ESSAYS IN MACRO-FINANCE AND POLITICAL ECONOMY

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Isaac Nelson Green

May 2022

ESSAYS IN MACRO-FINANCE AND POLITICAL ECONOMY

Isaac Nelson Green, Ph.D.

Cornell University 2022

In the chapters that follow, I present three empirical studies in which I utilize natural experiments in combination with a consideration of institutional frictions to analyze topics related to macro-finance and political economy. The three chapters touch on a disparate array of topics, united by a few common threads. In each study I seek to make a case for causal identification via the quasi-random assignment of some economic variable. In addition, each study explores some feature of the political or regulatory environment, and uses policy frictions as tools for generating this exogenous variation. Two of my studies touch on topics broadly related to macro-finance. In these first two chapters, I utilize frictions introduced by economic regulations to study monetary policy and financial stability. In my final chapter, I take interest in the political process itself, and the incentives that it creates for political actors who award contracts to firms, and grants to state and local governments.

In my first chapter, I study online-based "fintech" lenders in residential mortgage markets and I explore how their development of new technology has affected the transmission of monetary policy. I hypothesize that the proliferation of fintech lenders over the last decade has made it easier for borrowers to refinance their mortgages, particularly during periods of monetary easing, when interest rates are falling and borrowers have the strongest refinance incentives. For identification, I make use of the fact that the certification of non-bank lenders in the United States takes place at the state level, and I show that as new fintech firms entered the market and began to expand in the early 2010's, they gravitated toward states with the least restrictive requirements for licensing new non-bank lenders. I compare mortgage refinancing behavior across neighboring states with different levels of regulatory strictness and show that states with faster fintech entry saw stronger mortgage refinancing. I present additional evidence suggesting that stronger refinancing, amidst falling interest rates, allows for stronger consumption in counties with a substantial fintech presence. I

show that this amplification of stimulative interest rate policy is strongest in counties that are underserved by the traditional banking system.

In my second chapter, I study so-called "credit booms." Prior literature has shown an empirical relationship between rapid credit growth (relative to GDP) and a number of adverse economic outcomes. In order to isolate exogenous variation in credit growth, I study a set of policies called "credit ceilings" across an international panel of countries. Credit ceilings were implemented in these countries in the years after WWII and restricted the rate at which banks could extend new loans. As these countries modernized their monetary policies, they removed these credit ceilings. I show that aggregate credit grows sharply after these ceilings are removed, and I use ceiling removals as an instrument for credit growth in a linear-projections with instrumental variables (LPIV) setting to study the effects of credit growth on a number of macroeconomic and financial outcomes. I find that rapid credit growth is often accompanied by GDP and asset price growth in the short-run with sharp reversals in the medium-run, and often financial crises.

In my final chapter, I shift to a political economy setting, and again analyze incentives that underlie government policies. I ask whether the Electoral College system for conducting presidential elections in the United States creates incentives for policymakers to shift government funds to politically important regions of the country. I analyze federal government expenditures awarded to firms, local governments, and other agencies and assess whether government funds are disproportionately directed either toward politically important battleground states, or to states that strongly support the sitting president's political party. My approach centers on comparisons of spending in counties located next to one another on opposite sides of adjacent state borders. Using this approach, I find evidence suggesting that states which strongly support the winning candidate in an election receive more funding than states which voted for the losing candidate. I also find evidence suggesting that federal funds are directed toward politically important battleground states, particularly in an election year.

BIOGRAPHICAL SKETCH

Isaac Green is a doctoral candidate in finance, in his sixth year of study, at the Johnson Graduate School of Management in the SC Johnson College of Business. Prior to joining the business school at Cornell University, Isaac graduated with a Bachelors of Arts with High Honors in Economics from Oberlin College in Oberlin, Ohio.

ACKNOWLEDGEMENTS

I would like to thank the chair of my graduate committee, Matthew Baron, for his invaluable advice and his feedback on my research. Professor Baron is also a co-author on the working paper, titled "The Aftermath of Credit Booms: Evidence from Credit Ceiling Removals," which comprises the second chapter of this dissertation. I express my sincerest thanks to my graduate committee members: Murillo Campello, Justin Murfin, and Mani Sethuraman, who provided valuable support on my research, and helped me navigate the job market. I would also like to thank the faculty in the SC Johnson College of Business, and my fellow Ph.D. students in the finance program at Johnson for their valuable feedback. In particular, the third chapter of this dissertation is based on joint work I have been conducting with Ekaterina Potemkina. Our work on the intersection of firm behavior and federal government spending is ongoing, and the third chapter of this dissertation owes a great deal to her efforts.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1: Does Fintech Lending Amplify the Transmission of Monetary Policy?

1 Introduction

The residential mortgage market in the United States has changed substantially following the 2008 financial crisis. Traditional depository institutions, particularly the largest banks, have seen their share of new mortgage originations steadily shrink, as newer online-based lenders have entered the market and disrupted the traditional brick-and-mortar business model. These online-based non-bank lenders, which I call "fintech lenders" in this paper, use technology to substitute for the role traditionally played by human loan officers. These lenders have developed algorithms to score potential borrowers, generate customized interest rate quotes, and automatically search public records for relevant property information, among other advances.[1] This technology has been shown to improve the efficiency of mortgage markets in various ways. Fintech lenders are able to process mortgage applications and originate loans faster than other types of firms (see Fuster, Plosser, Schnabl, and Vickery, 2019) and may use "big-data" to better screen mortgage applicants (Buchak, Matvos, Piskorski, and Seru, 2018), potentially reducing cognitive biases that afflict human loan officers (Bartlett, Morse, Stanton, and Wallace, 2019).

While the existing literature mostly focuses on the microeconomic effects of fintech lending, the residential mortgage market has been shown to have important connections to the macroeconomy, particularly in its role in transmitting the effects of interest rate shocks to households (Bernanke and Gertler, 1995; Beraja, Fuster, Hurst, and Vavra, 2019; Di Maggio, Kermani, and Palmer, 2020; Greenwald, 2018). Given this link, it is likely that the rise of fintech mortgage lending has an effect on the transmission of monetary policy.

In this paper, I investigate this possibility. My central hypothesis is that fintech lenders amplify the effects of expansionary monetary policy by alleviating market frictions that impede mortgage refinancing when interest rates on new mortgages fall. By allowing a larger

[1] See NerdWallet, "What is an online mortgage?" https://www.nerdwallet.com/best/mortgages/online-mortgage-lenders, accessed 10/12/2021.

number of borrowers to take advantage of beneficial refinancing opportunities, fintech lenders may induce stronger consumer spending in the wake of interest rate declines, amplifying the effects of Fed policy changes.

I test this hypothesis using annual mortgage lending data from the Home Mortgage Disclosure Act (HMDA) database, in addition to monthly data from Fannie Mae, covering the post-crisis period (2010-2019). This time period coincides with the inception of a large number of new fintech lenders, as well as the rapid expansion of the few fintech lenders that existed in prior years. The HMDA dataset captures nearly the entire universe of residential mortgage loans extended during this period across all lenders (both fintech and non-fintech). I collapse these data into a county-year panel in order to exploit geographic variation across various types of mortgage lending.

I first establish that when mortgage rates fall, fintech lenders are associated with stronger refinancing activity. Specifically, I estimate an interaction regression, in which I regress annual county-level growth in mortgage refinance loans on a lagged measure of local fintech concentration, and the interaction of this variable with the spread between average coupon rates on outstanding fixed-rate mortgages and current 10-year Treasury yields. This mortgage spread captures the difference between prevailing market interest rates and rates paid by borrowers with outstanding mortgages, and thus gives a measure of the incentive to refinance. I consider two measures for assessing the extent of local fintech market penetration: a lagged count of the total number of fintech lenders originating loans in a given county and year, and the lagged share of fintech-originated refinance loans as a fraction of the county's total refinancing volume.

Consistent with my hypothesis, I find that when mortgage rates fall, refinance activity is stronger in counties with greater exposure to fintech lenders. Baseline regression results suggest that for every percentage point fall in mortgage rates, each additional fintech lender active in a county in year t-1 is associated with 1.3% stronger refinance growth in year t. Similarly, a 1% increase in a county's fintech market share in year t-1 is associated with an additional .2% of year-t refinance growth, for each percentage point fall in mortgage rates.

Counties with high and low levels of fintech market penetration may also differ from one another along other, unobserved dimensions, that relate to mortgage lending. One particular concern is that fintech lenders are attracted to counties that are poised for strong future refinance loan demand. To address endogeneity concerns, I adopt an identification approach that makes use of the geographic expansion of fintech lenders over time. Fintech lending is a relatively recent phenomenon, and at the beginning of my sample period, in 2010, most fintech lenders were quite small. Despite possessing online lending technology that would potentially allow them to originate mortgages across the entire country at a low cost, in the nascent stages of their development, most fintech lenders nonetheless originated mortgages in only a small handful of states, before gradually expanding. I conjecture that the staggered timing with which lenders tend to enter state mortgage markets is affected by the regulatory protocol governing non-bank mortgage lenders. Given that there is no equivalent of a national bank charter for non-depository institutions, fintech firms must become licensed by state-level regulators in each state in which they want to originate loans. In a logistic regression setting, using hand-collected data on state-level licensing requirements, I show that states with the most restrictive licensing requirements see slower fintech entry. Specifically, high licensing application costs, net-worth requirements, laws requiring the establishment of physical (i.e. brick-and-mortar) branch locations, and the number of other qualitative application requirements, all decrease the probability that a fintech lender will enter a particular state before a given year in the sample.

I then make use of this exogenous source of variation in fintech entry across states, and compare adjacent counties located across state borders from one another in neighboring states with different numbers of licensed fintech lenders. For each year of the sample, I identify all pairs of bordering states which differ in the number of fintech lenders they have licensed, labeling states with a greater number of fintech lenders than their neighbor as "treated" states. I then form my sample by retaining only counties located close (within 50 or 100 miles) to the shared border with their paired state. By doing so, I generate a sample of counties with similar demographics and housing markets, but with differential

access to fintech lending. Using a regression discontinuity framework, I find that refinance loan growth is between 1%-3.3% stronger in "treated" counties, located in states with more fintech lenders than their neighbor. These differences are far larger during years in which interest rates fall and the refinance incentive is high. A 1% widening of mortgage spreads amplifies the treatment effect by 1.5 to 6.4 percentage points, depending on the specification.

The goal of expansionary Fed policy is to stimulate the economy by inducing stronger consumer spending, business investment, or other economic activity. Thus, for fintech lending to amplify the effects of monetary policy, stronger refinancing activity must transmit to other outcomes. To study local consumer spending in the wake of expansionary monetary policy, I make use of store-level retail sales data from Nielsen. I aggregate this data to the county level, employing several filters to ensure the comparability of observations across counties, and the consistency of this measure of spending across time. Returning to my baseline interaction regression specification, I use my new measure of county-level retail spending to assess whether consumer spending growth is stronger in high-fintech jurisdictions amid falling interest rates. I find evidence consistent with this hypothesis: the addition of a single fintech lender to a county's mortgage market is predicted to raise retail consumption growth by .2%, an effect which doubles in magnitude after a 1% widening of mortgage interest rate spreads. Similar results are observable in other outcome variables related to local consumer demand shocks.

Prior research on fintech mortgage lending has proposed different mechanisms through which fintech lenders may reduce market frictions. If fintech lenders expand refinancing only because they are able to process more applications in a short time than other lenders, as described by Fuster et al. (2019), then there is no reason why the impact of fintech presence should vary substantially across regions. On the other hand, if fintech lenders are better at screening particular types of borrowers, or if online lending technology is more useful in places with less access to the brick-and-mortar banking system (such as remote or sparsely populated areas), then fintech lending may be more potent as a facilitator of monetary policy in particular areas of the country.

Continuing with my baseline interaction regression approach, I examine whether fintech-driven monetary transmission (in the form of fintech effects on refinancing and consumption growth) is stronger in regions where borrowers have limited access to traditional finance, either due to a history of lending discrimination or lack of physical access to brick-and-mortar bank branches. To do so, I interact my measures of local fintech presence with quartile indicator variables that capture where a county falls within the distribution of counties, as sorted by racial or ethnic composition, population density, or bank branch prevalence.

I find that conditional on a 1% widening of mortgage interest rate spreads, a unit increase in the number of fintech lenders active in a county at time t-1 predicts 1.9% stronger refinance loan growth in counties with the smallest share of White residents (i.e. in the bottom 25% of counties, sorted by White population shares), but only .6% stronger refinance growth in counties with the largest White population shares, a difference which is statistically significant at the 5% level. I find similar results when sorting counties by their share of Hispanic or Latino residents. Using similar tests, I show that high fintech presence is more strongly associated with refinance growth in regions with few bank branches, and low population densities. Moreover, the same correlations broadly hold when using consumer retail spending growth as the left-hand side variable, suggesting that there is pass-through from household balance sheets to other local outcomes.

This paper contributes to a growing body of literature on the interplay of housing and monetary policy (e.g. Chen, Michaux, and Roussanov, 2020; Drechsler, Savov, and Schnabl, 2019; Taylor, 2007), particularly those which relate to the so-called refinance channel of monetary policy, through which a decrease in interest rates generates increased refinancing and stronger consumer spending among those who refinance (see, e.g. Beraja et al., 2019; Di Maggio, Kermani, and Palmer, 2020; Eichenbaum, Rebelo, and Wong, 2018; Scharfstein and Sunderam, 2018). Relative to these papers, my study is unique in showing how the rise of a new class of intermediaries with different lending technology can influence the strength of this channel.

This paper is also related to a number of recent studies that have focused on the rapid emergence of fintech lenders over the past several years (see, e.g. Balyuk, Berger, and Hackney 2020; Chernenko, Erel, and Prilmeier 2019; Gopal and Schnabl 2020; Philippon 2016; Stulz 2019). The study which relates most closely to this paper is Fuster et al. (2019). Given their findings on the impact of fintech lenders in alleviating microeconomic frictions in lending markets, the authors speculate that these technological advances may enhance the effectiveness of Fed policy, and estimate a regression similar to my baseline analysis, showing a correlation between fintech lending and refinance credit growth. Relative to that study, my analysis focuses more closely on how the effect of fintech lending on credit growth varies alongside changes in interest rates, and across different time horizons. It is also novel in using state-level regulatory frictions to identify a causal channel in addition to presenting correlations. I also add new evidence suggesting that amid falling rates, consumption growth is stronger in the aftermath of fintech-induced refinancing, a result which is important in establishing the amplifying effects of fintech lenders on monetary transmission. Finally, my paper is novel in suggesting that in addition to broadly amplifying the transmission of monetary policy, fintech lending also expands the geographic reach of Fed policy through its unique ability to penetrate areas that elude the traditional banking system.

The rest of this paper is organized as follows. Section 2 introduces the data sources and key variables and discusses how fintech lending varies over regions and how it co-varies with county-level demographics. Section 3 presents the results of my baseline analysis on fintech presence and refinance credit growth. Section 4 delves further into identification of this effect and presents the cross-border analysis on fintech credit expansion. Section 5 examines the link between fintech lending and consumer spending growth. Section 6 shows how the strength of fintech-driven monetary policy transmission varies across regions according to county-level demographics and regional traits. Section 7 concludes.

2 Data

2.1 HMDA and other Data Sources

My primary source of information on mortgage lending activity is the Home Mortgage Disclosure Act (HMDA) database. Under HMDA, lenders meeting certain requirements must enter information annually on all residential mortgage loans that they originate. Residential loans include home loans for purchase, refinancing (including cash out refinancing), and home improvement. HMDA reporting requirements cover virtually all lenders, including non-bank mortgage originators, so the lending information in HMDA covers nearly the full universe of mortgage lending activity in the United States.[2]

HMDA data are reported at the loan-level. For each loan, identifying information is included for the lender, and demographic information on the borrower (e.g. race and annual income) is also generally available. Importantly, the location of the underlying property is also provided at the county level. Information on the purpose of each transaction, that is, whether a loan is for a home purchase, refinance, or for home-improvement, is also available, as is information on the presence or absence of underlying government guarantees (e.g. whether a loan is FHA or VA guaranteed). The majority of my analyses will make use of HMDA data, aggregated at the county level. Since HMDA is reported at the loan level, however, I can identify loan volumes by lender type, and thus calculate the volume of home refinance transactions originated by fintech firms (as well as other lenders) by local market. The county-level HMDA panel that I construct covers the years 2010-2019.[3]

I merge the county-level aggregated HMDA data with county-level data from a number of other sources. I obtain several useful variables from the United States Census Bureau. I obtain demographic information from the United States Census and American Community Survey (ACS), and obtain information on small business activity from the Census' County Business Patterns (CBP) database. I add data on county-level unemployment rates from the Bureau of Labor Statistics' (BLS) Local Area Unemployment data sets. I make use of Call Report data from the FDIC when matching banks and other lenders to the HMDA

[2] Exemptions to reporting requirements apply to small banks and other lenders with total assets below an annually announced threshold. Nonetheless, HMDA requirements apply to the majority of institutions and the vast majority of total loan volume, and can thus be see as the near-universe of mortgage loans in the United States.

[3] The HMDA data covers all counties in the US and Puerto Rico. I drop counties in Puerto Rico, and any counties with multiple years without refinance lending transactions.

data. I also acquire information from the FDIC's Summary of Deposits (SOD) database on the locations of bank branches. I use the SOD data to determine the number of bank branches located in each US county.

The primary advantage of the HMDA data is its completeness. Since HMDA-reported transactions represent a substantial majority of mortgage transactions, statistics drawn from the HMDA data are very likely to be reflective of the overall residential mortgage market. However, for some parts of the analysis in this paper (i.e. analyzing the speed of refinancing by fintech firms) it will be desirable to get a picture of mortgage market activity at a higher frequency. As such, in addition to the merged county-level HMDA panel, I also conduct a set of tests on data from the Fannie Mae Single Family Loan Performance dataset. The Fannie Mae data consists of information on roughly 35 million loans sold to and securitized by Fannie Mae between 2000 and 2020. All of the loans within this dataset are fully amortizing, full-documentation loans. I focus on loans with initial maturities between 15 and 30 years with fixed rates.[4] The Fannie Mae data provide information on loan performance at a monthly frequency, including information on the size of each monthly pay-down or pre-payment, the loan-age and time-to-maturity, and whether a loan is delinquent or in forbearance. Information is displayed for each loan in the portfolio until it fully amortizes, prepays, defaults, or, in rare cases, is removed from the Fannie Mae portfolio for other reasons.[5] The Fannie Mae data also contains static information, such as the interest rate at origination, borrower credit scores and LTV ratios, and whether the purchaser of the home underlying the mortgage is a first-time home buyer. Information on the location of each underlying property is given at the 3-digit ZIP code level. Since credit scores are an important determinant of mortgage market activity, I extract average annual credit scores from each ZIP code, which I merge with the county-level HMDA data. Using data from the department of Housing and Urban Development (HUD) "crosswalk" files, I translate the ZIP code data to the county level.[6]

In addition to the sources of data mentioned above, obtain information on weekly long-term bond yields from the United States Treasury, which I utilize when constructing monthly

[4]Since the dataset consists only of loans securitized by Fannie Mae, it does not contain information on so-called non-qualifying (formerly "sub-prime") mortgages originated to borrowers with low credit scores or high loan-to-value (LTV) ratios, nor does it contain information on so-called "jumbo" mortgages which have loan amounts above a pre-set conforming limit. It also excludes most FHA- and VA-guaranteed mortgages, which are primarily the province of Ginnie Mae.

[5]So-called "put-backs" occur when Fannie Mae removes a loan from an MBS it has issued and requires its originator to repurchase it. Such an event most frequently occurs when Fannie Mae finds that some of the information on the underlying mortgage documentation is fraudulent or misrepresented.

[6]Neither ZIP codes nor counties are subsets of one another (3-digit ZIP codes are somewhat larger than counties, on average). The mapping process uses information from HUD that details the percentage of a ZIP Code's addresses that lie within a particular county. Mapping ZIP codes to counties is thus not an exact procedure.

and annual mortgage spreads. Finally, I construct a measure of county-level consumer retail-spending using Retail Scanner Data provided by Nielsen. Additional details on these sources of data will be discussed in the ensuing sections, and in the appendix.

2.2 Fintech Lending

In Table 1.1, Panel A, I present information on the home refinance lending activity of the fintech firms in my sample. To identify fintech lenders in my sample, I begin by combining the sets of fintech lenders identified in Buchak et al. 2018 and Fuster et al. 2019. I make a few additions to this set of lenders, including SoFi Lending and Zillow Home Loans, both of which are associated with well known technology firms, and thus have the potential to incorporate big-data methods into their screening processes. I also add LenderFi, a more recent market entrant which uses technology extensively in the origination process. Fintech lending increases over time, both in dollar volume and as a percentage of total refinancing loan volume. Fintech loans make up roughly 4% of refinance lending activity in 2010, and gradually rise to roughly 15% of the market. In 2010, there were 12 fintech firms making loans, some of which were very small. This number rises to 22 by 2018. The largest fintech lender in the sample, by some margin, is Quicken Loans, which comprises roughly 60% of all fintech lending, by volume, in 2010, though this share drops gradually as the sample progresses. Panel B in displays information on fintech lending at the county level. At the beginning of the sample, in 2010, the median county has only two active fintech lenders with a market share of 4.3%, and the 90th percentile county has six fintech lenders and a fintech market share of 9.5%. By 2019, the median and 90th percentile counties have seven and sixteen fintech lenders, respectively, while the equivalent figures for fintech market share are 16.6% and 29.4%.

In order to assess whether regional markets with more fintech lending see stronger refinance credit growth, I also need to define a measure of regional fintech presence. Establishing a meaningful geographic footprint of lenders with little physical presence is a difficult un-

dertaking fraught with potential endogeneity issues. When observing where fintech firms make loans, it is unclear whether these firms lend in a particular region because of conscious decisions made by the firms themselves (e.g. to become licensed in a particular state, to target online advertisements toward, or give favorable interest rate terms to, potential borrowers in a particular region, etc.) or if they do so because of high borrower demand for fintech services in an area. I address these difficulties in two ways. First, I will consider multiple alternate measures for assessing the regional presence of fintech firms. I will discuss the benefits and drawbacks of each measure, and I will display baseline results with each of these across a number of specifications. Secondly, in Section 4, I will utilize an identification procedure that looks at differences in the number of licensed fintech lenders across states, utilizing potentially exogenous differences in regional fintech presence.

The first measure of regional fintech presence that I will utilize is a simple count of the number of fintech firms that originate a refinance loan within a given county and year. In regression specifications, I will refer to this variable as *FintechCount*. The most obvious drawback of the *FintechCount* variable in assessing an areas fintech exposure is that could grow larger by virtue of one or two fintech lenders originating a single loan in a county. It would be hard to argue that such activity would represent a meaningful increase in the extent to which fintech lenders pervade that county. On the other hand, if *FintechCount* is, on average, a meaningful proxy for the number of fintech lenders that are *willing* to originate loans in an area, then the count of fintech firms would be a useful measure of the effect that *access* to fintech lenders has on a local market.[7] While the count of active fintech originators may be a noisy measure of fintech activity, given the low variability of the measure, and the large number of US counties, there is hope that the large number of observations would uncover the average effects of fintech lender access on a regional market.

The second measure that I use to define fintech concentration is the regional market share of fintech lenders within the refinance segment of mortgage originations. That is, I divide

[7] This variable may do a good job of capturing the extent to which there is a competitive environment local market among fintech lenders. Even if fintech firms lower the costs associated with refinancing, they may not pass along the cost savings to consumers in the form of lower rates unless there is some competitive reason for them to do so. See Taylor (2007) for a discussion of these issues in an environment that does not explicitly analyze fintech behavior.

the volume of refinance loans originated by fintech lenders in a given county and year, by the total volume of refinance loans originated in that county and year. I refer to this variable as *FintechShare.* While *FintechShare* benefits from the fact that it is a more quantitatively rich measure of fintech activity than the count variable, using this variable (lagged by a year) to predict refinance credit growth faces its own issues. First, the denominator of *FintechShare* depends on the activity of all of the other intermediaries in a county. As such, fintech market shares can become large both if fintech firms expand their origination activity, or else if other intermediaries cut back. Buchak et al. 2018 argue that non-bank lenders (of which fintech firms are a subset) increased their market shares in areas with weak intermediaries that had to raise capital in the post-2010 period and subsequently cut back lending.[8] Thus, high fintech market shares could signal that the aggregate credit supply is contracting due to the retrenchment of other intermediaries.

Another issue with the *FintechShare* variable is that it may, in part, proxy for the effect it is trying to measure. In other words, since one hypothesis regarding fintech lenders is that they may be able to more easily reach borrowers that are less easily screened by banks, a high fintech market share in year t-1 may signal that fintech firms have already targeted and refinanced those borrowers, and that there are thus fewer borrowers available to be refinanced in subsequent years.

In Figure 1.1, I display the regional patterns associated with both of these measures of fintech lending. The two maps shown in the diagram display the average number of active fintech lenders, and the average market share of fintech refinance lending, at the county level, across the full sample from 2010-2019. As measured by the number of fintech firms active within a county, fintech lenders are most prominent in large urban and suburban areas with high populations. This tendency comes as little surprise, since areas with the largest housing markets might reasonably also be expected to have the largest volume of fintech loans. Since I will always control for county-level populations in my regression analyses, it will be instructive to look at where *FintechCount* is high relative to a county's population,

[8] Capital constraints for these firms were likely binding due to financial crisis-era losses and new regulatory capital requirements, which began to be implemented after the 2010 Dodd-Frank Act was passed.

as I will illustrate shortly. Regional fintech presence is somewhat different when looking at market shares. The counties where fintech lenders represent the largest fraction of the total refinancing market tend to be counties with low population densities, often in the western United States. States with large number of high fintech-share counties include Nevada, New Mexico, and Alaska, each of which cover large land areas with dispersed populations.[9]

In addition to these two these main measures of fintech concentration, for robustness, when conducting analyses using monthly Fannie Mae data, I will also look to estimate the effects of recent growth in the number of fintech refinance loans. This measure of fintech activity will not be subject to the effects of the idiosyncratic behavior of other intermediaries (as market shares would be) and since it is expressed as a growth rate, it will not be strongly correlated with county sizes. I will discuss this measure in further detail in the sections that follow.

2.3 What Explains Regional Fintech Concentration?

Fintech lending is likely to be explained by a number of supply- and demand-related factors, which also correlate with regional housing market activity. The purpose of this section is to better understand the factors that correlate with regional fintech activity and to introduce the reader to some of the observable factors I will need to control for when examining the relationship between fintech activity and credit growth.

To gauge the extent to which various county-level observables are able to explain regional variation in fintech lending, I begin by regressing my measures of regional fintech concentration on a number of county observables. Summary statistics on this county-level merged HMDA panel are shown in Table 1.2. In addition, simple pairwise correlations between groups of variables are displayed in Appendix Table 1.13. I estimate regressions of

[9] Each of these states also feature physical barriers (i.e. deserts in the case of Nevada and New Mexico, and tundra in Alaska) which might make them difficult to access and contribute to their being poorly connected to the traditional financial system.

the following form:

$$Fintech_i = \alpha + \beta X_i + \varepsilon_i,$$
$$where\ Fintech\ \epsilon\ \{FintechCount, FintechShare\} \tag{1}$$

where the subscript i indexes counties, *Fintech* measures the average level of fintech presence in a county across all sample years, and X is a vector of county-level observables, which are also averaged across sample years.

Table 1.3 displays the results of estimating (1) on various sets of county observables. Panel A displays results where the dependent variable is *FintechShare*, while Panel B displays results for the *FintechCount* specifications. I have five sets of county-level observables, categorized with different labels in the leftmost column of the table. The variables labeled as "HMDA Mortgage Variables" consist of a county's share of FHA guaranteed mortgages and so-called "jumbo" mortgages. FHA mortgages are a riskier segment of the market given to lower income borrowers. Existing evidence suggests that non-bank lenders target this loan segment, as they may have a regulatory advantage in originating these mortgages.[10] Jumbo mortgages consist of loans originated for amounts above the maximum amount for loans eligible for sale to the government sponsored mortgage agencies (GSEs).[11] Since non-bank lenders sell the vast majority of the mortgages they originate to one of the GSEs, areas where there is high demand for jumbo mortgages may deter fintech entry.

The HMDA mortgage variables, together with the list of "Demographic Variables" displayed in Table 1.3 comprise the set of baseline county-level control variables, which I will use in credit growth regressions in the following section.[12] The demographic variables include county populations, average incomes, unemployment rates, population density, and the employment to population ratio. These two sets of variables are available for a broad set of US counties, at an annual frequency. In addition to these variables, I have "ACS

[10] See Agarwal et al. (2020).

[11] The so-called GSEs: Fannie Mae, Freddie Mac, and Ginnie Mae.

[12] I will add average credit scores, calculated via the ZIP code level Fannie Mae data to complete the set of baseline right-hand side variables.

