Three essays on causal analysis of banking regulation and monetary policy

by

Haowei Tang

A thesis submitted to the
Faculty of Graduate and Postdoctoral Affairs
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Economics

Department of Economics
Carleton University
Ottawa, Ontario
September, 2022

The undersigned hereby recommends to the
Faculty of Graduate and Postdoctoral Affairs
acceptance of the thesis

Three essays on causal analysis of banking regulation and monetary policy

submitted by **Haowei Tang**

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Economics

Department of Economics
Carleton University
September, 2022

Abstract

As the second largest financial crisis after the "Great Depression", the 2007/8 financial crisis posed great challenges to policy makers, ranging from the liquidity problem faced by the banking sector to the key monetary policy rate hitting the effective zero lower bound. To respond to such challenges, new policies are adopted. In my dissertation, I conduct causal analysis to evaluate the effectiveness of some of the newly proposed regulations following the financial crisis. Particularly, I examine the effects of liquidity regulations on banks, and of conventional and unconventional monetary policies.

The first chapter is a joint work with Professor Lynda Khalaf. In this paper, we examine the impact of the Liquidity Coverage Ratio (LCR) on bank lending in the U.S, using a Difference-in-Difference framework with a variety of identification methods. We are particularly interested in treatment effect dynamics. In this context, the dynamic two-way fixed effect (TWFE) model is commonly used which consists in including dynamic indicators for time relative to treatment which allows for treatment adoption to vary across time. The coefficients on these indicators aim to track the evolution of treatment effects. However, recent econometric works suggest that TWFE estimators do not recover the hypothesized causal effect; severe bias cannot be ruled out even when treatment effect dynamics are homogeneous and in the absence of anticipatory behaviour. The underlying reasons for such failures can be summarized as follows: (i) parallel trend assumptions (PTAs) - some of which may often be implausible - with varying strengths lead to different interpretable causal effects; (ii) the interpretable causal effect is in fact an (unknown) linear combination (weighted average) of the indicator coefficients, (iii) the role of conditioning covariates and treatment anticipation. Available evidence on the LCR ratio is scarce and is restricted to standard event studies. In this paper, we compare standard dynamic TWFE estimates to recently proposed alternative specifications that allow us to introduce various group-time aggregation schemes. Results underscore the importance of defining clear interpretable parameters, allowing for conditioning on covariates. In general, we find no effects of the LCR on bank lending, and the assumptions embedded in the TWFE models translate into meaningful differences in empirical results.

The second chapter is also a joint work with Professor Lynda Khalaf. In this paper, we study the dynamic causal effects of a monetary policy shock on the US economy within the Local Projection - Instrumental Variable [LP-IV] framework. Our reassessment is motivated by the emerging concerns in the literature about popular IVs that are based on high-frequency identification. We approach related difficulties as follows. First we provide weak-instruments robust inference on the traditional LP-IV coefficient which we denote as the direct causal effect [DCE]. Second, we define, estimate and test an alternative response parameter, denoted as the total causal effect [TCE], that accounts for the inherent unobservable endogeneity

factor resulting from the first stage regression error. The TCE is identified whether the considered IVs are weak or strong. Our view is that both effects play an important role in capturing the net impact of a policy shock. In the context of two baseline empirical models with factor controls, results confirm that conventional 2SLS methods produce statistically insignificant responses at conventional levels. Using identification-robust approaches produces economically more plausible results, yet overall, we find that instruments are weakly informative on DCEs. Estimates of the TCEs provide critical insights which, in sharp contrast to DCE estimates, are mostly unchanged as additional credit spreads are considered. We find that outcomes can go in the opposite direction from what theory would predict on credit markets and the macroeconomy, which suggests that DCEs may miss important responses.

The third chapter is a joint work with Professor Hashmat Khan. In this paper, we examine the flow view of quantitative easing (QE) using monthly data on Federal Reserve's pre-announced asset purchases from the second and third rounds of QE. We determine both average and cumulative purchasing effects using structural VAR and local projection methods, respectively. For financial assets, we find that the purchasing shock increases the stock price index and the 10-year treasury yield, but it decreases the housing price index. For macro aggregates, there is no statistically significant average effects of the purchasing shock, however, we find that the accumulation of the purchases does increase both the industrial production and the consumer price index.

Preface

This thesis is comprised of three chapters. The first two chapters are joint with Professor Lynda Khalaf, and the third chapter is joint with Professor Hashmat Khan. I acknowledge the contribution of Professor Lynda Khalaf and Professor Hashmat Khan. I was fully involved in every step of the research. In all cases, my contribution is equal to that of my co-authors.

Contents

List of Tables

List of Figures

Chapter 1

Basel Liquidity Regulation and Bank Lending in the U.S.

1.1 Introduction

Banks face two major risks: capital risk and liquidity risk. Banks within G20 jurisdictions were only subject to capital regulations before the 2008 financial crisis. During that crisis, previously fully accessible money markets such as the repo market[1] dried up rapidly [Gorton and Metrick (2012)], and banks experienced difficulties because of the shortage of liquid assets even though they had a sufficient amount of capital. Consequently, liquidity regulations, such as the Liquidity Coverage Ratio[2] (LCR), were added to the Basel III accord, which was released as a response to the financial crisis to strengthen the resilience of the banking sector. While banking regulations are consequential for financial stability, liquidity regulations may restrain the provision of credit, at least to some degree.

In this paper, we evaluate the impact of the LCR on bank lending in the U.S. The reason for concern about lending relates to the definition of the LCR, which is equal to the ratio of the stock of High Quality of Liquid Assets (HQLA)[3] over the total net cash outflows over the next 30 calendar days. HQLA can be easily and immediately converted into cash with almost no loss to value during stress times. The numerator of the LCR ratio relates to the asset side of a bank's balance sheet, and the denominator is about the liability side. Theoretically, a bank can adjust the composition of either side of its balance sheet in order to comply with the LCR. It is important to note that loans are not part of HQLA. The minimum requirement for the LCR is 100% after full implementation of the regulation.[4]

As a matter of fact, recent crises provide motivation for concern about the potential effect of the LCR on bank lending. First, in September 2019, the interest rate in the U.S. repo market suddenly spiked to about 10% from 2%, implying significant shortages of cash. However, none of the largest US banks stepped in to make profits and stabilize the market. The Federal Reserve was forced to act as the "lender of the last resort". The sudden rise

[1] The repurchase, or repo, market is one of the major money markets, where corporations find short-term whole-sale fundings.

[2] https://www.bis.org/publ/bcbs238.htm

[3] Such as cash and T-bills.

[4] The full implementation was scheduled at 1 January 2019, and the LCR was phased-in gradually since 2015.

of the repo rate was a serious shock to the market not only because it was not expected, but also because it was reminiscent of the 2008 financial crisis. The Federal Reserve had to inject 75 billion dollars a day for four days to stabilize the market, and on October 11 2019, the Federal Reserve announced its plan to purchase about 60 billion dollars of Treasury bills per month and extend its repo operations at least until the second quarter of 2020.[5] Such reactions had not been seen for a decade. Many important figures in the financial industry pointed out that the LCR regulation possibly contributed to the shortage of liquidity in the repo market.[6][7] Amid the COVID-19 disruptions, the Federal Reserve announced an interim final rule to modify the LCR rule on May 5, 2020.[8] The interim final rule is an effort to make more loans available to the real economy through banking organizations. More recently, on March 19, 2021, the Federal reserve decided not to extend the change to the supplementary leverage ratio, or SLR, for bank holding companies.[9][10] This raises the discussion about whether or not the LCR should be implemented again, especially in 2022, when the Federal Reserve is raising interest rates and conducting Quantitative Tightening, which more broadly motivates our work on quantifying the effects of the LCR.

In this paper, we provide a formal causal analysis of the LCR on bank lending in the U.S with a state-of-the-art econometric method. Our motivation also draws upon the following study from the Federal Reserve Bank of New York. Roberts et al. (2022) apply the Differences-in-Differences (DiD) method to study the effect of the LCR on bank liquidity creation, which is measured by the Liquidity Mismatch Index of Bai et al. (2018), and the ratio of total loans to total assets of banks. They find a negative impact on both variables, yet suggest that demand and supply for credit are not separable from a causal perspective.[11] Since loans are not the only type of bank assets, banks can adjust loans and other assets as they reweigh the ratio from a dynamic perspective.[12] Therefore, and despite its motivational implications, the fact that the ratio of total loans to total assets declined does not necessarily imply reduction of total loans. The growth rate of total loans is also relevant, so the effect of the LCR on the provision of credit is still left to be examined. The current paper fills this gap.

Our contributions are two folds. Methodologically, we apply a recent econometric method developed by Callaway and Sant'Anna (2021) (hereafter CS(2021)). To the best of our knowledge, this is the first paper to use this method in the context of the LCR. Unlike the conventional two-way fixed effects method (TWFE), which is commonly used in the literature including Roberts et al. (2022), the CS(2021) method is unconstrained by the homogeneous (in covariates) treatment effects assumption and the no covariate-specific trends assumption, when pre-treatment covariates are needed. In section 1.4, we show that these two

[5] https://www.federalreserve.gov/econres/notes/feds-notes/what-happened-in-money-markets-in-september-2019-20200227.htm

[6] https://www.reuters.com/article/us-usa-fed-repo-banks/u-s-bankers-seize-on-repo-market-stress-to-push-for-softer-liquidity-rules-idUSKBN1W3288

[7] https://www.ft.com/content/45a9c196-e231-11e9-9743-db5a370481bc

[8] https://www.federalreserve.gov/newsevents/pressreleases/bcreg20200505a.htm

[9] https://www.federalreserve.gov/newsevents/pressreleases/bcreg20210319a.htm

[10] https://www.nytimes.com/2021/03/19/business/economy/federal-reserve-bank-leverage.html

[11] The section 8.2 in their paper talks about the separation of demand and supply for credit by using a different regression model.

[12] Table 2 in their paper lists all type of assets on a bank's balance sheet.

assumptions implicitly imposed in the TWFE model have large impacts on the results. Recently, Wooldridge (2021) proposes a modified TWFE framework that can address the above limitations of this method. We provide simulation evidence suggesting that the CS(2021) methodology closely corresponds to the proposed modification by Wooldridge (2021) in situations where the latter is applicable. In fact, this correction requires the addition of a large number of interaction terms which may be prohibitive in some contexts including our own. This said, the correction does not control for anticipation effects which we show matter importantly for the banking sector. Empirically, we find no significant impacts of the LCR on bank loan growth rates. Similar to Roberts et al. (2022)'s results, we find some effects on the ratio of loans to total assets. Our results also show that the effects of the LCR on the ratio of loans to total assets vary across different loan types.

Our paper contributes to an emerging and limited line of research on the effects of liquidity regulations. Results are scarce because liquidity regulation is recent. It has only received attention after the 2008 financial crisis. There are contributions from both major central banks and academia.

For central banks, one major source is the Bank for International Settlement (BIS), as many early stage impact studies have been conducted. BIS (2010b) examines macroeconomic impacts of the transition to stronger capital and liquidity requirements. Besides emphasizing the importance of the length of transition period, the paper predicts a median increase of 14 basis points in lending spreads and a median decrease in lending volume of 3.2 percent after four and a half years across countries. BIS (2010a) suggests that tighter capital and liquidity requirements on banks have long-term positive net benefits on aggregate economic output. BIS (2010c) uses information on 263 banks from 23 committee member jurisdictions; this includes 94 banks as group 1, which have Tier 1 capital in excess of 3 billion euros and the remaining 169 banks are in group 2. The paper concludes that banks are not satisfied with the LCR requirement on average, where group 1 banks have an LCR of 83 percent, and group 2 banks of 98 percent. Mordel (2018) provides a useful literature review of liquidity regulation.

Beyond the industry itself, Rochet (2004) shows that liquidity regulations could be a way to relieve central banks' pressure as the Lender of Last Resort. Perotti and Suarez (2011) develop a formal model for liquidity regulation, the model identifies both advantages and disadvantages of price and quantity policy instruments, and suggests a combination. Calomiris et al. (2015) study the supplementary effect of liquidity regulation over and above the capital regulation. Diamond and Kashyap (2016) show that banks tend to hold insufficient amount of liquid assets due to incomplete information. Banerjee and Mio (2018) study a liquidity regulation that is similar to the LCR and implemented in the UK. They estimate the causal effect of such a liquidity regulation on bank balance sheets, and find that the UK banks adjust both of their assets and liabilities in order to comply with the regulation, and that there is no evidence of shrinking the sizes of their balance sheet. Haan and van den End (2013) study the experience of the Dutch banks after Dutch government implemented a liquidity regulation similar to the LCR. They find that banks hold more liquid assets than their liquid liabilities, and more solvent banks tend to hold less liquid assets.

Overall, the existing literature suggests that banks would adjust both asset and liability sides of their balance sheets in order to comply with the LCR. Furthermore, the available literature to date suggests potential negative effects on bank lending in contrast to our main

finding, which reinforces the novelty of our contribution. Our findings result methodologically from allowing for and adequately capturing heterogeneous treatment effects.

The remainder of the paper is organized in the following order. Section 1.2 provides details about the identification and estimation of the two methods studied in this paper, including the simulation. Section 1.3, we discuss the dataset that is used for this study. In section 1.4, we present both methodological and empirical results. Section 1.5 concludes.

1.2 Identification and Estimation

The DiD method is built within the Potential Outcome model (POM) [Rubin (1974)]. There are two potential outcomes for each individual bank. Let $Y_{it}(1)$ be bank i's outcome (e.g the ratio of total loans to total assets) if it is covered by the LCR, while $Y_{it}(0)$ is bank i's outcome if it is not covered by the LCR. d_{it} is a dummy variable that indicates whether a bank is subject to the LCR at time t. Then, the observed outcome, Y_{it}, for each individual bank can be formalized as

$$Y_{it} = \begin{cases} Y_{it}(0) & \text{if } d_i = 0 \\ Y_{it}(1) & \text{if } d_i = 1 \end{cases}$$

so that

$$Y_{it} = Y_{it}(0) + d_i(Y_{it}(1) - Y_{it}(0)). \tag{1.1}$$

The causal effect of the LCR on the outcome of an individual bank is given by $(Y_{it}(1) - Y_{it}(0))$. The parameter of interest is the average treatment effect on the treated (ATT),

$$ATT = E[Y_{it}(1) - Y_{it}(0)|d_i = 1],$$

which helps to explain how much the banks restricted by the LCR are affected in terms of the outcome variables compared to what they would have experienced without the LCR.

The problem is that only one potential outcome can be observed for each individual bank. The observed outcome is called the factual outcome and the unobserved potential outcome is called the counterfactual outcome. A naive comparison of the average outcomes between the group of banks covered by the LCR and the group of banks not covered by the LCR is informative, but it does not have a causal interpretation. In a canonical 2 x 2 case, where there are only two groups (one is treated, one is not treated) and two time periods ($t = 0, 1$) this can be shown mathematically as

$$\begin{aligned}
&E[Y_{i1}|d_i = 1] - E[Y_{i1}|d_i = 0] \\
&= E[Y_{i1}(1)|d_i = 1] - E[Y_{i1}(0)|d_i = 0] \\
&= E[Y_{i1}(1)|d_i = 1] - E[Y_{i1}(0)|d_i = 1] + E[Y_{i1}(0)|d_i = 1] - E[Y_{i1}(0)|d_i = 0] \\
&= E[Y_{i1}(1) - Y_{i1}(0)|d_i = 1] + E[Y_{i1}(0)|d_i = 1] - E[Y_{i1}(0)|d_i = 0] \\
&= ATT + E[Y_{i1}(0)|d_i = 1] - E[Y_{i1}(0)|d_i = 0] \\
&= ATT + \text{Selection Bias},
\end{aligned}$$

where the second line is obtained by applying (1.1), then by adding and subtracting

$E[Y_{i1}(0)|d_i = 1]$ to get the third line. $E[Y_{i1}(0)|d_i = 1]$ is not directly observed, it is the average of counterfactual outcomes for the LCR covered banks. The above shows that the result of a naive comparison of an observed average difference cannot be interpreted causally because the causal effect, ATT, is confounded by the selection bias.

1.2.1 Identification challenges with conventional TWFE

Instead of simply comparing the average outcome between treated and untreated groups, the most popular method that has been used to identify ATT is to specify a TWFE model. In a canonical 2 x 2 case, the following model is commonly employed,

$$Y_{it} = \alpha_1 + \alpha_2 t + \alpha_3 d_i + \beta d_{it} + \epsilon_{it}, \tag{1.2}$$

where t is equal to 0 for the first period, and 1 for the second period. d_i is 1 if the individual bank is treated by LCR in the second period, 0 otherwise. $d_{it} = t * d_i$. ϵ_{it} is the error term. β is the parameter of interest.

A sufficient condition for β to identify ATT is the (unconditional) Parallel Trend assumption (PTA), which states that the average outcome of the treated group would have evolved in parallel to the average outcome of the comparison group, if there was no treatment.

Assumption 1 (Unconditional Parallel Trend Assumption)**.**

$$E[Y_{i1}(0) - Y_{i0}(0)|d_i = 1] = E[Y_{i1}(0) - Y_{i0}(0)|d_i = 0].$$

To see this, we simply need to realize that $\beta = E[Y_{i1}(1) - Y_{i1}(0)|d_i = 1]$. [13] However, this condition might be too stringent if there exist some covariates that are associated with outcome dynamics and are unbalanced between treatment and control groups. A notorious example is the Ashenfler's dip [Ashenfelter (1978)]. In fact, the selection into the training program depends on pre-treatment conditions, which can be represented as X_i, where X_i is a vector of observed characteristics, such as years of experience and education. Angrist and Pischke (2009) shows that it is very convenient to add time-varying covariates into (1.2), but the key challenge is to make sure these time-varying covariates are not affected by the treatment. Therefore, more commonly, empirical researchers introduce pre-treatment, or time-invariant covariates, X_i, into (1.2),

$$Y_{it} = \alpha_1 + \alpha_2 t + \alpha_3 d_i + \beta d_{it} + \gamma X_i + \epsilon_{it}, \tag{1.3}$$

where β is still the parameter of interest, and X_i is assumed to be uncorrelated with ϵ_{it}. In this case, the previous identification condition, Assumption 1, is no longer sufficient. Instead, the conditional PTA is needed.

Assumption 2 (Conditional Parallel Trend Assumption)**.**

$$E[Y_{i1}(0) - Y_{i0}(0)|d_i = 1, X_i] = E[Y_{i1}(0) - Y_{i0}(0)|d_i = 0, X_i].$$

[13] $E[Y_{i1}(1)|d_i = 1] = \alpha_1 + \alpha_2 + \alpha_3 + \beta, E[Y_{i1}(0)|d_i = 1] = E[Y_{i0}(0)|d_i = 1] + E[Y_{i1}(0) - Y_{i0}(0)|d_i = 0] = \alpha_1 + \alpha_2 + \alpha_3$.

Assumption 2 implies that, conditional on the pre-treatment covariates, the average outcomes of the treatment group and control group would be in parallel paths in the absence of the treatment. Since γ does not vary with either t or i, and time-invariant variables are differenced out when β is estimated, there are additional assumptions embedded in (1.3) in order for β to be identified as an ATT. The first assumption is homogeneous treatment effect for individuals with different pre-treatment conditions.

Assumption 3 (Homogeneous (in X_i) Treatment Effects)**.**

$$\beta = E[Y_{i1}(1) - Y_{i1}(0)|d_i = 1, X_i] = E[Y_{i1}(1) - Y_{i1}(0)|d_i = 1] = ATT.$$

Relating to the present paper, Assumption 3 entails the condition that banks with different balance sheet characteristics have the same LCR effects as long as they are treated. The second assumption is that there is no covariate-specific trend for both treatment and comparison groups.

Assumption 4 (No Covariates-specific Trends)**.** *For* $s = 0, 1$

$$E[Y_{i1}(s) - Y_{i0}(s)|d_i = s, X_i] = E[Y_{i1}(s) - Y_{i0}(s)|d_i = s].$$

In other words, it is assumed that outcome dynamics are homogeneous for individual banks with different pre-treatment conditions.[14] Abadie (2005) provides more details.

In practice, it is rare to have only two time periods. Traditionally, researchers apply the extended versions of (1.3) to allow more time periods,

$$Y_{it} = \alpha_i + \alpha_t + \beta d_{it} + \gamma X_i + \epsilon_{it}, \tag{1.4}$$

and,

$$Y_{it} = \alpha_i + \alpha_t + \sum_{e=-k}^{-2} \beta_e^{lead} d_{it}^e + \sum_{e=0}^{L} \beta_e^{lag} d_{it}^e + \gamma X_i + \epsilon_{it}. \tag{1.5}$$

(1.4) is called the "static" TWFE model, and (1.5) is the "dynamic" TWFE model. α_i and α_t are individual and time fixed effects. d_{it}^e is equal to 1 if the individual is e periods away from its initial treatment. X_i is a vector of pre-treatment covariates, and ϵ_{it} is the error term. K and L are positive integers. $e = -1$ is omitted as the reference period. In the static model, β is the parameter of interest, and it is often interpreted as an overall effect across time and individuals.

In many studies, researchers are not only interested in the overall treatment effect, but also in how the treatment effect evolves over time. Then (1.5) is usually used to answer such a question. The parameters of interest are $\beta_e^{lag}, e \geq 0$, and these parameters are interpreted as the treatment effect at different lengths of exposure. Since the period just before treatment implementation is omitted for normalization purpose, β_e^{lead} are often used to check the validity of the underlying PTA.

In addition to the implied assumptions of TWFE models that have been discussed previously, recent research has shown that practitioners need to be careful in interpreting these

[14] For example, with $s = 0$, $E[Y_{i1}(0) - Y_{i0}(0)|d_i = 0, X_i] = \alpha_1 + \alpha_2 + \gamma X_i - (\alpha_1 + \gamma X_i) = \alpha_2$. On the other hand, $E[Y_{i1}(0) - Y_{i0}(0)|d_i = 0] = \alpha_1 + \alpha_2 + \gamma E[X_i|d_{i1} = 0]) - (\alpha_1 + \gamma E[X_i|d_i = 0]) = \alpha_2$.

parameters as causal effects, especially when there are multiple treatment groups.[15] Papers such as Goodman-Bacon (2021), de Chaisemartin and D'Haultfœuille (2020), Athey and Imbens (2018), Borusyak and Jaravel (2017), and Sun and Abraham (2021) have shown that, in general, β and β_e^{lag} are weighted averages of some treatment effect parameters. These parameters do not recover ATT due to the negative weighting problem, which is mostly caused by having incorrect comparison groups. In particular, it is possible to have negative β while the treatment effects are in fact positive for all units.[16] Furthermore, these underlying weights can also be sensitive to other factors, such as the total number of time periods.

Despite recent critics of the TWFE model, Wooldridge (2021) shows that adequate corrections can restore its validity. By incorporating appropriate interaction terms between covariates, time periods, and treatment groups, researchers can allow for both a staggered design and heterogeneity within this model. We provide a simulation exercise to compare the performance of both methods in settings where there such comparisons are legitimate. Results confirm the arguments of Wooldridge (2021). This said, the TWFE adjustment of Wooldgrige does not directly allow for anticipation effects. In addition, the number of interaction terms can grow rapidly if there are many post treatment periods and covariates, which is our case.

1.2.2 Econometric solutions from CS(2021)

Theoretical advancements have only been made recently. In particular, CS(2021) have carefully studied the case with both covariates and anticipation effects. CS(2021)'s work builds on Heckman et al. (1997), Abadie (2005), and Sant'Anna and Zhao (2020) by extending their results to accommodate multiple groups and time periods.

We keep notations consistent with the original paper. The observed outcome for individual i becomes

$$Y_{i,t} = Y_{i,t}(0) + \sum_{g=2}^{T}(Y_{i,t}(g) - Y_{i,t}(0))G_{i,g},$$

where $G_{i,g} = 1$ if the individual bank is first treated at time g, then this individual bank belongs to treatment group g.[17] We will suppress the unit index i. The causal parameter of interest in their framework is called the group-time average treatment effect,

$$ATT(g,t) = E[Y_t(g) - Y_t(0)|G_g = 1].$$

Intuitively, this $ATT(g,t)$ basically breaks the multiple groups and time periods case into a family of canonical 2×2 cases. It is worth noting that no restrictions are imposed on treatment effect heterogeneity either across times or groups.

Besides standard assumptions needed for conditional DiD methods, there are two assumptions that are particularly noteworthy. The first one is the limited treatment anticipation assumption.

[15] This is called the staggered events.

[16] They focus on the cases where multiple groups getting treated at different time periods, and covariates are not included in the specifications. Baker et al. (2021) provides a good survey of this recent literature, also some simulation results to illustrate when TWFE works and when it does not work.

[17] In our case, there is only one treatment group, so the only g is equal to 17 (2015Q1).

Assumption 5 (Limited Treatment Anticipation)**.** *Let δ be the anticipation horizon,*

$$E[Y_t(g)|X, G_g = 1] = E[Y_t(0)|X, G_g = 1], t \leq g - \delta.$$

This assumption implies that there should be no treatment effects for time periods before 2013Q1 in our case. The second one is the conditional parallel trends based on a "Never-Treated Group".[18]

Assumption 6 (Conditional Parallel Trends Based on a "Never-Treated Group")**.**

$$E[Y_t(0) - Y_{t-1}(0)|G_g = 1, X] = E[Y_t(0) - Y_{t-1}(0)|C = 1, X], t \geq g - \delta.$$

$C = 1$ is used to represent the never treated group, and this group is always the comparison group in calculating all $ATT(g,t)$. Assumption 6 is the extension of Assumption 2. Specifically, Assumption 6 allows covariate specific trends in outcome dynamics, and this is crucial when the distributions of these pre-treatment covariates are not identical for different groups. It is important to emphasize that the identification of the $ATT(g,t)$ does not require Assumptions 3 and 4, in contrast to the TWFE. This is because the $ATT(g,t)$ are nonparametrically point-identified, and covariates are averaged out.

