Unexpected Supply Effects of Quantitative Easing and Tightening

Stefania D'Amico and Tim Seida*

May 8, 2022

Abstract

To analyze the evolution of the effects of quantitative easing (QE) and tightening (QT) across consecutive announcements, we focus on their unexpected component. Treasury yield sensitivities to QT supply surprises are on average larger than sensitivities to QE surprises, implying supply effects did not diminish during periods of market calm amid economic expansion. Yield sensitivities to later QE and QT surprises do not fall monotonically, thus supply shocks seemed to remain powerful. Finally, yield sensitivities are amplified by the amount of interest-rate uncertainty prevailing before announcements, implying that turning points in the balance sheet policy tended to elicit larger reactions.

Keywords: Balance sheet policy surprises, quantitative easing and tightening, asset supply effects

JEL codes: E43, E44, E52, E58

*D'Amico: Federal Reserve Bank of Chicago. E-mail: stefania.damico@chi.frb.org (corresponding author). Seida: Federal Reserve Bank of Chicago. E-mail: tseida@frbchi.org. For helpful comments, discussions, and suggestions we thank Gene Amromin, Gadi Barlevy, David Kelly, Thomas King, Thomas Laubach, Rich Rosen, Sam Schulhofer-Wohl, Min Wei, and seminar participants at the Federal Reserve Bank of Chicago. The views expressed here do not reflect official positions of the Federal Reserve.

1 Introduction

As demonstrated by the unfolding of monetary policy actions during the onset of the Covid-19 pandemic, the likelihood that the policy rate hits the zero-lower-bound (ZLB) has become significantly higher, making it very valuable for central banks to keep a variety of policy tools at their disposal. One of the available tools is balance sheet policy (BSP) and its most common form has been quantitative easing (QE), large-scale purchases of government-guaranteed securities.[1] QE has been used extensively by all major central banks during the last global financial crisis and, since March 2020, 17 central banks have announced QE interventions to support financial markets amid the pandemic.[2] Clearly, so far, the use of QE has been limited to ZLB periods, as this tool is not viewed yet as part of normal monetary policy. As argued by Bernanke (2020), the lack of experience with QE created a lot of uncertainty about its effectiveness over time and in normal financial conditions.

In this study, we try to shed some light on the evolution of asset price effects of the Federal Reserve's (Fed) BSP over time and across diverse economic and financial market conditions. That is, we study the state dependence of BSP effects. We label changes in asset supply triggered by the Fed's BSP during a tightening cycle, quantitative tightening (QT), and during an easing cycle, QE.[3] During both QT and QE announcements, the asset supply shift induced by the BSP can be either larger or smaller than expected, thus QT and QE are not necessarily synonyms of hawkish and dovish BSP surprises, respectively. The distinguishing characteristics of QT and QE are the economic and financial market conditions that led to the change in BSP and that might induce similar asset supply shocks to possibly have different announcement impacts. In what follows we examine this eventuality.

In particular, we investigate whether the sensitivity of Treasury yields to localized asset supply shocks changed across QE and QT as well as across earlier and later QE programs. Hence, our primary focus is the state dependence of the supply/scarcity channel of BSP. This type of analysis is relevant for several reasons. First, it helps us understand whether BSP has diminishing returns across subsequent programs. Second,

[1]Other forms of BSP include, for example, credit easing, that is, the purchase of corporate bonds and other privately-issued securities, as well as yield curve control, that is, the targeting of yield levels rather than asset supply.

[2]See Table 1 in Hartley and Rebucci (2020).

[3]To divide events into just two categories, we include in a "tightening cycle" also the beginning of reduction in policy accommodation, such as the tapering of asset purchases.

it helps us understand whether the impact of BSP is diminished during periods of financial market calm, such as those observed during QT. Third, it helps evaluate the impact of the unwinding of QE in the Treasury market, which has been hardly studied before. And last but not least, it speaks to macro-finance theory arguing that central banks' BSP should be effective mainly in the presence of tighter financial constraints or market dysfunctions.

Overall, the joint understanding of all of those BSP aspects can clarify the relative importance of the yield sensitivity versus the size of the asset supply shock for the *total supply effect* of BSP, approximated by the product of the yield sensitivity and supply change induced by the BSP.[4] If the yield sensitivity does not materially change across economic and financial market conditions, then what matters is the size of supply shocks. Most previous research on QE has not fully controlled for expectations about the level and composition of asset purchases, which is crucial to a well-identified BSP shock. That failure might have led to the conclusion that later rounds of QE were less effective than the first QE. Our premise, similarly to Joyce et al. (2011) and Cahill et al. (2013), is that later rounds of QE were better anticipated by market participants, and therefore the magnitude of the supply surprise—not necessarily the yield sensitivity—has become smaller, inducing smaller changes in asset prices around the formal BSP announcements. To correctly identify the asset price effects of later rounds of QE, it is necessary to focus on the unexpected component of the BSP announcements, that is, the amount and distribution of asset purchases that was not correctly anticipated by investors. To do so we use the Survey of Primary Dealers (SPD) compiled by the New York Fed before each FOMC. This is conceptually equivalent to using the unexpected change in the Federal Funds Rate (FFR) to identify the policy rate surprise in the case of conventional monetary policy.

Arguably, the role of investor anticipation is also very relevant during QT as, while both policy-makers and investors had gained experience with BSP during QE, none of them had experience with the unwinding of the Fed's large balance sheet. Turning points in BSP, as demonstrated by the episode of the "Taper Tantrum" in May-June 2013—when the eventuality of the Fed tapering the pace of asset purchases pushed longer-term rates higher by about 120 basis points over two months—imply that substantial uncertainty about the normalization process of the Fed's balance

[4] In equilibrium models, the yield sensitivity is determined by the derivative of the risk premium with respect to changes in the supply factor, formalized in equation 3.

sheet could have amplified asset price effects through both larger BSP surprises and increased interest rate uncertainty. This also suggests that asset supply shocks and interest rate uncertainty might interact, therefore in our study we will account also for this interaction. Importantly, differently from QE, QT is implemented away from the ZLB, which tends to compress interest rate volatility. Hence, during QT, interest rate volatility is usually higher, potentially playing a larger role.

Our first innovation is to develop measures of BSP surprises for both QE and QT to assess their relative impact in the Treasury market. Our second contribution consists of using a new identification procedure to quantify localized supply effects of QE and QT announcements. We build on the methodology of Cahill et al. (2013), who rely on *relative* yield changes of *individual* Treasury securities in small time windows around earlier QE announcements. But, differently from that study, our empirical strategy isolates the supply effects of QE and QT by exploiting *kinks* in the yield curve reaction to the supply shocks triggered by various BSP announcements. The kinks and their maturity locations on the yield curve are connected to the release of the operational details of the purchases/reinvestments implementation, which are decided before the formal BSP announcement, and are therefore independent of the price reaction on the day of the announcement. Importantly, the slope change around the kink should retrieve the causal effect of only the supply shock, as it seems reasonable to assume that kinks in the yield reaction cannot be caused by either a change in the expected policy-rate path or in aggregate duration risk, as those factors should affect yields with close maturities (i.e., those around the kinks) similarly and in a smooth fashion. Finally, our third novelty, is to control for the interactions between the asset supply surprise and the uncertainty about longer-term rates prevailing before the BSP announcement.

In each event, the Treasury yield sensitivity is obtained by dividing the slope change in the yield curve reaction around the kink by the size of the asset supply surprise. This normalization is necessary to test the state dependence of the yield sensitivity. Based on a range of estimates, we obtain four main findings. First, the yield sensitivity to the supply shock does not seem to fall monotonically across later QE and QT announcements. This suggests that the supply effects remained powerful over time. Second, the yield sensitivity to supply shocks during QT is at least as large as the yield sensitivity to supply shocks during QE, implying the supply effects did not diminish during periods of market calm or across monetary-policy cycles. Third, the impact of both QE and QT announcements seems amplified by the amount of interest-rate uncertainty prevailing before the announcement, implying that turning points in BSP tend to elicit larger reactions. Fourth, total supply effects are found to be sizable, accounting for about half of the overall impact of the Fed's BSP.

Our findings show that taking account of investor expectations and uncertainty about BSP ahead of each announcement delivers conclusions about the state dependence of QE effects quite different from previous studies, as most found QE to be powerful only during crises.[5] Further, most of our findings pose challenges to existing macro-finance models of QE, in which changes in asset supply have meaningful financial effects only under some form of market dysfunction or tighter financial constraints.[6] Indeed, we find that during the previous economic expansion, while the Fed was tightening policy amid good financial conditions, BSP shocks continued to have significant yield effects. Finally, this also shows that, similarly to the FFR surprises for conventional monetary policy, the magnitude of the BSP surprises is extremely important for the impact of unconventional monetary policy, suggesting that the transmission mechanism of these two policy tools might be more similar than previously thought.[7]

The above evidence leads to three main policy implications. First, careful forward guidance about the BSP and not just about the policy rate can help control financial market effects by calibrating the size of the asset supply shock. It is the size of such shock that matters for BSP effects as the yield sensitivity does not appear to be particularly state dependent, except for the case of elevated interest rate uncertainty. Second, BSP can still be effective in changing Treasury yields away from the ZLB and during normal market conditions, suggesting that BSP could effectively become part of the normal monetary policy toolkit. Finally, since supply effects are estimated to be a significant share of the overall yield change, and can be localized to segments of the yield curve, it is likely that BSPs such as yield curve control could be implemented successfully.

The rest of the paper is organized as follows. In the next section, we review the literature and provide a simple framework to understand the state dependence of yield changes. Section 3 describes our measure of the BSP surprise and the data. Section

[5]See for example, Altavilla and Giannone (2017), Bauer and Neely (2014), Bowman et al. (2015), de los Rios and Shamloo (2017), Droste et al. (2021), Greenlaw et al (2018), and Krishnamurthy and Vissing-Jorgensen (2011 and 2013).

[6]See for example, Curdia and Woodford (2011), He and Krishnamurthy (2013), Gertler and Karadi (2011, 2013), and Schabert (2015).

[7]This is in line with the evidence reported in Swanson (2021), who finds that conventional and unconventional monetary policies have comparable effects in financial markets.

4 summarizes the events used to estimate the Treasury yield sensitivities. Section 5 details our empirical strategy and its advantages. Section 6 focus on the baseline results. Section 7 analyzes the role of BSP uncertainty in amplifying yield sensitivities. Section 8 offers concluding remarks.

2 Related Literature

Our work relies on methods and insights from the monetary-policy event-study literature, the QE literature focused on the supply/scarcity channel, and macro-finance studies focused on the role of financial frictions in the transmission of unconventional monetary policy.

A number of papers estimate the impact of central banks' QE programs on interest rates using event studies, but except for Joyce et al. (2011) and Cahill et al. (2013), none focuses on constructing the unexpected component of BSP to better isolate its effects.[8] In most event studies (e.g., Gagnon et al. 2011; Krishnamurthy and Vissing-Jorgensen 2011, 2013), the total impact of BSP is computed combining high-frequency yield changes across selected QE-related events. This approach becomes increasingly more problematic after the introduction of the first QE program, as the central bank signaled its intentions well before formal policy announcements and strengthened the conditionality of the QE program to macroeconomic outcomes. As a consequence, the identification of the relevant events becomes extremely hard, as any economic news and data releases can alter market participants' expectations. If the set of relevant events selected for each program is not exhaustive, the evolution of investor expectations about BPS is not properly tracked, and therefore the asset price impact will not be estimated correctly. Controlling for pre-announcement market expectations using the SPD helps avoid these limitations. Importantly, relative to Cahill et al. (2013), we extend the computations of the BSP surprises to open-ended QE programs and to QT, which is key to understand the evolution of BSP effects across different economic and financial conditions.

Since our primary focus is to test whether BSP effects are state dependent, controlling for the evolution of BSP expectations is even more important. This is because we need to estimate the marginal effect of each program, that is, the ratio between the

[8] More recently, Hesse et al. (2018) and Kim et al. (2020) have shown the importance of controlling for the evolution of expectations about the central bank's balance sheet size to correctly evaluate the macroeconomic effects of later QE programs in the UK and US, respectively.

change in yields and the BSP shock, and analyze how it evolved over time. If we do not control for the expected component of the BSP, we would overstate the size of the BSP surprise, and therefore underestimate the magnitude of BSP impact, leading to an incorrect comparison across programs and to the wrong policy conclusions about the efficacy and conduct of BSP.

Our empirical strategy is conceptually related to studies of local supply effects (e.g., D'Amico and King, 2013 and Cahill et al., 2013 for the US; Joyce and Tong, 2012 and McLaren et al., 2014 for the UK), as we also rely on the fact that for investors it may have been harder to correctly anticipate the maturity distribution of asset purchases, generating supply surprises specific to maturity sectors of the yield curve. But, differently from previous studies, our methodology does not require us to compute the asset supply surprise for each individual security and control for proxies of other QE channels.[9] Since it is quite hard to derive security-level proxies of all the QE channels without recurring to some arbitrary assumptions, our approach is more robust. Further, while previous work has estimated scarcity/supply effects only for earlier QE, characterized by fixed-size programs, we extend the estimation of those effects to open-ended QE and QT. This is crucial to our understanding of later QE and the policy normalization process following QE.

Another strand of the QE literature relevant for our work consists of studies showing that the price impacts of QE announcements do not seem transitory in nature. If they were transitory due to, for instance, asset-specific liquidity effects, they would be less relevant for QE macroeconomic outcomes. Bernanke (2020) argues that if announcement effects were predictably temporary then, first, smart investors could profit by betting on reversal and, second, we should not observe cross-asset impacts around the announcements of Treasury-only QE, as those cross-asset impacts would be inconsistent with Treasury-specific liquidity effects. Indeed, Bernanke (2020) reports evidence showing that predicting reversals of QE announcement effects is not a money-making strategy,[10] and that QE's cross-asset impacts are similar to those

[9]The literature (e.g., D'Amico et al., 2012) has identified three main channels of QE's transmission mechanism. The supply/scarcity (a.k.a., portfolio balance) channel works by changing the available supply of the targeted assets and tends to affect more the yields of those assets and their close substitutes. The signaling channel works through changes in the expected future path of the policy rate. The duration-risk channel affects term premiums by changing the amount of aggregate duration risk held by private investors.

[10]For example, Neely (2016) shows that time series models that estimate a quick reversal of the effects of QE announcements do not predict well asset prices.

observed following conventional monetary policy announcements.[11] This, overall, indicates that it is reasonable to assume that the Treasury yield effects identified by event studies are viewed by investors as largely persistent.

Finally, our quest for understanding whether BSP effects are state dependent has been in part motivated by the growing literature on the interplay between market imperfections and unconventional monetary policy. In most macro-finance models, changes in asset supply matter for asset prices and the macroeconomy only in the presence of market imperfections that restrict arbitrage opportunities, such as: capital constraints on financial intermediaries, limited risk-bearing capacity, elevated transactions costs, and limited market participation. Curdia and Woodford (2011), He and Krishnamurthy (2013), Gertler and Karadi (2011, 2013), Schabert (2015), and Droste et al. (2021) among others, emphasize alternative imperfections in financial markets that allow central bank purchases and sales to affect asset prices and, in turn, economic activity and inflation. Overall, all these studies lead to similar conclusions about the financial and macro impact of QE: if frictions and distortions are smaller outside of a crisis, then the benefits of central bank asset purchases are diminished. In our work, estimates of a large yield sensitivity to QT supply shocks, which occurred in normal market conditions and during an economic expansion, cast doubts on the empirical validity of some of these theoretical mechanisms.

2.1 A Guiding Framework

To provide some basic intuition of why the yield impact of BSP can be state dependent and hence differ during QE and QT (or across earlier and later programs), we start with the equation for Treasury bond returns resulting from equilibrium models of the term structure that exploit the interaction between preferred habitat and limited risk-bearing capacity to allow for the quantity of bonds to play a role (e.g., Vayanos and Vila, 2021; King, 2019; Ray, 2019; Droste et al, 2021).

In particular, we focus on the equations derived from King (2019), which is the only study that modifies Vayanos and Vila (2021) to account for the ZLB. Since in the

[11] For example, Gilchrist and Zakrajšek (2013) and Gilchrist, Lopez-Salido, and Zakrajšek (2015) find large and significant effects of QE on corporate bond yields and credit risk in the US, D'Amico and Kaminska (2019) find significant and persistent effects of various rounds of QE on corporate bond yields for the UK, Rosa (2012) and Mamaysky (2018) find significant effects of QE on equities and equity-implied volatility for a few countries, and Di Maggio et al. (forthcoming) show that the Fed's QE programs that did not include MBS purchases reduced mortgage rates.

US, QE has been employed only at the ZLB, and QT is usually implemented after the policy rate lifts off from the ZLB, we think that it is crucial to utilize a model that accounts for bond price non-linearity induced by the ZLB.