Mortgage Variables" and "ACS Demographic Variables," which extend my set of observable characteristics. These variables are only available for larger counties with populations over 60,000. I refer to these variables, combined with the baseline controls, as the "full" set of right-hand side control variables, and I estimate separate regressions that include this full set of controls and the smaller set of counties.[13] The final set of county-level observables is labeled "Bank Branch Presence," and includes the number of bank branches per-capita and per square mile.

The results shown in Table 1.3 suggest that while the county-level maps in Figure 1.1 highlight differences between the *FintechCount* and *FintechShare* variables in measuring local fintech concentrations, the two measures correlate similarly with county-level observables, particularly once accounting for the strong co-variation between the count of fintech firms and a county's population. In the first column of each panel, I show how fintech lending activity varies with characteristics of a county's local mortgage market. Both measures have a strong positive association with the average share of FHA-guaranteed loans among refinance loans in a county, and both measures suggest that fintech lenders have a stronger presence in counties with greater home ownership.

In the second column of each panel, I look at fintech presence relative to the set of baseline controls. The results suggest that fintech firms originate loans in counties with wealthier borrowers and lower unemployment rates, but also with a smaller share of employed people relative to the total population. Fintech firms also originate loans in counties with lower population densities. In column (3) of each panel, I add demographic and mortgage market-related variables from the American Community Survey (ACS). The ACS data suggest that fintech presence is stronger in areas where a larger share of the population is 65 years old or above. On the county level then, there is little evidence that fintech firms target younger, technologically savvy borrowers. Instead, this finding suggests that fintech lending may be more appealing to those who have some experience taking out a mortgage, and thus may

[13] Limiting the sample to counties with populations above 60,000 excludes roughly the bottom two-thirds of counties, by population.

not need the extra hand-holding provided by interactions with human mortgage lenders.[14]

The ACS data also allow an assessment of the racial and ethnic composition of the counties in which fintech lenders originate mortgages. I include a county's White and Black population shares in the column (3) regressions.[15] Both measures of fintech presence suggest that a decrease in a county's white population, in favor an equal percentage increase in that county's Black population predict an increase in the level of fintech activity.[16] Additionally, both the number of active fintech firms and fintech market shares are strongly positively associated with a county's Hispanic or Latino share of the population. Taken as a whole, then, fintech firms tend to be more active in counties with larger concentrations of racial or ethnic minorities.[17]

Finally, in column (4), I estimate how fintech presence co-varies with the presence of bank branch locations. Both measures of fintech presence suggest that fintech lenders are more active in areas with fewer branches per-capita. The evidence with respect to the number of bank branches per square-mile is mixed, after controlling for population density.[18]

3 Interest Rate Declines and Fintech Credit Expansion: A Difference-in-differences Approach

In this section I begin to test the aggregate effects of fintech lending. I ask whether fintech firms amplify monetary policy by increasing the availability of mortgage refinance credit when interest rates decline. If fintech lenders are able to more efficiently process mortgage

[14] More experienced borrowers may also be more likely to refinance when it is optimal for them to do so, as suggested by Browne et al. (1996), suggesting borrower experience could be a confounding factor in regressions of refinance growth on fintech activity.

[15] This leaves members of the American Indian and Asian/Pacific Island communities, as well as those who identify as "Some Other Race" as the residual population.

[16] The signs of coefficients on the Black and White population share variables differ according to which measure of fintech presence is used, however the magnitudes yield similar intuition. The signs of the White and Black coefficients in column (3) in panel A suggests that fintech market shares are higher in counties with higher concentrations of those identifying as some race other than Black or White, while panel B suggests that the number of active fintech lenders is smaller in such places. However, both sets of regressions show smaller (more negative) coefficients for White population shares than for Black population shares.

[17] The census' treatment of race and ethnicity makes the analysis rather more confusing. On the US Census, respondents do not have an option of identifying as Hispanic or Latino when selecting a race, despite the fact that many Americans who identify as Hispanic or Latino would identify this as a racial categorization as well as their ethnicity. Hispanic/Latino origin is instead treated separately. Thus, under racial identity, many Hispanic/Latino Americans identify either as White, or as "Some other race."

[18] Bank branches per square-mile and population density are strongly collinear with a correlation coefficient of .913 (see Appendix Table 1.13).

applications and assess credit-worthiness, then the presence of fintech lenders in a local housing market could have an expansionary effect on the aggregate supply of credit. If, as suggested by Fuster et al. 2019, automated lending technology is more valuable when demand for credit is strong, then the activity of fintech lenders should have the strongest effects on aggregate credit provision in declining interest rate environments, when a larger proportion of borrowers have an incentive to refinance.

3.1 Baseline Specifications and Results

I first look to assess the correlation between fintech lending activity and the supply of home refinancing credit by exploiting regional variation in fintech lending activity. I ask whether the growth of mortgage refinance credit is stronger in local markets that have a more concentrated fintech presence, particularly when rates fall and the aggregate refinance incentive is strongest.

To assess the relationship between regional credit growth and fintech lending, I make use of the HMDA data, aggregated at the county level, to estimate the following regressions, where I use the subscripts i and t to index counties and years, respectively:

$$\Delta_1 Refivol_{i,t} = \alpha_t + \beta \cdot Fintech_{i,t-1} + \gamma \cdot Fintech_{i,t-1} \cdot \Delta_{avg} Rates_t + \delta \cdot Controls_{i,t-1} + \varepsilon_{i,t} \quad (2)$$

where I adopt the convention that for a variable, y,

$$\Delta_k y_{i,t} = y_{i,t} - y_{i,t-k} \; and, \; \Delta_{avg} y_{i,t} = y_{i,t} - \bar{y}_i$$

where the last y term in the final expression denotes the county average, across all sample years.

The variable denoted as *Refivol* is defined as the natural logarithm of the aggregate county-level volume of refinance loans originated in year t. Thus, the outcome variable is the log-change in total refinance credit in county i from year t-1 to year t. I will use *Refivol*

as of the year 2010 to weight observations in these regressions. The intuition is that refinance growth should be weighted more strongly in counties with larger housing markets. I choose the year 2010 because fintech still comprised a minimal portion of the overall market at this point, so choosing this year minimizes the potential effects of fintech lending on the weighting variable.[19]

I use the growth rate, rather than the level of refinancing primarily to account for the fact that different counties have very differently-sized housing markets. Without such a normalization, specifications without county fixed-effects would be highly suspect, since it would be a difficult task to control for time-invariant cross-county differences in housing market activity with only county observables. Moreover, using growth rates allows county fixed-effects, when added to a specification, to control for general time trends in the growth-rates of refinancing within a county, rather than simply the level of refinancing. Use of fixed effects in such a manner will help to alleviate concerns that fintech lenders simply enter counties where refinancing activity exhibits persistently strong growth during this period.

The intercept term, α in (2) has a time subscript, t, to denote the presence of year fixed-effects in all specifications. Since county-level refinancing growth rates are strongly correlated over time due to the aggregate interest rate environment and other macroeconomic fundamentals, it will be important to include year fixed-effects to isolate the cross-sectional differences in which I am primarily interested.[20]

The variable labeled *Fintech* denotes a county-level measure of fintech market presence, defined either as the *FintechCount* or *FintechShare* variables described earlier. These enter in lagged form in (2). The next key variable, labeled *Rates* is a measure of the aggregate refinance incentive at time t. This variable, though labeled *Rates* for intuition, is really an interest rate spread. It is defined as the difference between the average coupon interest rate (i.e. the rate at origination) of outstanding fixed-rate mortgage loans minus the yield to

[19] Choosing population rather than housing market size produces similar results.

[20] Time fixed-effects will also be important in dealing with the general upward trend in fintech activity over time. Removing the covariation between fintech activity and year fixed effects removes the potentially mechanical relationship that could arise between refinancing and fintech activity if, for example, refinancing growth happens to be strongest in the later years of the sample when fintech shares are also high.

maturity on 10-year US Treasury debt. This variable captures the difference between the average rate paid by those with a currently outstanding mortgage, and the level of market interest rates, and thus represents the potential cost savings associated with refinancing.

The coefficients on the two fintech variables have the following interpretation. The β coefficient reveals whether, in an average interest rate environment, the growth of refinancing credit is stronger in counties with a strong fintech presence. A positive coefficient on this variable would indicate that refinance growth is greater in high-fintech counties. The γ coefficient then describes how much this estimated fintech effect increases or decreases conditional on a 1% increase in the aggregate refinance incentive.[21] Positive loadings on both of these coefficients would suggest that aggregate county-level results mirror the microeconomic stylized facts which suggest that technological advances in mortgage lending are more valuable in strong housing demand environments. The *Rates* variable primarily moves around when market interest rates (i.e. Treasury yields) change. Thus, a widening spread should be thought of, intuitively, as indicating a falling interest rate environment.

In addition to the variables related to fintech presence, the above specification also controls for a number of county-level observables. I will display results from regressions that include both the so-called "baseline" controls and the "full" set of controls, as defined in Section 2. In addition to the true county-level controls mentioned, I also add to a county's mean borrower credit score, which I derive from Fannie Mae data. I will present results of specifications both with and without county-level fixed-effects.

The regression framework outlined in equation (2) can be thought of as analogous to a difference-in-differences approach. In this interpretation, the β coefficient can be thought of as giving the differences in refinancing activity between high- and low-fintech regions in normal times, while the γ coefficient details how this refinancing behavior changes in response to interest rate shocks. While it is likely that fintech firms enter county-level housing markets endogenously, it may be plausible that the deliberations that give rise to fintech market-entry in a county at time t-1 do not anticipate subsequent changes to the

[21] The year-over-year standard deviation of this spread is roughly .8%.

aggregate interest rate environment at time t. The identification assumption underlying the difference-in-differences interpretation would be that, in the absence of fintech activity, and controlling for county-level observables, high- and low-fintech counties would have similar growth of refinance credit in response to a change in the aggregate interest rate environment.

Table 1.4 displays the results of estimating equation (2) on the merged panel of county-level HMDA data, using the two measures of fintech market presence. Panel A displays coefficients on the two *FintechCount* variables (i.e. the pure count variable and its interaction with *Rates*), while Panel B shows coefficients for the *FintechShare* variables. The specifications across the four columns vary in the sets of observables that are included as controls, as well as in their inclusion or exclusion of county fixed-effects. Specifications with the "Full" set of right-hand side observables include variables from the American Community Survey data, which restricts the sample only to large counties, subsequently reducing the number of observations in these specifications.

The coefficients on the *FintechCount* variable, displayed in Panel A, suggest that the marginal addition of a single fintech lender to a county's residential mortgage market at time t-1 corresponds to between .4%-.8% stronger refinancing growth in the subsequent year. This relationship is strengthened in environments where interest rates are declining and the incentive for borrowers to refinance is high. The interaction coefficients in Panel A suggest that conditional on a 1% increase in the aggregate refinance incentive at time t (i.e. a 1% rise in the standardized *Rates* variable from equation (2)), the estimated effect of an additional fintech lender on refinancing rises by 1.4%-3.0%, suggesting a total effect on refinance credit growth of between 1.9%-3.8%. These results suggest that the estimated impact of an additional fintech lender at time t-1 grows roughly 5-7 times larger conditional on a 1% widening of interest rate spreads at time t.

The *FintechShare* coefficients in Panel B mirror the results in panel A. Panel B suggests that a 1% increase in fintech market shares in the refinance segment of the market, in year t-1, predicts between .36%-1.09% stronger refinance credit growth at time t. The largest *FintechShare* coefficient is obtained in the specification that includes the full set of controls

as well as county fixed-effects. The results across all specifications in Panel B suggest that the estimated effect of fintech firms gaining market share grows larger when the aggregate refinance incentive increases. The increase in the estimated fintech effect on refinance growth, conditional on a 1% widening of the aggregate refinance incentive, ranges from .23%-.60%. Given the *FintechShare* coefficients, the loadings on the *FintechShare*Rates* variable suggest that the effect of fintech lending on subsequent credit growth becomes at least 60% larger upon a 1% widening of interest rate spreads.

One potential issue with the *FintechCount* variable concerns its covariation with county population. It is apparent from Figure 1.1 that areas with larger populations, and housing markets, have a larger number of active fintech lenders. While I control for population in all specifications, if population has highly nonlinear effects on refinancing at the high -end of the distribution (for example, if the largest cities have sophisticated borrowers that are more attuned to interest rate movements), then this could have an impact on the results in specifications without county fixed-effects. While the inclusion of county fixed-effects solves this particular issue, the largest urban areas have very different housing markets than other counties, and it may thus be concerning if the association between fintech concentration and refinancing growth didn't hold in smaller counties. In Table 1.5 I display the results associated with the estimation of equation (2) across population-sorted sub-samples. Panel A suggests that the *FintechCount* coefficients are fairly consistent across the population distribution. The coefficient on the interaction term (i.e. *FintechCount*Rates*) remains stable from the second through fourth population quartiles, at a value of .016. Coefficients on the non-interacted fintech term are slightly weaker at the lower end of the population distribution, but not dramatically so.

In Table 1.14 I display results of two additional tests of the robustness of my results. First, I assess the possibility that fintech lenders enter counties while refinancing growth is already strong, because their business models prioritize the refinance sector of the market. In this case, the ability of the lagged count of fintech lenders to forecast refinancing growth might be due to the persistence of a temporary refinancing surge that began prior to or

contemporaneously with fintech firm arrival. I attempt to rule out this possibility by re-estimating a version of equation (2) that controls for the lagged growth of refinancing. I display the results of these specifications in panel A of Table 1.14 . The results across all specifications mirror the baseline results, with coefficient estimates that are quite similar in magnitude to those presented in the baseline analysis, in Table 1.4.

Next, I look at whether fintech presence differs from other that of intermediaries. It may be the case that entry into a local mortgage market by intermediaries of any type is broadly predictive of subsequent refinance activity. For example, an increase in the total number of intermediaries originating loans in a region may signify that the local market is becoming more competitive. These competitive effects may result in more borrowers being able to refinance. To consider this possibility, I estimate analogous regressions to those displayed in panel A of Table 1.4 (i.e. the *FintechCount* regressions) but instead of using *FintechCount* on the right-hand side, I include counts of other intermediaries. In panel B of Table 1.14 I show results of within-county estimates of equation (2) which include these additional intermediaries.[22] The three new intermediary-count variables, which I name *OtherNonbank*, *LargeBank*, and *SmallBank*, refer, respectively, to counts of non-bank non-fintech lenders, large bank lenders (with assets over $50 billion) and small bank lenders (with assets below $50 billion).

The results of this analysis suggest that fintech firms are indeed different than other types of intermediaries. In an average interest rate environment, the coefficients in Table 1.14 , panel B, suggest that the heightened presence of these other types of intermediaries actually forecasts lower subsequent refinancing volume. For each of these types of intermediaries, a widening of interest rate spreads diminishes this negative association, however, in most cases, a 1% increase in the refinance incentive would not be sufficient (or would be barely sufficient) to make the aggregate association positive. For example, the coefficients on *OtherNonbank* of -.003 and -.002, combined with the interaction effect of .001, suggest that even in the event that interest rate spreads widened by 1%, the presence of other non-bank intermediaries would still forecast between .1% and .2% lower refinance growth in the year

[22]Versions of these regressions without county fixed-effects are not displayed. The results of the analogous specifications, without fixed-effects, suggest that counties with more large non-bank lenders and large banks have subsequently stronger refinancing activity, though the results are much weaker than the *FintechCount* results. The presence of small banks negatively forecasts subsequent refinance activity in these specifications.

the interest rate shock took place.[23]

3.2 The Timing of Refinancing After Interest Rate Movements

The results shown up to this point illustrate that regions where fintech lenders are more active tend to have stronger refinancing growth, particularly in declining interest rate environment where the aggregate incentive to refinance is high. However, the HMDA data displays mortgage lending transactions only at an annual frequency, while intra-year variation in interest rates can be substantial. On average, the time between an initial submission of a mortgage application and the closing of a loan is 2-3 months. Thus, if strong refinancing growth in high-fintech regions has to do, in part, with the ability of fintech lenders to remove capacity constraints from local markets by more efficiently processing mortgage applications, then when refinancing incentives increase, the relative differences in refinancing growth between high- and low-fintech areas should manifest relatively quickly.

To get a sense of the timing of refinancing relative to interest rate movements, I make use of the Single-family Loan Performance dataset from Fannie Mae. The higher frequency of the Fannie Mae data allow for a more refined observation of the dynamics of fintech lending alongside interest rate changes. In Panel A of Appendix Table 1.15 I show that the results of my baseline analysis, from section Section 3.1, can be replicated over a shorter (three month) time horizon using this dataset. Next, I estimate impulse responses using local projections, as described by Jorda (2005). In particular, I estimate OLS regressions of the form

$$\begin{aligned}\Delta_h Refivol_{i,t+h} = \alpha_t^h + \sum_{k=0}^{2} \beta_k^h \cdot \Delta_1 Refivol_{i,t-k} + \gamma^h \cdot \Delta_3 Fintech_{i,t-1} \\ + \delta^h \cdot \Delta_3 Fintech_{i,t-1} \cdot \Delta_1 Rates_t + \lambda^h \cdot \Delta_1 Controls + \varepsilon_{i,t+h} \qquad (3)\end{aligned}$$

where h=1, 2, ...5 denote the time horizon, in months, over which the regressions are

[23] This estimate is derived from the sum of the *OtherNonbank* coefficient and the *OtherNonbank*Rates* coefficients.

estimated. The impulse responses are the dynamic evolution of the *Fintech* coefficients, γ^h and δ^h. In practice, I will plot the sum of these coefficients (and report both coefficients in appendix tables).

In contrast to equation (2), all variables are now expressed as growth rates, rather than levels. With monthly data, it is easier to control for several lags of past refinancing growth (i.e. without sacrificing several years of sample data). To the extent that there is a persistent fintech effect on refinance growth related to the overall level of fintech presence, this effect should generally be impounded in the lagged growth of refinance activity. Thus, this specification should better control for the correlation between very recent changes in fintech presence and refinance growth. In theory, if a stronger fintech presence leads to stronger refinancing, then once we control for recent refinance activity, the areas with the largest increases in refinance growth should be those with an increasing fintech presence. For the same reason, I opt to use changes in interest rate spreads in these specifications, rather than levels. Doing so allows me to better capture the timing of true interest rate shocks.[24]

Another important point is that when estimating each specification in differences, I am able to look at changes in fintech lending activity without concern about the differential size of housing markets across geographic areas. That is, I can re-define the fintech variable to be the log growth rate of the number of total fintech loans, without worrying that the sizes of these values simply pick up the sizes of regional housing markets. This is useful because the FNMA panel only explicitly identifies eight of the largest fintech firms in its panel. Smaller lenders, whose combined loan volume does not constitute at least 1% of national origination volume in a given year, are listed as "other" in the FNMA data.

Panels B and C of Appendix Table 1.15 display the impulse responses that result from these specifications, with panel B displaying results where *Fintechshare* is used as the right-hand side fintech variable, and panel C showing results where the growth rate of the number of fintech loans is used. In Figure 1.2 panels A and B, I plot the impulse responses associated

[24] If refinancing incentives have been large for some time, then lagged refinancing activity should also be high, and the growth of refinancing relative to its own growth in the recent past should not be particularly strong. It is instead when we see a dramatic change in recent interest rate spreads when refinancing should pick up relative to its recent behavior.

with the latter of these sets of specifications. The impulse responses suggest that the effect of fintech lending on refinance growth is both quick and persistent. The results in panel C of Appendix Table 1.15 suggest that for every one percent of growth in the number fintech loans from time t-3 to time t-1, and conditional on an interest rate shock in the form of a 1% widening of mortgage spreads from time t-1 to time t, we can expect .056% stronger growth in refinancing within the first month after the interest rate shock. This aggregate fintech effect rises to .128% after two months and .180% after three months, after which point the effect levels off. The effect is quite persistent, leveling off in the fourth month after the interest rate shock, and only declining slightly by the fifth month, from .187% to .179%. One reason for the persistence of the fintech effect may be that interest rates are also persistent, and many homeowners do not attempt to refinance right away, even when it is in there interest to do so (see, for example Stanton, 1995 and Agarwal, Rosen, and Yao, 2013). Thus, if interest rates decline at time t, it is likely that some borrowers who have a refinance incentive will nonetheless not immediately submit applications to refinance. If rates remain low for multiple months, some of these inattentive borrowers may ultimately refinance with a delay.[25]

In the final set of tests in this section, I look to assess whether the correlation between fintech lending and refinance credit growth is stronger amidst a spike in regional credit demand. When regional demand is strong, and local loan officers are overwhelmed by the surge in activity, lenders with more extensively automated lending technology should have an advantage in processing the onslaught of loan requests. To test this proposition, I develop a proxy for local refinance demand, which considers the total stock of Fannie Mae-guaranteed mortgages in a region with a refinance incentive. Early work on the mortgage-backed securities market by Richard and Roll (1989) suggests that mortgage pools exhibit very low pre-payment rates at issuance and then begin to pre-pay more extensively as underlying loans age, with mortgage pools becoming fully seasoned 30-60 months after initial issuance. As such, I proxy for changes in local refinance demand by looking at

[25]Thus, even if the average mortgage application is processed in 2-3 months, refinancing volume may remain high several months after an initial interest rate shock due to the delayed reaction of some borrowers in submitting refinance applications.

changes in the total stock of mortgages, aged 30-months or more, with a refinance incentive. Specifically, I define the local refinance incentive in ZIP code i as

$$OutstandingStock_i = log\left(\sum_k Ratespread_{i,k} \cdot Mortgagebalance_{i,k}\right)$$

where the sum is taken over all mortgages, k, with loan ages greater or equal to 30 months, and less than or equal to 240 months.[26] The variable *Ratespread* denotes the interest rate spread on each individual mortgage[27] while the variable *Mortgagebalance* is the remaining principal balance on each mortgage. The *OutstandingStock* variable gives the total value of outstanding fixed rate Fannie Mae-securitized mortgages in a 3-digit ZIP code weighted by their individual interest rates at origination (less current market interest rates). Intuitively, mortgages that carry higher interest rates have more to gain from refinancing, and the sum of outstanding balances on mortgages within a ZIP-code give a measure of the total supply of mortgage loans available to be refinanced. I estimate the following set of local projections,

$$\begin{aligned}\Delta_h Refivol_{i,t+h} = \alpha_t^h + \sum_{k=0}^{6} \beta_k^h \cdot \Delta_1 Refivol_{i,t-k} + \\ \sum_{\tau \epsilon T} \gamma_\tau^h \cdot 1_{t=\tau}^h \cdot \Delta_3 Fintech_{i,t-3} + \delta^h \cdot \Delta_3 OutstandingStock_{i,t} \\ + \eta^h \cdot \Delta_3 Fintech_{i,t-3} \cdot \Delta_3 OutstandingStock_{i,t} + \lambda^h \cdot Controls_{i,t} + \varepsilon_{i,t} \quad (4)\end{aligned}$$

for h=1,2,...5. In the above equation, I denote T as the set of all sample dates (i.e. Jan. 2010-Dec. 2019), and the term $1_{t=\tau}$ denotes an indicator function that takes a value of one at time t and zero otherwise. Thus, the above specification suggests that I allow lagged fintech activity to have separate coefficients at each date in the sample, in order to control for time-series variation in the effect of fintech lending as it relates to the interest rate environment. Thus, the interaction term $\Delta_3 Fintech_{i,t-3} \cdot \Delta_3 OutstandingStock_{i,t}$ looks at the strength of fintech lending as a predictor for refinancing growth only as it varies in

[26] As mortgages become highly seasoned, they become increasingly unlikely to pre-pay, as the remaining mortgage balance becomes low relative to the fixed costs associated with refinancing.

[27] That is, the interest rate at origination for each mortgage minus the prevailing 10-year Treasury yield

the cross-section. Intuitively, this specification asks, if we observe an increase in the local supply of mortgages with a refinance incentive in a subset of ZIP codes from time t-2 to time t, whether the estimated effects of recent fintech activity on refinance growth are stronger in these ZIP codes.[28]

Figure 1.2 panels C and D display impulse responses associated with these specifications. They depict the evolution of the interaction effect between fintech lending and the expansion of the local supply of mortgages with a refinance incentive. As the impulse responses illustrate, the association between fintech lending and local demand shocks is positive at all time horizons from 1-5 months. This suggests that when a local mortgage market sees an increase in the number of local homeowners with a refinance incentive, the benefits associated with having a stronger regional fintech presence become more substantial than they would otherwise be. The estimated fintech effect, interacted with the regional refinance incentive, is strongest in the third month after the local demand shock, and decays in months 4-5, remaining positive for the entire time horizon.

4 Identification of Fintech Effects Using a Cross-border Approach

Despite my attempt to control for factors that might simultaneously influence fintech regional fintech lending patterns and housing market activity, endogeneity remains a concern. For example, the use of online lending platforms is stronger amongst sophisticated borrowers who may be more likely to refinance when they have an incentive to do so. Calculating the optimal time to refinance is a complex problem, and an extensive literature suggests that many borrowers make mistakes when doing so, either refinancing too early, or waiting too long to do so (see, e.g. Agarwal, Driscoll, and Laibson, 2013; Deng and Quigley, 2012; Stanton, 1995). Existing research also suggest that borrower characteristics and experience play a role in determining the extent to which borrowers make refinancing mistakes (see LaCour-Little, 1999; Agarwal, Rosen, and Yao, 2013). There is no perfect measure for

[28] The cross-sectional demand proxy boils down to an assessment of whether a) new mortgage origination was high in a ZIP code 30 months ago, and b) whether these borrowers had relatively high rates at origination.

borrower experience or sophistication, making it difficult to control for the possibility that fintech uptake is higher among this class of borrowers.

To better identify a causal connection between fintech lending and credit supply expansion, in this section, I will attempt to isolate a source of exogenous variation in regional fintech presence. To do so, I make use of the staggered timing of market entry by fintech lenders in various state-level mortgage markets. In the first part of this section I will argue that state-level regulatory factors likely played a role in the timing with which firms entered various states. In the second part of this section I will discuss and present the results of my identification scheme, which makes use of comparisons of county-level credit growth amongst counties on opposite sides of state borders, in pairs of states that have differing numbers of active fintech lenders.

4.1 Regulatory Barriers and Fintech Market Entry

My approach for identification makes use of the staggered timing with which fintech lenders entered various state-level mortgage markets. While some fintech lenders were already well established prior to 2010, there were a number of new firms that only began to originate mortgages in 2011 or later. Among the fintech lenders that existed at the start of my sample, a number of these lenders were still fairly new and originated mortgages in only a few states. Over time, new fintech firms emerged, and existing lenders began to expand their regional presence. Figure 1.3 displays an example of how this process unfolded for an individual fintech lender, CashCall mortgage. While CashCall existed prior to the start of my sample in 2010, it had a fairly minimal presence at that time, originating loans only in California and Florida. In the following year, CashCall expanded across most of the western states, and a handful of others. By 2015, it had grown to serve almost all states.

In the aggregate, this process generated variation in the number of active fintech lenders across states, with some states accumulating a large number of fintech lenders very early on, while others saw a relatively late influx of fintech lenders. In Figure 1.4 I show the total

number of fintech lenders active in each state across the same years displayed in Figure 1.3. It is apparent that population appears to play a role in state-level differences, with California, Florida, and Texas maintaining relatively high numbers of active fintech lenders in all years.[29] Nonetheless, there are also more curious patterns. New York, and other states in the Northeast tend to have sparse fintech coverage, despite large populations, while states like Idaho, Oklahoma, and Colorado appear to have heavy fintech concentrations relative to their populations.

At first glance, it may be puzzling that newer fintech lenders did not immediately enter all or almost all state mortgage markets. After all, one advantage of online mortgage lending should be that it reduces the required physical presence of the lenders that possess this technology. In practice, however, there are a few reasons why we might observe this staggered entry across states. First, while online mortgage applications are the bedrock of their business models, a number of fintech lenders do hire loan officers that are available to meet with potential applicants. At a minimum, most mortgage loans still involve some in-person interaction upon closing of the loan. More importantly, unlike banks, which can apply for national charters, non-bank mortgage lenders must apply separately for licenses in each state. Each state has its own requirements for approving new mortgage lenders, with some states more stringent than others.

To examine the timing of fintech market entry in conjunction state-level regulatory factors, I collect information on state licensing requirements for non-bank mortgage lenders from the Nationwide Multistate Licensing System (NMLS). The NMLS is the system of record for non-depository financial services licensing. State regulators allow applicants to submit their application materials through the NMLS portal, enabling firms seeking to apply for licenses in multiple states to utilize a centralized system.[30] Via its website, the NMLS provides checklists complete with each state's individual licensing requirements. Using these

[29] It is relatively intuitive that state population should have some bearing on the location of fintech lending. Even if we are convinced of the theory that states' regulatory environments play a role in lenders' decisions regarding where to originate loans, lenders likely consider both the benefits of entering a state in along with the costs. Once state-level licensing costs are paid by a lender, larger states offer bigger markets in which to originate loans, generating potentially higher revenue per dollar paid in administrative costs.

