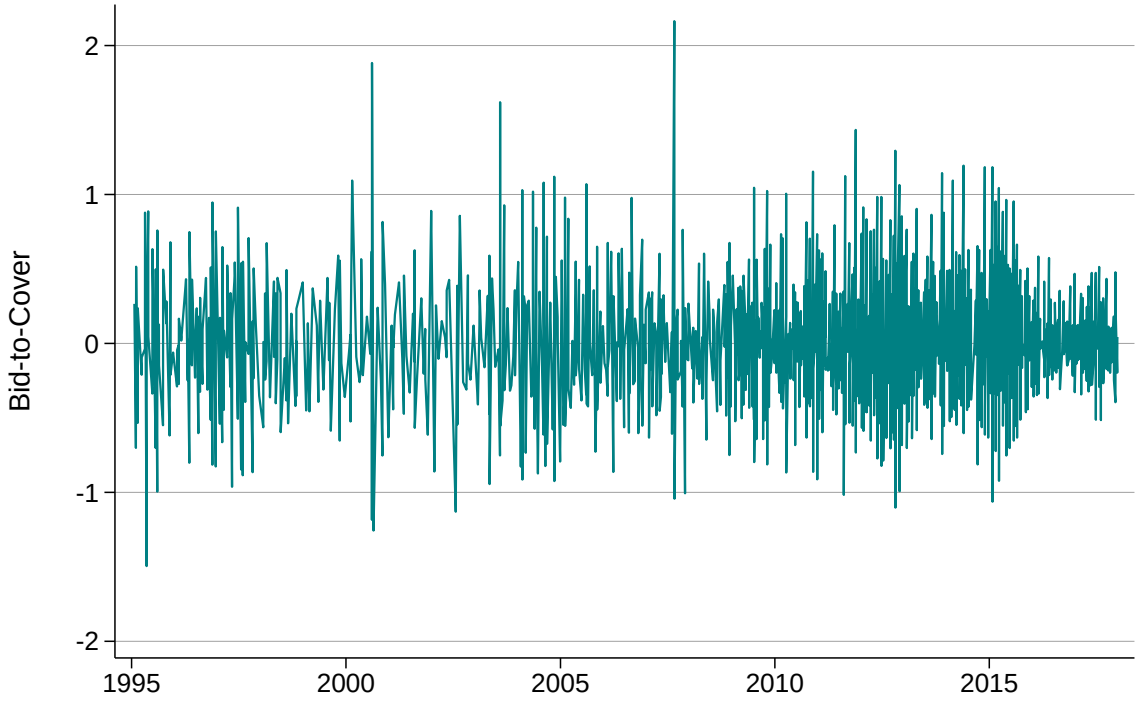

Figure B8: Time Series of Bid-to-Cover (Changes)

Notes: Plots of the times series of the change in the bid-to-cover. Changes in the bid-to-cover ratio are computed relative to the most recent note or bond auction.

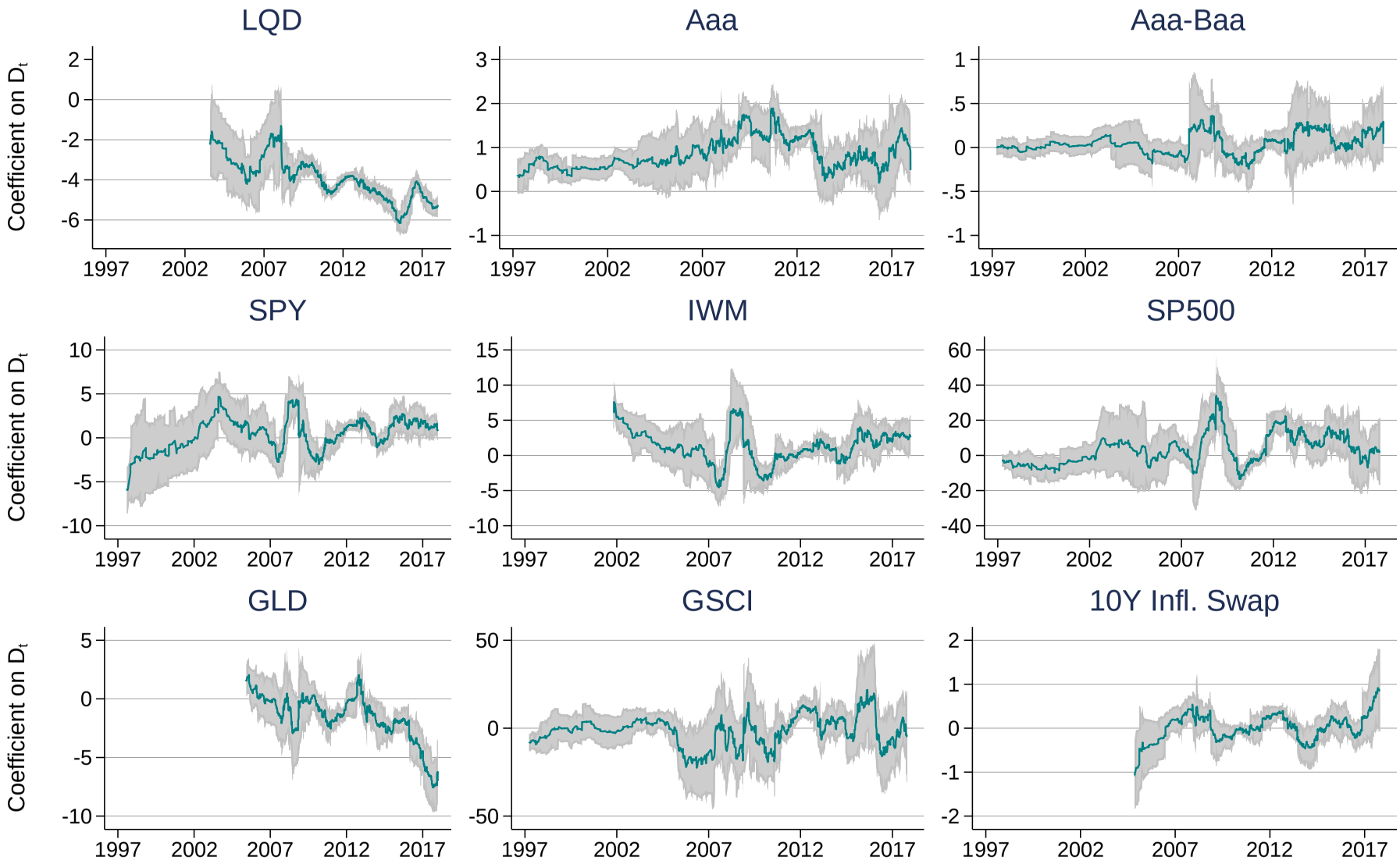

Figure B9: Rolling Regressions, Asset Prices

Notes: Rolling coefficient estimates of asset prices on auction demand shocks D_t (pooled across maturities). Each data point is from rolling estimates of the regressions reported in Table 3, using the 75 most recent auctions.

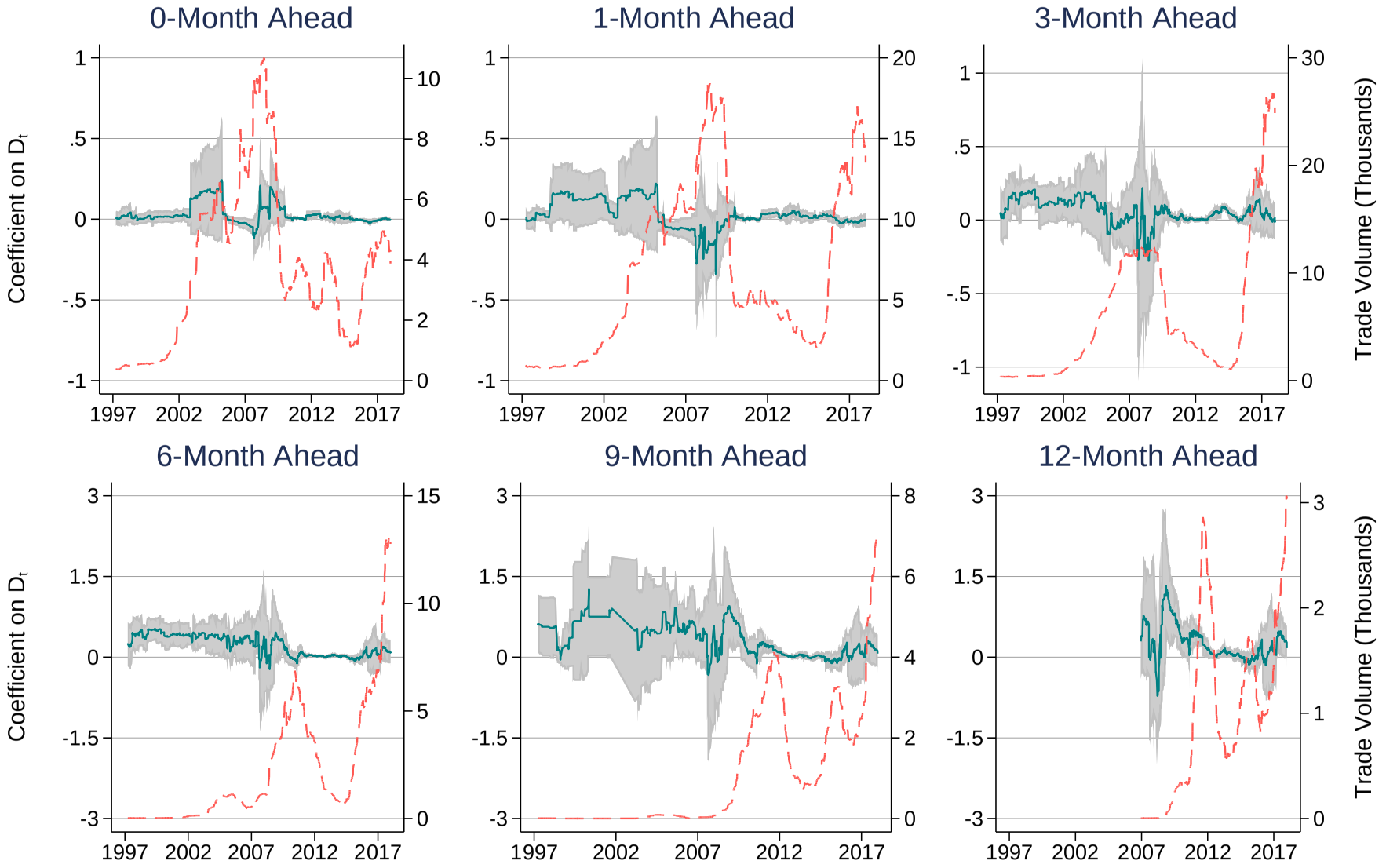

Figure B10: Rolling Regressions, Federal Funds Futures

Notes: Rolling coefficient estimates of h-month ahead expected federal funds rates on auction demand shocks D_t (pooled across maturities). Each data point presents rolling estimates using the 75 most recent auctions. The red line, indexed to the right y-axis, documents the average trade volume (in thousands) over the same 75-auction window

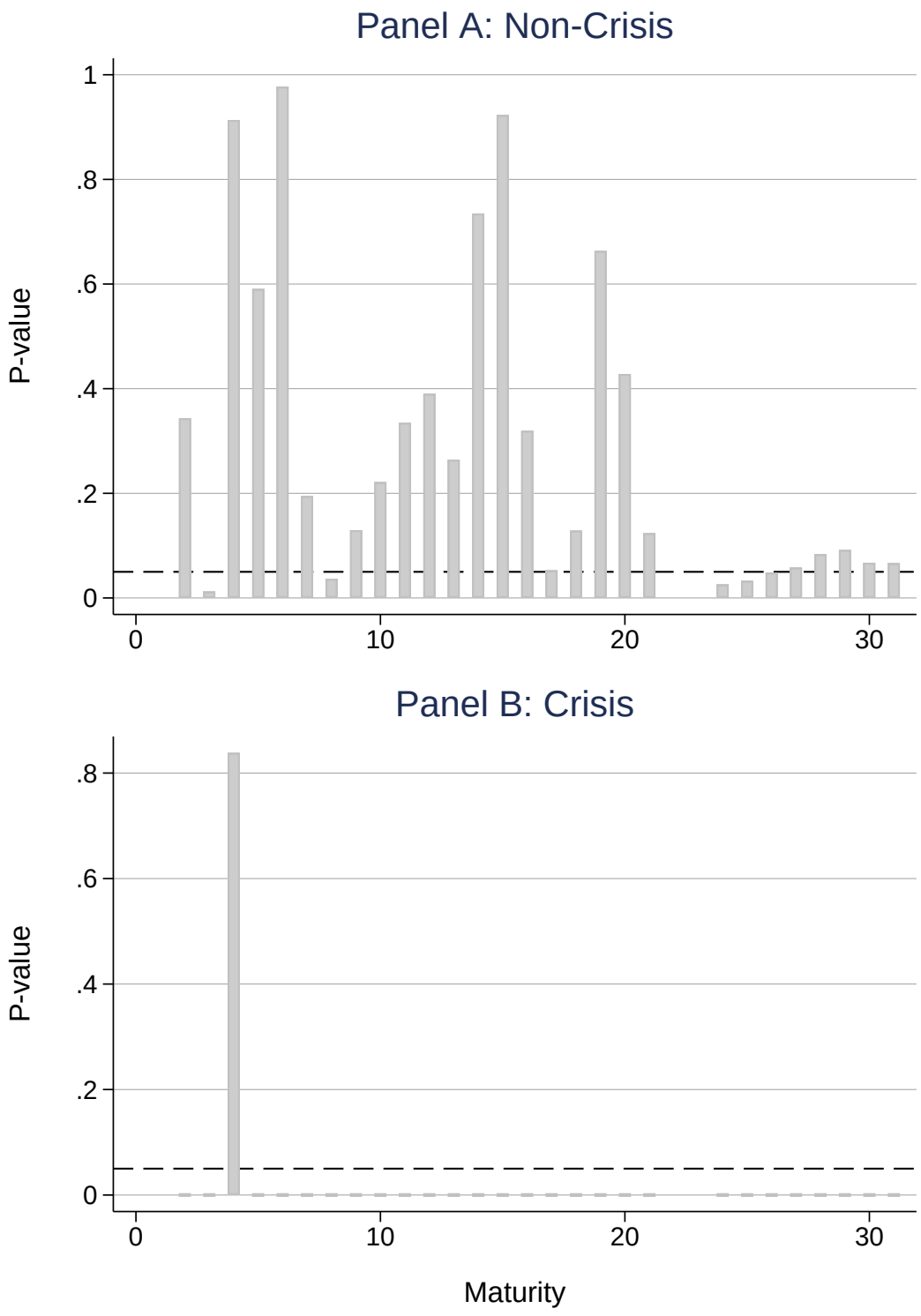

Figure B11: Localization Regression P-Values

Notes: P-values testing equality of coefficients from Figure 4.

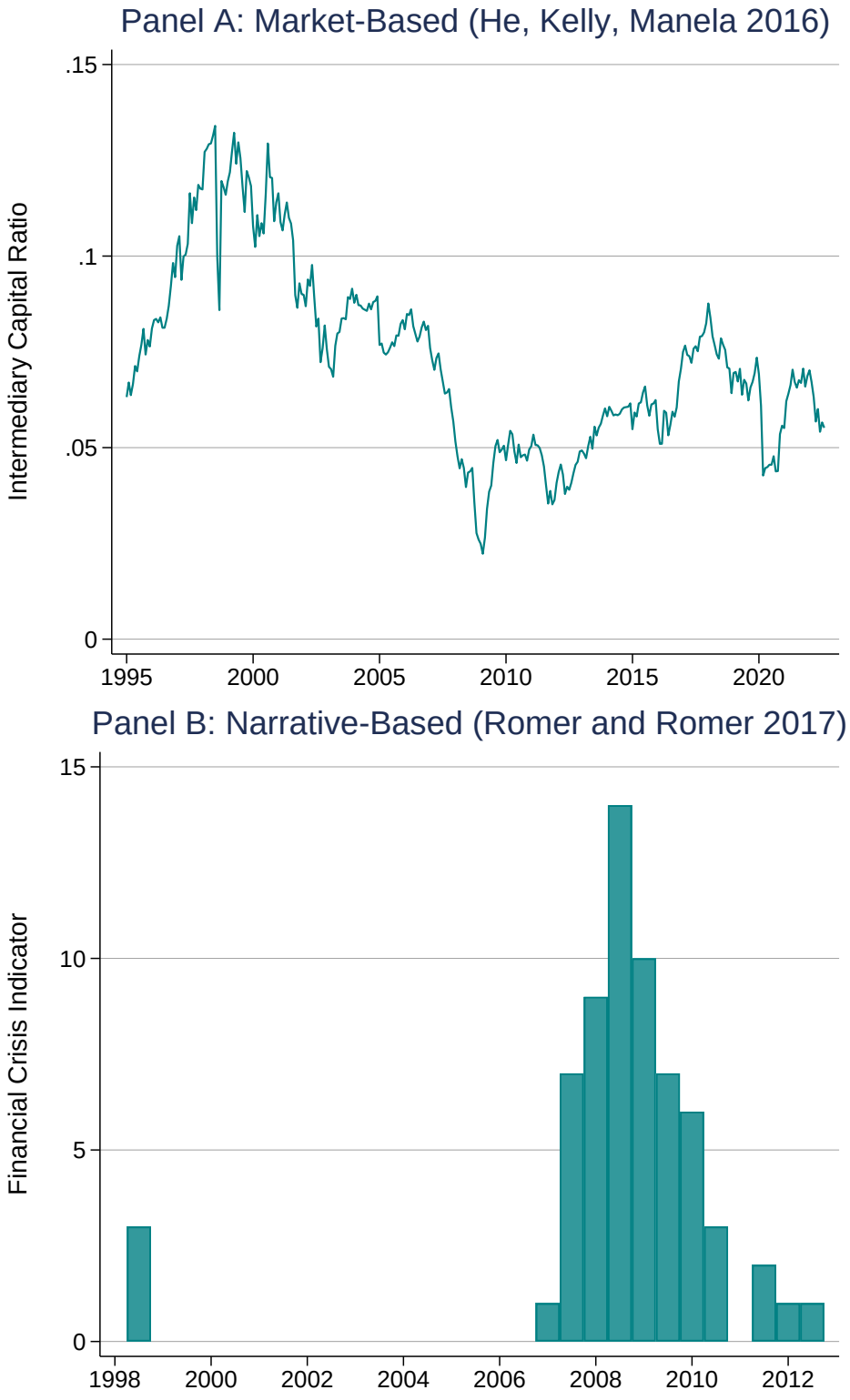

Figure B12: Alternative Measures of Financial Crises

Notes: Alternative measures of financial crises from He et al. (2016) (top panel) and Romer and Romer (2017) (bottom panel).