The authors have proposed inverse probability weighting (IPW), outcome regression (OR), and doubly robust (DR) estimands for the $ATT(g,t)$,

$$ATT^{nev}_{ipw}(g,t;\delta) = E[(\frac{G_g}{E[G_g]} - \frac{\frac{p_g(X)C}{1-p_g(X)}}{E[\frac{p_g(X)C}{1-p_g(X)}]})(Y_t - Y_{g-\delta-1})], \tag{1.6}$$

$$ATT^{nev}_{or}(g,t;\delta) = E[\frac{G_g}{E[G_g]}(Y_t - Y_{g-\delta-1} - m^{nev}_{g,t,\delta}(X))], \tag{1.7}$$

and

$$ATT^{nev}_{dr}(g,t;\delta) = E[(\frac{G_g}{E[G_g]} - \frac{\frac{p_g(X)C}{1-p_g(X)}}{E[\frac{p_g(X)C}{1-p_g(X)}]})(Y_t - Y_{g-\delta-1} - m^{nev}_{g,t,\delta}(X))], \tag{1.8}$$

where $m^{nev}_{g,t,\delta}(X) = E[Y_t - Y_{g-\delta-1}|X, C = 1]$, which is the outcome regression for the comparison group, and $p_g(X) = P(G_g = 1|X, G_g + C = 1)$, which is the probability that an individual gets treated starting at time period g conditional on pre-treatment covariates and belongs to either group g or the comparison group. $\frac{p_g(X)}{1-p_g(X)}$ helps give more weights to units in the never treated group that are similar in covariates to the units in the treatment group g. Note that $p_g(X)$ needs to be relatively far from 1, as the denominator of a ratio cannot be zero. This is the overlap condition. Therefore, we cannot include too many covariates so that the propensity score function can predict treatment perfectly. These three estimands are identical from an identification point of view,[19] and the key is to realize that $E[\frac{\frac{p_g(X)C}{1-p_g(X)}}{E[\frac{p_g(X)C}{1-p_g(X)}]})(Y_t - Y_{g-\delta-1})] = E[\frac{G_g}{E[G_g]}m^{nev}_{g,t,\delta}(X)]$.[20] Because of its specification robustness,

[18] The authors also consider the case where all units are eventually treated.

[19] They are not the same in terms of estimation/inference.

[20] Please refer to Appendix A CS(2021) for more details.

we will focus on the doubly robust estimand, $ATT_{dr}^{nev}(g,t;\delta)$, and this is also the one we calculate empirically. Note that the chosen reference time period is the period just before anyone can anticipate any treatments, $(g-\delta-1)$, and the comparison group is always the never treated group. This can be crucial as the negative weighting problem revealed by the recent literature is mostly caused by having incorrect comparison groups when the TWFE model is applied.

CS(2021) propose a two-step procedure to obtain the estimator for each $ATT_{dr}^{nev}(g,t;\delta)$,

$$\hat{ATT}_{dr}^{nev}(g,t,\delta) = E_n[(\hat{w}_g^{treat} - \hat{w}_g^{comp,nev})(Y_t - Y_{g-\delta-1} - \hat{m}_{g,t,\delta}^{nev}(X;\hat{\beta}_{g,t,\delta}^{nev}))] \tag{1.9}$$

where,

$$\hat{w}_g^{treat} = \frac{G_g}{E_n[G_g]}, \hat{w}_g^{comp,nev} = \frac{\frac{\hat{p}_g(X;\hat{\pi}_g)C}{1-\hat{p}_g(X;\hat{\pi}_g)}}{E_n[\frac{\hat{p}_g(X;\hat{\pi}_g)C}{1-\hat{p}_g(X;\hat{\pi}_g)}]}, E_n[Z] = \frac{1}{n}\sum_{i=1}^{n} Z_i.$$

The first step is to estimate the nuisance functions $m_{g,t,\delta}^{nev}(X)$, which is the evolution path of the outcome for the comparison group, and $p_g(X)$, which is the conditional probability that a bank belongs to treatment group g given pre-treatment covariates and that this bank is in either the treatment group g or the never-treated group (i..e comparison group).[21] We follow the suggestion of the authors. $\hat{m}_{g,t,\delta}^{nev}(X;\hat{\beta}_{g,t,\delta}^{nev})$ is the parametric estimator of $m_{g,t,\delta}^{nev}(X)$, where the linear model is used as the working model for the outcome regression,[22] and $\hat{\beta}_{g,t,\delta}^{nev}$ is the weighted least square estimator. $\hat{p}_g(X;\hat{\pi}_g)$ is the parametric estimator of $p_g(X)$, where the logistic regression model is used as the working model, and $\hat{\pi}_g$ is the inverse probability tilting estimator.[23]

CS(2021) also provide guidelines about how to aggregate all ATT(g,t) to answer different questions. In general, the aggregated parameters are weighted averages of ATT(g,t), and different weights can be used depending on the specific application. For the present paper, we are interested in the event study, or dynamic treatment effects. The weights suggested by the authors in this case focus on reflecting how treatment effects vary with the number of time periods that treated groups have been exposed to the treatment. In the graphs below, we report point and interval estimates of aggregated ATT(g,t) over g, which we denote as $AATT_t$. For further reference, the $(1-\alpha)$ interval estimates are denoted as $CI_a(AATT_t)$. α is the level of significance.

In addition to the ordinary standard errors and pointwise confidence intervals validated through weak convergence, authors have proposed a multiplier bootstrap procedure. This procedure enables users to calculate simultaneous confidence intervals that cover all parameters of interest jointly with a probability of at least $(1-\alpha)$. In the case of event study estimates, or $AATT_t$, simultaneous inference implies (asymptotically)

$$Pr[AATT_t \in CI_\alpha(AATT_t), t = 1, ..., T] \geq 1-\alpha.$$

In contrast, in the pointwise $(1-\alpha)$ confidence interval case, the overall probability of

[21] Again, in our case, there is only one treatment group and one comparison group, which is never treated.
[22] $E[Y_t|X, C=1] = X'\beta$.
[23] Please refer to Sant'Anna and Zhao (2020) for more details.

covering all parameters is not smaller than $(1 - T\alpha)$.[24] This is especially attractive for the current study because we are mostly interested in the treatment effect dynamics. Having simultaneously valid asymptotic inference helps prevent multiple-testing problems. This bootstrap procedure also allows for clustering.

1.2.3 Simulation

We conduct a small simulation exercise to demonstrate that the methods proposed by Wooldridge (2021) and CS2021 can recover the true ATT with an appropriate setup. The DGP[25] was made available in the DiD R package that was published together with CS2021. We use the case suitable for outcome regression only. Specifically, we include 140 individuals, with 29 time periods, and one covariate that follows a normal distribution with variance 1, and different means for treatment and control groups. This allow us to generate a situation similar to our empirical study. With 1,000 replications, the results are shown in Figure 1.1. The key message is that both methods can recover the true ATT if correct specifications are given, even in a relatively small N and large T situation.

Our results unify existing econometric methods that are relevant for our set-up. The number of interaction terms that are required to apply Wooldgrige's correction turned out to be prohibitive and associated TWFE regressions were infeasible. The results below are thus based on the CS2021. The latter as we have already emphasized allow for anticipation which is materially relevant in our setting in view of the policy announcements schedule that we duly exploit.

1.3 Data

This paper uses data retrieved from the Federal Deposit Insurance Corporation (FDIC).[26] Together with the Board of Governors of the Federal Reserve System (FRB) and the Office of the Comptroller of the Currency (OCC), they form the primary federal regulators.[27] The FDIC has a regulatory focus solely on U.S. commercial banks, or bank subsidiaries of bank holding companies (BHCs).[28] Figure 1.2 shows the percentage of banking assets held by commercial banks in the US, and it shows about 90 percent of loans in the banking industry are on commercial banks' balance sheets. Since the paper studies bank lending represented primarily by loans, this figure justifies the suitability of using data from the FDIC for our purpose.

There are in total more than 5,000 FDIC-Insured institutions, however, it might not be appropriate to include all of them for the analysis. This is because only BHCs that have 250 billion USD or more in total assets need to comply with the final rule of the LCR, and BHCs with more than 50 billion USD but less than 250 billion USD in total consolidated assets are

[24] Bolduc et al. (2010) provides more information.

[25] https://bcallaway11.github.io/did/articles/did-basics.html

[26] https://www7.fdic.gov/sdi/index.asp.

[27] https://www.fdic.gov/about/strategic/strategic/bankingindustry.html

[28] A bank holding company is a parent corporation that might hold more than one bank subsidiaries and other nonbank financial institutions.

subject to a modified, less strict version of the final rule. Depository institutions, which are subsidiaries of these covered BHCs, are subject to the LCR if they have at least 10 billion USD as consolidated assets.[29] Therefore, only large bank organizations are subject to the LCR, small and community banks are not. More importantly, Roberts et al. (2022) indicate that large balance sheet size difference should imply significant operational and characteristic differences for banks. Since the DiD method needs the control group to be similar to the treatment group, so that they are comparable, then only commercial banks with at least 3 billions USD of total assets in every quarter between 2011Q1 and 2018Q1 are included in our dataset.

We removed two commercial banks, State street bank and trust company and Wells Fargo financial national bank, from the original dataset. This is because their outstanding loan amounts suddenly decrease to zero around 2017 for one or two quarters and then back to similar levels as before. This creates huge loan growth rates, which lead to large standard errors in the estimation results. In total, there are 140 institutions included in this dataset, 27 of them are subject to the LCR, and the rest are not. We do not distinguish between banks that are subject to the full LCR and the modified LCR. Therefore, as long as a bank's BHC has a total asset value greater or equal to 50 billion USD[30] and the bank itself has a total asset value no less than 10 billion USD by 2014Q4, then the bank is subject to the LCR. Table 1.1 shows basic statistical information about the banks characteristics, which are the covariates in our estimations. The key information is that pre-treatment conditions are not balanced between treatment and control groups, and this is very important in terms of identification, which we will discuss in more details below. The dataset contains information about various loan types, and we focus on total loans, real estate loans, and commercial and industrial loans. The dataset contains each bank's quarterly balance sheet information ranging from 2011Q1 to 2018Q1. The starting time is chosen to avoid the 2008 financial crisis periods, and the ending time is selected because the LCR was modified to include less banks in 2018Q2.[31] The policy activation date is 2015Q1, this is the time period when the LCR became effective in the U.S. Since it was initially finalized by the BCBS in 2013Q1, then the banks had 2 years as the transition period, so that banks in the U.S can anticipate LCR implementation 8 quarters before it becomes effective.

1.4 Results

In this section, we first discuss our main results, which are based on the method proposed by CS(2021) with anticipation effects. Then, we also present results calculated by using the same method but without the anticipation effect. Lastly, we show results obtained with the conventional TWFE model. By comparing these three sets of results, we show that the additional assumptions of the TWFE model and the anticipation effect are consequential, and translate into meaningful differences empirically.

For all methods, we use exactly the same set of dependent variables and pre-treatment covariates. The dependent variables are ratios of loans to total asset values, and growth

[29] https://www.occ.treas.gov/news-issuances/federal-register/2014/79fr61440.pdf

[30] https://www.newyorkfed.org/medialibrary/media/research/banking_research/QuarterlyTrends2014Q4.pdf.

[31] https://www.federalreserve.gov/newsevents/pressreleases/files/bcreg20180706b1.pdf

rates[32] of different loans. As a reminder, we consider three types of loans, total loans, real estate loans, and commercial and industrial loans. Therefore, there are six left-hand side variables that we examine individually. We also include five pre-treatment covariates: total asset values growth rate (size growth), ratio of non-performing loans to total loans (NonPerloan/loan), ratio of non-performing loans to total asset values (NonPerloan/size), return on asset (roaq), and total risk-based capital ratio (rbcrwaj) for all methods employed.[33] They are used to capture each bank's balance sheet characteristics. The first period in calculation, 2011Q2, values of these covariates are used as the pre-treatment value.[34] This is the default choice of CS(2021)'s method.

Our first set of empirical results are obtained with CS(2021)'s method using the doubly robust estimator with Assumption 6. Results are reported in Figure 1.3 - 1.8. These figures can be divided into two sets. Figure 1.3 - 1.5 show every single group-time average treatment effect (i.e ATT(g,t))[35] estimated, and they are represented by the dots in the figures. The uniform (or, simultaneous) 95% confidence bands are presented by the vertical lines around each $ATT(g,t)$. Clustering is done at the individual bank level. Figure 1.6 - 1.8 depict the event study[36], that is, the dynamics of the average treatment effect by the length of exposure to LCR. The vertical segments represent joint coverages of 95% confidence intervals. Since there is only one treatment group in our case, the results of the event study are very similar to the results of the group-time average treatment effects. Across Figure 1.3 - 1.8, the red lines are the pre-treatment results, and they can be used as evidence against the parallel trend assumption if any of them are significantly different from zero, and the blue lines are post-treatment effects of the LCR. Panel (a) are the results generated when the ratio of each type of the loans over total asset values are left-hand side variables, and the dependent variables in Panel (b) are the growth rates of each type of loan. In general, we find statistically significant negative effect of the LCR on the ratio of the loans to total asset values, but such effect is not found with loan growth rates. This suggests that large banks have expanded their total asset values at a rate higher than the loan growth rates, but they did not slow down on lending. In other words, our results reveal a positive effect on liquidity that does not impede loan growth.

In addition, we also find heterogeneous treatment effects across loan types. This is most obvious among loan ratios.[37] Since the results for the ATT(g,t) and event studies are very similar as shown in the figures, we only focus on the event studies. Table 1.2 shows the dynamic effects of the LCR on different loan ratios.[38] Event time 0 is the treatment period, 2015Q1, and event time 12 is the last period in the dataset, 2018Q1. Results show

[32]The growth rates are calculated as natural log of first difference.

[33]We have also experimented with factor controls, by performing principle component analysis with all available performance ratios in the database. The ratios do not show high correlation, the first five principle components can only explain about 70 % of overall variations. Because of the limited number of observations and relatively long time periods, five covariates are selected as the bank characteristics that affect lending.

[34]The first time period, 2011Q1, in the dataset is omitted because growth rates are calculated.

[35]There is only one treatment group, which is treated in time period 17 (2015Q1). Therefore, the only group, indicated by g, is 17.

[36]The event study is generated by averaging ATT(g,t) across treatment groups for each time period. Please refer to CS(2021) for more details.

[37]There is no statistical significance for loan growth rates across all three types of loans.

[38]The results in this table are the same as the blue lines in Figure 1.6, 1.7, and 1.8.

that total loans make up about 5% less of the total asset values on average in 2015Q1 had the LRC not been implemented. Statistically significant effects are also found in 2015Q4, 2016Q2, and 2016Q3. The magnitudes of these effects are all around 5%. The LCR seems to have larger effects on real estate loans. As Table 1.2 and Figure 1.7 suggest, statistically significant effects are found in all post-treatment periods. The magnitudes are similar to the ratio of total loans to total asset values for 2015 and 2016. However, the treatment effects trends on the total loans ratio and the real estate loans ratio tend to diverge after 2017. The estimated effects of LCR on the real estate loans ratio extend to above 6% and stay statistically significant at 5% level, while the estimated effects of LCR on the total loans ratio decrease to approximately 3%, and they are no longer significant. For the commercial and industrial loans, we do not find statistically significant effects at any post-treatment periods for both of its ratio and growth rate, and the magnitudes are close to zero. However, significance is found in the pre-treatment periods for the ratio, so the results for the ratio need to be carefully interpreted.

Next, we discuss different results obtained with different identification strategies. Again, we focus on event studies with the ratio of total loans to total asset values as left-hand side variable. In the case with the assumptions of homogeneous (in X_i) treatment effects (i.e Assumption 3) effect and no covariates-specific trends (i.e Assumption 4) (hereafter, we refer to these two assumptions as the "additional" assumptions), we apply the dynamic TWFE model, (1.5). In terms of estimations with the TWFE model, it is not necessary to distinguish between conditional and unconditional cases when pre-treatment covariates are considered, because all time-invariant variables are differenced out before applying Ordinary Least Squares (OLS). Therefore, estimation is invariant to whether or not pre-treatment covariates are included. Clustering is at the individual bank level for all three cases, and the multiplier bootstrap method proposed in CS(2021) is used to generate the standard errors when aggregated ATT(g,t) are estimated. In the case without the "additional" assumptions, we apply the CS(2021) method, and obtain results in scenarios with anticipation effects and without anticipation effects.

Table 1.3 shows the estimated dynamics of treatment effects across the different assumptions used. It is clear to see that estimation results vary significantly when different assumptions are applied. The first two sets of results are obtained with TWFE and CS(2021)'s method, anticipation effects are not considered in both cases. The main difference between these two methods is whether the "additional" assumptions are imposed. We find no statistically significant effect of the LCR in the post-treatment periods if the TWFE model is applied, and the estimated ATT's are very close to zero. However, strong evidence is found in the pre-treatment periods against the PTA, as multiple estimated ATT's are significantly different from zero. Therefore, based on the TWFE model, it is difficult to draw any meaningful conclusion as strong evidence is found against the PTA. By using CS(2021)'s method without anticipation effects, the estimated ATT's are also very close to zero, and no effects are found in the post-treatment periods. In the pre-treatment periods, no evidence is found against the conditional PTA. Therefore, this method suggests that the LCR does not affect the total loans ratio. The last set of results are obtained by applying CS(2021)'s method with anticipation of 8 periods.[39] The estimated ATT's are about 10 times larger in

[39]Please refer to the previous section for the reason of 8 periods of anticipation.

the post-treatment periods and similar in the pre-treatment periods, relative to the other two sets of results. No evidence is found against the conditional PTA in the pre-treatment periods. Most importantly, LCR has significant effects on the ratio in the post-treatment periods, and such significance is not revealed if the other two methods are applied.

There is one possible modification in the model design. The Silicon Valley Bank became eligible to be treated with the LCR in 2017Q3, which is period 27 in our dataset. Since it is the only bank that joined the treatment group later than the rest treated banks, it should not be viewed as a threat to our main results. Nevertheless, it would make sense to check if our results are robust to such change in the model. Unfortunately, CS(2021)method does not allow for different anticipation periods for groups treated at different times, we are not able to implement the most natural design in this scenario. Instead, since all treated banks can anticipate the implementation of the LCR by 2013Q1, we create a pseudo-treatment group, which consists of both treatment groups and subjects to the LCR by 2013Q1. Therefore, anticipation is no longer allowed. Figure 1.9 - 1.11 show the event study results for each outcome variable. Despite small variations, results are very similar to previous event study figures. Our results thus seem consistent across these two model designs.

Overall, our findings can be summarized as following: methodologically, it is important to use the conditional PTA, and the "additional" assumptions embedded in the TWFE specification can translate into meaningful differences in applications. Anticipation effects also play an prominent role in our study. Empirically, we conclude that loans make up smaller portions of banks' assets in general, but variations exist across different loan types. We do not find evidence suggesting that banks have decreased lending, as loan growth rates are not significantly impacted by the LCR. Overall, results suggest that the LCR has had a positive effect on bank liquidity without affecting bank lending.

1.5 Conclusion

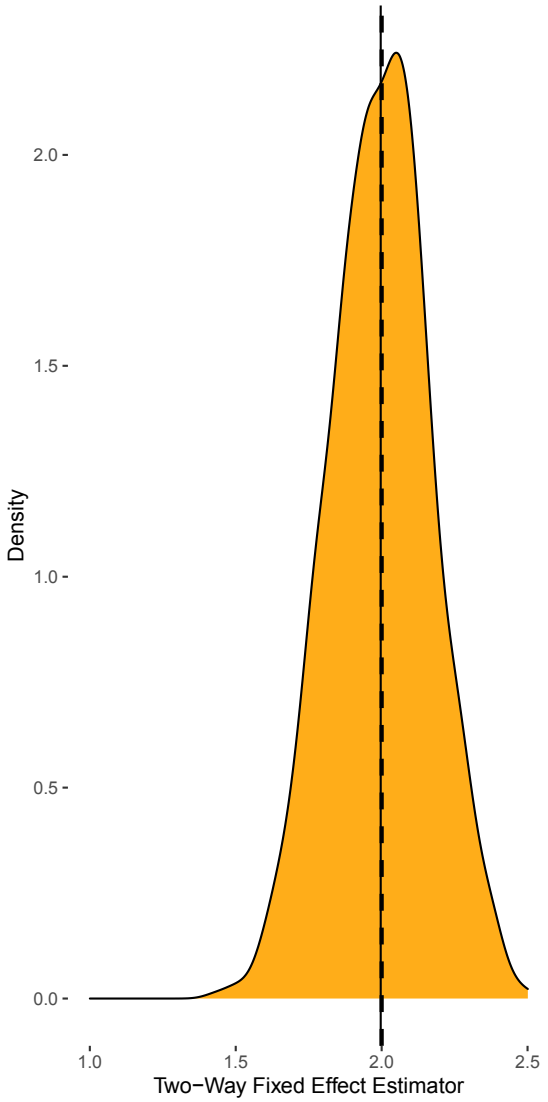
In this article, we examine the impact of the Liquidity Coverage Ratio (LCR) on bank lending in the U.S, using a Difference-in-Difference framework with a variety of identification methods. We are particularly interested in treatment effect dynamics. In this context, the dynamic two-way fixed effect (TWFE) model is commonly used. However, the embedded assumptions of this model, such as homogeneous treatment effects and no covariate-specific trends, become inapplicable when the distributions of pre-treatment covariates are not identical for different groups. Available evidence on the LCR ratio is scarce and is restricted to standard event studies. In this paper, we compare standard dynamic TWFE estimates to recently proposed alternative specifications that allow us to introduce various group-time aggregation schemes. Results underscore the importance of defining clear interpretable parameters, allowing for conditioning on covariates. In general, we find no effects of the LCR on bank lending, and the assumptions embedded in the TWFE models translate into meaningful differences in empirical results.

To conclude, we note that Dube et al. (2022) have recently attempted to link the Difference-in-Difference framework with possibly staggered treatments to the popular Local Projection method [Jordà (2005), Stock and Watson (2018), and Plagborg-Møller and Wolf (2021)]. This links to my second chapter and holds concrete promise for further econometric

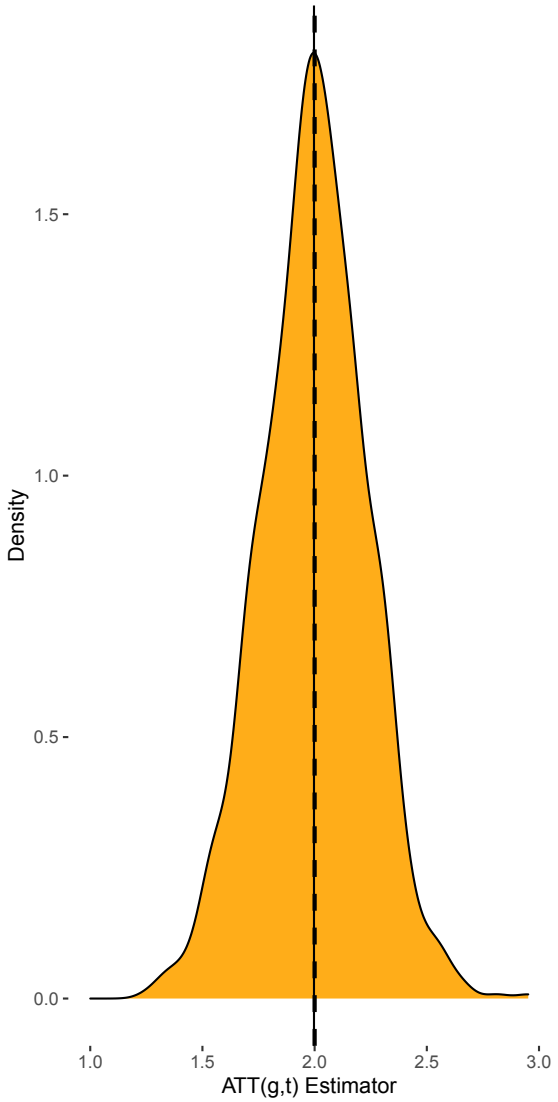
work beyond the scope of the current paper.

Figure 1.1: Simulation results

(a) Wooldridge (2021)

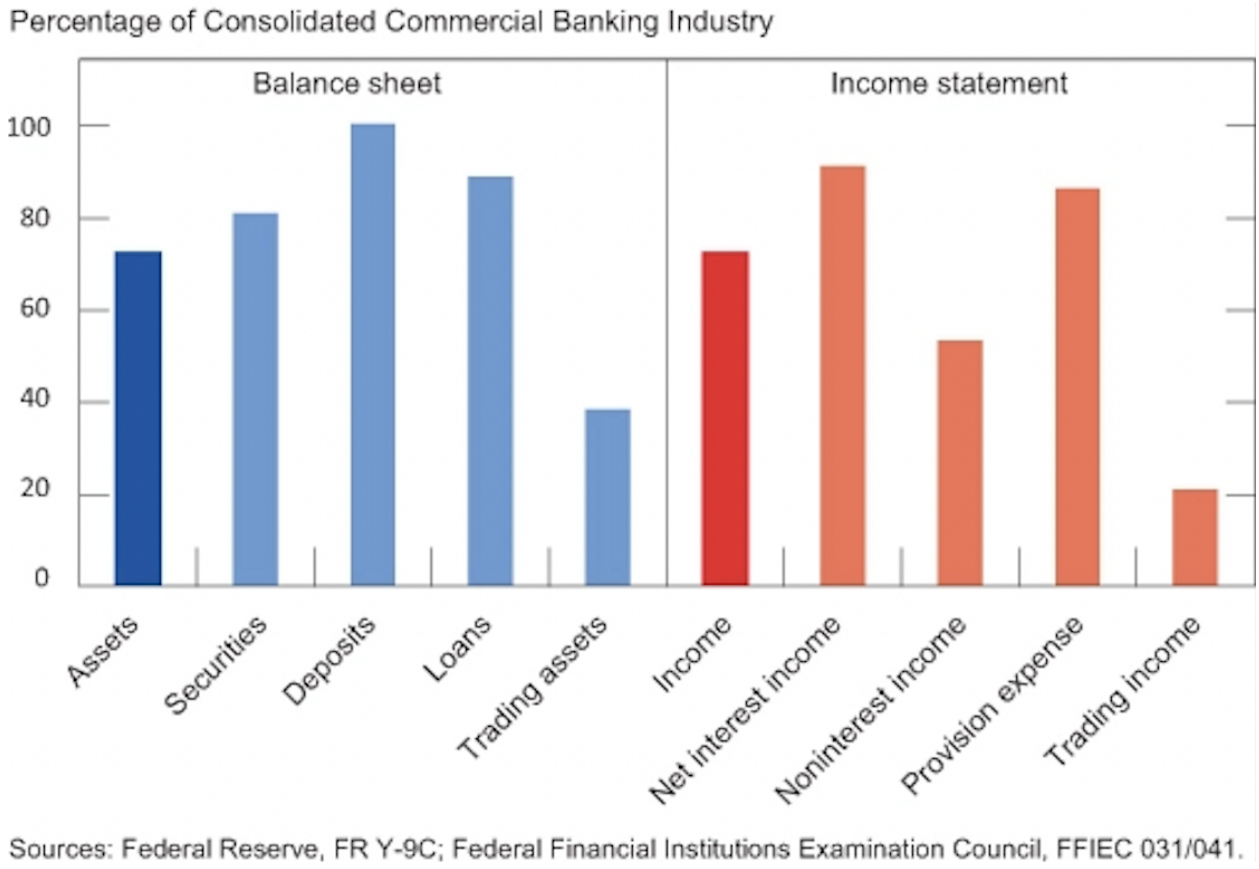


(b) Callaway and Sant'Anna (2021)



Note: The true ATT is set to be 2, which is represented by the solid vertical line. The dashed vertical line represents the average of the estimated ATT from the simulation. Similar to our situation with 140 individuals and 29 time periods, 1 covariate is included, with 1,000 replications. DGP was made available with CS(2021) in the DiD R package, we use the case suitable for outcome regression only.