The time t expected excess return (i.e., risk premium) on a τ-maturity bond is given by:

$$E_t\left[\frac{dP_t^\tau}{P_t^\tau}\right] - r_t dt = a_t \beta_t \sigma_{r_t}^2 A_t^\tau \int_0^T A_t^s ds \tag{1}$$

where P is the bond price, r is the short rate, a_t is the coefficient of risk aversion, β_t is the bond supply, $\sigma_{r_t}^2$ is the short-rate variance, and A_t^τ is defined as the derivative of the bond price with respect to r, which is non-negative and increasing with τ. In a model that accounts for the ZLB, A_t^τ does not admit an analytical solution, but King (2019) shows that its first-order approximation is given by:

$$A_t^\tau \approx \int_0^\tau e^{-ks} Pr\left(r_{t+s} > 0\right) ds, \tag{2}$$

which is the discounted stream of probabilities that r will be above the ZLB in each of the next τ periods. The more likely r is to be at the ZLB between time t and $t+\tau$, the smaller A_t^τ will be. This suggests one of the possible reasons why, away from the ZLB (i.e., during QT), the impact of changes in bond supply can be different than at the ZLB (i.e., during QE). The higher r is, the larger A_t^τ is, and hence the larger the change in the risk premium will be in response to a given change in bond supply:

$$\frac{\partial E_t\left[dP_t^\tau / P_t^\tau\right]}{\partial \beta} = a_t \sigma_{r_t}^2 A_t^\tau \int_0^T A_t^s ds. \tag{3}$$

Equation (3) shows two other important reasons why the impact of changes in bond supply can be state dependent. Not only the risk premium is more responsive to changes in supply when A_t^τ is larger, but also when a_t and σ_{r_t} are higher. In most macro-finance models of QE, limited risk-bearing capacity, or a variety of market imperfections that restrict arbitrage, are used to motivate smaller QE effects outside crisis periods. The basic idea being that BSP becomes less potent in normal times or away from the ZLB because some market frictions and distortions begin to wane. As distortions wane, arbitrage becomes more efficient, and the asset price effects of changes in asset supply either get smaller or disappear faster, and BSP becomes less effective. Here, this type of mechanism is captured by a_t: the higher a_t is, the more risk averse arbitrageurs are, the more segmented bond markets are, and hence the more responsive the risk premium will be to a given change in supply. The opposite is true when a_t is low. But, in equation (3), because of the ZLB, a_t is only one of the factors that can affect the potency of BSP.

The level of the short rate not only matters for A_t^τ, it also matters for σ_{r_t}. As shown in Figure 1 of King (2019), σ_{r_t} is increasing in r, hence near or at the ZLB, interest rates display little volatility, causing changes in bond supply to have a muted impact. This implies that, even if QE is conducted when risk aversion is elevated potentially being more potent, its impact can be attenuated by a smaller σ_{r_t}, induced by an extended period of very low r. This, in turn, indicates that away from the ZLB (higher r), during QT, the attenuating effect of low risk aversion can be more than offset by the magnifying effect of higher σ_{r_t} and higher A_t^τ. Further, since the variance and covariance of bond returns is given by:

$$cov_t\left[\frac{dP_t^\tau}{P_t^\tau}, \frac{dP_t^s}{P_t^s}\right] = A_t^\tau A_t^s \sigma_{r_t}^2, \tag{4}$$

it is easy to see how the effect of higher σ_{r_t} and A_t^τ get passed through to the variances of longer-term bond returns or yields. In general, not only the ZLB, but any event that increases interest rate uncertainty, $\sigma_{r_t}^2$, could make all yields' volatilities and risk premiums more sensitive to changes in supply.

In summary, we have identified three potential factors, A_t^τ, a_t, and σ_{r_t}, that can affect the asset price effects of BSP. Since each of these factors can influence the impact of QE and QT in opposite directions, depending on market conditions as well as interest rate level and uncertainty, which of these factors dominates on balance during each BSP program is ultimately an empirical question. We proceed to investigate this question more rigorously in the reminder of our study.

3 Balance Sheet Policy Surprises and Data

Financial markets are inherently forward looking and react only to the new information contained in policy announcements. Therefore, to correctly identify the asset price effects of BSP announcements it is necessary to focus on the unexpected component of those announcements. In the event-study analysis of *conventional* monetary policy, it has been the standard for two decades to focus on the unexpected component of the change in the FFR around FOMC announcements. In contrast, in the event-study analysis of *unconventional* monetary policy, distinguishing the expected and unexpected component of BSP is far from being the standard approach. Similar to the seminal works of Kuttner (2001) for the FFR and Gürkaynak et al (2005) for forward guidance, we show that isolating the policy surprise is very important for evaluating BSP such as QE and QT.

In particular, we control for the pre-announcement market expectations to estimate the *total* asset supply surprise, that is, the unexpected component of the total size of the announced program; and then, we exploit the programs' operational details to compute the *local* asset supply surprises, that is, the unexpected component local to the maturity sectors in which purchases were conducted. The quality of these surprises' measurement depends on the availability of market participants' expectations about: (i) the likelihood of each asset purchase program being announced, (ii) the size and maturity distribution of purchases, (iii) the monthly pace of purchases and program length for open-ended QE, and (iv) monthly redemption caps for QT implementation. Here, we focus on the computation of the total asset supply surprise, our most robust surprise measure as it requires minimal assumptions.[12] We leave to the Appendix the computation of the local supply surprise, which requires the description of each program's operational details used to derive the expected maturity distribution of purchases. This is based on the assumption that, at the time of each announcement, all primary dealers are familiar with past and current operational details, which seems quite reasonable as all of them participate at the Fed auctions.

As shown in Cahill et al. (2013), to compute the expected change in asset supply for each fixed-size program k, typical of earlier QE, it is sufficient to multiply the probability of a given program announcement (Pr_t) by the expected size (Q) of the program, conditional to the announcement occurring:

$$E_t\left[BSP^k\right] = Pr_t^k * E_t\left[Q^k|announcement\right]. \tag{5}$$

[12] Our main assumption is that, in the case of QT, elevated uncertainty about the policy normalization process led market participants to heavily discount longer-term changes to the BSP, and hence we do not use information about expected changes in the SOMA portfolio beyond the one-year horizon.

However, a more nuanced approach is needed for BSP announcements pertaining to open-ended programs (typical of later QE) and to the implementation of QT. Such approach is novel to our study.

For open-ended programs in which, instead of announcing the total amount of asset purchases, policymakers announce a specific pace of monthly purchases that would continue until a certain threshold (e.g., a substantial improvement in the outlook for the labor market) is crossed, the expected change in asset supply is obtained by multiplying the probability of the program announcement by the expected monthly pace (q_m) and expected program length, summarized by the total number of months (M) in the purchase period:

$$E_t\left[BSP^k\right] = Pr_t^k * E_t\left[q_m^k|announcement\right] * E_t\left[M^k|announcement\right]. \qquad (6)$$

For QT, expectations about monthly redemption caps (c_m), which limit the amount of maturing securities the Fed would reinvest at auctions to gradually shrink the balance sheet, have to be subtracted from the portfolio share expiring each month (S_m^e):

$$E_t\left[BSP^k\right] = Pr_t^k * \left[S_m^e - E_t\left(cap_m^k|announcement\right)\right] * E_t\left[M^k|announcement\right]. \qquad (7)$$

Since only the amount exceeding the caps get reinvested, higher caps imply less reinvestment at auctions and hence a faster balance sheet run-off.

The unexpected (U) component or BSP surprise is simply given by the difference between the BSP actually announced and the expected BSP:

$$BSP_t^{U,k} = BSP_t^k - E_t\left[BSP^k\right]. \qquad (8)$$

Market expectations about all the variables in equations 5-7 are included in the SPD compiled by the New York Fed's Open Market Trading Desk (the Desk) *before* each FOMC announcement. Often, in a "flash"survey, the Desk also asks primary dealers to update their responses immediately *after* an FOMC meeting to gauge how expectations have changed due to new information. We use both regular and flash SPD to capture shifts in market expectations about the Fed's BSP. Clearly, the availability of a flash survey makes the measurement of the surprise more precise, as in this case the unexpected component is given by the difference in pre- and post-announcement BSP expectations:

$$BSP^{U,k}_{t+\delta} = E_{t+\delta}\left[BSP^k\right] - E_t\left[BSP^k\right], \tag{9}$$

where δ indicates the time interval between the pre- and post-FOMC SPD (i.e., the flash survey). This way of measuring the BSP surprise is particularly important for open-ended QE and QT as, in the case of these policies, expectations keep being updated during the course of multiple FOMC announcements, until the program's conclusion is telegraphed in advance.

3.1 Data

The detailed nature of the SPD data is the linchpin to our surprise computation. The results of the SPD have been public since 2011 but we have access to them since 2009. To illustrate the richness of the data, we focus on three examples from QT, which has never been analyzed before.[13] However, the Appendix describes the details of the SPD data and surprise computation for each QE/QT event used in this study.

The first example is the June 2013 FOMC announcement, part of the "Taper Tantrum" episode. For this meeting, the Desk conducted both a pre- and post-FOMC survey on June 10 and June 24, respectively. As summarized in Table 1, from these two surveys, for both the Treasury (top panel) and MBS (bottom panel) purchases, we obtained the expected monthly pace of purchases in billions and the expected start of tapering, the first reduction in the pace of purchases highlighted in yellow. From the June-10 to the June-24 SPD, the expected start of tapering moved from December 2013 to September 2013, suggesting that the FOMC indicated an earlier-than-expected reduction in the pace of purchases. Further, as shown in the last row of each panel, the change in the expected monthly pace of purchases from the June-10 to the June-24 SPD (a straightforward application of equation 9) implies a cumulative \$27.5bn hawkish Treasury surprise and a \$38bn hawkish MBS surprise, respectively, due to the smaller-than-expected amount of purchases in each asset class.

The second example is the June 2017 FOMC announcement, which detailed the intended redemption cap schedule to gradually reduce the size of the balance sheet. The announced caps for Treasuries and MBS are depicted in red in the top and middle panels of Figure 1, respectively. The expected Treasury and MBS cap sizes reported in

[13]Since the details of the data and computations for fixed-size programs have been previously described in Cahill et al. (2013), we relegate them to the Appendix.

Table 1: Expectations about the monthly pace of purchases ($bn) in June-2013 SPD

Treasuries	Jun13	Jul13	Aug13	Sep13	Oct13	Nob13	Dec13	Jan14	Feb14	Mar14	Apr14	May14	Jun14
June 10	45	45	45	45	45	45	30	25	20	15	10	5	0
June 24	45	45	45	40	35	32.5	30	25	20	15	10	5	0
Change				**-5**	**-10**	**-12.5**							

MBS	Jun13	Jul13	Aug13	Sep13	Oct13	Nob13	Dec13	Jan14	Feb14	Mar14	Apr14	May14	Jun14
June 10	40	40	40	40	40	40	30	25	20	15	0	0	0
June 24	40	40	40	35	33	29	25	20	15	10	5	0	0
Change				**-5**	**-7**	**-11**	**-5**	**-5**	**-5**	**-5**	**+5**		

the SPD are shown in blue. By comparing the red and blue lines, it is easy to see that the announced (red) Treasury cap schedule and terminal size turned out to be a bit higher than expected, implying overall less reinvestments and therefore less Treasury purchases than anticipated. In contrast, for MBS, the terminal cap size turned out to be lower than expected, implying more reinvestments and therefore larger MBS purchases than anticipated. The third and final example is the March 2019 FOMC announcement, which ended the balance sheet runoff by detailing the reduction in cap size. As seen in the bottom panel of Figure 1, overall, the announced caps (red) were lower than the expected caps (blue). Further, while primary dealers expected the cap reduction to begin in September 2019 and end sometime in 2019:Q4 (i.e., the cap size reaches zero), the FOMC announced that the cap reduction would start in May 2019 and stop in September 2019, implying larger-than-expected reinvestments and thus Treasury purchases.

In using the SPD median expectations as a measure of market expectations, it is worth bearing in mind that the SPD might not reflect the marginal investor's expectations very closely. Most of the survey's respondents are sophisticated investors and attentive "Fed watchers," likely less liable to surprise than the broad investor community. In some instances, comparison across different surveys has confirmed that primary dealers' thinking tends to be more homogeneous and more in tune with policymakers' thinking, potentially biasing downward our measure of the asset supply surprise.[14] Nevertheless, since we are more interested in comparing the magnitude of Treasury yield sensitivities across different events rather than just measuring their absolute size, we still favor the use of the SPD as it provides the views of the same type of investors consistently across FOMC-related events. More importantly, since

[14] For example, we have compared some of the SPD questions to similar questions in the Bloomberg survey and Blue Chip Survey, when similar questions were available, which however happened rarely as SPD questions tend to be more granular.

the primary dealers are among the largest investors that have the ability to trade and affect prices within the small time windows considered in our event-study, they tend to be the "representative investors" in those time windows around BSP announcements.

Finally, to measure and illustrate the Treasury market reaction, we use a new dataset consisting of intraday price quotes on all outstanding U.S. nominal Treasury securities for the 2009-2019 period from Thompson Reuters Tick History. On average, we have high-frequency information for about 200 securities at each point in time over a sample period of more than 10 years. These data allow us to capture individual yield changes in small time windows around BSP announcements, which are crucial to our new identification procedure described in Section 5. Usually the event window starts 15 minutes before the announcement and ends 15 minutes after the announcement, but it can extend up to two hours after the announcement, either to capture the reaction to the Fed Chair's press conference if during the Q&A session investors received additional information, or to give market participants more time to process relatively more complex BSP statements, for instance, addendums provided with the FOMC decision.[15]

4 Events providing an ideal testing ground

Across both QE and QT episodes, we use the following criteria to select events that can help identify the sensitivity of Treasury yields to a supply shock: (i) The presence of a sharp kink in the yield curve reaction to the announcement, (ii) availability of enough information to measure the supply surprise from the SPD, and (iii) the absence of significant policy-rate surprises. This last condition has been harder to meet because of the implicit or explicit forward guidance in every FOMC statement, but in Section 5, we discuss in detail why this should not be an issue for our identification.

A potential concern might be that focusing on events characterized by kinks could bias our estimates upward if those kinks were a manifestation of extreme market segmentation and lack of liquidity. But we do not think this is the case as, for example, all major QT announcements are characterized by kinks in the yield curve reaction and took place during periods of market calm and in the absence of market dysfunctions.

[15]We look for all available traded quotes that are the closest to the beginning and end of the event window, and we drop a given CUSIP from our sample if no quote exists in the 30 minutes before the start time or after the end time. This should ensure both enough reaction time and a precise identification of the reaction to a specific event.

As we describe below, it seems that the formation of kinks depends on the *granularity* of the information received during the announcements, rather than on market segmentation or other factors that could amplify the supply effects. In other words, kinks seem to be the result of very targeted trading executed by well informed investors, as demonstrated by the elaborate expectations of primary dealers described in the previous section. Further, as we will show in Section 6, in the case of earlier QE programs, our estimates of the supply effect are not larger than those reported in other studies (e.g., D'Amico and King, 2013; Cahill et al., 2013) that do not exploit kinks in their approach.

Finally, to support our claim that most of the events considered in this study, especially QT announcements, occurred during periods of normal market functioning, Figure 2 shows the time series of widely-used measures of Treasury market liquidity, limits to arbitrage, and financial constraints, over the sample period relevant to our analysis, 2009-2019. The top left panel plots the 10-year on-the-run premium, measured by the spread between the yields on the off-the-run and on-the-run 10-year notes (Gürkaynak, Sack, and Wright, 2007), which is a standard measure of liquidity for nominal Treasury securities (see for example Adrian, Fleming and Vogt, 2017). The next four panels display additional Treasury price anomalies documented in the literature: the 10-year TIPS-Treasury bond puzzle of Fleckenstein, Longstaff and Lustig (2014), proxied with the wedge between the 10-year inflation swap rate and the 10-year TIPS break-even rate; the off-the-run note-bond spread of Musto, Nini and Schwarz (2017), derived as the difference between the average yields of off-the-run Treasury notes and bonds with maturity as close as possible to 10 years; the 10-year TIPS liquidity premium implied by the term-structure model of D'Amico, Kim and Wei (2018); and the average absolute nominal yield curve fitting errors, which can be interpreted as a measure of limits to arbitrage (Hu, Pan, and Wang, 2013).[16] Finally, the bottom right panel depicts changes in broker-dealer leverage, obtained using the Fed Flow of Funds data, which captures the balance sheet capacity of financial intermediaries (Adrian, Etula, and Muir, 2014). As funding constraints tighten, balance sheet capacity falls and intermediaries are forced to deleverage by selling assets at fire sale prices. It is easy to note that after 2012, all the measures summarized in Figure 2 are back to their normal levels, some even as early as 2010, and balance sheet capacity growth is mostly

[16] The yield curve fitting errors are derived from the DTSM in Pancost (2021) and have been kindly provided by the author.

positive during QT quarters. This corroborates the idea that kinks are not the result of Treasury market dysfunctions, and that certainly during QT such dysfunctions cannot be at the origin of the supply effects that we estimate in Section 5.