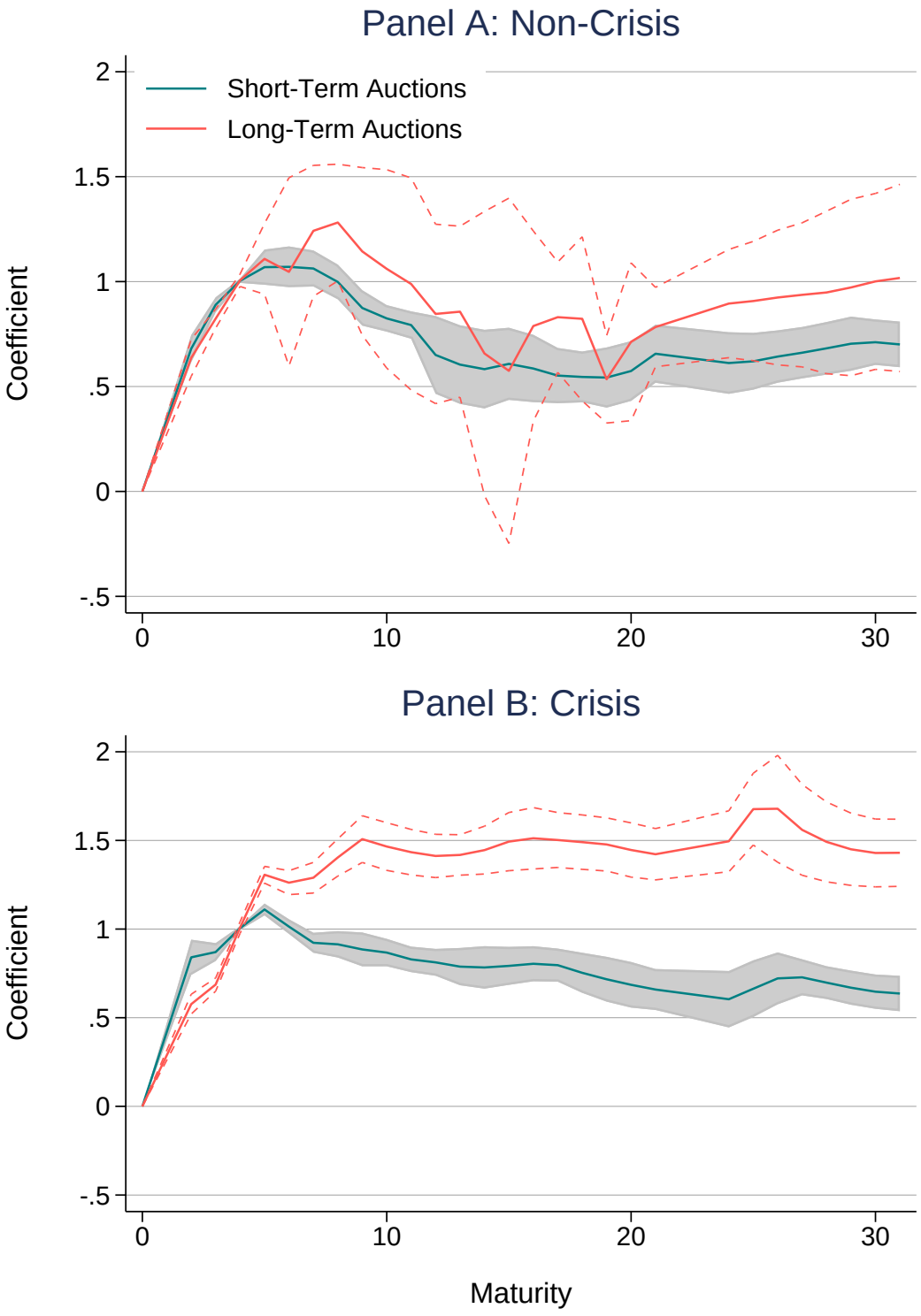

Figure B13: Localization Robustness: Market-Based Crises

Notes: Estimates of regression equation (4), identifying financial crises as periods in which the aggregate capital ratio from He et al. (2016) is low (≤ 0.055).

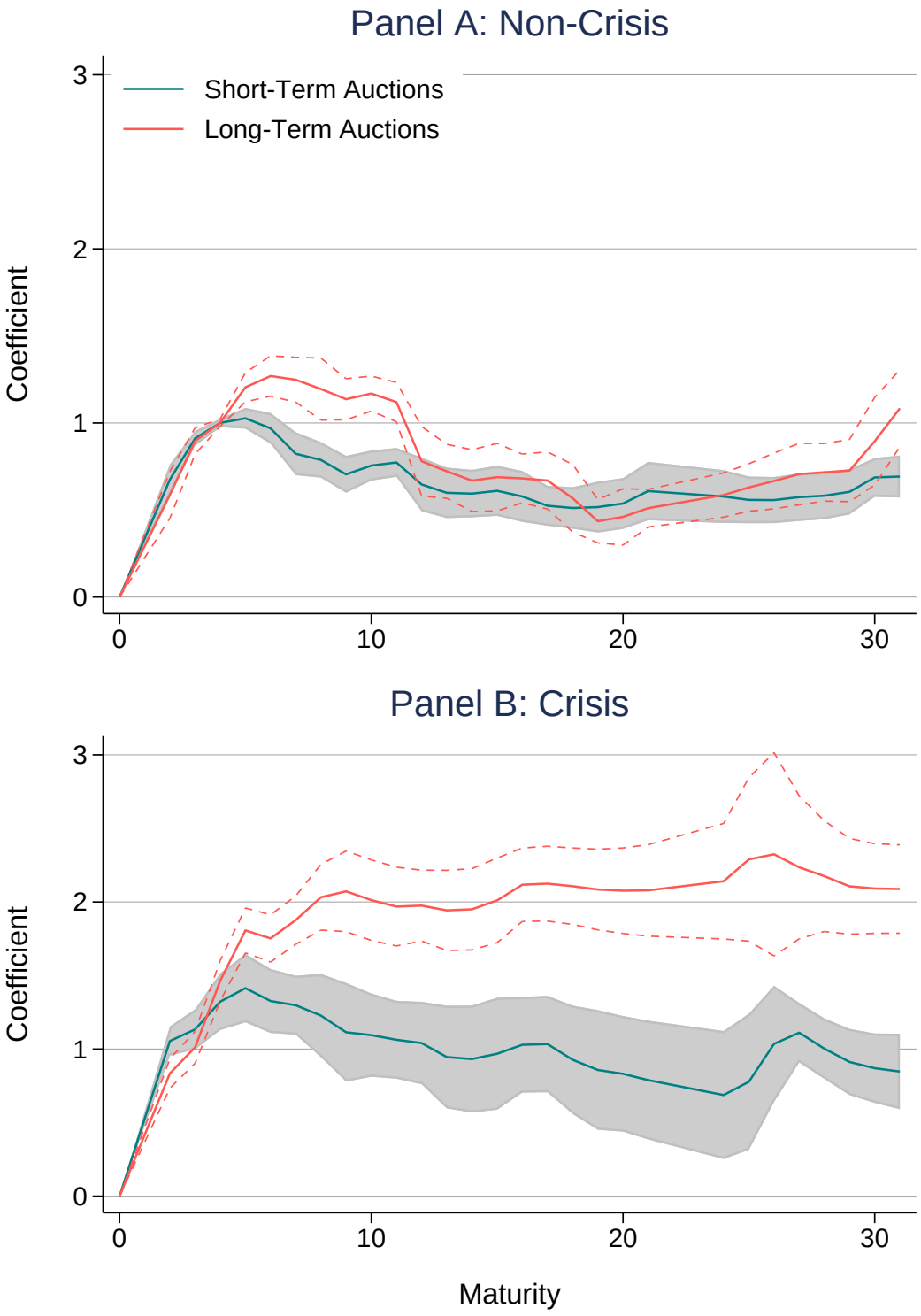

Figure B14: Localization Robustness: Market-Based Crises (Continuous)

Notes: Alternative specification of regression equation (4), where the crisis measure is treated as continuous (as measured by the aggregate capital ratio from He et al. (2016)). The measure is inverted and transformed to range from 0 to 1, and truncated so that values below 0.04 are set to 1 while values above 0.075 are set to 0.

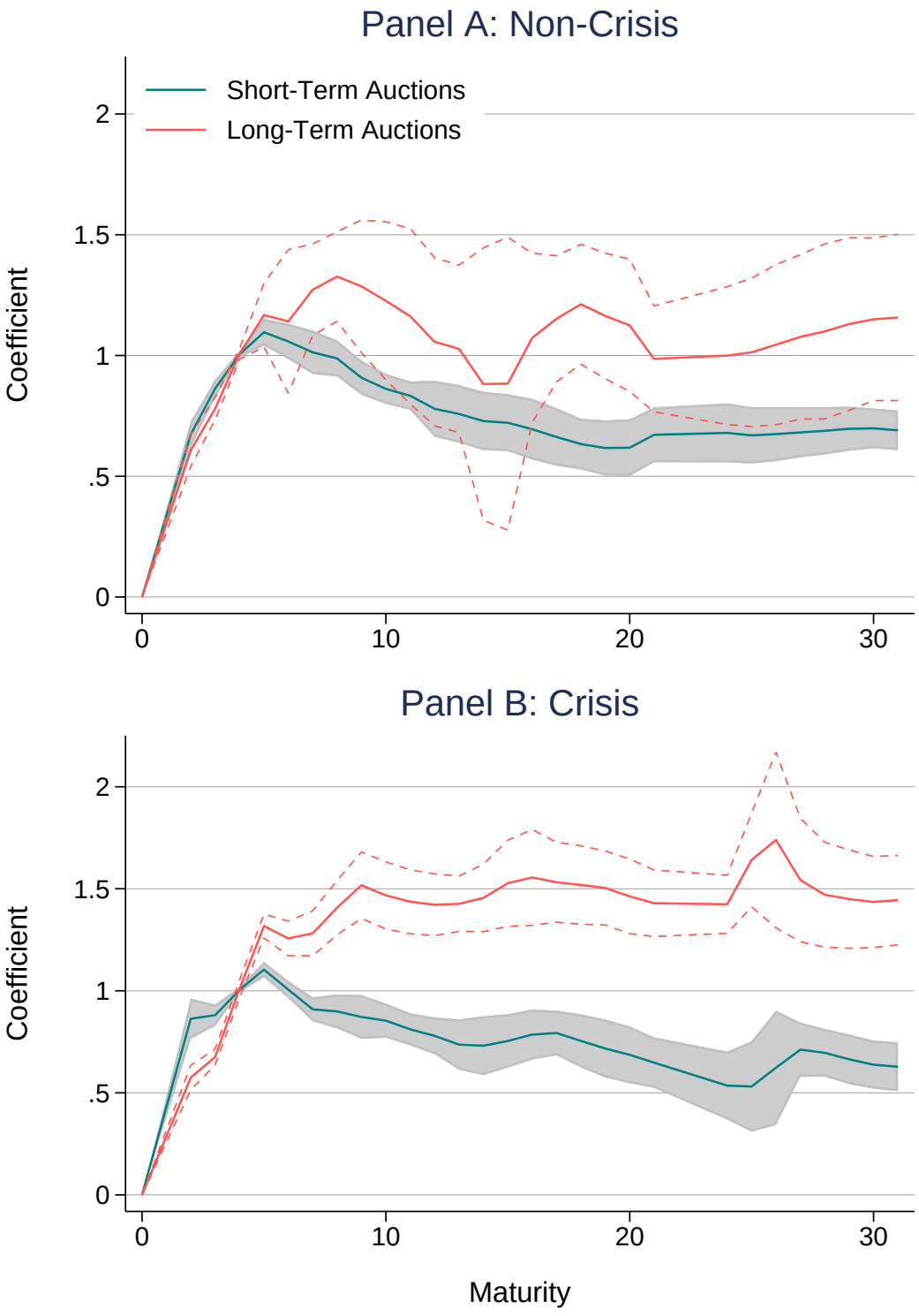

Figure B15: Localization Robustness: Narrative-Based Crises

Notes: Estimates of regression equation (4), identifying financial crises as periods identified by Romer and Romer (2017).

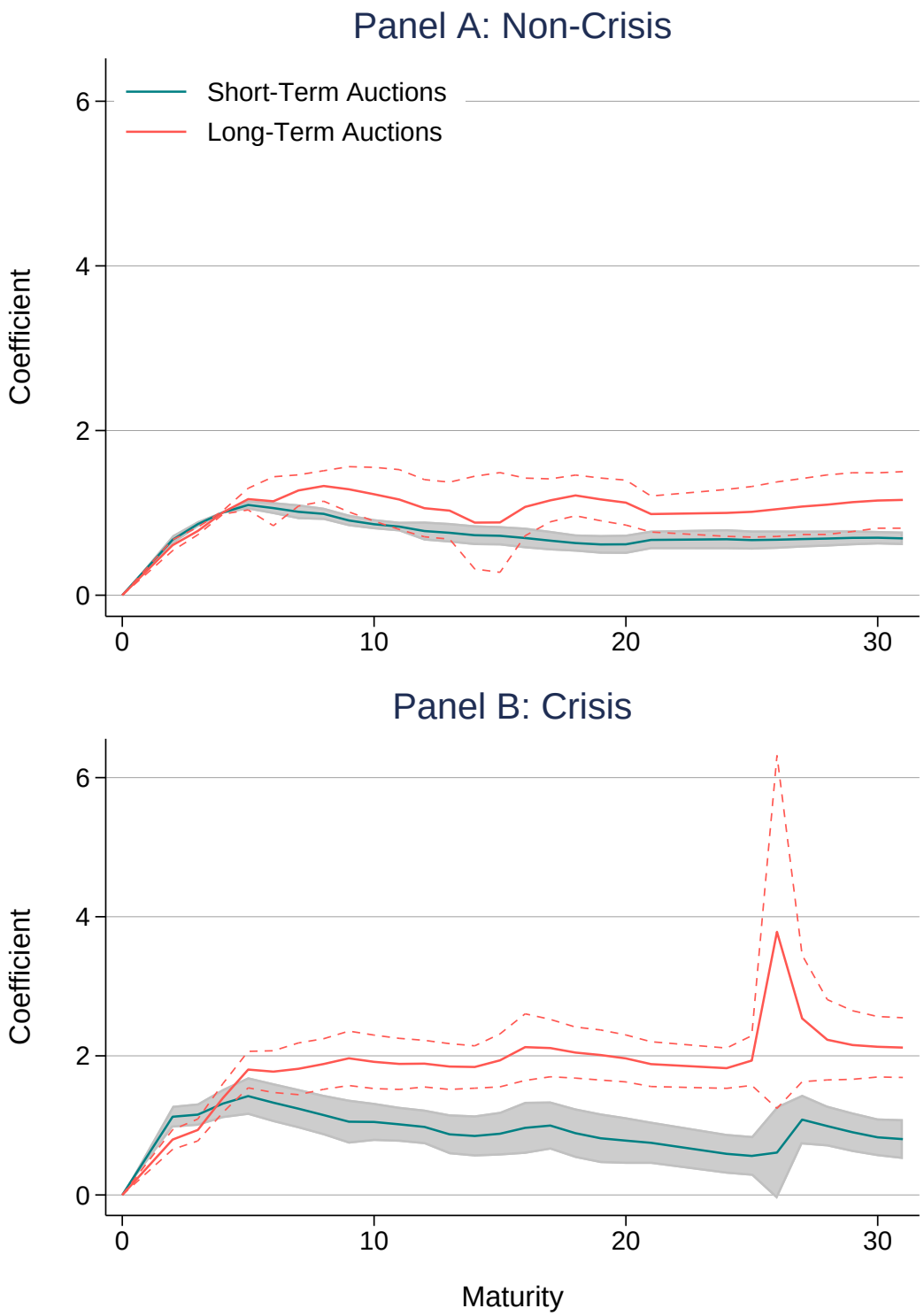

Figure B16: Localization Robustness: Narrative-Based Crises (Continuous)

Notes: Alternative specification of regression equation (4), where the crisis measure is treated as continuous (as measured by Romer and Romer (2017)). The measure is transformed to range from 0 to 1, and truncated so that values above 10 are set to 1.

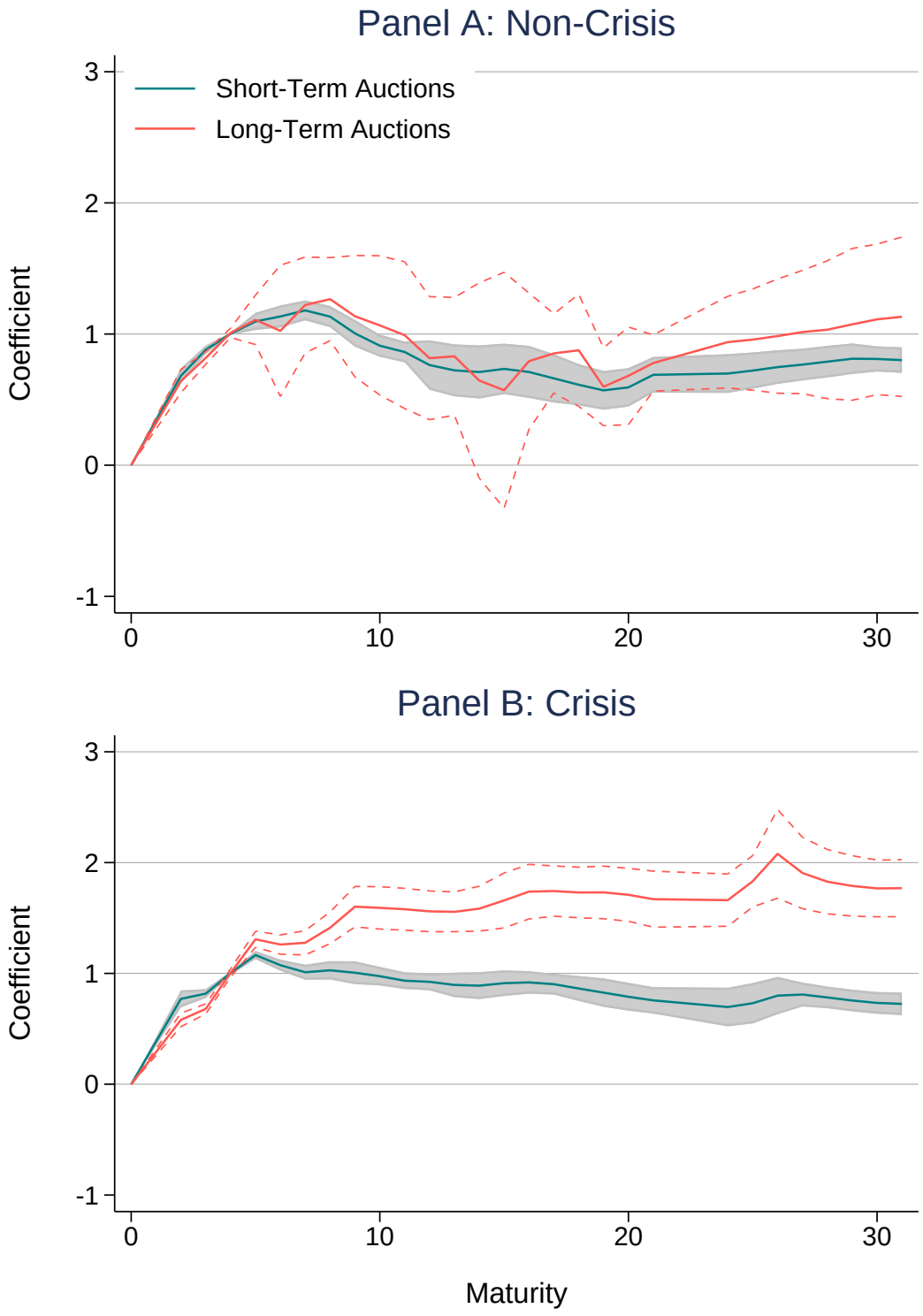

Figure B17: Localization Robustness: Long/Short Auctions

Notes: Estimates of regression equation (4), where short-maturity auctions are those with maturities 7 years or below, while long-maturity auctions are 10 years and above.

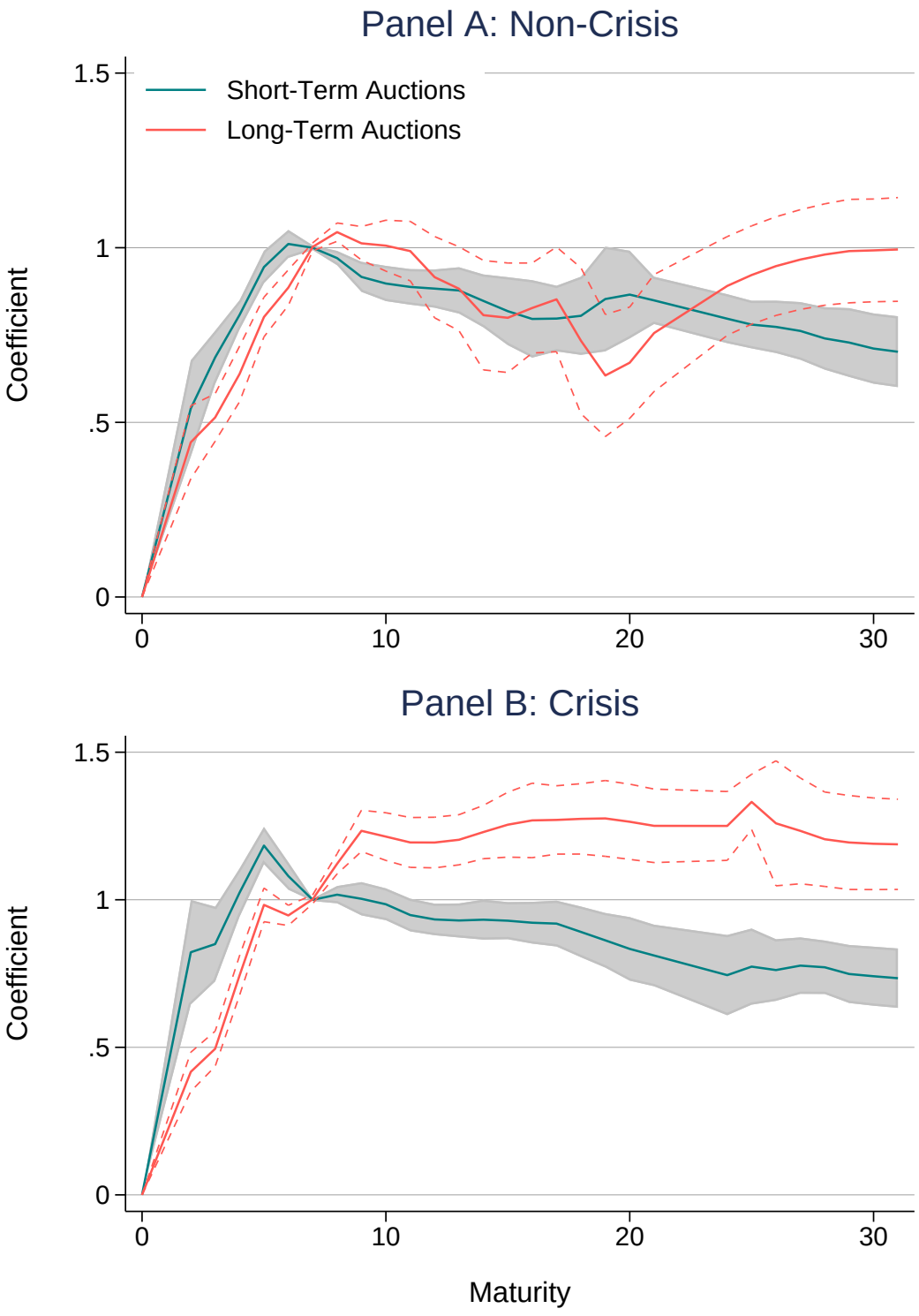

Figure B18: Localization Robustness: τ^*

Notes: Estimates of regression equation (4), where $\tau^* = 6$.