[30] After a firm submits an application, state regulators can provide commentary and give updates on their application statuses. At the time of writing, all states utilized the NMLS for licensing mortgage loan originators.

checklists, I compile information on the licensing requirements in each state. I then use this information to create four quantitative variables meant to capture the costliness of each state's requirements.

The first of these variables is the dollar value of the application fees associated with submitting a licensing request. These application fees are not, in themselves, likely to constitute a significant deterrent for a firm seeking to enter a state mortgage market. Application fees average only $1,001 and hit a maximum of $3,000. However, since application fees are used to recoup the costs incurred by state regulators when they review an application, application fees are likely a proxy for the length and extensiveness of the review process. Since a non-trivial portion of the review process involves responding to follow-up questions from regulators, application fees are likely an indicator for the time and labor costs associated with the application process that are not captured by the written application requirements.

The second variable consists of the minimum net worth, in dollars, required for mortgage lenders. Like the application fees, net worth requirements seem unlikely to bind for moderately sized firms. Minimum required net worth averages under $100,000 and tops out at $1 million.[31] However, when applying for a license, firms must submit financial statements proving that they meet these minimum requirements, and are subject to periodic review once entering a state. This suggests that the documentation and verification costs associated with proving satisfactory net worth are likely to be higher in states where requirements are stricter.

The third variable is a dummy variable that takes a value of one if a state is a so-called "brick and mortar" state. Brick and mortar states require that mortgage originators maintain a physical branch presence within the state. This requirement runs counter to the business model of many fintech lenders, which focuses on maintaining a minimal branch presence and originating most loans via online platforms. As such, I separate this variable from other qualitative application requirements.

Finally, I take a simple count of other mortgage licensing requirements in each state.

[31] The relatively small average level reflects the fact that a number of states do not have any minimum net worth requirement.

These include such requirements as submitting audited financials, completing criminal background checks and credit history checks for upper management and minority shareholders, maintaining a minimum number of employees certified as licensed (individual) lenders in the state[32], submitting plans for anti-money laundering compliance, and submitting documentation certifying access to a bank-provided line of credit.[33]

To assess the extent to which state-level barriers of entry are able to predict the timing with which fintech lenders enter state mortgage markets, I use a logistic regression approach to estimate the probability that a lender will be active in a state mortgage market as of a given year, t. I use firm-level HMDA data to generate the set of years in which each lender enters a given state's market and estimate specifications which assume that the probability a lender is active in a given state assumes the following form:

$$Log\left(\frac{Prob_{i,j,t}^{Active}}{1-Prob_{i,j,t}^{Active}}\right) = \alpha_t + \beta \cdot log(Population_{j,t-1}) + \gamma \cdot Regulations_j \tag{5}$$

where $Prob^{Active}$ denotes the probability that a firm will operate in state i in year t. The dependent variable takes a value of one if firm i has began originating loans in state j by year t, and a value of zero otherwise. In setting up the data, a firm does not enter into the regression until the first year of its existence so equation (5) is estimated on an unbalanced panel. This specification assesses the strength of the association between a state's regulatory environment and the number of firms that enter that state early on in the sample. If a firm enters a state in 2011, for example, the dependent variable for that firm-state pair would take a value of zero in 2010, and would be equal to one from 2011-2019. If that same firm enters a different state for the first time in 2016, the firm-state dependent variable sequence would be six zeroes (from 2010-2015) followed by four ones (from 2016-2019). Thus, by construction, this specification places heavy weights on the states that received an influx

[32]So-called "qualified individual" requirements vary by state, but generally require that some number of high-level employees either become licensed at the individual level, to become mortgage brokers within the state, or submit documents suggested that they have been certified in another state. These requirements also mandate certification that high-level employees have achieved a certain minimum number of years working in the mortgage industry.

[33]Credit line requirements vary across states, with some states requiring that mortgage originators maintain access to a minimum line of credit of a particular size, while others merely require proof of access to credit. As such, I do not code this as a separate quantitative variable.

of fintech lenders early in their histories. The variable *Regulations*, listed above, denotes a vector containing combinations of the four regulatory variables defined previously.

In Table 1.6, I show the results of estimating equation (5) on various combinations of regulatory variables. Each coefficient is an estimate of how a unit change to a covariate affects the log-odds ratio associated with the probability of state-level market entry. Across all specifications, each of the regulatory variables has a negative coefficient. The strongest regulatory predictor of a firm's probability of early entry in a given state is the state's application cost. This suggests that these application costs are indeed a proxy for the level of time and scrutiny applied to an application by state regulators. The coefficients on the application cost variable suggest that a marginal dollar increase in application costs decreases the market-entry odds-ratio by between 24-34%, a magnitude which would seem inconceivable if application costs did not serve as a proxy for some latent factor. The next strongest effect, in terms of magnitude, is the brick and mortar dummy variable. While including the other regulatory covariates diminishes the level of statistical significance of the brick and mortar variable somewhat, it maintains significance at the 10% level, at a minimum, across all specifications. States with brick and mortar requirements are estimated to have log-odds of firm entry between .17 and .32 lower in a given year.[34] The number of qualitative application requirements also appears to have a substantial predictive effect on the pace of firm market entry into states, with coefficients between -.04 and -.065 across all specifications. This suggests that the addition of more licensing paperwork in the form of financial statement submissions, and background check documents, as well as other requirements, have a substantial effect on the timing with which firms enter state mortgage markets.

4.2 Cross-border Approach and Results

I utilize an identification approach that only makes comparisons across pairs of bordering

[34] The states with brick and mortar requirements are Arizona, Missouri, Nevada, and Texas. In Texas, the brick and mortar requirement is apparently easier to circumvent than in these other states, though I do not explicitly model this distinction.

states. In particular, I focus on comparisons of counties located close to the borders of adjacent states with differential levels of fintech activity. The underlying idea is that by focusing on pairs of bordering states, and on counties located close to their adjacent borders, the impact of regionally correlated but unobserved factors relevant to housing market outcomes should be much less severe than it would be when comparing geographically distant regions. Given state-level licensing requirements, a lender that is licensed to operate in one state, but not another, must "artificially" refrain from lending beyond the border of the state in which it is licensed, inducing a potential discontinuity.

My empirical approach takes the form of a cross-border regression discontinuity framework. The use of state borders as a tool for studying similar regions facing differential policy environments dates back, at least, to Card and Krueger (1994), and state-border regression discontinuity frameworks have been used by Holmes (1998) and by Pence (2006). Unlike these studies, the effects I seek to identify are not static. Rather than studying the effects of a set of policies, per se, I look to study the effects of fintech market entry which arises both from the creation and expansion of new and young fintech lenders, and from static regulatory factors. As such, my study seeks to exploit time-varying differences in fintech presence while controlling for time-invariant differences across states. Recent research that has made use of a state-border scheme alongside time and regional fixed-effects includes Campello, Gao, and Xu (2019), Ljunquist and Smolyansky (2018), and Moretti and Wilson (2017).

To construct my sample I begin by identifying the number of fintech firms operating in each state by year, using HMDA data, as in Figure 1.4. I then identify all pairs of states which share a border and which contain a differential number of fintech lenders in a given year. For each state pair, I label the state with a larger number of fintech lenders as the "treated" state, and the adjacent state with which it is paired as the "control" state.[35] I then identify the set of counties in both the treated and untreated state that lie within a given distance of their shared state border, with the distance between a county and border

[35] Alaska and Hawaii will not appear in the analysis, as they do not share borders with any other states.

calculated using the population centroids of each county. I will estimate specifications where I use distances of 50 and 100 miles from a state border as cutoff points for a county's inclusion in the sample.[36]

Figure 1.4 gives further insight into the construction of the sample. In that figure, I select several state pairs that appear in the sample in various years. Since the sample of treated and control states changes from year to year as fintech firms expand into new states, this diagram should be thought of as illustrating pairs of states at a single point in time. In the figure, the counties that are shaded red denote the set of counties, in treated states, which are located within 50 miles of the state's border with its paired control state. Counties highlighted in gray represent control state counties within 50 miles of the shared border. Each county that appears in the sample references one or more specific border. For example, counties in the northeastern corner of Texas are treated counties that reference the Texas-Oklahoma border. Counties in the Southwestern part of Texas reference the Texas-New Mexico border. Some shaded counties in the Northwest panhandle of Texas are located within 50 miles of both Oklahoma and New Mexico. Such counties would appear in the sample twice.[37]

The key assumption underlying my identification scheme is that unobserved factors that influence mortgage lending activity in border counties of adjacent states are continuous (as a function of border distance) across state borders, and are time invariant. This implies that fintech firms do not choose to enter one state, and not its neighboring state, for reasons related to differential lending opportunities in the border counties of those states. My empirical approach is insensitive to large unobserved differences between non-neighboring states (since it only compares adjacent states), to differential lending opportunities in adja-

[36] Given this methodology, a single state can be in both the treated sample and in the set of "control" (or paired) states if it shares a border with one state that has fewer licensed fintech firms than it does, and shares a border with another state with more fintech lenders. However, the set of counties within that state that reference those two state borders will, in general be different, though some overlap is possible.

[37] It is also worth noting that some counties in western states are located on state borders but remain unshaded. By and large, this is because county distances are counted using population centroids of counties, and these unshaded counties have large towns located more than 50 miles from the state border. Loving County, in southwestern part of Texas, next to New Mexico, is an example of a county excluded from the sample for a different reasons. Namely, it is small enough that it does not any have mortgage lending transactions in a number of sample years, and is dropped from the sample.

cent states outside of border counties,[38] and even to average differences in adjacent states' border counties that do not change over time. The empirical specifications I estimate will assume the following form, where subscripts i, j, b, and t index counties, states, borders, and time respectively:

$$\Delta_1 Refivol_{i,j,b,t} = \alpha_t + \sum_{\nu \epsilon B} \beta_\nu \cdot 1_{b=\nu} \cdot distance_{i,b} + \\ \gamma \cdot Treat_{j,t} + \delta \cdot Treat_{j,t} \cdot Rates_t + \lambda \cdot Controls_{i,t-1} + \mu_{(j,b)} + \varepsilon_{i,j,b,t} \quad (6)$$

In the specifications above, the variable *Treat* assumes a value of one if county i is located within a state, j, which is a treated state with respect to the border, b. The *Rates* variable is as defined in equation (2). I let B denote the set of all state borders in the sample. The term *distance* denotes a county's distance from border b and the distance variable soaks up unobservable differences between counties that vary as a function of their distance from the border, with a separate distance coefficient estimated for each state pair. In versions of the equation (6) regressions that use 100 mile bandwidths for sample selection, I will also include a squared border distance term. The term $\mu_{(j,b)}$ denotes a set of fixed effects for borders or for states, which will vary across the set of regressions that I estimate.

I will also make one additional sample refinement in some of the specifications I estimate. Specifically, I will display results of regressions in which I exclude border counties that contain large metropolitan areas (defined as any of the top 100 metropolitan areas, by population, residing within 50 miles of a state border). Doing so potentially improves sample selection by excluding counties that are likely to have very different characteristics than other counties near the border, and thus may improve the level of comparability between counties located in the treatment and control groups in a given year.

In Table 1.7 I show results from estimates of various versions of equation (6). The coefficients on the *Treat* variable suggests that in an average interest rate environment,

[38] For example, this methodology would not lead to bias if a fintech firm chose to enter one state and not another neighboring state for reasons having to with strong lending opportunities in an urban area outside of border counties.

refinancing growth is between 1% and 3.3% stronger in areas with a larger number of active fintech lenders. The largest coefficients on the *Treat* variable are obtained in specifications that exclude the largest metro areas from the set of border counties. Much of this increase in the treatment effect in these specifications comes from the exclusion of the New York City metropolitan area which had particularly strong refinancing activity in some years. The *Treat*Rates* coefficients suggest that the treatment effect of having a larger number of fintech lenders than adjacent states increases in declining interest rate environments. A 1% widening of mortgage spreads is estimated to increase the treatment effect by between 1.4% and 6.5%, effects which are statistically significant at the 1% level across all specifications. The exclusion of large metropolitan areas tends to diminish the estimated size of the effect of the interaction between interest rates and fintech presence, as the interaction effect is roughly four percentage points larger in specifications which include the full set of counties. This suggests that some of the large fintech refinance effects in years when interest rates decline are driven by high refinancing activity in dense urban areas.

5 Does Fintech Credit Expansion Boost Local Economies?

When interest rates decline, the ability to refinance has the potential to increase the wealth of existing mortgage borrowers by allowing them to reduce their interest costs, extract equity from their homes when it is relatively attractive to do so, or extend the maturity of their loans and reduce their monthly payments. In this section, I look to assess whether, by inducing more borrowers to refinance, fintech lenders promote subsequent gains in household spending in regions where they are active. Existing literature has uncovered a robust link between mortgage credit, refinancing, and household consumption plans (e.g. Campbell and Cocco, 2003; Hurst and Stafford, 2004; Koijen, Van Hemert, and Van Nieuwerburgh, 2009). However, it is not self evident that fintech lending will spur stronger consumption growth. The speed and intensity of consumption gains in the wake of an expansion in home refinance credit depends on the regional distribution of home equity (Beraja et al., 2019), home prices

(Mian, Sufi, and Rao, 2014), income, and marginal propensities to consume (Agarwal et al., 2020; Mian and Sufi, 2011). It is thus important to investigate the empirical association between fintech lending and household spending to shed light on the macroeconomic effects of fintech presence.

If fintech lending generates consumption effects as a result of refinancing activity, such spending increases may bolster other economic outcomes. If wealth gains stemming from refinancing activities lead consumers to spend more on goods and services produced within their local area, then we may see local employment growth and business expansion as a side-effect of stronger consumption activity. For example, Mian, Sufi, and Verner (2020) find that credit supply shocks that fuel consumer credit growth foster employment gains and business expansion in local non-tradable goods. In the final part of this section I will evaluate the correlation between fintech credit expansion and local business growth.

5.1 Fintech Lending and Local Retail Spending

In order to investigate whether fintech lending has an effect on consumption, I will need a measure of consumer spending that is likely to capture the behavior of local consumers (i.e. those who reside in a given county) so that it can plausibly be linked to local mortgage refinancing activity. I will also need to consider spending categories which are likely to see an uptick in demand amidst a wave of mortgage refinancing. Some prior studies on local consumption effects of housing outcomes have focused on automobile expenditures (e.g. Mian, Sufi, and Rao 2014; Agarwal et al. 2020) and on broad measures of credit card spending. While large durable expenditures on automobiles are a plausible response to home refinancing for those who take out large amounts of home equity, those who refinance merely to lower their monthly mortgage payments are unlikely to save enough to afford a car that they otherwise would not have purchased.

In order to capture new spending that might plausibly arise from refinancing activity, I instead focus on a measure of local retail expenditures. The measure of retail spending

that I create comes from records of purchases at a large sample of mass market retail establishments that upload their purchases to a centralized database. Mass market retailers, (commonly referred to as "big-box" retailers) sell a wide variety of general purpose consumer goods such as home furnishings and kitchen supplies, clothing, electronics, automotive accessories (e.g. motor oil), school supplies, beauty products, and many other items. In comparison to measures of spending that focus on large durable consumption goods, local retail expenditures are more likely to respond elastically to relatively modest increases in income.[39]

I base my measure of local consumption on information from Nielsen's Retail Scanner dataset. Nielsen collects weekly sales information from a panel of over 30,000 national retail store locations. The dataset is referred to as Retail Scanner data because the dataset is populated when items are brought to a check-out counter and their product codes are scanned via a laser scanner. The dataset is populated with highly granular information the products purchased, allowing for the calculation of sales volume by product categories. Importantly, the database reports the geographic location of each store in the panel, and each purchase is indexed by its individual store location, allowing for precise identification of where each purchase took place.[40]

The key outcome variable in my analysis of local consumption is the log change in local retail spending from time t to time t+1. In order to build a meaningful measure of consumption growth for a given year, t+1, I begin by identifying the sample of stores that operate within a county. Since the Nielsen panel changes from year to year, I consider only stores that were active in both year t, and in year t+1, and I further refine this sample by keeping only stores that reported at least 26 weeks of data in both years. Keeping only sets of stores that are active in both years allows for an assessment of expenditure growth which is unaffected by changing panel composition. Further, I annualize spending at each store, thereby eliminating potential sample biases resulting from stores that report data less frequently than others. I also consider only counties that contain five or more retail store

[39] Moreover, in contrast to broader measures (e.g. credit card expenditures) which assess purchases at a broader array of merchants (e.g. in the hotel and tourism sector, and specialty/niche retailers) local mass market retailers appear, intuitively, to generate large portions of their expenditures from consumers living outside of the local areas that they serve. Most counties of sufficient size have their own "big box" retailers. It is unlikely that consumers would have to drive far in order to find a purveyor of generic household goods.

[40] To my knowledge, this is the first study that has used Nielsen scanner data as a broad proxy for local consumption. Other studies have used Nielsen data as a proxy for local spending on certain product categories. See, e.g. Cotti et al. (2021)

locations in an attempt to minimize the extent to which results in a given county might be unrepresentative of true county-level retail purchase activity. For verification, I ensure that all of my consumption results are robust to considering a smaller sample constructed only from retail store locations that are active for all years of the study. Doing so ensures that results are not driven by the changing sample composition across years.[41]

I build two measures of county-level retail expenditure growth using the Nielsen data. The first measure, which I refer to as "total retail spending" (or *TotalRetail*) consists of spending within all product codes outside of food/groceries and pharmacy/medicine. I exclude these two categories of spending, as food and medicine are appear to be the categories least likely to respond elastically to changes in income.[42] I refer to my second measure of retail expenditures as "discretionary retail spending" (or *DiscretionaryRetail*); to construct it, I begin with my measure of total retail spending and remove personal hygiene/toiletries, school supplies, car maintenance and repair products, and home maintenance and repair products. The idea behind this measure is to further refine the definition of retail spending to include categories of purchases most likely to respond elastically to modest changes in consumer income.

To examine the correlations between fintech activity and consumption growth I return to the functional form of my baseline analysis and ask whether consumption growth is stronger in the year after a fintech-led wave of refinancing activity. Specifically, I estimate

$$\begin{aligned}\Delta_1 Spending_{t+1} = \alpha_t + \beta \cdot Fintech_{i,t-1} \\ + \gamma \cdot Fintech_{i,t-1} \cdot \Delta_{avg} Rates_t + \delta \cdot Controls_{i,t-1} + \varepsilon_{i,t+1} \\ Spending \, \epsilon \, \{TotalRetail, \, DiscretionaryRetail\}\end{aligned} \tag{7}$$

relative to the set-up of equation (2), the only distinctions are the change to the dependent

[41] I will include year fixed-effects in all regressions. However, part of the analysis asks about whether the effects of fintech activity are stronger in years in which interest rates decline. In principle these coefficients could be biased if strong spending growth in a particular year is driven by a large influx of new panel participants (i.e. retail stores) that were active in two consecutive years (so as to be included in the calculation for consumption growth in that year) but not in other years.

[42] Of course, it is plausible that consumers could substitute to more expensive categories of food and to be less sensitive to sales/coupons, etc. However, I remain comfortable with the assumption that grocery purchases are relatively inelastic compared to other retail categories.

variable, and the time horizon, where I now look at the year after an interest rate shock. The idea underlying this is that for borrowers that refinance, it will likely take some time for the effect of lower monthly payments to outweigh the initial outlay of refi fees paid to the lender up front.

Table 1.8 displays the results of estimating equation (7) using the *FintechCount* measure for fintech market activity. In Appendix Table 1.16, I display analogous results with the *FintechShare* measure of fintech activity. Results for both measures of consumption are shown. Results are fairly consistent across both measures of consumption growth. Depending on the inclusion or exclusion of county fixed-effects, the results suggest that the presence of an additional fintech lender at time t-1 forecasts between .09% and .2% stronger consumption growth. Coefficients on the *FintechCount* variable are statistically significant at the 1% level across all specifications. Given previous results on refinancing, suggesting that the marginal addition of a fintech lender corresponds to between .4%-.8% stronger refinancing growth, the results in Table 1.8 suggest that consumption gains are between 10% and 50% as large as the initial refinancing surge. While the high end of this estimate seem excessively large, prior research has suggested that consumption growth effects can be large when homeowners extract equity from their homes (e.g. Mian and Sufi, 2011; Agarwal et al., 2020).

The interest rate interaction terms are also large, suggesting that when refinancing is particularly high, in the wake of widening spreads between outstanding mortgages and prevailing market interest rates, consumption growth is larger in the year after this wave of refinancing occurs. Coefficient estimates suggest that a 1% widening of rate spreads increases the fintech effect by between .18% and .23%. This suggests that the effect of fintech presence becomes two to three times stronger in a falling interest rate environment. The *FintechShare* results tell a largely similar story, though the same patterns only hold in regressions that contain county fixed-effects.[43]

[43] Appendix Table 1.16 suggests that a 1% increase in the refinance market share of fintech lenders at time t-1 predicts between .04% and .31% stronger retail consumption growth at time t+1. Interest rate interaction terms are inconsistently signed across specifications, with values ranging between -.05 and .07, with positive interaction effects obtaining in specifications with county fixed effects.

5.2 Fintech Lending and Local Business Growth

If the presence of fintech lending brings about an increase in consumption among those who refinance, and if these consumers tend to spend locally, then the positive consumption effects of fintech-induced refinancing might also bring about positive outcomes for regional businesses. The effects of these local spending surges would likely be most impactful for small businesses that cater to local consumers. Larger businesses with a broader reach, and those in industries like tourism and manufacturing that have customer bases outside of the areas in which their production is located, would likely gain little from a boom in local consumption. Thus, to investigate whether there is an association between fintech activity and local business activity, I focus on small businesses operating within the non-tradable sector. According to Bahadir and Gumus (2016) and Mian, Sufi, and Verner (2020) businesses in the non-tradable sector should be most affected by credit expansions that target households, as is the case in this study.

My primary source of information on small business activity comes from the US Census Bureau's County Business Patterns (CBP) database. The database contains employment and payroll information at the county level for a number of industry groups, as classified by the North American Industry Classification System (NAICS) codes of each industry group. I follow the industry classification procedure of Mian and Sufi (2014b) in order to identify industries that belong to the non-tradable sector. I further refine this group of businesses to include only establishments with 100 or fewer employees. The key outcome variables I will focus on within the CBP data are total employment and the number of small business establishments in a county. If local consumption grows stronger as a result of a positive wealth shock to home owners, local businesses that serve these consumers may expand, bolstering employment. Stronger consumption may also encourage new businesses to enter a local market, or forestall the demise of businesses on the margin of failing, leading to a greater number of small business establishments.

Stronger local consumption may also fuel the level of investment by small businesses, as some of these businesses expand to keep up with growing demand. To generate a viable

proxy of small business investment at the county level, I utilize data on small business lending from Community Reinvestment Act (CRA) disclosures. I focus on the total volume, and the total number of loans granted to businesses with assets under $1 million.

To determine whether there is evidence consistent with the notion that fintech credit expansion fosters stronger growth for small businesses, I run analogous specifications to the consumption growth models of the previous sub-section, with the only modification being the left-hand side outcome variable. I therefore estimate

$$\begin{aligned}\Delta_1 Smallbus_{t+1} &= \alpha_t + \beta \cdot Fintech_{i,t-1} \\ &\quad + \gamma \cdot Fintech_{i,t-1} \cdot \Delta_{avg} Rates_t + \delta \cdot Controls_{i,t-1} + \varepsilon_{i,t+1} \\ Smallbus &\,\epsilon\, \{EmploymentNonTradable,\, EstabNonTradable,\, LoanCount,\, LoanVol\}\end{aligned} \tag{8}$$

where *EmploymentNonTradable* gives non-tradable sector employment, *EstabNonTradable* denotes the number of small business establishments in the non-tradable sector, *LoanCount* gives the total number of small business loans to firms with under $1 million in total assets, and *LoanVol* gives the total dollar volume of such loans. All variables are expressed in log form, so that growth-rates are measured as log differences.

Table 1.9 displays the results of these regressions, with coefficients on the *FintechCount* variables. The results suggest that small business activity tends to pick up in the year after a fintech-induced refinancing surge. Coefficients on the *FintechCount* variable in columns (1)-(4) suggests that the presence of an additional fintech lender at time t-1 predicts between .3% and .6% stronger growth in the number of small business establishments in the non-tradable sector at time t+1 and between .3% and .4% stronger growth in small business employment. The interaction of *FintechCount* with widening mortgage interest rate spreads is also positive and statistically significant in these specifications. The interest rate interaction term coefficients in these regressions tend to be stronger in specifications that include county fixed-effects. In these specifications, in columns (2) and (4), a 1% widening in interest rate spreads amplifies the fintech effect by .9% and 1%, respectively, for small

business establishment growth and employment growth.

Columns (5)-(8) suggest that small business lending is stronger in regions with a more concentrated fintech presence. An additional fintech lender at time t-1 forecasts between .3% and .5% stronger growth in the total number of small business loans extended between time t and t+1. Interaction terms have positive and significant coefficients in all four of these specifications, suggesting that small business lending growth is stronger in years that follow large expansions of fintech-induced refinance credit. To the extent that small business lending is a proxy for the level of investment undergone by these businesses, the results in columns (5)-(8) suggest that small local businesses invest more in places in which fintech lenders expand the supply of credit.

6 Does Fintech Lending Help Fed Policy Reach Underserved Communities?

In the last portion of this paper I ask whether the effect of fintech lending on the transmission of monetary policy varies across geographical regions and communities. If fintech lenders affect markets by alleviating microeconomic frictions faced by potential borrowers, then fintech lending may have the strongest impact on monetary transmission in places in which these frictions are most binding.

One market imperfection that may be improved by the presence of fintech lenders is the poor access to credit extended to minority borrowers (see, e.g. Bayer, Ferreira, and Ross, 2018; Bostic, 1996; Browne et al., 1996; Cheng, Lin, and Liu, 2015; Ghent, Hernandez-Murillo, and Owyang, 2014). Discriminatory effects can arise from cognitive biases, from a so-called "taste for discrimination" among White loan originators (Becker, 1957), or from the use of race as a proxy for economically important traits for which obtaining reliable information is costly (Ladd, 1998). Recent research by Bartlett et al. (2019) suggests that fintech lenders are less discriminatory than other categories of lenders. They find, broadly, that loans to racial and ethnic minorities carry higher interest rates than loans extended to similarly situated White borrowers, but that these interest rate mark-ups are lower on

fintech loans. Moreover, they find no evidence that fintech lenders reject minority applicants at a higher rate than White borrowers. They interpret this latter finding as evidence that fintech lenders alleviate cognitive biases of human loan officers.[44]

Fintech lenders may also help overcome borrowing constraints in geographic regions where borrowers have limited access to the traditional financial system due to the scarcity of brick-and-mortar bank branches. Existing research suggests that potential borrowers' local access to financial services is an important determinant of real outcomes (e.g. Burgess and Pande, 2005; Cetorelli and Strahan, 2006; Gilje, 2019; Jayaratne and Strahan, 1996). Rural areas with sparse populations often have few bank branches, and the transportation and time costs associated with going through the face-to-face mortgage lending process at physical branch locations are likely to be higher for residents of these areas. Recent research by Erel and Liebersohn (2020) on small business loans have suggested that fintech lenders are particularly active in regions with little access to the traditional banking system and are more likely to expand the overall supply of small business credit in these regions.

Given the research on fintech lenders and access to finance, it is plausible that fintech lenders broaden the Fed's reach by allowing the refinance channel of monetary policy to operate efficiently in areas where financial frictions dampen its transmission. In the remainder of this section, I will test these hypotheses by looking within and across counties. I will ask whether, within a county, minority populations gain more access to credit when fintech firms enter, and whether this association is stronger during monetary expansions. I also ask, across counties, whether the effect of fintech lending on credit growth and retail spending is stronger in sparsely populated areas, regions with few bank branches, and regions with significant minority populations.

[44]Since GSE guarantees eliminate credit risk in the conforming segment of the mortgage market, the authors argue that higher rejection rates for minority applicants among non-fintech lenders are likely inconsistent with profit maximizing behavior. They cite the fact that minority borrowers continue to pay relatively high interest rates (as compared to White borrowers) as evidence that fintech algorithms may use variables correlated with race strategically, to proxy for lenders' likely market power. For example, variables correlated with race may inform borrowers' likelihood to shop around among many different lenders, or proxy for the likelihood that the borrower lives in a so-called "financial desert" with limited access to financial services.