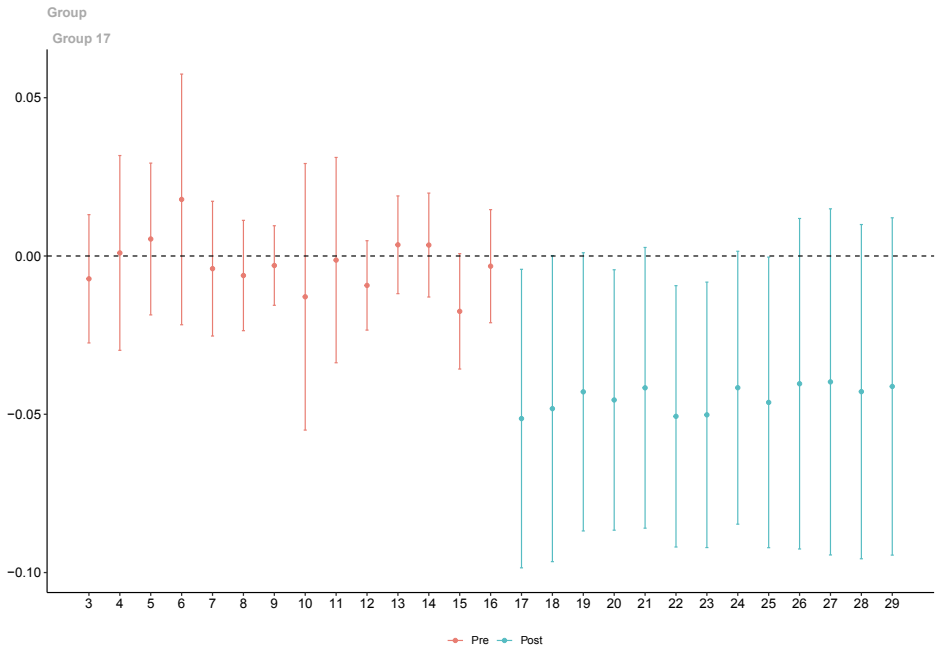
Figure 1.2: Commercial bank subsidiaries



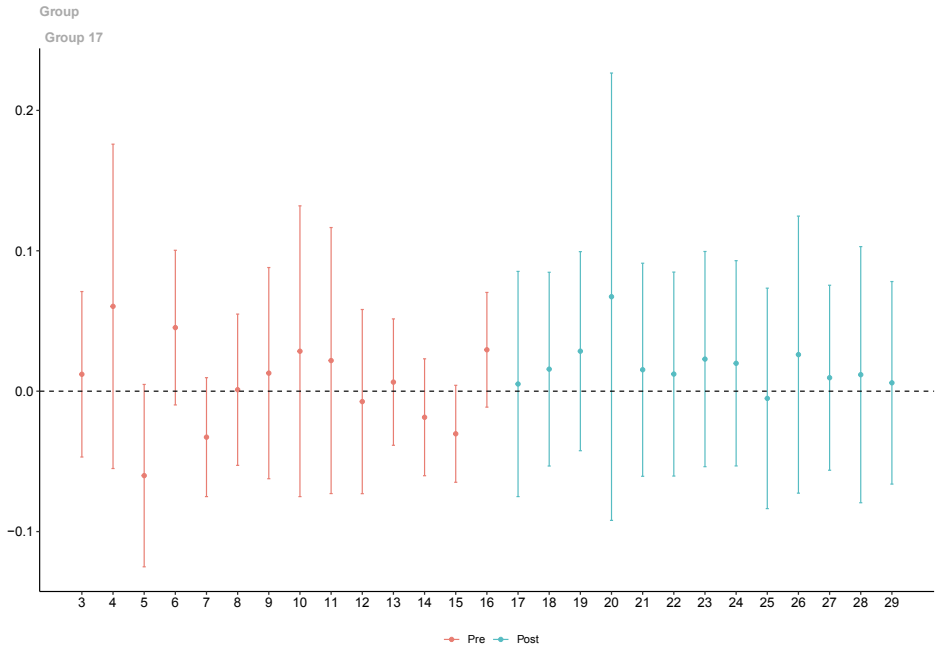
Note: This chart is obtained from a Federal Reserve Bank of New York article, which can be found here at https://libertystreeteconomics.newyorkfed.org/2012/10/tracking-the-us-banking-industry.html.

Figure 1.3: ATT(g,t) with anticipation, total loans.

(a) loan ratio, group.



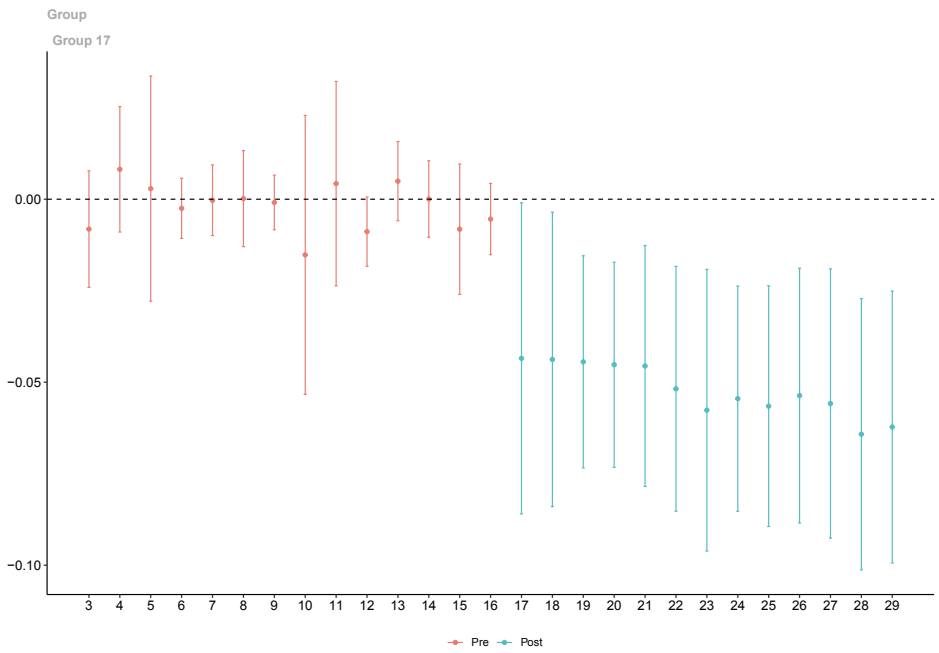
(b) loan growth, group.



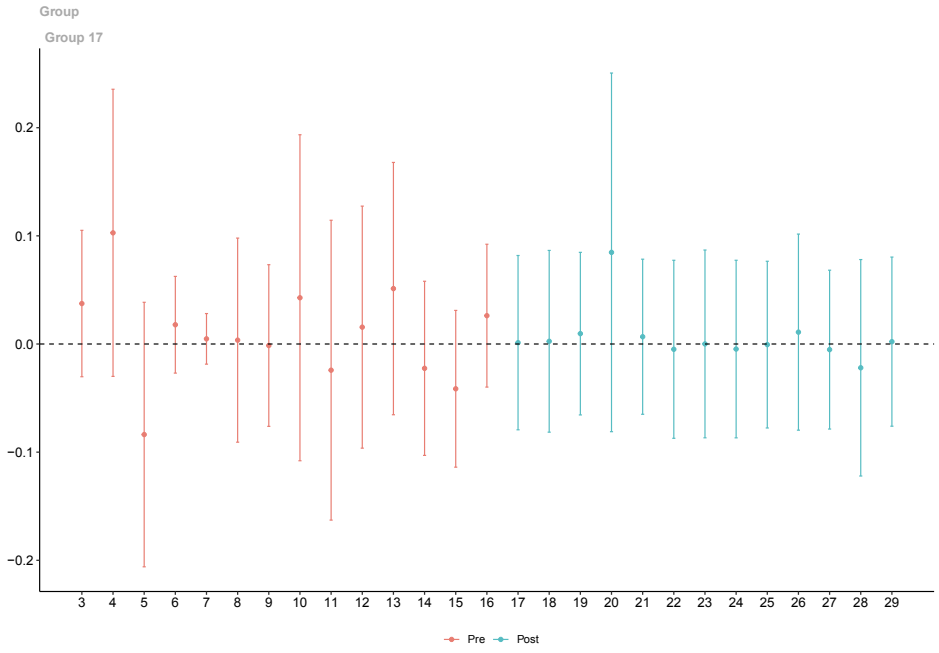
Notes: Each dot is a point estimate of the ATT for the treatment group (group 17) at each time period (i,e ATT(g,t)). They are estimated with the doubly robust estimator under conditional parallel trends assumption. Vertical lines around the dots are 95 % simultaneous confidence intervals allowing clustering at the individual bank level. The red lines are for the pre-treatment period, while the blue ones are for the post-treatment periods.

Figure 1.4: ATT(g,t) with anticipation, real estate loans.

(a) loan ratio, group.



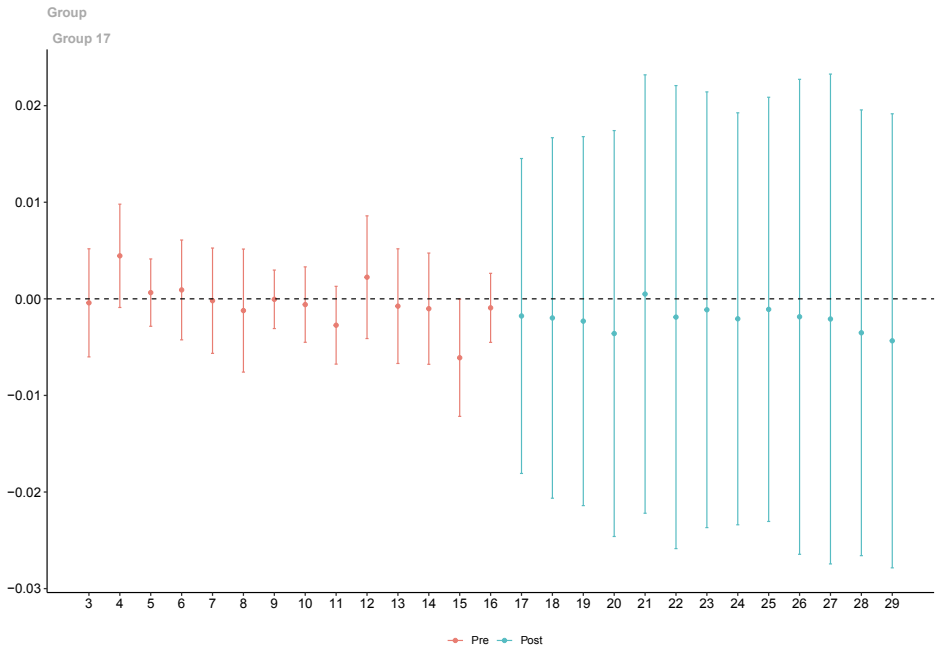
(b) loan growth, group.



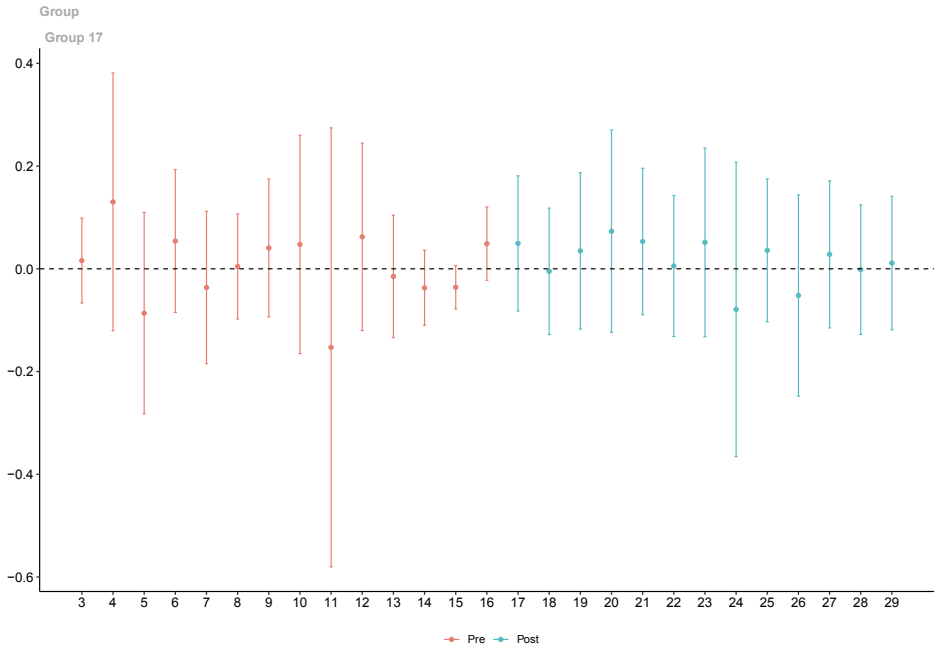
Notes: Each dot is a point estimate of the ATT for the treatment group (group 17) at each time period (i,e ATT(g,t)). They are estimated with the doubly robust estimator under conditional parallel trends assumption. Vertical lines around the dots are 95 % simultaneous confidence intervals allowing clustering at the individual bank level. The red lines are for the pre-treatment period, while the blue ones are for the post-treatment periods.

Figure 1.5: ATT(g,t) with anticipation, commercial and industrial loans.

(a) loan ratio, group.



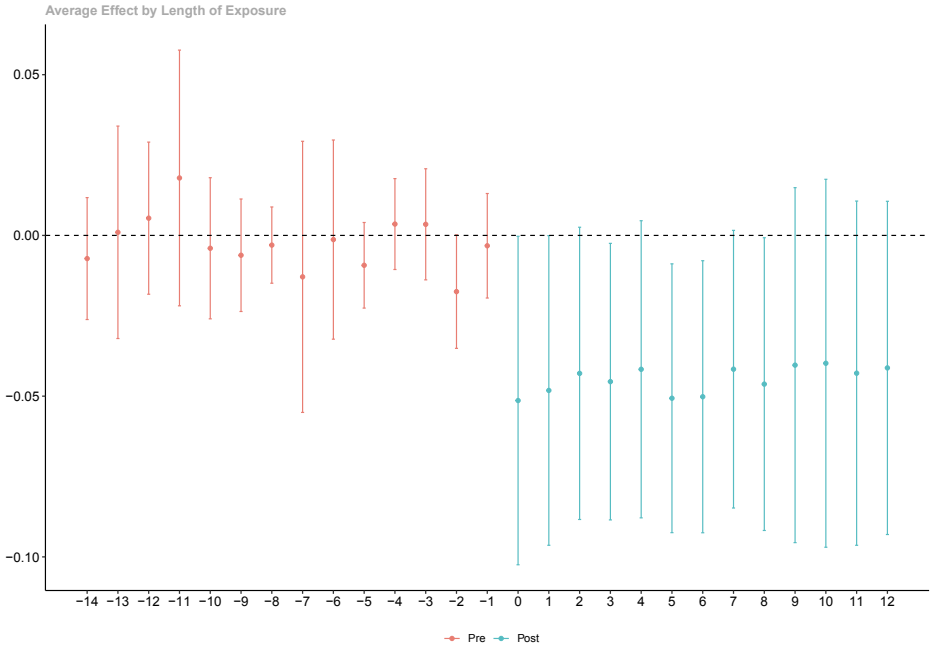
(b) loan growth, group.



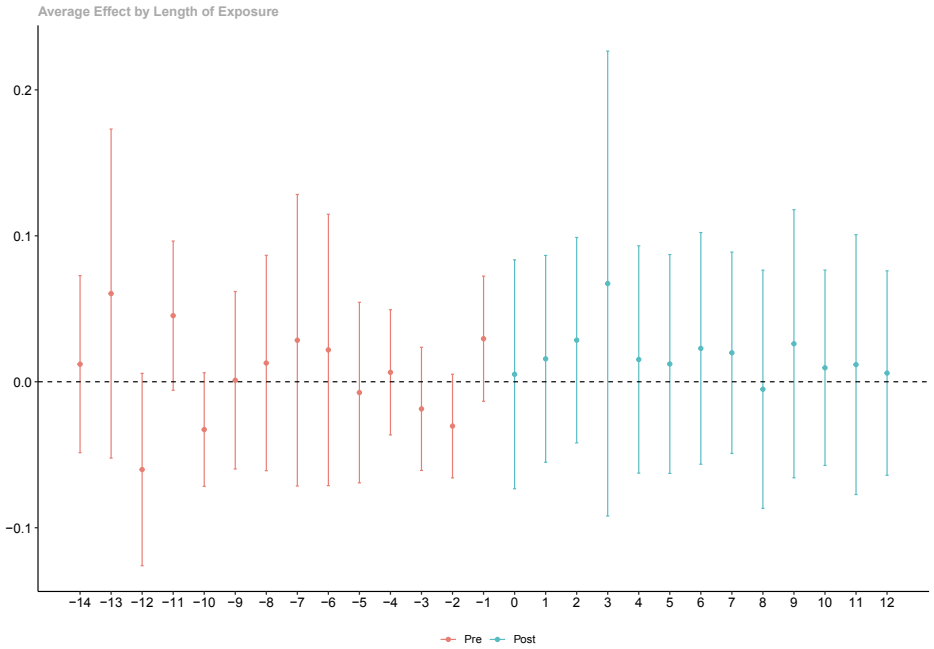
Notes: Each dot is a point estimate of the ATT for the treatment group (group 17) at each time period (i,e ATT(g,t)). They are estimated with the doubly robust estimator under conditional parallel trends assumption. Vertical lines around the dots are 95 % simultaneous confidence intervals allowing clustering at the individual bank level. The red lines are for the pre-treatment period, while the blue ones are for the post-treatment periods.

Figure 1.6: Event study with anticipation, total loans.

(a) loan ratio, event study.



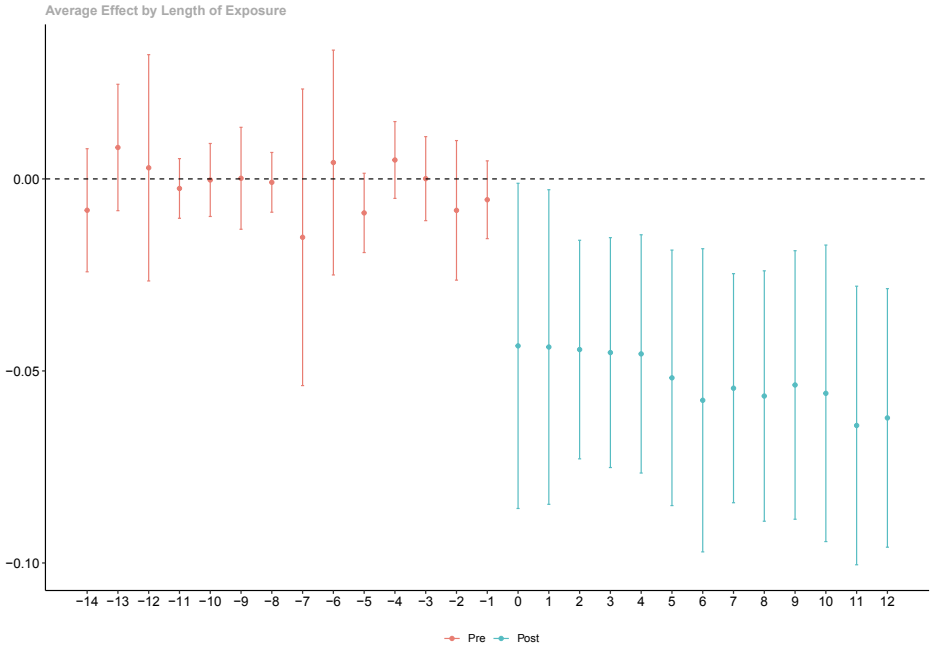
(b) loan growth, event study.



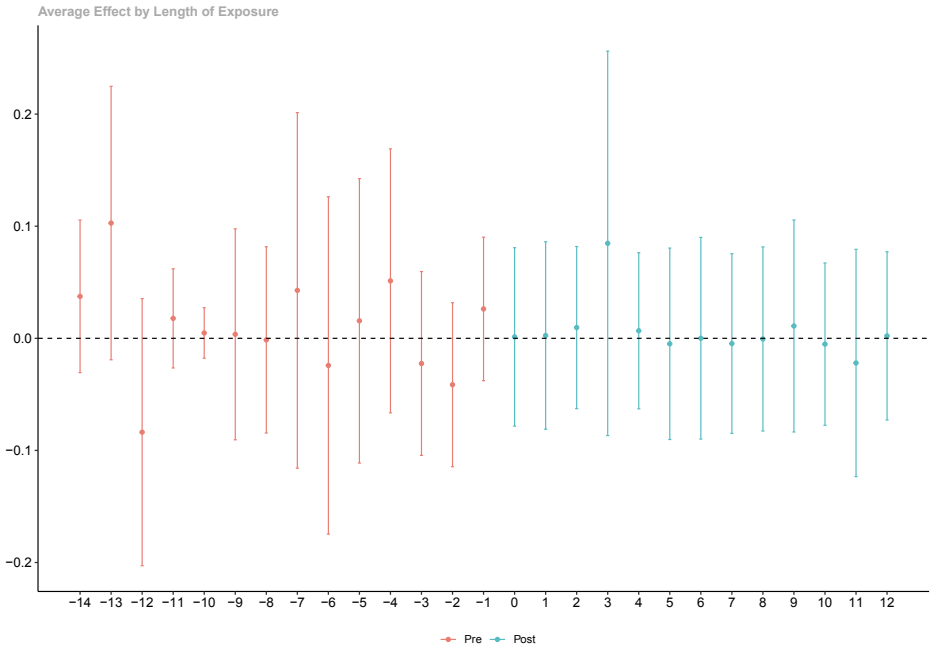
Notes: Each dot is a point estimate of the aggregated ATT for the treatment group by length of exposure to the LCR. Vertical lines around the dots are 95 % simultaneous confidence intervals allowing clustering at the individual bank level. The red lines are for the pre-treatment period, while the blue ones are for the post-treatment periods.

Figure 1.7: Event study with anticipation, real estate loans.

(a) loan ratio, event study.



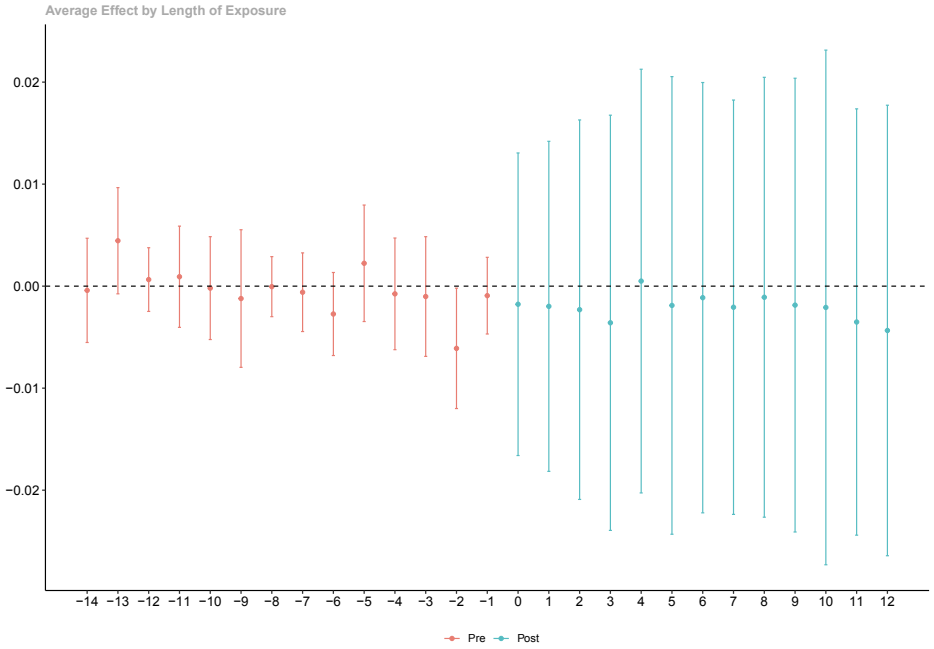
(b) loan growth, event study.



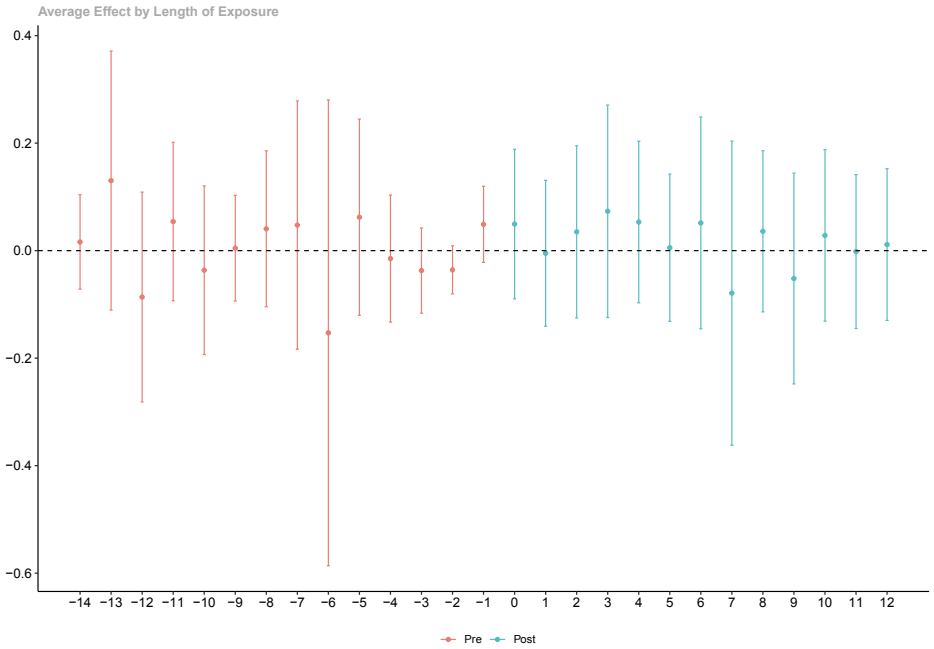
Notes: Each dot is a point estimate of the aggregated ATT for the treatment group by length of exposure to the LCR. Vertical lines around the dots are 95 % simultaneous confidence intervals allowing clustering at the individual bank level. The red lines are for the pre-treatment period, while the blue ones are for the post-treatment periods.

Figure 1.8: Event study with anticipation, commercial and industrial loans.

(a) loan ratio, event study.