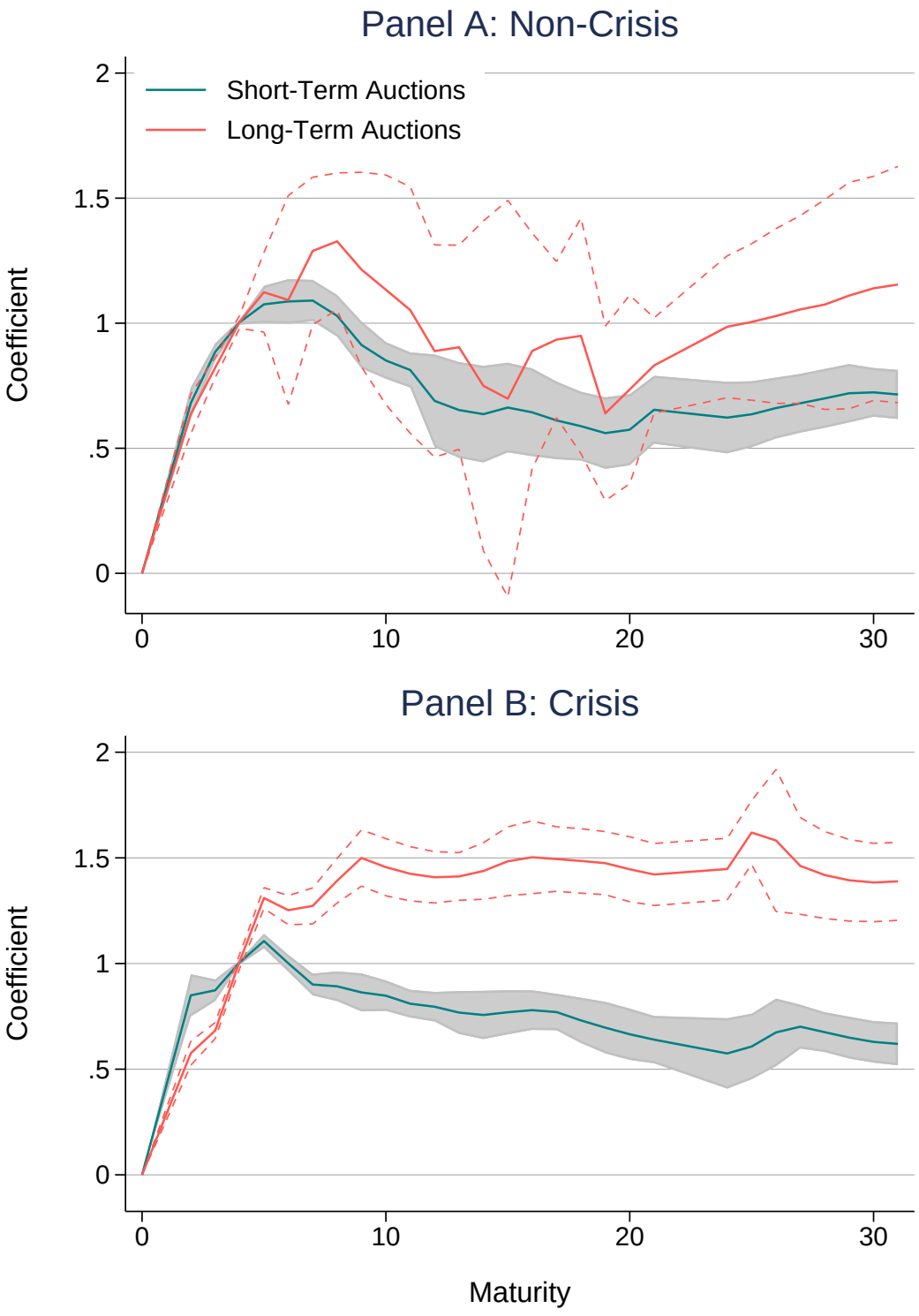

Figure B19: Localization Robustness: No QE Event Weeks

Notes: Estimates of regression equation (4), dropping auctions that occurred during the weeks of QE announcements.

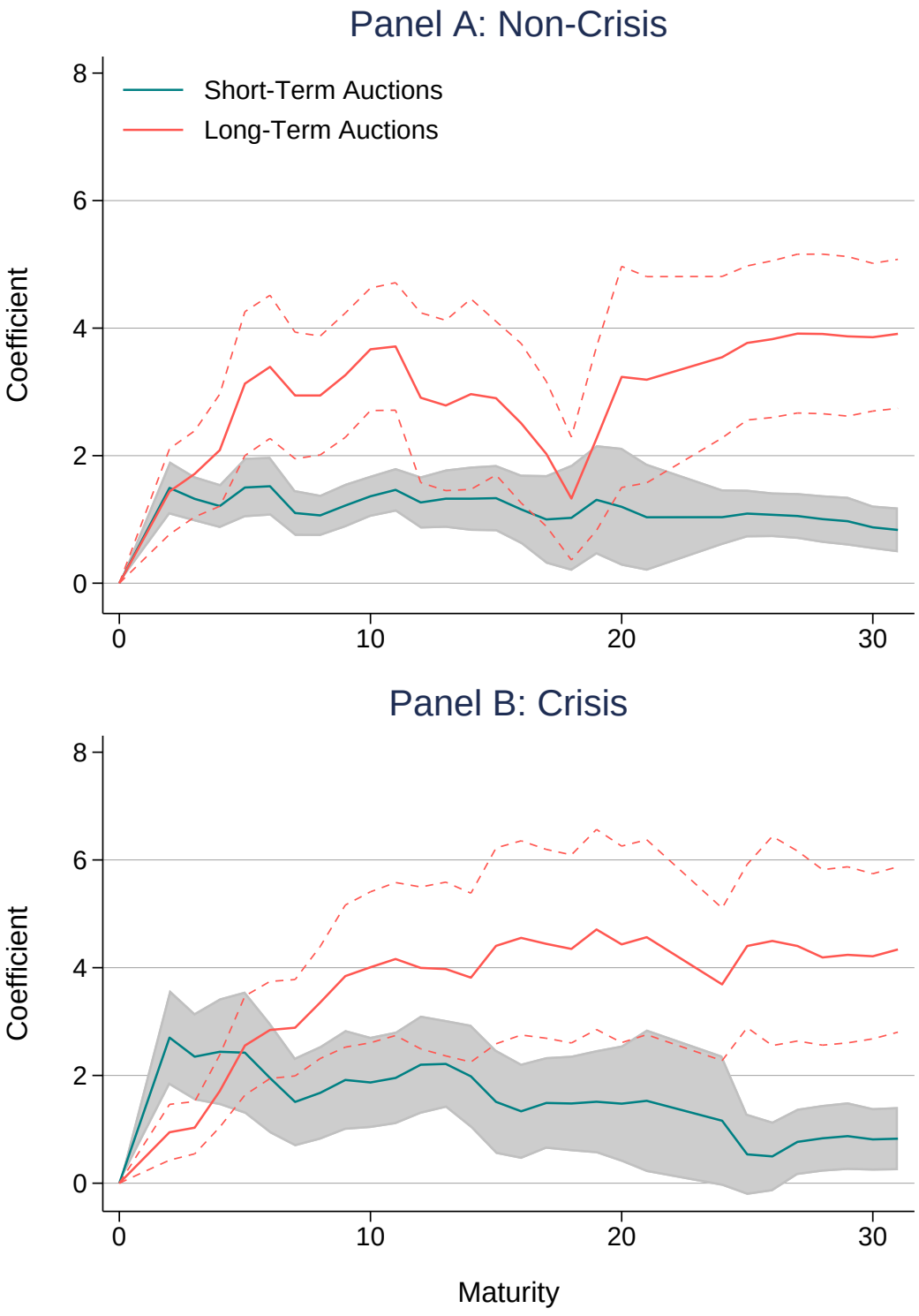

Figure B20: Alternative Localization: Bid-to-Cover (Levels)

Notes: Estimates of the alternative localization regression (5), using the bid-to-cover ratio (in levels).

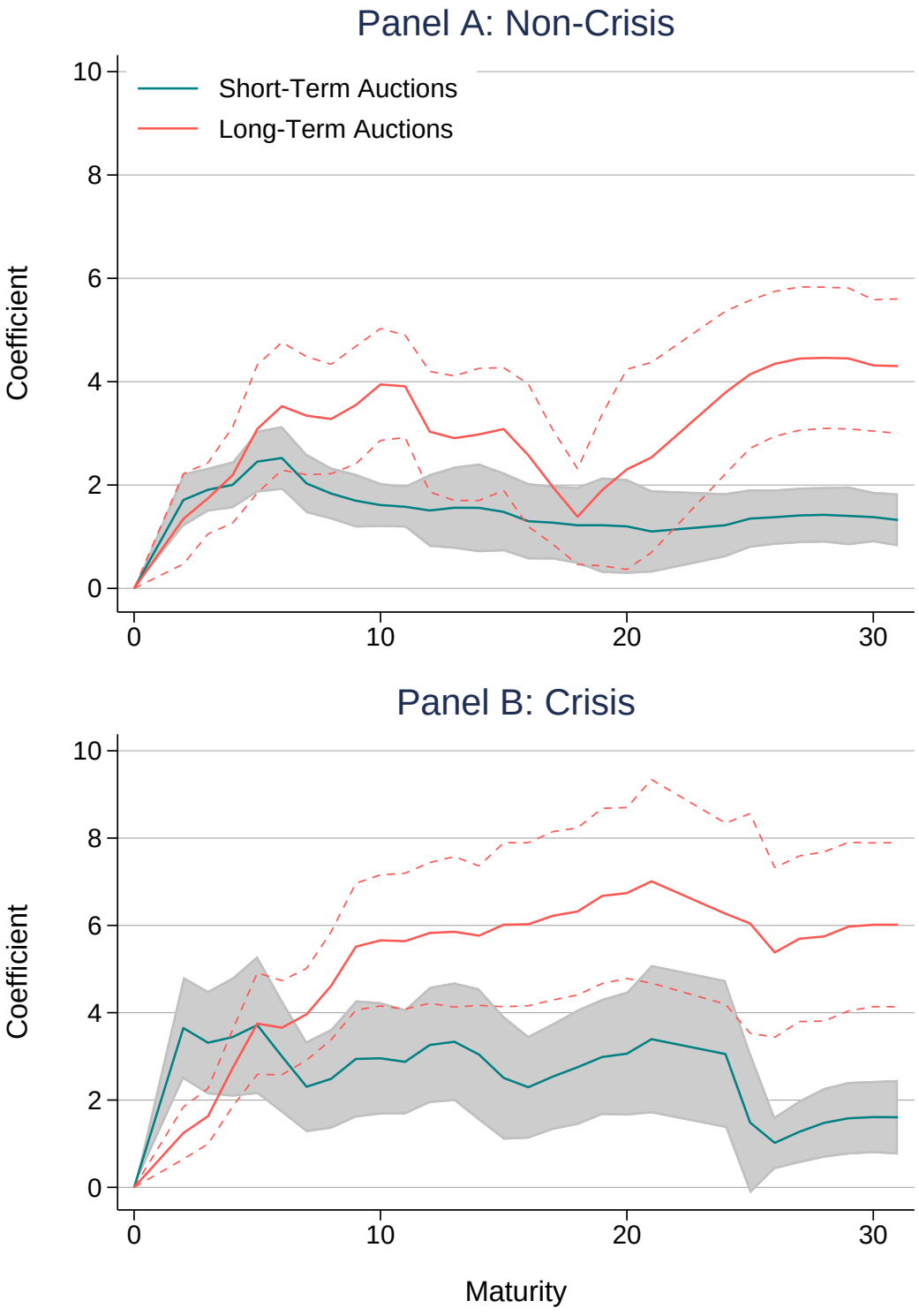

Figure B21: Alternative Localization: Bid-to-Cover (Residualized)

Notes: Estimates of the alternative localization regression (5), using the residualized bid-to-cover ratio (residuals from an AR(4) model).

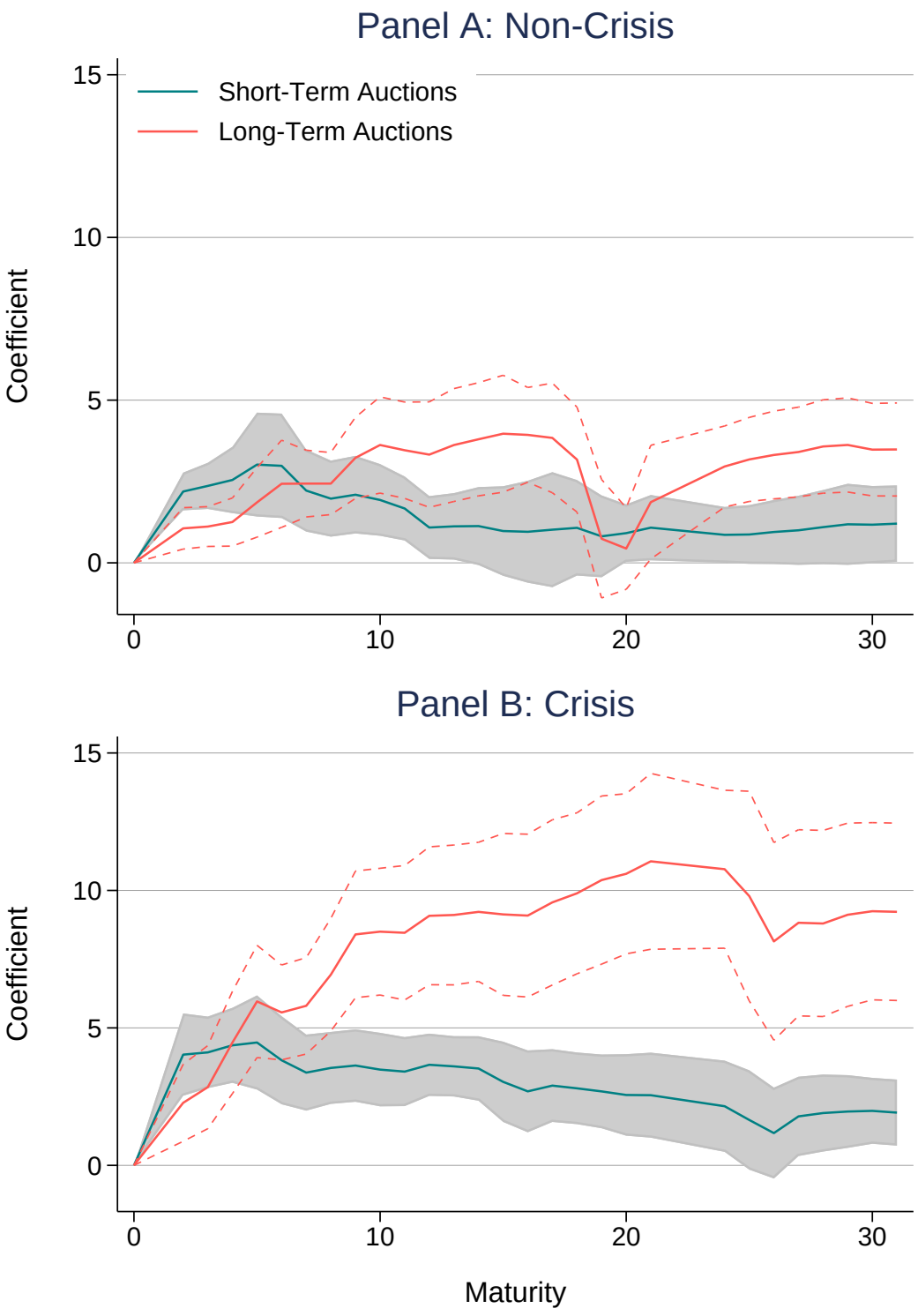

Figure B22: Alternative Localization: Indirect Bidders

Notes: Estimates of the alternative localization regression (5), using the bid-to-cover ratio of Indirect Bidders (in changes).

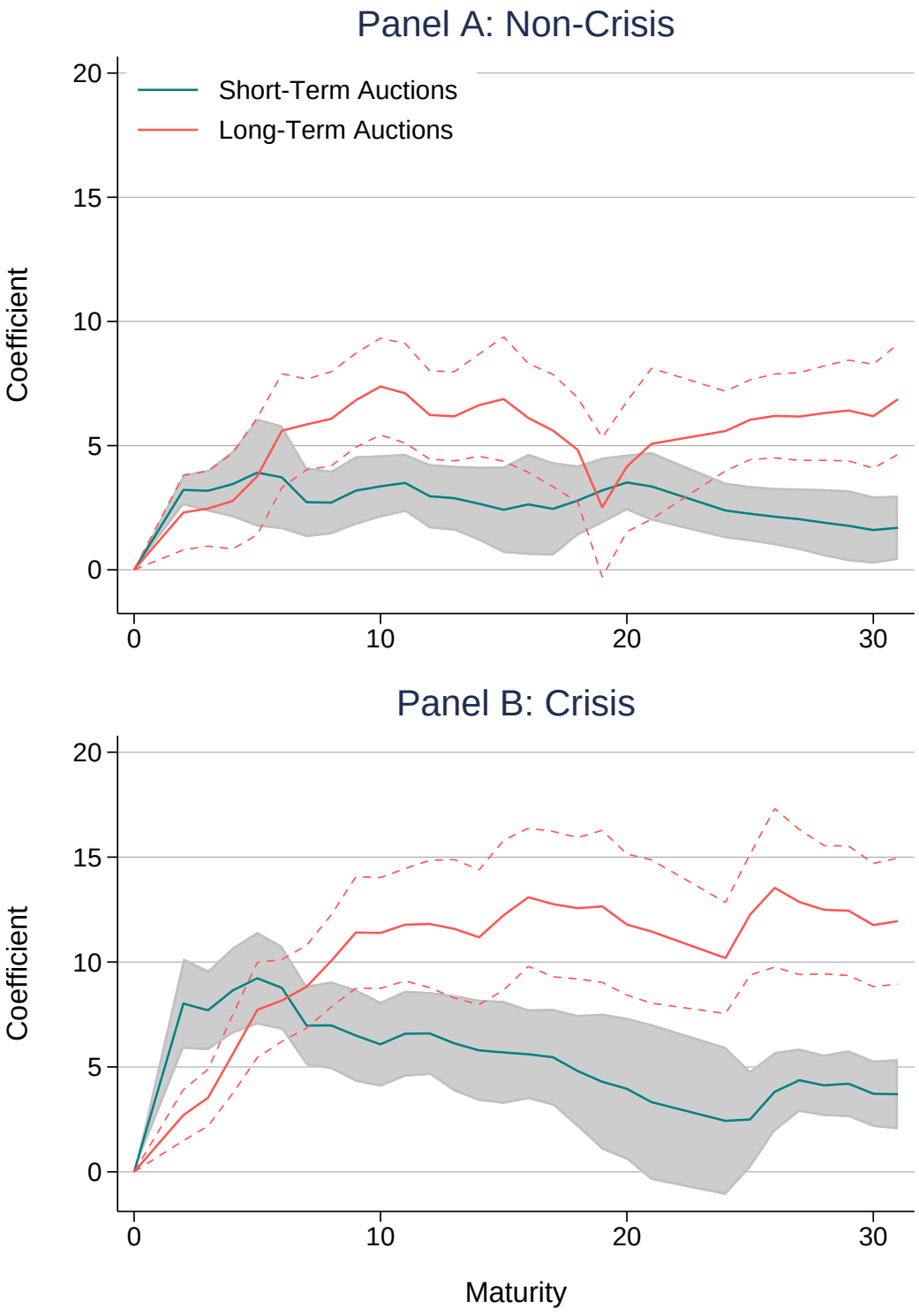

Figure B23: Alternative Localization: Indirect Bidders (Levels)

Notes: Estimates of the alternative localization regression (5), using the bid-to-cover ratio of Indirect Bidders (levels).