4.1 During QE

The first large-scale asset purchase (LSAP1) announcement (March 18, 2009):

At 2:15 p.m. the FOMC announced its decision to increase purchases of agency MBS by $750 billion, agency debt by $100 billion, and to purchase up to $300 billion of longer-term Treasury securities over the subsequent six months. Without prior notice, the Desk announced at 2:44 p.m. that the Treasury purchases would be concentrated in the 2-to-10-year maturity sector. According to market commentaries and the SPD, those LSAP programs were larger than anticipated. In particular, according to the SPD, Treasury purchases had a 52% likelihood of being announced, leading to a dovish Treasury supply surprise of $142.5 billion; while, the MBS purchases were largely expected, leading to a dovish MBS supply surprise of $217.5 billion (see the Appendix for the surprise computation using eq. 5). In the 30-minute window around the FOMC announcement, as shown by the red dots in panel A of Figure 3, longer-term yields declined by more and plateaued at around −45 basis points at the 15-year maturity and beyond. Following the Desk announcement, shifts in market expectations about the maturity distribution of Treasury purchases created a sharp kink in the yield curve reaction, shown by the blue dots, as investors assigned much smaller probability to the purchase of securities with maturities above 10 years. This caused a price reversal among those securities, which created a kink around the 10-year maturity—the upper bound of the maturity range targeted by LSAP1 purchases.

The Reinvestment policy announcement (August 10, 2010):

At 2:15 p.m., the FOMC announced that it would keep the face value of the Fed's portfolio holdings constant by reinvesting principal payments from agency debt and agency MBS in longer-term Treasury securities. At 2:45 p.m., the Desk indicated that it would again concentrate Treasury purchases in the 2-to-10-year maturity sector, and it would refrain from purchasing securities for which the Fed already held large concentrations. According to market commentaries and the SPD, the reinvestment program was largely unanticipated, creating a dovish supply surprise of about $185.6 billion (see the Appendix); and, the novel aspect of the Desk statement implied that securities already held by the Fed in high concentration had a lower probability of being bought. At that time, the largest and smallest fractions of Treasury holdings were in the 10-to-17-year and 4-to-5.5-year maturity sectors, respectively.[17] Following the FOMC announcement, as shown by the red dots in panel B of Figure 3, the initial yield curve reaction was quite similar in shape to the one prevailing after the LSAP1 announcement. Following the Desk statement, as shown by the blue dots, securities with maturity longer than 10 years reversed part or all of their initial yield declines as they had a lower probability of being purchased. In contrast, yields of securities in the 4-to-6-year sector, which had a relatively higher probability of being bought, decreased further. This pattern created two kinks, around the 5- and 6.5-year maturities, respectively.

The first maturity extension program (MEP1) announcement (September 21, 2011):

At 2:23 p.m. the FOMC announced its intention to extend the average maturity of its security holdings by purchasing $400 billion of Treasury securities with remaining maturity between 6 and 30 years and selling an equal amount with remaining maturity of 3 years or less. According to the SPD, these relative supply changes were larger than expected, creating a dovish $146.5 billion supply surprise in longer-term Treasuries (see the Appendix). Following the announcement, as shown in panel C of Figure 3, yields increased at the front-end of the curve and decreased at the long-end of the curve. This pattern created a kink around the 3-year maturity, roughly the threshold between the maturity sectors targeted for sales and those targeted for purchases.

The MEP2 announcement (June 20, 2012):

At 12:30 p.m. the FOMC announced its decision to extend the MEP1 through the end of 2012, resulting in the purchase and sale of $267 billion of additional Treasury securities. The Desk contemporaneously released the operational details, which stated the suspension of the practice of rolling over maturing Treasury securities into new issues at auction. Although the MEP extension was in part expected by the SPD respondents, the decision to redeem securities, which allowed the Fed to further increase its longer-term purchases, was unexpected. Overall, this created a dovish $174.75 billion surprise in longer-term Treasuries (see the Appendix). Just like during

[17] The Fed's portfolio held at the time of the announcement the following shares of privately-held Treasury amount outstanding: about 13% in the 1.5-4-year maturity sector, 7% in the 4-5.5-year maturity sector, 14% in the 5.5-7-year maturity sector, 13% in the 7-10-year maturity sector, 27% in the 10-17-year, and 14% in the 17-30-year.

the MEP1 announcement, Treasury yields increased at the short end of the yield curve and decreased at the long end, generating again a kink around the 3-year maturity, shown in panel D of Figure 3.

4.2 During QT

It should be stressed that, in the case of QT announcements, the location of the kink is not simply connected to a new operational detail released during the announcement, but it is also related to the Desk's purchase rules in place since December 2013 and the reinvestment rules in place since December 2015. Those rules, by determining the maturity distribution of asset purchases/reinvestments, affect also the distribution of the asset supply surprise across maturity sectors of the yield curve. In general, the kink being the peak of the yield curve reaction, should be at the edges of the maturity sector characterized by the largest asset supply surprise. And, based on our computation of the local supply surprises in the Appendix, this is mostly the case in Figure 3. But, for the June 2013 announcement, the match is not perfect, as the kink is closer to the 5-year maturity rather than the 7-year maturity. This might suggest that, during that episode, other effects are operating and concentrating the trading in the sector perceived to be the most liquid.

The Post-Taper-Tantrum FOMC Meeting (June 19, 2013):

According to Bloomberg Newsfeed, QE-related remarks during the Fed Chair's press conference gathered more attention than the FOMC statement, as they were viewed as more hawkish than expected. Hence, our event study is centered around the time of those remarks. Specifically, market participants were very focused on the Chair's 2:38 p.m. remarks indicating that, if the economy evolved as expected by the Committee, the Committee anticipated "moderating the monthly pace of purchases" in the latter part of 2013 and ending purchases in mid-2014. Further, the Chair also indicated that the majority of the Committee did not expect to sell agency MBS during the policy normalization process. The change in market expectations induced by these remarks were already described in Table 1. As a result, Treasury yields increased in the 30-minute window around the Chair's key remarks, as illustrated in panel E of Figure 3, which exhibits a kink around the 5-year maturity.

The FOMC announcement delaying the start of tapering (September 18, 2013):

Despite market-wide expectations that the Fed was going to start tapering LSAP3 in September 2013, the FOMC "decided to await more evidence that progress will be sustained before adjusting the pace of its purchases." This decision had the exact opposite effect of the June 2013 FOMC announcement, as most primary dealers shifted again the expected start of tapering from September 2013 to December 2013, as shown in the responses to the pre- and post-FOMC SPD. Such a shift implied a larger-than-expected pace of purchases for three additional months, creating a dovish $95 billion Treasury supply surprise and a dovish $119.5 billion MBS supply surprise. The yield curve reaction to the dovish surprises is shown in panel F of Figure 3, which exhibits a kink around the 6-year maturity.

The FOMC announcement including the Normalization Addendum (June 14, 2017):

The June 2017 FOMC statement was accompanied by a Normalization Addendum providing investors with precise details about the monthly redemption caps' sizes and schedule. The difference between the announced and expected caps has already been illustrated in Figure 1. For Treasuries, that difference provided a hawkish supply surprise of about $78 billion that pushed yields higher, creating a kink around the 3-year maturity, shown in panel G of Figure 3. In contrast, for MBS, it created an $18 billion dovish surprise.

The FOMC announcement including the Balance Sheet Normalization Principles and Plans (March 20, 2019):

The FOMC announced the phasing out of the balance sheet reduction, which market participants viewed as occurring faster than they had anticipated, implying a larger terminal size of the Fed's balance sheet. This faster-than-expected tapering of QT, already illustrated in the bottom panel of Figure 1, provided a dovish supply surprise of about $52 billion; that is, the opposite effect of the faster-than-expected tapering of QE described in the June 2013 FOMC. As shown in panel H of Figure 3, following the announcement, Treasury yields decreased and formed a kink around the 5-year maturity.

5 Empirical Strategy

To identify and quantify the Treasury yield effect of the asset supply surprise, we use the yield curve reaction depicted in each panel of Figure 3. In particular, we exploit the relative yield changes of individual Treasury securities around the kink, hence our empirical strategy is conceptually similar to the regression kink design (RKD) described in Card et al. (2015, 2016). The idea of the RKD is to examine the slope of the relationship between the outcome of interest (changes in yields) and the assignment variable (maturity) at the exact location of the kink. Provided that securities on either side of the kink are "similar," any kink in the outcome can be attributed to the treatment effect of the policy variable (change in asset supply). To ensure the similarity condition across securities, we restrict the estimation to securities whose maturity is within three years of the kink, and in the robustness analysis we reduce the maturity distance from the kink to two years.

For each of the eight events described in the previous section, we estimate the following specification:

$$\Delta Y_{i,t} = \alpha + \beta_1 * (\tau_{i,t} - K_t) + \beta_2 * D_i * (\tau_{i,t} - K_t) + u_{i,t}, \tag{10}$$

where, $\Delta Y_{i,t}$ is the yield change of security i within a narrow time-interval t around the announcement, $\tau_{i,t}$ is the maturity of security i, K_t is the kink location in the maturity range, and D_i is a dummy variable that takes the value of one to the right of the kink (i.e., for maturities larger than the kink's maturity) and thus allows the slope to change at the kink.[18] Since the dependent variable is the shift in each yield due to the announcement, β_2 captures whether, on average, such shift is larger or smaller to the right of the kink. This should depend only on whether, on average, securities to the right of the kink will be bought in larger or smaller amounts than expected, relative to securities to the left of the kink.

β_2 identifies only the causal effect of the BSP surprise if everything evolves smoothly across the kink threshold except the derivative of the BSP surprise with respect to maturity. That is, the unexpected change in asset supply should exhibit a discrete jump across the two maturity sectors adjacent to the kink. These discrete jumps are an artifact of the intended maturity distribution of purchases announced in the BSP's operational details, reported in the Appendix for each event. This implies that, other potential BSP channels, such as changes in the policy rate path (signaling) and duration risk should not jump at the kink threshold in the small time window around the FOMC announcement. In other words, both the policy rate path and duration

[18]Higher order polynomials could be added to the baseline specification. However, the addition of a quadratic term increased the R-squared only marginally, except for June 2017 (from 0.62 to 0.75), so we stick to the linear specification.

risk can change as a result of the announcement, but as long as they change smoothly across similar maturities they do not affect the estimate of β_2.

Theoretically, these identification conditions are supported by any dynamic term-structure model, included that of Vayanos and Vila (2021), as the expectation hypothesis implies that the signaling channel is continuous along the maturity dimension and duration risk is continuous and monotonic along the maturity dimension. In contrast, Proposition 7 of Vayanos and Vila (2021) shows that supply effects can be discontinuous along the maturity dimension, what the authors call "full localization:" the case in which a supply shock originating at a certain maturity affects only the yields of that maturity.[19]

Empirically, the identification conditions are supported by the following observations. As already mentioned, the maturity of the securities adjacent to the kink is very similar as we limit each security's maturity distance from the kink to either three or two years, hence duration risk should not affect β_2. Regarding the signaling effect, it is very unlikely that it plays any role in the slope change around the kink for several reasons. First, most of the kinks are at or beyond the 4-year maturity, where the impact of the signaling channel tends to be quite small. Usually, changes in the policy-rate path affect the most short-term securities.[20] Second, during MEP1 and MEP2, when the kink is closer to the 3-year maturity, shorter-term yields went up and not down in response to a dovish monetary policy surprise. Third, changes in the expected policy-rate path can hardly create such abrupt kinks at a specific maturity. Calendar-based forward guidance could create such kinks, but such explicit forward guidance is absent from the FOMC statements used in our analysis.

Finally, for each event, the kink's precise location is determined using a gradient-based algorithm described in Section 2 of the Appendix, which delivers the following estimates: 10.035, 6.74, 2.625, 2.965, 5.07, 6.175, 2.145, and 4.635.

Our methodology has some important advantages relative to previous event studies of QE. First, it does not require combining yield changes across selected events leading to a BSP announcement. Second, as it can be noted from equation (10), the estimation of β_2 is independent of the measure of changes in asset supply or any other QE channel. Hence, relative to previous work that has already exploited operational changes in the maturity distribution of purchases to better identify QE supply effects (Cahill et al.,

[19]In their framework, demand effects are equivalent to our supply effects, as QE is modeled as an increase in the demand for Treasury securities by the Fed.

[20]See Swanson (2015), Greenwood, Hanson, and Vayanos (2015), and King (2019).

2013; McLaren et al., 2014), our methodology, not only does not require us to assume a specific distribution of the BSP surprise across individual securities, but also it does not require us to control for proxies of other QE channels. In other words, this new approach based on kinks is able to fully control for all explanations alternative to scarcity/supply effects.

In Cahill et al. (2013) and McLaren et al. (2014), the yield sensitivities to shifts in the local supply cannot be estimated without a measure of the surprise for each individual security. Arbitrary assumptions are inevitable in the computation of security-level supply surprises, as there is not such granularity of information about market expectations, creating variation in the value of the coefficient estimates across alternative assumptions. This is why we first estimate β_2 and then we normalize it by (i) either the total BSP surprise that hardly requires any assumptions, or (ii) the local BSP surprise, which being specific to an entire maturity sector rather than to each individual security within the sector, requires fewer arbitrary assumptions. Also, differently from those two studies, we do not need to control for surprises in other QE channels such as duration, signaling, and liquidity. In the study of McLaren et al. (2014), controlling for the level of duration is not sufficient to control for the duration and/or signaling channels, which require controlling for unexpected changes in duration risk and policy rates. Cahill et al. (2013) do control for unexpected changes in duration risk, but since it is done at the security level, it again requires some arbitrary assumptions.

6 Results

Since our primary goal is to analyze whether the yield impact of a BSP shock is state dependent, we focus first on the cross-sectional regressions that deliver for each event a separate β_2, which is then normalized by the surprise specific to that event. In the next section, we shift our focus to panel regressions that pool all QE and QT events together, allowing us to obtain a more robust estimate of the average β_2 across multiple BSP announcements.

6.1 Baseline

The results from our baseline specification are summarized in Tables 2 and 3, with the first column reporting the estimates of the local supply effect, β_2, for each FOMC announcement. In inspecting Table 2, it is worth bearing in mind two aspects of the estimates. First, the size of the β_2 coefficients is not normalized yet by the magnitude of the asset supply surprise, hence a comparison of their relative sizes is not warranted. Second, the sign of the coefficient depends on the relative size of the supply surprises in the maturity sectors adjacent to the kink. Specifically, if the maturity sector after the kink is perceived as being affected by a less (more) dovish surprise than the maturity sector before the kink, then the coefficient will be positive (negative). This implies that the sign of β_2 does not reflect the overall easing or tightening provided by the BSP decisions.

The main takeaways of Table 2 are purely statistical. The t-statistics reveal that the slope changes around the kink are highly significant for all events.[21] The adjusted R-squared are quite large, varying between 45 and 95 percent, indicating that the supply effect is an important driver of the yield reaction around BSP announcements. To visualize this first set of results, Figures 4 and 5 show the fit of the yield reaction in the 3-year maturity windows around the kink for QE and QT, respectively. For each event, the maturity at the kink is rescaled to equal zero, and the left and right panels illustrate the estimated slopes to the left and right of the kink. The difference between the right and left slope estimates is equivalent to β_2 in equation (10).[22]

Next, the Treasury yield *sensitivities* to a \$100 billion supply surprise are obtained dividing β_2 by the size of the asset supply surprise and multiplying by 100, and they are reported in the last two columns of Table 3. For each announcement, we provide a range of values for the estimated sensitivity, which depends on whether we use either the total or the local asset supply surprises to normalize β_2, shown in the second and third column of Table 3. Since these surprise measures rely on different assumptions about the degree of market segmentation, we believe that providing a lower (LB) and upper bound (UB) for the estimates is a more robust approach.

Specifically, the LB is obtained dividing β_2 by the total supply surprise, which does not require us taking a stance on the degree of market segmentation, as it implicitly assumes that the unexpected change in supply affects all Treasury securities, independently of their maturity. This provides the most conservative estimates of the yield

[21] Recall from Section 4.1 that in the case of the Reinvestment announcement there are two kinks. Our baseline estimates are obtained using the second kink, at the 6.75-year maturity, but for robustness we also use the first kink at the 5.1-year maturity and obtain very similar results. That is, β_2 equals 1.974, t-stat equals 11.4, and adjusted R^2 equals 0.839.

[22] Since for illustrative purposes those fitted lines have been obtained estimating two separate OLS regressions, while our main results are based on the single regression described in equation (10), β_2 is not exactly equal to the difference in slopes.

impact of a BSP shock, but not necessarily the most pertinent to the kink approach, which requires dividing β_2 by the difference in the BSP surprises around the kink. This is why the UB is obtained dividing β_2 by the local supply surprise, which being equal to the relative supply changes in the maturity sectors adjacent to the kink, implies a higher degree of market segmentation, that is, the securities around the kink react only to supply changes close to the kink.

Further, while in Table 3, the BSP surprises include only unexpected changes in Treasury holdings; in Table 4, the total asset supply surprises include also unexpected changes in MBS holdings, which provides an additional set of Treasury yield sensitivities, based on the alternative assumption of perfect substitutability between Treasuries and MBS.