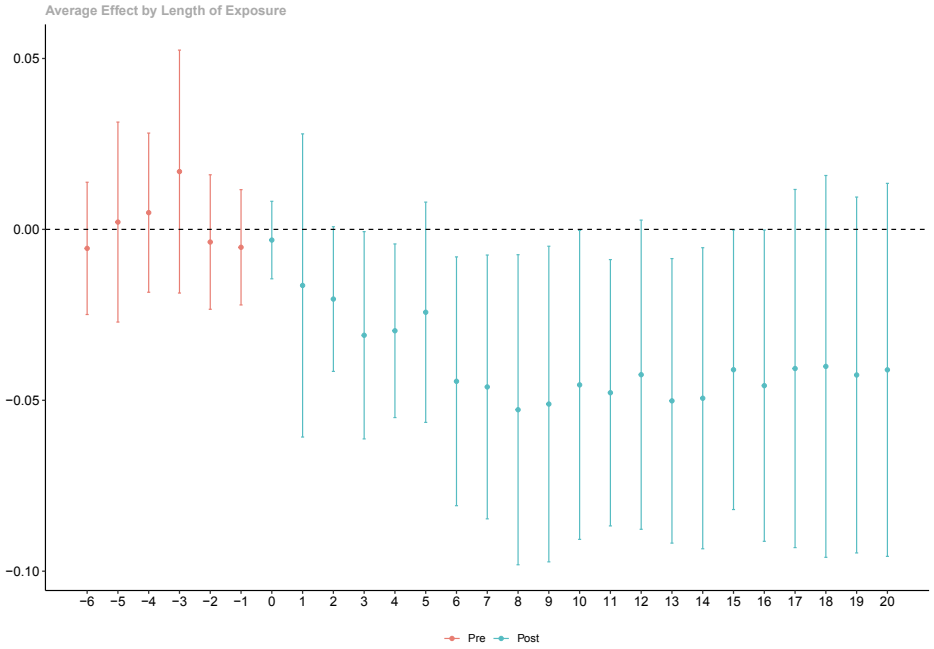
(b) loan growth, event study.



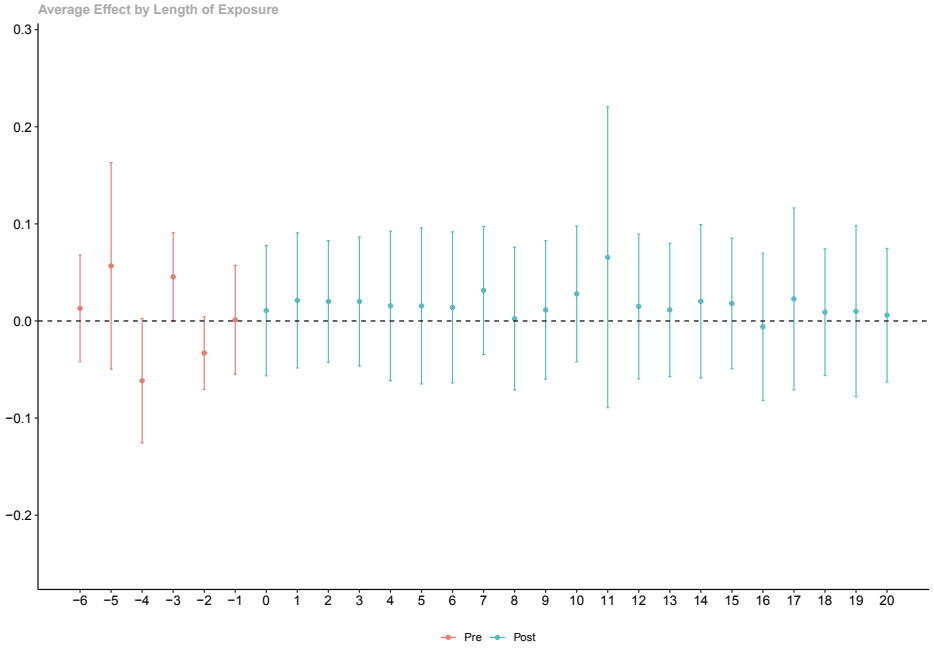
Notes: Each dot is a point estimate of the aggregated ATT for the treatment group by length of exposure to the LCR. Vertical lines around the dots are 95 % simultaneous confidence intervals allowing clustering at the individual bank level. The red lines are for the pre-treatment period, while the blue ones are for the post-treatment periods.

Figure 1.9: Event study with two treatment groups, total loans.

(a) loan ratio, event study.



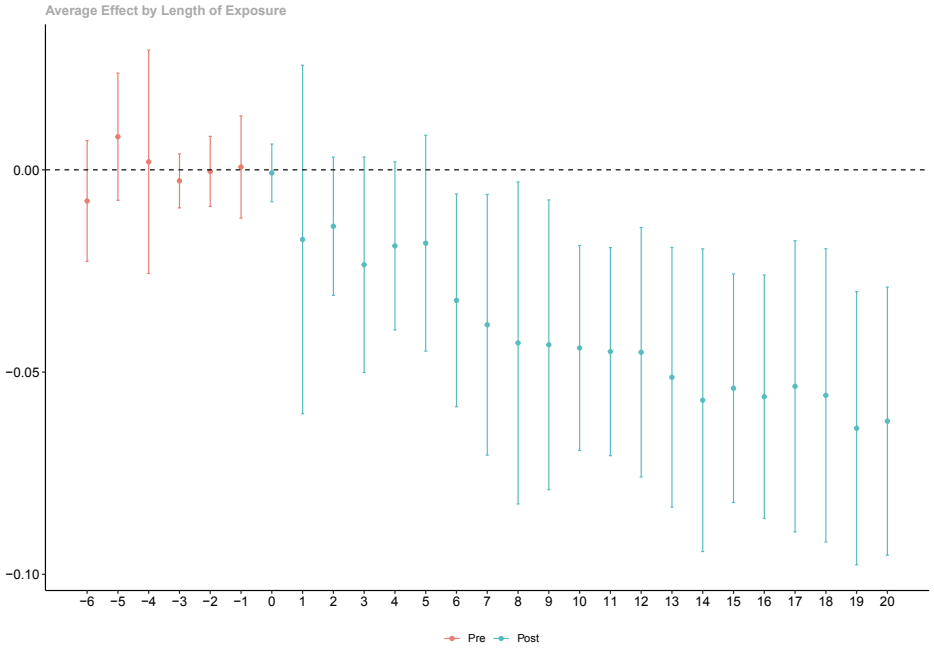
(b) loan growth, event study.



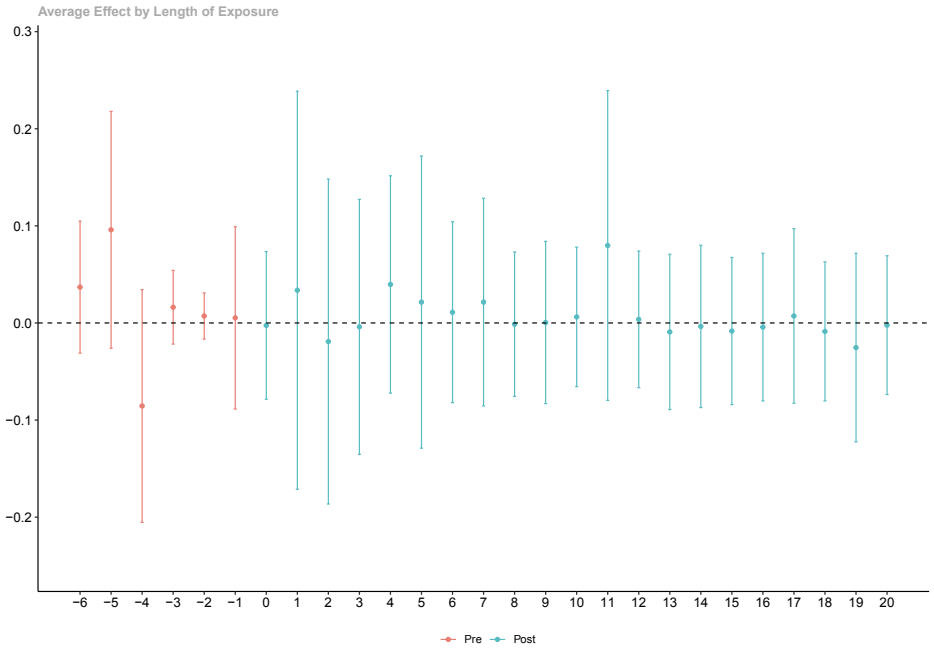
Notes: Each dot is a point estimate of the aggregated ATT for the treatment group by length of exposure to the LCR. Vertical lines around the dots are 95 % simultaneous confidence intervals allowing clustering at the individual bank level. The red lines are for the pre-treatment period, while the blue ones are for the post-treatment periods.

Figure 1.10: Event study with two treatment groups, real estate loans.

(a) loan ratio, event study.



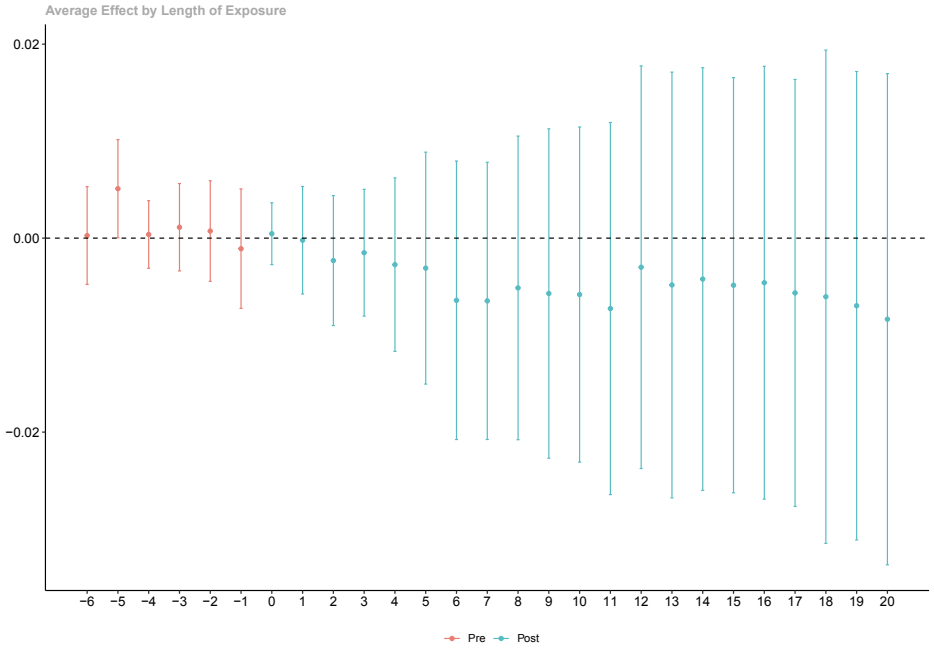
(b) loan growth, event study.



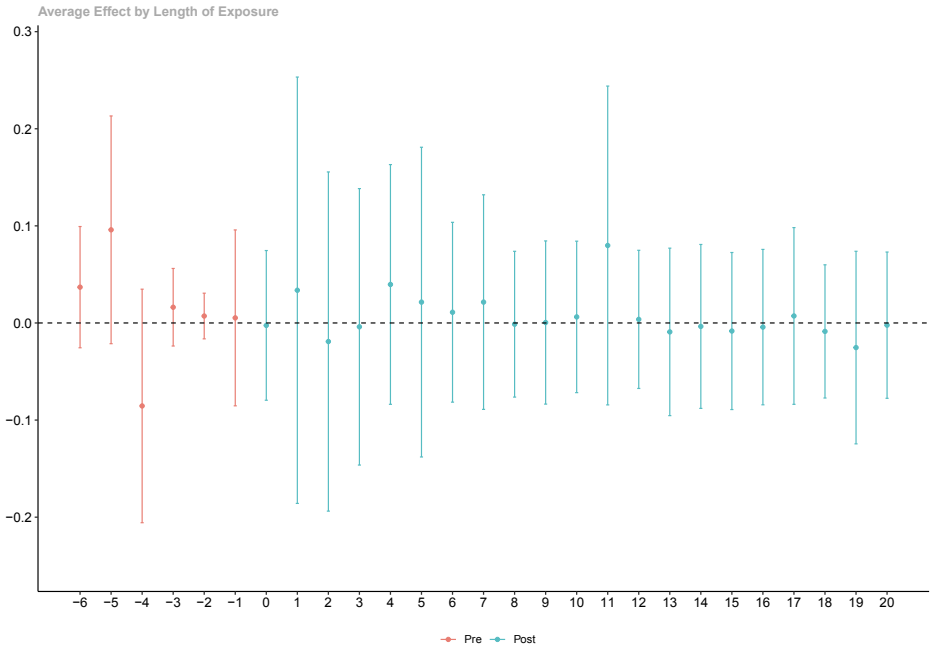
Notes: Each dot is a point estimate of the aggregated ATT for the treatment group by length of exposure to the LCR. Vertical lines around the dots are 95 % simultaneous confidence intervals allowing clustering at the individual bank level. The red lines are for the pre-treatment period, while the blue ones are for the post-treatment periods.

Figure 1.11: Event study with two treatment groups, commercial and industrial loans.

(a) loan ratio, event study.



(b) loan growth, event study.



Notes: Each dot is a point estimate of the aggregated ATT for the treatment group by length of exposure to the LCR. Vertical lines around the dots are 95 % simultaneous confidence intervals allowing clustering at the individual bank level. The red lines are for the pre-treatment period, while the blue ones are for the post-treatment periods.

Table 1.1: Summary Statistics of covariates, mean @ 2011Q2

	Untreated	Treated
size growth	0.00918	0.0211
NonPerloan/loan	0.0383	0.0296
NonPerloan/size	0.0234	0.0185
roaq	1.12	1.45
rbcrwaj	17.3	16.5

Notes: Authors' calculations. Data from the FDIC Statistics on Depository Institutions at 2011 Q2, which is the first period in the dataset used for estimation. The five pre-treatment covariates included in all models are total asset values growth rate (size growth), ratio of non-performing loans to total loans (NonPerloan/loan), ratio of non-performing loans to total asset values (NonPerloan/size), return on asset (roaq), and total risk-based capital ratio (rbcrwaj).

Table 1.2: Event studies

	Event time												
	0	1	2	3	4	5	6	7	8	9	10	11	12
	Ratio of total loans to total asset values												
ATT	-0.0514*	-0.0482	-0.0429	-0.0455*	-0.0416	-0.0507*	-0.0502*	-0.0416	-0.0463	-0.0404	-0.0398	-0.0429	-0.0412
Std.Err	0.0185	0.0196	0.017	0.0162	0.0183	0.0165	0.0175	0.017	0.0181	0.0193	0.0218	0.0206	0.0209
	Ratio of real estate loans to total asset values												
ATT	-0.0435*	-0.0438*	-0.0444*	-0.0452*	-0.0456*	-0.0518*	-0.0576*	-0.0545*	-0.0565*	-0.0537*	-0.0558*	-0.0642*	-0.0622*
Std.Err	0.0166	0.0161	0.0114	0.0118	0.0128	0.0133	0.0157	0.0122	0.0127	0.0147	0.0152	0.0141	0.0154
	Ratio of commercial and industrial loans to total asset values												
ATT	-0.0018	-0.002	-0.0023	-0.0036	0.0005	-0.0019	-0.0011	-0.0021	-0.0011	-0.0019	-0.0021	-0.0035	-0.0043
Std.Err	0.0055	0.0055	0.0063	0.0072	0.0076	0.0083	0.008	0.0076	0.0081	0.0082	0.0088	0.0082	0.0084

Notes: * indicates statistical significance at 5%. There are 140 banks in the dataset.

Table 1.3: Estimated treatment effect dynamics with different assumption

	TWFE		Aggregated ATT(g,t)		Aggregated ATT(g,t)	
	ATT	Std. Error	ATT	Std. Error	ATT	Std. Error
2011Q3	0.0383	0.0227	-0.0072	0.0075	-0.0072	0.0078
2011Q4	0.0463*	0.0200	0.001	0.0121	0.001	0.012
2012Q1	0.0486*	0.0191	0.0054	0.0096	0.0054	0.0094
2012Q2	0.0513**	0.0183	0.0179	0.0152	0.0179	0.0148
2012Q3	0.0511**	0.0176	-0.004	0.0079	-0.004	0.008
2012Q4	0.0454*	0.0183	-0.0062	0.0064	-0.0062	0.0066
2013Q1	0.0402*	0.0157	-0.003	0.0046	-0.003	0.0048
2013Q2	0.0250*	0.0120	-0.0129	0.0152	-0.0129	0.0173
2013Q3	0.0223	0.0128	-0.0013	0.0116	-0.0013	0.0125
2013Q4	0.0117	0.0102	-0.0093	0.0049	-0.0093	0.0049
2014Q1	0.0169*	0.0080	0.0035	0.0057	0.0035	0.0055
2014Q2	0.0240**	0.0077	0.0035	0.0066	0.0035	0.0064
2014Q3	0.0033	0.0062	-0.0175	0.0062	-0.0175	0.0067
2014Q4	0	0	-0.0032	0.0068	-0.0032	0.0068
2015Q1	-0.0058	0.0052	-0.0057	0.0051	-0.0514*	0.0192
2015Q2	-0.0027	0.0067	-0.0044	0.0065	-0.0482*	0.0182
2015Q3	0.0023	0.0085	0.0028	0.0092	-0.0429	0.0176
2015Q4	-0.0013	0.0074	-0.007	0.0069	-0.0455*	0.0154
2016Q1	0.0033	0.0082	-0.0016	0.0073	-0.0416	0.0174
2016Q2	-0.0049	0.0081	-0.0111	0.0074	-0.0507*	0.0164
2016Q3	-0.0039	0.0087	-0.0085	0.0091	-0.0502*	0.0173
2016Q4	0.0038	0.0092	-0.004	0.0103	-0.0416	0.0165
2017Q1	-0.0005	0.0098	-0.0081	0.011	-0.0463*	0.0163
2017Q2	0.0047	0.0126	-0.0115	0.0132	-0.0404	0.0212
2017Q3	0.0054	0.0134	-0.012	0.0132	-0.0398	0.022
2017Q4	0.0024	0.0127	-0.0119	0.0129	-0.0429	0.0216
2018Q1	0.0049	0.0128	-0.0105	0.014	-0.0412	0.0207
PTA	Conditional		Conditional		Conditional	
Additional assumptions	Yes		No		No	
Cluster at	Individual level		Individual level		Individual level	
Anticipation periods	0		0		8	

Notes: * indicates statistical significance at 5%, ** indicates statistical significance at 1%, and *** indicates statistical significance at 0.1%. Dependent variable is the ratio of total loans to total asset values. 2014Q4 is used as the reference period in TWFE model. The first time period, 2011Q1 is omited because growth rates are used. CS(2021) method does not produce anything for the second time period, 2011Q2, because the authors use the "short difference", $(Y_t - Y_{t-1})$, instead of the "long difference", $(Y_t - Y_{t-\delta-1})$, to estimate ATT(g,t) for all pre-treatment periods, $t < g - \delta$. All standard errors are obtained by clustering at individual bank level, in addition, standard errors of aggregated ATT(g,t)'s are obtained through multiplier bootstrap procedure proposed by CS(2021). There are 140 banks in the dataset.

Chapter 2

Monetary policy surprises: robust dynamic direct and total causal effects

2.1 Introduction

Evaluating the effects of monetary policies is an enduring challenge in macroeconomics. The enormous modern literature on the subject builds on the structural vector autoregression (SVAR) approach of Sims (1980), which originates from a much older literature on system of equations; see for example Kilian and Lütkepohl (2017) and Stock and Watson (2018) for recent reviews. An important feature in this strand of literature is that exogenous variation internal to the considered system is exploited to identify dynamic causal effects. Sources of exogenous variation obtained through restrictions on variables within the considered system are commonly defined as "internal instruments". In contrast, the modern microeconometric literature on causal inference has achieved influential advances by using variations outside the considered system to generate so-called "external instruments". The local projection instrumental variable (LP-IV) approach [Ramey (2016), Stock and Watson (2018)] allows one to easily use such instruments to estimate dynamic causal effects.

In this paper we build on the LP-IV approach to assess the effects of a monetary policy shock on the US economy. Following Gertler and Karadi (2015), we pay particular attention to the transmission channel through credit markets. Methodologically, we focus on bringing the rigor of weak-instruments robust methods to this literature. We first produce identification-robust Anderson Rubin (AR) confidence sets for the LP-IV response parameter.[1] For further reference, we denote the latter as the "direct effect". In addition, we use recent econometric methods [Doko Tchatoka and Dufour (2014), Doko Tchatoka and Dufour (2020), Beaulieu et al. (2022a), Beaulieu et al. (2022b)] to account for omitted variables. Specifically, we introduce an alternative LP-IV response parameter, which we denote as the "total effect", that extends the partial derivative foundations of LP-IV to account for unobservables in the first-stage regression. In contrast to the direct LP-IV effect, the total LP-IV effect is identified whether considered IVs are weak or strong. It is important to note that the "total effect" should be interpreted as the collective impact of the "direct effect" of the monetary policy shock and other missing factors that could either amplify or offset the

[1] Anderson and Rubin (1949); see Dufour (1997), Staiger and Stock (1997), Dufour and Taamouti (2005), and the surveys by Stock et al. (2002), Dufour (2003), Mikusheva (2013) and Andrews et al. (2019).

"direct effect".

So far, the literature on monetary policy evaluation seems to have focused on the direct effect, within a specified system of equations and given a set of controls. While there is an awareness that models (because of their specificity) will omit some variables, several historical episodes underscore the consequences of incomplete specifications. In particular, it is now widely acknowledged that the pre-2008 literature missed on the interactions between the financial sector and the real macro-economy, which has obscured warning signs about the looming crisis; see for example Jordà et al. (2013), Kuo et al. (2020), Inoue and Rossi (2021) and references therein. The earlier dot-com episode illustrates similar weaknesses that have led to unexpected effects with reference to theory [Stigum and Crescenzi (2007)]: following the Federal Reserve unprecedented interest-rate cuts in 2001, the stock market fell, credit spreads widened, the dollar value rose, and commercial and industrial lending did not recover until 2003; as rates were raised in 2004, economic activity strengthened, the stock market rose, credit spreads tightened, the dollar appreciated and lending standards relaxed. Other examples of omitted variables include fiscal policy in a model designed to study monetary policy, or foreign variables in a closed economy model; see Consolo et al. (2009).

Such arguments have traditionally led to model augmenting, possibly through factor LP-IV, which may or may not address the problem [Ramey (2016)]. Alternatively, Inoue and Rossi (2021) introduce the concept of functional LP, where the IVs can be functions. In this paper, we propose to parametrize the effect of missing intervening factors based on the endogeneity-corrected linear structural modelling approach of Doko Tchatoka and Dufour (2014). This approach leads to an LP-IV regression that is always identified. Said differently, we validate projections on an endogenous covariate while controlling for the instruments underlying LP-IV, as analyzed by Plagborg-Møller and Wolf (2021).

Empirically, we apply the above to revisit the analysis from Stock and Watson (2018) and Gertler and Karadi (2015). Our results reveal notable differences between direct and total effects, that underscore the impact of unobservables. While with few exceptions, most direct effects are insignificant (at conventional levels), a number of total effects which rely on the same instrumental information as their direct counterparts, are in fact significant, with noticeable sign reversals. In addition, total effects are robust to increasing the dimension of the underlying VAR, in stark contrast to direct effects which seem unstable when more information is included in the model. Robustness pertains to estimates and related inference. For example, the contemporaneous total effects of the excess bond premium are estimated to be -0.31% and statistically significant at 1 % level with and without the additional credit cost variables. This result implies that the financial market becomes looser as the net impact right after a monetary policy tightening. Moreover, the estimated total effects also confirm the predictability of the excess bond premium on output and price either with or without additional variables.

Our paper contributes to a growing literature in macroeconometrics, which uses external sources of variations to identify dynamic causal effects. External instruments can be used with or without an intervening VAR step. LP-IV refers to the latter case, building on the Local Projection approach of Jordà (2005); see Ramey (2016) on LP-IV in the context of monetary policy evaluation, and Jordà et al. (2015) and Ramey and Zubairy (2018), among other empirical applications. In contrast, the SVAR-IV methodology refers to structural VARs identified by external instruments, as introduced by Stock (2008). Such methods

have been used empirically by Stock and Watson (2012), Mertens and Ravn (2013), Gertler and Karadi (2015), and Caldara and Christophe (2017), among others. Stock and Watson (2018) and Plagborg-Møller and Wolf (2021) discuss both LP-IV and SVAR-IV in detail, and provide linkages between these methods. Recent related econometric works have focused on estimation concerns [Dake et al. (2022), Herbst and Johannsen (2021), Bruns and Lütkepohl (2022)], whereas contributions on inference are scarce particularly in the presence of possibly weak-IV [Ganics et al. (2021), Montiel Olea and Plagborg-Møller (2021), Montiel Olea et al. (2021)]. The concept of total effects originates from Doko Tchatoka and Dufour (2014) which we adapt here to the LP-IV framework.

2.2 Framework and econometric methods

The idea of using LP-IV to identify dynamic causal effects is rooted in the modern microeconometrics use of instrumental variables [Imbens (2014)]. In the potential outcomes context, an individual has two outcomes, one is observed and one is counterfactual, depending on whether the individual is treated or not. When assignment to treatment is random, the average treatment effect is defined as the difference between the average value of outcomes from treated individuals and the average value of outcomes from untreated individuals:

$$E[Y|W = 1] - E[Y|W = 0],$$

where Y denotes the outcome variable, and W is the treatment status. If linearity and homogeneous treatment effect are assumed, then the average treatment effect can be estimated by ordinary least squares (OLS) in the regression

$$Y = \alpha + \beta W + u,$$

where β is the parameter of interest and random assignment implies that $E[u|W] = 0$.

In observational studies, random assignment is rarely satisfied. Consequently, covariates are included in the linear model leading to the regression

$$Y = \alpha + \beta W + \gamma X + u,$$

where X includes all variables that are correlated with W and Y. In the latter regression, the conditional mean independence assumption, that is $E[u|W, X] = E[u|X]$, will restore the interpretation of β as the average treatment effect. Inclusion of covariates implies that the treatment is *as good as randomly assigned*, with proper controls. This strategy may also not suffice, which requires instrumentation. An IV denoted Z should be strongly correlated with W and satisfy the conditional independence assumption.

The above framework can be translated into macroeconometrics. Following the notations of Stock and Watson (2018), let $\epsilon_{1,t}$ be the mean zero structural shock associated with variable 1 at time t. The structural shock corresponds to random treatment in the microeconometrics jargon. If the structural shock is observed, and linearity and stationarity are assumed[2], then

[2]Homogeneous treatment effect is assumed throughout the paper.

the average treatment effect of variable 1 on variable 2, h periods ahead, can be obtained by OLS estimation of $\beta_{h,21}$ in the regression

$$Y_{2,t+h} = \beta_{h,21}\epsilon_{1,t} + u_{t+h}. \tag{2.1}$$

The simplicity of this framework has spurred a large literature that attempts to measure $\epsilon_{1,t}$ directly; see Stock and Watson (2018) for details and further references. The assumption of observable shocks is of course questionable. Instead, the popular structural vector moving average (SVMA) framework provides alternative foundations to define dynamic causal effects.