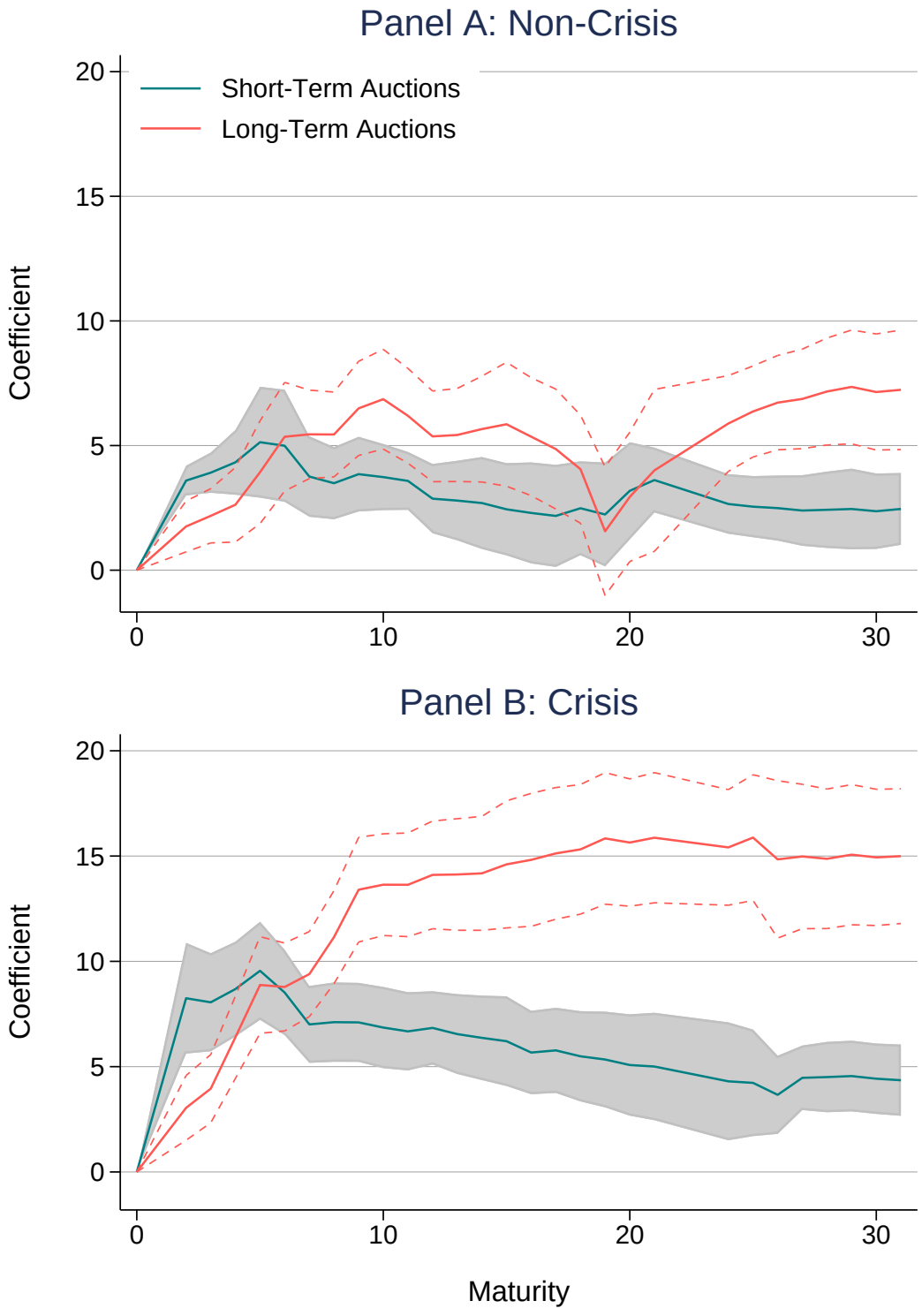

Figure B24: Alternative Localization: Indirect Bidders (Residualized)

Notes: Estimates of the alternative localization regression (5), using the residualized bid-to-cover ratio of Indirect Bidders (residuals from an AR(4) model).

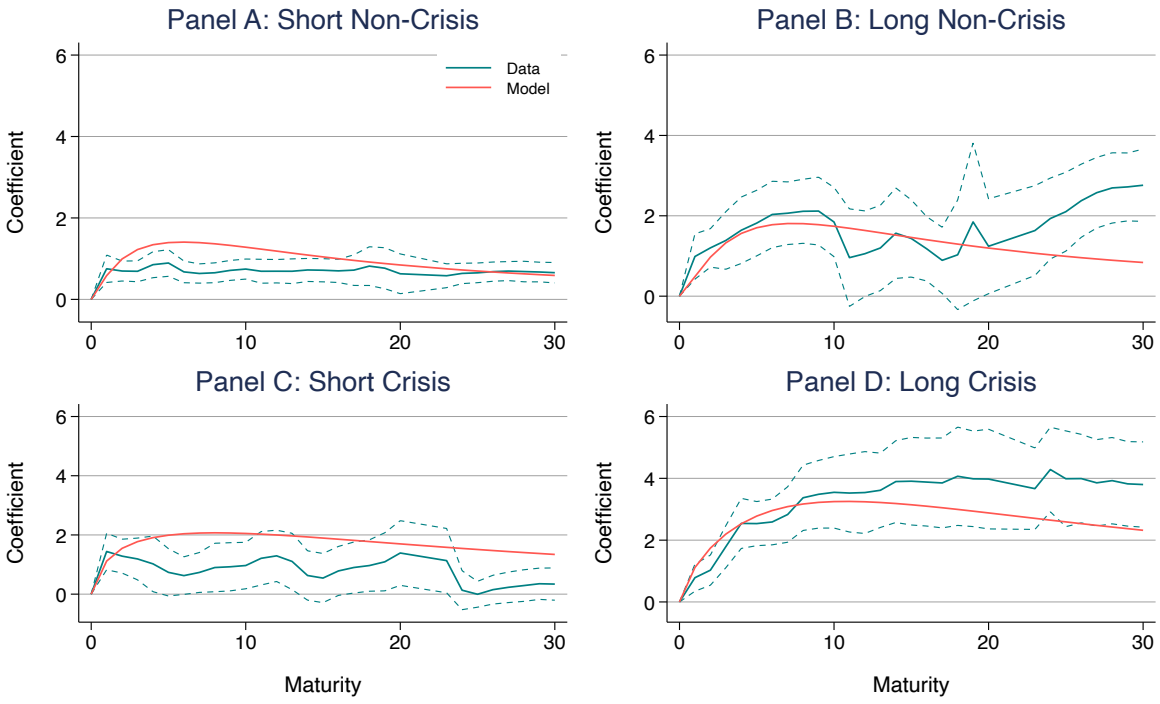

Figure B25: Model Fit: Alternative Localization

Notes: Plots of the model-implied and empirical coefficients from the alternative localization regression equation (5). Panels A and B plot the coefficients from the model calibrated to non-crisis periods, while Panels C and D plot the coefficients from the model calibrated to crisis periods.

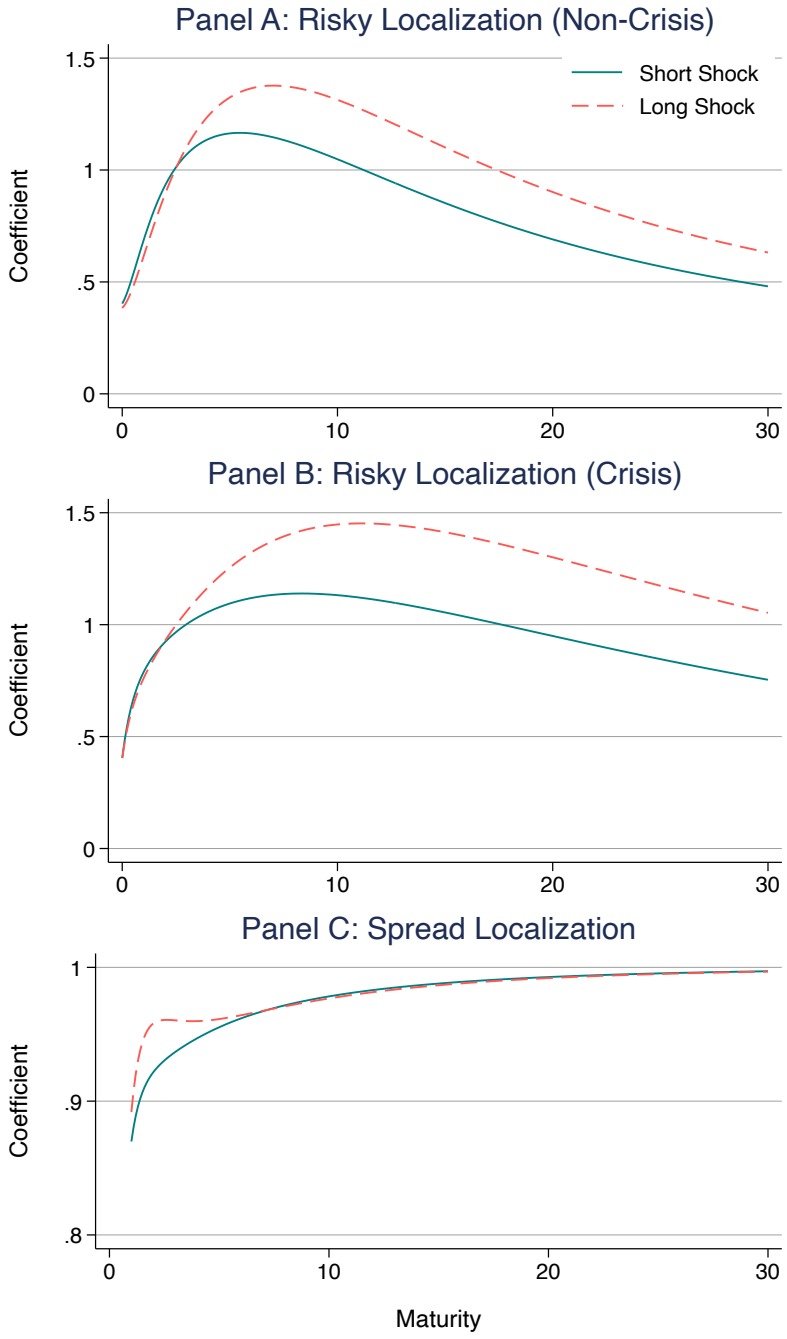

Figure B26: Model-Implied Risky Borrowing Localization

Notes: Plots of various measures of localization for risky borrowing rates in the model. Panels A and B plot the response of the risky yield curve $\tilde{y}_t^{(\tau)}$ to the short and long riskless demand factors $\beta_t^{(s)}, \beta_t^{(\ell)}$ from the model calibrated to non-crisis and crisis periods, respectively. The response is then scaled by the reaction of the riskless rate $y_t^{(\tau^*)}$ for $\tau^* = 3$, analogously to Panel B of Table 4 (although note the coefficients in the data are for log price changes of various ETFs, not yields). Hence this measures localization in maturity space for risky borrowing rates (rather than riskless rates) in the model. We see that the gap between the two relative responses widens in Panel B compared to Panel A, showing evidence of increased maturity localization in crisis periods, consistent with the results of Panel B in Table 4. Panel C reports another measure of risky localization in the model: we again compute the response of the risky yield curve $\tilde{y}_t^{(\tau)}$ to the short and long riskless demand factors, but this time scale the response by the reaction of the riskless rate $y_t^{(\tau)}$ for each corresponding maturity (rather than a fixed maturity τ^*). We then take the ratio of these responses in the model calibrated to the crisis periods relative to the non-crisis calibration. A value less than one therefore implies a weakened pass-through of riskless demand shocks to risky borrowing rates (relative to the same riskless borrowing rates) in the crisis vs. non-crisis calibration. In response to both short and long demand factors, Panel C shows that these responses always lies below one, consistent with our findings in Panel A of Table 4.

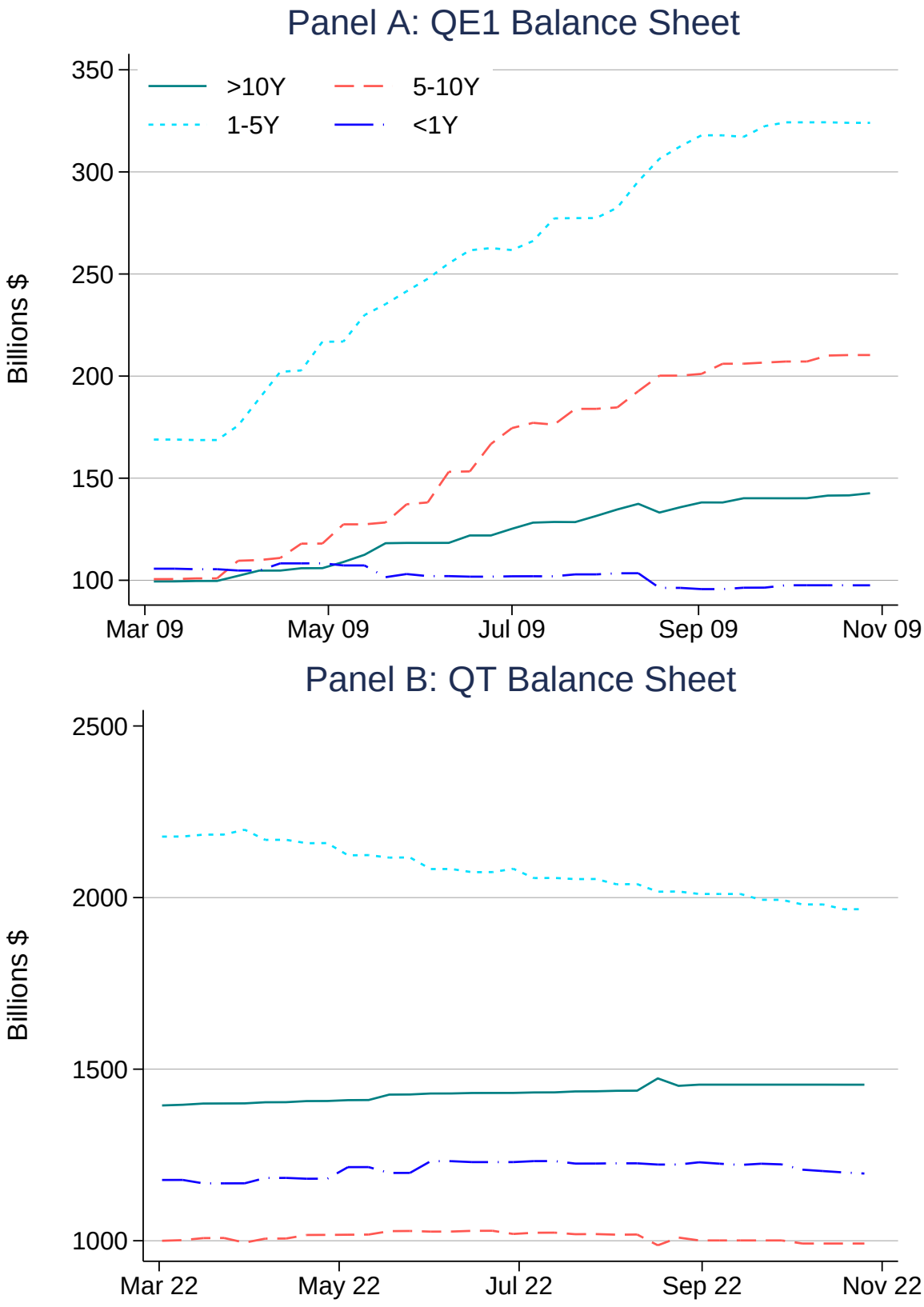

Figure B27: Fed Treasury Holdings: QE1 vs. QT

Notes: Changes in the Fed's holdings of Treasuries following QE1 (Panel A) and QT (Panel B). Following QE1, the Fed increased holdings of both intermediate- and long-term Treasuries. To date, the reduction in Treasury holdings following QT has concentrated on intermediate maturities.

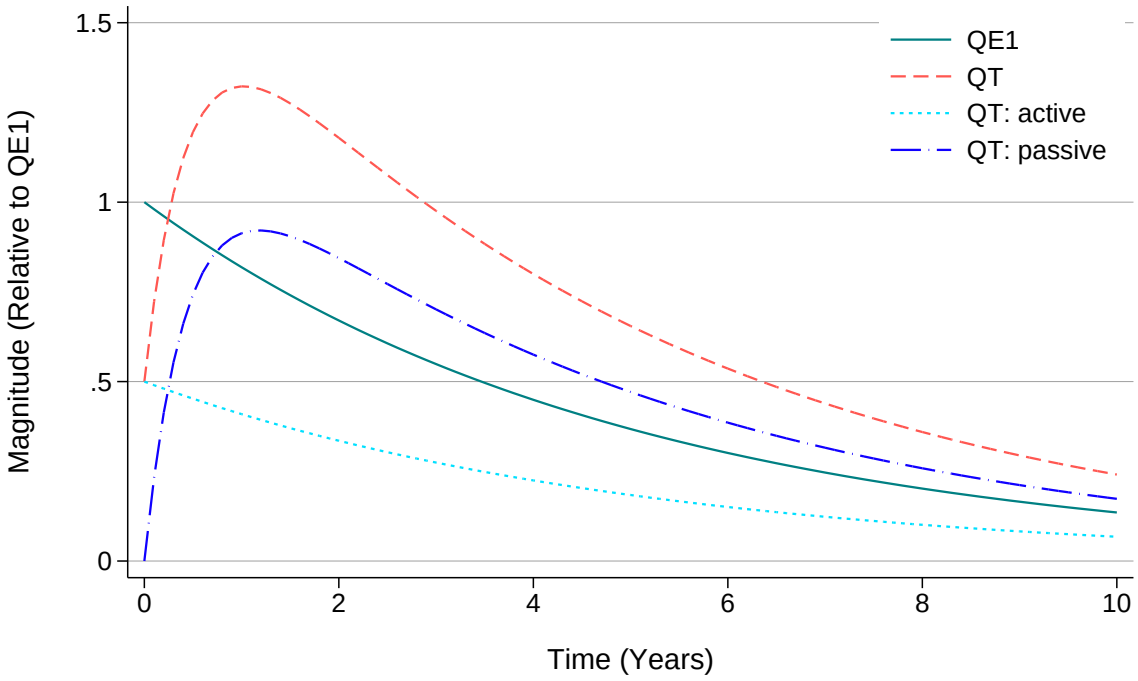

Figure B28: QE1 vs. QT Shocks

Notes: Comparison of the QE1 and QT shocks in the model (β_t^{QE}, solid line; and β_t^{QT}, dashed line). The magnitude is relative to the initial size of QE1 (and with signs flipped so that all numbers are positive). The dotted and dashed-dotted lines plot the contribution of the "active" and "passive" components of the QT shock.