By observing the last two columns of Table 3, it is easy to note that, on average, the yield sensitivity during QT events is larger than the yield sensitivity during QE events, implying that the impact of asset supply shocks did not diminish in normal market conditions and away from the ZLB. Further, both the lower and upper bound of the Treasury yield sensitivities do not fall monotonically over time, indicating that there is no evidence of diminishing returns of the supply/scarcity channel across subsequent BSP announcements. We reach the same conclusions about the evolution of the Treasury yield sensitivity even when we account for the MBS supply surprise in the total asset supply surprise, as shown in the last column of Table 4. We consider the LB of the estimated Treasury yield sensitivities in Table 3 as our reference estimates, since they are normalized by the total supply surprise that relies on minimal assumptions. It is reassuring, however, that the main message stays the same when we use alternative measures of the BSP surprises.

Overall, our results suggest that, in evaluating the relative efficacy of BSP announcements, it is very important to account for the size of the BSP shock, because doing so leads to conclusions and policy implications quite different from those reached in previous QE studies, which document decreasing impacts of QE announcements across subsequent programs and conclude that later rounds of QE were less effective than the first QE.[23] We also think that our novel findings are made more striking by the fact that are specific to the supply/scarcity channel, that is, the only QE channel that in macro-finance models requires some form of market dysfunction to work and is

[23] For example, Altavilla and Giannone (2017), Bauer and Neely (2014), Bowman, Londono, and Sapriza (2015), de los Rios and Shamloo (2017), Greenlaw et al. (2018), and Krishnamurthy and Vissing-Jorgensen (2011 and 2013).

predicted to have larger effects during financial crises.[24] In contrast, not only we find that the supply/scarcity channel remained powerful across vastly different financial market conditions, but its effects were apparently larger during the economic expansion. This indicates, for example, that in a model like that one used in Section 2.1, the amplifying effects of higher A_t^τ and σ_{r_t} dominate on average the attenuating effect of lower a_t. That is, variation in the level and uncertainty of interest rates is more important than variation in risk-bearing capacity or other forms of market stress.

Interestingly, the evidence reported in Table 3 and 4 also shows that the magnitude of the asset supply surprise is not the only factor driving the Treasury yield effects of BSP, as the yield sensitivities still display notable variation over time (despite being normalized by the BSP surprise). First, they seem to become particularly large at turning points of the BSP. For example, toward the end of QE (June 2013), as the Fed announced its intention to reduce the amount of policy accommodation provided through the balance sheet, and again toward the end of QT (March 2019), as the Fed announced its intention to stop tightening monetary policy and stabilize the size of its long-run portfolio. Due to the lack of experience with the balance sheet normalization process, those turning points could have been characterized by elevated BSP uncertainty that amplified the yield reaction. In Section 7, we explore this eventuality. Second, some variation in the yield sensitivity could be due to variation in the kink's location, which determines the average maturity of the yields affected by the surprise. In other words, in LSAP1 we estimate supply effects around the 10-year maturity, but in QT we mostly estimate supply effects around the 5-year maturity. This would matter if preferred habitat or limits to arbitrage vary across segments of the yield curve, which cannot be excluded. But, we observe the largest yield sensitivities for events where the kink is located around the 3- and 5-year maturities, which are among the most liquid segments and usually not the preferred habitat of institutional investors. Third, convexity hedging could also be an amplifying factor. That is, following a positive interest rate shock due to the BSP announcement, as in June 2013 and 2017, agency MBS prepayments are expected to decline, and hence MBS average duration in investors' portfolios extends. Consequently, investors tend to sell longer-term Treasuries to shed duration fast from their portfolio, and such Treasuries' selloff pushes prices down and interest rates further up, amplifying the initial shock.

[24]Differently from the supply/scarcity channel, the signaling and duration-risk channels do not rely on any form of limits to arbitrage. See for more detail Greenwood and Vayanos (2014), King (2019), and Bauer and Rudebusch (2014).

We believe that this could have been a relevant amplifying factor of the yield reaction to the June 2013 and June 2017 QT announcements.

Finally, Tables 5, 6, and 7 report the results of our robustness checks, which consist of re-estimating the β_2 coefficients and yield sensitivities using securities with maturity distance within two years from the kink rather than three years. As can be noted, the results are qualitatively very similar to those of our baseline specification. We have also re-done the estimation using securities whose duration rather than maturity is within two years from the kink, not shown for brevity, and the results are again qualitatively similar, confirming our main conclusions. Finally, we have re-estimated all our regressions using a 2-hour window around all the announcements, as in the case of the MEP2, June-2013, and June-2017 events, the time window was tighter to allow a better identification. Even if the estimates of β_2 (not shown for brevity) become a bit smaller for the MEP2 and June-17 announcements, while are little changed for the June-2013 announcement, the patterns of the LB and UB of the yield sensitivity convey the same message as the baseline results.

7 Interest-rate uncertainty as magnifying force

In this section, we investigate whether the Treasury yield impact is amplified by investor uncertainty about BSP, measured by volatilities derived from swaption contracts[25] written on the 10-year swap rate and with expiration ranging from 1 month to 10 years.[26] Differently from other interest rate derivatives, swaptions are very liquid for long-term rates and at long horizons, those most affected by BSP actions. In particular, QE-related shocks are estimated to have their largest impact at the 7-year maturity and beyond, while forward-guidance shocks have their largest impact between the two- and five-year maturity (e.g., Swanson, 2015; Greenwood, Hanson, and Vayanos, 2015; and King, 2019).

The red bars in each panel of Figure 6 summarize changes in the term structure of investor uncertainty about the 10-year rate over the six weeks *prior* to each FOMC

[25] A swaption is an option on an interest rate swap that gives the owner the right to enter an interest rate swap at a predetermined fixed rate when the option expires. That is, the owner of a payer swaption pays the fixed rate and receives a sequence of Libor rates at semiannual intervals over the life of the swap.

[26] Swaptions have been used in other studies to analyze how interest rate uncertainty varies around macroeconomic and monetary policy announcements. See for example Fornari (2004) and Hattori, Schrimpf, and Sushko (2016).

meeting. Panel A shows that before the LSAP1 announcement, the 10-year rate uncertainty increased substantially up to 4 years ahead. This is perhaps not surprising as there was a lot of uncertainty about the possibility of a Treasury purchase program that, as already discussed in Section 4.1, was not correctly anticipated by many market participants. Ahead of the subsequent QE announcements, as shown in the remaining left panels, the term structure of the 10-year rate uncertainty moved lower, most likely indicating that Fed communications about later QE programs had improved somewhat.

In contrast, ahead of the first QT announcement (panel E), investor uncertainty about the 10-year rate increased considerably, especially at shorter horizons. This is by far the largest increase in BSP uncertainty across the eight events, and it occurred during the Taper-Tantrum period, when investors seemed very uncertain about QE tapering. Similarly, also prior to the September 2013 announcement that delayed the start of tapering, the 10-year rate uncertainty increased somewhat, but much less than in the Taper-Tantrum period. Finally, ahead of the last two QT events (panels G and H), the term structure of the 10-year rate uncertainty declined, most likely because the BSP normalization process was well under way.

The blue bars in Figure 6 illustrate instead how the term structure of the 10-year rate uncertainty changed *on the day of* each BSP announcement. As shown in panel E, following the June 2013 announcement, the 10-year rate uncertainty spiked at very short horizons. This sharp increase is a bit surprising as the FOMC communication was in line with the Fed Chair Bernanke's testimony in front of Congress a month earlier. However, since market-based measures of uncertainty are contaminated by risk premiums, it is possible that part of the increase is due to higher risk aversion. This, in turn, would suggest that investors did not view favorably the upcoming tapering of QE purchases and possibly considered it premature. Further, also following the June 2017 FOMC meeting, the 10-year rate uncertainty increased a bit over the near term.

Overall, Figure 6 shows that the Fed's pre- and post-FOMC communication about BSP decisions could have been important for fluctuations in longer-term interest rate uncertainty, which might have played a role in the investors' reaction to BSP announcements. To investigate this possibility, we modify our empirical specification to account for investor uncertainty about upcoming BSP actions.

In particular, to capture whether market uncertainty about the 10-year rate was *unusually elevated ahead of* a specific FOMC meeting, we proceed in three steps.

First, for each forecasting horizon (i.e., swaption expiration), we compute the average 10-year rate uncertainty prevailing over the 10 days prior to the meeting.[27] Second, to collapse the term-structure of these average implied volatilities into one number for each FOMC meeting, we compute their weighted sum using weights that are inversely related to the length of the forecasting horizon, so that near-term volatilities get a higher weight and we still preserve the information contained in the entire term structure of uncertainty. (We use two versions of this uncertainty variable, one with a maximum horizon of 5 years and one with a maximum horizon of 10 years). Finally, this proxy of BSP uncertainty is normalized dividing by the average implied volatility prevailing in the year before the FOMC meeting and subtracting one. In this way, a value equal to 0 indicates that BSP uncertainty in the 10 days prior to the FOMC was not elevated relative to the previous year, while a large and positive (negative) value indicates that BSP uncertainty was particularly higher (lower) relative to the previous year. This normalization is necessary to account for the average interest-rate uncertainty that characterized a certain state of the economy, and in particular to control for the unusual uncertainty of the global financial crisis. Our measures of BSP uncertainty ahead of each FOMC announcement are summarized in Table 8. It is interesting to note that, following the normalization, the event characterized by the highest BSP uncertainty is the September 2013 announcement and not the June 2013 announcement. The latter, however, together with the LSAP1 and MEP1 announcements, remains one of the four events with somewhat unusually elevated BSP uncertainty.

Then, we pool together all eight BSP events in one panel and estimate three different specifications for our panel regression. First, a specification identical to equation (10) but with event fixed effects, as the estimated β_2 of this regression provides us with the average "supply/scarcity effect" across all eight BSP announcements and is, therefore, a useful term of comparison before the inclusion of BSP uncertainty. Second, we augment the first specification by interacting the two main regressors with our proxy of the BSP uncertainty, BPS_unc:

$$\begin{aligned}\Delta Y_{i,t} &= \alpha + \beta_1 * (\tau_{i,t} - K_t) + \beta_2 * D_i * (\tau_{i,t} - K_t) + \beta_3 * (\tau_{i,t} - K_t) * BSP\_unc + \\ &\quad \beta_4 * D_i * (\tau_{i,t} - K_t) * BSP\_unc + u_{i,t}, \end{aligned} \tag{11}$$

[27] For robustness, we have also computed the average over 5 or 20 days prior to each meeting and results are very similar, therefore are not shown for brevity.

where β_4 measures the amplifying effect of the excess 10-year interest-rate uncertainty ahead of the BSP announcement, with BPS_unc being calculated using the swaption-implied volatilities up to the 5-year horizon. Finally, the last specification is the same as in equation (11) but BPS_unc is calculated using the swaption-implied volatilities up to the 10-year horizon.

As shown in Table 9, on average, the supply/scarcity effect of a BSP announcement is of about -2.34 basis points per \$110bn, which is the average size of the BSP surprise across our events, and all coefficients are statistically significant; but, if ahead of the FOMC meeting investor uncertainty about the 10-year interest rate is unusually elevated, then the average effect is of about -7.8 basis points, as β_4 equals almost -5.5 basis points and it is statistically significant. This implies that elevated interest rate uncertainty makes the supply effect of a given BSP announcement about three times as large, explaining a significant amount of variation in the estimates of the β_2 coefficients and related yield sensitivities in Table 2 and 3.

Further, as shown in the last two rows of Table 9, those estimates are unchanged if we use our second proxy of uncertainty that considers changes in implied volatilities up to 10 years ahead. Finally, as shown in Table 10, the coefficient estimates are also very little changed when we restrict the maturities to 2 years around the kink, which should be expected if the duration risk played no role in the estimates resulting from the baseline specifications. This also indicates that the magnitude of the supply/scarcity effect, varying between 2.3 and 7.8 basis points per \$110*bn* BSP surprise, is quite meaningful because it does not include the influence of the signaling and duration-risk channels, found to be significant in the QE literature.[28]

However, it is possible that BSP uncertainty might be capturing some of the effects related to market functioning, as states characterized by high risk aversion and distressed market conditions can be associated to high interest rate volatility. Considering the relevance of this new result, not only for the magnitude of BSP effects but also for theoretical models of QE, it is important to disentangle the impact of BSP uncertainty from that of market dysfunction. To this purpose, we augment equation (11) with the interaction between the two main regressors and measures of market

[28]See for example Bauer and Rudebusch (2014), King (2019), and Greenwood and Vayanos (2014).

functioning (mf):

$$\Delta Y_{i,t} = \alpha + \beta_1 * (\tau_{i,t} - K_t) + \beta_2 * D_i * (\tau_{i,t} - K_t) + \beta_3 * (\tau_{i,t} - K_t) * BSP\_unc + \beta_4 * D_i * (\tau_{i,t} - K_t) * BSP\_unc + \beta_5 * (\tau_{i,t} - K_t) * mf + \beta_6 * D_i * (\tau_{i,t} - K_t) * mf + u_{i,t} \quad (12)$$

where β_6 captures the amplifying effect of abnormal market conditions ahead of each BSP announcement. This is because mf is proxied by either the 10-year GSW on-the-run premium or the yield curve fitting errors (shown in Figure 1), and is computed as the BPS_unc. That is, for each mf proxy, we take the average over the 10 days prior to the FOMC meeting, divide it by the average prevailing in the year prior to the same meeting, and subtract one. Hence, similarly to the BSP uncertainty measure, a value far from 0 indicates that market conditions in the 10 days prior to the FOMC were abnormal relative to the previous year. The two measures of market functioning are reported in Table 11.

The top panel of Table 12 shows the results obtained using the on-the-run premium to proxy mf, while the bottom panel reports the results obtained using the yield curve fitting errors to proxy mf. In both cases, we estimate two specifications because we use the 5- and 10-year horizon BPS_unc, precisely as in Table 9. In all the regressions, the number of observations is 861 and the adjusted R-square is 0.98. In the case of the on-the-run premium, the coefficients on the additional interaction terms (β_5 and β_6) are not statistically significant and they do not affect at all the magnitude of β_4, suggesting that the impact of BSP uncertainty is distinct from that of Treasury market distortions such as the on-the-run premium. In the case of the yield curve fitting errors, β_4 gets smaller (from about 5.5bp to 3.7bp), but its size and statistical significance are still notable, corroborating the idea that the amplifying influence of BSP uncertainty is a factor separate from market frictions.

7.1 Economic Relevance of Estimated Supply Effects

To better understand the relevance of our estimated supply effects, first, we quantify the total supply effects of each QE program implemented by the Fed before the current pandemic, and second, we compare those effects to the estimates obtained in other studies, which include all BSP channels' impact. For each QE program, the total supply effect is obtained multiplying the estimated Treasury yield sensitivity to one-billion supply shock by the amount of Treasury purchases (reported in parentheses in the first column of Table 13).[29] The resulting total supply effects are summarized in the second column Of Table 13. Over time there have been numerous studies that quantify how the Fed's purchase programs have affected longer-term interest rates. The analysis varies from event studies to regression-based analysis and term structure model estimates. The third column of Table 13 reports the studies for which we were able to retrieve the overall impact of *only* the Treasury component of each QE program, which is comparable to our estimates; and the fourth column summarizes the average effect of each program across the available studies. Overall, as shown in the last row of Table 13, our estimates of the supply effect account for about half of the overall QE effect estimated in the literature (65bps out of 129.5bps), suggesting that the magnitude of the supply channel is sizable and economically relevant relative to the other channels of the BSP.

To put our estimates of the size of the supply effect further into context, rather than focusing only on the BSP conducted in the US, we also compare our results to the international evidence available on the supply/scarcity channel of QE. Even if QE has been carried out in major advance economies since 2008, there are not many studies that isolate the contribution of the supply channel. Among the few that, to the best of our knowledge, have reported estimates of the supply/scarcity effect, there is significant variation. Similarly to us, McLaren et al. (2014) find that the supply effect accounts for about 40-60% of the total impact of the Bank of England's 2009-12 QE programs. In particular, the first £200bn of purchases are estimated to have reduced gilt yields with maturity between 5- and 25 years by 93bp, with local supply effects accounting for 48bp. In contrast, Altavilla et al. (2015) find that supply effects account for a small share of the 65bp reduction in yields estimated for the 2015 asset purchase program (about €1.0tr) of the European central bank (ECB). As pointed out by the authors, differently from the US and UK, the ECB targeted long-term sovereign securities characterized by different degrees of creditworthiness, and therefore it is not surprising that the credit and duration risk channels account for the bulk of the total effect. Indeed, Eser and Schwaab (2016) report that the effects of the 2010-11 ECB asset purchase program for Greece, Italy, Portugal, and Spain were quite large (around 150-200bp) but not unreasonable, as yields had reached levels as high as 20% at the

[29] Specifically, we take β_2 from Table 9 and divide it by $110bn$, the average surprise across all events, and then we multiply it by the total amount of Treasury purchases. In the case of LSAP1, we use the sum of β_2 and β_4, considering the consensus in the QE literature about the exceptional circumstances that characterized this program.

end of 2011.