2.2.1 The LP-IV direct dynamic causal effect

An observed vector of macroeconomic variables $\mathbf{Y}_t = (Y_{1,t}, ..., Y_{n_Y,t})'$, assumed to be linear and stationary, has the SVMA representation

$$\mathbf{Y}_t = \mathbf{B}(L)\boldsymbol{\epsilon}_t,$$

where L is the lag operator and $\mathbf{B}(L) = \mathbf{B}_0 + \mathbf{B}_1 L + \mathbf{B}_2 L + ...$ $\mathbf{B}_h$ is $n_Y \times n_\epsilon$ where n_ϵ can be larger than n_Y. $\Sigma_\epsilon = \boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t'$ is assumed to be positive definite to rule out non-varying shocks. $\boldsymbol{\epsilon}_t$ = (ϵ_{1t}, ϵ_{2t}, ... ,$\epsilon_{n_\epsilon t}$)' contain all shocks and measurement errors. In this context, the structural shock is unobserved, so any measured shock is, at best, an incomplete proxy. However, if a proxy can be assumed to be exogenous relative to the VAR and strongly correlated with the true structural shock, then this proxy can serve as an external instrument.

One complication that is specific to macroeconometrics in this case is that the scale of the shock is unknown. Indeed, (2.1) still holds if $\epsilon_{1,t}$ is multiplied by a positive c and $\beta_{h,21}$ is multiplied by c^{-1}. Scale ambiguity thus requires some normalization, which includes the *unit effect* definition: one unit increase in $\epsilon_{1,t}$, increases $Y_{1,t}$ by one unit, that is:

$$\beta_{0,11} = 1. \tag{2.2}$$

Using the SVMA, (2.1) and the unit effect normalization leads to

$$Y_{1,t} = \epsilon_{1,t} + u_t, \tag{2.3}$$

where u_t is a linear combination of all current and past shocks except $\epsilon_{1,t}$. This allows one to view $Y_{1,t}$ as a policy indicator in empirical analysis, rather than a direct measurement of the policy in question. In our empirical analysis, $Y_{1,t}$ refers to a monetary policy indicator.

Substituting (2.3) into a regression of the form (2.1) gives

$$Y_{i,t+h} = \beta_{h,i1}Y_{1,t} + u^h_{i,t+h}, \tag{2.4}$$

where $u^h_{i,t+h} = \{\epsilon_{t+h}, ..., \epsilon_{t+1}, \epsilon_{2:m,t}, \epsilon_{t-1}...\}$, using the $\{.\}$ notation introduced by Stock and Watson (2018) to represent linear combination. Since the error terms from both (2.3) and (2.4) contain the current and past shocks except $\epsilon_{1,t}$, $Y_{1,t}$ is endogenous which calls for instrumentation. Let Z_t be a vector of instruments; Stock and Watson (2018) define the dynamic causal effects in the context of (2.4), which is denoted as the LP-IV regression,

through the following conditions:

$$
\begin{aligned}
&(1) \quad E[\epsilon_{1,t} Z_t'] = \alpha' \neq 0 && \text{(relevance)} \\
&(2) \quad E[\epsilon_{\bullet,t} Z_t'] = 0 && \text{(contemporaneous exogeneity)} \\
&(3) \quad E[\epsilon_{t+j} Z_t'] = 0 \text{ for } j \neq 0 && \text{(lead/lag exogeneity)}
\end{aligned}
$$

where $\epsilon_{\bullet,t}$ includes all shocks at time t except $\epsilon_{1,t}$. The first two conditions are standard. The lead/lag exogeneity assumption, which is needed since $u^h_{i,t+h}$ contains shocks from all leads and lags, can be particularly strong. This motivates inclusion of covariates in practice, which we denote as w_t. By projecting on the latter, the LP-IV regression yields

$$
Y^{\perp}_{i,t+h} = \beta_{h,i1} Y^{\perp}_{1,t} + u^{h\perp}_{i,t+h}, \tag{2.5}
$$

where $u^{h\perp}_{i,t+h} = \{\epsilon^{\perp}_{t+h}, ..., \epsilon^{\perp}_{t+1}, \epsilon^{\perp}_{2:m,t}, \epsilon^{\perp}_{t-1} ...\}$ and for any vector s_t, $s^{\perp}_t = s_t - \text{Proj}(s_t|w_t)$. In this case the LP-IV conditions become:

$$
\begin{aligned}
&(1) \quad E[\epsilon^{\perp}_{1,t} Z^{\perp\prime}_t] = \alpha' \neq 0, \\
&(2) \quad E[\epsilon^{\perp}_{\bullet,t} Z^{\perp\prime}_t] = 0, \\
&(3) \quad E[\epsilon^{\perp}_{t+j} Z^{\perp\prime}_t] = 0 \text{ for } j \neq 0.
\end{aligned}
$$

The underlying justification is that the external instrument might satisfy lead/lag exogeneity after including certain covariates. When these three conditions are satisfied, $\beta_{h,i1}$ captures the dynamic causal effects . The standard two-stage Least Square (2SLS) estimator denoted $\hat{\beta}_{h,i1}$ has been suggested for this regression; it is however now well known that weak-IVs compromise the asymptotic properties of this estimator.

The LP-IV regression (2.5) will generally not require invertibility, unless lags of $\boldsymbol{Y}$ are needed as controls. This condition which is commonly required in typical SVAR applications subsumes that shocks can be explained by the observable $\boldsymbol{Y}$s, which entails that the considered VAR system is complete. While avoiding invertibility is desirable, lead/lag exogeneity remains a crucial and possibly restrictive requirement.

2.2.2 The LP-IV total dynamic causal effect

The above linear IV framework leads us to the popular simultaneous equations framework, which we will adopt to simplify our exposition thereafter. We thus reconsider (2.5) with first-stage regression (projected on w_t)

$$
Y^{\perp}_{1,t} = \Pi_2' Z^{\perp}_t + V^{\perp}_{1,t}, \qquad V^{\perp}_{1,t} = V_t - \text{Proj}(V_t|w_t) \tag{2.6}
$$

where $E[V^{\perp}_{1,t}|Z^{\perp\prime}_t] = 0$. In this context, and building on Doko Tchatoka and Dufour (2014) and Plagborg-Møller and Wolf (2021), we introduce the total causal effect through the following projection

$$
u^{h\perp}_{i,t+h} = a_{h,i1} V^{\perp}_{1,t} + e^{\perp}_{i,t+h}, \qquad e^{\perp}_{i,t+h} = e_{i,t+h} - \text{Proj}(e_{i,t+h}|w_t) \tag{2.7}
$$

where $E[e^{\perp}_{i,t+h}|V^{\perp}_{1,t}] = 0$.

Substituting (2.7) into (2.5) yields the incomplete regression

$$Y^{\perp}_{i,t+h} = \beta_{h,i1} Y^{\perp}_{1,t} + a_{h,i1} V^{\perp}_{1,t} + e^{\perp}_{i,t+h}, \tag{2.8}$$

where the endogeneity factor $V^{\perp}_{1,t}$ intervenes as latent variable. Viewed generally, the latter can include missing factors, latent shocks, expectation errors, *etc.* Importantly, $e^{\perp}_{i,t+h}$, by its definition through (2.7), satisfies the conditional independence assumption

$$E[e^{\perp}_{i,t+h}|V^{\perp}_{1,t}, Y^{\perp}_{1,t}] = 0. \tag{2.9}$$

This allows us to combine (2.6) and (2.8) into the regression

$$Y^{\perp}_{i,t+h} = \theta_{h,i} Y^{\perp}_{1,t} + \pi'_{h,2} Z^{\perp}_{t} + e^{\perp}_{i,t+h}, \tag{2.10}$$

where $\theta_{h,i} = \beta_{h,i1} + a_{h,i1}$ and $\pi_{h,2} = -\Pi_2 a_{h,i1}$, and (2.9) identifies $\theta_{h,i}$. Throughout the paper, we follow the terminology in Doko Tchatoka and Dufour (2014) and refer to $\beta_{h,i1}$ as the *direct* (here, dynamic & causal) effects and $\theta_{h,i}$ as the *total*(dynamic & causal) effects.

The distinction between $\beta_{h,i1}$ and $\theta_{h,i}$ can be important when $Y_{1,t}$ is a policy indicator (and is thus endogenous). Nevertheless, (2.9) legitimizes the use of a traditional "*local projection on an endogenous covariate*" as we propose in (2.10), to recover dynamic causal effects "*while controlling for confounding variables*"; the terminology is from Plagborg-Møller and Wolf (2021), specifically in footnote 10, with reference to Jordà et al. (2013). In fact, (2.10) fits within formulation (1) in Plagborg-Møller and Wolf (2021) which, as shown by these authors, corresponds to a recursive SVAR on the system *augmented by the instrument*, with the instrument ordered first and $Y_{1,t}$ second.

For further reference, and for clarity regarding the latter ordering, let

$$y_t = (Z_t, Y_{1,t}, Y_{2,t}, ..., Y_{n_Y,t})'$$

and assume that w_t consists of lags of y_t. The main results of Plagborg-Møller and Wolf (2021) allows us to further clarify the difference between $\beta_{h,i1}$ and the $\theta_{h,i}$ parameter that we introduce. When (all)[3] lags of y_t are considered as controls, $\beta_{h,i1}$ corresponds [see Corollary 1 from Plagborg-Møller and Wolf (2021)] to a *relative* impulse response. Formally, it corresponds to the ratio of two coefficients in the context of the augmented system associated with y_t: the numerator is the impulse response of $Y_{i,t+h}$ with respect to an innovation in Z_t, and the denominator is the impulse response of $Y_{1,t}$ with respect to this same innovation. In contrast, within this same augmented system, $\theta_{h,i}$ is (up to a proportionality constant) the response of $Y_{i,t+h}$ with respect to an innovation in $Y_{1,t}$.

Thus defined, $\theta_{h,i}$ accounts for the same instrumental information (*i.e.* the first stage regression) as $\beta_{h,i1}$, but in contrast to the latter, the former is identified whether the instrument is strong or not. Furthermore, since $a_{h,i1}$ is unknown, $\theta_{h,i}$ accounts for the effects of endogeneity regardless of its extent, which also preserves the role of $Y_{1,t}$ as a proxy for $\epsilon_{1,t}$,

[3]These results hold "in population", which explains the infinite numbers of lags that are involved theoretically. Finite lag length have associated costs (as usual), which will not alter our interpretative discussion here.

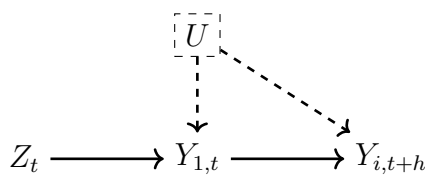
since endogeneity of $Y_{1,t}$ (within (2.5)) is embedded in the definition of $\theta_{h,i}$ through $a_{h,i1}$.

2.2.3 Underlying complete model assumptions

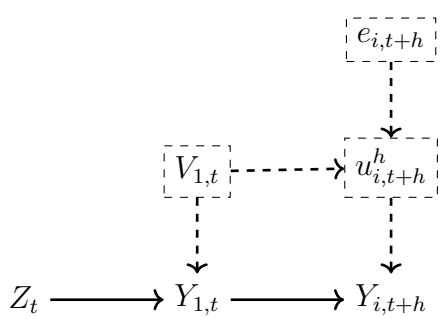
In contrast to standard SVAR analysis, the definition of both direct and total effects above does not require invertibility as defined in the following sense: $\epsilon_{1,t}$ can be linearly determined as a function of *only* current and past y_ts. It is important to note that invertibility here refers to the *augmented* information set that includes (or "internalizes") Z_t. Invertibility is restrictive because it entails the assumption that the considered model is complete.

Here, the LP-IV conditions (specifically (3) above), which can be formulated through a projection of Z_t on $\epsilon_{1,t}$ controlling for (all) lags of y_t, allows for measurement error as a confounder [see equation 17, under Assumption 4 in Plagborg-Møller and Wolf (2021)]. In view of (2.3), this implies that the first stage regression (2.6) error will also allow for such an unobserved confounder. This reinforces the usefulness of $\theta_{h,i}$ which - in contrast to $\beta_{h,i1}$ - embed the confounder's dynamic effects.

While the measurement error in question is assumed, under Assumption 4 of Plagborg-Møller and Wolf (2021), to be independent of all structural shocks at all leads and lags, (2.6)-(2.7) allows for unobservables broadly, provided an instrument can be found. To help further illustrate the implications of such confounders, we include two Directed Acyclic Graphs (DAGs) of IV (Imbens (2014), Imbens (2020) and references therein). We exclude covariates to ease presentation. In each graph, the arrows (directed edges) represent causality directions, and dashed lines show causality originates from unobserved variables.



The first DAG captures a system of equations (2.5) and (2.6) (abstracting from controls). U is the unobserved confounder. By using exogenous variation from Z_t, we are able to identify part of the causal effects of $Y_{1,t}$ to $Y_{i,t+h}$, and we call these type of effects the direct effects.



The second DAG captures a system of equations (2.8) and (2.6), with realization of (2.7), where $V_{1,t}$ contains missing and unobserved information. While the first DAG shows part of the effects of $Y_{1,t}$ to $Y_{i,t+h}$ identified through variation in Z_t, the second DAG shows that additional effects from $Y_{1,t}$ to $Y_{i,t+h}$ can be captured if $V_{1,t}$ is considered. The sum of the effects obtained through variations in both Z_t and $V_{1,t}$ portray the total effects.

2.3 Empirical Analysis

This section presents our empirical reassessment of the model considered by Stock and Watson (2018) to study the transmission of monetary policy on the macroeconomy and the financial sector in the US. The empirical model itself originates from Gertler and Karadi (2015) who consider a VAR with four variables, using monthly data from 1990m1 to 2012m6: the one-year US Treasury yield (R), industrial production (IP), the consumer price index (P), and the excess bond premium (EBP) as constructed by Gilchrist and Zakrajšek (2012). IP and P are growth rates measured in percentage points, and R and EBP are measured in percentage points at the annual rate. In what follows, we denote by Y all the considered variables. Our first set of results correspond to the four-dimensional case where Y includes R, IP, P and EBP. Subsequently, we add the mortgage spread (M) and three-month commercial paper spread ($COMM$). Unless further specified, significance in our discussion thereafter refers to standard levels (10, 5 and 1%).

The instrument is constructed based on the changes in the three months ahead Federal Funds futures rates within a 30 minutes window around FOMC announcements [Giirkaynak et al. (2005)]. This approach, that is commonly known as "High Frequency Identification" (HFI), has attracted increasing interest in view of its fundamental empirical implications. Ramey (2016) points out that as constructed by Gertler and Karadi (2015), the resulting instrument has a moving average structure. Several issues associated with its validity and strength are raised by Bauer and Swanson (2022) and Amir-Ahmadi et al. (2022), and alternative improvements are proposed. For benchmarking purposes, our analysis relies on this instrument as originally proposed, and we refer to it as Z in subsequent tables (in line with our econometric framework). Controls include four lags of Y, four lags of Z, and four lags of the first three principle components of the FRED-MD dataset as constructed by McCracken and Ng (2015), which we refer to as the three-dimensional variable F.

For ease of discussion in what follows, we will refer to the direct causal effect $\beta_{h,i}$ [in (2.5)] as the **DCE,** and its total effect analogue $\theta_{h,i}$ [in (2.10)] as the **TCE**. Estimates of both effects on each variable use the same above defined HFI-based instrument and the same control variables. Discrepancies between estimates can thus be attributed to the $a_{h,i1}$ parameter [in (2.8)] that captures the *endogeneity factor*, that is, the effect of the unobservable first stage regression disturbance $V_{1,t}^{\perp}$. Interpretation-wise, recall that R is used as the monetary policy indicator; the responses thus pertain to monetary policy tightening. Furthermore, the unit-effect normalization implies that the contemporaneous effect ($h = 0$) of the considered shock on R is set to 1. Reporting-wise, our estimates are summarized in confidence set form. In contrast with 2SLS estimates of the DCEs or the OLS-based estimates of the TCEs, the AR DCE sets are not necessarily symmetric nor can they be summarized in an interval form based on a point estimate with companion standard errors. We thus report confidence sets for all considered methods to facilitate comparisons. Newey-West HAC corrections are used throughout, with 24 lags following Stock and Watson (2018).

We begin our analysis with Table 2.1 that replicates the LP-IV results from Table 1 in Stock and Watson (2018), on the above defined four-variate VAR.[4] As pointed out by Stock and Watson (2018), results broadly illustrate the importance of covariates, which in this case

[4]Our reported confidence intervals match the standard errors from Table 1 of Stock and Watson (2018).

include factor controls. Indeed, the first stage f-statistics in column 1 are lower than the threshold suggested by Stock et al. (2002); in contrast, the first stage f-statistics in columns 2 and 3, that account for covariates, exceed that threshold. Low first stage f-statistics raise weak instruments concerns. In addition, the LP-IV exogeneity conditions are more likely to be met when lags of Z are included in the regression. This said, the above threshold has recently been revised upwards [Lee et al. (forthcoming)], and by a large margin. Such revisions reinforce the usefulness of our AR-based inference; refer to Table 2.2 robust p-values associated with tests of $\beta_{h,i} = 0$ and Table 2.3 for companion confidence sets for $\beta_{h,i}$.

The first major take-away of the above results is that 2SLS obliterates the information content of the considered (and popular) instrument. Indeed, there are very few statistically significant estimates across all three columns of Table 2.1. This is a disconcerting finding that calls into question the standard perspective on the effects of monetary policy, to the extent the literature has contemplated. Following the account of Andrews et al. (2019) and references therein, we argue that weak instruments are a more likely explanation. This view is supported by our results using the AR method: our reassessment in Table 2.2 produces stronger and more plausible significance test results, in sharp contrast to 2SLS-based methods.

The most prominent change occurs with P and EBP: DCEs are significant for P at 24 months into the future, and for EBP, contemporaneously and 6 months after the shock. The DCE point estimate, that can be interpreted as the least rejected value [Dufour et al. (2006), Dufour et al. (2010), Dufour et al. (2013)], is actually covered by our AR confidence set.[5] The results from Table 2.1 can then be interpreted as follows: monetary policy tightening decreases inflation by 2.6% within two years and increases EBP by 0.8% on impact and 1.66% within a half-year, which signals a deterioration in risk appetite or credit market sentiment. Relative to 2SLS that produces insignificant financial and economic outcomes, our AR-based estimates thus yield some - *albeit only a few* - DCEs that are significantly in line with general theoretical predictions.

Methodologically, such contrasts provide further useful insights on the trade-off between systems-based and the LP-IV single-equations approach. Insignificant findings such as those obtained in Table 2.1 can be attributed to power concerns relative to more efficient full system based methods. Stock and Watson (2018) explain such issues and propose an alternative SVAR-IV approach that nevertheless remains strong-IV based and requires invertibility, that is, a complete model assumption. Montiel Olea et al. (2021) provide a robust system-based alternative. Instead, we argue that single-equation LP-IV which avoids a restrictive complete model assumption, can inform on macro-theory as long as weak-IV robust methods are used.

Moving beyond significance tests, let us further examine the AR-based confidence sets for DCEs in Table 2.3. Reported results follow the format of Table 2.1: column 1 excludes covariates while column 3 includes all the covariates considered by Stock and Watson (2018). For each coefficient, we report the confidence set obtained through grid search based robust test inversion and its associated search set. 1000 points are screened within the search set. The default search set is 2 times the Wald-type confidence interval as suggested by Finlay et al. (2016). If no closed confidence interval is found within the default set, we broaden our search set within 20 times the original Wald-type confidence interval and report this

[5] Refer to Table 2.3 which we will discuss with further details below.

modification in the following format [*search set*] (20).

The most salient result from Table 2.3 is that no single confidence set is bounded in column 1, whereas all sets are closed intervals after some (or all) covariates are added. This observation confirms the discussion above on the importance of covariates. The role of factor controls is again worth emphasizing here. In addition, even though AR confidence intervals in column 3 are for the most part fairly tight, they tend to be roughly 2 times wider than the confidence intervals generated through 2SLS. Estimation uncertainty is thus clearly understated with results available in the literature so far.

Taken collectively, our results on DCEs, using the considered conventional baseline specification and variables, support the emerging literature (cited above) that seeks to improve the relevance of the instrument. We thus proceed to analyze the TCEs; these parameters are identified whether the instrument is relevant or not, eschewing one of the main difficulties that plague inference on the DCEs. Overall, our comparison of estimates of the former versus the latter suggests that: (i) Z clearly contains important information on the outcomes that embed the endogeneity factor, and (ii) the endogeneity component can reverse the direction of outcomes, enough to counteract the intended effects of monetary policy.

To see this, compare Table 2.2 to Table 2.4, focusing on results that include all controls. In contrast to DCEs, we find that TCEs at several horizons are statistically significant. However, several TCEs depict opposite responses which suggests that DCEs may miss important factors.

There is one occasion where TCEs conform to estimates that are generally regarded as benchmarks in the literature, namely with reference to R. In contrast to DCEs that are insignificant, the sign and statistical significance of TCEs confirm that monetary policy is effective as expected on R, at the 6 months horizon and 1 year into the future. A consensus has in fact emerged regarding the ability of the Federal Reserve to control short term interest rates [see for example Bundick and Smith (2020), Stigum and Crescenzi (2007) and references therein]. The short term interest rates tend to track the federal funds rate, which is a major monetary policy tool in our sample; see Figure 2.1 for descriptive supportive arguments.[6]

In contrast, several TCEs on EBP, P and IP have opposite signs relative to their DCE counterparts (even when the latter are significant), which signals that the endogeneity factor is not to be ignored. Importantly, TCEs on these variables can significantly deviate from theory expectations. Our estimates of TCEs on EBP are especially informative: despite positive DCEs for $h = 0$ and 6, the instantaneous TCE is negative and significant at 1% level which suggests that credit market sentiment may eventually not deteriorate in response to monetary policy tightening. Yet this negative response is short-lived, since the TCEs at further horizons are insignificant. In line with the literature on credit risks [see Gertler and Karadi (2015)], our results suggest that EBP tends to respond swiftly to the monetary policy shock, since it reflects risk appetite in the corporate bond market.

The sign reversal relative to the DCE is consistent with an improvement in financial conditions which itself is not consistent with typical expectations. As illustrated by Gilchrist and Zakrajšek (2012) and the literature thereafter EBP is a forward-looking measure that

[6]Further discussion of the relationship between interest rates and the federal funds rate can be found at https://www.stlouisfed.org/on-the-economy/2017/october/increases-fed-funds-rate-impact-other-interest-rates .

contains predictive information about the real economy. Here we find a significant and positive TCE on IP at the 6 months horizon, which reflects favorable credit market sentiment. This effect permeates into P though at the 24 months horizon: the TCE is significant at 10% in this case which should not be discounted, since estimates at longer horizons are known to be less precise.

Following Gertler and Karadi (2015), we next refine our analysis of credit costs, by expanding Y to include two widely studied credit spreads, namely the above defined M and $COMM$. These two additional spreads[7] incorporate information from three important financial markets: EBP reflects information about long term credit cost to non-farm businesses; in parallel, M is relevant to housing finance, whereas $COMM$ pertains to short term funding to corporations [Gertler and Karadi (2015)]. Table 2.5 presents our six-variables based TCE estimates. We only report results that include all control variables, which in addition to the four-variable based set include 4 lags of M and 4 lags of $COMM$. HAC standard errors are again obtained with 24 lags.

Expanding the system leads to several noteworthy findings. First, estimated TCEs are largely unchanged. This holds in particular for the effects that are statistically significant at 5%. Strikingly, the coefficient estimates are almost the same in these cases. The only difference between Table 2.4 and Table 2.5 is that the TCEs on IP are slightly more persistent once we add further data on spreads. Importantly, the role of EBP is preserved as discussed above.

Secondly, we find that instantaneous TCEs on M and $COMM$ conform to those on EBP. However, the patterns diverge after 6 months: our estimated expansionary TCE on M lasts for about a year, whereas TCEs on $COMM$ reverse to contraction at the 6 months horizon. Such dynamics are consistent with the underlying times to maturity. The commercial paper market is a short term funding market[8], in contrast to mortgage lending; see Figure 2.2 for illustrative evidence. As pointed out above, the Federal Reserve has good control over short term interest rates, but long term interest rates can be affected by a large variety of factors [Stigum and Crescenzi (2007)].

Two further results emerge when comparing DCEs which we report in Table 2.6 to the TCEs, in the expanded system. The above discussed concerns with 2SLS do not improve with the addition of further spread data, which lends further credibility to our weak-instruments arguments. Our discussion thereafter only focuses on AR-based inference.

In this context, most interesting is our finding that DCEs are less robust than TCEs to variable addition. While the latter are largely unchanged, the DCEs on EBP at horizons 0 and 6 change from 0.82% and 1.66% [see Table 2.1] to 0.71% and 0.9% [see Table 2.6]. Moreover, while the DCE on IP is significant at 10% within a one year horizon and a four-variable system, its p-value increases to 42% when the additional spreads are controlled for. This provides direct support for our proposed TCEs, particularly in the presence of factor controls. The latter do not seem to suffice as proxies for missing variables when DCEs are concerned. Instead, we document tangible successes through our proposed endogeneity factor embedded within the TCEs, which we show to be strikingly robust across the two

[7] Spreads refer to the yield difference between certain fixed-income securities and Treasuries with similar maturity.

[8] According to Stigum and Crescenzi (2007), the commercial paper market is a short term funding market in general, with average maturity of 45 days. In our case the maturity is three months.

considered systems.

To summarize, there are three main features to carry away from our reassessment of the estimates from Stock and Watson (2018) that can be relevant more generally. First, our results are consistent with the literature that underscores the importance of the EBP as a financial driver. However, we find that risk appetite in credit markets may not swiftly abate in response to monetary tightening, and resulting expansionary impacts may eventually permeate into output and price down the road. These findings are revealed through our proposed TCEs, which embed the LP-IV coefficients originally defined by Stock and Watson (2018) (denoted in the present paper as the DCEs) *plus* the impact of unobservables in the first-stage regression.

Secondly, we have consistently found that estimates of the DCEs are plagued by weak-instruments which suggest seemingly insignificant responses. While efforts to derive improved instruments are worthy research endeavors, our proposed TCEs that rely on the same HFI approach can inform reliably about consequential policy outcomes. Crucially, in contrast to DCEs, estimates of TCEs are largely unchanged as credit spreads beyond EBP are factored in. This robustness to using a larger system is an important point when complete model assumptions lack support.

Finally, and crucially, our results illustrate how direct impacts can be considerably reversed, delivering sometimes net outcomes that depart from standard theory expectations. Our definition of the TCE is broader than its DCE counterpart. The latter exploits exogenous changes in the policy indicator (here R) alone. Instead, the endogeneity factor which is embedded into the TCE, has the potential to capture, more broadly, other latent changes. These are captured through the disturbance to the first stage regression, which inevitably intervenes once the policy indicator is treated as a proxy that is observed with error and needs to be instrumented.