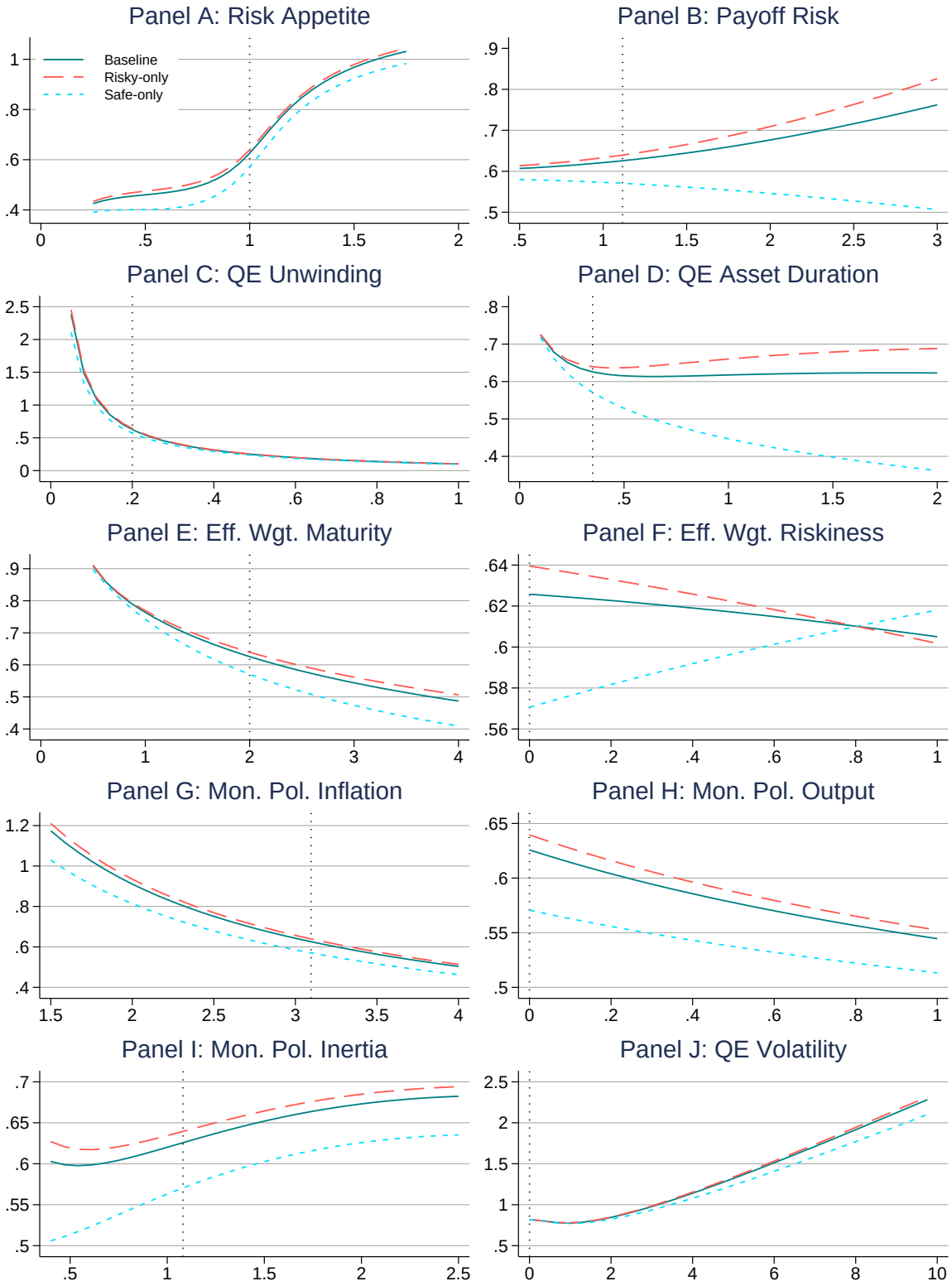

Figure B29: QE Sensitivity Analysis: Inflation

Notes: Response of inflation to a QE shock as a function of various model parameters. This figure conducts the same experiments as Figure 8, but plots the response of inflation π_t rather than the output gap x_t.

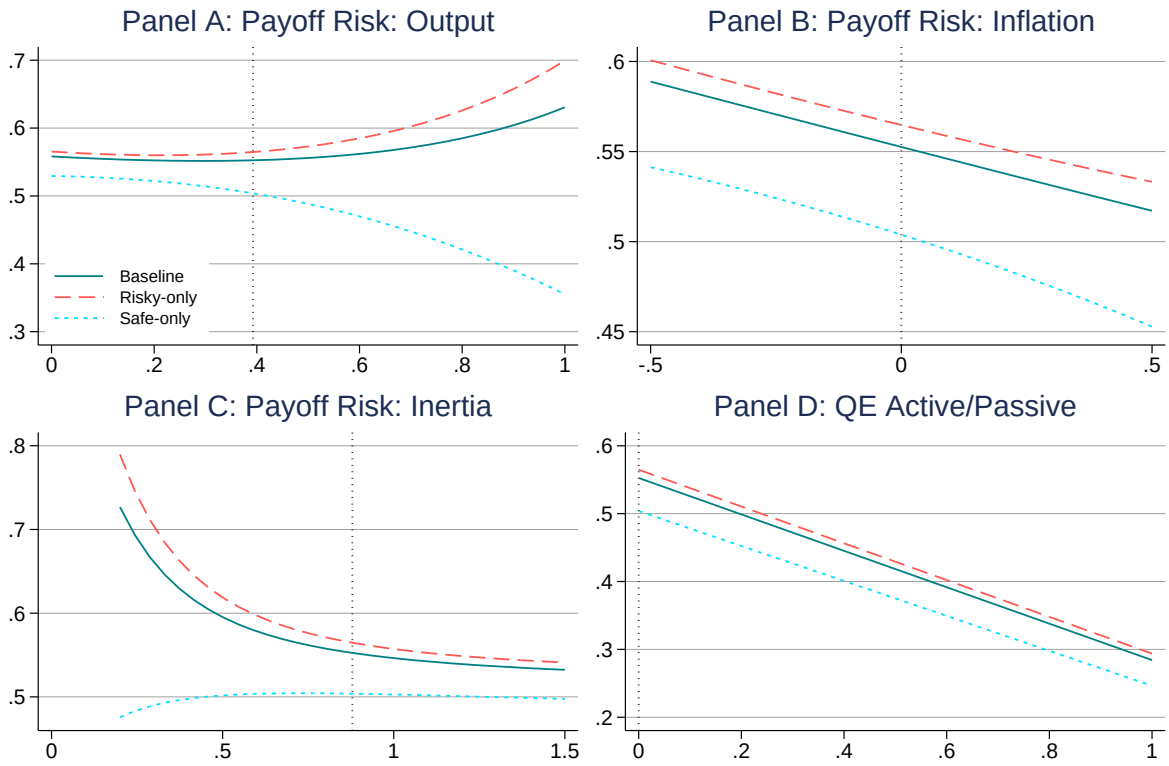

Figure B30: Additional QE Sensitivity Analysis: Output

Notes: Response of output to a QE shock as a function of additional model parameters (as in Figure 8). Panel A: payoff process d_t response to the output gap (ψ_x). Panel B: payoff process d_t response to inflation (ψ_π). Panel C: payoff process d_t inertia (κ_d). Panel D: active and passive QE weights (λ).

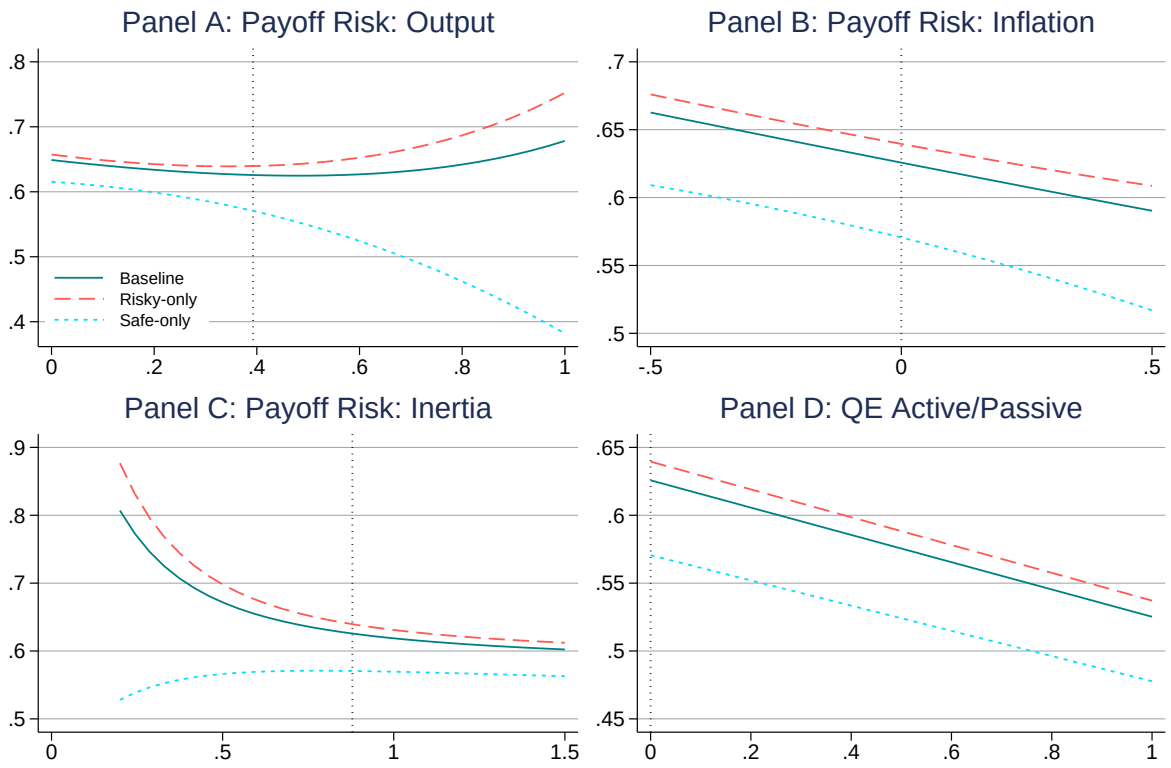

Figure B31: Additional QE Sensitivity Analysis: Inflation

Notes: Response of inflation to a QE shock as a function of additional model parameters. This figure conducts the same experiments as Figure B30, but plots the response of inflation π_t rather than the output gap x_t.

Table B1: Asset Price Reactions to Demand Shocks (IV Specification)

Asset type	Estimate	Std. Err.	Obs.	F-stat	Sample
	(1)	(2)	(3)	(4)	(5)
Panel A: Corporate and Private Debt					
LQD	-4.30***	(0.34)	830	48.8	2002-2017
HYG	0.33	(0.52)	678	36.3	2007-2017
MBB	-1.25***	(0.24)	662	37.3	2007-2017
VMBS	-1.45**	(0.69)	371	40.8	2009-2017
Corp. Aaa†	1.19***	(0.23)	1039	82.1	1995-2017
Corp. Baa†	1.05***	(0.22)	1039	82.1	1995-2017
Corp. C†	-0.18	(0.56)	973	68.0	1997-2017
Panel B: Equities					
SPY	0.87	(0.96)	1033	75.1	1995-2017
IWM	1.27	(1.09)	876	52.5	2000-2017
SP500†	5.99	(5.69)	973	79.8	1995-2016
Russell 2000†	5.49	(6.28)	973	79.8	1995-2016
Panel C: Swaps, Commodities, and Spreads					
GLD	-2.78***	(0.96)	775	38.2	2004-2017
GSCI†	-3.79	(4.66)	973	79.8	1995-2016
10Y Infl. Swap†	-0.18	(0.18)	724	38.6	2004-2016
2Y Infl. Swap†	-0.42	(0.32)	724	38.6	2004-2016
LIBOR-OIS†	0.23	(0.15)	737	41.4	2003-2016
Auto CDS†	-0.51	(5.48)	733	41.4	2004-2016
Bank CDS†	-0.07	(0.22)	733	41.4	2004-2016
VIX†	0.72	(8.09)	1039	82.1	1995-2017
Panel D: Federal Funds Futures					
1-Month Ahead†	0.06	(0.04)	1039	82.1	1995-2017
2-Month/1-Month Slope†	-0.04	(0.03)	1039	82.1	1995-2017

Notes: Results for IV regressions of asset price changes on demand shocks D_t. The table repeats the regressions from Table 3, but instruments the demand shocks D_t with the change in the bid-to-cover ratio. First-stage F-statistics are reported in column (3). Newey-West (9 lags) standard errors in parentheses. ***, **, * denote statistical significance at 1, 5 and 10 percent.

Appendix C Microfoundation Details

In a rational expectations equilibrium, the state variables $\mathbf{y}_t$ and non-predetermined (jump) variables $\mathbf{x}_t$ evolve according to

$$\mathrm{d}\mathbf{y}_t = \mathbf{h}(\mathbf{y}_t)\,\mathrm{d}t + \mathbf{r}(\mathbf{y}_t)\,\mathrm{d}\mathbf{B}_t \tag{C1}$$

$$\mathbf{x}_t \equiv \mathbf{x}(\mathbf{y}_t) \tag{C2}$$

where we have collected all Brownian terms into a vector $\mathbf{B}_t$. The functions $\mathbf{h}(\mathbf{y}_t), \mathbf{r}(\mathbf{y}_t)$ are of course endogenous, but taken as given by optimizing agents (households, firms, and arbitrageurs) in the model. The non-predetermined (jump) variables are a function of the state $\mathbf{x}(\mathbf{y}_t)$, and this endogenous mapping is also taken as given by agents.

Given the vector of Brownian components $\mathbf{B}_t$, we can decompose the dynamic processes of the model into usual drift and diffusion terms (determined in equilibrium). We use the following notation:

$$\frac{\mathrm{d}P_t^{(\tau)}}{P_t^{(\tau)}} = \mu_t^{(\tau)}\,\mathrm{d}t + \boldsymbol{\sigma}_t^{(\tau)}\,\mathrm{d}\mathbf{B}_t \tag{C3}$$

$$\frac{\mathrm{d}\tilde{P}_t^{(\tau)}}{\tilde{P}_t^{(\tau)}} = \tilde{\mu}_t^{(\tau)}\,\mathrm{d}t + \tilde{\boldsymbol{\sigma}}_t^{(\tau)}\,\mathrm{d}\mathbf{B}_t \tag{C4}$$

$$\mathrm{d}F_t \equiv \mu_t^F\,\mathrm{d}t + \boldsymbol{\sigma}_t^F\,\mathrm{d}\mathbf{B}_t \tag{C5}$$

Equations (C3) and (C4) give the dynamics of the realized returns for safe and risky assets, respectively. From these expressions, we also can write the effective borrowing rate faced by households as

$$\begin{aligned}\int_0^T \left(\eta(\tau)\frac{\mathrm{d}P_t^{(\tau)}}{P_t^{(\tau)}} + \tilde{\eta}(\tau)\frac{\mathrm{d}\tilde{P}_t^{(\tau)}}{\tilde{P}_t^{(\tau)}}\right)\mathrm{d}\tau &= \int_0^T \left(\eta(\tau)\mu_t^{(\tau)} + \tilde{\eta}(\tau)\tilde{\mu}_t^{(\tau)}\right)\mathrm{d}\tau\,\mathrm{d}t \\ &\quad + \int_0^T \left(\eta(\tau)\boldsymbol{\sigma}_t^{(\tau)} + \tilde{\eta}(\tau)\tilde{\boldsymbol{\sigma}}_t^{(\tau)}\right)\mathrm{d}\tau\,\mathrm{d}\mathbf{B}_t \\ &\equiv \hat{r}_t\,\mathrm{d}t + \hat{\boldsymbol{\sigma}}_t\,\mathrm{d}\mathbf{B}_t \end{aligned} \tag{C6}$$

Finally, equation (C5) is the process for household transfers. All of the drift and diffusion terms will be functions of the state variables $\mathbf{y}_t$ in equilibrium, which we derive now.

C.1 Aggregation and Market Clearing

In a symmetric equilibrium, $Y_{j,t} = Y_t, P_{j,t} = P_t, N_{j,t} = N_t, \pi_{j,t} = \pi_t$. Note also that the price level is (locally) non-stochastic:

$$\mathrm{d}P_t = P_t \pi_t \, \mathrm{d}t$$

Because the price adjustment costs are paid to households, market clearing implies:

$$Y_t = C_t = N_t \tag{C7}$$

The fiscal authority passively accommodates the central bank: for instantaneous short-term bonds (with $\tau = 0$) the total supply ($S_t^{(0)}$) is determined endogenously such that the targeted policy rate i_t is the equilibrium outcome. We assume the fiscal authority simply balances the budget each period, paying interest i_t on short-term bond holdings and levying lump-sum taxes $P_t T_t^H$ on households. Thus we have $P_t T_t^H = -S_t^{(0)} i_t$. Note that the fiscal authority therefore does not supply any long-term bonds, so that $S_t^{(\tau)} = 0$ for $\tau > 0$. Thus, in a zero-net supply equilibrium for $\tau > 0$, asset markets clear when arbitrageurs and habitat funds take opposite positions:

$$X_t^{(\tau)} + Z_t^{(\tau)} = 0, \quad \tilde{X}_t^{(\tau)} + \tilde{Z}_t^{(\tau)} = 0$$

and for $\tau = 0$, arbitrageurs and τ-maturity habitat funds holdings of short-term bonds are given by:

$$\begin{aligned} W_t - \int_0^T X_t^{(\tau)} \, \mathrm{d}\tau \\ W_t^{(\tau)} - Z_t^{(\tau)} + \eta(\tau) A_t \\ \tilde{W}_t^{(\tau)} - \tilde{Z}_t^{(\tau)} + \tilde{\eta}(\tau) A_t \end{aligned}$$

Summing across arbitrageurs and habitat investors and imposing market clearing for $\tau > 0$ gives the total demand for short-term bonds:

$$S_t^{(0)} = W_t + \int_0^T W_t^{(\tau)} \, \mathrm{d}\tau + \int_0^T \tilde{W}_t^{(\tau)} \, \mathrm{d}\tau + A_t$$

Thus, the fiscal authority's per-period budget constraint is given by

$$P_t T_t^H = -\left[W_t + \int_0^T W_t^{(\tau)} \, \mathrm{d}\tau + \int_0^T \tilde{W}_t^{(\tau)} \, \mathrm{d}\tau + A_t \right] i_t$$

Nominal firm profits and price adjustment costs are given by:

$$\begin{aligned}\int_0^1 \Pi_{j,t}\,\mathrm{d}j + \int_0^1 \Theta(\pi_{j,t})\,\mathrm{d}j &= P_t \int_0^1 Y_{j,t}\left(1 - w_t - \frac{\theta}{2}\pi_{j,t}^2\right)\mathrm{d}j \\ &\quad + \int_0^1 P_{j,t}Y_{j,t}\frac{\theta}{2}\pi_{j,t}^2\,\mathrm{d}j \\ &= P_tY_t - \mathcal{W}_tY_t\end{aligned}$$

where $w_t = \frac{\mathcal{W}_t}{P_t}$ is the real wage.

Total transfers from the financial sector:

$$\begin{aligned}\mathrm{d}W_t + \int_0^T \left(\mathrm{d}W_t^{(\tau)} + \mathrm{d}\tilde{W}_t^{(\tau)}\right)\mathrm{d}\tau &= \left[W_t + \int_0^T \left(\mathrm{d}W_t^{(\tau)} + \mathrm{d}\tilde{W}_t^{(\tau)}\right)\mathrm{d}\tau + A_t\right] i_t\,\mathrm{d}t \\ &\quad - A_t \int_0^T \left(\eta(\tau)\frac{\mathrm{d}P_t^{(\tau)}}{P_t^{(\tau)}} + \tilde{\eta}(\tau)\frac{\mathrm{d}\tilde{P}_t^{(\tau)}}{\tilde{P}_t^{(\tau)}}\right)\mathrm{d}\tau\end{aligned}$$

Hence, total nominal transfers to the household sector are given by

$$\mathrm{d}F_t = [P_tY_t - \mathcal{W}_tY_t]\,\mathrm{d}t - A_t \int_0^T \left(\eta(\tau)\frac{\mathrm{d}P_t^{(\tau)}}{P_t^{(\tau)}} + \tilde{\eta}(\tau)\frac{\mathrm{d}\tilde{P}_t^{(\tau)}}{\tilde{P}_t^{(\tau)}}\right)\mathrm{d}\tau$$

Combining with the household wealth dynamics, we obtain:

$$\begin{aligned}\mathrm{d}A_t &= 0 \\ \mathrm{d}a_t &= -a_t\pi_t\,\mathrm{d}t\end{aligned}$$

Thus, if initial household wealth $A_0 = 0$, then $\mathrm{d}A_t = 0, \mathrm{d}a_t = 0 \implies A_t = 0, a_t = 0$ as well. In this case, we have $\boldsymbol{\sigma}_t^F = 0$, which the household takes as given. Moreover, if initial arbitrageur wealth $W_0 = 0$ and habitat fund wealth $W_0^{(\tau)} = 0, \tilde{W}_0^{(\tau)} = 0$, then we also have that the supply of short-term bonds $S_t^{(0)} = 0$.

C.2 Optimality Conditions

Households: In addition to the aggregate state variables given by equation (C1), household wealth is also an idiosyncratic state variable. Defining real wealth $a_t \equiv \frac{A_t}{P_t}$, the household budget constraint (7) and Ito's Lemma imply (dropping notational dependence on time t for convenience)

$$\mathrm{d}a = \left(wN - C + a(\hat{r} - \pi) + \frac{1}{P}\mu^F\right)\mathrm{d}t + \left(a\hat{\boldsymbol{\sigma}} + \boldsymbol{\sigma}^F\right)\mathrm{d}\mathbf{B}$$

where the real wage $w \equiv \frac{\mathcal{W}}{P}$ and we have defined the drift and diffusion terms from equations (C6) and (C6). From the household lifetime utility function and budget constraint given by equations (6) and (7), we form the household HJB equation for the value function $V(a; \mathbf{y}) \equiv V$:

$$\rho V = \max_{C,N} u(C, N, \Psi) + V_a \left(wN - C + a(\hat{r} - \pi) + \frac{1}{P}\mu^F \right) + V_{\mathbf{y}^\top} \mathbf{h}(\mathbf{y}) + S^H$$

$$S^H \equiv \frac{1}{2} \operatorname{Tr} \begin{bmatrix} a\hat{\boldsymbol{\sigma}} + \boldsymbol{\sigma}^F \\ \mathbf{r}(\mathbf{y}) \end{bmatrix}^\top \begin{bmatrix} V_{aa} & V_{a\mathbf{y}} \\ V_{a\mathbf{y}^\top} & V_{\mathbf{y}\mathbf{y}^\top} \end{bmatrix} \begin{bmatrix} a\hat{\boldsymbol{\sigma}} + \boldsymbol{\sigma}^F \\ \mathbf{r}(\mathbf{y}) \end{bmatrix}$$

where subscripts denote the gradient of a given function.