8 Conclusions

Exploiting kinks in the yield curve reaction within narrow time-intervals around BSP announcements, we estimate the Treasury yield sensitivity to asset supply surprises induced by unexpected changes in QE and QT programs, with the latter being hardly studied before. By including in our 10-year sample (2009-2019) BSP events that span very diverse macroeconomic and financial market conditions, we can examine the state dependence of the supply effect of BSP and whether such policy has diminishing returns.

Our analysis indicates that BSP's supply effects do not fall monotonically across later QE and QT announcements and do not get smaller in normal financial market conditions and away from the ZLB. This is quite the opposite of what is predicted by most macro-finance models of central bank asset purchases. Further, we also show that not only investor expectations but also investor uncertainty about BSP is very important for the Treasury yield effects of QE and QT. This is relevant for thinking about different mechanisms that could be included in modelling approaches of BSP to produce outcomes closer to the evidence reported in this paper. Certainly, explicitly accounting for the ZLB (King, 2019) and BSP uncertainty (Droste et al., 2021) seem important steps in the right direction. But, most likely, the key is to model imperfect asset substitutability (salient for the impact of QE) independently of market imperfections that limit arbitrage. This, for instance, can be achieved by having that all agents in the model are risk-averse and optimize a portfolio comprising of multiple assets, but have different preferences for those assets (e.g., Koijen and Yogo, 2019; Koziol and Riedler, 2021).

Finally, our findings have relevant implications for the conduct of monetary policy. First, they seem to suggest that the usefulness of BSP might not be limited to exceptional circumstances such as the global financial crisis and the Covid-19 pandemic. Second, they indicate that it could be very valuable to provide forward guidance about the BSP, and not just about the policy rate. This is because such forward guidance, by allowing the Fed to better shape expectations and uncertainty about future asset supply, could make BSP more efficient, that is, could deliver larger and more persistent asset price effects with potentially smaller changes to the central bank's balance sheet.

Bibliography

Adrian, T., E. Etula, and T. Muir, 2014. Financial Intermediaries and the Cross-Section of Asset Returns. Journal of Finance 66 (6), 2557-2596.

Adrian, T., M. Fleming, and E. Vogt, 2017. An Index of Treasury Market Liquidity: 1991-2017. Federal Reserve Bank of New York Staff Report 827.

Altavilla, C., G. Carbonia, and R. Motto, 2021. Asset Purchase Programs and Financial Markets: Lessons from the Euro Area. International Journal of Central Banking 17 (4), 1-48.

Altavilla, C., and D. Giannone, 2017. The Effectivness of Non-Standard Monetary Policy Measures: Evidence from Survey Data. Journal of Applied Econometrics 32, 952-964.

Bauer, M. D., and C. J. Neely, 2014. International Channels of the Fed's Unconventional Monetary Policy. Journal of International Money and Finance 44, 24-46.

Bauer, M. D., and G. D. Rudebusch, 2014. The signaling channel for Federal Reserve bond purchases. International Journal of Central Banking, 10 (3): 233-89.

Bernanke, B. S., 2020. The New Tools of Monetary Policy. American Economic Review, 110 (4): 943-83.

Bonis, B., J. Ihrig, and M. Wei, 2017. The Effect of the Federal Reserve's Securities Holdings on Longer-term Interest Rates. FEDS Notes. Washington: Board of Governors of the Federal Reserve System, April 20.

Bowman, D., J. Londono, and H. Sapriza, 2015. U.S. Unconventional Monetary Policy and Transmission to Emerging Market Economies. Journal of International Money and Finance 55, 27–59.

Cahill, M. E., S. D'Amico, C. Li, and J. S. Sears, 2013. Duration Risk versus Local Supply Channel in Treasury Yields: Evidence from the Federal Reserve's Asset Purchase Announcements, FEDS working paper 2013-35.

Card D., Lee David S., Pei Z., Weber Andrea., 2015. Inference on Causal Effects in a Generalized Regression Kink Design. Econometrica, Vol. 83, No. 6, 2453–2483.

Card D., Lee David S., Pei Z., Weber Andrea, 2016. Regression Kink Design: Theory and Practice. NBER Working Paper 22781.

Curdia, V. and M. Woodford, 2011. The Central-Bank Balance Sheet as an Instrument of Monetary Policy. Journal of Monetary Economics, 58 (October), 54-79.

D'Amico, S., English, W., López-Salido, D., Nelson, E., 2012. The Federal Reserve's Large-Scale Asset Purchase Programs: Rationale and Effects. Economic Journal 122, November, 415-446.

D'Amico, S. and I. Kaminska, 2019. Credit Easing versus Quantitative Easing: Evidence from Corporate and Government Bond Purchase Programs. Bank of England Staff Working Paper, 825, September.

D'Amico, S., D. H. Kim, and M. Wei, 2018. Tips From TIPS: The Informational Content Of Treasury Inflation-Protected Security Prices. Journal of Financial and Quantitative Analysis, 53(1): 395-436.

D'Amico, S., 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, 425-448.

de los Rios, A. D., and M. Shamloo, 2017. Quantitative Easing and Long-Term Yields in Small Open Economies. IMF Working Paper 17/212.

Di Maggio, M., A. Kermani, and C. Palmer (forthcoming). How Quantitative Easing Works: Evidence on the Refinancing Channel. Review of Economic Studies.

Ederington, L. H., Lee, J. H., 1996. The Creation and Resolution of Market Uncertainty: The Impact of Information Releases on Implied Volatility. Journal of Financial and Quantitative Amalysis 31, No 4, 513-39.

Engen, E. M., T. Laubach, and D. Reifschneider, 2015. The MacroeconomicEffects of the Federal Reserve's Unconventional Monetary Policies. Finance and Economics-Discussion Series 2015-005. Washington: Board of Governors of the Federal Reserve System.

Eser, F, and B. Schwaab, 2016. Evaluating the impact of unconventional monetary policy measures: Empirical evidence from the ECB's Securities Markets Programme. Journal of Financial Economics, joint with Fabian Eser, Vol. 119 (1), 147–167.

Fleckenstein, M., F. A. Longstaff, and H. Lustig, 2014. The TIPS Treasury Bond Puzzle. The Journal of Finance, 69(5): 2151-2197.

Fornari, F., 2004. Macroeconomic Announcements and Implied Volatilities in Swaption Markets. BIS Quarterly Review, September, 79-86.

Gagnon, J., Raskin, M., Remache, J., Sack, B., 2011. Large-scale asset purchases by the Federal Reserve: did they work? International Journal of Central Banking 7 (1): 3-43.

Gertler, M. and P. Karadi, 2011. A Model of Unconventional Monetary Policy. Journal of Monetary Economics, 58 (October), 17-34.

Gertler, M. and P. Karadi, 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, January.

Gilchrist, S., Lopez-Salido, D., and Zakrajšek, E., 2015. Monetary Policy and Real Borrowing Costs at the Zero Lower Bound. American Economic Journal: Macroeconomics 7(1), 77-109.

Gilchrist, S. and Zakrajšek, E., 2013. The Impact of the Federal Reserve's Large-Scale Asset Purchase Programs on Corporate Credit Risk. Journal of Money, Credit and Banking, 2013, vol. 45, issue s2, 29-57.

Greenlaw D., Hamilton J. D., Harris E., West K. D. A skeptical view of the impact of the Fed's balance sheet. NBER Working Paper 24687.

Greenwood, R., Hanson S., Vayanos, D., 2015. Forward Guidance in the Yield Curve: Short Rates versus Bond Supply. NBER Working Paper No. 21750.

Greenwood, R., Vayanos, D., 2014. Bond supply and excess bond returns. Review of Financial Studies 27, 663-713.

Gürkaynak R. S., B. Sack, E. Swanson, 2005. Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements. International Journal of Central Banking 1(1), 55-93.

Gürkaynak, R. S., B. Sack, and J. H.Wright, 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.

Hartley, J. and Rebucci, A, 2020. An Event Study of COVID-19 Central Bank Quantitative Easing in Advanced and Emerging Economies. Working Paper, May 22.

Hattori, M., A. Schrimpf, and V. Sushko, 2016. The Response of Tail Risk Perceptions to Unconventional Monetary Policy. American Economic Journal: Macroeconomics, 8 (2), 111-36.

He, Z. and A. Krishnamurthy, 2013. Intermediary Asset Pricing. The American Economic Review, 103 (February), 732-770.

Hesse, H., B. Hofmann, and J. M. Weber, 2018. The macroeconomic effects of asset purchases revisited. Journal of Macroeconomics 58(C), 115-138.

Hu, G. X., J. Pan, and J. Wang, 2013. Noise As Information For Illiquidity. The Journal of Finance, 68(6): 2341-2382.

Joyce, M. A. S., Lasaosa A., Stevens I., Tong M., 2011. The Financial Market Impact of Quantitative Easing in the United Kingdom. International Journal of Central Banking, vol. 7(3), 113-161.

Joyce, M. A. S., Tong, M., 2012. QE and the Gilt market: a disaggregated Analysis. Economic Journal 122, November, 348-384.

Kim, K., T. Laubach, and M. Wei, 2020. Macroeconomic Effects of Large-Scale Asset Purchases: New Evidence. FEDS Working Paper No. 2020-47.

King, T. B., 2019. Expectation and duration at the effective lower bound. Journal of Financial Economics, Volume 134, Issue 3, 736-760.

Koijen, R. S. and M. Yogo, 2019. A demand system approach to asset pricing. Journal of Political Economy 127(4), 1475–1515.

Koziol, T. and J. Riedler, 2021. Scaling, unwinding and greening QE in a calibrated portfolio balance model. ZEW Discussion paper, 21-086.

Krishnamurthy, A., Vissing-Jorgensen, A., 2011. The effects of quantitative easing on interest rates: channels and implications for policy. Brookings Papers on Economic Activity 43, 215-87.

Krishnamurthy, A., Vissing-Jorgensen, A., 2013. The Ins and Outs of Large Scale Asset Purchases. Kansas City Federal Reserve Symposium on Global Dimensions of Unconventional Monetary Policy, 2013

Kuttner, K., 2001. Monetary Policy Surprises and InterestRates: Evidence from the Fed Funds Futures Market. Journalof Monetary Economics, 523–44.

McLaren, N., Banerjee, R. N., and Latto, D., 2014. Using changes in auction maturity sectors to help identify the impact of QE on gilt yields. The Economic Journal 124, May 453–479.

Meaning, J., and Zhu, F., 2012. The Impact of Federal Reserve Asset Purchase Programs: Another Twist. BIS Quarterly Review, Bank for International Settlements, March.

Musto, D., G. Nini, and K. Schwarz, 2017. Notes on Bonds: Illiquidity Feedback During the Financial Crisis. Review of Financial Studies, vol. 31(8), pages 2983-3018.

Neely, C., 2016. How Persistent Are Unconventional Monetary Policy Effects?. Federal Reserve Bank of St. Louis.

Pancost, A., 2021. Zero-Coupon Yields and the Cross-Section of Bond Prices. The Review of Asset Pricing Studies, Volume 11, Issue 2, 209–268.

Rosa, C., 2012. How "unconventional" are large-scale asset purchases? The impact of monetary policy on asset prices. Federal Reserve Bank of New York Staff Reports, 560, May.

Schabert, A., 2015. Optimal Central Bank Lending. Journal of Economic Theory, 157 (January), 485-516.

Swanson, E. T., 2011. Let's Twist Again: A High-Frequency Event-Study Analysis of Operation Twist and its Implications. Brookings Papers on Economic Activity, p. 151-188, Spring.

Swanson, E. T., 2015. Measuring the Effects of Unconventional Monetary Policy on Asset Prices. In Albagli, Elias, Diego Saravia, and Michael Woodford (eds.), Series on Central Banking, Analysis and Economic Policies 24: Monetary Policy through Asset Markets Lessons from Unconventional Measures and Implications for an Integrated World (Santiago, Chile: Banco Central de Chile, 2016), 105-130.

Swanson, E. T., 2021. Measuring the Effects of Federal Reserve Forward Guidance and Asset Purchases on Financial Markets. Journal of Monetary Economics, 118 32-53.

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

Table 2: Treasury Yield Impact to the Right of the Kink

	β_2 in bps	T-Stat	Adj R-Sq	N. of obs.
LSAP1	2.278	5.12	0.783	27.0
Reinvestment	1.131	9.32	0.712	70.0
MEP1	-4.701	-22.9	0.869	97.0
MEP2	-1.572	-11.9	0.748	94.0
Jun13 FOMC	-2.977	-34.1	0.946	138
Sep13 FOMC	3.353	7.67	0.450	106
Jun17 FOMC	-2.277	-20.7	0.720	170
Mar19 FOMC	1.387	16.6	0.801	159

Table 3: Treasury Yield Sensitivity to a $100 billion BSP Surprise

	β_2 in bps	Total Surp.	Local Surp.	Sensit. LB	Sensit. UB
LSAP1	2.278	143	74.7	1.59	3.048
Reinvestment	1.131	186	77.5	0.61	1.46
MEP1	-4.701	147	127	3.209	3.711
MEP2	-1.572	175	117	0.900	1.343
Jun13 FOMC	-2.977	27.5	11.3	10.8	26.2
Sep13 FOMC	3.353	95.0	39.2	3.530	8.56
Jun17 FOMC	-2.277	78.2	12.0	2.912	19.0
Mar19 FOMC	1.387	50.8	5.63	2.730	24.6

Table 4: Treasury Yield Sensitivity to a $100 billion BSP Surprise including MBS

	β_2 in bps	Treasury + MBS Surprise	Sensitivity with MBS
LSAP1	2.278	360	0.633
Reinvestment	1.131	139	0.813
MEP1	-4.701	147	3.209
MEP2	-1.572	175	0.900
Jun13 FOMC	-2.977	65.5	4.545
Sep13 FOMC	3.353	215	1.563
Jun17 FOMC	-2.277	60.2	3.782
Mar19 FOMC	1.387	50.8	2.730

Table 5: Treasury Yield Impact at the Kink, 2-year bandwidth

	β_2 in bps	T-Stat	Adj R-Sq	N
LSAP1	2.719	3.477	0.626	18.0
Reinvestment	1.839	9.745	0.728	47.0
MEP1	-5.115	-15.7	0.773	74.0
MEP2	-1.247	-6.333	0.499	75.0
Jun13 FOMC	-2.655	-23.8	0.894	95.0
Sep13 FOMC	2.364	2.677	0.072	70.0
Jun17 FOMC	-2.412	-16.2	0.710	142
Mar19 FOMC	0.748	8.857	0.522	103

Table 6: Treasury Yield Sensitivity to a $100 billion BSP Surprise, 2-year bandwidth

	β_2 in bps	Total Surp.	Local Surp.	Sensit. LB	Sensit. UB
LSAP1	2.719	143	74.7	1.908	3.639
Reinvestment	1.839	186	77.5	0.991	2.374
MEP1	-5.115	147	127	3.492	4.038
MEP2	-1.247	175	117	0.714	1.065
Jun13 FOMC	-2.655	27.5	11.3	9.655	23.4
Sep13 FOMC	2.364	95.0	39.2	2.488	6.033
Jun17 FOMC	-2.412	78.2	12.0	3.084	20.1
Mar19 FOMC	0.748	50.8	5.63	1.473	13.3

Table 7: Treasury Yield Sensitivity to a $100 billion BSP Surprise including MBS, 2-year bandwidth

	β_2 in bps	Treasury + MBS Surprise	Sensitivity with MBS
LSAP1	2.719	360	0.755
Reinvestment	1.839	139	1.321
MEP1	-5.115	147	3.492
MEP2	-1.247	175	0.714
Jun13 FOMC	-2.655	65.5	4.054
Sep13 FOMC	2.364	215	1.102
Jun17 FOMC	-2.412	60.2	4.006
Mar19 FOMC	0.748	50.8	1.473

Table 8: Measure of Uncertainty about BSP. Average uncertainty about the 10-year rate prevailing over the 10 days prior to the meeting, computed using the entire term-structure of swaption-implied volatility over a maximum horizon (MaxH) of either 5 or 10 years (5Y or 10Y)

MaxH	**LSAP1**	**Reinv**	**MEP1**	**MEP2**	**Jun2013**	**Sept2013**	**Jun2017**	**Mar2019**
5 Y	0.096	-0.203	0.018	-0.092	0.149	0.306	-0.136	-0.129
10 Y	0.095	-0.199	0.019	-0.093	0.146	0.299	-0.133	-0.128

Table 9: Impact of Investor Uncertainty about BSP on Treasury Yield Sensitivity (3-year bandwidth)

	Intercept	β_1	β_2	β_3	β_4	**Adj. R-Sq**	**N. of obs.**
Point Est	9.718	1.311	-2.344				861
T-Stat	(67.9)	(26.6)	(-26.9)			0.980	
Point Est	8.893	1.281	-2.373	2.998	-5.489		861
T-Stat	(59.3)	(27.7)	(-28.7)	(11.7)	(-10.8)	0.983	
Point Est	8.891	1.283	-2.377	3.061	-5.617		861
T-Stat	(59.3)	(27.8)	(-28.8)	(11.7)	(-10.9)	0.983	