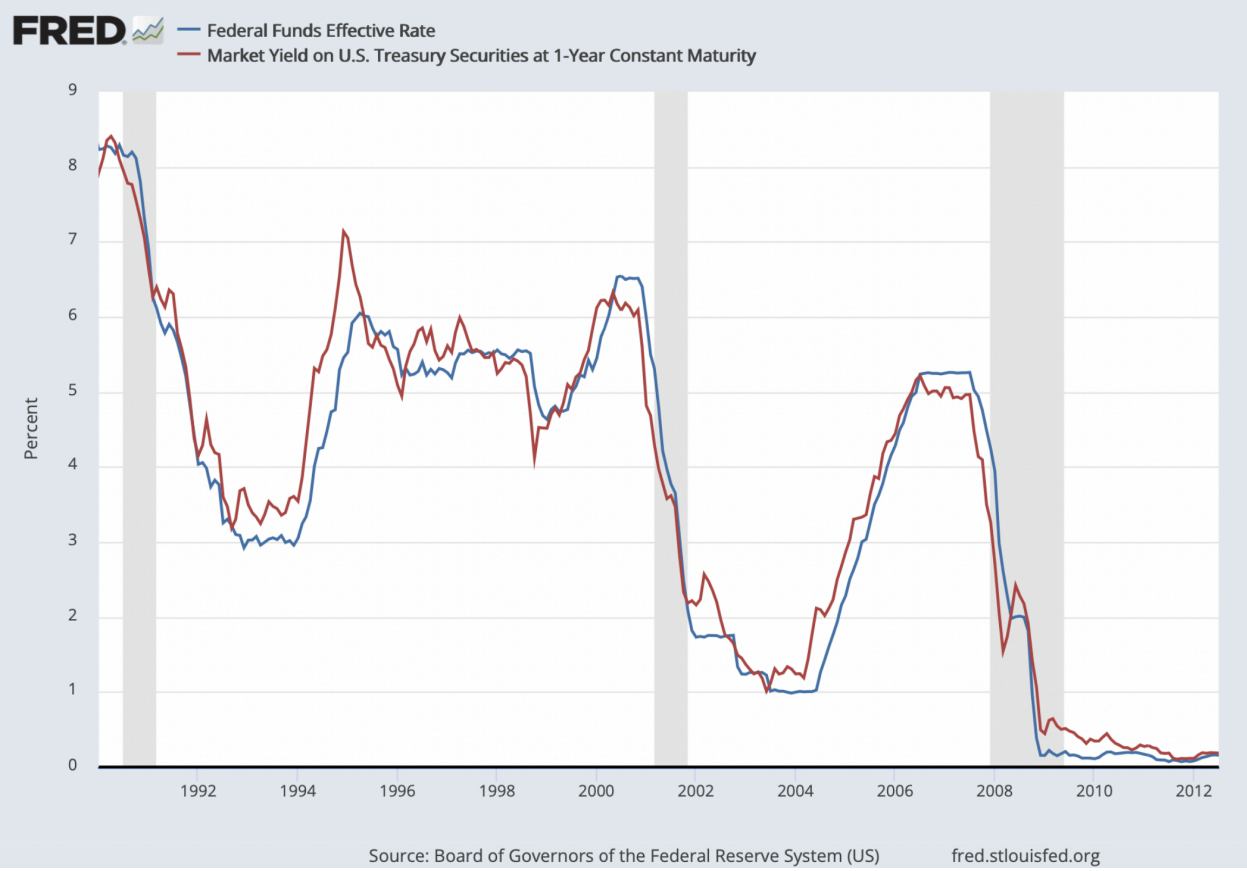
2.4 Conclusion

In this paper, we study the dynamic causal effects of a monetary policy shock on the US economy. With the Local Projection-IV framework, we are able to produce identification-robust results on the traditional direct causal effects of a monetary policy shock. We show that the conventional 2SLS method understates the estimation uncertainty. Moreover, despite the insights we gain from examining the direct effects, the direct effects ignore potentially missing and unobserved information, and such information can have great influence on the results. To provide a more complete picture of the responses of the economy to a monetary policy shock, we formally define and assess the total causal effects of a monetary policy shock. We find the total effect can explain undesired or unintended effects of a monetary policy shock.

Methodologically, our contributions can be summarized as follows. Direct effects will preserve causal inference on dynamic responses with no complete-model assumptions (i.e. no invertibility requirements) provided weak-identification robust methods are applied. This said, robust methods will lack power when instrumental information can be weak. In contrast, the total causal responses that we propose are always identified, retain the partial derivative foundations of LP, account for the same instrumental information as the direct

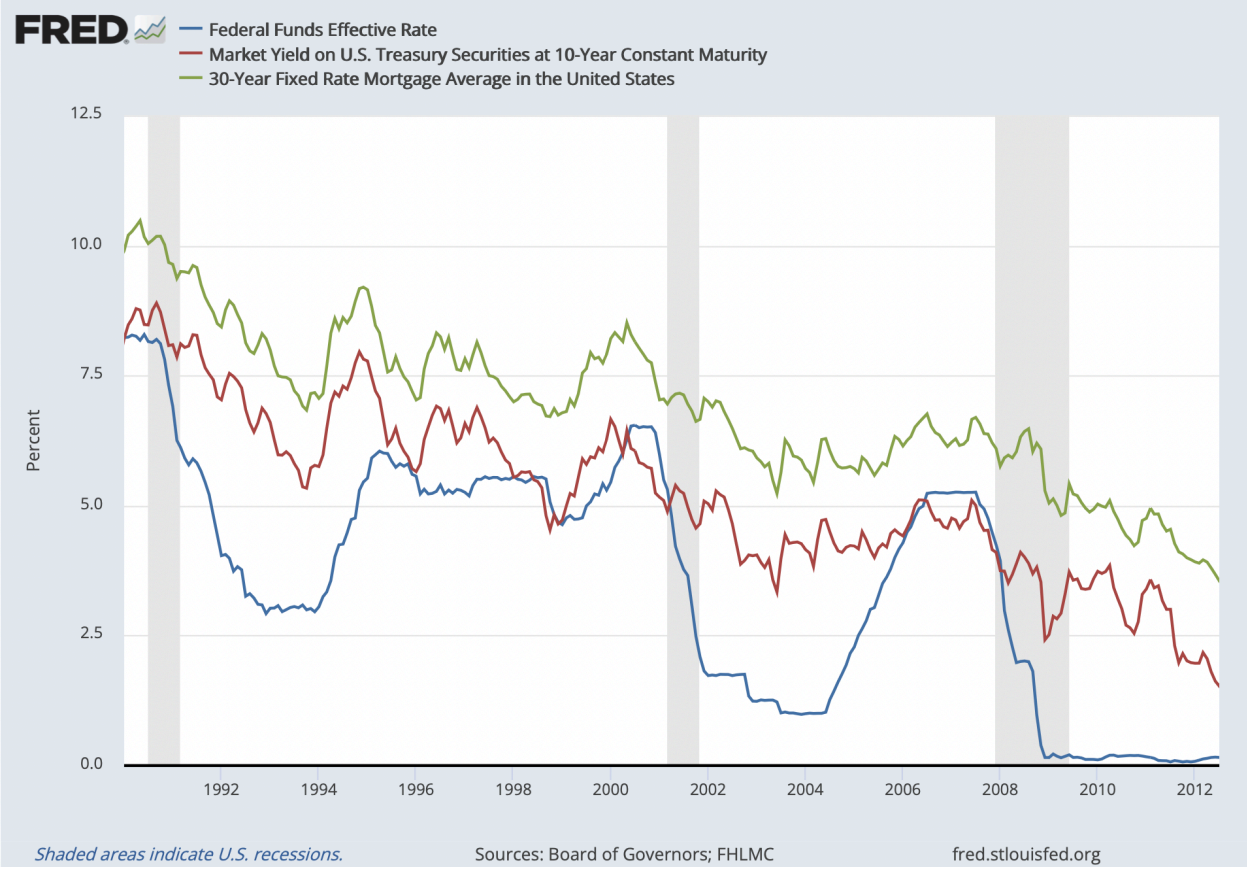
effects and control for unobservables in the first-stage regression. These advantages are concretised empirically, using the 4 or six-variable models of Gertler and Karadi (2015). Whereas 2SLS responses are insignificant at usual levels, AR-based inference on the direct effects produces stronger and economically more plausible results. In contrast to direct effects, our estimates of the total effects following a contractionary monetary shock are compatible with an expansion in economic and long-term credit market activities. Taken collectively, these results confirm that unobservable effects play an important role in shaping the net impact of a policy shock.

Figure 2.1: Correlation between the federal funds rate and one-year Treasury yield



Notes: Time spans between 1991m1 and 2012m6.

Figure 2.2: Correlation between the federal funds rate and long term interest rates



Notes: Time spans between 1991m1 and 2012m6.

Table 2.1: Stock and Watson (2018) Table 1

		LP-IV (2SLS)		
	lag(h)	(1)	(2)	(3)
R	0	1.00(0.00)	1.00(0.00)	1.00(0.00)
		[ ]	[ ]	[ ]
	6	-0.07(1.33)	1.12(0.52)**	0.67(0.57)
		[-2.68, 2.54]	[0.10, 2.14]	[-0.44, 1.79]
	12	-1.05(2.50)	0.78(1.01)	-0.12(1.06)
		[-5.94, 3.85]	[-1.21, 2.77]	[-2.19, 1.95]
	24	-2.09(5.64)	-0.80(1.52)	-1.57(1.47)
		[-13.14, 8.97]	[-3.77, 2.17]	[-4.46, 1.31]
IP	0	-0.59(0.71)	0.21(0.40)	0.03(0.56)
		[-1.98, 0.80]	[-0.57, 0.99]	[-1.06, 1.12]
	6	-2.15(3.41)	-3.80(3.15)	-4.05(3.67)
		[-8.82, 4.52]	[-9.97, 2.36]	[-11.26, 3.15]
	12	-3.60(6.19)	-6.70(4.70)	-6.85(5.49)
		[-15.74, 8.54]	[-15.91, 2.50]	[-17.62, 3.90]
	24	-2.99(10.19)	-9.51(7.68)	-8.13(7.58)
		[-22.96, 16.98]	[-24.56, 5.53]	[-22.99, 6.73]
P	0	0.02(0.07)	-0.08(0.25)	-0.04(0.25)
		[-0.11, 0.15]	[-0.57, 0.41]	[-0.52, 0.44]
	6	0.16(0.42)	-0.39(0.52)	-0.79(0.84)
		[-0.67, 0.98]	[-1.41, 0.63]	[-2.42, 0.85]
	12	-0.26(0.87)	-1.35(1.03)	-1.37(1.23)
		[-1.97, 1.45]	[-3.37, 0.67]	[-3.79, 1.05]
	24	-0.88(3.07)	-2.26(1.31)*	-2.58(1.69)
		[-0.689, 5.13]	[-4.83, 0.31]	[-5.89, 0.73]
EBP	0	0.51(0.61)	0.67(0.40)*	0.82(0.49)*
		[-0.69, 1.70]	[-0.11, 1.46]	[-0.15, 1.79]
	6	0.22(0.30)	1.33(0.81)	1.66(1.04)
		[-0.37, 0.81]	[-0.25, 2.91]	[-0.37, 3.69]
	12	0.56(0.91)	0.84(0.65)	0.91(0.80)
		[-1.22, 2.33]	[-0.43, 2.11]	[-0.65, 2.48]
	24	-0.44(1.29)	0.94(0.65)	0.85(0.76)
		[-0.97, 2.09]	[-0.34, 2.22]	[-0. 63, 2.34]
Controls		None	4 lags of (Z, Y)	4 lags of (Z, Y, F)
First stage f^{Hom}		1.7	23.7	18.6
First stage f^{HAC}		1.1	15.5	12.7

Notes: Authors' calculations. Significance level: * 10%, **5%, ***1% Newey-West standard errors with 24 lags are in parentheses. For the controls, Z is the instrument. Y includes R, IP, P, and EBP. F includes first three principle components of the FRED-MD dataset.

Table 2.2: P-values

		LP-IV (2SLS)			AR		
	lag(h)	(1)	(2)	(3)	(1)	(2)	(3)
R	0	N/A	N/A	N/A	NA	NA	NA
	6	0.956	0.032821	0.239041	0.9529	0.03246	0.29
	12	0.675	0.4422	0.908872	0.3474	0.4983	0.9062
	24	0.712	0.60027	0.28705	0.2606	0.5869	0.1915
IP	0	0.408	0.5932	0.96141	0.00806	0.6	0.9616
	6	0.528	0.22787	0.27113	0.1062	0.1855	0.1713
	12	0.561	0.15493	0.21305	0.09962	0.06374	0.06342
	24	0.769	0.216522	0.284714	0.5999	0.1575	0.1586
P	0	0.765	0.755	0.86463	0.7474	0.7456	0.8626
	6	0.711	0.4531	0.34692	0.7089	0.4532	0.3017
	12	0.765	0.1904	0.2694	0.683	0.1165	0.1789
	24	0.774	0.0858	0.12856	0.5873	0.02565	0.0315
EBP	0	0.406	0.09674	0.0976	0.01563	0.008684	0.005389
	6	0.468	0.101597	0.1095	0.3473	0.007723	0.004448
	12	0.539	0.1959	0.252	0.1827	0.09869	0.136
	24	0.734	0.1498	0.2593	0.3762	0.0549	0.1548
Controls		None	4 lags of (Z, Y)	4 lags of (Z, Y, F)	None	4 lags of (Z, Y)	4 lags of (Z, Y, F)

Notes: Authors' calculations. P-values are calculated with Newey-West HAC with 24 lags. For the controls, Z is the instrument. Y includes R, IP, P, and EBP. F includes first three principle components of the FRED-MD dataset.

Table 2.3: Confidence Sets

		2SLS/Direct effect			AR		
	lag(h)	1	2	3	1	2	3
R	0						
	6	[-2.6938, 2.54661]	[.097557, 2.14445]	[-.444894, 1.79265]	[... , .778201] U [1.71883, ...]	[.114973, 2.70074]	[-.927567, 1.96175]
					[-52.2788, 52.1316](20)	[-.925889, 3.16789]	[-1.56366, 2.91142]
	12	[-5.96076, 3.86692]	[-1.20738, 2.76927]	[-2.19365, 1.9513]	[... , .354384] U [1.49111, ...]	[-2.74988, 2.47271]	[-4.72667, 1.33101]
					[-98.9421, 96.8483](20)	[-3.19571, 4.75759]	[-41.5707, 41.3283](20)
	24	[-13.1834, 9.01149]	[-3.76551, 2.17496]	[-4.45574, 1.31368]	entire grid	[-5.33239, 2.17199]	[-7.86599, .565794]
					[-223.13, 218.958](20)	[-6.73574, 5.1452]	[-59.2653, 56.1232](20)
IP	0	[-1.98485, .807768]	[-.566559, .991848]	[-1.06251, 1.11636]	[... ,-.090019] U [.53383, ...]	[-.777934, 1.13771]	[-1.76808, 1.09782]
					[-28.4111, 27.234](20)	[-1.34576, 1.77105]	[-2.15194, 2.20579]
	6	[-8.84944, 4.54852]	[-9.97323, 2.36349]	[-11.2553, 3.14802]	[... , .12082] U [2.2585, ...]	[-13.9682, 1.86335]	[-19.4806, 1.28093]
					[-135.622, 131.321](20)	[-16.1416, 8.53185]	[-148.087, 139.98](20)
	12	[-15.7906, 8.5837]	[-15.9137, 2.50433]	[-17.6222, 3.90492]	[... , .04215] U [1.98646, ...]	[-24.9753, .319604]	[-33.3636, .252411]
					[-246.399, 239.193](20)	[-25.1228, 11.7134]	[-222.13, 208.413](20)
	24	[-23.0464, 17.0613]	[-24.5584, 5.53118]	[-22.9893, 6.72687]	entire grid	[-36.0491, 3.52822]	[-43.5289, 2.27985]
					[-402.436, 396.451](20)	[-39.6032, 20.576]	[-305.293, 289.031](20)
P	0	[-.114374, .15539]	[-.56896, .412494]	[-.522031, .438397]	entire grid	[-.896603, .364846]	[-.806122, .447531]
					[-2.66712, 2.70813](20)	[-1.05969, .903221]	[-1.00225, .918612]
	6	[-.669705, .980995]	[-1.41496, .630609]	[-2.42475, .849905]	entire grid	[-1.74974, .777013]	[-3.74084, .710593]
					[-16.2887, 16.6](20)	[-2.43774, 1.65339]	[-4.06208, 2.48723]
	12	[-1.98026, 1.45817]	[-3.37418, .666674]	[-3.78703, 1.05266]	entire grid	[-5.0791, .300611]	[-6.35705, .522184]
					[-34.5118, 33.9897](20)	[-5.3946, 2.6871]	[-49.7641, 47.0297](20)
	24	[-6.91654, 5.15606]	[-4.8285, .307105]	[-5.88687, .73285]	entire grid	[-7.27292,-.250666]	[-10.3298,-.257788]
					[-121.114, 119.354](20)	[-7.3963, 2.87491]	[-68.7742, 63.6202](20)
EBP	0	[-.690954, 1.70593]	[-.118209, 1.46489]	[-.146102, 1.7854]	[... , -.42476] U [.101125, ...]	[.150394, 2.40064]	[.181615, 3.15911]
					[-23.3725, 24.3874](20)	[-15.1576, 16.5043](20)	[-18.4954, 20.1347](20)
	6	[-.37544, .816043]	[-.255675, 2.91143]	[-.365541, 3.693]	entire grid	[.281687, 4.84689]	[.404321, 6.82323]
					[-11.6493, 12.0899](20)	[-30.3432, 32.999](20)	[-38.9217, 42.2491](20)
	12	[-1.22925, 2.34585]	[-.429378, 2.11013]	[-.646109, 2.47534]	entire grid	[-.128146, 3.35446]	[-.178985, 4.3829]
					[-35.0539, 36.1705](20)	[-1.69913, 3.37988]	[-30.2999, 32.1291](20)
	24	[-2.97748, 2.09882]	[-.336142, 2.22314]	[-.626326, 2.3359]	entire grid	[-.004383, 3.53096]	[-.242333, 4.14615]
					[-50.9955, 50.1169](20)	[-24.6493, 26.5363]	[-28.7675, 30.4771](20)
Controls		None	4 lags of (Z, Y)	4 lags of (Z, Y, F)	None	4 lags of (Z, Y)	4 lags of (Z, Y, F)

Notes: Authors' calculations. For the controls, Z is the instrument. Y includes R, IP, P, and EBP. F includes first three principle components of the FRED-MD dataset. The default search set is 2 times of the Wald-type confidence interval, which is centered around the IV point estimate. 1000 points are searched in each search set.

Table 2.4: Total Effects

		LP-IV (Total Effects)		
	lag(h)	(1)	(2)	(3)
R	0	1.00(0.00)	1.00(0.00)	1.00(0.00)
		[ ]	[ ]	[ ]
	6	0.91(0.04)***	1.58(0.22)***	1.39(0.14)***
		[0.84, 0.99]	[1.14, 2.02]	[1.11, 1.67]
	12	0.76(0.09)***	2.07(0.35)***	1.60(0.21)***
		[0.58, 0.94]	[1.39, 2.76]	[1.19, 2.01]
	24	0.40(0.20)**	1.08(0.65)*	0.26(0.50)
		[0.01, 0.80]	[-0.20, 2.35]	[-0.72, 1.24]
IP	0	0.04(0.04)	0.20(0.18)	0.12(0.15)
		[-0.03, 0.11]	[-0.16, 0.56]	[-0.18, 0.42]
	6	0.14(0.24)	1.35(0.56)**	1.20(0.46)***
		[-0.33, 0.60]	[0.25, 2.45]	[0.30, 2.11]
	12	0.11(0.38)	2.16(1.17)*	1.66(1.04)
		[-0.64, 0.86]	[-0.14, 4.46]	[-0.37, 3.70]
	24	-0.02(0.51)	3.76(3.25)	2.10(2.49)
		[-1.03, 0.98]	[-2.61, 10.14]	[-2.78, 6.98]
P	0	0.02(0.01)**	-0.07(0.007)	-0.13(0.06)
		[0.00, 0.04]	[-0.20, 0.06]	[-0.13, 0.10]
	6	0.12(0.06)**	-0.11(0.38)	-0.06(0.34)
		[0.00, 0.24]	[-0.85, 0.63]	[-0.73, 0.61]
	12	0.18(0.10)*	0.16(0.23)	0.34(0.22)
		[-0.02, 0.37]	[-0.30, 0.62]	[-0.10, 0.78]
	24	0.17(0.16)	0.37(0.30)	0.39(0.20)*
		[-0.14, 0.47]	[-0.22, 0.95]	[-0.01, 0.79]
EBP	0	-0.06(0.05)	-0.28(0.10)***	-0.31(0.11)***
		[-0.15, 0.04]	[-0.47, -0.09]	[-0.54, -0.09]
	6	-0.004(0.04)	-0.08(0.15)	-0.008(0.13)
		[-0.09, 0.08]	[-0.37, 0.22]	[-0.27, 0.26]
	12	0.04(0.04)	-0.35(0.31)	-0.20(0.27)
		[-0.04, 0.11]	[-0.96, 0.26]	[-0.73, 0.33]
	24	0.10(0.05)*	0.03(0.14)	0.13(0.14)
		[-0.003, 0.20]	[-0.25, 0.31]	[-0.14, 0.41]
Controls		None	4 lags of (Z, Y)	4 lags of (Z, Y, F)

Notes: Authors' calculations. Significance level: * 10%, **5%, ***1% Newey-West standard errors with 24 lags are in parentheses For the controls, Z is the instrument. Y includes R, IP, P, and EBP. F includes first three principle components of the FRED-MD dataset.

Table 2.5: Total Effects with more credit spreads

	LP-IV (Total Effects), with additional credit spreads				
	lag(h)	θ	P-value	95% Confidence Interval	
R	0	1			
	6	1.38	0	1.1174537	2.01982772
	12	1.62	0	1.22500919	2.01982772
	24	0.31	0.41	-0.4280963	1.04092358
IP	0	0.10	0.505	-0.1989794	1.04092358
	6	1.27	0.007	0.35106729	1.04092358
	12	2.07	0.02	0.33566867	1.04092358
	24	2.76	0.11	-0.6066807	6.12649551
P	0	0.01	0.88	-0.0878179	0.10271085
	6	-0.03	0.94	-0.6974398	0.10271085
	12	0.43	0.06	-0.0224272	0.88615585
	24	0.48	0.03	0.05525114	0.88615585
EBP	0	-0.31	0.005	-0.5246085	-0.0942168
	6	-0.08	0.57	-0.3415304	0.18886037
	12	-0.29	0.15	-0.6947978	0.10781089
	24	0.06	0.62	-0.1824136	0.30648459
M	0	-0.20	0	-0.281484	-0.1234958
	6	-0.11	0.03	-0.208356	-0.0080572
	12	-0.28	0.009	-0.4809164	-0.0699751
	24	-0.02	0.84	-0.1677798	0.13602557
COMM	0	-0.15	0.04	-0.2897923	-0.0050712
	6	0.18	0.05	0.00310919	0.36150726
	12	-0.04	0.53	-0.1530382	0.0785341
	24	-0.02	0.86	-0.1754545	0.1461086

Notes: Authors' calculations. The lagged values of M and COMM are also included as controls. P-values are calculated with Newey-West HAC with 24 lags.

Table 2.6: Direct Effects with more credit spreads

		2SLS				AR		
	lag(h)	β	P-value	95% CI		P-value	95% CI	
R	0	1						
	6	0.93	0.17	-0.3885704	2.24448065	0.23	-0.946018	2.512
	12	0.59	0.6	-1.6340715	2.81132819	0.64	-3.99472	2.32407
	24	0.80	0.5	-1.4997645	3.09109829	0.48	-1.63993	4.70181
IP	0	0.08	0.89	-0.9666377	1.11939664	0.89	-1.32266	1.30837
	6	-1.63	0.5	-6.3600304	3.10874014	0.47	-10.2509	2.63957
	12	-2.27	0.5	-8.827749	4.27949822	0.42	-17.625	2.318
	24	1.58	0.71	-6.7499667	9.91284104	0.72	-11.2618	11.0887
P	0	-0.11	0.69	-0.6287648	0.41572914	0.67	-1.09978	0.322153
	6	-1.18	0.21	-3.0359831	0.6729474	0.11	-5.3768	0.19216
	12	-1.84	0.16	-4.4117263	0.7242763	0.03	-8.27015	-0.147148
	24	-2.77	0.13	-6.3082092	0.77796644	0.02	-12.1992	-0.424343
EBP	0	0.71	0.11	-0.153999	1.57760359	0.01	0.139801	2.77447
	6	0.90	0.14	-0.3025573	2.09948422	0.01	0.153085	4.09638
	12	0.32	0.53	-0.6787478	1.31943618	0.47	-0.419724	2.50054
	24	-0.07	0.87	-0.9407989	0.79439745	0.87	-1.27168	1.05581
M	0	0.32	0.18	-0.1425262	0.78299501	0.05	0.014507	1.40418
	6	0.20	0.45	-0.3187846	0.72530477	0.42	-0.308856	1.09162
	12	0.28	0.43	-0.4108768	0.97492409	0.33	-0.203492	1.93278
	24	-0.17	0.52	-0.6947815	0.3528463	0.52	-0.978448	0.426778
COMM	0	0.59	0.07	-0.0509672	1.23832383	0	0.193598	2.18109
	6	0.60	0.11	-0.131734	1.3233055	0.01	0.115142	2.41641
	12	0.36	0.06	-0.0106579	0.73225761	0	0.100519	1.23088
	24	0.23	0.25	-0.1597722	0.61545362	0.15	-0.05928	1.13576

Notes: Authors' calculations. Data from the . The lagged values of M and COMM are also included as controls.P-values are calculated with Newey-West HAC with 24 lags.

Chapter 3

Revisiting the *Flow* View of Quantitative Easing: Evidence from Asset Purchases

3.1 Introduction

Large scale asset purchases or quantitative easing (QE) is among the new tools of monetary policy that central banks can use to achieve their mandates when faced with a near-zero constraint on the short-term interest rate [Bernanke (2021)]. Over the past decade there is a large body of research that seeks to determine the effects of QE, with several insightful and detailed surveys [Krishnamurthy and Vissing-Jorgensen (2011), Joyce et al. (2012), Gagnon and Sack (2018), Rudebusch (2018), Rossi (2020), Bernanke (2021), Bhattarai and Neely (2022)]. However, there remains considerable debate on the effectiveness of QE [Greenlaw et al. (2018)]. One aspect of this debate is whether QE is more effective when implemented through a large change in the stock of financial assets on the balance sheet as in Fed's initial phase of QE1 (2008-2009) or when implemented through flow of purchases as in the subsequent phases of QE2 (2010-2011) and QE3 (2012-2014) [Gagnon and Sack (2018), Bernanke (2021)].