6.1 Within-county Analysis of Fintech Lending and Credit Composition

I first turn to my within-county analysis of the effects of fintech lending. I look at the association between fintech market presence and the composition of refinance credit. If fintech lenders have a particular advantage in screening racial and ethnic minority applicants, then the presence of fintech lenders should allow more of these borrowers to obtain credit. This effect should be strongest in declining interest rate environments when financial technology is most beneficial. To investigate this possibility, I make use of the county-level HMDA dataset, breaking down total refinance loans by loan-type. In particular, I sort total refinance loans within a county by the borrower's race and ethnicity. I also examine whether riskier loan segments (FHA loans and loans secured by junior liens) expand alongside fintech entry. I estimate the following set of within-county regressions:

$$\begin{aligned}\Delta_1 Reficomposition_{i,t} = \alpha_{1,t} + \alpha_{2,i} + \beta \cdot Fintech_{i,t-1} \\ + \gamma \cdot Fintech_{i,t-1} \cdot \Delta_{avg} Rates_t + \delta \cdot Controls_{i,t-1} + \varepsilon_{i,t}\end{aligned} \tag{9}$$

where *Reficomposition* is the percentage of total home refinance loans (by volume) that to go borrowers within a particular category. The coefficients $\alpha_{1,t}$ and $\alpha_{2,i}$ denote the presence of year and county fixed effects, respectively.

I first look at how fintech lending co-varies with the racial composition of borrowers. I examine the association between fintech lenders, and the share of non-white borrowers among all refinance credit within a county. I label a loan as going to a non-White borrower if the primary applicant on the loan lists a race other than White/Caucasian on his or her loan application.[45] For the purposes of this analysis, I exclude borrowers who do not list a race, or who are listed as belonging to "some other race."[46] I exclude those in the "some other race" category due to evidence suggesting that a substantial percentage of those identifying

[45] Mortgage applications can have multiple co-applicants. I consider only the race of the primary applicant.

[46] Thus, I ignore the phenomenon documented in Agarwal et al. (2020) that a larger percentage of fintech borrowers do not list their race on their loan applications, presumably because racial anonymity is easier to achieve via an exclusively online approval process. While I imagine that this phenomenon is likely to lead to an understatement of the shift toward non-White borrowers when fintech firms enter a market (due to the presumed greater incentive of non-White borrowers to remain anonymous), this is speculative.

as belonging to this category also identify as Hispanic/Latino, a group which I will study separately.[47] Borrowers who identify as Black or African American make up the plurality of the non-White racial group and appear to drive the results discussed below.

I display the associations between fintech lending, interest rates, and credit composition in Table 1.10. I display a similar table, presenting results of an analysis that uses *FintechShare* rather than *FintechCount* in (7). In column (1) I present the results of the analysis where *RefiComposition* is the non-White share of refinance loans. The coefficient on the *FintechCount* variable is .002, suggesting that within a county, a unit increase in the number of active fintech lenders at time t-1 forecasts a .2% increase in the share of refinance loans originated to non-White borrowers in the subsequent year. This effect becomes stronger if interest rate spreads widen at time t. Conditional on a 1% widening of mortgage interest rate spreads at time t, the estimated effect of a marginal increase in fintech lenders rises by .0049. That is, a unit increase in the number of active fintech lenders at time t-1 would instead be associated with a .69 percentage-point increase (=.2%+.49%) in the share of mortgage loans going to non-White borrowers at time t.

I next look at analogous results for Hispanic or Latino borrowers. I display these results in column (2) of Table 1.10. In a similar fashion to the results for non-White borrowers, the results in column (2) suggest that a strong fintech presence precedes a sharp increase in the share of Hispanic or Latino-identifying borrowers who receive refinance loans. The *FintechCount* variable in this regression is .0017 suggesting that an additional fintech lender is associated with an increase, by .17%, in the share of loans within a county that go to Hispanic or Latino borrowers. The interaction of *FintechCount* with *Rates* has a coefficient of .0028 suggesting that a widening of interest rate spreads by 1% predicts a .28 percentage point amplification of the fintech effect.

I next consider the possibility that fintech lenders change the composition of lending in favor of riskier categories of loans. Buchak et al. (2018) suggest that non-bank lenders[48]

[47] See, for example, https://www.npr.org/2021/09/30/1037352177/2020-census-results-by-race-some-other-latino-ethnicity-hispanic

[48] This applies to non-bank lenders generally, rather than to fintech lenders specifically.

may have a comparative advantage in riskier loan segments due to the regulatory advantage they enjoy over depository institutions. I thus consider the possibility that FHA-loans, which are issued to lower income borrowers, are facilitated by increased fintech presence. I also consider the possibility that loans secured by junior liens on a property also see stronger origination upon fintech entry. Columns (3) and (4) display results of the FHA-share and junior lien regressions, respectively. FHA loans exhibit inconclusive results in the sense that the *FintechCount* and interaction coefficients display opposite signs. The fintech coefficients in the junior lien regressions are both positively-signed, suggesting a positive association between fintech lending and riskier loans. Since junior loans against junior liens are associated with borrowers extracting equity from their houses, these results suggest that fintech lenders may help borrowers take advantage of their home equity for the purposes of consumption, particularly in falling interest rate environments when the terms associated with doing so are likely to be favorable.

6.2 Do Fintech Lenders Amplify Monetary Policy in Underserved Areas? Assessing the Cross-county Evidence

I next address the question of how the effects of fintech lending vary across regions according to the demographic make-up and geographical features of these areas. I first assess whether estimated effects of fintech firms on credit growth are stronger in regions with larger shares of racial and ethnic minorities and in areas with sparser populations or fewer bank branches. I then ask whether these same regions see stronger consumption growth in the wake of these refinancing booms. To uncover the cross-regional variation of fintech effects, I will estimate regressions of the form

$$\Delta_1 Refivol_{i,t} = \alpha_t + \sum_{j=1}^{4} \beta_j \cdot 1_{i\epsilon Q_j} \cdot Fintech_{i,t-1} + \sum_{j=1}^{4} \gamma_j \cdot (1_{i\epsilon Q_j} \cdot Fintech_{i,t-1} \cdot \Delta_{avg} Rates_t) + \delta \cdot Controls_{i,t-1} + \varepsilon_{i,t} \quad (10)$$

where the term $1_{i\epsilon Q_j}$is an indicator function that takes a value of one if county i is in the jth quartile of the distribution of counties, as sorted by a particular demographic or geographic trait, and zero otherwise. Thus, I display fintech effects across each quartile of the population to show how the strength of fintech effects varies across regions.

The retail consumption data are not available for the full set of counties included in the HMDA panel. The retail consumption data covers roughly 30% of US counties, most of which are on the larger end of the population distribution. Thus, rather than divide this sample into quartiles according to county-level traits, I instead extend the diff-in-diff specification from equation (2) into a triple-diff framework that looks at the interaction between fintech effects and regional traits. These regressions will take the form

$$\begin{aligned}\Delta_1 Spending_{i,t+1} = \alpha_t + \beta \cdot Trait_{i,t-1} + \gamma \cdot Fintech_{i,t-1} + \delta \cdot Fintech_{i,t-1} \cdot \Delta_{avg} Rates_t + \\ \eta \cdot Fintech_{i,t-1} \cdot Trait_{i,t-1} + \theta \cdot Trait_{i,t-1} \cdot \Delta_{avg} Rates_t + \\ \lambda \cdot Fintech_{i,t-1} \cdot Trait_{i,t-1} \cdot \Delta_{avg} Rates_t + \mu \cdot Controls_{i,t-1} + \varepsilon_{i,t} \qquad (11)\end{aligned}$$

where the *Spending* variable is as defined in equation (7) and *Trait* refers to one of the demographic or geographic characteristics on which I sort counties. In these specifications, the η and λ coefficients will be of primary interest. They will reveal how the effect of fintech presence, and the interaction between fintech lenders and interest rate shocks varies according to the characteristics of individual counties.

I first turn my attention to the effects of fintech market presence on credit growth, as it varies across counties, and ask whether estimated fintech effects are stronger in counties with larger racial minority populations. Analogously to the previous section, I sort counties based on their shares of White residents. In Table 1.11, in the first column, labeled "% White," I show the results of estimating equation (10) on the HMDA sample sorted on the basis of counties' White population shares. The *FintechCount* coefficients in that table show that the estimated effect of fintech lending on refinance credit growth is larger in counties

where White residents make up a smaller portion of the population. The *FintechCount* coefficient for the quartile of counties with the smallest White population share attains a value of .006, suggesting that a unit increase in active fintech lenders predicts .6% stronger refinance credit growth in counties with the fewest White residents. This coefficient remains fairly stable across the first three quartiles of the distribution of counties and declines substantially, to .001, for the top quartile. This phenomenon is mirrored across the set of *Fintech*Rates* interaction terms. The interaction coefficients decline monotonically across the county distribution, however the degree of this decline is fairly mild until reaching the top quartile of the population. The interaction effects decline from .013 to .005 from the first to fourth quartiles. The differences between first and fourth quartile coefficients are -.005 and -.008 for the *FintechCount* and interaction coefficients, respectively; the first of these differences is significant at the 5% level, while the latter attains only marginal statistical significance at the 10%. The results are broadly supportive of the notion that the presence of fintech lenders in a local market is more meaningful in markets with a larger proportion of minority borrowers. However, differences are not meaningful between counties outside of the top quartile.[49]

I next examine how fintech effects vary according to the Hispanic or Latino population of a local market. Column (2) of Table 1.11 displays these results. In this regression, the first quartile denotes the set of counties with the smallest Hispanic or Latino share of the population. From the first quartile to the top quartile the *FintechCount* coefficients grow from .003 to .006, a difference which is statistically significant at the 1% level. A similar pattern is observable in the interaction coefficients, which rise from .002 to .009 from the first to fourth quartiles, suggesting that the fintech-induced transmission of interest rate shocks to credit growth is stronger in counties with a higher Hispanic/Latino population share. I interpret these results as broadly consistent with the notion that part of the advantage of fintech lenders in propagating the effects of monetary policy lies in their ability to overcome

[49] Many of these counties are likely to be located in rural areas with smaller populations. While I control for population and population density on the right-hand side of all regressions, it is unclear whether these counties differ from more diverse counties along other dimensions than race.

discrimination or cognitive biases that hamper lenders.

I next ask whether fintech lending can help Fed policy reach rural areas, and counties that are poorly connected to the banking system. In column (3) of Table 1.11 I examine how fintech credit growth effects vary by population density. For both the *FintechCount* and interaction coefficients, a monotonic pattern is observable across the distribution of counties, with fintech effects achieving their maximum potency in the most sparsely populated counties. Coefficients on the *FintechCount* variable steadily drop from .010, in the most sparsely populated areas to .006 in the densest counties, suggesting that the amount of additional credit growth predicted by a marginal increase in active fintech firms drops by .4% as we move from sparsely populated to densely populated counties. Meanwhile, the interest rate interaction terms drop from .018 to .011 from Q1 to Q4. The differences between the Q1 and Q4 coefficients are significant at the 1% level for both the count and interaction coefficients.

I display the results of regressions that sort counties by the accessibility of bank branches in columns (4)-(5) of Table 1.11. Column (4) displays results where counties are sorted by branches per-capita. The availability of fintech mortgage lenders may be more valuable in counties with fewer branches per capita if the small number of branches signals a greater likelihood that bank staff will become unable to process all outstanding mortgage applications in a timely manner when demand is high. The coefficients in column (4) appear consistent with this hypothesis. The *FintechCount* coefficient progresses from a high of .009 in counties with the smallest number of branches per capita to .001 (and a statistically insignificant coefficient) in counties with the smallest number of fintech lenders. The difference between coefficients in the first and fourth quartiles are significant at the 1% level. This suggests that in counties in which borrowers have many alternatives to fintech lenders and traditional finance is readily available, the expansionary effects of fintech lenders on the credit supply are muted. Results are similar when counties are sorted by the number of branches per square mile.[50]

[50] However, in this specification, the interaction between fintech presence and interest rate spreads is strongest in the the set of counties with the largest number of branches, on a county-size adjusted basis, a result which is not consistent with the

In the final part of this analysis, I look at whether retail spending growth patterns are consistent the credit growth results. If the aggregate results with respect to retail spending growth, presented in section 5, are indeed driven by the effects of refinancing on borrower wealth, then we might expect consumption growth to be strongest in the counties where refinance growth is strongest. An important caveat to this, however, would be that consumption patterns should only mirror refinance growth if the demographic and geographic traits upon which I sort counties also have no bearing on borrowers' marginal propensities to consume. For example, if sparsely populated rural areas contain residents that are more frugal and less likely to consume out of positive income shocks, then the consumption response could theoretically be muted in these counties, even if fintech firms do cause credit expansions.

With this caveat in mind, I turn to the results of the consumption analysis, with results displayed in Table 1.12. I first assess whether consumption growth is stronger in areas with smaller White population shares. The results of this analysis are shown in column (1). In this specification, only the coefficient on the *FintechCount*Rates*White* variable is suggestive of a negative relationship between a county's White population share and its consumption growth. That is, in an average interest rate environment, it appears that the racial make-up of a county has no impact of the strength of fintech presence on future consumption growth (given a positive but insignificant coefficient of .0003 on the *FintechCount* variable). However, a 1% increase in interest rates is estimated to generate a negative fintech effect (with a coefficient of -.0088 on the triple-diff coefficient).

Mirroring the refinance results, the estimated fintech effect on consumption growth also appears to be stronger in areas with large Hispanic populations. The second column of Table 1.12 displays an interaction coefficient between *FintechCount* and a county's Hispanic population of .0011, suggesting that consumption growth related to a marginal increase in the number of active fintech lenders is .11 percentage points higher in a county with a 100% Hispanic population than in a county with a 0% Hispanic share. This disparity becomes more

notion that fintech firms transmit monetary policy in areas where traveling to a branch is most costly.

dramatic in falling interest rate environments. The coefficient on the triple-diff interaction term, *FintechCount*Hisp.*Rates* is .0132. This suggests that the strong effects that fintech lenders have on credit growth in counties with large shares of Hispanic borrowers also find their way into spending growth.

In columns (3) and (4) I show how the consumption growth effects of fintech lenders co-vary with bank branch presence. Column (3) displays the branch per capita results. The two interaction coefficients (*FintechCount*Brnch./Pop.* and *FintechCount*Rates*Brnch./Pop.*) detail how the estimated effect of fintech presence changes as we move from counties with zero bank branches to counties with a bank branch for every resident (which of course do not exist). The first of these interaction terms has a coefficient of -.0067, which suggests that moving from a county without branches to one that is fully saturated with branches induces a weakening by .67 percentage points of the effect of a marginal increase in fintech lenders on retail consumption growth. The triple-diff interaction coefficient attains a value of -.0256, suggesting that the differential effect of fintech lending on consumption growth in counties with few bank branches is more pronounced in the year following an interest rate shock. This suggests that fintech lenders are particularly adept at amplifying the stimulative effects of Fed policy in counties where individuals have less access to brick-and-mortar bank branches. These results are mirrored, in column (4), by the interaction regressions sorted by branches per square mile. Again, the interaction terms are negative and statistically significant at the 1% level, suggesting that fintech lending effects on consumption growth are stronger in areas where residents have to travel long distances in order to reach a bank branch.

7 Conclusion

Fintech lenders facilitate the transmission of monetary policy. In falling interest rate environments, when mortgage borrowers have strong incentives to refinance, the presence of fintech lenders increases the aggregate availability of credit. A high regional concentration

of fintech lenders also predicts stronger retail spending growth in the year after a widening of interest rate spreads. I find evidence linking the expansionary effects of fintech lending to the ability of these lenders to alleviate various microeconomic frictions in credit markets. The fact that fintech lending predicts rapid credit growth, within 1-3 months after an interest rate shock, gives credence to the idea expressed by Fuster et al. (2019) that fintech lenders alleviate capacity constraints when loan demand is particularly strong, perhaps because automated technology is less easily overwhelmed than human loan officers.

However, the regional patterns associated with fintech monetary amplification suggest that this is not the entire story. The ability of fintech lenders to expand the supply of credit is more pronounced in areas where residents are poorly served by the traditional banking system, either because there simply aren't many physical bank branches, or because traditional banks are worse at accurately evaluating residents' creditworthiness. Thus, the macroeconomic evidence on fintech-induced credit growth is consistent with evidence uncovered by Buchak et al. (2018) and Bartlett et al. (2019). Importantly, these findings suggest fintech lending improves the Fed's ability to stimulate the economy in a downturn, particularly in areas where households' responses to monetary easing are typically muted, due to poor access to financial services.

As fintech lending and the mortgage market evolve, time will tell exactly how the changing technological and institutional environments in this market affect economic policymakers. There has yet to be a true financial crisis in the fintech era, making it unclear how this new class of intermediaries will react in a crisis environment. Since these lenders rely on short-term funding from traditional banks, rather than from more stable deposits, it is conceivable that an evaporation of liquidity in the midst of a credit crunch could have a severe impact on fintech firms. Nonetheless, since fintech lenders sell the mortgages they originate quickly, and are less leveraged than the traditional banking sector, perhaps they are less prone to adverse short-term funding conditions than their bank counterparts. The evidence in this paper suggests, however, that provided these lenders maintain adequate funding sources during a crisis, that their technological advantages provide the Fed with additional ammunition

in responding to downturns. Fast and convenient online lending options appear to bolster the supply of credit and allow a broader range of potential borrowers to benefit from the refinancing channel of monetary policy.

References

Agarwal, S., G. Amromin, S. Chomsisengphet, T. Landvoigt, T. Piskorski, A. Seru, and V. Yao, (2020). "Mortgage refinancing, consumer spending and competition: evidence from the Home Affordable Refinancing Program," *Review of Economic Studies*, forthcoming.

Agarwal, S., J. C. Driscoll, and D. Laibson, (2013). "Optimal mortgage refinancing: a closed form solution," *Journal of Money, Credit, and Banking*, vol. 45(4), pages 591-622.

Agarwal, S., R. Rosen, and V. Yao (2013). "Why do borrowers make mortgage refinancing mistakes?" Working paper, No. 2013-02, Federal Reserve Bank of Chicago.

Bahadir, B., and I. Gumus (2016). "Credit decomposition and business cycles in emerging market economies," *Journal of International Economics*, vol. 103(C), pages 250-262.

Balyuk, T., A. N. Berger, and J. Hackney (2020). "What is fueling fintech lending? The role of banking market structure," Emory University Working Paper.

Bartlett, R., A. Morse, R. Stanton, and N. Wallace (2019). "Consumer-lending discrimination in the fintech era", NBER Working Papers 25943.

Bayer, P., F. Ferreira, and S. L. Ross (2018). "What drives racial and ethnic differences in high-cost mortgages? The role of high-risk lenders," *Review of Financial Studies*, vol. 31(1), pages 175-205.

Becker, G. (1957). "The economics of discrimination," Chicago: University of Chicago Press.

Beraja, M., A. Fuster, E. Hurst, and J. Vavra (2019). "Regional heterogeneity and the refinancing channel of monetary policy," *Quarterly Journal of Economics*, vol. 134(1), pages 109-183.

Bernanke, B., and M. Gertler (1995). "Inside the black box: the credit channel of monetary policy transmission," *Journal of Economic Perspectives*, vol. 9(4), pages 27-48.

Bostic, R. (1996). "The role of race in mortgage lending: revisiting the Boston Fed study," Federal Reserve Board of Governors Working Paper.

Browne, L. E., J. McEneany, A.H. Munnell, and G. M. B. Tootell (1996). "Mortgage lending in Boston: interpreting HMDA data," *American Economic Review*, vol. 86(1), pages 25-53.

Buchak, G., G. Matvos, T. Piskorski, and A. Seru (2018). "Fintech, regulatory arbitrage and the rise of shadow banks," *Journal of Financial Economics*, vol. 130(3), pages 453-483.

Burgess, R., and R. Pande (2005). "Do rural banks matter? Evidence from the Indian social banking experiment," *American Economic Review*, vol. 95(3), pages 780-795.

Campbell, J. Y., and J. Cocco (2003). "Household risk management and optimal mortgage choice," *Quarterly Journal of Economics*, vol. 188(4), pages 1449-1494.

Campello, M., J. Gao, and Q. Xu, (2019). "Personal taxes and firm skill hiring: evidence from 27 million job postings," Kelley School of Business Research Paper No. 19-35.

Card, D., and A. Krueger, (1994). "Minimum wages and employment: a case study of the fast-food industry in New Jersey and Pennsylvania," *American Economic Review*, vol. 84(4), pages 772-293.

Cetorelli, N., and P. E. Strahan (2006). "Finance as a barrier to entry: bank competition and industry structure in local U.S. markets," *Journal of Finance*, vol. 61(1), pages 437-461.

Chen, H., M. Michaux, and N. Roussanov (2020). "Houses as ATMs: mortgage refinancing and macroeconomic uncertainty," *Journal of Finance*, vol. 75(1), pages 323-375.

Cheng, P., Z. Lin, and Y. Liu (2015). "Racial discrepancies in mortgage interest rates," *Journal of Real Estate Finance and Economics*, vol. 51(1), pages 101-120.

Chernenko, S., I. Erel, R. Prilmeier (2019). "Why do firms borrow directly from non-banks?" NBER Working Papers 26458.

Cotti, C. D., C. J. Courtemanche, J. C. Maclean, E. T. Nesson, M. F. Pesko, N. Tefft (2021). "The effects of e-cigarette taxes on e-cigarette prices and tobacco product sales: evidence from retail panel data," NBER Working Papers 26724.

Deng, Y., and J. M. Quigley (2012). "Woodhead behavior and the pricing of residential mortgages," NUS Institute of Real Estate Studies Working Paper.

Di Maggio, M., A. Kermani, and C. J. Palmer, (2020). "How quantitative easing works: evidence on the refinancing channel," *Review of Economic Studies*, vol. 87(3), pages 1498-1528.

Drechsler, I., A. Savov, and P. Schnabl (2019). "How monetary policy shaped the housing boom," NBER Working Papers 25649.

Eichenbaum, M., S. T. Rebelo, A. Wong (2018). "State dependent effects of monetary policy: the refinancing channel," NBER Working Papers w25152.

Erel, I., and J. Liebersohn, (2020). "Does fintech substitute for banks? Evidence from the Paycheck Protection Program," NBER Working Papers 27659.

Fuster, A., M. Plosser, P. Schnabl, and J. Vickery (2019). "The role of technology in mortgage lending," *Review of Financial Studies*, vol. 32(5), pages 1854-1899.

Ghent, A. C., R. Hernandez-Murillo, and M. T. Owyang (2014). "Differences in subprime loan pricing across races and neighborhoods," *Regional Science and Urban Economics*, vol. 48(C), pages 199-215.

Gilje, E. P. (2019). "Does local access to finance matter? Evidence from U.S. oil and natural gas shale booms," *Management Science*, vol. 65(1), pages 1-18.

Gopal, M. and P. Schnabl (2020). "The rise of finance companies and fintech lenders in small business lending," New York University Working Paper.

Greenwald, D. (2018). "The mortgage credit channel of macroeconomic transmission," MIT Sloan Research Paper No. 5184-16.

Holmes, T. J., (1998). "The effect of state policies on the location of manufacturing: evidence from state borders," *Journal of Political Economy*, vol. 106(4), pages 667-705.

Hurst, E., and F. Stafford, (2004). "Home is where the equity is: mortgage refinancing and household consumption," *Journal of Money, Credit, and Banking*, vol. 36(6), pages 985-1014.

Jagtiani, J., and C. Lemieux (2018). "Do fintech lenders penetrate areas that are underserved by traditional banks?" Journal of Economics and Business, vol. 100(C), pages 43-54.

Jayaratne, J., and P. E. Strahan (1996). "The finance-growth nexus: evidence from bank branch deregulation," *Quarterly Journal of Economics*, vol. 111(3), 639-670.

Jordà, O. (2005). "Estimation and inference of impulse responses by local projections," *American Economic Review*, vol. 95(1), pages161-182.

Koijen, R., O. Van Hemert, and S. Van Nieuwerburgh (2009). "Mortgage timing," *Journal of Financial Economics*, vol. 93(2), pages 292-324.

LaCour-Little, M. (1999). "Another look at the role of borrower characteristics in predicting mortgage prepayments," *Journal of Housing Research*, vol. 10(1), pages 45-60.

Ladd, H. F. (1998). "Evidence on discimination in mortgage lending," *Journal of Economic Perspectives*, vol. 12(2), pages 41-62.

Ljunquist, A., and M. Smolyansky, (2018). "To cut or not to cut: on the impact of corporate taxes on employment and income," Working Paper.

Mian, A., and A. Sufi (2011). "House prices, home equity-based borrowing, and the household leverage crisis," *American Economic Review*, vol. 101(5), pages 2132-2156.

Mian, A., and A. Sufi (2014a). "House price gains, and U.S. household spending, from 2002 to 2006," NBER Working Papers 20152.

Mian, A., and A. Sufi (2014b). "What explains the 2007-2009 drop in employment?" *Econometrica*, vol. 82(6), pages 2197-2223.

Mian, A., A. Sufi, and K. Rao (2014). "Household balance sheets, consumption, and the economic slump," *Quarterly Journal of Economics*, vol. 128(4), pages 1687-1726.

Mian, A., A. Sufi, and E. Verner (2020). "How does credit supply expansion affect the real economy? The productive capacity and household demand channels," *The Journal of Finance*, vol. 75(2), pages 949-994.

Moretti, E., and D. Wilson, (2017), "The effect of state taxes on the geographical location of top earners: evidence from star scientists," *American Economic Review*, vol. 107(7), pages 1858-1903.

Pence, K. M., (2006). "Foreclosing on opportunity: state laws and mortgage credit," *Review of Economics and Statistics*, vol. 88(1), pages 177-182.

Philippon, T. (2016). "The fintech opportunity," NBER Working Papers 22476.

Richard, S. F., and R. Roll (1989). "Pre-payments on fixed-rate mortgage-backed securities," *Journal of Portfolio Management*, vol. 15(3), pages 73-82.

Scharfstein, D., and A. Sunderam (2018). "Market power in mortgage lending and the transmission of monetary policy," Working Paper, Harvard University.

Stanton, R. (1995). "Rational prepayment and the valuation of mortgage-backed securities," *Review of Financial Studies*, vol. 8(3), pages 677-708.

Stulz, R. (2019). "FinTech, BigTech, and the future of banks," NBER Working Papers 26312.

Taylor, J. B. (2007). "Housing and monetary policy," NBER Working Papers 13682.

Tables and Figures

Table 1.1

This exhibit presents summary statistics on the refinance lending behavior of fintech firms in the county-level HMDA sample. Panel A describes fintech lending by year. The second column, labelled "Total Fintech Refis," displays the aggregate volume of refinance credit supplied by fintech firms, in millions of dollars, in each year of the sample. The third column displays the share of total refinancing credit originated by fintech lenders. The fourth column displays the mean quantity of refinance credit originated by an individual fintech firm, while the fifth column shows the amount originated by the largest lender in the sample. The sixth column gives a count of the number of fintech firms in the sample each year. Panel B displays summary statistics at the county level. I display mean, median, and 90th percentile levels of fintech activity. The second through fourth columns describe the number of fintech lenders that operate in individual counties (e.g. the median column displays the median number of fintech lenders that originate a home refinance loan in a given county during each year of the sample). Columns 5-7 display the total value of fintech refinance loans, in millions, at the county level, while columns 8-10 show market shares of fintech lenders in county-level markets for home refinance credit.

Panel A
Fintech Summary Statistics by Year

Year	Total Fintech Refis ($ Millions)	Fintech Share of Total Refis	Mean Fintech Firm Refis	Max Fintech Firm Refis	Count of Fintech Firms
2010	41,393	.043	3,449	24,987	12
2011	44,969	.051	2,811	27,147	15
2012	105,596	.074	5,866	65,045	17
2013	112,763	.107	6,265	68,966	17
2014	75,461	.150	3,972	45,230	17
2015	115,988	.151	6,105	59,833	17
2016	152,853	.162	7,643	71,050	18
2017	111,323	.186	5,060	57,049	20
2018	87,871	.142	3,661	48,716	22
2019	182,840	.153	7,618	101,503	22

Panel B
Fintech Summary Statistics by County and Year

	Count of Fintech Firms			Fintech Lending ($ Millions)			Fintech Market Share		
Year	Mean	Median	90th Pct.	Mean	Median	90th Pctl.	Mean	Median	90th Pct.
2010	2.68	2	6	13.01	1.21	21.59	.050	.043	.095
2011	3.89	3	8	14.10	1.49	23.97	.060	.053	.109
2012	4.75	4	10	33.15	2.23	54.29	.069	.061	.121
2013	5.60	5	11	35.34	3.00	61.72	.105	.098	.177
2014	5.91	5	12	23.74	2.66	42.99	.168	.161	.270
2015	6.39	6	13	36.44	3.23	57.77	.166	.157	.264
2016	6.92	6	13	47.94	3.77	75.45	.170	.164	.267
2017	7.27	6	15	34.90	3.37	59.52	.197	.192	.307
2018	7.53	6	15	27.51	3.05	49.36	.187	.170	.306
2019	7.88	7	16	57.26	4.43	89.32	.180	.166	.294

Table 1.2

This table shows summary statistics for the county-level HMDA dataset merged with demographic and economic information from the US Census, and other sources. The sumary statistics show the average, median, standard deviation, 1st and 3rd quartiles and the 90th percentile value of each variable. The "Observations" column displays the number of non-missing observations in the merged data set. "Total Loans" shows the total volume of all loans (for purchase or refinance) in millions of dollars, while "Total Refis" describes the volume of refinance loans. "FHA Share" describes the share of refinance loans that are FHA-guaranteed, while "Jumbo Share" describes the share of loans that exceed the conforming limit. "No. Bank Branches" gives the number of physical brick-and-mortar bank branch establishments in the county, as given by the FDIC's Summary of Deposits data. "Branches Per Cap." gives the total number of bank branches on a per capita basis, while "Branches/Mi Sq." gives the total number of bank branches per square mile. "Pct. w/ Mortgage" describes the share of a county's home owners with an outstanding mortgage balance, while "Pct. Renting" gives the proportion of a county's households that rent their homes. "Pct. Black," "Pct. White," and "Pct. Hispanic" give the share of a county's population that identify as Black or African American, White, or Hispanic/Latino of any race. "Pct. College Degree" describes the proportion of the population with at least a bachelor's degree, while "Pct. Over 65" describes the proportion of a county's population over the age of 65.