Table 10: Impact of Investor Uncertainty about BSP on Treasury Yield Sensitivity (2-year bandwidth)

	Intercept	β_1	β_2	β_3	β_4	**Adj. R-Sq**	**N. of obs.**
Point Est	9.058	1.131	-2.105				624
T-Stat	(59.9)	(15.2)	(-15.9)			0.984	
Point Est	8.852	1.154	-2.189	1.061	-3.004		624
T-Stat	(53.8)	(15.6)	(-16.5)	(2.46)	(-3.62)	0.985	
Point Est	8.850	1.155	-2.192	1.093	-3.093		624
T-Stat	(53.8)	(15.6)	(-16.5)	(2.48)	(-3.66)	0.985	

Table 11: Measure of Market Functioning (MMF). Average 10-year on-the-run premium and yield curve fitting errors over the 10 days prior to the meeting relative to previous year

MMF	**LSAP1**	**Reinv**	**MEP1**	**MEP2**	**Jun2013**	**Sept2013**	**Jun2017**	**Mar2019**
OTR	0.523	-0.231	-0.321	-0.637	0.773	0.721	-0.241	-0.786
F.E.	0.179	-0.004	-0.136	0.412	-0.219	-0.155	-0.067	0.1903

Table 12: Impact of BSP Uncertainty and Market Functioning on Treasury Yield Sensitivity

	Intercept	β_1	β_2	β_3	β_4	β_5	β_6
OTR premium							
Point Est	8.79	1.21	-2.33	3.17	-5.74	0.02	0.08
T-Stat	(53.4)	(19.8)	(-24.04)	(11.9)	(-11.2)	(0.67)	(1.46)
Point Est	8.79	1.21	-2.33	3.24	-5.87	0.02	0.08
T-Stat	(53.4)	(19.9)	(-24.09)	(11.9)	(-11.3)	(0.69)	(1.45)
Fit. Errors							
Point Est	9.21	1.27	-2.37	1.69	-3.66	-1.95	2.72
T-Stat	(58.4)	(28.0)	(-29.3)	(4.96)	(-6.08)	(-5.71)	(5.37)
Point Est	9.20	1.27	-2.37	1.72	-3.76	-1.95	2.70
T-Stat	(58.4)	(28.0)	(-29.3)	(4.97)	(-6.11)	(-5.70)	(5.33)

Table 13: Total Supply Effects versus Overall QE Impact Estimated in the Literature

Treasury BSP	Supply Effects (bps)	Other Studies All Channels (bps)	Average Effect (bps) Across Studies
		47–Gagnon et al. (2011)	
LSAP 1		41–Krishnamurthy and Vissing-Jørgensen (2011)	
($300bn)	21	30–D'Amico and King (2013)	37
		35–D'Amico et al. (2012)	
		34–Bonis, Ihrig, and Wei (2017)	
		18–Krishnamurthy and Vissing-Jørgensen (2011)	
LSAP 2		55–D'Amico et al. (2012)	
($600bn)	13	21–Meaning and Zhu (2011)	24
		15–Swanson (2011)	
		12–Bonis, Ihrig, and Wei (2017)	
MEP		22–Hamilton and Wu (2012)	
($667bn)	14	17–Meaning and Zhu (2012)	22
		28–Bonis, Ihrig, and Wei (2017)	
LSAP 3	17	60–Engen, Laubach, and Reifschneider (2015)	45.5
($790bn)		31–Bonis, Ihrig, and Wei (2017)	
Total	65		129.5

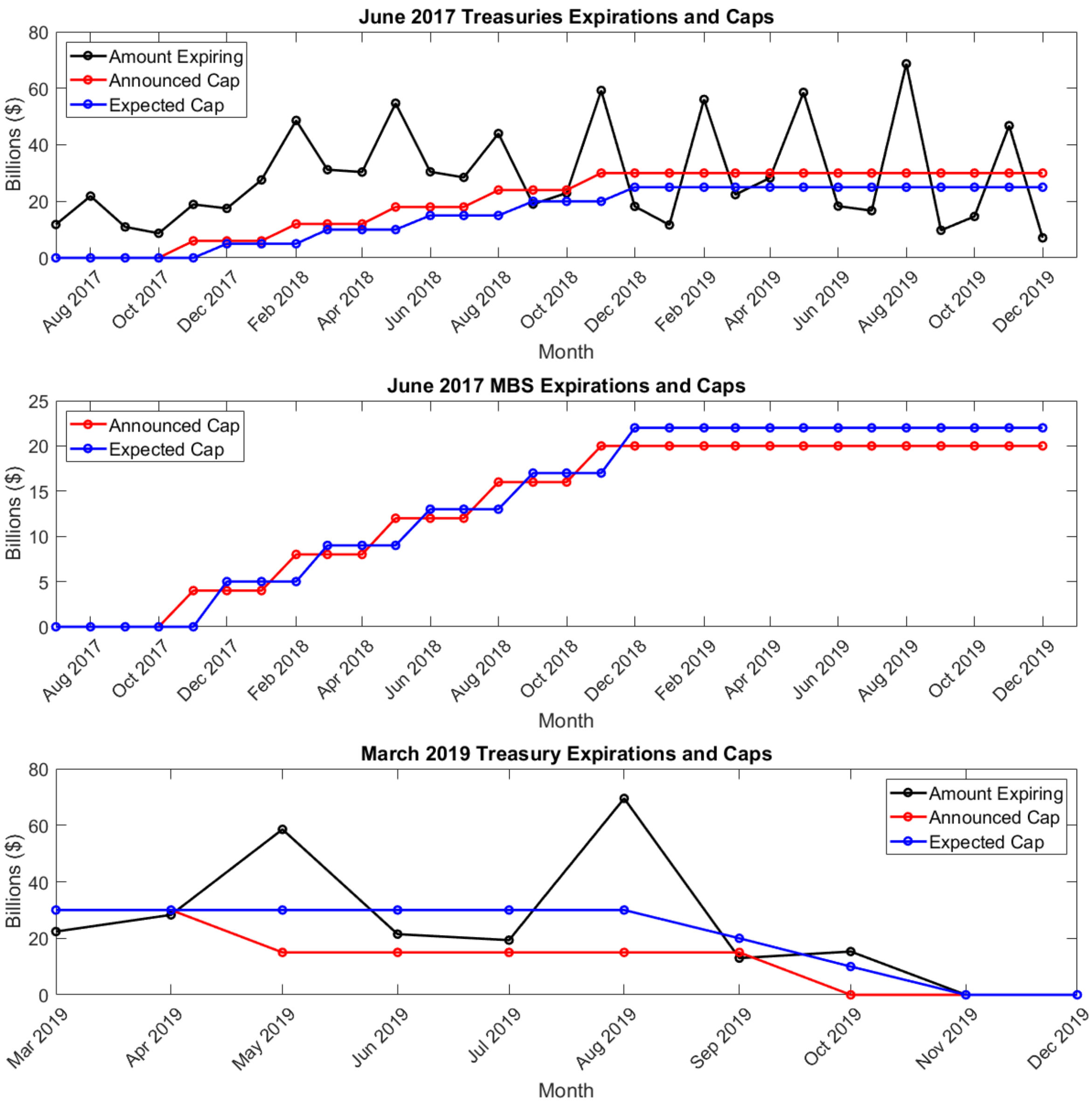

Figure 1: Treasury and MBS Expirations and Caps. The top panel shows the June 2017 expected and announced cap schedules for Treasury securities, together with the amount of Treasury securities expiring each month. The middle panel shows the June 2017 expected and announced cap schedule for agency MBS. The bottom panel shows March 2019 expected and announced cap schedules for Treasury securities together with the amount of Treasury securities expiring each month.

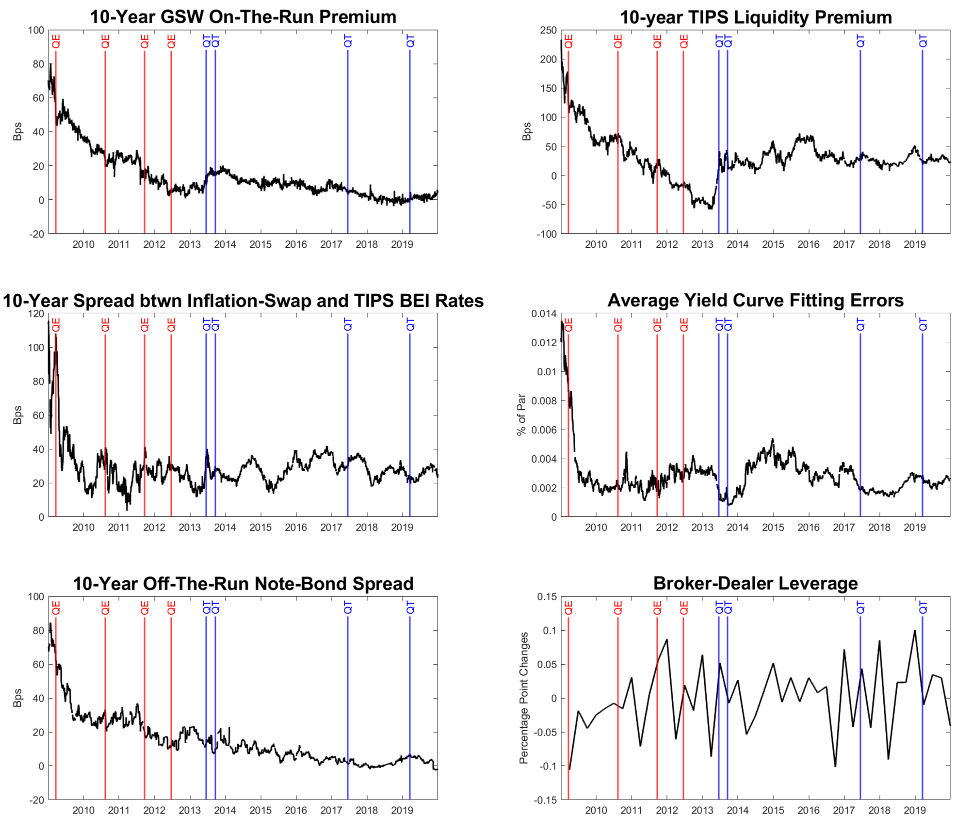

Figure 2: Time series of widely-used measures of Treasury market liquidity, limits to arbitrage, and financial constraints. Sample period: 2009-2019. The vertical red lines denote the time of QE announcements and the blue lines QT announcements. The top left panel plots the 10-year on-the-run premium, measured as in Gürkaynak, Sack, and Wright, (2007). The next four panels display: the 10-year TIPS-Treasury bond puzzle of Fleckenstein, Longstaff and Lustig (2014), proxied with the wedge between the 10-year inflation swap rate and the 10-year TIPS break-even rate; the off-the-run note-bond spread of Musto, Nini and Schwarz (2017), derived as the difference between the average yields of off-the-run Treasury notes and bonds with maturity as close as possible to 10 years; the 10-year TIPS liquidity premium implied by the term-structure model of D'Amico, Kim and Wei (2018); and the average absolute nominal yield curve fitting errors, which can be interpreted as a measure of limits to arbitrage (see for example Hu, Pan, and Wang, 2013). Finally, the bottom right panel depicts changes in broker-dealer leverage, obtained using the Fed Flow of Funds data, which captures the balance sheet capacity of financial intermediaries (Adrian, Etula, and Muir, 2014).

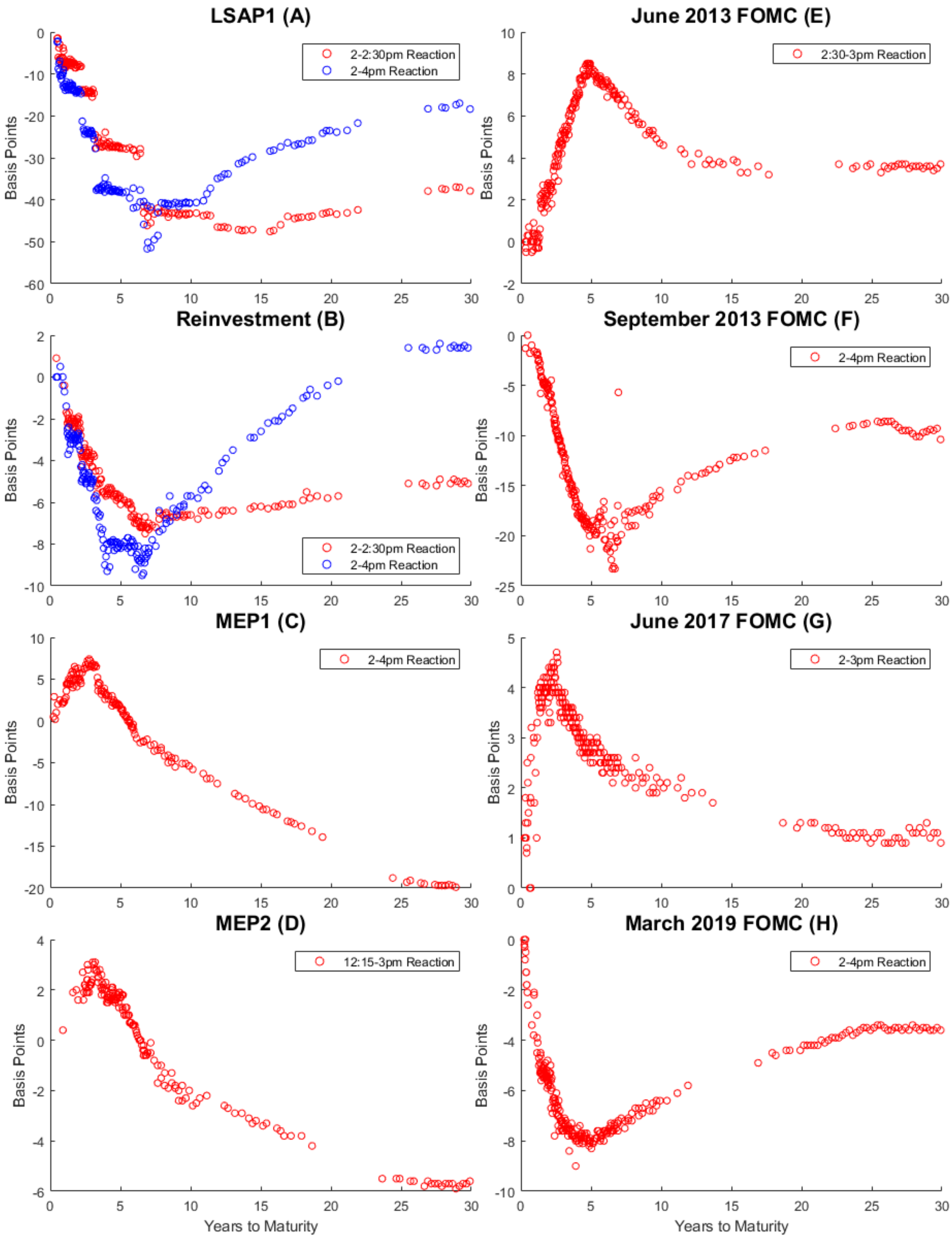

Figure 3: Yield curve reactions to QE announcements (left column) and QT announcements (right column). Each dot represents the change in a Treasury security's yield in the small time window around the FOMC announcement.