The optimality conditions are standard:

$$\Psi C^{-\varsigma} = V_a \tag{C8}$$

$$\Psi N^{\phi} = w V_a \tag{C9}$$

The envelope theorem combined with Ito's Lemma imply that the co-state V_a evolves according to

$$\begin{aligned} \mathrm{d}V_a &= (\rho + \pi - \hat{r})\, V_a \,\mathrm{d}t + \left(V_{aa}(a\hat{\boldsymbol{\sigma}} + \boldsymbol{\sigma}^F) + V_{a\mathbf{y}^\top} \mathbf{r}(\mathbf{y}) \right) \mathrm{d}\mathbf{B} \\ &= (\rho + \pi - \hat{r})\, V_a \,\mathrm{d}t + V_{a\mathbf{y}^\top} \mathbf{r}(\mathbf{y})\, \mathrm{d}\mathbf{B} \end{aligned} \tag{C10}$$

where the second line uses the conditions derived above that in equilibrium, $a\hat{\boldsymbol{\sigma}} + \boldsymbol{\sigma}^F = 0$.

Firms: Define the relative price of firm j as $\tilde{P}_j = \frac{P_j}{P}$. Then imposing CES market clearing, the firm problem can be written as

$$\begin{aligned} \max E_0 \int_0^\infty e^{-\rho t} Q Y \left(\tilde{P}_j^{1-\epsilon} - w \tilde{P}_j^{-\epsilon} - \frac{\theta}{2} \pi_j^2 \right) \mathrm{d}t \\ \text{s.t.} \quad \mathrm{d}\tilde{P}_j = \tilde{P}_j (\pi_j - \pi)\, \mathrm{d}t \end{aligned}$$

where since households own firms, $Q \equiv V_a$ is the real stochastic discount factor of the household.

We first solve the firm problem with no price adjustment costs and no fluctuations in desired markups. When $\theta = 0$ and $\epsilon = \bar{\epsilon}$, we have the standard solution

$$P_j^n = \frac{\bar{\epsilon}}{\bar{\epsilon} - 1} \mathcal{W}^n \implies w^n = \frac{\bar{\epsilon} - 1}{\bar{\epsilon}} \tag{C11}$$

where the superscript n denotes outcomes under flexible prices. Combined with the household optimality conditions (C8) and (C9) as well as market clearing conditions

(C7), we find that output under flexible prices is given by

$$Y^n = \left(\frac{\bar{\epsilon}-1}{\bar{\epsilon}}\right)^{\frac{1}{\varsigma+\phi}} \tag{C12}$$

When $\theta > 0$, we form the firm HJB equation for the value function $U(\tilde{P}_j; \mathbf{y}) \equiv U$:

$$\rho U = \max_{\tilde{\pi}_j} QY\left(\tilde{P}_j^{1-\epsilon} - w\tilde{P}_j^{-\epsilon} - \frac{\theta}{2}\pi_j^2\right) + U_{\tilde{P}_j}\tilde{P}_j(\pi_j - \pi) + U_{\mathbf{y}^\top}\mathbf{h}(\mathbf{y}) + S^F$$
$$S^F \equiv \frac{1}{2}\operatorname{Tr}\begin{bmatrix}0\\ \mathbf{r}(\mathbf{y})\end{bmatrix}^\top \begin{bmatrix}U_{\tilde{P}_j\tilde{P}_j} & U_{\tilde{P}_j\mathbf{y}}\\ U_{\tilde{P}_j\mathbf{y}^\top} & U_{\mathbf{y}\mathbf{y}^\top}\end{bmatrix}\begin{bmatrix}0\\ \mathbf{r}(\mathbf{y})\end{bmatrix}$$

The firm optimality conditions are

$$\theta QY\pi_j = U_{\tilde{P}_j}\tilde{P}_j \tag{C13}$$

The envelope theorem combined with Ito's Lemma imply that the co-state $U_{\tilde{P}_j}$ evolves according to

$$\mathrm{d}U_{\tilde{P}_j} = \left((\rho - (\pi_j - \pi)U_{\tilde{P}_j})U_{\tilde{P}_j} - QY\left[(1-\epsilon) + \epsilon w\tilde{P}_j^{-1}\right]\tilde{P}_j^{-\epsilon}\right)\mathrm{d}t + U_{\tilde{P}_j\mathbf{y}^\top}\mathbf{r}(\mathbf{y})\,\mathrm{d}\mathbf{B}$$

In a symmetric equilibrium, all firms make the same choices so that $\pi_j = \pi$ and $\tilde{P}_j = 1$. We therefore have

$$\mathrm{d}U_{\tilde{P}_j} = QY\left(\rho\theta\pi - [(1-\epsilon) + \epsilon w]\right)\mathrm{d}t + U_{\tilde{P}_j\mathbf{y}^\top}\mathbf{r}(\mathbf{y})\,\mathrm{d}\mathbf{B} \tag{C14}$$

Arbitrageurs: Using equations (C3) and (C4), we rewrite the arbitrageur budget constraint (11) as

$$\mathrm{d}W = \left[Wi + \int_0^T \left(X^{(\tau)}(\mu^{(\tau)} - i) + \tilde{X}^{(\tau)}(\tilde{\mu}^{(\tau)} - i)\right)\mathrm{d}\tau\right]\mathrm{d}t$$
$$+ \left[\int_0^T \left(X^{(\tau)}\boldsymbol{\sigma}^{(\tau)} + \tilde{X}^{(\tau)}\tilde{\boldsymbol{\sigma}}^{(\tau)}\right)\mathrm{d}\tau\right]\mathrm{d}\mathbf{B}$$

The optimality conditions with respect to holdings $X^{(\tau')}, \tilde{X}^{(\tau')}$ are given by

$$\frac{\partial E\,\mathrm{d}W}{\partial X^{(\tau')}} = \frac{a}{2}\frac{\partial Var\,\mathrm{d}W}{\partial X^{(\tau')}}$$
$$\frac{\partial E\,\mathrm{d}W}{\partial \tilde{X}^{(\tau')}} = \frac{a}{2}\frac{\partial Var\,\mathrm{d}W}{\partial \tilde{X}^{(\tau')}}$$

Thus, the optimality conditions with respect to holdings $X^{(\tau')}$ imply

$$\frac{\partial E\,\mathrm{d}W}{\partial X^{(\tau')}} = \left[\mu^{(\tau')} - i\right]\mathrm{d}t\,, \quad \frac{\partial Var\,\mathrm{d}W}{\partial X^{(\tau')}} = 2\int_0^T \left(X^{(\tau)}\boldsymbol{\sigma}^{(\tau)} + \tilde{X}^{(\tau)}\tilde{\boldsymbol{\sigma}}^{(\tau)}\right)\mathrm{d}\tau\left[\boldsymbol{\sigma}^{(\tau')}\right]^\top \mathrm{d}t$$
$$\implies \mu^{(\tau')} - i = a\int_0^T \left(X^{(\tau)}\boldsymbol{\sigma}^{(\tau)} + \tilde{X}^{(\tau)}\tilde{\boldsymbol{\sigma}}^{(\tau)}\right)\mathrm{d}\tau\left[\boldsymbol{\sigma}^{(\tau')}\right]^\top$$

and analogously the optimality conditions with respect to holdings $\tilde{X}^{(\tau')}$ imply

$$\tilde{\mu}^{(\tau')} - i = a\int_0^T \left(X^{(\tau)}\boldsymbol{\sigma}^{(\tau)} + \tilde{X}^{(\tau)}\tilde{\boldsymbol{\sigma}}^{(\tau)}\right)\mathrm{d}\tau\left[\tilde{\boldsymbol{\sigma}}^{(\tau')}\right]^\top$$

C.3 Steady State

The deterministic steady state is found by setting the diffusion terms $\mathbf{r}(\mathbf{y}) = 0$. We focus on the deterministic steady state with zero inflation and where output is at potential: $\bar{\pi} = 0, \bar{Y} = \bar{Y}^n$. Combining market clearing conditions, household optimality conditions (C8) and (C9), and firm conditions (C11) we find

$$\bar{w} = \left(\frac{\bar{\epsilon} - 1}{\bar{\epsilon}}\right), \quad \bar{Y} = \bar{w}^{\frac{1}{\varsigma+\phi}}$$

We further assume that the habitat demand in the steady state is zero:

$$\overline{Z^{(\tau)}} = -\alpha(\tau)\log\overline{P^{(\tau)}} - \overline{\beta(\tau)} = 0$$
$$\overline{\tilde{Z}^{(\tau)}} = -\tilde{\alpha}(\tau)\log\overline{\tilde{P}^{(\tau)}} - \overline{\tilde{\beta}(\tau)} = 0$$

Hence, arbitrageur optimality conditions imply:

$$\overline{\mu^{(\tau')}} = \bar{r}, \quad \overline{\tilde{\mu}^{(\tau')}} = \bar{r}$$

Thus from equation (C6), we have

$$\bar{\hat{r}} = \int_0^T \left(\eta(\tau)\overline{\mu^{(\tau')}} + \tilde{\eta}(\tau)\overline{\tilde{\mu}^{(\tau')}}\right)\mathrm{d}\tau$$
$$= \bar{r}\int_0^T \left(\eta(\tau) + \tilde{\eta}(\tau)\right)\mathrm{d}\tau = \bar{r}$$

Hence we must have that the long-run central bank target is given by

$$\bar{r} = \rho$$

so that the dynamics of the household co-state (C10) are consistent with a zero output gap, zero inflation deterministic steady.

C.4 Log-linearization

From the household and firm optimality conditions (along with market clearing conditions), we have that

$$c = y \equiv \log Y = \varsigma^{-1}(-\log V_a + \log \Psi), \quad w = Y^{\varsigma+\phi}$$
$$\pi = \frac{1}{\theta} U_{\tilde{P}_j} \frac{1}{QY}, \quad QY = V_a^{1-\varsigma^{-1}} \Psi^{\varsigma^{-1}}$$

Denote the dynamics of the discount factor shock by

$$\mathrm{d}\Psi = \mu^{\Psi}\,\mathrm{d}t + \boldsymbol{\sigma}^{\Psi}\,\mathrm{d}B$$

Then using the household co-state equation (C10), we can derive the dynamics of the output gap $x \equiv \log\left(\frac{Y}{Y^n}\right)$:

$$E\,\mathrm{d}x = \varsigma^{-1}\left((\hat{r} - \pi - \rho) - \frac{1}{\Psi}\mu^{\Psi} + S^x\right)\mathrm{d}t$$
$$S^x \equiv \frac{1}{2}\,\mathrm{Tr}\begin{bmatrix} V_{a\mathbf{y}^\top}\mathbf{r} \\ \boldsymbol{\sigma}^{\Psi} \end{bmatrix}^\top \begin{bmatrix} \varsigma^{-1}V_a^{-2} & 0 \\ 0 & \varsigma^{-1}\Psi^{-2} \end{bmatrix} \begin{bmatrix} V_{a\mathbf{y}^\top}\mathbf{r} \\ \boldsymbol{\sigma}^{\Psi} \end{bmatrix}$$

We log-linearize around the deterministic steady state ($\mathbf{y} = \bar{\mathbf{y}}$ derived above and no risk $\mathbf{r}(\mathbf{y}) = 0$). Note that to a first-order, the term $S^x \approx 0$. Hence we have

$$E\,\mathrm{d}x \approx \varsigma^{-1}\left(\hat{r} - \pi - \bar{r} - z_x\right)\mathrm{d}t$$

where the aggregate demand shock is defined as $z_x \equiv \frac{1}{\Psi}\mu^{\Psi}$.

Using the firm co-state dynamics (C14), we find the dynamics of inflation:

$$E\,\mathrm{d}\pi = \Bigg\{\rho\pi - \frac{1}{\theta}\left[(1-\epsilon) + \epsilon e^{x(\varsigma+\phi)}\frac{\bar{\epsilon}-1}{\bar{\epsilon}}\right]$$
$$- \pi\left[(1-\varsigma^{-1})(\rho + \pi - \hat{r}) - \varsigma^{-1}\Psi^{-1}\mu^{\Psi} + S^{(QY)}\right] + S^{\pi}\Bigg\}\mathrm{d}t$$
$$S^{\pi} = \frac{1}{2}\,\mathrm{Tr}\begin{bmatrix} U_{\tilde{P}_j\mathbf{y}^\top}\mathbf{r} \\ \boldsymbol{\sigma}^{(QY)} \end{bmatrix}^\top \begin{bmatrix} 0 & -\frac{1}{\theta}(QY)^{-2} \\ -\frac{1}{\theta}(QY)^{-2} & 2\pi(QY)^{-2} \end{bmatrix} \begin{bmatrix} U_{\tilde{P}_j\mathbf{y}^\top}\mathbf{r} \\ \boldsymbol{\sigma}^{(QY)} \end{bmatrix}$$
$$S^{(QY)} \equiv \frac{1}{2}(QY)\,\mathrm{Tr}\begin{bmatrix} V_{a\mathbf{y}^\top}\mathbf{r} \\ \boldsymbol{\sigma}^{\Psi} \end{bmatrix}^\top \begin{bmatrix} \varsigma^{-1}(\varsigma^{-1}-1)V_a^{-2} & \varsigma^{-1}(1-\varsigma^{-1})V_a^{-1}\Psi^{-1} \\ \varsigma^{-1}(1-\varsigma^{-1})V_a^{-1}\Psi^{-1} & \varsigma^{-1}(\varsigma^{-1}-1)\Psi^{-2} \end{bmatrix} \begin{bmatrix} V_{a\mathbf{y}^\top}\mathbf{r} \\ \boldsymbol{\sigma}^{\Psi} \end{bmatrix}$$
$$\boldsymbol{\sigma}^{(QY)} \equiv (1-\varsigma^{-1})(QY)V_{a\mathbf{y}^\top}\mathbf{r} - \varsigma^{-1}(QY)\Psi^{-1}\boldsymbol{\sigma}^{\Psi}$$

To a first-order, the term $S^{\pi} \approx 0$. Further, a first-order approximation gives

$$\begin{aligned}
\pi(1-\varsigma^{-1})\left(\rho+\pi-\hat{r}\right) &\approx 0 \\
\pi\varsigma^{-1}\Psi^{-1}\mu^{\Psi} &\approx 0 \\
\pi S^{(QY)} &\approx 0 \\
(1-\epsilon)+\epsilon e^{x(\varsigma+\phi)}\frac{\bar{\epsilon}-1}{\bar{\epsilon}} &\approx (\bar{\epsilon}-1)(\varsigma+\phi)x-\frac{\epsilon-\bar{\epsilon}}{\bar{\epsilon}}
\end{aligned}$$

Then we have

$$E\,\mathrm{d}\pi \approx (\rho\pi-\delta x-z_{\pi})\,\mathrm{d}t$$

where the aggregate cost-push shock is defined as $z_{\pi} \equiv -\frac{\epsilon-\bar{\epsilon}}{\bar{\epsilon}}$, and where $\delta \equiv \frac{(\bar{\epsilon}-1)(\varsigma+\phi)}{\theta}$.

Appendix D Numerical Solution Algorithm

Using the functional forms assumed in Section 4.4, we can solve the model by assuming that $T \to \infty$ and using Laplace transforms.

The differential equations (25) are characterize the coefficient functions $\mathbf{A}(\tau), \tilde{\mathbf{A}}(\tau)$. Define the Laplace transform $\mathcal{A}(s) \equiv \mathcal{L}\left\{\mathbf{A}(\tau)\right\}(s)$. Then equation (25) implies:

$$\begin{aligned} s\mathcal{A}(s) + \mathbf{M}\mathcal{A}(s) - \frac{1}{s}\mathbf{e}_i = \mathbf{0} \\ \implies \mathcal{A}(s) = \left[s\mathbf{I} + \mathbf{M}\right]^{-1}\left[\frac{1}{s}\mathbf{e}_i\right] \end{aligned}$$

using the fact that $\mathbf{A}(0) = \mathbf{0}$. Similarly, we have

$$\begin{aligned} s\tilde{\mathcal{A}}(s) + \mathbf{e}_d + \mathbf{M}\tilde{\mathcal{A}}(s) - \frac{1}{s}\mathbf{e}_i = \mathbf{0} \\ \implies \tilde{\mathcal{A}}(s) = \left[s\mathbf{I} + \mathbf{M}\right]^{-1}\left[\frac{1}{s}\mathbf{e}_i - \mathbf{e}_d\right] \end{aligned}$$

using the fact that $\tilde{\mathbf{A}}(0) = -\mathbf{e}_d$.

Note that the derivative with respect to the Laplace parameter s is given by

$$\begin{aligned} \mathcal{A}'(s) &= \left[s\mathbf{I} + \mathbf{M}\right]^{-1}\left[-\frac{1}{s^2}\mathbf{e}_i - \mathcal{A}(s)\right] \\ \tilde{\mathcal{A}}'(s) &= \left[s\mathbf{I} + \mathbf{M}\right]^{-1}\left[-\frac{1}{s^2}\mathbf{e}_i - \tilde{\mathcal{A}}(s)\right] \end{aligned}$$

Additionally, define $\mathbf{X}(\tau) \equiv \mathbf{A}(\tau)\mathbf{A}(\tau)^\top$. Note that from equation (25) we can write

$$\begin{aligned} \mathbf{A}'(\tau)\mathbf{A}(\tau)^\top + \mathbf{A}(\tau)\mathbf{A}'(\tau)^\top + \mathbf{M}\mathbf{X}(\tau) + \mathbf{X}(\tau)\mathbf{M}^\top - \mathbf{e}_i\mathbf{A}(\tau)^\top - \mathbf{A}(\tau)\mathbf{e}_i^\top = \mathbf{0} \\ \iff \mathbf{X}'(\tau) + \mathbf{M}\mathbf{X}(\tau) + \mathbf{X}(\tau)\mathbf{M}^\top - \mathbf{e}_i\mathbf{A}(\tau)^\top - \mathbf{A}(\tau)\mathbf{e}_i^\top = \mathbf{0} \end{aligned}$$

Define the Laplace transform $\mathcal{X}(s) \equiv \mathcal{L}\left\{\mathbf{X}(\tau)\right\}(s)$. Then we have

$$\left[\frac{1}{2}s\mathbf{I} + \mathbf{M}\right]\mathcal{X}(s) + \mathcal{X}(s)\left[\frac{1}{2}s\mathbf{I} + \mathbf{M}\right]^\top = \mathbf{e}_i\mathcal{A}(s)^\top + \mathcal{A}(s)\mathbf{e}_i^\top$$

since $\mathbf{X}(0) = \mathbf{0}$. This is a Lyapunov equation. A sufficient condition for a unique solution $\mathcal{X}(s)$ is if all the eigenvalues of $\left[\frac{1}{2}s\mathbf{I} + \mathbf{M}\right]$ have positive real parts.

Analogously, define $\tilde{\mathbf{X}}(\tau)$ and $\tilde{\mathcal{X}}(s)$, and the same steps as above imply

$$\left[\frac{1}{2}s\mathbf{I} + \mathbf{M}\right]\tilde{\mathcal{X}}(s) + \tilde{\mathcal{X}}(s)\left[\frac{1}{2}s\mathbf{I} + \mathbf{M}\right]^\top = \mathbf{e}_i\tilde{\mathcal{A}}(s)^\top + \tilde{\mathcal{A}}(s)\mathbf{e}_i^\top + \mathbf{e}_d\mathbf{e}_d^\top$$

since $\tilde{\mathbf{X}}(0) = (-\mathbf{e}_d)(-\mathbf{e}_d^\top) = \mathbf{e}_d\mathbf{e}_d^\top$.

With this notation, we have that

$$\int_0^T \alpha(\tau;\alpha_0,\alpha_1)\mathbf{A}(\tau)\mathbf{A}(\tau)^\top \,\mathrm{d}\tau = \alpha_0\mathcal{X}(\alpha_1)$$
$$\int_0^T \theta(\tau;\theta_0,\theta_1)\mathbf{A}(\tau)^\top \,\mathrm{d}\tau = -\theta_0\theta_1^2\mathcal{A}'(\theta_1)^\top$$

and similarly for terms involving $\tilde{\mathbf{A}}(\tau)$.