Summary Statistics at the County-level							
	Observations	Mean	Q1	Median	Q3	90th Pct.	St. Dev.
Total Loans	31862	560.3	15.45	54.48	227.0	1014	2626
Total Refis	31862	275.7	6.837	24.11	101.5	442.6	1505
FHA Share	31862	0.271	0.182	0.255	0.339	0.434	0.130
Jumbo Share	31862	0.048	0	0.023	0.058	0.125	0.080
Population	31210	102218	11167	25995	68053	206295	326268
Avg. Wage	31096	36490	30590	34800	40180	47,300	9,654
Employment/Pop.	31210	0.444	0.397	0.445	0.492	0.531	0.072
Unemployment Rate	31210	0.063	0.041	0.057	0.08	0.104	0.029
Pop. Density	31209	269.2	17.17	45.26	117.35	396.3	1781
No. Bank Branches	30987	29.99	5	11	23	61	75.50
Branches Per Cap.	30987	0.466	0.279	0.381	0.543	0.829	0.304
Branches/Mi Sq.	30987	0.086	0.008	0.018	0.041	0.119	0.623
Poverty Rate	8179	0.144	0.104	0.140	0.178	0.214	0.056
Pct. w/ Mortgage	7937	0.634	0.594	0.625	0.673	0.725	0.069
Pct. Renting	7747	0.142	0.019	0.037	0.265	0.336	0.141
Pct. Black	7746	0.110	0.025	0.065	0.147	0.271	0.124
Pct. White	7872	0.807	0.766	0.836	0.891	0.931	0.130
Pct. Hispanic	8179	0.115	0.037	0.066	0.135	0.263	0.132
Pct. College Degree	8179	0.293	0.210	0.278	0.358	0.451	0.114
Pct. Over 65	8179	0.152	0.126	0.148	0.171	0.197	0.042

Table 1.3

This exhibit shows coefficients and standard errors from estimating equation (1). Panel A expresses fintech presence as the average fintech refinance market share, from 2010-2019. Panel B expresses fintech presence as the average total number of fintech lenders that originate refinance loans in a county. The variables labeled as "HMDA Mortgage Variables" consist of mortgage market characteristics taken from the HMDA data. "Demographic Variables" consist of county-level information for which data are available for the majority of counties. "Avg. Wage" is the average income for employed persons in a county; "Employment/Pop." is a county's employment/population ratio. "ACS Mortgage Variables" and "ACS Demographic Variables" are taken from the American Community Survey (ACS). The ACS Demographic variables describe a county's poverty rate, and the racial and ethnic composition of its residents. The variables labeled as "Bank Branch Presence" come from the FDIC's Summary of Deposits (SOD) database. In both panels, standard errors are listed in parentheses. Coefficient significance levels of 10%, 5%, and 1% are denoted by *, **, and ***, respectively.

Panel A

Regressing Average Fintech Refinance Share on County-level Characteristics

		(1)	(2)	(3)	(4)
HMDA Mortgage Variables	FHA Share	.148***	.169***	.134***	.155***
		(.006)	(.004)	(.008)	(.005)
	Jumbo Loan Share	.116***	-.053***	-.041***	-.060***
		(.010)	(.007)	(.011)	(.007)
Demographic Variables	Log Population		.003***	.005***	-.0001
			(.001)	(.001)	(.001)
	Avg. Wage		.078***	.032***	.069***
			(.003)	(.005)	(.003)
	Employment/Pop.		-.464***	-.227***	-.426***
			(.009)	(.023)	(.010)
	Unemployment Rate		-1.348***	-.914***	-1.382***
			(.022)	(.034)	(.022)
	Log Pop. Density		-.009***	-.004***	-.011***
			(.001)	(.001)	(.001)
ACS Mortgage Variables	Pct. Mortgage	-.038***		-.025**	
		(.010)		(.010)	
	Pct. Renting	-.241***		-.126***	
		(.005)		(.006)	
ACS Demographic Variables	Poverty Rate			.040**	
				(.016)	
	Pct. White			-.077***	
				(.008)	
	Pct. Black			-.035***	
				(.009)	
	Pct. Hispanic			.087***	
				(.006)	
	Pct. College Degree			.017**	
				(.009)	
	Pct. Over 65			.442***	
				(.020)	
Bank Branch Presence	Branches Per Cap.				-.021***
					(.001)
	Branches Per Mi Sq.				.036***
					(.004)

Panel B Regressing Fintech Firm Count on County-level Characteristics		(1)	(2)	(3)	(4)
HMDA Mortgage Variables	FHA Share	3.614***	2.184***	3.165***	1.077***
		(.387)	(.126)	(.421)	(.125)
	Jumbo Loan Share	18.325***	4.653***	3.471***	5.541***
		(.630)	(.191)	(.610)	(.189)
Demographic Variables	Log Population		2.347***	1.923***	1.963***
			(.019)	(.061)	(.020)
	Avg. Wage		1.144***	1.244***	1.156***
			(.071)	(.253)	(.071)
	Employment/Pop.		-7.047***	-2.010	-3.395***
			(.246)	(1.261)	(.252)
	Unemployment Rate		-45.66***	-53.34***	-45.86***
			(.600)	(1.871)	(.586)
	Log Pop. Density		-.112***	-.451***	.102***
			(.015)	(.044)	(.017)
ACS Mortgage Variables	Pct. Mortgage	4.943***		-.111	
		(.646)		(.556)	
	Pct. Renting	-13.03***		-7.201***	
		(.306)		(.330)	
ACS Demographic Variables	Poverty Rate			-3.889***	
				(.904)	
	Pct. White			1.053**	
				(.427)	
	Pct. Black			5.333***	
				(.508)	
	Pct. Hispanic			6.676***	
				(.324)	
	Pct. College Degree			-.157	
				.473	
	Pct. Over 65			12.35***	
				(1.100)	
Bank Branch Presence	Branches Per Cap.				-1.364***
					(.033)
	Branches Per Mi Sq.				-2.330***
					(.110)

Table 1.4

This exhibit displays the results of estimating equation (2) using the merged county-level HMDA sample. Panel A displays results where the *Fintech* variable from equation (2) is defined as the lagged count of fintech firms active within a county. The table displays the coefficients on the *FintechCount* variable and the interaction between *FintechCount* and the *Rates* variables, as described in equation (2). Panel B displays results where the right-hand side *Fintech* variable is the lagged refinance market share of fintech firms in the county (*FintechShare*). In each panel, the columns display results across specifications which differ on the basis of the set of included controls, and the inclusion or exclusion of county fixed-effects. The set of controls labeled as "Baseline" includes the set of mortgage market and demographic/economic controls labelled as "HMDA Mortgage Variables" and as "Demographic Variables" in Table 1.3, plus average credit scores from Fannie Mae data. The specifications where the set of controls is labeled as "Full," adds to the baseline set of controls by including information from the American Community Survey (i.e. the variables labeled as "ACS Mortgage Variables" and "ACS Demographic Variables" in Table 1.3). I list the number of observations and adjusted R-squared values for each regression at the bottom of each of the panels. Adjusted R-squared values for the county fixed-effects specifications are calculated net of fixed-effects (i.e. they are estimated after the data has been de-meaned by county). I list Huber-White standard errors in parentheses beneath each coefficient. Significance at the 10%, 5%, and 1% levels are denoted by *, **, and ***, respectively.

Panel A				
Dependent Variable: First Difference of Log Refi Volume				
	(1)	(2)	(3)	(4)
FintechCount	.006***	.004***	.004***	.008***
	(.001)	(.001)	(.001)	(.002)
FintechCount*Rates	.013***	.016***	.023***	.030***
	(.001)	(.004)	(.001)	(.002)
Controls	Baseline	Full	Baseline	Full
N	27760	6484	27760	6484
Adj. R-squared	.897	.912	.555	.849
Year Fixed Effects	✓	✓	✓	✓
County Fixed Effects			✓	✓

Panel B				
Dependent Variable: First Difference of Log Refi Volume				
	(1)	(2)	(3)	(4)
FintechShare	.361***	.475***	.571***	1.09***
	(.030)	(.064)	(.053)	(.117)
FintechShare*Rates	.234***	.598***	.455***	.603***
	(.055)	(.113)	(.058)	(.121)
Controls	Baseline	Full	Baseline	Full
N	27760	6484	27760	6484
Adj. R-squared	.896	.912	.501	.830
Year Fixed Effects	✓	✓	✓	✓
County Fixed Effects			✓	✓

Table 1.5

This exhibit depicts results from the estimation of equation (2) on subsamples of the county-level HMDA dataset, sorted into quartiles based on county population. Panel A depicts results that use the *FintechCount* variable as the key regressor, while Panel B shows analogous results for the *FintechShare* variable. Tables present coefficients on the *Fintech* variable and the *Fintech*Rates* interaction term as described in equation (2). All regressions are estimated with the set of "Baseline" controls (and do not include the extended ACS variables). Each column contains results for a different population subsample, with the labels Q1-Q4 indicating the population quartile referenced in each specification. Huber-White standard errors are listed beneath each coefficient and significance at the 10%, 5%, and 1% levels are denoted by *, **, and ***, respectively.

Panel A				
Dependent Variable: Log Refi Volume (First-Difference)				
	Q1	Q2	Q3	Q4
FintechCount	.002	.001	.005***	.005***
	(.002)	(.002)	(.001)	(.001)
FintechCount*Rates	.010	.016***	.016***	.016***
	(.010)	(.002)	(.002)	(.002)
N	6822	6973	6981	6984
Adj. R-squared	.276	.573	.733	.857
Year Fixed Effects	✓	✓	✓	✓

Panel B				
Dependent Variable: Log Refi Volume (First-Difference)				
	Q1	Q2	Q3	Q4
FintechShare	.200***	.382***	.458***	.478***
	(.045)	(.055)	(.054)	(.059)
FintechShare*Rates	-.263***	-.039	.028	.439***
	(.083)	(.094)	(.096)	(.109)
N	6822	6973	6981	6984
Adj. R-squared	.475	.701	.803	.910
Year Fixed Effects	✓	✓	✓	✓

Table 1.6

This table displays results of estimating equation (5) using data from HMDA. Data are expressed at the state-year-firm level. Each column of this table denotes a different specification, containing various combinations of regulatory policy variables. The dependent variable takes a value of one if a firm has entered state j by year t, and a value of zero otherwise. The first regulatory variable, "Application Cost" is the cost, in dollars, of submitting an application to become a licensed mortgage originator in state j. "Qualitative Reqs Count," is a simple count of the number of significant qualitative requirements on a state's application checklist. "Net Worth Requirement" is the minimum level of capital required of non-bank lenders. "Brick & Mortar" is a dummy variable that takes a value of one if state j requires licensed lenders to maintain a physical branch presence in the state. The row labeled "AIC" reports the Akaike Information Criterion associated with the regression. *, **, and *** represent statistical significance (via a Z-statistic) at the 10%, 5%, and 1% levels, respectively. All models are estimated via a maximum likelihood approach.

Predicting Firm Market Entry by State Logistic Regression Results								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log Pop.	.415***	.337***	.350***	.356***	.373***	.382***	.388***	.403***
	(.027)	(.027)	(.028)	(.028)	(.028)	(.029)	(.029)	(.029)
Application Cost	-.346***					-.278***	-.265***	-.241***
	(.042)					(.046)	(.047)	(.047)
Qualitative Reqs Count		-.065***	-.064***	-.068***	-.068***	-.040***	-.040***	-.045***
		(.010)	(.010)	(.010)	(.010)	(.011)	(.011)	(.011)
Net Worth Requirement				-.001***	-.001***			-.001***
				(.000)	(.000)			(.000)
Brick and Mortar			-.263***		-.319***		-.172*	-.252**
			(.097)		(.097)		(.099)	(.100)
N	8544	8544	8544	8544	8544	8544	8544	8544
AIC	8808	8830	8825	8818	8810	8796	8795	8786
Fixed Effects:								
Year	✓	✓	✓	✓	✓	✓	✓	✓

Table 1.7

This table displays the results of the cross-border analysis of section 4. Each column displays coefficients attained from estimating a version of equation (6). Here, I display the coefficients on *Treat* and the *Treat*Rates* interaction terms from these models. The specifications displayed here vary along three dimensions, including the distance cutoff used for sample inclusion (with "Bandwidth" denoting this maximum distance, in miles), the set of fixed effects, and the inclusion or exclusion of counties containing large urban areas (defined as MSAs in the top 100 by population) as denoted by the "Sample" row.

Cross-Border Analysis of US States Dependent Variable: First Difference of Log Refi Volume								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treat	.010	.024**	.020***	.033***	.014**	.019*	.018***	.033***
	(.007)	(.010)	(.006)	(.009)	(.007)	(.011)	(.006)	(.010)
Treat*Rates	.055***	.055***	.015***	.014***	.063***	.064***	.020***	.020***
	(.005)	(.005)	(.004)	(.004)	(.003)	(.003)	(.003)	(.003)
Sample	All	All	Excl. Metros	Excl. Metros	All	All	Excl. Metros	Excl. Metros
Bandwidth (Mi.)	50	50	50	50	100	100	100	100
Fixed Effects:								
Year	✓	✓	✓	✓	✓	✓	✓	✓
State	✓	✓	✓	✓	✓	✓	✓	✓
Border		✓		✓		✓		✓

Table 1.8

This table displays results from estimating versions of equation (7), with each column of the table presenting results from a separate specification. The dependent variables are total retail spending (columns (1) and (2)) and discretionary retail spending (columns (3) and (4)), as defined in Section 5.1. Within a dependent variable, the specifications differ in their inclusion of county fixed effects (with columns (2) and (4) displaying results where these fixed-effects are included). The rows of the table display coefficients of the *FintechCount* variable and the *FintechCount*Rates* interaction term. Huber White standard errors are displayed in parentheses beneath each coefficient. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Response of Retail Spending to Fintech Presence Dependent Variables: First Difference of Log Retail Spending				
	Total Retail		Discretionary Retail	
	(1)	(2)	(3)	(4)
FintechCount	.0020***	.0009**	.0020***	.0009
	(.0001)	(.0004)	(.0002)	(.0004)
FintechCount*Rates	.0018***	.0019***	.0021***	.0023***
	(.0003)	(.0003)	(.0003)	(.0003)
N	10766	10766	10766	10766
Adj. R-squared	.622	.429	.602	.427
Year Fixed Effects	✓	✓	✓	✓
County Fixed Effects		✓		✓

Table 1.9

This table displays the association between fintech lending and local small business outcomes at the county level, with results generated via estimation of equation (8). Each column represents a different regression specification, which differ on the basis of their outcome variable and the inclusion or exclusion of county fixed-effects. The table reports coefficients on the *FintechCount* and *FintechCount*Rates* variables. "Estab. NonTr." is defined as the number of small business establishments in the non-tradable sector. "Emp. NonTr." describes the total employment in non-tradable sector small businesses. "Loan Count" expresses the total number of small business loans (to businesses with assets under $1 million), while "Loan Vol" is the the dollar volume of such loans. All variables are expressed as annual growth rates (i.e. one-year log differences). Standard errors are displayed in parentheses beneath each coefficient. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

	Response of Local Small Business Expansion to Fintech Presence							
	Dependent Variables Expressed as Log Differences (t to t+1)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Estab NonTr.	Estab NonTr.	Emp. NonTr.	Emp. NonTr.	Loan Count	Loan Count	Loan Vol	Loan Vol
FintechCount	.003***	.006***	.004***	.004***	.003***	.005***	.001	.006***
	(.000)	(.000)	(.001)	(.001)	(.000)	(.001)	(.001)	(.001)
FintechCount* Rates	.002***	.009***	.004***	.012***	.002***	.002***	.005***	.004**
	(.000)	(.000)	(.001)	(.001)	(.000)	(.001)	(.001)	(.001)
N	24611	24611	24611	24611	21563	21563	21563	21563
Adj. R-squared	.771	.059	.720	.063	.267	.061	.036	-.001
Fixed Effects:								
Year	✓	✓	✓	✓	✓	✓	✓	✓
County		✓		✓		✓		✓

Table 1.10

This table displays results from estimating equation (9) using a number of dependent variables. Each column displays results from a different specification, where each specification differs on the basis of its outcome variable. Dependent variables are expressed as shares of total county-level refinance credit, and first differences are taken. Thus, "Non-White Share" refers to the first difference in the share of refinance loans that went to non-White borrowers, with the share calculated as the total volume of loans to non-white borrowers divided by the total volume of refinance loans for the county. "Hispanic Share" is analogously defined for Hispanic/Latino borrowers. "FHA Loans" refers to the share of FHA guaranteed loans. "Junior liens" refers to loans backed by a subordinate lien on the property (i.e. not a first lien mortgage). Coefficients are displayed for the *FintechCount* and *FintechCount*Rates* interaction terms. Standard errors are displayed beneath each coefficient. *, **, and *** denote statistical significance at the 10%, 5%, and 1%, levels, respectively.

Refinance Credit Composition Regressions				
Dependent Variables: Percentage Point Change of Refinance Composition				
	Non-White Share	Hispanic Share	FHA Loans	Junior Liens
	(1)	(2)	(3)	(4)
FintechCount	.0020***	.0017***	-.0004***	.0002***
	(.0002)	(.0001)	(.00001)	(.0001)
FintechCount*Rates	.0049***	.0028***	.0010***	.0002***
	(.0002)	(.0001)	(.0001)	(.0001)
N	27760	27760	27760	27760
Adj. R-squared	.026	.014	.068	.010
Year Fixed Effects	✓	✓	✓	✓
County Fixed Effects	✓	✓	✓	✓

Table 1.11

This table displays results generated by estimating equation (10). Each column of the table displays results of a different variant of equation (10), where the sample is sorted into quartiles according to a different county-level characteristic. The dependent variable in each equation is the log growth of refinancing from time t-1 to time t. The rows of the table display coefficients associated with the *FintechCount* variable interacted with quartile indicator functions (variables which take a value of one if a county is in a given quartile of the distribution, as sorted by a particular trait, and zero otherwise) and the *FintechCount*Rates* interaction terms, also interacted with quartile indicators (the *FintechCount*Rates* interaction terms are denoted as "Count*Rates" below, to save space). The term "Q1" in the leftmost column denotes the quartile indicator associated with the bottom quartile of the distribution; "Q2" represents the second quartile, and so on. The first column, labeled "% White" indicates that counties are sorted based on their White population. The next column, labeled "% Hispanic," displays results where counties are sorted according to their percentage of Hispanic/Latino residents. The third column, labeled "Pop. Density," displays results from the population density-sorted sample. The last two columns display results where counties are sorted by the number of bank branch locations that they contain. "Branches/Pop." denotes the number of branches per capita, while "Branches/Mi Sq." denotes the number of bank branches per square mile. Standard errors are shown in parentheses beneath each coefficient. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Refinance Results Sorted by County Traits					
Dependent Variable: Log Refi Volume (First-Difference)					
	% White	% Hispanic	Pop. Density	Branches/Pop.	Branches/Mi Sq.
	(1)	(2)	(3)	(4)	(5)
FintechCount*Q1	.006***	.003***	.010***	.009***	.007***
	(.001)	(.001)	(.001)	(.001)	(.001)
FintechCount*Q2	.007***	.005***	.008***	.001	.004***
	(.001)	(.001)	(.001)	(.001)	(.001)
FintechCount*Q3	.006***	.006***	.007***	.005***	.002***
	(.001)	(.0004)	(.001)	(.001)	(.001)
FintechCount*Q4	.001	.006***	.006***	.001	.003*
	(.002)	(.0004)	(.0004)	(.003)	(.002)
Count*Rates*Q1	.013***	.002***	.018***	.009***	.012***
	(.001)	(.001)	(.003)	(.001)	(.001)
Count*Rates*Q2	.012**	.009***	.016***	.013***	.008***
	(.001)	(.001)	(.002)	(.001)	(.001)
Count*Rates*Q3	.010***	.009***	.016***	.007***	.004***
	(.001)	(.001)	(.001)	(.002)	(.001)
Count*Rates*Q4	.005	.009***	.011***	.006	.015***
	(.004)	(.001)	(.001)	(.006)	(.003)
Count Q4-Q1	-.005**	.003***	-.004***	-.008**	-.004***
Count*Rates Q4-Q1	-.008*	.007***	-.007***	-.003	.003
N	27760	27760	27760	27760	27760
Adj. R-squared	.903	.914	.899	.902	.900

Table 1.12

This table displays results from estimating equation (11). The dependent variable is the log change in discretionary spending, with discretionary spending described in Section 5.1. Each column shows results from a different specification, where specifications vary based on the county-level characteristic interacted with fintech variables. Column (1) displays results where fintech variables are interacted with the percentage of White residents in a county. The "*FintechCount*White*Rates*" variable denotes the "triple-diff" interaction between interest rate spreads, the count of fintech lenders, and the White population share of a county. Similarly, "Hisp." denotes a county's Hispanic/Latino population share, with "*FintechCount*Hisp.*Rates*" denoting the triple-diff interaction between the local count of fintech lenders, interest rate spreads, and a county's Hispanic/Latino population. The terms "Brnch./Pop" and "Brnch./Mi. Sq." denote the the number of bank branches per-capita, and the number of bank branches per square mile. Coefficients relating to these quantities are displayed in columns (3) and (4), respectively. Standard errors are displayed in parentheses beneath each coefficient. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Heterogeneous Consumption Responses to Fintech Dependent Variable: First Difference of Log Discretionary Spending	(1)	(2)	(3)	(4)
FintechCount	.0016***	.0016***	.0034***	.0019***
	(.0003)	(.0002)	(.0002)	(.0002)
FintechCount*Rates	.0086***	-.0006	.0090***	.0024***
	(.0014)	(.0004)	(.0007)	(.0003)
FintechCount*White	.0003			
	(.0004)			
FintechCount*White *Rates	-.0088***			
	(.0017)			
FintechCount*Hisp.		.0011***		
		(.0003)		
FintechCount*Hisp. *Rates		.0132***		
		(.0017)		
FintechCount *Brnch./Pop.			-.0067***	
			(.0005)	
FintechCount *Brnch./Pop.*Rates			-.0256***	
			(.0022)	
FintechCount *Brnch./Mi. Sq.				-.0001***
				(.00003)
FintechCount* Brnch./Mi. Sq.*Rates				-.0003***
				(.0001)
N	10766	10766	10766	10766
Adj. R-squared	.605	.606	.610	.604

Figure 1.1

This figure displays the geographic profile of fintech activity from 2010-2019. Counties are shaded to reflect sample averages of the number of active fintech lenders (on top) and the fintech market share of refinancing loans (on the bottom). Darkly shaded areas represent counties with high fintech activity, while lighter-shaded areas have less fintech activity. Small counties with inconsistent market activity (i.e. counties without mortgage refinancing in every year of the study) are dropped from the sample. Legend labels on top display the range of average fintech counts for counties shaded in a given color, while legend labels on bottom give the equivalent range for fintech market shares.

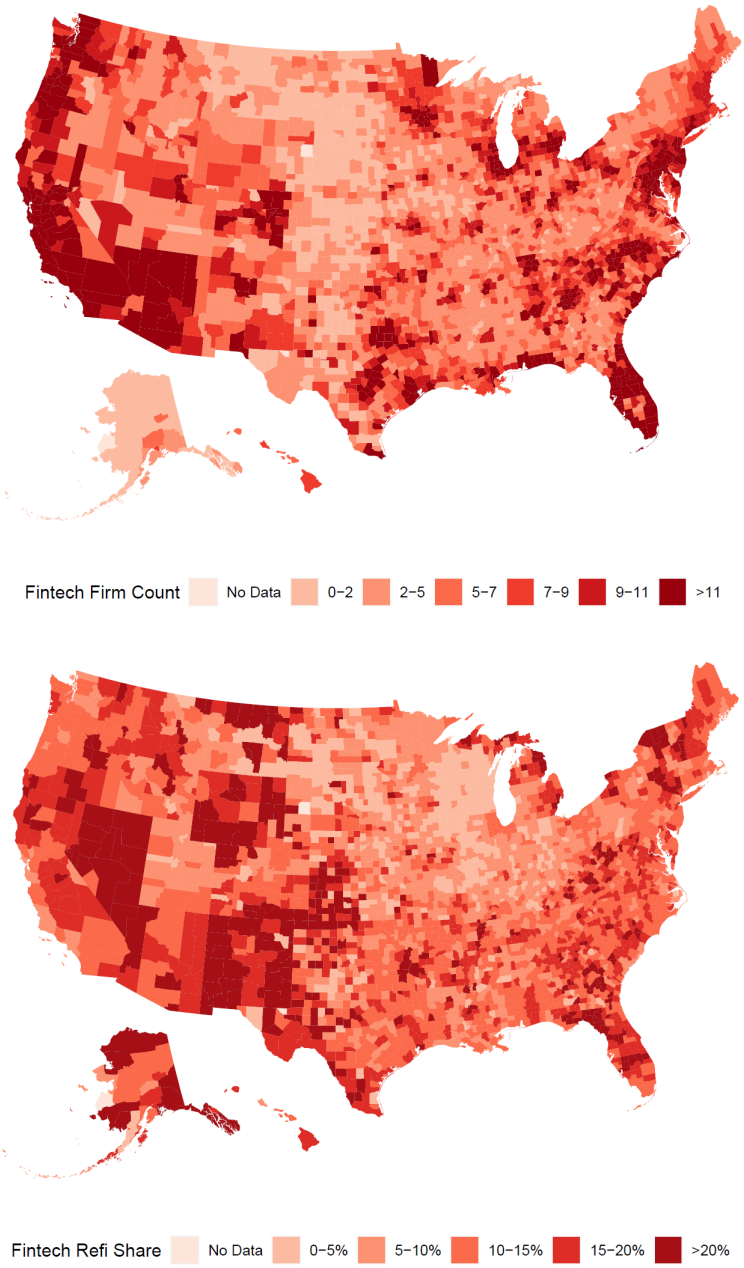


Figure 1.2

This figure displays impulse responses generated from estimating equations (3) and (4) using data from the Fannie Mae Single Family Loan Performance dataset. Panels A and B of this figure show impulse responses from equation (3). The impulse responses plot the behavior of refinancing activity in response to an increase in fintech lending. The blue line in panels A and B plots the sum of the coefficients γ^h and δ^h on the *Fintech* variable and the *Fintech*Rates* interaction. Impulse responses are plotted over time horizons ranging from 1-5 months. Panel A shows impulse responses estimated from equations which omit ZIP code fixed-effects, while Panel B includes these fixed-effects. Dotted blue lines display the 95 percent confidence intervals for the impulse responses. Panels C and D display impulse responses generated from equation (4). They display the coefficients on the interaction between the *Fintech* and *OutstandingStock* variables, over 1-5 month time horizons.