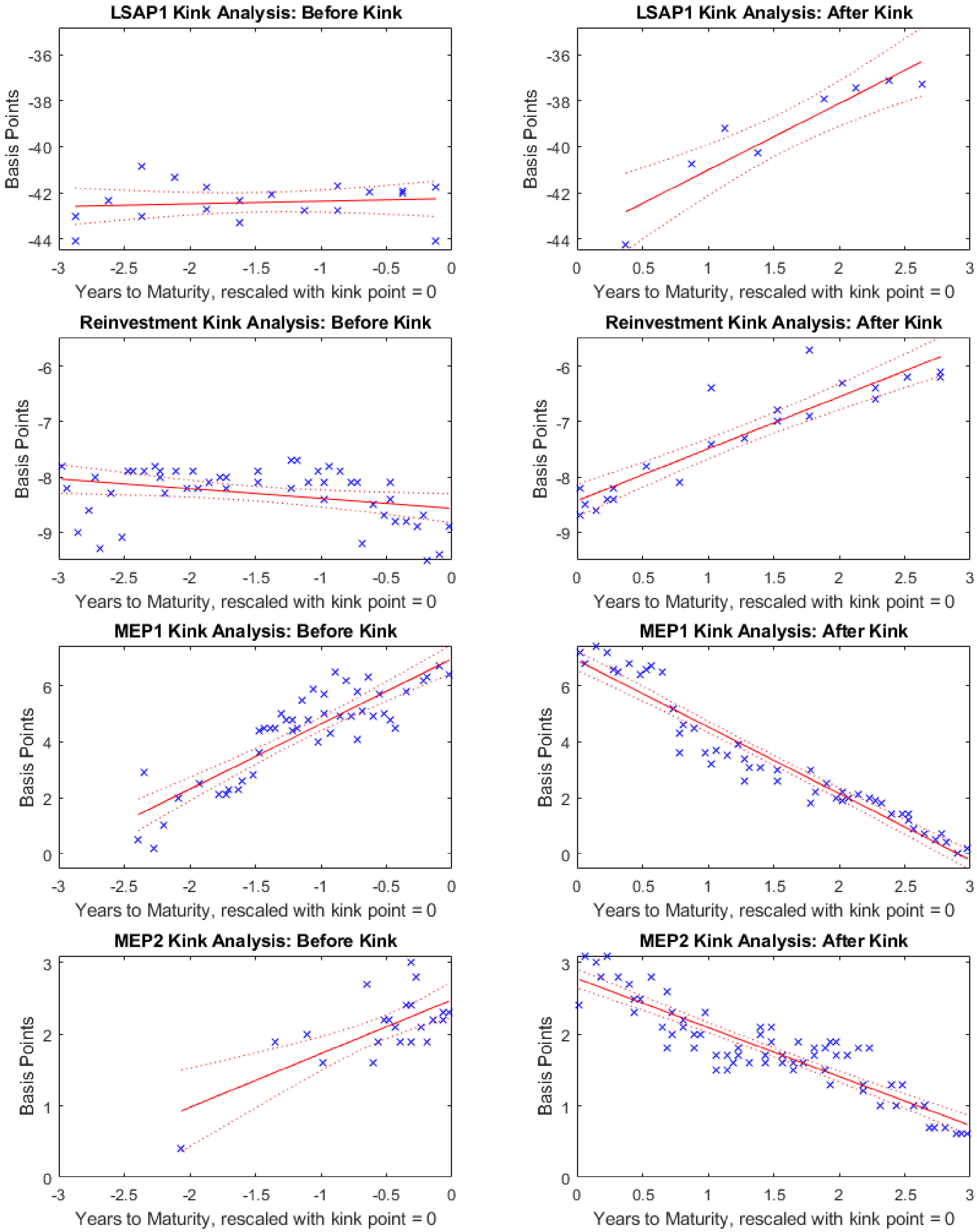

Figure 4: QE Kink Analysis. Each row shows the yield reactions and the fitted line to the left and right of the kink for each event.

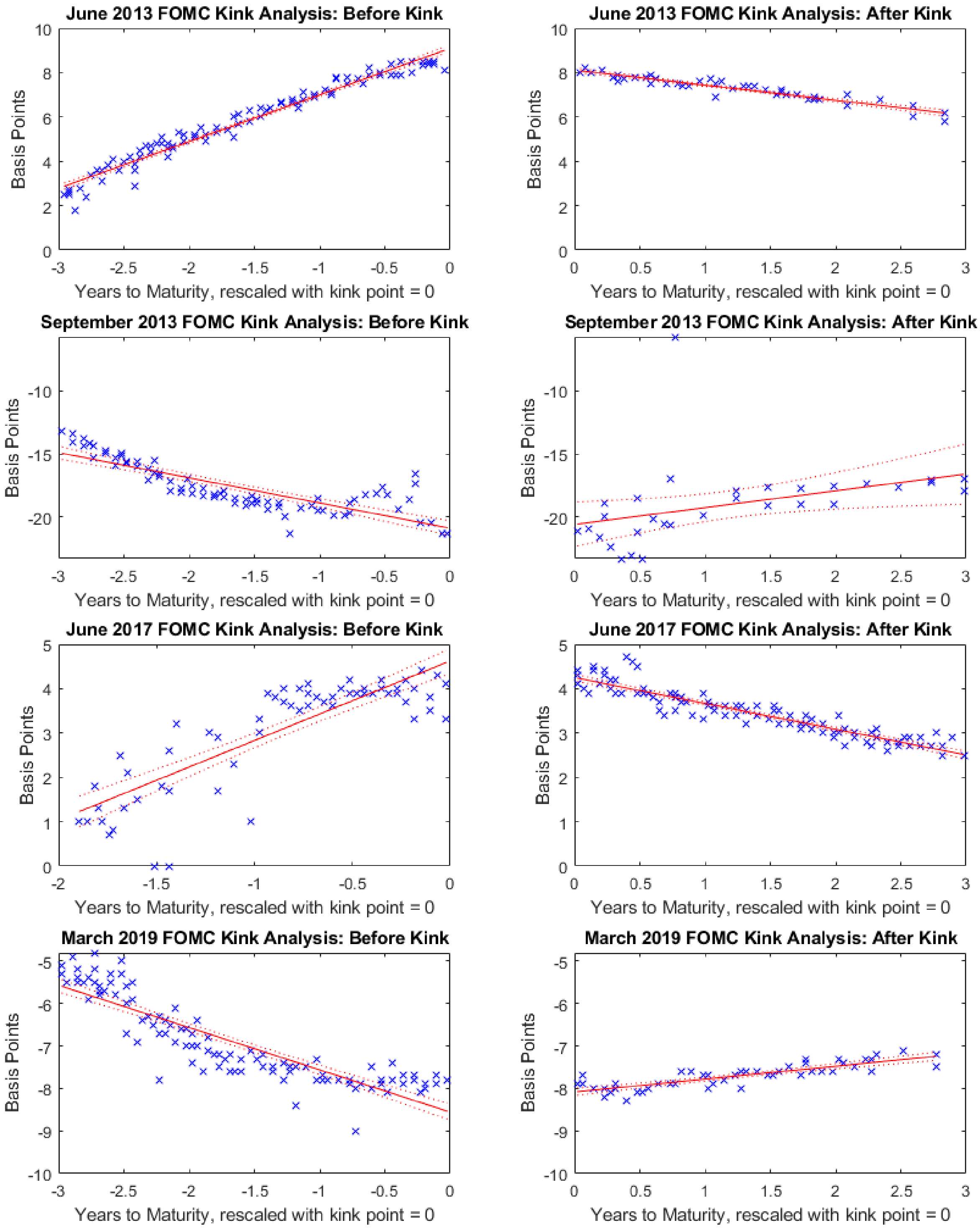

Figure 5: QT Kink Analysis. Each row shows the yield reactions and the fitted line to the left and right of the kink for each event.

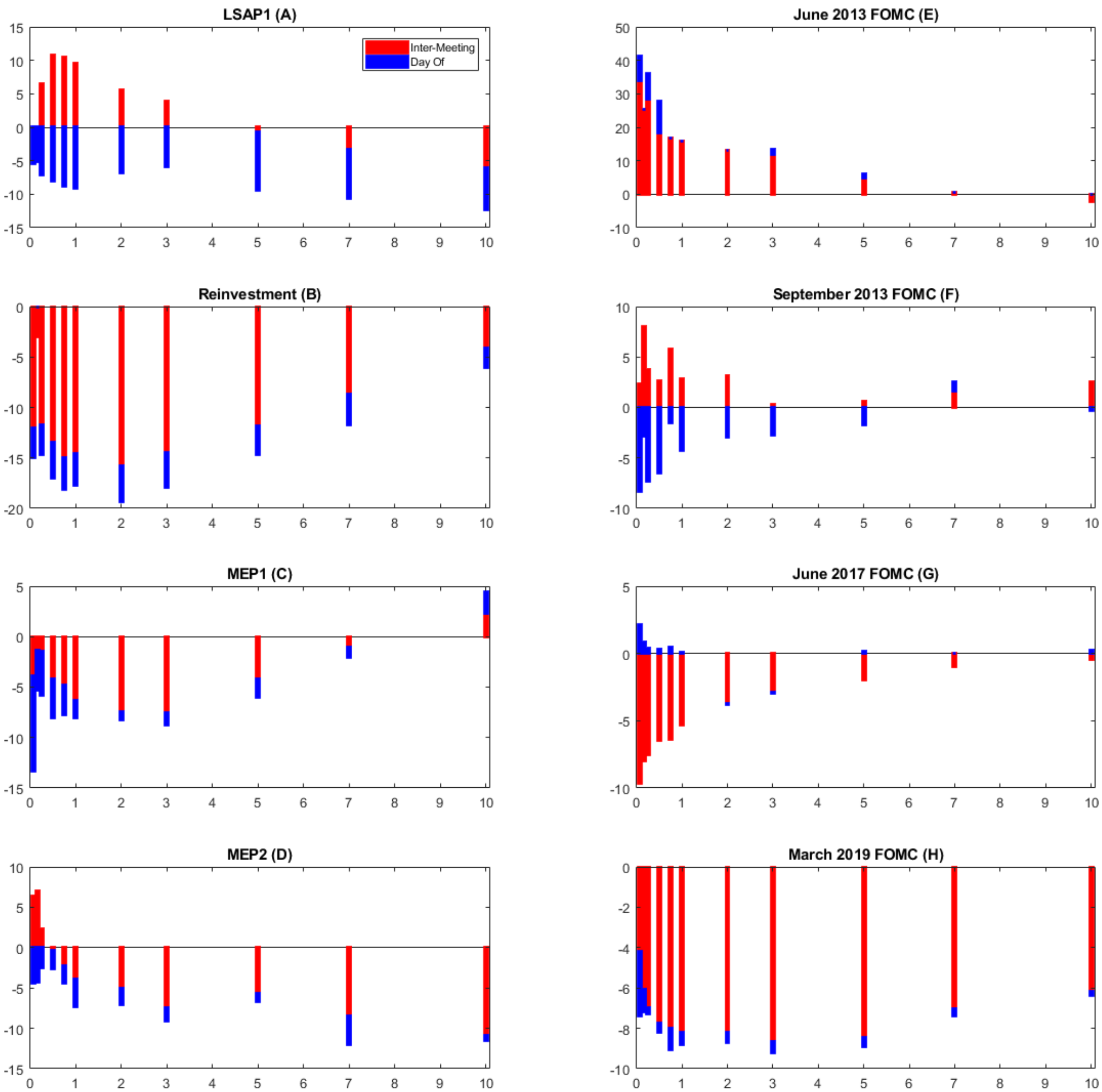

Figure 6: Evolution of the Term Structure of the 10-Year Rate Uncertainty Ahead and On-the-Day of each BSP announcement. Each panel shows the change in the swaption-implied volatility of the 10-year rate across different forecasting horizons in the six weeks preceding the day of the announcement (red bars) and on the day of the FOMC announcement (blue bars).

Appendix

I. Computation of BSP Surprises

In this Appendix, we describe in more detail the surprise computation for each event. In particular, we provide an upper and lower bound for the size of the BSP surprise and, since the surprise is the denominator of the yield sensitivity estimate, an upper/lower bound for the surprise corresponds to a lower/upper bound for the yield sensitivity estimate. In the case of the Treasury-only surprises, our bounds rely on two alternative assumptions about the degree of market segmentation in the Treasury market. Each has its own limitations, but together they provide a reasonable range for the true value of the surprise.

The upper bound is given by the total size of the Treasury surprise at announcement. This method takes no stance on the degree of market segmentation. The lower bound is given by the size of the surprise local to the kink, obtained from the relative supply changes only in the maturity buckets adjacent to the kink. This implies a high degree of market segmentation, as it assumes that only the supply surprises local to the kink trigger the yield reaction around the kink (i.e., a tight relation between the maturities of the bonds' quantities and prices). Table 2 and Table 5 in the main text contain the surprise estimates used to compute the yield sensitivities.

The bounds in Table 2 and Table 5 only take into account the supply surprise in the Treasury market. Since at times agency securities (i.e., agency MBS and agency debt) were included in the Fed's BSP, we also perform a robustness check using a version of the total surprise that includes unexpected changes in the supply of both Treasury and agency securities. This implies that large investors view Treasuries and agency securities as perfect substitutes. Table 3 and Table 6 in the main text contain these alternative surprise estimates.

LSAP1

The March 2009 SPD asked about the likelihood that the Fed would announce a purchase program related to Treasuries, agency debt, and agency MBS. However, it did not ask about the expected size of those programs, so we make the conservative assumption that the dealers perfectly forecasted the announced size of the program.

Hence, for each asset class, the surprise is obtained by simply multiplying the announced size by the median probability of the Fed *not* announcing purchases of that asset class. Table A1 reports those median probabilities, actual sizes, and surprises (in billions), which measure the upper bound of LSAP1 surprises.

Table A1: LSAP1 Surprises

Asset Class	$Prob$(NO announcement)	Size ($bn)	Surprise ($bn)
Treasuries	0.475	300	142.5
Agency Debt	0.30	100	30
Agency MBS	0.25	750	187.5

To compute the surprise local to the kink we need some additional assumptions. First, we assume that before the FOMC announcement, investors expected Treasury purchases to be distributed over the entire maturity range and to be proportional to the amount outstanding in each maturity sector. This seems quite reasonable considering the absence of any guidance prior to the FOMC meeting. Second, since the Desk indicated that it would concentrate Treasury purchases in the 2- to 10-year sector, we assume that market participants interpreted the Desk's guideline as about 80% of purchases being allocated to that sector. Reportedly, this interpretation was quite common among market participants at that time. Third, we use the Fed's auction maturity sectors (reported in the first row of Table A2) released by the Desk before the start of the actual purchases but after the FOMC meeting. These three assumptions imply that, after the Desk announcement on the day of the FOMC, investors recalibrated their expectations as follows: 80% of purchases would occur in the 2- to 10-year maturity range, distributed within that range proportionally to the amount outstanding in each auction maturity sector. Similarly, the remaining 20% of purchases would be allocated to the 10- to 30-year range proportionally to the amount outstanding in the two auction maturity sectors included in that range. Table A2 contains the pre- and post-announcement expected maturity distribution of purchases. The LSAP1 local supply surprise is computed as $2 * [(22.1 * 300 - 17.8 * 157.5) - (4.5 * 300 - 8 * 157.5)]/100 = \$74.73bn$, using only the supply changes in the two maturity sectors adjacent to the kink located at the 10-year maturity.

Table A2: Maturity Distribution of LSAP1 Purchases

Fed Auction Sectors	1.5-2.5Y	2.5-4Y	4-5.5Y	5.5-7Y	7-10Y	10-17Y	17-30Y
Expected	17.4%	20%	18%	8.5%	17.8%	8%	10.3%
Announced	9.7%	25%	22.3%	10.6%	22.1%	4.5%	5.8%

Reinvestment

The August 2010 SPD asked about the probability of the Fed announcing the reinvestment of proceeds from agency MBS either 1) into Treasuries and Agency MBS, and 2) into only Treasuries. The median probability reported for option 1 was 32.5%, while for option 2 it was 10%. Publicly available information indicated that about $206.25bn of agency MBS proceeds were set to roll off the Fed's balance sheet over the next six months (the same horizon as LSAP1). Similar to LSAP1, we obtain the surprise for each asset class by multiplying the size ($206.25bn) by the median probability of the Fed *not* announcing a program related to that asset class. Therefore, the total surprise is $185.6bn for Treasuries only and $139.2bn for Treasury and agency securities together, as seen in Table A3.

Table A3 Reinvestment Surprises

Asset Class	$Prob$(NO announcement)	Actual Size ($bn)	Surprise ($bn)
Treasuries	0.90	206.25	185.6
Treasuries or Agencies	0.675	206.25	139.2

For the computation of the local surprise, we assume the expected maturity distribution of the reinvestment purchases to be the same as the actual maturity distribution of LSAP1 purchases (second row of Table A4), as before the August 2010 FOMC meeting investors did not receive any new information that could lead them to expect a different maturity distribution. Similar to LSAP1, on the day of the FOMC, the Desk indicated that it would "concentrate purchases in the 2- to 10-year sector," but also announced that it would "refrain from purchasing securities for which there is heightened demand or of which the SOMA already holds large concentrations." To account for the change in market expectations caused by this statement, we use security-level data to identify any Treasury securities for which the Fed already held 25% or more of the total amount outstanding. Since these securities had a lower likelihood of being purchased, we adjust each sector's expected share of purchases in inverse proportion to the amount already owned by the Fed. This implies that, for example, since the Fed already held a relatively higher concentration of securities in the 10- to 17-year maturity range, the expected share of purchases allocated to that sector is reduced by a larger factor, as seen in Table A4. The reinvestment local supply around the kink is computed as $2 * [(26.12 * 206.25 - 24.48 * 20.6) - (7.3 * 206.25 - 24.13 * 20.6)]/100 = \$77.49bn$.

Table A4 Maturity Distribution of Reinvestment Purchases

Fed Auction Sectors	1.5-2.5Y	2.5-4Y	4-5.5Y	5.5-7Y	7-10Y	10-17Y	17-30Y
Expected	6.29%	25.87%	24.48%	24.13%	4.55%	10.49%	4.20%
Announced	2.42%	58.48%	26.12%	7.30%	2.61%	1.23%	1.83%

MEP1

The September 2011 SPD (question 7b) reports a 78% median probability of the Fed increasing the duration of its portfolio through an "Operation Twist" at some point over the next two years. Further, conditional on the MEP being announced, dealers expected the size of the program to be \$325bn. As a result, the expected total purchases were $0.78 * \$325\text{bn} = \$253.5\text{bn}$. In contrast, the Fed announced a \$400bn MEP, creating a total Treasury surprise of \$146.5bn. There was no MBS component in MEP1.