With a slight abuse of notation, define the following matrices:

$$\begin{aligned}
\mathcal{X} &\equiv \alpha_0\mathcal{X}(\alpha_1) \\
\tilde{\mathcal{X}} &\equiv \tilde{\alpha}_0\tilde{\mathcal{X}}(\tilde{\alpha}_1) \\
\mathcal{Y} &\equiv \begin{bmatrix} \vdots \\ -\theta_{0k}\theta_{1k}^2\mathcal{A}'(\theta_{1k})^\top \\ \vdots \end{bmatrix} \\
\tilde{\mathcal{Y}} &\equiv \begin{bmatrix} \vdots \\ -\tilde{\theta}_{0k}\tilde{\theta}_{1k}^2\tilde{\mathcal{A}}'(\tilde{\theta}_{1k})^\top \\ \vdots \end{bmatrix}
\end{aligned}$$

Using the parameterization of $\eta(\tau), \tilde{\eta}(\tau)$ defined in equation (32), we define:

$$\begin{aligned}
\mathcal{N} &\equiv \int_0^T \eta(\tau)\mathbf{A}(\tau)\,\mathrm{d}\tau = -\eta_1^2\mathcal{A}'(\eta_1) \\
\tilde{\mathcal{N}} &\equiv \int_0^T \tilde{\eta}(\tau)\tilde{\mathbf{A}}(\tau)\,\mathrm{d}\tau = -\tilde{\eta}_1^2\tilde{\mathcal{A}}'(\tilde{\eta}_1) \\
\mathcal{V} &\equiv \eta_0\mathcal{N} + (1-\eta_0)\tilde{\mathcal{N}}
\end{aligned}$$

All of these objects are implicitly determined by $\mathbf{M}$. Hence we can write the equation characterizing $\hat{\mathbf{A}}$ and $\mathbf{M}$ (equations (26) and (27)), as the solution of a root-finding problem, as shown in the proof of Proposition 1:

$$f(\hat{\mathbf{A}};\mathbf{M};a) = \begin{bmatrix} \mathbf{e}_i + \left(\mathbf{\Gamma}(\hat{\mathbf{A}})^\top - \mathbf{M}\right)\mathcal{V} - \hat{\mathbf{A}} \\ \mathrm{vec}\left\{\mathbf{\Gamma}(\hat{\mathbf{A}})^\top - a\left\{\mathcal{Y} - \mathcal{X} + \tilde{\mathcal{Y}} - \tilde{\mathcal{X}}\right\}\mathbf{\Sigma} - \mathbf{M}\right\} \end{bmatrix}$$

We utilize a continuation algorithm to solve the model numerically, which also serves as an equilibrium selection device. First, for $a_0 \equiv 0$ we have a known solution. Next, given a solution $\hat{\mathbf{A}}_s, \mathbf{M}_s$ for $a = a_s$, we compute the solution $\hat{\mathbf{A}}_{s+1}, \mathbf{M}_{s+1}$ for $a_{s+1} \equiv a_s + \varepsilon_s$ using $\hat{\mathbf{A}}_s, \mathbf{M}_s$ as an initial value for a numerical root-finding algorithm. Finally, we stop when $a_S \equiv \bar{a}$. When the step size ε_s is small, this allows for a stable numerical approach to solving the model.

Note that there are numerous ways to improve the efficiency of our numerical approach, including common tools such as adaptive step sizes. However, in practice solving the model is efficient enough that setting fixed, small step sizes and using off-the-shelf root-finding numerical packages works well for the purposes of stability of the solution algorithm.

In order to compute second moments in the model, note that the long-run variance-covariance of the state $Var[\mathbf{y}_t] \equiv \mathbf{\Sigma}^\infty$ solves the following matrix equation

$$\mathbf{\Gamma}\mathbf{\Sigma}^\infty + \mathbf{\Sigma}^\infty\mathbf{\Gamma}^\top = \mathbf{\Sigma} \implies \text{vec}\,\mathbf{\Sigma}^\infty = (\mathbf{\Gamma} \oplus \mathbf{\Gamma})^{-1}\,\text{vec}(\mathbf{\Sigma})$$

where vec is the vectorization operator, and $\oplus$ denotes the Kronecker sum, defined for any two squares matrices A, B with respective dimensions $(a \times a), (b \times b)$ as $A \oplus B \equiv I_a \otimes B + A \otimes I_b$, and where $\otimes$ is the Kronecker product. Then the auto-covariance of the state is given by

$$Cov[\mathbf{y}_{t+s}, \mathbf{y}_t] = \exp(-\mathbf{\Gamma}s)\mathbf{\Sigma}^\infty$$

and computing covariances of any model object is immediate due to the linearity in the model. For instance, the standard deviation of τ-year yields is given by

$$\sqrt{Var\left[y_t^{(\tau)}\right]} = \frac{1}{\tau}\sqrt{\mathbf{A}(\tau)^\top \mathbf{\Sigma}^\infty \mathbf{A}(\tau)}$$

using the relationship between riskless yields and prices: $y_t^{(\tau)} = -\frac{1}{\tau}P_t^{(\tau)}$. Note for risky assets, we compute yields as $\tilde{y}_t^{(\tau)} = -\frac{1}{\tau}\frac{\tilde{P}_t^{(\tau)}}{D_t}$.

Appendix E Localization Proofs

Formal definitions and terminology: To compare the effects of demand factors j and k, we compare their relative effects on bond prices. Suppose demand factor k is more concentrated on long-maturity bonds compared to demand factor j ($\theta_1^j > \theta_1^k$). Consider the effect of factor k on long-maturity bonds relative to its own effects on short-maturity bonds. If this relative movement is the same as factor j, then we say that demand factors have "global" effects on the yield curve. If instead the relative effect of factor k on long-maturity bonds is larger than that of factor j (and similarly, the relative effect of factor j on short-maturity bonds is larger than that of factor k), we say that demand factors have "local" effects on the yield curve. To formalize this, define

$$\mathcal{B}_j(s,t) \equiv \frac{\mathcal{A}_j(s)}{\mathcal{A}_j(t)} \tag{E1}$$

This vector is the weighted average response of bond prices with Laplace frequency parameter s, relative to the weighted average response of bond prices with Laplace frequency parameter t. The weights in the Laplace transform are given by $e^{-s\tau}$, so $s > t$ implies that $\mathcal{A}_j(s)$ has more weight on short-maturity bonds than $\mathcal{A}_j(t)$. Hence, if $s > t$ and $\mathcal{B}_j(s,t) > \mathcal{B}_k(s,t)$, then we have that the j-demand factor has a relatively larger impact on short-maturity yields, compared to the k-demand factor. If $s < t$ and $\mathcal{B}_j(s,t) > \mathcal{B}_k(s,t)$, then we have that the j-demand factor has a relatively larger impact on long-maturity yields, compared to the k-demand factor.

Hence, demand factors have "global" effects if $\mathcal{B}_j(s,t) = \mathcal{B}_k(s,t)$ for all demand factors j,k and frequency weights s,t. Demand factors have "local" effects if $\mathcal{B}_j(s,t) > \mathcal{B}_k(s,t)$ for demand factors j,k such that $\theta_1^j > \theta_1^k$ and frequency weights $s > t$.

Before stating and proving our results formally, we preview and discuss the results. Proposition E1 shows that demand shocks have global effects when arbitrageurs approach risk-neutrality ($a \to 0$) or when demand factors are risk-free ($\sigma_\beta = 0$). Proposition E1 further shows that the first-order effects of increasing risk aversion do not lead to a localization of demand shocks. That is, in a neighborhood around $a = 0$, increasing risk aversion does change the global nature of demand shocks.

Proposition E2 shows that the second-order effects of increasing risk aversion lead to the localization of demand shocks. That is, for large enough values of risk aversion $a \gg 0$, demand shocks become localized. Proposition E2 also shows that with infinite risk aversion, demand shocks become fully localized. Note that this is a "discontinuous" result; the model is no longer arbitrage-free.

Assumptions:

1. The "general equilibrium" effects of imperfect arbitrage are not too large near $a = 0$: $\lim_{a\to 0} \mathrm{d}^n/\mathrm{d}a^n \, (\mathbf{\Gamma}) \approx 0$.

2. The state dynamics and correlation matrices are diagonal: $\mathbf{\Gamma} = \text{diag}\begin{bmatrix}\dots & \gamma_{j,j} & \dots\end{bmatrix}$ and $\mathbf{\Sigma} = \text{diag}\begin{bmatrix}\dots & \sigma^2_{j,j} & \dots\end{bmatrix}$.

3. Demand factors are identical besides the location of $\theta(\tau)^k$. That is, $\theta^j_0 = \theta^k_0$, $\gamma_{j,j} = \gamma_{k,k}$, and $\sigma^2_{j,j} = \sigma^2_{k,k}$ for $j, k > 1$.

4. There are no risky asset habitat demand traders: $\tilde{\alpha}(\tau) = 0$ and $\tilde{\theta}(\tau) = 0$.

Assumption 1 and 2 are for analytical tractability (and for instance hold trivially in partial equilibrium models such as Vayanos and Vila (2021)). Note that Assumption 1 does not imply that demand factors have no macroeconomic implications, but only assumes that changes in risk aversion in a neighborhood around risk-neutrality do not lead to changes in the dynamics of the model. Assumption 3 implies that besides the location in maturity space, demand factors are symmetric. Assumption 4 is also for tractability, and allows us to focus only on localization effects across maturities.

In order to simplify notation, define $\lim_{a=0} \frac{\mathrm{d}^n}{\mathrm{d}a^n}(x) \equiv \mathrm{d}^n_0 x$ for any model object x. Assumption 1 implies that as $a \to 0$,

$$\begin{aligned}\mathbf{M} &\to \mathbf{\Gamma}^\top \\ \mathrm{d}_0\mathbf{M} &= \{\mathcal{X} - \mathcal{Y}\}\,\mathbf{\Sigma} \\ \mathrm{d}^n_0\mathbf{M} &= n\left\{\mathrm{d}^{n-1}_0\mathcal{X} - \mathrm{d}^{n-1}_0\mathcal{Y}\right\}\mathbf{\Sigma}\end{aligned}$$

and hence we can recursively solve for all derivatives with respect to a at $a = 0$.

With these assumptions, we have that as $a \to 0$, $\mathbf{M}$ is also diagonal and so is $s\mathbf{I} + \mathbf{M}$. To simplify notation, define the vector and matrix equations

$$[\mathbf{g}(s)]_i \equiv \frac{1}{s + \gamma_{i,i}}, \quad [\mathbf{G}(s)]_{i,j} \equiv \frac{1}{s + \gamma_{i,i} + \gamma_{j,j}} \tag{E2}$$

and note that Assumption 3 implies that for all $j, k, m, n > 1$:

$$\mathbf{g}_j(s) = \mathbf{g}_k(s), \quad \mathbf{G}_{j,k}(s) = \mathbf{G}_{m,n}(s)$$

Using this notation, we have that

$$\lim_{a\to 0} \mathcal{A}(s) \equiv \mathcal{A}^0(s) = \frac{1}{s} \cdot \mathbf{g}(s) \circ \mathbf{e}_i = \begin{bmatrix}\frac{1}{s(s+\gamma_{11})} & 0 & \dots & 0\end{bmatrix}^\top \tag{E3}$$

$$\lim_{a\to 0} \mathcal{A}'(s) \equiv \mathcal{A}^{0\prime}(s) = -\mathbf{g}(s) \circ \left[\frac{1}{s^2}\mathbf{e}_i + \mathcal{A}(s)\right] = \begin{bmatrix}-\frac{2s+\gamma_{11}}{s^2(s+\gamma_{11})^2} & 0 & \dots & 0\end{bmatrix}^\top \tag{E4}$$

where $\circ$ denotes the Hadamard (element-wise) product. Further,

$$\mathrm{d}_0^n \mathcal{A}(s) = -\mathbf{g}(s) \circ \sum_{k=0}^{n-1} \binom{n}{k} \mathrm{d}_0^{n-k} \mathbf{M} \mathrm{d}_0^k \mathcal{A}(s) \tag{E5}$$

$$\mathrm{d}_0^n \mathcal{A}'(s) = -\mathbf{g}(s) \circ \left(\mathrm{d}_0^n \mathcal{A}(s) + \sum_{k=0}^{n-1} \binom{n}{k} \mathrm{d}_0^{n-k} \mathbf{M} \mathrm{d}_0^k \mathcal{A}'(s) \right) \tag{E6}$$

We also can characterize the limit of $\mathcal{X}(s)$:

$$\lim_{a \to 0} \mathcal{X}(s) \equiv \mathcal{X}^0(s) = \mathbf{G}(s) \circ \left(\mathbf{e}_i \mathcal{A}(s)^\top + \mathcal{A}(s) \mathbf{e}_i^\top \right) \tag{E7}$$

and recursively the derivatives are found from:

$$\mathrm{d}_0^n \mathcal{X}(s) = \mathbf{G}(s) \circ \left(\mathbf{e}_i \mathrm{d}_0^n \mathcal{A}(s)^\top + \mathrm{d}_0^n \mathcal{A}(s) \mathbf{e}_i^\top \sum_{k=0}^{n-1} \left(\mathrm{d}_0^{n-k} \mathbf{M} \mathrm{d}_0^k \mathcal{X}(s) + \mathrm{d}_0^k \mathcal{X}(s) \mathrm{d}_0^{n-k} \mathbf{M}^\top \right) \right) \tag{E8}$$

In particular, equations (E4) and (E7) imply:

$$\lim_{a \to 0} [\mathcal{X}]_{i,j} = \begin{cases} 2\alpha_0 \mathcal{A}_1^0(\alpha_1) \mathbf{G}_{1,1}(\alpha_1) & \text{if } i = j = 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\lim_{a \to 0} [\mathcal{Y}]_{i,j} = \begin{cases} -\theta_0^j (\theta_1^j)^2 \mathcal{A}_1^{0\prime}(s) & \text{if } i > 1, j = 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\implies [\mathrm{d}_0 \mathbf{M}]_{i,j} = \begin{cases} 2\sigma_{1,1}^2 \alpha_0 \mathcal{A}_1^0(\alpha_1) \mathbf{G}_{1,1}(\alpha_1) & \text{if } i = j = 1 \\ \sigma_{1,1}^2 \theta_0^j (\theta_1^j)^2 \mathcal{A}_1^{0\prime}(s) & \text{if } i > 1, j = 1 \\ 0 & \text{otherwise} \end{cases}$$

So $\mathrm{d}_0 \mathbf{M}$ is zero everywhere except the first column, which we denote $\mathrm{d}_0 \mathbf{M}_1$. Then

$$\mathrm{d}_0 \mathcal{A}(s) = -\mathbf{g}(s) \circ \mathrm{d}_0 \mathbf{M}_1 \cdot \mathcal{A}_1^0(s) \tag{E9}$$

$$\mathrm{d}_0 \mathcal{A}'(s) = -\mathbf{g}(s) \circ \mathrm{d}_0 \mathbf{M}_1 \circ \left(\mathcal{A}_1^{0\prime}(s) - \mathbf{g}(s) \cdot \mathcal{A}_1^0(s) \right) \tag{E10}$$

and

$$\mathrm{d}_0^2 \mathcal{A}(s) = -\mathbf{g}(s) \circ \left(\mathrm{d}_0^2 \mathbf{M}_1 \cdot \mathcal{A}_1^0(s) + 2\mathrm{d}_0 \mathbf{M}_1 \cdot \mathrm{d}_0 \mathcal{A}_1^0(s) \right) \tag{E11}$$

$$\mathrm{d}_0^2 \mathcal{A}'(s) = -\mathbf{g}(s) \circ \left(\mathrm{d}_0^2 \mathbf{M}_1 \cdot (\mathcal{A}_1^{0\prime}(s) - \mathbf{g}_1(s) \mathcal{A}_1^0(s)) + 2\mathrm{d}_0 \mathbf{M}_1 \cdot (\mathrm{d}_0 \mathcal{A}_1^{0\prime}(s) - \mathbf{g}_1(s) \mathrm{d}_0 \mathcal{A}_1^0(s)) \right) \tag{E12}$$

The following Lemma derives some useful algebraic properties of the model.

Lemma E1. *For $n \geq 2$, the ratio*

$$\frac{\mathrm{d}_0^n \mathcal{A}_j(s)}{\mathrm{d}_0 \mathcal{A}_j(s)} = n \frac{\mathrm{d}_0^{n-1} \mathcal{A}_1(s)}{\mathcal{A}_1^0(s)} + \frac{\mathrm{d}_0^n \mathbf{M}_{j,1}}{\mathrm{d}_0 \mathbf{M}_{j,1}} + \frac{\sum_{k=1}^{n-2} \binom{n}{k} \left[\mathrm{d}_0^{n-k} \mathbf{M} \mathrm{d}_0^k \mathcal{A}(s)\right]_j}{\mathrm{d}_0 \mathbf{M}_{j,1} \mathcal{A}_1^0(s)} \tag{E13}$$

Proof. We have that

$$\frac{\mathrm{d}_0^n \mathcal{A}_j(s)}{\mathrm{d}_0 \mathcal{A}_j(s)} = \frac{\left[\sum_{k=0}^{n-1} \binom{n}{k} \mathrm{d}_0^{n-k} \mathbf{M} \mathrm{d}_0^k \mathcal{A}(s)\right]_j}{\mathrm{d}_0 \mathbf{M}_{j,1} \mathcal{A}_1^0(s)}$$

Since $\mathrm{d}_0 \mathbf{M}$ is zero everywhere besides the first column, we have that $\mathrm{d}_0 \mathbf{M} \mathbf{x}(s) = \mathrm{d}_0 \mathbf{M}_1 \cdot \mathbf{x}_1(s)$ for any vector function $\mathbf{x}(s)$. Hence

$$\frac{\left[\mathrm{d}_0 \mathbf{M} \mathrm{d}_0^n \mathcal{A}(s)\right]_j}{\mathrm{d}_0 \mathbf{M}_{j,1} \mathcal{A}_1^0(s)} = \frac{\mathrm{d}_0^n \mathcal{A}_1(s)}{\mathcal{A}_1^0(s)}$$

which is independent of j.

Since $\mathcal{A}^0(s)$ is zero everywhere besides the first element, we have that $\mathbf{Q} \mathcal{A}^0(s) = \mathbf{Q}_1 \cdot \mathcal{A}_1^0(s)$ for any matrix $\mathbf{Q}$, where $\mathbf{Q}_1$ is the first column of $\mathbf{Q}$. Hence

$$\frac{\left[\mathrm{d}_0^n \mathbf{M} \mathcal{A}^0(s)\right]_j}{\mathrm{d}_0 \mathbf{M}_{j,1} \mathcal{A}_1^0(s)} = \frac{\mathrm{d}_0^n \mathbf{M}_{j,1}}{\mathbf{M}_{j,1}}$$

which is independent of s.

In particular, equation (E13) implies that for $n = 2$, the ratio is additively separable in the demand factor index j and the Laplace frequency weight s. □

Proposition E1 ("Global" Demand Shocks)**.** *Demand shocks have "global" effects under (near) risk-neutrality, or when demand is risk-free.*

(a) For any demand factors j, k and Laplace frequency parameters s, t:

$$\lim_{a \to 0} \left(\mathcal{B}_j(s,t) - \mathcal{B}_k(s,t)\right) = 0$$

(b) Assume $\mathbf{\Sigma}_{j,j} = 0$ for all demand factors j (but $\sigma_i^2 > 0$ and $a > 0$). For any demand factors j, k and Laplace frequency parameters s, t, we have that:

$$\mathcal{B}_j(s,t) = \mathcal{B}_k(s,t)$$

(c) For any demand factors j, k and Laplace frequency parameters s, t, we have that:

$$\lim_{a \to 0} \frac{\mathrm{d}}{\mathrm{d}a} \left(\mathcal{B}_j(s,t) - \mathcal{B}_k(s,t)\right) = 0$$

Proof. (a) For $j > 1$, $\mathcal{A}_j(s) = 0$ when $a = 0$, so apply L'Hopital's rule to equation (E1):

$$\lim_{a \to 0} \mathcal{B}_j(s,t) = \frac{\mathrm{d}_0 \mathcal{A}_j(s)}{\mathrm{d}_0 \mathcal{A}_j(t)} \tag{E14}$$

Equation (E9) implies that the ratio:

$$\frac{\mathrm{d}_0 \mathcal{A}_j(s)}{\mathrm{d}_0 \mathcal{A}_j(t)} = \frac{\mathbf{g}_j(s)\mathcal{A}_1^0(s)}{\mathbf{g}_j(t)\mathcal{A}_1^0(t)}$$

and from Assumption 3, $\mathbf{g}_j(s) = \mathbf{g}_k(s)$ for any $j > 1, k > 1$. Hence:

$$\frac{\mathrm{d}_0 \mathcal{A}_j(s)}{\mathrm{d}_0 \mathcal{A}_j(t)} = \frac{\mathrm{d}_0 \mathcal{A}_k(s)}{\mathrm{d}_0 \mathcal{A}_k(t)}$$

and the result follows.