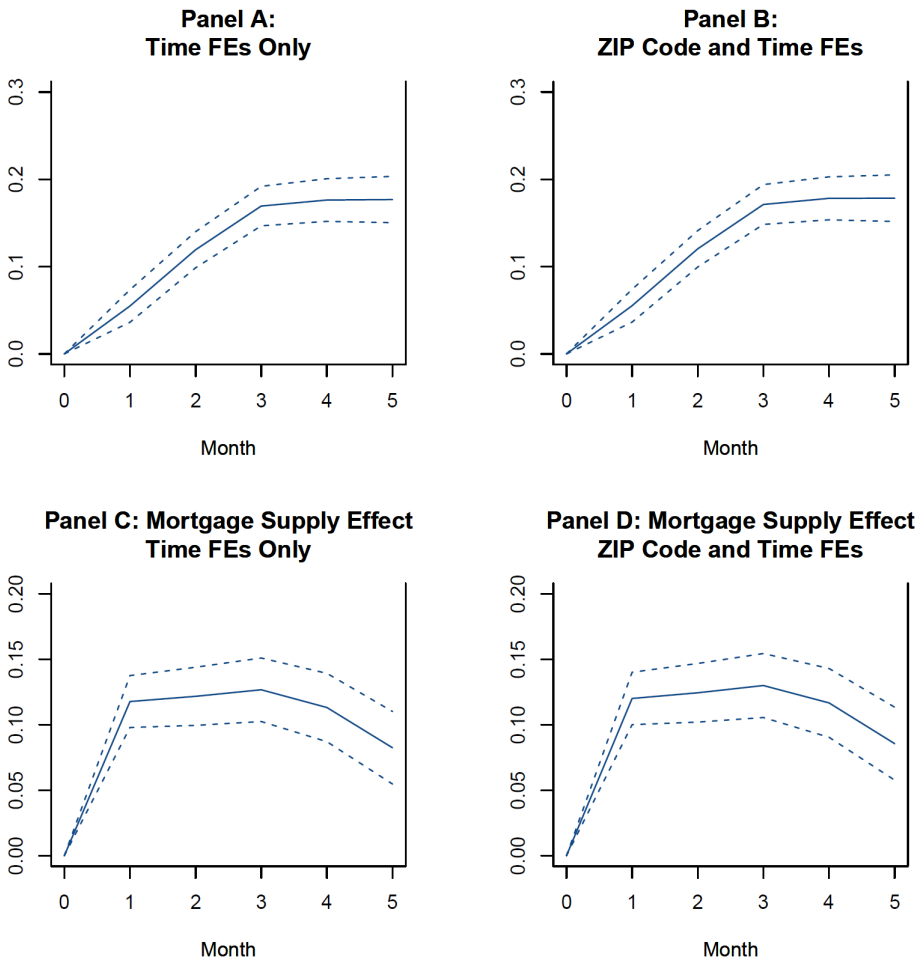


Figure 1.3

This figure gives an example of the progression with which some fintech lenders (i.e. those that were smaller and less established early on in the sample) entered state mortgage markets. The pictures below show the timing of market entry by a single firm, CashCall Mortgage, during the first half of the sample, from 2010-2015 (at which point it had begun originating mortgage loans in almost every state). The figure was generated using loan-level information from the Home Mortgage Disclosure Act database. States shaded in red represent the set of states in which CashCall originated a positive dollar value of mortgages in a given year, while those in white represent states in which CashCall did not make any loans.

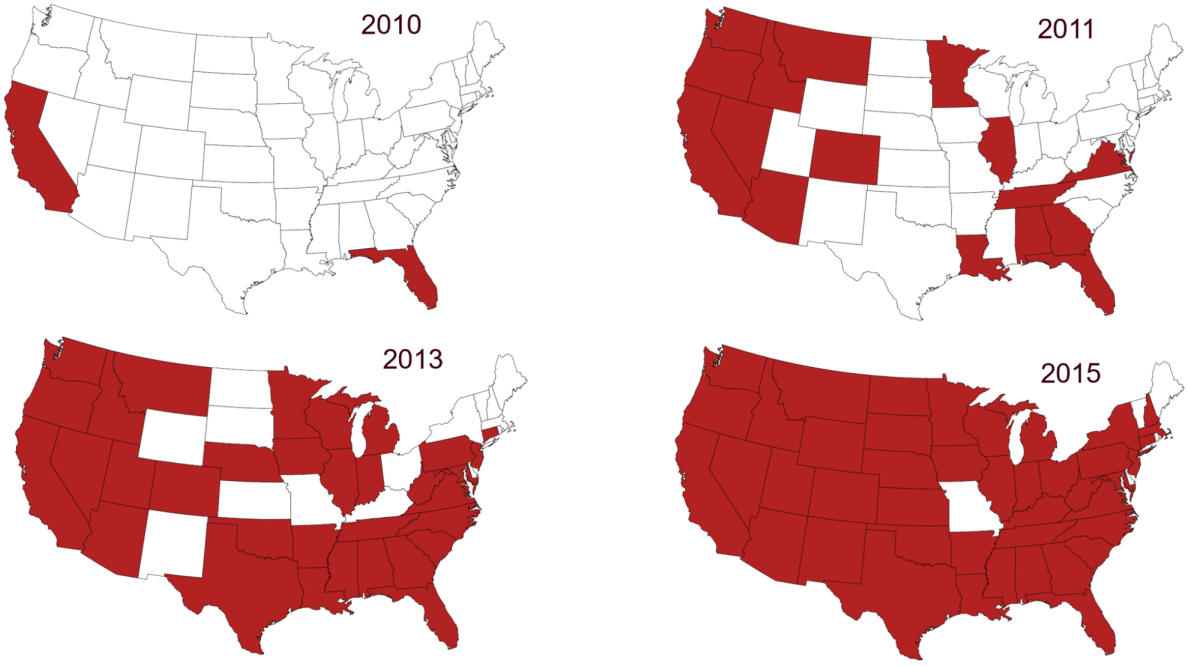


Figure 1.4

This figure displays the aggregate analogue of Figure 1.3. Itshows the total number of fintech firms that have entered each state, by year, matching the years displayed in Figure 1.3. States with darker shading have a larger number of active fintech firms. The figure was generated using loan level HMDA data. The numbers next to each legend label give the number of firms in states colored with a given hue. In 2013, the dispersion in state-level fintech counts is rather wide, an accordingly, states with 15 or more fintech lenders are labeled as ">=15."

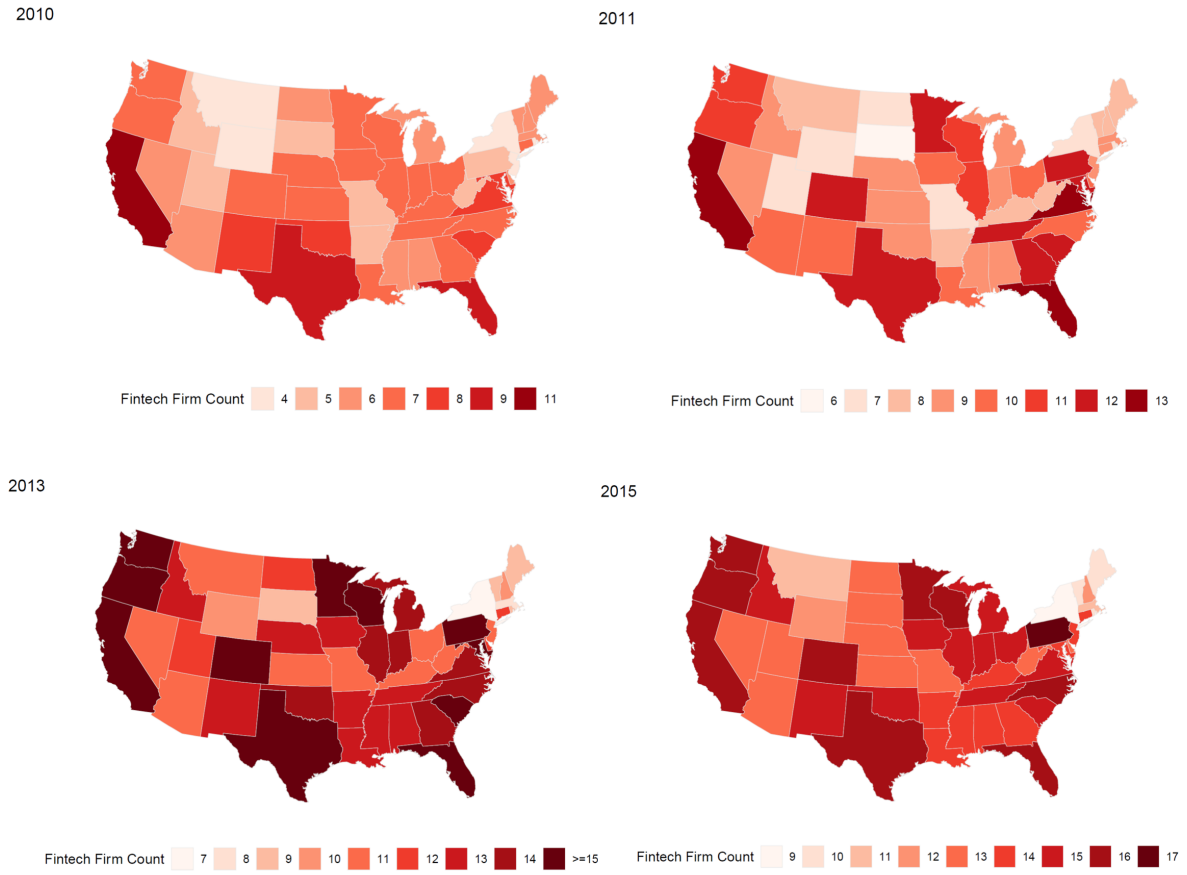
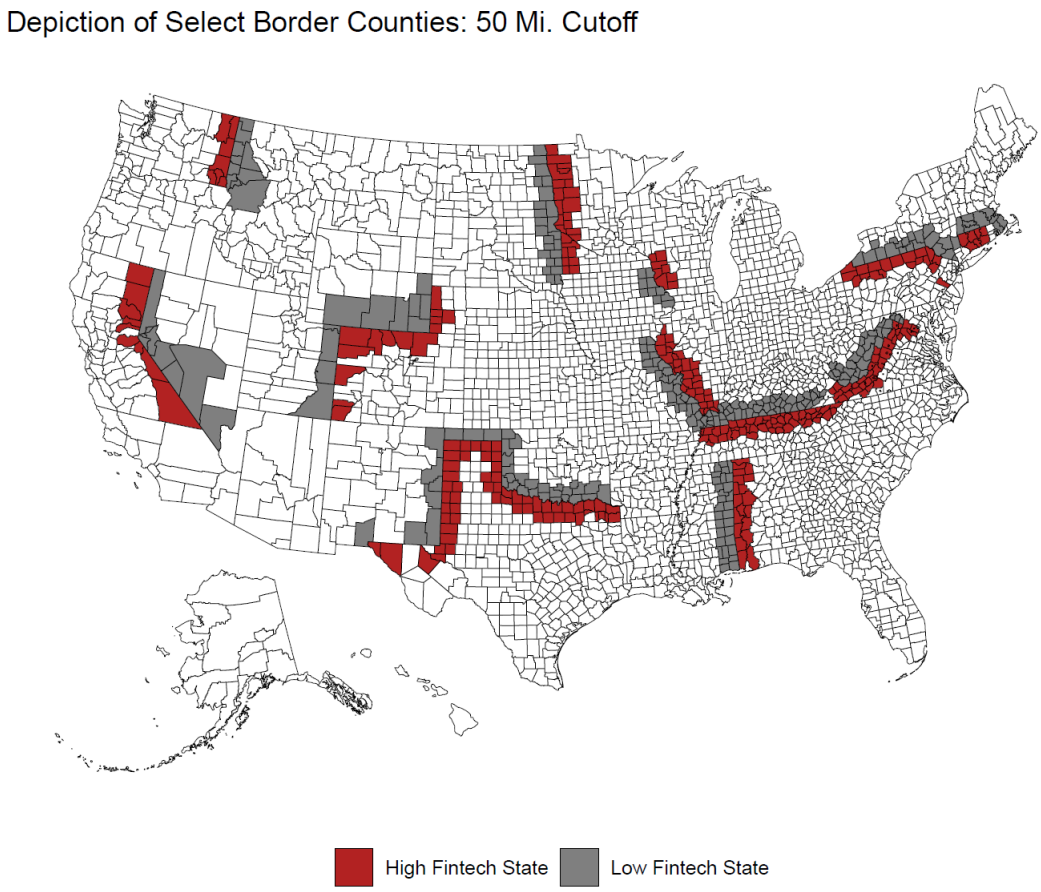


Figure 1.5

This figure gives intuition for the empirical approach discussed in Section 4.2 and estimated via equation (6). It illustrates how sample counties are selected given a group of treated and control states. State pairs are selected by identifying the set of bordering states that have different numbers of active fintech firms in a given year. To generate this figure, I selected a group of treatment and control states that appear in the sample on a number of occasions, though not necessarily in all years. Counties residing in a treated state in this example are shaded in red, while counties in a paired control state are shaded in gray. Only counties located within 50 miles of their paired state border are shaded. Border distances are determined using population centroids in each county. Some counties located on state borders (in western states with large counties) are nonetheless excluded from the sample if most of their population lives more than 50 miles away from the shared border.



Appendix

Table 1.13

This table displays pairwise correlations between variables that comprise the merged county-level HMDA panel. Panel A shows pairwise correlations between the main variables that comprise the "Baseline" sample, which includes county-level demographic information from sources other than the American Community Survey. The full set of correlations of these variables is split over two panels, panel A parts I and II, due to space considerations. Variable names follow the naming conventions of Table 1.2 and Table 1.3. Panel B displays pairwise correlations between the main variables comprising the extended set of controls, including obsevables derived from American Community Survey data. These correlations are, again, split over two panels. See Table 1.2 and Table 1.3 in the set of main exhibits for variable definitions.

Panel A-I
Correlations: Baseline Controls Sample (W/o Census ACS Data)

	Ft. Refi	Ft. Count	Pop.	Wage	Unem.	Emp./Pop.
Fntch. Share	1					
Fntch. Count	0.220	1				
Log-Pop.	-0.052	0.750	1			
Avg. Wage	0.131	0.391	0.361	1		
Unemp.	-0.256	-0.248	0.067	-0.240	1	
Emp./Pop.	-0.099	0.125	0.018	0.263	-0.575	1
Pop. Density	-0.028	0.142	0.255	0.267	-0.008	0.064
FHA Share	0.303	0.093	0.030	-0.014	-0.048	-0.192
Jumbo Shr.	-0.099	0.179	0.094	0.181	-0.161	0.188
Brnch/Pop.	-0.085	-0.457	-0.564	-0.175	-0.244	0.358
Brnch/Mi sq.	-0.037	0.106	0.198	0.263	-0.012	0.072

Panel A-II
Correlations: Baseline Controls Sample (Continued)

	Dens.	FHA	Jumbo	Br./Pop.	Br./Mi
Pop. Density	1				
FHA Share	-0.070	1			
Jumbo Shr.	0.176	-0.368	1		
Brnch/Pop.	-0.074	-0.189	0.077	1	
Brnch/Mi sq.	0.913	-0.079	0.180	-0.030	1

Panel B-I Correlations: Full Controls Sample	Ft. Refi	Ft. Count	Poverty	% Mtge.	% Rent
Fntch. Share	1				
Fntch. Count	0.544	1			
Poverty Rate	-0.082	-0.266	1		
Pct. Mtge.	-0.229	-0.111	-0.099	1	
Pct. Rental	-0.550	-0.516	0.314	0.336	1
Pct. Black	0.028	0.088	0.322	0.015	0.087
Pct. White	-0.087	-0.175	-0.319	-0.020	-0.108
% Hispanic	0.153	0.242	0.206	0.021	0.059
Pct. Over 65	0.310	0.137	-0.140	-0.207	-0.336
Pct. College	-0.133	0.198	-0.380	0.154	0.034

Panel B-II Correlations: Full Controls Sample (Continued)	% Black	% White	% Hisp.	% >65	% Coll.
Fntch. Share					
Fntch. Count					
Poverty Rate					
Pct. Mtge.					
Pct. Rental					
Pct. Black	1				
Pct. White	-0.779	1			
% Hispanic	-0.104	-0.188	1		
Pct. Over 65	-0.214	0.317	-0.201	1	
Pct. College	0.027	-0.084	-0.143	-0.219	1

Table 1.14

This exhibit displays the results of several robustness tests of the baseline difference-in-difference analysis. Panel A displays results analogous to Panel A of Table 1.4, which are estimated from a version of equation (2) which controls for the lagged growth of refinancing. Panel B displays within-county results similar to Panel A of Table 1.4, which use counts of other intermediaries rather than of fintech firms as the key right-hand side variables. These specifcations take the form $\Delta_1 Refivol_{i,t} = \alpha_t + \beta \cdot Intermediary_{i,t-1} + \gamma \cdot Intermediary_{i,t-1} \cdot \Delta_{avg} Rates_t + \delta \cdot Controls_{i,t-1} + \varepsilon_{i,t}$ where Intermediary is one of *OtherNonbank*, *LargeBank*, or *SmallBank*. Huber White standard errors are listed in parentheses beneath each coefficient. Significance at the 10%, 5%, and 1% levels are given by *, **, and ***, respectively.

Panel A: Controlling for Lagged Refi Growth Dependent Variable: Log Refi Volume (First-Difference)	(1)	(2)	(3)	(4)
FintechCount	.005***	.003***	.012***	.008***
	(.0004)	(.0008)	(.0009)	(.002)
FintechCount*Rates	.012***	.011***	.016***	.014***
	(.0005)	(.001)	(.0006)	(.001)
Controls	Baseline	Full	Baseline	Full
N	24656	5772	24656	5772
Adj. R-squared	.909	.921	.566	.855
Year Fixed Effects	✓	✓	✓	✓
County Fixed Effects			✓	✓

Panel B: Falsification Tests with Non-Fintech Intermediaries Dependent Variable: Log Refi Volume (First-Difference)						
	(1)	(2)	(3)	(4)	(5)	(6)
OtherNonbank	-.003***	-.002***				
	(.0002)	(.0002)				
OtherNonbank*Rates	.001***	.001***				
	(.0001)	(.0001)				
LargeBank			-.011***	-.001		
			(.001)	(.001)		
LargeBank*Rates			.012***	.009***		
			(.001)	(.001)		
SmallBank					-.004***	-.002***
					(.0002)	(.0002)
SmallBank*Rates					.001***	.0003***
					(.0001)	(.0001)
Controls	Baseline	Full	Baseline	Full	Baseline	Full
N	27760	6484	27760	6484	27760	6484
Adj. R-squared	.557	.850	.556	.845	.558	.867
Year Fixed Effects	✓	✓	✓	✓	✓	✓
County Fixed Effects	✓	✓	✓	✓	✓	✓

Table 1.15

This exhibit displays correlations between fintech activity and refinance credit growth using monthly data from Fannie Mae. Panel A displays *FintechShare* coefficients from estimating an equation of the form $\Delta_3 Refivol3mo_{i,t} = \alpha_t + \beta \cdot Fintech_{i,t-1} + \gamma \cdot Fintech_{i,t-1} \cdot \Delta_{avg} Rates_{i,t} + Controls_{i,t-1} + \varepsilon_{i,t}$. This is a 3-month regression analogue of equation (2), so that the dependent variable, *Refivol3mo*, is the total refinancing activity over a three month period (months t through t+2), and *FintechShare* is defined as the market share of fintech lenders over a 3-month period (from time t-3 to t-1). Panels B and C depict impulse responses generated from local projections. Specifically, they plot values of fintech coefficients from estimating equation (3) for time horizons of 1-5 months (i.e. values of h=1...5). Panel B displays these coefficients for the *FintechLoanGrowth* variable, which is a percent change in the total number of loans originated by fintech firms.

Panel A: Baseline Regressions with FNMA Data Dependent Variable: First Difference of Log Refi Volume	(1)	(2)
FintechShare	.033***	.041***
	(.012)	(.013)
FintechShare*Rates	.045**	.039**
	(.019)	(.020)
Year-Month FEs	✓	✓
ZIP-Code FEs		✓

Panel B: Local Projection Impulse Responses with Market Shares Dependent Variable: First Difference of Log Refi Volume	(t+1)	(t+2)	(t+3)	(t+4)	(t+5)
FintechShare	.028**	.040***	.055***	.079***	.087***
	(.014)	(.015)	(.017)	(.018)	(.019)
FintechShare*Rates	.048**	.085**	.114***	.147***	.153***
	(.020)	(.023)	(.025)	(.026)	(.028)
Year-Month FEs	✓	✓	✓	✓	✓
ZIP-Code FEs	✓	✓	✓	✓	✓

Panel C: Impulse Response with Growth in Fintech Loans Dependent Variable: First Difference of Log Refi Volume	(t+1)	(t+2)	(t+3)	(t+4)	(t+5)
FintechLoanGrowth	.001	.008***	.009***	.009***	.001
	(.002)	(.002)	(.002)	(.002)	(.002)
FintechLoanGrowth*Rates	.055***	.120***	.171***	.178***	.178***
	(.010)	(.011)	(.012)	(.013)	(.014)
Year-Month FEs	✓	✓	✓	✓	✓
ZIP-Code FEs	✓	✓	✓	✓	✓

Table 1.16

This table shows results from estimating versions of equation (7). The dependent variables are total retail spending and discretionary retail spending, with each column of the table representing a separate specification. Columns (1) and (2) display regression results where total retail spending is the outcome variable, while columns (3) and (4) depict results where discretionary retail is the dependnt variable. The table displays results of specifications that both include and exclude county fixed effects. The rows of the table display coefficients of the *FintechShare* variable and the *FintechShare* interaction with interest rate spreads as detailed in equation (7) and as described in section 3. Standard errors are displayed in parentheses beneath each coefficient. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Response of Retail Spending to Fintech Presence Dependent Variables: First Difference of Log Retail Spending				
	Total Retail		Discretionary Retail	
	(1)	(2)	(3)	(4)
Lagged Fintech	.044***	.288***	.049***	.312***
Market Share	(.011)	(.020)	(.011)	(.021)
Lagged Share*Rates	-.047**	.062***	-.048***	.066***
	(.019)	(.020)	(.021)	(.021)
N	10766	10766	10766	10766
Adj. R-squared	.611	.431	.591	.426
Year Fixed Effects	✓	✓	✓	✓
County Fixed Effects		✓		✓

Table 1.17

This table displays results from estimating equation (9). Each column of the table shows results from a different specification; each specification differs on the basis of the dependent variable used in the estimation. Dependent variables are expressed as shares of total county-level refinance credit, and first differences are taken. Thus, "Non-White Share" refers to the first difference in the share of refinance loans that went to non-White borrowers, with this share calculated as the total volume of loans to non-white borrowers divided by the total volume of refinance loans for the county. "Hispanic Share" is analogously defined for Hispanic/Latino borrowers. "FHA Loans" refers FHA guaranteed loans. "Junior liens" refers to refinances of loans backed by subordinate liens (i.e. not a first lien mortgage). Standard errors are displayed beneath each coefficient. *, **, and *** denote statistical significance at the 10%, 5%, and 1%, levels, respectively.

Refinance Credit Composition Regressions				
Dependent Variables: Percentage Point Change of Refinance Composition				
	FHA Loans	Junior Liens	Hispanic Share	Non-White Share
	(1)	(2)	(3)	(4)
FintechShare	.113***	.010***	.013**	.0001
	(.008)	(.003)	(.006)	(.001)
FintechShare*Rates	.249***	.062***	.099***	.040***
	(.009)	(.009)	(.006)	(.009)
N	27760	27760	27760	27760
Adj. R-squared	.050	.011	.020	.005
Year Fixed Effects	✓	✓	✓	✓
County Fixed Effects	✓	✓	✓	✓

Chapter 2: The Aftermath of Credit Booms: Evidence from Credit Ceiling Removals

1. Introduction

Do credit booms sow the seeds of their own demise, as Minsky (1974) hypothesized? Several recent papers have presented evidence in support of Minsky's hypothesis (Borio and Lowe, 2003; Schularick and Taylor, 2012; Greenwood and Hanson, 2013; Mian, Sufi, and Verner 2017; Lopez-Salido, Stein, and Zakrajsek, 2017). Underlying this hypothesis is the idea that extended stretches of financial calm can help inflate credit booms and booms in real estate and other asset classes. These booms eventually collapse, leading to banking crises and deep recessions. Moreira and Savov (2017), and Greenwood, Hanson, and Jin (2019) model these boom-bust dynamics with time-varying "credit market sentiment" arising from over-extrapolation of default rates in the recent past.

This paper analyzes these theories by studying the removal of "credit ceilings" in an international panel of countries from 1950 to 2016. "Credit ceilings," as we refer to them throughout the paper, were systems of rigid direct controls on bank credit, enacted in several countries in the post-World War II period. These policies imposed maximum permissible growth rates of bank loans each year, at a time when corporate bond markets or nonbank lending were, by law, either highly circumscribed or nonexistent. As the global financial system modernized, and as international sentiment tended toward financial deregulation in the 1970s and 1980s, these countries gradually dismantled their credit ceilings, allowing bank credit to expand freely. We hypothesize that the removal of strict, administratively imposed caps on bank lending are likely to precipitate sharp accelerations in the extension of bank credit, as previously suppressed lending is unshackled. The resulting growth of credit has the potential to boost economic conditions and asset prices in the short-term, and thus to facilitate optimistic sentiment in credit markets. We thus view these credit ceiling removals as a useful natural experiment for evaluating whether exogenous shocks to the availability of credit can, themselves, generate boom-bust cycles of the type discussed

by Minsky and modeled by later authors.

To study the effects of credit booms in the context of ceiling removals, we assemble macroeconomic panel data on 39 economies over the period 1950 to 2018, and identify 13 countries which remove credit ceilings. We implement a local projections-instrumental variable (LPIV) framework of Jordà, Schularick and Taylor (2020). Our approach centers on the use of credit ceilings removals as an instrument for the growth of credit relative to GDP, while simultaneously controlling for the normal feedback between credit cycles and the real economy. In the first stage of our analysis, we find that credit ceiling removals are followed by a large and sudden increase in domestic bank-credit relative to GDP. We observe substantial increases in credit to GDP ratios in 12 out of 13 cases, with credit-to-GDP increasing by an average of about 8 percentage points and reaching an average peak after three years. These increases in credit are mainly found among institutions and credit types most affected by deregulation.

In the second stage, we find that in the short run, this sharp credit growth coincides with increases in GDP growth, investment, asset prices, and real estate construction (the boom), while in the medium run, these macroeconomic and asset price booms are often followed by sharp reversals, and often a banking crisis (the bust). Specifically, we find that if quantities are measured relative to the credit boom peak at t = 3 years after a credit ceiling removal, real GDP declines by 1%, bank stocks decline by 10%, house prices decline by 2%, and residential investment declines by 3%, on average, over the subsequent five years. Following11 out of 13 credit ceiling removals, a banking crisis occurs within five years after the initial credit boom. These effects appear to be unique to the removals of credit ceilings, are not observed following other types of deregulation.

We then document three phenomena associated with the aftermath of credit booms. The first phenomenon is the "calm before the storm" described in theory by Greenwood, Hanson, and Jin (2019), that the business cycle starts to turn before the credit cycle does, as banks continue to expand their lending even after GDP has started to decline. Specifically, we find that after credit ceilings are removed, GDP hits its cyclical peak on average 1-2 years prior

to the subsequent peak of the credit cycle. This pattern, whereby GDP ebbs before credit growth falters, occur in 9 of the 11 cases where there is a clear downturn in real GDP within the ten years following credit ceiling removals. The second phenomenon we call "successive bubbles", as several different types of asset price booms (residential real estate, commercial real estate, and stock prices) inflate and peak in succession. We argue that bankers and investors often chase other lending opportunities and asset classes, in the final stages of a boom, after one asset class starts to deflate. The third phenomenon is the "irreversibility" of credit booms, as credit booms "take on a life of their own" and are resistant to regulatory efforts to reimpose control. Indeed, we document that in the six countries in which credit controls are reimposed within five years of ceiling removal, credit continues to rapidly expand in three and is only partially moderated in the other three. Moreover, all six countries still experience banking crises within five years of the new restrictions.

In addition to uncovering these three phenomena after credit booms, our paper expands the existing literature in two important ways. First, while a large body of work provides historical evidence that credit booms tend to precede banking crises, economic downturns, and asset price crashes (Schularick and Taylor, 2012; Greenwood and Hanson, 2013; Baron and Xiong, 2017; Mian, Sufi, Verner, 2017; Lopez-Salido, Stein, and Zakrajsek, 2017)—the evidence in these papers is mainly correlational, with the notable exception of Mian, Sufi, and Verner (2019), discussed below. As a result, this has led some to question to what extent these patterns are causal (Santos and Veronesi, 2018; Gomes, Grotteria, and Wachter, 2018; Gomes, Grotteria, and Wachter, 2019).[51] In contrast, our paper takes an instrumental variables approach using credit ceiling removals, and we discuss our identification strategy, along with potential limitations, below and also in Section 4.

[51] Santos and Veronesi (2021) show that a number of the stylized facts on high leverage and subsequent downturns can be obtained in a purely frictionless model in which investors have heterogeneous endowments and risk-preferences. Gomes, Grotteria, and Wachter (2018) similarly show that risky lending and increased leverage can arise as a result of risk-shifting by banks, in a model with government guaranteed deposits. The authors suggest that in good times, banks take smaller risks to protect the expected economic rents that they earn as a result of deposit insurance, but increase risk as the probability of a downturn increases, and franchise values decline. Gomes, Grotteria, and Wachter (2019) show that the measure of issuer quality from Greenwood and Hanson (2013) is driven, in the time-series, by low-quality firms that rapidly repay their debt during downturns (rather than by excessive issuance of these firms during credit booms). They show, in a frictionless model with heterogenous firms possessing differential investment opportunities and exposures to rare economy-wide risks, that rapid credit growth and higher issuance (or lower repayments) by low-credit quality borrowers empirically forecast economic downturns, even though these downturns are exclusively driven by real investment opportunities rather than credit frictions.