Question 8c in the same SPD shows that dealers correctly anticipated that the MEP would be implemented over a six-month period, and by selling Treasuries with remaining maturity less than 3 years to buy Treasuries with remaining maturity between 7 and 30 years. This information is very useful for the computation of the local surprise. We obtain the expected maturity distribution of purchases for MEP1 by proportionally reallocating to the 6- to 30-year sector the share of purchases previously allocated to the 1.5- to 6-year sector. Then we use the actual maturity distribution of purchases released by the Desk contemporaneously to the FOMC announcement, listed in Table A5.[30] As a result, the MEP1 local surprise is $[(100 - 32.99) * 400 - (100 - 42.23) * 253.5]/100 = \$126.67\text{bn}$.

[30] The announced weights do not sum to 100 because 3% of purchases were allocated to TIPS, which we do not include here. Thus, to compute the local surprise, we rescale the announced weights to sum to 100, the results of which are displayed in the third row of Table A.5.

Table A5 Maturity Distribution of MEP1 Purchases

Fed Auction Sectors	0-3Y	3-6Y	6-8Y	8-10Y	10-20Y	20-30Y
Expected	-100%	0%	44.23%	44.23%	3.85%	7.69%
Announced	-100%	0%	32.99%	32.99%	4.12%	29.90%

MEP2

In the June 2012 SPD, dealers reported a 45% median probability of extending the MEP at some point over the next 2 years (question 7b). At the time, the Fed held \$205bn in short-term Treasuries whose maturity was expected to fall below 3 years over the next six months. It is conceivable that this publicly available information was used by market participants to form their expectations ahead of the MEP2 announcement. Thus, the expected amount of purchases under the MEP2 is $0.45 * \$205bn = \$92.25$bn. However, the Desk's statement indicated that it would stop rolling over maturing Treasuries at auction, allowing the Fed to further increase its long-term purchases from \$205bn to \$267bn. Thus, the \$267bn program created a \$174.75bn total Treasury surprise (\$267bn minus \$92.25bn). There was no MBS component in MEP2. To compute the local surprise, we simply apportion the total surprise according to the Desk's purchase/sale weights, which did not change from the MEP1 weights listed in Table A5. We compute the local surprise as $[(100 - 32.99) * 174.75]/100 = \$117.1$bn.

June 2013 FOMC: Post-taper tantrum

In the case of this event, the Desk conducted both a pre- and post-FOMC survey on June 10 and June 24, respectively. Question 6a from these two surveys shows the expected start of tapering moving forward from December 2013 to September 2013 (the first reduction in the pace of purchases is highlighted in Table A6). Further, as shown in the last row of each panel in Table A6, the change in the expected monthly pace of purchases from the June 10 to the June 24 SPD implies a \$27.5bn hawkish Treasury surprise and a \$38bn hawkish MBS surprise, respectively–a smaller-than-expected amount of purchases in each asset class. To compute the local surprise, we allocate the Treasury total surprise across the Fed auction maturity sectors according to the Desk's purchase weights already in place during LSAP3, shown in Table A.7.[31] Then, we take the difference between the surprise in the 7-10-year sector and each of

[31] https://www.newyorkfed.org/markets/longertermtreas_faq_12122012.html

the surprises in the adjacent maturity sectors, and sum those two differences: $(8.22 - 0.57) + (8.22 - 4.54) = \$11.34$bn.

In computing those surprises, we do not use information from question 6b, which asks about expected changes in the SOMA portfolio beyond the one-year horizon, as we posit that elevated uncertainty about QE tapering and the policy normalization process led market participants to heavily discount longer-term changes to the balance sheet policy.

Table A6 Expectations about the monthly pace of purchases ($bn) in SPD from June 10 to June 24, 2013

Treasuries	Jun13	Jul13	Aug13	Sep13	Oct13	Nob13	Dec13	Jan14	Feb14	Mar14	Apr14	May14	Jun14
June 10	45	45	45	45	45	45	30	25	20	15	10	5	0
June 24	45	45	45	40	35	32.5	30	25	20	15	10	5	0
Change				**-5**	**-10**	**-12.5**							

MBS	Jun13	Jul13	Aug13	Sep13	Oct13	Nob13	Dec13	Jan14	Feb14	Mar14	Apr14	May14	Jun14
June 10	40	40	40	40	40	40	30	25	20	15	0	0	0
June 24	40	40	40	35	33	29	25	20	15	10	5	0	0
Change				**-5**	**-7**	**-11**	**-5**	**-5**	**-5**	**-5**	**+5**		

Table A7 June 2013 Local Supply Surprises

Fed Auction Sectors	4-4.75Y	4.75-5.75Y	5.75-7Y	7-10Y	10-20Y	20-30Y
Desk's Weights	11%	12%	16%	29%	2%	27%
Surprise ($bn)	3.12	3.40	4.54	8.22	0.57	7.65

September 2013 FOMC: Tapering feint

Also for this event, the Desk conducted both a pre- and post-FOMC survey on September 9 and September 23, respectively. Question 1a from the September 9 SPD indicated that dealers expected the FOMC to announce QE3 tapering at the upcoming meeting. In contrast, the FOMC announced it was not yet ready to reduce its monthly pace of purchases. This announcement had the exact opposite effect of the June 2013 FOMC announcement: question 8 from both the September 9 and September 23 SPD shows the expected start of tapering moving backward from September 2013 to December 2013 (the first reduction in the pace of purchases is highlighted in Table A8). Further, as shown in the last row of each panel in Table A8, the change in the expected monthly pace of purchases from the September-9 to the September-23 SPD implies a

$95bn dovish Treasury surprise and a $119.5bn dovish MBS surprise, respectively–a larger-than-expected amount of purchases in each asset class. To compute the local surprise, again we allocate the Treasury total surprise across the Fed auction maturity sectors according to the Desk's LSAP3 weights, shown in Table A9. Then, just as for June 2013, we take the difference between the surprise in the 7-10-year sector and each of the surprises in the adjacent maturity ranges, and sum those two differences: $(28.40 - 1.96) + (28.40 - 15.67) = \$39.18$bn.

Consistent with the previous event, in computing those surprises, we do not include data from question 8c about changes in SOMA from 2015 through 2018, as we posit again that elevated uncertainty about QE tapering and the policy normalization process led market participants to heavily discount longer-term balance sheet policy decisions.

Table A8 Expectations about the monthly pace of purchases ($bn) in SPD from Sep 09 to Sep 23, 2013

Treasuries	Sep13	Oct13	Nob13	Dec13	Jan14	Feb14	Mar14	Apr14	May14	Jun14	Jul-Dec14
Sep 9	40	35	32.5	30	25	20	15	15	7.5	0	0
Sep 23	45	45	45	40	35	30	25	20	15	10	5
Change	**+5**	**+10**	**+12.5**	**+10**	**+10**	**+10**	**+10**	**+5**	**+7.5**	**+10**	**+5**
MBS	Sep13	Oct13	Nob13	Dec13	Jan14	Feb14	Mar14	Apr14	May14	Jun14	Jul-Dec14
Sep 9	35	35	30	25	25	20	15	10	5	0	-5
Sep 23	40	40	40	35	35	30	25	20	16.5	13	20
Change	**+5**	**+5**	**+10**	**+10**	**+10**	**+10**	**+10**	**+10**	**+11.5**	**+13**	**+25**

Table A9 September 2013 Local Supply Surprises

Fed Auction Sectors	4-4.75Y	4.75-5.75Y	5.75-7Y	7-10Y	10-20Y	20-30Y
Desk's Weights	11%	12%	16%	29%	2%	27%
Surprise ($bn)	10.77	11.75	15.67	28.40	1.96	26.44

June 2017 FOMC: Redemption caps announced

At the June 2017 meeting, the FOMC released the policy normalization addendum detailing the intended redemption cap schedule to gradually reduce the size of its portfolio. We compare the FOMC's announced caps to the expected caps in questions 4c and 4e of the SPD and obtain a month-by-month surprise from the start of the cap implementation through December 2019 (see Figure 4 in main text). We believe it is sufficient to stop in December 2019 because dealers' expectation of SOMA's size at year-end 2019 (question 5 in the SPD) closely matches their expectation of SOMA's size at year-end 2025 (question 7), suggesting dealers anticipated the Fed's portfolio run-off would be completed by 2019. Comparing the expected and announced caps gives a \$125bn hawkish Treasury surprise and a \$18bn dovish MBS surprise because the MBS/Treasury caps were lower/higher than expected, respectively.

However, putting together publically available information from the Fed's System Open Market Account (SOMA), the Treasury's Monthly Statement of the Public Debt (MSPD), and the Securities Industry and Financial Markets Association (SIFMA) (as described in the Addendum at the end of this Appendix) reveals that, in certain months, the amount of Treasuries set to expire would be below the expected redemption caps. This implies that there is no surprise in those months, making the actual Treasury hawkish surprise estimate (\$78.2bn) smaller than the one obtained by only comparing the actual and expected caps (\$125bn). Figure 4 illustrates how the expected redemptions in those months were not affected by the FOMC's higher-than-expected redemption caps. Since we assume that primary dealers use all the available information, we estimate a \$78.2bn hawkish Treasury surprise, and a \$18bn dovish surprise for MBS. The maturity distribution of the \$78.2bn hawkish Treasury surprise is displayed in Table A10.

Since the Fed rolled over maturing securities at auction, for this event the maturity ranges are determined by the maturities of Treasury issuance; thus, to compute the local surprise consistently across events, we need to combine the 2- and 3-year maturities into one range, the 5- and 7-year maturities into one range, and the 10- and 30-year maturities into one range. As a result, the local surprise is computed as follows: $[(16.47+13.59)-(12.48+14.54)]+[(16.47+13.59)-(12.99+8.14)] = \$11.97$bn.

Table A10 June 2017 Local Supply Surprises

Maturity	2Y	3Y	5Y	7Y	10Y	30Y
Surprise (\$bn)	12.48	14.54	16.47	13.59	12.99	8.14

March 2019 FOMC: QT tapering

At the March 2019 meeting, the FOMC announced a reduction in its redemption caps, implying larger asset purchases than before. Before the meeting, primary dealers were already expecting an accommodative adjustment to QT (question 5a), but they underestimated the dovishness of the Fed's policy move. In particular, they expected the initial reduction in caps to occur in September 2019 and the end of QT to occur sometime in 2019:Q4. In contrast, the FOMC announced that the cap reduction would start in May 2019 and that QT would end in September 2019, implying a $75bn dovish Treasury surprise (see Table A11). Adding SOMA, MSPD, and SIFMA data, however, produces a slightly lower $50.8bn dovish surprise because Treasury redemptions were sometimes below the caps (see Figure 5 in the main text).

Table A11 Expectations about the monthly caps ($bn) before and after March 2019 FOMC

	Mar19	Apr19	May19	Jun19	Jul19	Aug19	Sep19	Oct19	Nov19	Dec19
Pre	30	30	30	30	30	30	20	10	0	0
Pre	30	30	15	15	15	15	15	0	0	0
Change			**-15**	**-15**	**-15**	**-15**	**-5**	**-10**		

Since, similar to the June 2017 event, this event's maturity ranges are determined by the maturities of Treasury issuance, to compute the local surprise consistently across events we need to combine the 2- and 3-year maturities into one range, the 5- and 7-year maturities into one range, and the 10- and 30-year maturities into one range. As a result, the local surprise is computed as follows: $[(10.41+8.405)-(9.948+10.27)]+[(10.41+8.405)-(7.071+4.706)]=\$5.63$bn, as seen in Table A12.

Table A12 March 2019 Local Supply Surprises

Maturity	2Y	3Y	5Y	7Y	10Y	30Y
Surprise ($bn)	-9.948	-10.27	-10.41	-8.405	-7.071	-4.706

Additional Addendum on SOMA Surprise Methodology for June 2017 and March 2019

The Treasury sells its notes and bonds in six different maturities: 2, 3, 5, 7, 10, and 30 years. It sorts these six maturities into two different monthly auctions: a mid-month auction selling 3, 10, and 30-year Treasuries, and an end-month auction selling 2, 5, and 7-year Treasuries. After ending QE3 in October 2014, the Fed held the size of its balance sheet constant by reinvesting principal payments from its maturing Treasuries into new Treasury notes and bonds at Treasury auctions. Then, in late 2017, the Fed began to reduce the balance sheet by redeeming some principal payments from its maturing Treasury securities each month—but only up to a monthly cap. To quantify the policy surprise, we use unexpected changes in this monthly cap schedule released at the FOMC announcements in June 2017 and March 2019.

The Fed reinvests principal payments from its maturing Treasuries according to an entirely deterministic mechanism which has been in place since December 2015.[32] To obtain the total dollar amount to be reinvested at Treasury auctions in a given month, the Fed adds up the dollar amount of Treasuries set to mature during that month and then subtracts the FOMC-directed cap on Treasury redemptions. Changes in the FOMC redemption cap therefore directly affect the total reinvestment amount. Two key numbers then determine the allocation of these reinvestments across the six Treasury maturities. The first is the proportion of Treasuries maturing at the mid-month auction (3, 10, and 30-year) versus those maturing at the end-month auction (2, 5, and 7-year) within each month. The Fed splits its total pool of reinvestment dollars between these two auctions according to this first proportion. Then, at each auction, the Fed allocates reinvestment purchases across an auction's three maturities according to the maturity-level proportion of Treasury issuance within each auction.

Take a hypothetical example from the Federal Reserve Bank of New York where $21bn of SOMA's Treasury portfolio matures and the FOMC cap is $6bn for that month. The Fed will therefore reinvest $15bn during the month. Assume mid-month Treasuries (3, 10, and 30-year) constitute $7bn of the maturing $21bn, and end-month Treasuries (2, 5, and 7-year) comprise the remaining $14bn. The Fed thus allocates 33% of the $15bn (so $5bn) to reinvesting at the mid-month auction and 67% of the $15bn ($10bn) to reinvesting at the end-month auction. Next, at each auction, the Fed's allocation across an auction's three maturities depends on the maturity-level proportions of Treasury issuance within each auction. Assume that for the end-month auction, the Treasury announces offerings of $25bn for 2Y Treasuries, $15bn for 5Y, and $10bn for 7Y. Thus, the Fed reinvests 50% of the $10bn ($5bn) in 2Y Treasuries, 30% of the $5bn ($3bn) in 5Y Treasuries, and 20% of the $5bn ($2bn) in 7Y Treasuries.

The pre-determined nature of the Fed's reinvestment mechanism allows market actors to anticipate SOMA monthly purchases ahead of time and also enables us to measure market expectations of reinvestment purchases. The Fed releases weekly CUSIP-level snapshots of its SOMA portfolio holdings, while the Treasury Department

[32] https://www.newyorkfed.org/markets/treasury-rollover-faq-12-16-2015.html

releases a Monthly Statement of the Public Debt (MSPD) containing CUSIP-level data on issuance date and amount outstanding. Combining the SOMA and MSPD data reveals the monthly reinvestment amount in dollars, in addition to the auction-level proportions of Treasuries maturing each month. The Securities Industry and Financial Markets Association (SIFMA) releases monthly data on Treasury issuance volumes broken out by maturity. Adding in the SIFMA data produces an approximation of market expectations for issuance volumes and therefore the maturity-level proportions of Treasuries within each auction. For each of the six maturities, we recreate the amounts investors expected the Fed to purchase each month as part of its reinvestment policy. Then, when the FOMC announces changes in redemption caps, we look at the maturity-specific expectations before and after the announcement, quantifying the six maturity-specific surprises in dollars by taking the difference at each maturity. As a result, we quantify how much the market expectations of Fed purchases/sales changed because of a given balance sheet policy adjustment.

II. Algorithm to identify kink's precise location

For each event, the kink's precise location is determined using a gradient-based algorithm that refines the search of the kink starting from an initial guess. This guess consists of the maturity where we expect the kink to appear based on the operational details of the Fed's BSP implementation; for instance, the 10-year maturity in the case of LSAP1, as explained in Section 4.1.[33] Then, the gradient of the yield curve reaction is analyzed along the maturity spectrum that extends a few years to the left and right of the initial guess, that is, we repeatedly estimate the local gradient within rolling windows and the point where the gradient changes sign is selected as the kink. We favor this procedure for the following reason. In theory, the location of the kink could be strictly dictated by the operational details, which being determined before the announcement are independent of the asset price reaction following the announcement. In practice, it would be unreasonable to expect a perfect mapping between kink's locations and operational details, because such mapping would require an extreme form of market segmentation across adjacent maturity sectors. This could probably be a realistic assumption for the first QE, which was announced at the height

[33]The kink's location is usually close to the edge of the maturity range receiving more or less purchases than expected because of purchase/reinvestment practices. The mapping between the maturity bound and the kink location should be less precise the less segmented the market is.

of the global financial crisis. Indeed, in the case of LSAP1, we do have a nearly perfect mapping between the operational detail and the kink's location. But, for subsequent events, and especially QT announcements, this does not have to be the case. Therefore, we search for the change in gradient in the proximity of a reasonable guess. The implicit assumption is that investors reveal their preferences about securities' degree of substitutability through their trading and we take it as given.