A useful corollary of this result is that $\lim_{a \to 0} \mathcal{B}_j(s,t) > 0$, since $\mathcal{A}_1(s) > 0$ and $\mathbf{g}(s) > 0$ (element-wise).

(b) When demand is risk-free, $\mathbf{\Sigma}$ is zero everywhere except for the first element (equal to σ_i^2). Hence, we have that:

$$a \cdot \{\mathcal{X} - \mathcal{Y}\}\mathbf{\Sigma} = a\sigma_i^2 \cdot \begin{bmatrix} \mathcal{X}_1 - \mathcal{Y}_1 & \mathbf{0} & \dots & \mathbf{0} \end{bmatrix}$$

where $\mathcal{X}_1, \mathcal{Y}_1$ are the first columns of $\mathcal{X}$ and $\mathcal{Y}$. Since $\mathbf{\Gamma}$ is diagonal, for any $a > 0$, $\mathbf{M} \equiv \mathbf{\Gamma}^\top + a \cdot \{\mathcal{X} - \mathcal{Y}\}\mathbf{\Sigma}$ is zero everywhere except the first column and diagonal. Therefore, the $j > 1$ element of $\mathcal{A}(s)$ is given by

$$\begin{aligned} \mathcal{A}_j(s) &= \frac{1}{s}\left[(s\mathbf{I} + \mathbf{M})^{-1}\right]_{j,1} = \frac{a\sigma_i^2(\mathcal{Y}_{j,1} - \mathcal{X}_{j,1})}{s(s + \gamma_{j,j})(s + \gamma_{1,1} + a\sigma_i^2\mathcal{X}_{1,1})} \\ \implies \mathcal{B}_j(s,t) &= \frac{t(t + \gamma_{j,j})(t + \gamma_{1,1} + a\sigma_i^2\mathcal{X}_{1,1})}{s(s + \gamma_{j,j})(s + \gamma_{1,1} + a\sigma_i^2\mathcal{X}_{1,1})} \end{aligned}$$

Then the result follows since $\gamma_{j,j} = \gamma_{k,k}$ for $j, k > 1$ (by Assumption 3).

(c) Differentiating equation (E1) (and applying L'Hopital's rule and removing terms equal to zero in the limit) gives

$$\mathrm{d}_0 \mathcal{B}_j(s,t) = \frac{\mathrm{d}_0^2 \mathcal{A}_j(s)\mathrm{d}_0 \mathcal{A}_j(t) - \mathrm{d}_0^2 \mathcal{A}_j(t)\mathrm{d}_0 \mathcal{A}_j(s)}{2\mathrm{d}_0 \mathcal{A}_j(t)^2}$$

Equation (E14) and the corollary from the proof of (a) implies

$$\lim_{a \to 0} \mathcal{B}_j(s,t) = \frac{\mathrm{d}_0 \mathcal{A}_j(t)}{\mathrm{d}_0 \mathcal{A}_j(s)} = \frac{\mathrm{d}_0 \mathcal{A}_k(t)}{\mathrm{d}_0 \mathcal{A}_k(s)} = \lim_{a \to 0} \mathcal{B}_k(s,t) > 0$$

Hence:

$$\frac{\mathrm{d}_0\mathcal{A}_j(t)}{\mathrm{d}_0\mathcal{A}_j(s)}\mathrm{d}_0\mathcal{B}_j(s,t) = \frac{1}{2}\left(\frac{\mathrm{d}_0^2\mathcal{A}_j(s)}{\mathrm{d}_0\mathcal{A}_j(s)} - \frac{\mathrm{d}_0^2\mathcal{A}_j(t)}{\mathrm{d}_0\mathcal{A}_j(t)}\right)$$
$$\frac{\mathrm{d}_0\mathcal{A}_j(t)}{\mathrm{d}_0\mathcal{A}_j(s)}\mathrm{d}_0\mathcal{B}_k(s,t) = \frac{1}{2}\left(\frac{\mathrm{d}_0^2\mathcal{A}_k(s)}{\mathrm{d}_0\mathcal{A}_k(s)} - \frac{\mathrm{d}_0^2\mathcal{A}_k(t)}{\mathrm{d}_0\mathcal{A}_k(t)}\right)$$

From the Lemma E1, we have that

$$\frac{1}{2}\frac{\mathrm{d}_0^2\mathcal{A}_j(s)}{\mathrm{d}_0\mathcal{A}_j(s)} = \frac{1}{2}\frac{\mathrm{d}_0^2\mathbf{M}_{j,1}}{\mathrm{d}_0\mathbf{M}_{j,1}} + \frac{\mathrm{d}_0\mathcal{A}_1(s)}{\mathcal{A}_1^0(s)}$$

which is additively separable in terms which are a function of j and s. The result follows.

□

Proposition E2 ("Localized" Demand Shocks)**.** *Demand shocks have "local" effects far from risk-neutrality, and are fully localized under infinite risk aversion.*

(a) Suppose $\mathbf{\Sigma}_{j,j} > 0$*. For any demand factors* j, k *and Laplace frequency parameters* s, t*:*

$$\lim_{a\to 0} \operatorname{sgn} \frac{\mathrm{d}^2}{\mathrm{d}a^2}(\mathcal{B}_j(s,t) - \mathcal{B}_k(s,t)) = \operatorname{sgn}(s-t)(\theta_1^j - \theta_1^k)$$

(b) When risk aversion $a = \infty$*:*

$$\frac{\mathrm{d}\log P_t^{(\tau)}}{\mathrm{d}\beta_t^k} = \frac{\theta^k(\tau)}{\alpha(\tau)}, \quad \frac{\mathrm{d}\log P_t^{(\tau)}}{\mathrm{d}i_t} = 0$$

Proof. (a) Differentiating equation (E1) (and applying L'Hopital's rule and removing terms equal to zero in the limit) gives

$$\mathrm{d}_0^2\mathcal{B}_j(s,t) = \frac{1}{6\mathrm{d}_0\mathcal{A}_j(t)^3}\left(3\mathrm{d}_0^2\mathcal{A}_j(t)^2\mathrm{d}_0\mathcal{A}_j(s) - 3\mathrm{d}_0^2\mathcal{A}_j(s)\mathrm{d}_0^2\mathcal{A}_j(t)\mathrm{d}_0\mathcal{A}_j(t) + 2\mathrm{d}_0^3\mathcal{A}_j(s)\mathrm{d}_0\mathcal{A}_j(t)^2 - 2\mathrm{d}_0^3\mathcal{A}_j(t)\mathrm{d}_0\mathcal{A}_j(t)\mathrm{d}_0\mathcal{A}_j(s)\right)$$

Using the results from Prop E1, we can scale and rewrite as:

$$\frac{\mathrm{d}_0\mathcal{A}_j(t)}{\mathrm{d}_0\mathcal{A}_j(s)}\mathrm{d}_0^2\mathcal{B}_j(s,t) = \frac{1}{3}\left(\frac{\mathrm{d}_0^3\mathcal{A}_j(s)}{\mathrm{d}_0\mathcal{A}_j(s)} - \frac{\mathrm{d}_0^3\mathcal{A}_j(t)}{\mathrm{d}_0\mathcal{A}_j(t)}\right) + \frac{1}{2}\left(\frac{\mathrm{d}_0^2\mathcal{A}_j(t)}{\mathrm{d}_0\mathcal{A}_j(t)}\right)\left(\frac{\mathrm{d}_0^2\mathcal{A}_j(t)}{\mathrm{d}_0\mathcal{A}_j(t)} - \frac{\mathrm{d}_0^2\mathcal{A}_j(s)}{\mathrm{d}_0\mathcal{A}_j(s)}\right)$$
$$\frac{\mathrm{d}_0\mathcal{A}_j(t)}{\mathrm{d}_0\mathcal{A}_j(s)}\mathrm{d}_0^2\mathcal{B}_k(s,t) = \frac{1}{3}\left(\frac{\mathrm{d}_0^3\mathcal{A}_k(s)}{\mathrm{d}_0\mathcal{A}_k(s)} - \frac{\mathrm{d}_0^3\mathcal{A}_k(t)}{\mathrm{d}_0\mathcal{A}_k(t)}\right) + \frac{1}{2}\left(\frac{\mathrm{d}_0^2\mathcal{A}_k(t)}{\mathrm{d}_0\mathcal{A}_k(t)}\right)\left(\frac{\mathrm{d}_0^2\mathcal{A}_j(t)}{\mathrm{d}_0\mathcal{A}_j(t)} - \frac{\mathrm{d}_0^2\mathcal{A}_j(s)}{\mathrm{d}_0\mathcal{A}_j(s)}\right)$$

and note that the final parenthetical term in each expression is identical.

Focusing on the third-order terms, from Lemma E1 we have that

$$\frac{1}{3}\frac{\mathrm{d}_0^3\mathcal{A}_j(s)}{\mathrm{d}_0\mathcal{A}_j(s)} = \frac{\mathrm{d}_0^2\mathcal{A}_1(s)}{\mathcal{A}_1^0(s)} + \frac{1}{3}\frac{\mathrm{d}_0^3\mathbf{M}_{j,1}}{\mathbf{M}_{j,1}} + \frac{\left[\mathrm{d}_0^2\mathbf{M}\mathrm{d}_0\mathcal{A}(s)\right]_j}{\mathrm{d}_0\mathbf{M}_{j,1}\mathcal{A}_1^0(s)}$$

The final term in the sum can be written

$$\frac{\left[\mathrm{d}_0^2\mathbf{M}\mathrm{d}_0\mathcal{A}(s)\right]_j}{\mathrm{d}_0\mathbf{M}_{j,1}\mathcal{A}_1^0(s)} = \sum_{m=1}^{J}\left[\frac{[\mathrm{d}_0^2\mathbf{M}]_{j,m}}{\mathrm{d}_0\mathbf{M}_{j,1}}\right]\cdot\left[\frac{\mathrm{d}_0\mathcal{A}_m(s)}{\mathcal{A}_1^0(s)}\right] \equiv \mathbf{u}_j^\top\mathbf{v}(s)$$

and hence

$$\frac{1}{3}\left(\frac{\mathrm{d}_0^3\mathcal{A}_j(s)}{\mathrm{d}_0\mathcal{A}_j(s)} - \frac{\mathrm{d}_0^3\mathcal{A}_j(t)}{\mathrm{d}_0\mathcal{A}_j(t)}\right) - \frac{1}{3}\left(\frac{\mathrm{d}_0^3\mathcal{A}_k(s)}{\mathrm{d}_0\mathcal{A}_k(s)} - \frac{\mathrm{d}_0^3\mathcal{A}_k(t)}{\mathrm{d}_0\mathcal{A}_k(t)}\right) = \left[\mathbf{u}_j - \mathbf{u}_k\right]^\top\left[\mathbf{v}(s) - \mathbf{v}(t)\right]$$

Further, from Lemma E1,

$$\begin{aligned}
\left(\frac{\mathrm{d}_0^2\mathcal{A}_j(t)}{\mathrm{d}_0\mathcal{A}_j(t)} - \frac{\mathrm{d}_0^2\mathcal{A}_k(t)}{\mathrm{d}_0\mathcal{A}_k(t)}\right) &= \frac{\mathrm{d}_0^2\mathbf{M}_{j,1}}{\mathrm{d}_0\mathbf{M}_{j,1}} - \frac{\mathrm{d}_0^2\mathbf{M}_{k,1}}{\mathrm{d}_0\mathbf{M}_{k,1}} \\
&\equiv \left[\mathbf{u}_j - \mathbf{u}_k\right]_1 \\
\frac{1}{2}\left(\frac{\mathrm{d}_0^2\mathcal{A}_j(t)}{\mathrm{d}_0\mathcal{A}_j(t)} - \frac{\mathrm{d}_0^2\mathcal{A}_j(s)}{\mathrm{d}_0\mathcal{A}_j(s)}\right) &= \frac{\mathrm{d}_0\mathcal{A}_1(t)}{\mathcal{A}_1^0(t)} - \frac{\mathrm{d}_0\mathcal{A}_1(s)}{\mathcal{A}_1^0(s)} \\
&\equiv -\left[\mathbf{v}(s) - \mathbf{v}(t)\right]_1
\end{aligned}$$

Therefore, we have that

$$\frac{\mathrm{d}_0\mathcal{A}_j(t)}{\mathrm{d}_0\mathcal{A}_j(s)}\left(\mathrm{d}_0^2\mathcal{B}_j(s,t) - \mathrm{d}_0^2\mathcal{B}_k(s,t)\right) = \sum_{m=2}^{J}\left[\mathbf{u}_j - \mathbf{u}_k\right]_m\left[\mathbf{v}(s) - \mathbf{v}(t)\right]_m$$

Dealing with each vector separately: first, we have that

$$\begin{aligned}
\mathbf{v}(s) = \frac{\mathrm{d}_0\mathcal{A}(s)}{\mathcal{A}_1^0(s)} &= -\mathbf{g}(s)\circ\mathrm{d}_0\mathbf{M}_1 \\
\implies \mathbf{v}(s) - \mathbf{v}(t) &= -\mathrm{d}_0\mathbf{M}_1\circ\left(\mathbf{g}(s) - \mathbf{g}(t)\right)
\end{aligned}$$

Next, the j,m element of $\mathrm{d}_0^2\mathbf{M}$ is

$$[\mathrm{d}_0^2\mathbf{M}]_{j,m} = \left(\alpha_0[\mathrm{d}_0\mathbf{X}]_{j,m} + \theta_0^j(\theta_1^j)^2\mathrm{d}_0\mathcal{A}'_m(\theta_1^j)\right)\sigma_{m,m}^2$$

(and recall that $\theta_0^1 = 0$). We have that

$$\frac{[\mathrm{d}_0\mathbf{X}]_{j,m}}{\mathrm{d}\mathbf{M}_{j,1}} = \begin{cases} -\left(\mathbf{g}_j(\alpha_1)\mathcal{A}_1^0(\alpha_1) + \mathcal{X}_{1,1}^0(\alpha_1)\right)\mathbf{S}_{j,1}(\alpha_1) & \text{if } m = 1 \\ 0 & \text{otherwise} \end{cases}$$
$$= \frac{[\mathrm{d}_0\mathbf{X}]_{k,m}}{\mathrm{d}\mathbf{M}_{k,1}}$$

where the second equality follows from Assumption 3.

Recall $\mathrm{d}_0\mathbf{M}_{j,1} = \sigma_i^2\theta_0^j(\theta_1^j)^2[\mathrm{d}_0\mathcal{A}'(\theta_1^j)]_1$, hence we can write

$$-\frac{\theta_0^j(\theta_1^j)^2}{\mathrm{d}\mathbf{M}_{j,1}}\mathrm{d}_0\mathcal{A}'(\theta_1^j) = -\mathrm{d}_0\mathbf{M}_1 \circ \mathbf{g}(\theta_1^j) \circ \left(1 + \mathbf{g}(\theta_1^j)\frac{\theta_1^j}{1 + \theta_1^j\mathbf{g}_1(\theta_1^j)}\right)$$
$$\equiv -\mathrm{d}_0\mathbf{M}_1 \circ \mathbf{h}(\theta_1^j)$$

Therefore,

$$[\mathbf{u}_j - \mathbf{u}_k] = -\mathrm{d}_0\mathbf{M}_1 \circ \left(\mathbf{h}(\theta_1^j) - \mathbf{h}(\theta_1^k)\right)$$

which implies

$$[\mathbf{u}_j - \mathbf{u}_k] \circ [\mathbf{v}(s) - \mathbf{v}(t)] = (\mathrm{d}_0\mathbf{M}_1)^2 \circ (\mathbf{h}(\theta_1^j) - \mathbf{h}(\theta_1^k)) \circ (\mathbf{g}(s) - \mathbf{g}(t))$$

Hence, the sign of any element of the above vector is determined by the (element-wise) behavior of the functions $\mathbf{g}, \mathbf{h}$. From equation (E2), the function $\mathbf{g}_m(s)$ is decreasing in s for any element m. Further,

$$\mathbf{h}_m(x) = \frac{3x^2 + 2(\gamma_{m,m} + \gamma_{1,1}) + \gamma_{m,m}\gamma_{1,1}}{(x + \gamma_{m,m})^2(2x + \gamma_{1,1})}$$
$$\implies \mathbf{h}_m'(x) = -\frac{2x(x + \gamma_{1,1})(\gamma_{1,1} + \gamma_{m,m} + 3x)}{(x + \gamma_{m,m})^3(2x + \gamma_{1,1})^2}$$

so $\mathbf{h}_m(x)$ is also decreasing in x for any element m. Thus, we have that

$$\mathrm{sgn}\left[(\mathbf{h}_m(\theta_1^j) - \mathbf{h}_m(\theta_1^k)) \cdot (\mathbf{g}_m(s) - \mathbf{g}_m(t))\right] = \mathrm{sgn}(\theta_1^j - \theta_1^k) \cdot (s - t)$$

and therefore the sum over the $m \geq 2$ elements has the same sign. The result follows.

(b) If $a = \infty$, arbitrageurs invest their entire wealth in the risk-free rate and hold no long-maturity bonds: $X_t^{(\tau)} = 0$ for all τ and t. Thus market clearing implies that $Z_t^{(\tau)} = 0$. The results follow from differentiating preferred habitat demand with respect to any demand factor β_t^k or short rate i_t and setting to zero.

□

Appendix F Data Sources

This data appendix describes in detail the data sources used in this paper.

Treasury

Information regarding the details of each Treasury auction is available from the Treasury website TreasuryDirect. Historical data on announcements and results are available starting in 1979, and are continuously updated. Access is freely available through a web API, described here.

Additional data on auction allocations by investor type are provided by the Treasury here. Historical data is available starting in 2000, and are continuously updated.

The Treasury produces the High Quality Market (HQM) data set on corporate yield curves, and is available here. HQM data can also be accessed through Federal Reserve Economic Data (FRED) (whose web API is described here). HQM FRED data mnemonics are denoted by `HQMCB[XX]` where `XX` denotes maturity.

GovPX/Chicago Mercantile Exchange (CME)

Intraday (tick-level) data on the secondary market for Treasuries is available through GovPX, currently maintained by the Chicago Mercantile Exchange (CME). GovPX data is available via subscription through Wharton Research Data Services (WRDS), described here. Otherwise, GovPX data can be purchased from the CME's DataMine repository. Our GovPX data covers the period 1995 through 2017.

A previous version of this paper used intraday (tick-level) data on Treasury futures to construct auction demand shocks. This data is available for purchase directly from the Chicago Mercantile Exchange (CME) through the DataMine repository.

NYSE Trade and Quote (TAQ)

Intraday (tick-level) data on equities is available from the NYSE Trades and Quotes (TAQ) database. TAQ data via subscription through Wharton Research Data Services (WRDS), described here.

Refinitiv Datastream

Daily data on federal funds futures, equities, inflation swaps, commodity prices, spreads, and credit default swaps are obtained via subscription to Datastream, currently maintained by Refinitiv and described here.

Federal Reserve

Daily zero-coupon Treasury yields are from Gürkaynak et al. (2007); the data set is currently maintained by the Federal Reserve, available here. Data on nominal yields are available starting in 1961, and are continuously updated.

Balance sheet data for the Federal Reserve are available in the data release Factors Affecting Reserve Balances, available here. Balance sheet data can also be accessed also be accessed through Federal Reserve Economic Data (FRED). The FRED data mnemonics are denoted by `TREAS[XX]` and `MBSXX` for Treasury and MBS holdings, respectively, and where `XX` denotes maturity.

Industrial production data is available in the data release Industrial Production and Capacity Utilization, available here. Industrial production data can also be accessed through Federal Reserve Economic Data (FRED). The FRED data mnemonic is denoted by `INDPRO`.

Statistics on trading volume in the Commercial paper market are available from the Federal Reserve, and available here.

Bereau of Economic Analysis (BEA)

Personal Consumption Expenditures (PCE) price index data is available from the BEA's Personal Income data release, avilable here. PCE index data can also be accessed through Federal Reserve Economic Data (FRED). The FRED data mnemonic is denoted by `PCEPI`.

Securities Industry and Financial Markets Association (SIFMA)

Statistics on trading volume in Treasury and corporate fixed income markets are available from the Securities Industry and Financial Markets Association (SIFMA). These statistics are available here.