Second, it is unclear from this prior literature which types of deregulations are associated with subsequent credit booms and busts. While Mian, Sufi, Verner (2019) study bank branching deregulation, which is unique to the U.S., their paper leaves open a broader question of why credit booms inflate in other countries. An earlier literature on international financial crises shows that financial crises tend be preceded by deregulations (e.g., Kaminsky and Reinhart, 1999). Our paper goes further by building a new database of other types of financial deregulations and showing the key role of credit ceiling removals across a variety of developing and advanced economies. Lastly, Mian, Sufi, Verner (2019) leave open the question of whether rapid credit expansion may itself generate economic and financial instability, or whether credit supply shocks simply make households more vulnerable to economic downturns originating elsewhere. As the authors study a single business cycle, across US states with varying degrees of deregulation, their empirical approach does not show whether credit growth itself increases the probability of financial instability. In contrast, our study suggests that credit ceiling removals can themselves lead to credit booms that subsequent go bust.

Our identification strategy rests on the assumption that credit ceiling removals are generally exogenous events uncorrelated with the business cycle or other policy shifts that may have also affected lending and financial stability. Our arguments in support of this assumption are as follows. First, as we further describe in Section 2, credit ceilings were mainly removed for ideological reasons, usually a desire to modernize monetary regimes (to use interest rates and open market operations rather than quantitative controls) and a desire to enhance economic competition in the banking sector; they were generally not implemented due to macroprudential or business cycle considerations. Their particular timing often coincided with a political shift: for example, the removal credit controls in the U.K. in 1971 coincided with the election of Edward Heath's Conservative government, which implemented a broader deregulatory agenda.

Second, as we show, credit ceiling removals coincide exactly in time with sharp inflection points in bank credit, as these events are nearly all followed by rapid accelerations in lending.

The timing does not suggest that credit ceiling removals were implemented either in response to a downturn (to stimulate the economy) or after a boom had already started (to cater to banks seeking to take advantage of lending opportunities in a strong economy). Our local projection methodology likewise helps control for the normal endogeneity between the business cycle and credit cycle. In the context of our model, the rapid rise in bank credit coincident with credit ceiling removals appears as a break in prior trends.

Third, we show that these credit ceiling removals were distinct in time from other types of deregulatory policy (e.g., deregulations of interest rates, capital controls, foreign bank entry) and that the rapid rise in credit corresponds most closely in time with credit ceiling removals versus other types of deregulatory policy. This analysis helps isolate the effect of credit ceiling removals from other types of deregulations, which generally occurred in the same decade, though several years apart, as we document.

Lastly, we show these credit ceiling removals tended to more substantially affect certain institutions (e.g., large banks) and types of loans (e.g. real estate loans), helping us to isolate the consequences of these credit ceiling removals. Although this analysis is limited to a few countries where the policies were known to affect only some institutions or loan types—and where disaggregated data is available on these classes of institutions and loan types—this evidence is consistent with the rapid credit increase being driven by the credit ceiling removal.

Our study is related to Farhi and Werning (2016), Korinek and Simsek (2016), and Schmitt-Grohé and Uribe (2016) who construct macroeconomic models in which credit booms can boost the macroeconomy in the short-run but lead to financial instability in the medium-run. Our study is also related to Aikman, Bush, and Taylor (2016) who study the business cycle consequences of quantitative credit controls (a broader class of policies, among which credit ceilings are an important member) in the U.K. in the postwar period and find that, while quantitative credit controls reduced bank lending, there is mixed evidence on whether they affected output and inflation. Monnet (2014) studies the role of quantitative credit controls in postwar France as a key policy lever of the Bank of France

in managing inflation and employment. Other studies analyzing the effects of quantitative credit controls in various other countries include Romer and Romer (1993) for the U.S., Glocker and Towbin (2015), Sonoda and Sudo (2016), and Monnet (2016).

Our analysis proceeds as follows. In section 2 we describe the credit ceiling policies and the broader regulatory system in 13 countries that enacted these policies after WWII. We outline why we believe credit ceiling removals were not systematically related to broader events in the macroeconomy or financial system. Section 3 discusses our data. In section 4 we use an LPIV approach to study the link between credit growth, asset prices, and the macroeconomy. Section 5 discusses three new phenomena that we relate to credit booms. In section 6 we discuss other financial reforms in the countries we study and show that credit ceiling removals appear to coincide more closely in time with inflection points in credit growth than any other type of policy reform. Section 7 links our results to the recent discussion of macroprudential policy, and we also briefly discuss how China has prominently used quantitative credit controls, with varying degrees of success, to manage their macroeconomy in the aftermath of the 2007-8 global financial crisis.

2. Institutional Setting

In this paper, we study the imposition and the subsequent removal of a set of policies referred to here as "credit ceilings." Credit ceilings were implemented in a number of countries (hereafter referred to as "credit ceiling countries") in the decade immediately following the World War II, and took the form of tight restrictions on the quantity of loans and other forms of credit that could be extended by financial institutions over a particular time period. In identifying credit ceiling countries, we maintain the narrow criteria that only explicit norms, which specified a maximum growth rate for loans or other bank-held assets, would qualify as credit ceilings for our purposes. Statutory reserve requirements and other portfolio restrictions which placed implicit caps on lending do not qualify as credit ceilings for the purposes of this study. We maintain this narrow definition in order to isolate the set of

credit policies most likely to place a binding constraint on the aggregate supply of credit, and which operate in a straightforward way through a credit supply channel.[52][53]

Our appendix contains extensive documentation of the dates of credit ceiling removals, descriptions of how credit ceilings were implemented in each individual country in our sample, and commentary on policymakers' motivations for implementing and removing credit ceilings and other major financial policies. To understand how credit ceilings functioned in the credit ceiling countries in our sample, we gather a number of primary sources, including government budget documents and central bank reports, and accounts written by contemporaneous scholars analyzing such policies while they were in place. We use this section to summarize some general insights and to highlight common themes that unite the financial reform processes undertaken by a number of these countries.

Across the countries we study, credit ceilings were generally implemented by central banks via a mixture of formal decrees and informal guidance. In Sweden, for example, "[b]ank actions were continuously scrutinized by the Riksbank and views on proper bank behavior were communicated in weekly meetings between the governor and representatives of the major banks. As one result of these meetings, the banks would commit to keep their lending within certain limits. It was only in 1974, however, that a law was passed giving the Riksbank the right to impose legally binding regulations" (Englund, 2015). Sweden also provides an example of how these ceilings could vary across institutions and loan types: Hodgman (1976) writes, "The Riksbank imposed a ceiling on the rate of expansion of bank loans for purposes other than house building...In 1974 the ceiling rate on loan expansion over a 12 month period was 18 per cent for commercial banks, 10 per cent for savings banks and 13 per cent for cooperative banks." In our analysis we exploit some of these differences

[52] One caveat is some of the policies that we label as credit ceilings were technically labelled as supplementary reserve requirements. That is, a country might maintain a primary reserve requirement, which is rarely if ever changed. At the same time, they might set a maximum rate of loan expansion (i.e. a credit ceiling) accompanied by an additional reserve requirement. This additional requirement would serve as a penalty for exceeding the credit ceiling. Banks exceeding the ceiling would then be required to maintain an additional quantity of non-interest-bearing reserves at the central bank which would vary as a function of the amount by which it exceeded the ceiling. This supplementary reserve requirement would often be so high (particularly against the backdrop of moderate to high inflation) that it effectively maintained a hard credit ceiling, even if banks were technically allowed exceed it.

[53] Portfolio requirements, and other forms of directed credit that primarily target the composition of bank asset holdings, rather than the quantity, similarly do not qualify as credit ceilings for our purposes.

across institutions and loan-types in furtherance of our argument that credit ceiling removals were the cause of the credit growth with which they are empirically associated. We provide a further description of how these policies operated across institutions and loan types, in section 6 and in the appendix.

Understanding the motivation underlying the implementation of these policies is important in determining whether credit ceiling reforms might reasonably be considered exogenous to the broader credit cycle. Credit ceilings were implemented not as macroprudential policies (that is, for financial stability, a purpose which is almost never mentioned in contemporaneous accounts) but to keep inflation in check by directly controlling the growth of monetary aggregates. For example, French monetary authorities formalized and tightened their credit ceiling policies in 1972, in the midst of a bout of inflation. A French Banking regulator, the National Credit Council, noted in its 1974 annual report that "[t]he policy of restrictive credit, put in place at the end of 1972, has, to a large extent, contributed to limiting the growth of the money supply to 15% over the course of 1973...the main objective of the monetary policy that followed in 1974 was to slow down the rise in prices appreciably, without causing too marked a fall in economic activity." A general barrier faced by a number of countries that used credit ceilings was that money markets had not sufficiently developed for central banks to use open market operations to fine tune the money supply over the short-term.[54]

While the countries that implemented credit ceilings after WWII all broadly used these policies to control inflation, there are also a number of idiosyncratic historical factors that led these countries to adopt credit ceilings. Thus, it is does not appear that the set of countries that adopted credit ceilings did so for a set of systematic reasons which would set them apart from the other countries in our study, without these ceilings. For example, the Bank of Japan used quantitative restrictions on lending in part because it did not have the independent authority to set most interest rates.[55] The heavily controlled interest rate

[54] In addition to being a tool to control inflation, credit ceilings and other quantitative controls of lending were also related to government efforts in many countries to use credit policy in furtherance of national priorities, such as ensuring governments' access to cheap sources of funding, and channeling funds to priority sectors, such as agriculture and exports.

[55] It was able to adjust the rate at which banks could borrow from its discount window, but these would to transmit to other

regime was largely the purview of the Ministry of Finance, rather than the Bank of Japan, and so the BOJ did not have the ability to control inflation via interest rates and open market operations (see Rhodes and Yoshino, 2007).[56]

Over the course of two decades, from 1971 to the early 1990s, these credit ceilings were removed in all the countries we study. The complete list of credit ceiling countries, and the dates that these ceilings were removed, is shown in Table 2.1. Further notes on these policies can be found in the appendix. As can be seen from Table 2.1, although we identify only 13 countries with credit ceilings, the set of countries represents a fairly diverse swath of large economies during the postwar period and there is thus little evidence to suggest that the sample is biased in favor of countries with particular traits.[57] Our sample of credit ceiling countries includes mostly advanced economies, but also includes some emerging markets; it contains countries from every continent with the exception of North America; and it contains countries with highly regulated financial sectors as well as those with more market-oriented policies.[58]

For all the credit ceiling countries, we can identify formal credit ceilings in place for several years, and in most cases, more than a decade prior to their repeal. These policies were not completely static during the time they existed, and there were periods in which these restrictions were made more and less binding. Nonetheless, for most of the countries we study, we view the dates of their removal as being sharply defined.[59] Nonetheless, for most of the countries we study, we view the dates of their removal as being sharply defined.[60]

rates in the economy that were controlled by the MOF.

[56]The Ministry of Finance, in turn, set rates with an eye toward channeling cheap credit to industry, and was thus not principally concerned with inflation. See Suzuki (1987) for further discussion.

[57]Our empirical approach of using country fixed effects means that we generally focus within-country variation in our outcome variables, mitigating concerns about underlying differences between countries.

[58]We do not list the dates that these policies were established. The reason for this is that several of the countries that enacted credit ceilings initially did so in an informal manner, where central bankers conveyed lending preferences privately to banks, even before statutory authority was given to set binding credit ceilings (as in the case of Sweden, discussed above).

[59]In a small number of cases, we were able to identify sources suggesting that a country had quantitative credit guidelines for some stretch of time, but where there were insufficient details on the specifics of these policies to determine whether they would constitute credit ceilings for our purposes. Our sample is thus not necessarily exhaustive of the set of countries globally that had some form of credit ceilings.

[60]In a few countries, the exact deregulation dates can be slightly ambiguous. For example, in France, the central bank dismantled credit ceilings by first announcing, in 1984, its long-term intention to remove credit ceilings, then lowering the penalty for exceeding the ceilings, in 1985, before finally removing the ceilings altogether in 1987. In this case, we choose 1987 as the date of deregulation. In Japan, credit ceilings ("window guidance") were legally abolished in the 1990s but were effectively discontinued in the 1980s, with some debate among scholars about the exact year they stopped being enforced. In most cases other than these, however, credit ceilings were removed all at once, and there is little ambiguity surrounding the relevant dates.

We view the removal of credit ceilings as being random events in the sense that they were not systematically related to the state of the macroeconomy at the time they were removed, or to anticipated future economic outcomes. Instead, credit ceilings were generally removed for ideological reasons. For example, in the UK, liberalization came in 1971 after a surprise win by the Conservative party over the incumbent Labour party. The Conservatives tended more toward economic liberalism than Labour, campaigned on a pledge of broad economic deregulation, and when in power sought to foster greater competitiveness in financial markets. In a note written to officials at Her Majesty's Treasury, Bank of England executive John S. Fforde opined that "prolongation of the present system [of credit controls] is inconsistent with the Bank's fundamental and correct view that the shape of the banking industry. . . should not be notably subordinated to the requirements of monetary policy. Banking, as a legitimate commercial activity, often inconveniences the Government of the day. There is accordingly a persistent temptation to convert the banks into mere slaves of official policy. We have always said, and rightly, that this is a temptation that must be resisted" (Goodhart, 2014). In Japan, officials were also concerned about the distortions credit controls imposed on credit markets and the extent to which prolonged credit controls could essentially freeze the market shares of individual banks, with a former Bank of Japan official noting that "if strong window advice [i.e. a credit ceiling] is continued for long periods, there emerge disequilibria among financial institutions between those that are subject to controls and those that are not. In addition, the lending shares within one type of financial institution tend to become fixed" (Suzuki, 1987).

There is also some evidence that policymakers, perhaps mistakenly, assumed that credit ceilings were rendered ineffective by financial modernization, and no longer served a purpose. For example, in Japan, according to Shigehara (1991), "given a significant progress in the de-regulation of financial transactions both domestic and international, Japanese non-bank borrowers' access to alternative sources of finance and innovations in financial engineering could negate the effect of compulsory control on the volume of domestic bank lending." It should be noted that strong credit growth in the wake of credit ceiling removals in these

countries casts doubt on the notion that financial innovations had rendered ceilings to be completely ineffective.

Lastly, we should note that credit ceiling removals were separate in time from other types of financial deregulation. For example, in Sweden, interest rate restrictions were liberalized in 1978; restrictions on capital flows, issuance for private sector bonds, and foreign ownership of Swedish equity shares were removed in 1980; foreign banks were allowed to enter in 1986; restrictions on foreign exchange controls and foreign asset holdings were removed in 1989, and bank branching deregulations occurred in 1990. Figure 2.9 in Appendix II is a chart of the various types of deregulations in each of the countries, showing as in Sweden that credit ceiling removals were separate in time from other types of financial deregulation. In Section 6, we formally analyze the timing of these other types of deregulation and show that the specific credit booms we study correspond most closely in time with credit ceiling removals versus other types of deregulatory policy.

3. Data and Summary Statistics

Our main analysis focuses on country-level data from an unbalanced panel consisting of 39 countries. Our dataset contains macroeconomic and financial variables measured at an annual frequency over the years 1950-2016. The full list of countries that appears in the sample is available in the appendix alongside further details on variable construction.

Our main data source is Baron, Verner, and Xiong (2020), from which we obtain country-level data on bank credit, GDP, inflation, returns on bank equity indices, and the authors' list of banking crisis dates. We also make use of an alternative set of banking crisis dates from Reinhart and Rogoff (2009) for robustness. We add data on house prices and real residential fixed investment. Data on house prices come from Jorda, Schularick, and Taylor (2017), Global Financial Data, and central bank websites. Data on real residential investment is constructed using data from the OECD and from CEIC. In our analysis, all variables are expressed in real terms and are deflated using data on CPI inflation from Baron, Verner,

and Xiong (2020). For our analysis, bank stock prices, GDP, house prices, and residential investment are normalized relative to their 1980 values, so that their 1980 levels assume a value of 100. When data for a particular variable and country begins after 1980, that country's observations are normalized relative to the first year that data for the variable-country pair appears in the dataset. Such normalization ensures a relatively equal weighting of observations across countries. In a small number of cases, observations in the early part of the sample were generated via linear interpolation of the surrounding years. Such interpolations were made when there were no more than two consecutive missing years, and it was ensured that no data were interpolated in years surrounding credit ceiling removals.

Our key independent variable is the bank credit to GDP ratio. Our aggregate credit variable comprises all credit extended by domestic banks to residents (i.e. consumers and non-financial firms) with the exception of foreign currency loans. Aggregate credit data come from Baron, Verner and Xiong (2020), who in turn gather data from Jorda, Schularick, and Taylor (2017), the BIS long credit series, and a various of newly transcribed historical sources. Following the literature on credit booms we scale the aggregate credit variable by domestic GDP in order to focus on credit growth in excess of GDP growth.

Summary statistics on the variables are presented in Table 2.2. Table 2.2 shows an upward trend in all variables, including the bank credit to GDP ratio. Most of the secular growth in this variable has occurred since 1970, as noted by Schularick and Taylor (2012), with growth rates picking up even more substantially in much of the world after 1990.

We compile information on the existence and removal dates of government-imposed credit ceilings in a number of countries in our sample, as described in Section 2. After identifying the list of countries with credit ceilings and finding the dates of their removal, we also construct a larger database with timelines of various types of significant financial deregulations in these countries.

Finally, we collect some additional data which we use for robustness checks and for analysis of the three phenomena we study related to the aftermath of credit booms. We collect data on commercial real estate prices and on the level of corporate investment in real

estate. Commercial real estate data comes from the BIS commercial property database[61], from central banks and other government sources such as Statistics Sweden, as well as from a number of academic papers written at the time that credit ceiling policies were in effect. Corporate investment in real estate is similarly drawn from a variety of sources including OECD, CEIC, and central bank sources. Both of these data series give us limited data coverage, and as such we do not include them in the main LPIV analysis, given the more limited sample. However, we analyze this data around credit ceiling removals in the relevant parts of our analysis.

4. Credit Ceiling Removals and Subsequent Credit Booms and Busts

In this section, having identified a set of 13 countries that implemented and then removed credit ceilings, we introduce and estimate the LPIV specification of Jordà et al. (2020). Our approach uses credit ceiling removals to instrument for credit growth while using several auto-regressive terms to control for the broader state of the macroeconomy. We highlight the behavior of several real variables (e.g. GDP growth and residential fixed investment) and depict the evolution of financial conditions (e.g. bank stock prices and real estate prices) in the wake of credit booms, showing how these booms eventually go bust. These results are robust to a variety of specifications, including reduced form estimates and event studies.

4.1 Credit Growth and Macroeconomic Variables Around Credit Ceiling Removals

We first examine the behavior of country-level credit to GDP ratios in periods surrounding credit ceiling removals. We find evidence consistent with the notion that credit ceiling removals led to credit booms in the short- to medium-term.

Before beginning our formal econometric analysis using an LPIV framework, we first plot credit-to-GDP in each country with credit ceiling removals, which provides a preliminary

[61] For more information on the construction of these series, see https://www.bis.org/statistics/pp_commercial.htm, and Eurostat (2017): "Commercial property price indicators: sources methods and issues," Publications Office of the European Union.

visual presentation of our results. Figure 2.4 in Appendix II plots the behavior of these ratios over ±5-year windows surrounding credit ceiling removals for each of the 13 credit ceiling countries. Table 2.8 in Appendix II similarly shows country-level data for the credit ceiling countries, showing changes in credit to GDP ratios in the three years after credit ceiling removals as compared to the country-level averages for all three-year windows outside of the post-liberalization periods. In 12 of the 13 countries, credit grows faster in the three years after credit ceilings are removed than it does during the average of all other three-year windows in that country's credit history.[62] Table 2.3 quantifies these initial findings on credit growth. On average, credit to GDP ratios increase by 10.9% in the three-year period that follows credit ceiling removals. The average rate of credit expansion during all other periods is only 3.7%.

There does not seem to be a systematic relationship between credit ceiling removals and macroeconomic conditions at the time these removals took effect. As illustrated by the individual country plots in Figure 2.4, in many countries, such as Australia, Italy, South Africa, and Portugal, credit was decreasing in the leadup to credit ceiling removals. In other countries, such as Japan, Sweden, Chile, and France, credit was relatively flat or increasing moderately. There is thus little evidence that countries were in the midst of a boom prior to ceiling removal. There are also no instances in which a country removed ceilings in the middle of a financial crisis or a major recession. There is thus little to suggest that the financial problems experienced by countries after deregulation were already underway prior to liberalization. After liberalization however, credit grows rapidly in almost every country in the sample. In some countries (e.g. Argentina, Norway, France) credit growth begins in the very year that ceilings are removed, while others (Sweden, Portugal, the UK, New Zealand, and Italy) see credit growth take off in the year after liberalization. As we will show in Section 6, no other set of policy reform in these countries sees this kind of rapid credit growth within a short window of their enactment.

[62] The lone exception is Austria, which is the only country in which we do not see a sharp rise in credit in the years after credit ceilings are removed. It is also one of only two countries that does not experience a financial crisis in the ten-year period after ceiling removal.

In order to summarize the progression of credit and other macroeconomic variables, across all the countries in our sample, we estimate local projections of Jordà (2005) to trace out the impulse responses of various outcomes to removals of credit ceilings. Specifically, we consider regressions of the following form:

$$y_{i,t+h} = \alpha + \beta^h * Liberalize_{i,t} + \sum_{k=1}^{4} \gamma_k^h * y_{i,t-k} + \sum_{k=1}^{4} \delta_k^h * X_{i,t-k} + \varepsilon_{i,t+h}^h \tag{1}$$

For h = 0, 1, ... 12.

Here, y is one of the response variables of interest, which we describe below; the subscript i represents one of the 39 countries in our sample (i.e. the 13 credit ceiling countries, and the 26 other "control" countries) while the t subscript represents the year. *Liberalize* is an indicator variable which takes a value of one for country i, and year t, if country i removes its existing credit ceilings in year t. The *Liberalize* variable takes a value of zero in all other country-year pairs. Thus, for the 26 countries in the sample which never enact credit ceilings of the form described in Section 2, the *Liberalize* variable will assume a value of zero in all years.

The response variables y that we consider in this initial specification are real GDP, bank stock prices, real house prices, and real residential fixed investment. We also initially choose to estimate our regressions in levels, as recommended by Hamilton (2018). Each of these variables, with the exception of credit to GDP ratios, are expressed in log form in the estimating equations. For control variables X, we include up to three separate macroeconomic controls, including three lags of GDP, inflation, and credit to GDP ratios.[63] We include lagged values of the response variable y on the right-hand side of the equation, as we want to ensure that the coefficient on our *Liberalize* variable captures only innovations to the country-level time-series of a given variable that occur after ceilings are removed. If

[63] We include inflation in our set of controls because the primary purpose of instituting credit ceilings was to establish control over inflation. Thus, it is perhaps plausible that these ceilings were removed only when inflation was deemed to be dormant. The reasons for including lagged GDP and lagged credit to GDP ratios are somewhat more obvious. We ultimately seek to determine the effect of credit growth on the broader economy. If credit ceilings are removed in a way that coincides with the underlying state of the macroeconomy (for instance if credit ceilings are removed in the midst of broader economic booms, or if they are removed when growth is low with the intention of spurring credit and GDP growth) it will be important to control for these factors.

credit ceiling removals tend to coincide with periods of strong economic growth, we do not want to incorrectly attribute strong growth after ceiling removals to effects that result solely due to variables' persistent character.

Figure 2.1 shows the dynamic response of credit to GDP ratios to credit ceiling removals by plotting the sequence β^h in the set of equations above for h = 1, 2, ... 12. We adopt the convention of normalizing the estimated impulse response by subtracting from each coefficient the estimated coefficient from the time zero regression. Red lines in each picture trace out our point estimates for the effects of credit ceiling removals on each of the response variables, while dashed blue lines depict 95% confidence intervals. The standard errors used to compute confidence intervals are robust to clustering at the country level.

Figure 2.1 shows a rise in credit growth in the years after credit ceilings are removed. We estimate that removal of credit ceilings raises credit to GDP ratios by more than 7% in the three years after ceiling removal, relative to the counterfactual case where ceilings remain in place.[64] Credit to GDP ratios remain substantially elevated up to five years after ceilings are removed but begin to contract thereafter. In total, we see a contraction by roughly 5% from years 5-12, though we do not see an aggregate reversal.

The same cannot be said of the other variables. Real GDP remains relatively unchanged in the first two years after ceiling removal but begins to drop at a relatively steady rate thereafter. After 12 years, we find that GDP levels are roughly 20% lower as a result of credit ceiling removals. In this plot, we notice that the upper bound of the 95% confidence interval crosses below zero, indicating significantly lower GDP growth in the years after ceilings are removed than we would expect in similar periods without ceiling removals. Similar patterns are observable in the other variables as well, although these variables, if anything, present stronger evidence of boom-and-bust cycles. In general, these variables exhibit similar timing, with respect to their peak values, as credit to GDP. Bank stock prices, residential investment, and house prices, all see initial increases in the three years after ceilings are removed and then experience declines thereafter. Of these, house prices have perhaps the most dramatic collapses. After their peak, real estate prices decline by roughly 20% over the ensuing nine years. The general pattern of real estate price boom and

[64] This is a very similar result to our earlier "non-parametric" analysis.

bust cycles is visible across a number of individual countries, including the UK, Norway, Sweden, France, and Australia. As we will show later, these aggregated results perhaps understate or obscure the disaggregated episodes, since some credit cycles turned more quickly than others. Bank stock prices and residential real estate investment exhibit similar patterns to house prices; of these variables, bank stocks have the most precipitous single-year drop (7 years after ceilings are removed) but rebound in the years thereafter.

4.2 Estimating the Effects of Credit Ceiling Removals in a LPIV Approach

We now turn to our more formal econometric analysis, in which we explore the dynamic relationship between credit ceiling removals, credit growth, and macroeconomic outcome variables using a local projections-instrumental variables (LPIV) approach of Jordà, Schularick, and Taylor (2020). For our analysis, we formalize our conception of credit cycles, modeling them as a combination of an initial three-year "boom" phase, and a "bust" phase from years 4-12. In a first stage, we use credit ceiling removals to instrument for credit growth during the initial boom phase and then trace out the effect of fitted credit growth on macroeconomic outcomes during the bust.[65]

Specifically, in the first stage we instrument for credit growth using the *Liberalize* variable:

$$\Delta_3 Credit-to-GDP_{i,t} = \alpha^{s1} + \beta^{s1} * Liberalize_{i,t}$$
$$+ \sum_{k=1}^{4} \gamma_k^{s1} * \Delta_1 Credit-to-GDP_{i,t-k} + \sum_{k=1}^{4} \delta_k^{s1} * \Delta_1 X_{i,t-k} + \varepsilon_{i,t}^{s1} \quad (2)$$

where the *s1* , superscripts attached to the coefficients and error term in the above equation indicated that these are "stage 1" coefficients. We then use the predicted values of *Credit-to-GDP* generated from equation (2) to estimate

[65] We noted previously that high levels of credit growth are apparent up to t=5. Our results are not substantially different if we use year five as the peak of the credit boom rather than year three.

$$\Delta_h y_{i,t+3} = \alpha^{s2} + \beta^{s2,h} * \Delta_3 Credit - \hat{}to - GDP_{i,t}$$
$$+ \sum_{k=1}^{4} \gamma_k^{s2,h} * \Delta_1 y_{i,t-k} + \sum_{k=1}^{4} \delta_k^{s2,h} * \Delta_1 X_{i,t-k} + \varepsilon_{i,t+3+h}^{s2,h} \quad (3)$$

where the *s2* superscript above each coefficient and the error term indicates that these are our "stage 2" estimates. The "hat" marker over the *Credit-to-GDP* variable indicates that these are predicted values of the credit to GDP ratio generated from estimating (2). We adopt the same set of control variables that we used for our analysis in the previous section. Our dependent variables are now expressed as differences, so that we can assess how their values progress relative to their levels at the height of the credit boom. We adopt the notational convention that for any variable y, $\Delta_h y_{i,t} = y_{i,t+h} - y_{i,t}$; thus, *Credit-to-GDP* is constructed by subtracting the time zero value of a country's credit to GDP ratio from its value at time three.[66] It should be noted that the X vector above includes lagged values of a country's credit to GDP ratio as well as lagged GDP and inflation.

Figure 2.2 plots the impulse responses, $\beta^{s2,h}$ obtained from equation (3), for each of the variables of interest in the period after a credit boom. The impulse responses are defined as the responses to a one hundred basis point increase in a country's credit to GDP ratio. We leave horizontal space on the left-hand side of each plot, with a solid vertical bar at t=3 as a reminder that the impulse responses show the behavior of each variable after a credit shock, which is assumed to take place during the previous three years. The vertical axes are expressed in decimal form, so that a value of .01, for example, would indicate that a 1% increase in a country's credit to GDP ratio over the preceding three years would generate a 1% increase in the dependent variable.