The bulk of the empirical literature has focused on determining the announcement effects of QE without distinguishing between the stock and flow aspects [Gagnon et al. (2011), Bauer and Rudebusch (2014), Weale and Wieladek (2016), Kuttner (2018), Ihrig et al. (2018), Swanson (2021), Jarociński (2021), among others]. The rationale of examining the announcement effects is that the effects could have been fully priced in the markets at the time of the announcements because of rational expectations and the efficiency of financial markets [Rossi (2020)]. There is, however, evidence that actual purchases and not just announcements can affect financial markets. D'Amico and King (2013) estimated large effects of purchases during QE1 ($300 billion of Treasury purchases associated with a 30 basis point reduction in yields). More recently, Vissing-Jorgensen (2021) provides evidence that the effects of actual purchases during QE could even be independent of the announcement effects and possibly even larger. Similarly, Bernardini and Conti (2021), also investigate the purchasing effects in European countries.

The goal of our paper is to investigate flow view of the purchasing effects of QE . In contrast to D'Amico and King (2013) and Vissing-Jorgensen (2021), we do not rely on event study and high-frequency identification. We take advantage of the pre-determined nature of

QE2 and QE3 to identify a central bank purchasing shock with a simple timing restriction, and are better able to capture delayed effects. We use monthly data that covers QE2 and QE3 programs, and Structural VAR (SVAR) and Local Projection (LP) techniques to determine the average and cumulative effects, respectively. Since the estimation of cumulative effects is based on instrumental variable (IV) method we report the identification-robust Anderson-Rubin results.

We find that there are statistically significant purchasing effects on various assets prices as well as on macro aggregates. The stock price index and the 10-year treasury yield increase after a purchasing shock, but there is a decrease in the housing price index. For macro aggregates, there are no statistically significant average effects of the purchasing shock. By contrast, we find significantly positive cumulative effects of purchasing shock on both the industrial production and the consumer price index.

3.2 Data

Our dataset includes a set of 6 variables in monthly frequency. These are Federal Reserve balance sheet total size from the System Open Market Account (SOMA), S&P 500 price index from Yahoo Finance, Consumer Price Index (CPI), Industrial Production, 10-year treasury yield, and the national Freddie Mac House Price Index. We divide all nominal variables in dollars by the Personal Consumption Expenditures (PCE) to convert to real units. The variables are expressed as percentages. The time period for our analysis is between March 2010 and August 2017, which covers the periods after QE1 and before the Federal Reserves' balance sheet normalization.

3.3 Framework and Identification Assumption

In this section, we discuss the econometric methodologies that we apply to study both average and cumulative growth effects of the purchasing shock. In order to have a causal interpretation of the purchasing shock, identifying exogenous shocks is a necessary first step. As illustrated in Kilian and Lütkepohl (2017), researchers can either use creative ways to measure exogenous shocks to the variable of interest directly, or use SVAR to estimate such shocks. We follow the approach of using SVAR to recover shocks to SOMA purchases, or the purchasing shock.

We make two important assumptions similar to Stock and Watson (2018) throughout the current paper. First, we concentrate on linear models, that is we assume that the conditional expectation function is linear, so that conditional expectation can be replaced by the linear projection. Second, we assume homogeneous treatment effect.

Next, we make two assumptions that are similar to Plagborg-Møller and Wolf (2021). First, a Structural Moving Average (SVMA) model, that is

$$Y_t = \Theta(L)\epsilon_t, \tag{3.1}$$

where, $\Theta(L)$ is an $n \times m$ matrix where m should be greater or equal to n, as ϵ_t contains all current and past shocks and measurement errors. Second, the invertibility condition, which

implies that we can recover structural shock $\epsilon_{1,t}$, or the purchasing shock, from current and past Y_t.[1] We are interested in the propagation of the purchasing shock, $\epsilon_{1,t}$, on macro aggregates Y_t.

Since both estimating average effects of the purchasing shock from SVAR and cumulative effects from LP rely on the recursive structure, it is important to consider why such identification assumption might be feasible. The reason is that both QE2 and QE3 purchasing programs were announced in a way such that the monthly purchases are "predetermined". For example, for QE2, it was announced as "The FOMC announces its intent to purchase a further $600 billion of longer-term Treasury securities by the end of the second quarter of 2011, a pace of about $75 billion per month." Therefore, during the ongoing program, the average size of monthly purchases should not be a response to the market condition. This is our key identifying assumption, which allows us to place the purchasing flow in SOMA as the first variable in the VAR.

3.3.1 Structural VAR

To study the average effects of the purchasing shock, we start with a VAR(p) model

$$Y_t = \alpha + \sum_{j=1}^{p} A_j Y_{t-j} + u_t, \tag{3.2}$$

where u_t is the projection residual. Suppose $\sum_u \equiv E[u_t u_t']$, and $\sum_u = BB'$, where B is lower triangular matrix obtained from Cholesky decomposition, then we can have the recursive SVAR as

$$A(L)Y_t = \alpha + B\epsilon_t. \tag{3.3}$$

Let $C(L) = A(L)^{-1}$, then $\Theta(L) = C(L)B$ and $u_t = B\epsilon_t$, where $n = m$. With timing restriction, that is $Y_{1,t}$, the growth rate of SOMA, is predetermined to all other variables in the SVAR, we are able to estimate the purchasing shock and corresponding impulse responses. The recursive SVAR gives the estimated shocks a causal interpretation.

We focus on monthly bivariate models. There are two reasons to do so. First, as we are using the semi-structural model, in which only the shock to SOMA purchases, $\epsilon_{1,t}$, is both statistically and economically identified, and we are only interested in other variables' responses to such shock. Given the pre-determinedness of $Y_{1,t}$ with respect to other macro aggregates, we are able to order the Federal Reserve balance sheet variable first in SVAR models. As illustrated in Edelstein and Kilian (2009), there is no loss of generality by using a bivariate model as adding more variables into the bivariate model does not change the response of the existed variable to $\epsilon_{1,t}$, asymptotically [Kilian (2011)]. Second, since we restrict observations to only the time periods when the pre-determinedness assumption makes sense, then we are left with 91 observations. It can be difficult to fit all variables together with enough lags. The bivariate models allow us to focus on responses of each variable at a time, importantly, with enough lags considered. As illustrated in Kilian and Lütkepohl (2017), having large number of lags can help capture the delays in the responses

[1] We are using a semi-structural model, because only $\epsilon_{1,t}$ is both statistically and economically identified. More details are provided in Section 3.4

to the shock of interest. This is particularly important for macro aggregates.[2]

The bivariate recursive structure is

$$\begin{bmatrix} u_{1,t} \\ u_{i,t} \end{bmatrix} = \begin{bmatrix} b_{11}^0 & 0 \\ b_{i1}^0 & b_{ii}^0 \end{bmatrix} \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{i,t} \end{bmatrix}, \tag{3.4}$$

where the $\epsilon_{1,t}$ is identified as the purchasing shock. i is used to represent one of the five variables, that are real stock price growth rate, real housing price growth rate, industrial production index growth rate, CPI growth rate, and the 10-year treasury yield. It is important to note that $\epsilon_{i,t}$ does not have economical interpretation.

3.3.2 Local Projection Method

In addition to the average effects of the purchasing shocks, we are also interested in estimating the cumulative growth effect of the shocks, that is the cumulative growth effects of the variable of interest from accumulation of assets on Federal Reserve balance sheet. To study such cumulative growth effects, we rely on the LP method [Jordà (2005)].

Again, we start from (3.1), and single out the first shock

$$Y_t = \Theta_1(L)\epsilon_{1,t} + \{\epsilon_{2:n,t}, \epsilon_{t-1}, ...\}, \tag{3.5}$$

and the regressions can be run separately for each horizon h as

$$Y_{i,t+h} = \Theta_{h,i1}(L)\epsilon_{1,t} + \delta(L)Y_{t-1} + \{\epsilon_{t+h}, ..., \epsilon_{t+1}, \epsilon_{2:n,t}\}, \tag{3.6}$$

where lagged Y's are added to get smaller standard errors. $\{.\}$ is used to represent a linear combination. (3.4) shows the timing restriction that we want to maintain. By imposing the unit-effect normalization [Stock and Watson (2018)], (3.4) becomes

$$\begin{bmatrix} u_{1,t} \\ u_{i,t} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ b_{i1}^0 & 1 \end{bmatrix} \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{i,t} \end{bmatrix},$$

which gives that $u_{1,t} = \epsilon_{1,t}$. Then, (3.6) is equivalent to

$$Y_{i,t+h} = \Theta_{h,i1}(L)u_{1,t} + \delta(L)Y_{t-1} + \{\epsilon_{t+h}, ..., \epsilon_{t+1}, \epsilon_{2:n,t}\},$$

and by the Frisch-Waugh-Lovell Theorem, we can have the LP as

$$Y_{i,t+h} = \Theta_{h,i1}(L)Y_{1,t} + \delta(L)Y_{t-1} + e_{i,t+h}^h, \tag{3.7}$$

where $e_{i,t+h}^h = \{\epsilon_{t+h}, ..., \epsilon_{t+1}, \epsilon_{2:n,t}\}$. Here, LP method has the flexibility to add additional

[2] We impose a VAR lag order of 6 for all the bivariate models. This is because the Akaike Information Criterion suggests 4 lags for all the models and we choose the number of lags slightly larger than recommended lags as the Akaike Information Criterion tends to be the minimum as suggested by Kilian and Lütkepohl (2017). Our results are not sensitive to lag order choice.

lagged variables as controls, so we are not restricted to the bivariate case.[3]

Based on Ramey and Zubairy (2018) and Fieldhouse et al. (2018), the cumulative growth effect on Y_i relative to cumulation of Y_1 during a period of h, can be estimated using

$$\sum_{j=0}^{h} Y_{i,t+j} = \lambda_h + \rho_h \sum_{j=0}^{h} Y_{1,t+j} + \delta(L)Y_{t-1} + \omega_h, \tag{3.8}$$

with $Y_{1,t}$, which is the current period growth rate of SOMA, used as the instrument for $\sum_{j=0}^{h} Y_{1,t+j}$.[4] We choose to include 4 lags of each control variable. The control variables are lagged values of the real growth rate of SOMA, the real growth rate of stock price, the real growth rate of housing price, the growth rate of industrial production, the growth rate of CPI, and the 10-year treasury yield. Newey-West standard errors are generated with 12 lags.

3.4 Empirical Results

3.4.1 SVAR with average effects

We estimate the responses to a standard deviation increase in the SOMA. Figure 3.3 shows the estimated structural impulse response functions from the bivariate models (3.4). In each case, the estimated impulse responses are reported together with 95 percent confidence intervals, constructed using the standard boostrapping method. Given the timing restriction and assumptions illustrated above, these impulse responses are interpreted in the sense of dynamic causal effects from Stock and Watson (2018), each reflects the average effects of a purchasing shock over the sample.

The shock increases the stock price index by about 0.9 percent, and the rise in the index is immediate. The effect peaks on impact and fades away quickly within a month, suggesting that while the purchasing shock has strong impact on the stock price index, the effects are short-lived. Additionally, we find the purchasing shock has significant effects on the 10-year treasury yield. Figure 3.3 shows that the 10-year treasury yield rises by roughly 7 basis points on impact, and the effects stay around 10 basis point and statistically significant for approximately 3 months. This finding is similar to Hamilton (2018) and Greenlaw et al. (2018), where they find that the QE announcements only lowers the 10-year treasury yield initially (in the Fall of 2008 and 18 March 2009), and the overall impact of the announcements is to raise, not lower, the 10-year treasury yield. This will be further discussed in the next subsection when we talk about the cumulative growth effects.

There are no statistically significant effects of the purchasing shock on the housing price index on impact. We find the housing price index falls by roughly 5 basis points at 13 months into the future. This is likely because the 10-year treasury yield serves as the benchmark rate for the mortgage rates in the U.S. Then, rising 10-year treasury yield leads to rising mortgage rates, which puts downward pressure onto the housing market. For macro

[3] According to Plagborg-Møller and Wolf (2021), adding more controls in LP is equivalent to expand the VAR system in our case.

[4] For stock variables, the cumulative amount should be calculated as $Y_{t+h} - Y_{t-1}$, instead of $\sum_{j=0}^{h} Y_{t+j}$

aggregate variables, there is no evidence shows that the purchasing shock has statistically significant effects on either industrial production and headline inflation.

3.4.2 LP with cumulative growth effects

In addition to the average effects of the purchasing shock, it is also important to know the effects of the accumulation of the purchases. Such cumulative growth effect is captured by ρ_h in (3.8). According to Ramey and Zubairy (2018), ρ_h is interpreted as the cumulative growth gain of Y_i relative to the cumulative growth of the purchases during a giving period.

Tables 3.1 - 3.5 show the estimated cumulative growth effects of the purchases on different variables for a number of periods into the future. The use of a instrumental variable in (3.8) emphasizes the importance of instrument relevance. As suggested by Andrews et al. (2019), in addition to the 2SLS results, we also report identification-robust Anderson-Rubin (AR) confidence sets, since we only have a single instrument. The grid search set is 2 times of the wald-type confidence intervals and 1000 points are estimated in the search set [Finlay et al. (2016)]. Overall, since all AR confidence sets are closed, and the confidence sets are similar to the Wald-type intervals, then we should be safe from the weak instrument problem. In addition, there is evidence that suggests that the confidence intervals generated by the 2SLS method have wrong centers. For example, for the cumulative growth effects on CPI at 12 months into the future, the p-value produced by the 2SLS is 0.544, while the P-value provided by the AR method is 0.035. Moreover, all of the IV point estimates are covered our confidence sets, and the IV point estimates can be interpreted as the least rejected values [Dufour et al. (2006), Dufour et al. (2010), Dufour et al. (2013)]. Therefore, we focus on the P-values generated by the AR method hereafter.

Table 3.1 shows the estimated cumulative stock price index growth relative to the cumulative SOMA growth for a given period. Similar to the findings of the average effects, the accumulation of the purchases raise the stock price index, and the effects are concentrated within half of a year. On impact, a one percent expansion of the SOMA causes roughly 1.04% appreciation of the stock price index. Such effect quickly dies out to only 0.56% after three months.

For macro aggregates, we find different cumulative growth effects comparing to the average effects. Table 3.2 shows that the cumulative growth of the SOMA raises the industrial production index after 9 months. For 1% increase in the SOMA, the industrial production index grows by 5 basis points for a 9 months period, and almost 10 basis points for a 12 months period. Similar cumulative growth effects are also found for the CPI as in Table 3.3, every 1% increase of the SOMA induces roughly 3 basis points increase for CPI for both 9 months and 1 year periods.

Table 3.4 shows that the accumulation of SOMA holdings raise 10-year treasury yields. Similar to the average effects, the response of the cumulative growth of the 10-year treasury yield is statistically significant, and the cumulative effects diminish with longer the time periods. For a 1% increase in the SOMA, the within month cumulative growth rate of the 10-year treasury yield is about 14 basis points. The magnitude of the response quickly retreats to only about 4 basis over a 3 months period and only 1 basis points for a 1-year period. Such effects on 10-year treasury yield are not new, similar cumulative effects from LSAP announcements to the 10-year treasury yield have also been by Greenlaw et

al. (2018), in which the authors use couple different approaches to isolate Federal Reserve QE news shock and show that the news shock only decreases 10-year treasury yield for a short initial period of QE1, but has positive cumulative overall impact on QE1, QE2, and QE3 programs. Since our sample only includes QE2 and QE3 programs, then the positive effects we find are not surprising. In addition, the corner stone of our identification is the predeterminedness assumption, not high frequency identification that is commonly used in the announcement effect literature, we are free from the misspecification concern raised in this article.[5]

The housing price index falls for the periods of 9 months and 1 year. As shown in Table 3.5, there is no statistical significance found for periods shorter than 9 months. For each percentage increase of the SOMA, the housing price index falls by 9 basis points for the 9 months period and almost 19 basis points for the 1 year period. Such effects can be linked to the increases in the 10-year treasury yield that we observe in Table 3.4. We present two pieces of evidence. First, we conduct a simple cross correlation exercise on raw series of 10-year treasury yield and real growth rate of housing price index. Figure 3.4 shows the most dominant cross correlations occur at $h = -3$, this is a strong evidence that 10-year treasury yield leads housing price index. Second, Figure 3.5 shows that 30-year fixed mortgage rate tracks 10-year treasury yield closely. Therefore, an increase in the 10-year treasury yield can lead to a housing demand shock as mortgage rate goes up. Jarociński and Smets (2008) among others provide evidence that housing demand shock has significant effects on residential investment and housing price. Such relationship between 10-year treasury yield and housing price is also noticed by investors.[6]

In summary, there are three main findings. First, our results show consistent effects of the purchasing activities from the Federal Reserve on financial variables, both on average and cumulatively. Second, we find that the cumulation of the purchases has statistically significant effects on macro aggregates, and such significance cannot be found with the average effects. Lastly, we find little evidence of the weak instrument problem, nevertheless, the AR method still recovers more statistical significance compared to the results generated by the 2SLS method.

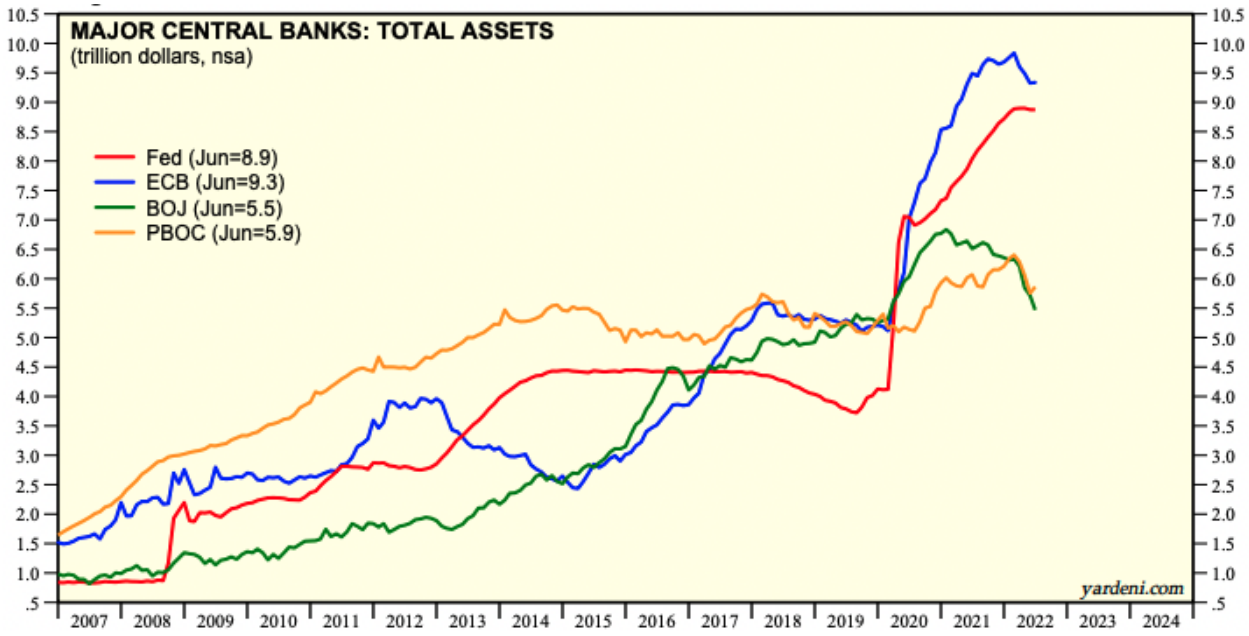
3.5 Conclusion

We examine the *flow* view of quantitative easing (QE) using monthly data on Federal Reserve's pre-announced asset purchases from the second and third rounds of QE. We determine both average and cumulative purchasing effects using structural VAR and identification-robust local projection methods, respectively. For financial assets, we find that the purchasing activity increases the stock price index and the 10-year treasury yield, but it decreases the housing price index. For macro aggregates, there is no statistically significant average effects of the purchases, however, we find that the accumulation of the purchases does increase both the industrial production and the consumer price index.

[5]https://voxeu.org/article/effectiveness-large-scale-asset-purchases

[6]For example, a recent article https://fortune.com/2022/06/11/housing-market-is-once-again-watching-mortgage-rates-spike/

Figure 3.1: Total assets of major central banks



Source: Figure 1 in Yardeni and Quintana (2022). Haver Analytics

Figure 3.2: Balance sheet size and stock market index

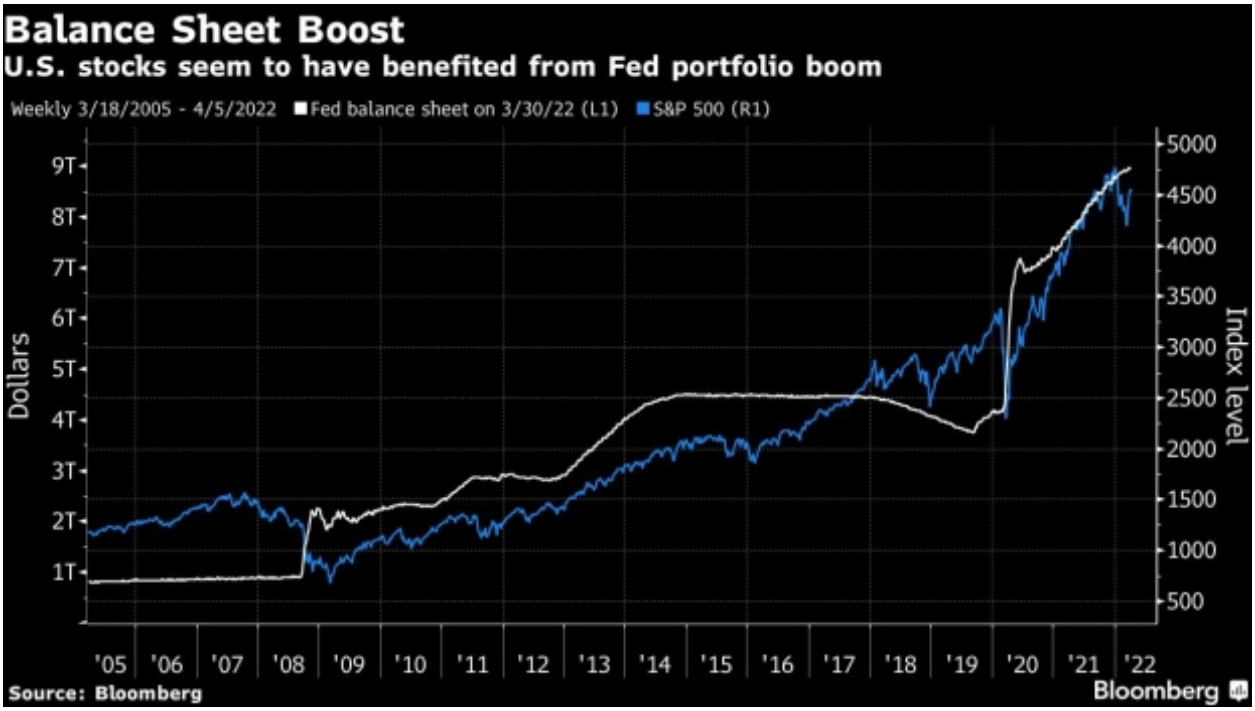
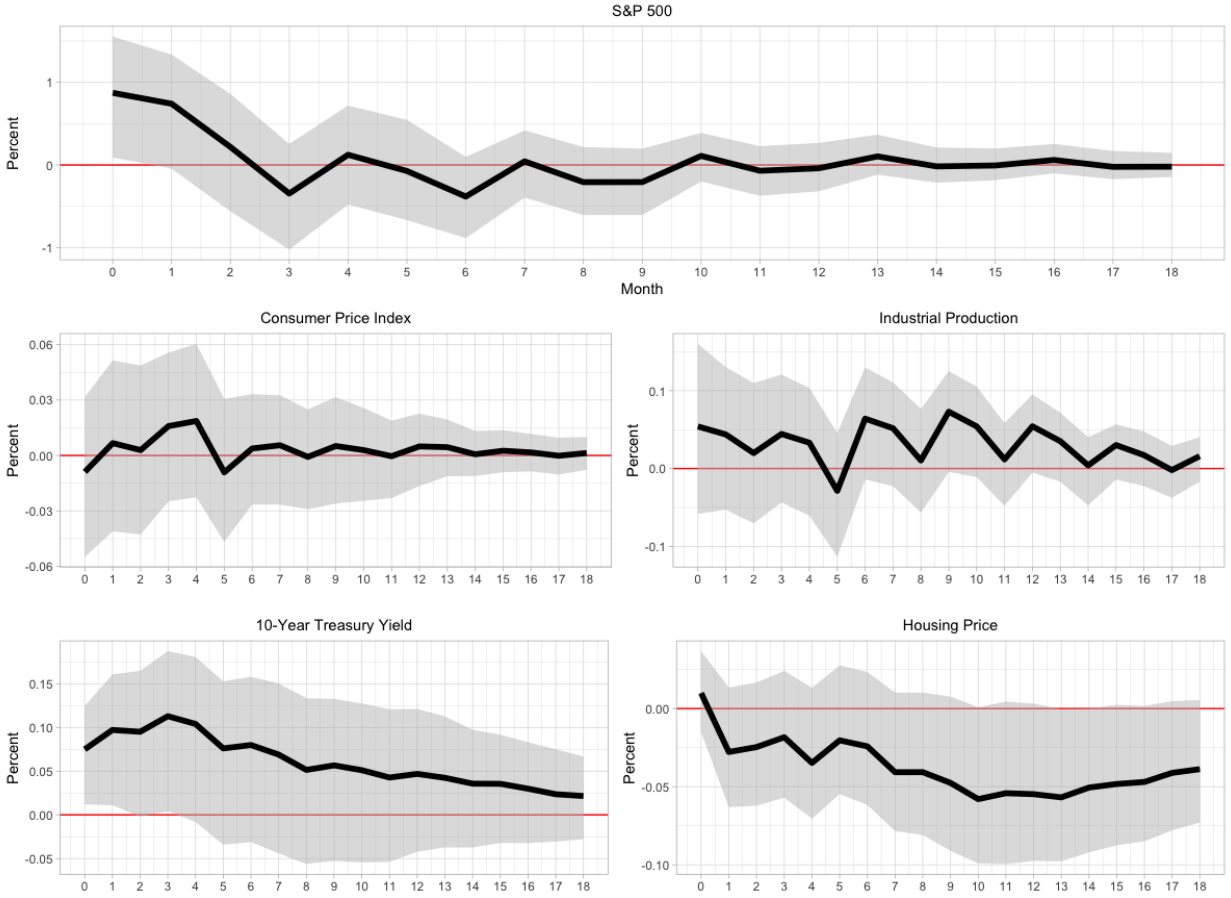
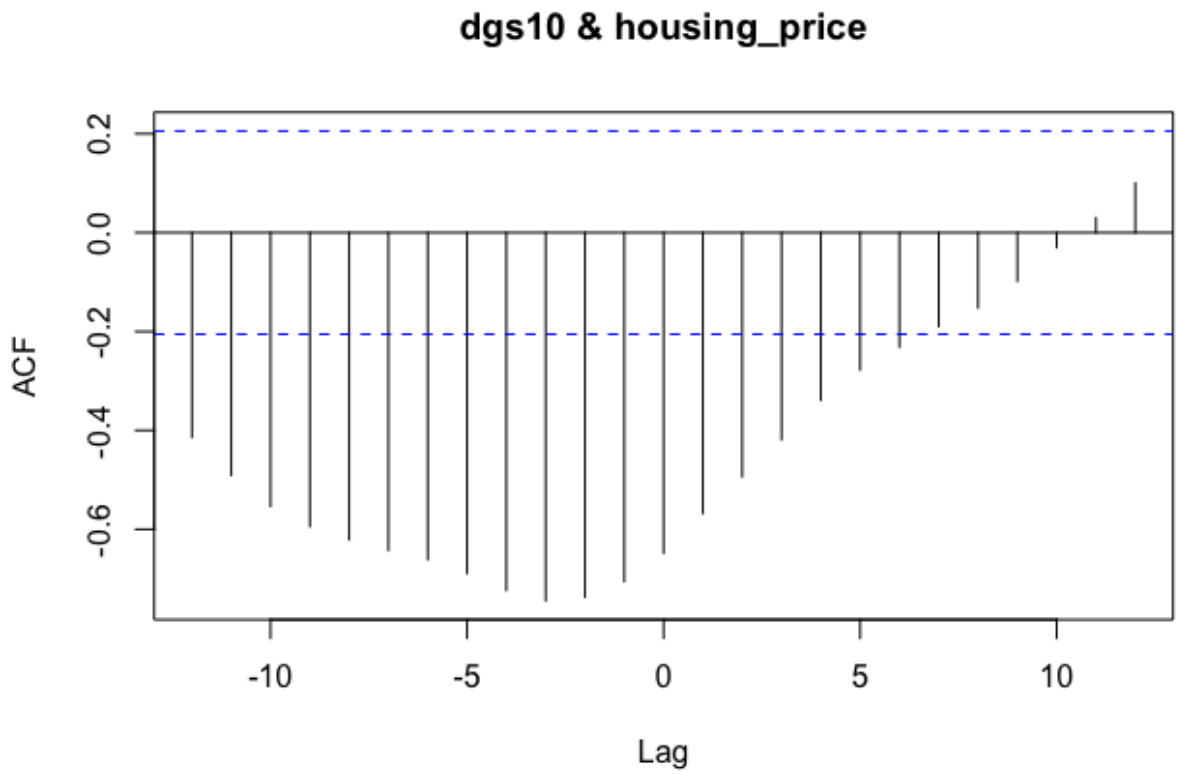


Figure 3.3: Impulse response functions from bivariate VARs



Notes: We impose a VAR lag order of 6 for all the bivariate models. In each case, the estimated impulse responses are reported together with 95 percent confidence intervals, constructed using the standard boostrapping method.

Figure 3.4: Cross correlation



Notes: dgs10 represents the 10-year treasury yield; housing _ price represents the housing price index.

Figure 3.5: 10-year treasury yield and mortgage rate

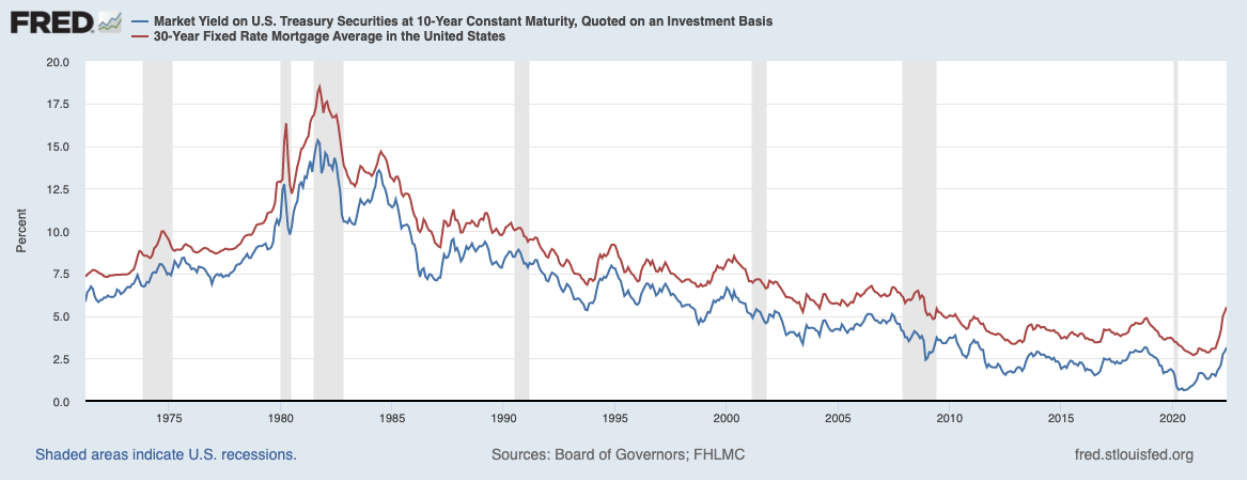


Table 3.1: Cumulative growth effects on stock price

Stock Price							
	2SLS					AR	
lag	point estimate	std err	P-value	95% ci		P-value	AR 95% CS
0	1.0428	0.42536	0.017102	0.2091117	1.87647975		
3	0.55518	0.11228	6.86E-06	0.3351187	0.775243	0	[0.342388, 0.790002]
6	0.38901	0.07859	7.40E-06	0.2349745	0.54304427	0	[0.230195, 0.549674]
9	0.14682	0.09667	0.134902	-0.0426615	0.33629491	0.2022	[-0.109994, 0.308034]
12	0.15034	0.07757	0.058391	-0.0017006	0.30238518	0.1112	[-0.047816, 0.290057]

Notes: The estimated cumulative growth effects during different periods are the point estimates. Newey-West standard errors with 12 lags.

Table 3.2: Cumulative growth effects on industrial production

Industrial Production							
	2SLS					AR	
lag	point estimate	std err	P-value	95% ci		P-value	AR 95% CS
0	0.046404	0.046404	0.560975	-0.1091738	0.20198198		
3	0.027918	0.020513	0.178781	-0.0122866	0.06812256	0.2109	[-0.018203, 0.064058]
6	0.02428	0.01712	0.161821	-0.0092783	0.05783464	0.1948	[-0.014754, 0.055383]
9	0.054399	0.020707	0.011292	0.01381307	0.09498462	0.0413	[0.002966, 0.091369]
12	0.09734	0.03535	0.008249	0.02805232	0.16662936	0.0197	[0.019521, 0.171276]

Notes: The estimated cumulative growth effects during different periods are the point estimates. Newey-West standard errors with 12 lags.

Table 3.3: Cumulative growth effects on consumer price index

CPI							
	2SLS					AR	
lag	point estimate	std err	P-value	95% ci		P-value	AR 95% CS
0	-0.030809	0.037156	0.41023	-0.1036329	0.04201468		
3	0.019096	0.025833	0.46276	-0.0315362	0.06972902	0.4459	[-0.028039, 0.075152]
6	0.0241316	0.0185872	0.199605	-0.0122986	0.06056183	0.1932	[-0.012554, 0.063005]
9	0.036884	0.014259	0.012526	0.00893721	0.06483079	0.006	[0.010867, 0.071517]
12	0.03405	0.01727	0.054364	0.00019374	0.06789648	0.0353	[0.002396, 0.077486]

Notes: The estimated cumulative growth effects during different periods are the point estimates. Newey-West standard errors with 12 lags.

Table 3.4: Cumulative growth effects on 10-year yield

10-Year Treasury Yield							
	2SLS					AR	
lag	point estimate	std err	P-value	95% ci		P-value	AR 95% CS
0	0.1405345	0.0494674	0.00611	0.0435802	0.23748873		
3	0.0451198	0.0191352	0.02177	0.00761553	0.08262409	0.0099	[0.010356, 0.086791]
6	0.0378	0.010379	0.0006	0.01745653	0.05814312	0.0003	[0.018129, 0.060485]
9	0.0235824	0.0087681	0.009594	0.00639722	0.04076763	0.0197	[0.004625, 0.041576]
12	0.01221	0.004294	0.00649	0.00379383	0.02062654	0.0258	[0.002016, 0.020517]

Notes: The estimated cumulative growth effects during different periods are the point estimates. Newey-West standard errors with 12 lags.

Table 3.5: Cumulative growth effects on housing price

Housing Price							
	2SLS					AR	
lag	point estimate	std err	P-value	95% ci		P-value	AR 95% CS
0	0.0151707	0.0226187	0.50493	-0.0291611	0.05950252		
3	-0.030609	0.034438	0.37777	-0.0981072	0.03688831	0.3506	[-0.107229, 0.030605]
6	-0.054601	0.038422	0.16094	-0.1299073	0.02070578	0.1233	[-0.144004, 0.013695]
9	-0.09189	0.04587	0.05038	-0.1817921	-0.0019851	0.0131	[-0.21662,-0.017194]
12	-0.18647	0.06604	0.00685	-0.3159058	-0.0570311	0.0001	[-0.38056,-0.083074]

Notes: The estimated cumulative growth effects during different periods are the point estimates. Newey-West standard errors with 12 lags.

Bibliography

Abadie, Alberto (2005) 'Semiparametric Difference-in-Differences Estimators.' *Review of Economics Studies* pp. 1–19

Amir-Ahmadi, Pooyan, Christian Matthes, and Mu-Chun Wang (2022) 'Understanding Instruments in Macroeconomics - A Study of High-Frequency Identification.' *Working paper*

Anderson, Theodore W., and Herman Rubin (1949) 'Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations.' *The Annals of Mathematical Statistics* pp. 46–63

Andrews, Isaiah, James H. Stock, and Liyang Sun (2019) 'Weak Instruments in Instrumental Variables Regression: Theory and Practice.' *Annual Review of Economics* pp. 727–753

Angrist, Joshua, and Jörn-Steffen Pischke (2009) *Mostly harmless econometrics: An empiricist's companion*

Ashenfelter, Orley (1978) 'Estimating the Effect of Training Programs on Earnings.' *Review of Economics and Statistics* pp. 47–57

Athey, Susan, and Guido W. Imbens (2018) 'Design-Based Analysis in Difference-in-Differences Settings with Staggered Adoption.' *Working paper*

Bai, Jennie, Arvind Krishnamurthy, and Charles-Henri Weymuller (2018) 'Measuring Liquidity Mismatch in the Banking Sector.' *Journal of Finance* pp. 51–93

Baker, Andrew C., David F. Larcker, and Charles C. Y. Wang (2021) 'How Much Should We Trust Staggered Difference-In-Differences Estimates?' *Working paper*

Banerjee, Ryan N., and Hitoshi Mio (2018) 'The Impact of Liquidity Regulation on Banks.' *Journal of Financial Intermediation* pp. 30–44

Bauer, Michael D., and Eric T. Swanson (2022) 'A Reassessment of Monetary Policy Surprises and High-Frequency Identification.' *IMFS Working Paper Series 165, Goethe University Frankfurt, Institute for Monetary and Financial Stability (IMFS)*

Bauer, Michael D., and Glenn D. Rudebusch (2014) 'The Signalling Channel for Federal Reserve Bond Purchases.' *International Journal of Central Banking* pp. 233–289

Beaulieu, Marie-Claude, Jean-Marie Dufour, Lynda Khalaf, and Olena Melin (2022a) 'Identification-Robust Benchmark Neutrality Tests: the Case of Catastrophe Bond Mutual Funds.' *Working paper*

Beaulieu, Marie-Claude, Lynda Khalaf, Maral Kichian, and Olena Melin (2022b) 'Finite Sample Inference in Multivariate Instrumental Regressions with an Application to Catastrophe Bonds.' *Econometric Reviews*

Bernanke, Ben S. (2021) 'The New Tools of Monetary Policy.' *American Economic Review* pp. 943–83

Bernardini, Marco, and Antonio M. Conti (2021) 'Assessing the Flexible Implementation of the ECB's Pandemic Assets Purchases.' *Banca D'Italia Covid-19 Note*

Bhattarai, Saroj, and Christopher J. Neely (2022) 'An Analysis of the Literature on International Unconventional Monetary Policy.' *Journal of Economic Literature* pp. 527–597

BIS (2010a) 'An Assessment of the Long-term Economic Impact of Stronger Capital and Liquidity Requirements .' *Bank for International Settlement*

_ (2010b) 'Assessing the Macroeconomic Impact of the Transition to Stronger Capital and Liquidity Requirements.' *Bank for International Settlement*

_ (2010c) 'Results of the Comprehensive Quantitative Impact Study .' *Bank for International Settlement*

Bolduc, Denis, Lynda Khalaf, and Clément Yélou (2010) 'Identification Robust Confidence Sets Methods for Inference on Parameter Ratios and their Application to Estimating Value-of-Time.' *Journal of Econometrics* pp. 317–327

Borusyak, Kirill, and Xavier Jaravel (2017) 'Revisiting Event Study Designs.' *Working paper*

Bruns, Martin, and Helmut Lütkepohl (2022) 'Comparison of local projection estimators for proxy vector autoregressions.' *Journal of Economic Dynamics and Control* p. 104277

Bundick, Brent, and A. Lee Smith (2020) 'Policymakers Have Options for Additional Accommodation: Forward Guidance and Yield Curve Control.' *Federal Reserve Bank of Kansas City Economic Bulletin*

Caldara, Dario, and Kamps Christophe (2017) 'The Analytics of SVARs: A Unified Framework to Measure Fiscal Multipliers.' *The Review of economics studies* pp. 195–218

Callaway, Brantly, and Pedro H. C. Sant'Anna (2021) 'Difference-in-Differences with Multiple Time Periods.' *Journal of Econometrics* pp. 200–230

Calomiris, Charles W., Florian Heider, and Marie Hoerova (2015) 'A Theory of Bank Liquidity Requirements.' *Columbia Business School Research Paper No. 14-39*

Consolo, Agostino, Carlo A. Favero, and Alessia Paccagnini (2009) 'On the Statistical Identification of DSGE Models.' *Journal of Econometrics* pp. 99–115

Dake, Li, Mikkel Plagborg-Møller, and Christian K. Wolf (2022) 'Local Projections vs. VARs: Lessons From Thousands of DGPs.' *Working paper*

D'Amico, Stefania, and Thomas B. King (2013) 'Flow and Stock Effects of Large-Scale Treasury Purchases.' *Journal of Financial Economics* pp. 425–448

de Chaisemartin, Clément, and Xavier D'Haultfœuille (2020) 'Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects.' *American Economic Review* pp. 2964–2996

Diamond, Douglas W., and Anil K. Kashyap (2016) 'Liquidity Requirements, Liquidity Choice, and Financial Stability.' *Handbook of Macroeconomics* pp. 2263–2303

Doko Tchatoka, Firmin, and Jean-Marie Dufour (2014) 'Identification-Robust Inference for Endogeneity Parameters in Linear Structural Models.' *Econometrics Journal* pp. 165–187

_ (2020) 'Exogeneity Tests, Incomplete Models, Weak Identification and Non-Gaussian Distributions: Invariance and Finite-Sample Distributional Theory.' *Journal of Econometrics* pp. 390–418

Dube, Arindrajit, Daniele Girardi, Òscar Jordà, and Alan M. Taylor (2022) 'A Local Projections Approach to Difference-in-Differences Event Studies.' *Working paper*

Dufour, Jean-Marie (1997) 'Some Impossibility Theorems in Econometrics with Applications to Structural and Dynamic Models.' *Econometrica* pp. 1365–1387

_ (2003) 'Identification, Weak Instruments, and Statistical Inference in Econometrics.' *Canadian Journal of Economics* pp. 767–808

Dufour, Jean-Marie, and Mohamed Taamouti (2005) 'Projection-Based Statistical Inference in Linear Structural Models with Possibly Weak Instruments.' *Econometrica* pp. 1351–1365

Dufour, Jean-Marie, Lynda Khalaf, and Maral Kichian (2006) 'Inflation Dynamics and the New Keynesian Phillips Curve: An Identification-Robust Econometric Analysis.' *Journal of Economic Dynamic and Control* pp. 1707–1728

_ (2010) 'On the Precision of Calvo Parameter Estimates in Structural NKPC Models.' *Journal of Economic Dynamic and Control* pp. 1582–1595

_ (2013) 'Identification-Robust Analysis of DSGE and Structural Macroeconomic Models.' *Journal of Monetary Economics* pp. 340–350

Edelstein, Paul, and Lutz Kilian (2009) 'How Sensitive Are Consumer Expenditures to Retail Energy Prices.' *Journal of Monetary Economics* pp. 766–779

Fieldhouse, Andrew, Karel Mertens, and Morten O. Ravn (2018) 'The Macroeconomic Effects of Government Asset Purchases: Evidence from Postwar US Housing Credit Policy.' *The Quarterly Journal of Economics* pp. 1503–1560

Finlay, Keith, Leandro Magnusson, and Mark Schaffer (2016) 'WEAKIV: Stata Module to Perform Weak-Instrument-Robust Tests and Confidence Intervals for Instrumental-Variable (IV) Estimation of Linear, Probit and Tobit Models.' *Statistical Software Components*

Gagnon, Joseph E., and Brian Sack (2018) 'QE: A User's Guide.' Technical Report 18-19, Peterson Institute For International Economics

Gagnon, Joseph, Matthew Raskin, Julie Remache, and Brian Sack (2011) 'The Financial Market Effects of the Federal Reserve's Large-Scale Asset Purchases.' *International Journal of Central Banking* pp. 3–43

Ganics, Gergely, Atsushi Inoue, and Barbara Ross (2021) 'Confidence Intervals for Bias and Size Distortions in IV and Local Projections-IV Models.' *Journal of Business and Economic Statistics* pp. 307–324

Gertler, Mark, and Peter Karadi (2015) 'Monetary Policy Surprises, Credit Costs, and Economic Activity.' *American Economics Journal: Macroeconomcis* pp. 44–76

Giirkaynak, Refet S., Brian Sack, and Eric T. Swanson (2005) 'Do Actions Speak Louder than Words? The Response of Asset Prices to Monetary Policy Actions and Statements.' *International Journal of Central Banking* pp. 55–93

Gilchrist, Simon, and Egon Zakrajšek (2012) 'Credit Spreads and Business Cycle Fluctuations.' *American Economic Review* pp. 1692–1720

Goodman-Bacon, Andrew (2021) 'Difference-in-Differences with Variation in Treatment Timing.' *Journal of Econometrics* pp. 254–277

Gorton, Gary B., and Andrew Metrick (2012) 'Securitized Banking and the Run on Repo.' *Journal of Financial Economics* pp. 425–451

Greenlaw, David, James D. Hamilton, Ethan Harris, and Kenneth D. West (2018) 'A Skeptical View of the Impact of the Fed's Balance Sheet.' *NBER, Working paper 24687*

Haan, Leo de, and Jan Willem van den End (2013) 'Bank liquidity, the Maturity Ladder, and Regulation.' *Journal of Banking and Finance*

Hamilton, James D. (2018) 'The Efficacy of Large-Scale Asset Purchases When the Short-term Interest Rate is at its Effective Lower Bound.' *Brookings Papers on Economic Activity*

Heckman, James J., Hidehiko Ichimura, and Petra E. Todd (1997) 'Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme.' *The Review of Economics Studies* pp. 605–654

Herbst, Edward P., and Benjamin K. Johannsen (2021) 'Bias in Local Projections.' *Working paper, Finance and Economics Discussion Series, Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board, Washington, D.C.*

Ihrig, Jane, Elizabeth Klee, Canlin Li, Min Wei, and Joe Kachovec (2018) 'Expectations about the Federal Researve's Balance Sheet and Term Structure of Interest Rates.' *International Journal of Central Banking* pp. 341–390

Imbens, Guido W. (2014) 'Instrumental Variables: An Econometrician's Perspective.' *Statistical Science* pp. 323–358

_ (2020) 'Potential Outcome and Directed Acyclic Graph Approaches to Causality: Relevance for Empirical Practice in Economics.' *Journal of Economics Literature* pp. 1129–79

Inoue, Atsushi, and Barbara Rossi (2021) 'A New Approach to Measuring Economic Policy Shocks, with an Application to Conventional and Unconventional Monetary Policy.' *Quantitive Economics* 12, 1085–1138

Jarociński, Marek (2021) 'Estimating Fed's Unconventional Policy Shocks.' *European Central Bank Working Paper Series*

Jarociński, Marek, and Frank Smets (2008) 'House Prices and the Stance of Monetary Policy.' *European Central Bank Working Paper Series 891*

Jordà, Òscar (2005) 'Estimation and Inference of Impulse Responses by Local Projections.' *American Economic Review* pp. 161–182

Jordà, Òscar, Moritz Schularick, and Alan M. Taylor (2013) 'When Credit Bites Back.' *Journal of Money, Credit and Banking* pp. 3–28

_ (2015) 'Betting the House.' *Journal of International Economics* pp. S2–S18

Joyce, Michael, David Miles, Andrew Scott, and Dimitri Vayanos (2012) 'Quantitative Easing and Unconventional Monetary Policy – an Introduction.' *The Economic Journal* pp. F271–F288

Kilian, Lutz (2011) 'Structural Vector Autoregressive'

Kilian, Lutz, and Helmut Lütkepohl (2017) *Structural Vector Autoregressive Analysis* (Cambridge: Cambridge University Press,)

Krishnamurthy, Arvind, and Annette Vissing-Jorgensen (2011) 'The Effects of Quantitative Easing on Interest Rates: Channels and Implications for Policy.' *Brookings Papers on Economic Activity* pp. 215–265

Kuo, Chung-Hun, Atsushi Inoue, and Barbara Rossi (2020) 'Identifying the Sources of Model Misspecification.' *Journal of Monetary Economics* pp. 1–18

Kuttner, Kenneth N. (2018) 'Outside the Box: Unconventional Monetary Policy in the Great Recession and Beyond.' *Journal of Economic Perspectives* pp. 121–46

Lee, David S., Justin McCrary, Marcelo J. Moreira, and Jack R. Porter (forthcoming) 'Valid t-ratio Inference for IV.' *American Economic Review*

McCracken, Michael W., and Serena Ng (2015) 'FRED-MD: A Monthly Database for Macroeconomic Research.' *Federal Reserve Bank of St.Louis Working Paper Series*

Mertens, Karel, and Morten O. Ravn (2013) 'The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States.' *American Economic Review* pp. 1212–1247

Mikusheva, Anna (2013) 'Survey on Statistical Inferences in Weakly-identified Instrumental Variable Models.' *Applied Econometrics* pp. 117–131

Montiel Olea, José L., and Mikkel Plagborg-Møller (2021) 'Local Projection Inference is Simpler and More Robust Than You Think.' *Econometrica* pp. 1789–1823

Montiel Olea, José L., James H. Stock, and Mark W. Watson (2021) 'Inference in Structural Vector Autoregressions Identified With an External Instrument.' *Journal of Econometrics* pp. 74–87

Mordel, Adi (2018) 'Prudential Liquidity Regulation in Banking—A Literature Review.' *Bank of Canada Staff Discussion Paper*

Perotti, Enrico, and Javier Suarez (2011) 'A Pigovian Approach to Liquidity Regulation.' *IDEAS Working Paper Series*

Plagborg-Møller, Mikkel, and Christian K. Wolf (2021) 'Local Projections and VARs Estimate the Same Impulse Response.' *Econometrica* pp. 955–980

Ramey, Valerie A. (2016) 'Macroeconomic shocks and their propagation.' *Handbook of Macroeconomics* pp. 71–162

Ramey, Valerie A., and Sarah Zubairy (2018) 'Government Spending Multipliers in Good Times and in Bad: Evidence from U.S. Historical Data.' *Journal of Political Economey* pp. 850–901

Roberts, Daniel, Asani Sarkar, and Or Shachar (2022) 'The Costs and Benefits of Liquidity Regulations.' *Federal Reserve Bank of New York Staff Reports*

Rochet, Jean-Charles (2004) 'Macroeconomic Shocks and Banking Supervision.' *Journal of Financial Stability* pp. 93–110

Rossi, Barbara (2020) 'Identifying and Estimating the Effects of Unconventional Monetary Policy: How To Do It and What Have We Learned?' *Econometrics Journal* pp. 1–32

Rubin, Donald B. (1974) 'Estimating Causal Effects of Treatments in Randomized and Non-randomized Studies.' *Journal of Educational Psychology*

Rudebusch, Glenn D. (2018) 'A review of the fed's unconventional monetary policy.' *FRBSF Economic Letter* pp. 1–5

Sant'Anna, Pedro H. C., and Jun Zhao (2020) 'Doubly Robust Difference-in-Differences Estimators.' *Journal of Econometrics* pp. 101–122

Sims, Christopher A. (1980) 'Macroeconomics and reality.' *Econometrica* pp. 1–48

Staiger, Douglas, and James H. Stock (1997) 'Instrumental Variables Regression with Weak Instruments.' *Econometrica* pp. 557–586

Stigum, Marcia, and Anthony Crescenzi (2007) *Stigum's Money Market,* 4 ed. (McGraw-Hill Education)

Stock, James H. (2008) 'What's New in Econometrics: Time Series.' *Lecture 7, Short course lectures, NBER Summer Institute*

Stock, James H., and Mark W. Watson (2012) 'Disentangling the Channels of the 2007–2009 Recession.' *Brookings Papers on Economic Activity* pp. 81–135

_ (2018) 'Identification and Estimation of Dynamic Causal Effects in Macroeconomics Using External Instruments .' *The Economic Journal* pp. 917–948

Stock, James H., Jonathan H. Wright, and Motohiro Yogo (2002) 'A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments.' *Journal of Business and Economic Statistics* pp. 518–529

Sun, Liyang, and Sarah Abraham (2021) 'Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects.' *Journal of Econometrics* pp. 175–199

Swanson, Eric T. (2021) 'Measuring the Effects of Federal Reserve Forward Guidance and Asset Purchases on Financial Markets.' *Journal of Monetary Economics* pp. 32–53

Vissing-Jorgensen, Annette (2021) 'The Treasury Market in Spring 2020 and the Response of the Federal Reserve.' *Journal of Monetary Economics* pp. 19–47

Weale, Martin, and Tomasz Wieladek (2016) 'What Are the Macroeconomic Effects of Asset Purchases?' *Journal of Monetary Economics* pp. 81–93

Wooldridge, Jeffrey M. (2021) 'Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Difference-in-Differences Estimators.' *Working paper*

Yardeni, Edward, and Mali Quintana (2022) 'Central Banks: Monthly Balance Sheets.' *Yardeni Research*