Figure 2.2 panel A shows how credit evolves in the years after the initial boom. We can see that after a three-year credit supply shock, credit to GDP ratios are gradually mean reverting. While credit remains relatively flat for the first few years after the initial boom,

[66] We do not difference the control variable inflation, since inflation is already expressed as a first difference of consecutive years' CPI levels.

in years 4-9 after the initial credit boom (or 7-12 years after the beginning of the boom) credit recedes and total growth over the period is negative, suggesting that credit booms eventually reverse.

These results, and those below, are also shown in tabular form in Table 2.4, which reports the estimates β^h of the coefficients on the *Credit-to-GDP* variable in equations (3) and (4). For the sake of brevity, the table presents coefficients beginning four years after the credit boom (h=4 in the equations above, or seven years after the *beginning* of the credit boom). Standard errors, in parentheses beneath each coefficient, are again robust to clustering.

Turning back to Figure 2.2, Panel B illustrates the dynamics of aggregate output after a credit boom. The impulse response suggests that GDP growth turns negative immediately after the peak of the credit boom, declining at a relatively stable rate across the nine years after the shock. By year four after the credit boom, the response is statistically significant at the 5% level. By the ninth year after the credit boom, the coefficient on Credit-to-GDP reaches a level of -1.673, suggesting a 1% shock to Credit-to-GDP over a three-year window leads to a 1.67% reduction in GDP in the subsequent nine years.[67]

Panels C-E of Figure 2.2 show some of the phenomena that accompany the fall in output that we observe in the aftermath of credit booms. Panel C shows a large decline in bank stock prices that occurs four years after the credit boom. Baron and Xiong (2017) present similar evidence across an international panel and a long time-series showing an association between credit expansion and subsequent bank stock price declines. Theoretical work by Holmstrom and Tirole (1997) and He and Krishnamurthy (2013) suggest that shocks to the net-worth of financial intermediaries can be an important channel in propagating economic distress. The β^h coefficient for h=4 in the bank stock price regressions is -7.445, suggesting that bank stock prices fall more than 7% for every 1% increase in credit to GDP ratios occurring during the initial credit boom.

[67] To put this figure into perspective, consider that the removal of credit ceilings led to roughly 7% growth to a country's credit to GDP ratio. Suppose that a country, in the absence of a credit boom, would be expected to grow at an annual rate of 4%. After nine years, we would then normally expect this country's total output to be 42.3% higher than it was at the beginning of this period. If, instead, this country experienced a credit boom in the years preceding this period, of the same size as the average credit boom generated by credit ceiling removals, then we would expect total economic growth over the nine years after the credit boom to be only 30.6% (or 42.3 – 7*(-1.673)). Thus, this average-sized credit boom would wipe out roughly 28% of the growth that this country would have otherwise expected over a nine-year period.

Panels D and E show the dynamics of real estate prices and residential fixed investment following a credit boom. Recent work by Jordà, Schularick, and Taylor (2015) and Mian, Sufi, and Verner (2017) suggests that housing booms and busts can be an important driver of financial crises, either because banks hold substantial quantities of real estate loans on their balance sheets or because housing wealth is an important part of the household balance sheet. Panel D shows that house prices exhibit a relatively steady decline in the years after a credit boom, with the largest single-year price drops coming five and seven years after the initial credit boom. The cumulative negative effects of the credit boom on house prices attain statistical significance, at the 5% level, by year eight after the boom. The results of estimating the effects of credit growth on residential real estate investment echo the results that we find in house prices, with large drops in residential investment levels in years five and six after the credit boom, followed by relatively flat growth in the ensuing years.

The results we present in Table 2.4 show that credit booms also have the ability to forecast banking crises. Using lists of banking crises assembled by Baron, Verner, and Xiong (2020) and Reinhart and Rogoff (2009), we define a new variable that captures the total number of banking crises that occur in a country over a particular period of time, with crisis dates defined according to one of these lists. The time windows for which we construct these variables are the same windows that we look at when constructing our impulse responses. Thus, our crisis variables are defined as the total number of banking crises that occur in a country in expanding time windows of length 4-9 years, and the regression coefficients on these variables should be interpreted as the marginal contribution of a 100-basis point credit supply shock to a country's expected number of financial crises over a given window. As shown in Table 2.4 panel E, which uses the Reinhart and Rogoff (2009) crisis list to construct the expected crisis variable, the coefficient for the h=4 regression is 4.97, which suggests that an increase in a country's credit to GDP ratio of 1% increases the expected number of financial crises that the country will experience in the four years after the credit boom by .0497. The β^h coefficient in the Reinhart and Rogoff crisis regressions peaks at a value of 7.55 for the year h=6. Based on this coefficient, a credit supply shock of 7% over a

three-year credit boom would increase the number of crises that the country could expect to suffer in the six years after the credit boom by 7*.0755=.528.[68]

The analogous results in Table 2.4 panel F suggests a strong association between credit booms and future banking crises. For both the Baron, Verner, and Xiong (2020) crisis list and the Reinhart and Rogoff (2009) crisis list, we observe large coefficient values on the Credit-to-GDP variable, suggesting that credit booms predict large increases in the expected number of crises a country will experience. For both sets of variables, we obtain coefficients that are statistically significant at the 5% level over the entire time horizon that we study. In total, 11 of the 13 countries that dismantle credit ceilings in our sample experience a financial crisis in the 10 years after ceilings are removed. The only countries that do not experience financial crises are Austria, which also does not experience any kind of credit boom, and Portugal. In the 11 countries that do experience crises, we observe a total of 14 distinct financial crises, with three countries (Argentina, Chile, and South Africa) experiencing two crises in the years after liberalization. The occurrence of such a large number of crises is striking since banking crises are relatively rare events in the larger sample.

5. Newly Uncovered Phenomena Associated with Credit Booms

In this section we look more closely at the periods surrounding credit ceiling removals. We newly document three empirical patterns on the behavior of key variables over the boom-and-bust phases of the credit cycle.

5.1 The Asynchronous Nature of Business and Credit Cycles: The Calm Before the Storm

The first of the phenomena that we present is the calm before the storm phenomenon of Greenwood, Hanson, and Jin (2019). While their paper is mainly theoretical, the authors

[68] Note that an increase of .52 in the expected number of crises a country will experience does not necessarily suggest a 52% increase in the probability of a single crisis. It could, for example, correspond to a 26% increase in the probability that a country will suffer two crises and a 0% change that the country will suffer from a single crisis.

note that in the prelude to the 2008 financial crisis in the United States, credit continued to grow strongly well into mid-2008, even as GDP had already started to decline. They show that this sequence of events has also played out across a number of smaller recessions in the U.S. In the model that they develop, investors form beliefs, in part, by extrapolating past market outcomes (i.e. the occurrence of defaults) in addition to looking at firm fundamentals (i.e. cash flows). In such a setting, credit markets can become detached from fundamentals in the late part of the credit cycle, when cash flows recede but do not yet trigger defaults.

We seek to investigate whether the calm before the storm phenomenon is visible across the credit booms of our study. Plots for each country of real GDP and real credit growth, in windows surrounding credit ceiling removals, are plotted in Figure 2.5, in Appendix II. Black vertical bars mark the years of credit ceilings removals, while vertical gray bars mark banking crises.

To systematically analyze turning points of business cycles and credit cycles, we use the following procedure of Hamilton (2018) to effectively detrend variables and then identify peaks. We begin by regressing GDP and credit on their four most recent lagged values, in separate regressions for each variable and credit ceiling country. We then compute forecast errors from these regressions and label as the start of downturns the first year in a sequence of years where we observe either two consecutive negative forecast errors, or one large negative forecast error, defined as observations belonging to the bottom 10th percentile for that country.[69] We define business cycle or credit cycle peaks as the last year before the beginning of a downturn in GDP or credit, respectively.

We use two approaches to analyze the timing relationship between business cycle and credit cycle peaks. The first is simply measuring the average number of years between the peaks of these cycles in the years immediately following a credit ceiling removal. We assign a positive value to the difference in the number of years between cyclical peaks if the business cycle peaks before the credit cycle, and we assign this difference a negative value

[69] We can define the beginning of a business cycle or credit cycle upturn (i.e. the transition from a contraction to an expansion) analogously, by looking for the first of two consecutive positive forecast error, or a single large positive forecast error. However, our analysis focuses more heavily on identifying downturns.

if the credit cycle peaks first. The results of this exercise are reported in the first column of Table 2.5 panel A. On average, business cycles tend to hit their peak 1.30 years before credit cycles. Some examples in which the calm before the storm phenomenon is easily visible in country-level data include Argentina, where the business cycle hits its peak three years before the credit cycle (1980 vs. 1983), and Norway, where GDP hits its peak in 1986, four years before credit recedes.

Our second approach involves looking at windows around financial crises occurring after the initial credit boom. Since 11 of the 13 credit ceiling countries experienced banking crises in the 10 years after ceiling removal, we use the dates of these banking crises, as given by the Baron, Verner, and Xiong (2020) crisis list to anchor our analysis. We choose the business cycle and credit cycle turning points that most closely coincide with these dates as the relevant cyclical turning points for evaluating the calm before the storm phenomenon. We consider this second approach for evaluating the distances between turning points, even if less algorithmic, to be the more natural of the two approaches.[70] When we apply this new approach to measuring distances between peaks, the average distance between credit cycle and business cycle turning points increases slightly. The average distance between peaks is 1.86 years, closely matching the historical experience of the 2008 financial crisis in the United States where these two cycles peaked two years apart. More notably, it lowers the volatility in the distances between peaks. The upshot of this refinement is that the average distance between peaks of business and credit cycles attains strong statistical significance. Of the 11 countries, only one of these countries (Japan) saw the credit cycle turn before the business cycle; there were two cases in which the cycles were perfectly synchronized, while the remaining eight countries saw business cycles which hit their peaks between 1-4 years before credit showed signs of a downturn. These results are noteworthy because, as noted by Greenwood, Hanson, and Jin (2019), in most models that feature representative investors with rationally formed beliefs, the business cycle and credit cycle are one in the

[70] It allows for the removal of two countries, Austria and Portugal, where there was no clear boom and bust cycle following deregulation, and it allows us to focus on the largest (and most obvious) economic shocks emerging in the wake of credit ceiling removals.

same.

5.2 Successive Bubbles

Can investor sentiment spill over from one asset market to another? Do bankers and investors often chase other lending opportunities and asset classes, in the final stages of a boom after one asset class starts to deflate? In this section we present evidence that appears to suggest the affirmative to both these questions. Specifically, we study markets for both residential and commercial real estate and find that rapid price appreciation in one of these two asset classes tends to be followed by a later-emerging price boom in the other. Then, as the credit cycle turns, we often observe continued price growth in this secondary asset class even after the initial price decline in the earlier forming bubble, before prices ultimately plummet across both asset classes. We find similar evidence of successive booms in the quantities of investment in residential and commercial real estate.[71][72]

Figure 2.6 in Appendix II presents plots of prices in residential and commercial real estate markets for each individual country, while Figure 2.7 presents analogous plots for investment quantities. While a number of these countries experience strong price growth in commercial real estate prior to the removal of credit ceilings, in six out of seven cases for which we have sufficient data, residential real estate prices do not show evidence of rapid price appreciation in the pre-liberalization period. After ceilings are removed, however, price gains appear to spill over into residential real estate markets, perhaps because the removal of credit restrictions allowed bankers greater freedom to chase new opportunities after commercial real estate markets began to overheat. In France, for example, a residential real estate boom begins in 1988, one year after ceilings are removed, and several years after commercial real estate prices begin their rapid ascent. In Australia, Norway, Italy, and Sweden, we similarly see smaller price booms in residential real estate, following credit

[71] Our analysis sometimes suffers from limitations with respect to the availability of data. In many of the countries in our sample, long time-series on commercial real estate prices are simply unavailable.

[72] It should be noted that the data on real estate prices and real estate investment do not necessarily reflect prices and investment on directly comparable structures. For example, commercial real estate price data often reflects office prices in a country's largest cities, while corporate real estate investment data covers investment in a broader range of structures that includes office buildings, factories, airplane hangars, and certain types of infrastructure.

ceiling removals and earlier spikes in commercial real estate prices. After the credit cycle begins to enter its contractionary phase, both residential and commercial real estate markets experience price collapses.

To examine the timing of these peaks, for each country where we have residential and commercial real estate data, we define peaks as the highest value that these series obtain at any point during the sixteen-year period beginning five years before ceiling removals and ending ten years after.[73] Using this approach, we find that the two series tend, on average, to peak 0.75 years apart, a difference which is statistically significant when tested against the null hypothesis that the two series have cyclical peaks that coincide.[74] While the differences in peak dates is not large, it is notable in a number of these cases that prices continue to rise in the late-developing market, even in the midst of price collapses in the earlier-developing market. Italy, Norway, and France, all see residential real estate prices continue to grow, even after commercial real estate prices see contractions of more than 10%.

The evidence that we collect from the price data is reinforced by the data in investment quantities.[75] We define investment peaks as the last year prior to the largest price decline that occurs in the years surrounding the financial crisis. When applying this approach, we again find data suggesting asynchronous peaks in the residential and commercial real estate markets. We find that the beginnings of investment booms, in the buildup to financial crises, tend to occur 0.66 years apart and that peak investment in these markets tends to occur roughly a year apart.

Taken together, we interpret our findings on real estate prices and quantities as consistent with the notion that, as markets in that asset class begin to overheat, and lending standards become increasingly loose, lenders may shift to a new asset. Such an account is anecdotally

[73] One could make the case that New Zealand is tougher to identify. The paths of residential and commercial real estate prices do not entirely align with the patterns we observe more generally, as both assets see strong growth in the years prior to credit ceiling removals, while residential real estate prices see a smaller price contraction a year before the financial crisis in 1987.

[74] We test the null hypothesis that the mean difference between cyclical peaks of the earlier-forming (generally commercial real estate) and late forming bubble is zero years.

[75] The investment data are somewhat more volatile than the price data, with cycles that are less clear in illustrating stark booms and busts in many cases. We thus use banking crisis dates to anchor our investigations of cyclical turning points, in a similar structure to section 5.1. We and look for large price declines that occur within a short window surrounding financial crises in each country. We thus exclude Austria from this analysis, since it does not experience a crisis.

consistent with the experience of bankers in the lead-up to the 2008 financial crisis, when residential real estate price growth drove the credit cycle, but began to subside as early as 2006, with stock prices and commercial real estate prices maintaining robust growth until 2007 and 2008, respectively.

5.3 "Irreversibility" of Credit Booms

We next present evidence suggesting that once credit ceilings are removed, the ensuing credit boom can "take on a life of its own" and be difficult to contain. Given the length of time it takes for a credit boom to inflate and then reverse, one tempting conclusion would be that if policymakers were to take early steps aimed at gradually bringing credit back down, that they could avoid the perilous after-effects of credit cycle downturns. Unfortunately, there appears to be little evidence that reimposing credit restrictions is effective in reversing the underlying forces that contribute to financial instability.

Of the 13 countries that remove credit ceilings in our sample, six of these countries reimpose new restrictions in the years after removal. Table 2.5 panel C shows the list of countries that instituted new restrictions in the years after credit ceilings were removed and displays the number of years of credit growth that these countries experienced in the years after new controls were put in place.

It should be noted that none of these countries fully restored the regulatory regimes that were in place prior to ceiling removal.[76] We also do not know, in all cases, what specific factors regulators were responding to when they reimposed these partial controls. However, in a number of cases, narrative evidence suggests that regulators were responding to greater than anticipated credit growth and potentially overheating financial markets. In the UK for example, after robust credit growth in the two years after ceiling removal, the Bank of England imposed new rules requiring banks to post additional reserves at the central bank if deposit growth exceeded a certain threshold. In Italy, after robust credit growth in the

[76] We also do not know, in all cases, what specific factors regulators were responding to when they reimposed these partial controls.

two years after deregulation, authorities partially reinstated credit ceilings in 1986, with the understanding that such restrictions would be temporary, rather than a permanent fixture of the regulatory regime. These restrictions were subsequently removed after a year. Norway announced new supplementary reserve requirements in 1986, two years after deregulation, and Japan re-issued lending guidance for the largest "city banks" in 1989.

The evidence appears mixed on whether these new restrictions were narrowly successful in dampening credit growth. When assessing the likely effects of these policies more broadly, however, it appears clear that they did little to stem the tide of the broader credit cycle or to prevent crises. While the ratio of credit to GDP decreased in Italy, in 1986, when temporary ceilings were put in place, this did not appear to stop credit growth in its tracks or to reverse the broader trend toward higher credit and greater financial instability. After temporary credit controls were removed, credit grew even more rapidly, and the banking system experienced a crisis in 1990 followed by weak growth and poor credit conditions in the subsequent years. After the UK installed new regulations in 1973, credit growth remained relatively flat in the following year. However, the financial system suffered a crisis in 1974 and credit contracted in the following year. After Japan instituted credit restrictions in 1989, credit subsequently receded relative to GDP (though growth continued in real terms) but the economy went into a recession shortly thereafter and there was a devastating financial crisis in 1991. Finally, credit growth in Norway continued to surge after supplementary restrictions were mandated in 1986. Finance companies and other non-bank institutions began to fail in 1987 and there was a wave of commercial bank failures in 1990.[77]

The historical experience of countries that install new restrictions after initial periods of financial deregulation suggests that it is difficult for regulators to fine tune the financial system using selective credit controls. There are a number of reasons this may be the case.

[77] The other two countries that imposed controls after removing credit ceilings did so in the midst of crises. Argentina reinstalled some restrictions on credit growth, while South Africa placed new restrictions on capital flows, with the assumption that one of the problems leading to their crisis in 1985 was unchecked growth of credit from abroad. Since these policies were installed during or after banking crises, disentangling their effects from the banking crises that were already underway is impossible. Nonetheless, it is noteworthy that both of these countries slid into crisis for a second time, shortly after their first crisis episodes.

It may be that regulators simply overestimate their ability to contain bank credit and are insufficiently aggressive with policies meant to reverse the tide of credit booms. In Italy, and the UK, credit barely budged when new controls on banks were put into place, and credit surged in Norway in spite of new controls. This suggests that there may be some momentum to credit growth once a boom is underway. Moreover, if deregulation in these instances worked primarily by unleashing buoyant sentiment on markets, then containing downturns may not be as simple as restraining credit, even if new policies are effective in doing so. If bankers extend credit in part by irrationally extrapolating prior credit market outcomes, as in Greenwood, Hanson, and Jin (2019), then the strong growth and low default rates that emerge shortly after deregulation may further boost sentiment, making it even harder to control markets as time goes on.

6. Additional Analyses Linking Credit Ceiling Removals to Subsequent Credit Booms

In this section we look to solidify our key argument that credit ceiling removals were the true cause of credit growth that we subsequently observe, and that this explosion of credit generated the ensuing economic declines. We address two possible sources of concern with our identification. The first is that credit ceiling removals may not have been the true driver of credit growth during the period we study, and that other structural changes, such as increased financial globalization or strong household credit demand, caused credit growth and also generated financial instability. If, for example, rapid credit growth during the 1970s and 1980s stemmed from consumer credit demand shocks rather than supply factors, one could argue that credit growth was not a true cause of later downturns, but only served to amplify them through a household balance sheet channel, as argued by Mian, Sufi, and Verner (2017). In the first part of this section, we will argue further that credit ceiling removals were the underlying cause of ensuing credit growth by showing that the types of credit and institutions that were most severely constrained under the credit ceilings were

the institutions that most rapidly expanded credit in the wake of deregulation. In the second part of this section, we will address a second concern, namely that some other set of government policies, adopted alongside credit ceiling removals, generated the correlations we observe. We address this possibility by systematically studying other financial policies adopted in credit ceiling countries.

6.1 Analyzing Types of Loans and Financial Institutions Most Affected by Credit Ceilings

In this section we attempt to make the causal interpretation of our empirical evidence more convincing by showing that in some of the credit ceiling countries, we can identify institutions that were differentially affected by credit ceiling policies. We use the knowledge of how policies were geared toward specific institutions to improve our identification of policy-created credit booms. We show that those institutions identified ex-ante as more constrained by credit ceilings were the institutions that drove credit growth after liberalization. In another instance, we show that institutions that remained constrained to a larger degree after deregulation did not expand credit as freely after ceilings were removed. The main results of our analysis are presented in Figure 2.3. Vertical lines in these figures mark the dates credit ceilings are removed.

We begin with a further discussion of credit policies in Norway. Our analysis centers on a comparison of private commercial banks with government-controlled state banks ("Statsbankene"). While the system of credit controls on the private banks was removed in 1984, the state-banks, regulated via a different mechanism, were still subject to government control.[78]

We plot the credit supplied by state banks and the private commercial banks in Norway from 1977-1991 in panels A and B of Figure 2.3. Figure 2.3 panels A and B show that prior to the removal of credit ceilings state-bank and private bank credit largely grew in tandem with one another: the annual credit supplied by the two types of institutions hovered within

[78]The state-banks were a central part of the financial system beginning in the years after WWII. The state-banks were instrumental in carrying out the government's lending policies, helping to ensure that credit was channeled to sectors of the economy that the government deemed to be high priority, or which were believed to be inadequately served by private credit markets. While lending by the state-banks was geared toward particular sectors of the economy, taken as a whole, state-bank lending appeared to cover a relatively broad and diverse swath of the economy, from small-business loans to consumer lending.

a narrow band, until 1984, when credit ceilings were removed. However, after 1984, credit rapidly accelerates among private banks but remains steady at state-banks, as can be seen in Figure 2.3. The fact that these two types of credit grew largely in parallel with one another before liberalization, and that state-bank credit continued to grow according to the same linear trend (as seen in panel B) post-liberalization, suggests that credit ceilings were likely binding for these institutions prior to reform. Rapid acceleration of private-bank credit after liberalization is consistent with the notion that deregulation generated a credit boom among the private banks.[79]

Another country where we can observe differential effects of credit ceilings is Australia. Among the credit ceiling countries we study, Australia appeared to have one of the more advanced financial systems, with non-bank financial institutions that included merchant banks (investment banks and securities firms), finance companies, and permanent building societies (which financed home loans).[80] Credit ceilings only affected the trading banks (the largest commercial banks) and the savings banks but not the non-bank financial institutions mentioned above.

Panel C of Figure 2.3 plots the total outstanding credit supplied by banks and non-bank financial institutions (NBFIs), both deflated by real GDP, between 1975 and 1995.[81] Credit ceilings were removed in 1982.[82] The values depicted in the plot are normalized, so that the outstanding credit of both banks and NBFIs assumes a value of 100 in 1982.[83]

[79] It is also unlikely that the explosion in credit by private banks merely reflected a compositional shift in institutions supplying credit to the economy. That is, it is likely that private banks didn't merely reclaim a share of the lending that would have otherwise been supplied by other institutions. Although the share of credit supplied by institutions regulated outside of the system of credit ceilings, such as finance companies and insurance companies, had been increasing prior to the removal of credit ceilings, likely at the expense of private banks, in 1980 these firms combined to make up only 11.4% of total outstanding credit. By 1985, after deregulation, their share had actually increased to 12.4%. Since credit expansion at private banks was quite large, and since the non-bank sector was quite small, it is unlikely that the expansion of bank credit only reflected a compositional shift in the supply of credit rather than a real expansion of the aggregate credit supply.

[80] While banks remained central to the financial system, there was opportunity for non-bank entities, particularly the specialty mortgage companies to expand at the expense of banks whenever credit controls on banks were particularly rigid. According to Hall (1987): "largely as a result of regulatory 'straightjackets', savings banks' share of total financing declined markedly between 1953 and 1982 from 20 per cent of total assets of all financial institutions to 13 per cent. This performance contrasts sharply with that of their major competitors, the permanent building societies and credit unions."

[81] Normalizing by real GDP "compresses" the graph somewhat, but does not change the overall pattern, relative to displaying total real credit without deflating by GDP, since both the bank credit and NBFI credit series are divided by the same GDP number.

[82] Identification of the entities subject and not subject to ceilings may not be perfect here. Among the banks, were institutions, like the state banks, which were not regulated by the federal government and thus not subject to credit ceilings. Thus, the outstanding bank credit numbers depicted in Figure 2.3 panel C likely included credit of institutions that were not covered by credit ceilings.

[83] While this normalization facilitates easy comparisons between banks and NBFIs in the years following deregulation, the

Panel C shows that, in the decade prior to the removal of credit ceilings, the NBFIs gradually but steadily gained market share.[84] The years after liberalization saw a stark reversal of this trend. In the five years after credit ceilings were removed, credit supplied by banks, relative to GDP, increased by 64.0% relative to its 1982 value, with the NBFI credit to GDP ratio increasing by only 17.4%.[85]

In the final piece of this analysis, we turn to the case of Sweden, for which we compare credit expansion by the largest commercial banks to credit extended by smaller regional and savings banks. In Sweden, for reasons described below, credit ceilings made it even more difficult for mid-sized and smaller banks to lend than the larger institutions. Thus, after the credit ceiling removals, the increase in credit growth is mainly driven by these previously constrained banks.

In the 1970's and 1980's credit markets were largely dominated by the three largest commercial banks: Handelsbanken, Skandinaviska Enskilda Banken (SEB), and Post-och Kreditbanken (PK-Banken). National policy favored these larger banks. For example, Englund (2015) notes that PK-Banken was state-owned and that its growth was a priority of the Social Democrat regime that gained control of the government in 1982. Additionally, the credit ceilings put in place by the Riksbank seemed to favor credit expansion by larger banks. For example, according to Hodgman (1976), when the credit ceilings in were formalized in 1974, the maximum rate of credit expansion for commercial banks was 18% over a 12-month period, while the smaller savings and cooperative banks were limited to 10% and 13% credit expansion, respectively. Moreover, while the credit ceilings remained in place in Sweden until 1985, certain other restrictions, which had been more rigid for the larger banks than for small banks, had already been lifted earlier in the 1980's.[86] Thus, by 1985, deregulation

picture would not be substantially different if we instead plotted the raw dollar values of credit of these institutions.

[84] In September of 1976, the NBFIs had 15.4 billion Australian dollars (AUDs) of credit outstanding, compared to 18.6 billion supplied by the banks. By the time credit ceilings were removed in 1982, the NBFIs had narrowly surpassed the banks, holding 41.1 billion AUDs of outstanding credit, compared to 39.1 AUDs for banks.

[85] A glance at the plot of Australian credit in Figure 2.4 in Appendix II shows that in the aggregate, Australia experienced a large credit boom after ceilings were removed in 1982. From 1982 to 1990, credit to GDP increased from around 23% to 50% of GDP. Meanwhile, the share of credit provided by the institutions most constrained under the regulatory regime that preceded the reforms of the 1980's increased rapidly at the same time. If increased credit provision after 1982 was solely the result of a positive shock to credit demand, there would be little reason why the total credit provided by banks and by NBFIs, which had moved in lockstep during the eight years before deregulation, would diverge so sharply after 1982. Similarly, it is unlikely that deregulation changed only the allocation of credit across institutions, without changing the aggregate amount of credit provided, since total credit boomed after 1982.

[86] For example, according to Hodgman (1976), liquidity ratio requirements, which were lifted in 1983, were set at higher rates

was somewhat further along for the larger banks than for smaller ones.

Panel D of Figure 2.3 shows outstanding loans of the three largest Swedish banks mentioned above (shown in red) as compared to loan provision by a collection of smaller regional and savings banks (shown in blue). While the aggregate real supply of outstanding credit in Sweden increased by over 15% from 1985-1986, outstanding loans of the three largest banks actually decreased slightly. In contrast, loans supplied by the smaller banks increased dramatically in 1986, expanding by more than 30%. Ultimately, both the larger and smaller banks played a large role in the credit boom, as large-bank credit expanded by roughly 70% from 1985-1989. Anecdotal evidence suggests that the results are even more stark when comparing banks, which were all subject to ceilings to some degree, to the finance companies and other "gray market" providers of credit, which were not.[87]

6.2 Relationship of Credit Ceiling Removals to Other Financial Deregulation Policies

In previous sections, we have made note of the fact that credit ceiling removals were often a part of a broader set of financial deregulations, though the process of financial deregulation often spanned more than a decade and major financial reforms, other than credit ceiling removals, generally occurred at separate times. In this section, we conduct a more thorough analysis of the other policy changes in financial markets that occurred during the period we study.

There are two factors that motivate this analysis. The first is related to identification concerns. A potential concern is that that the credit booms we observed were precipitated by policies other than credit ceilings, of which credit ceiling removals were merely a part. (In some sense, such a confounding variable might not matter to the ultimate conclusions of our study. Suppose, for example, that credit ceiling removals coincided nearly perfectly

for the largest commercial banks than for the smaller commercial banks. The savings banks and agricultural credit associations tended to face even lower liquidity ratio requirements. Hodgman puts the liquidity ratios in 1974 at 30% for larger commercial banks, 24% for smaller commercial banks, 27% for the post office bank, and 20% for savings banks and agricultural credit associations.

[87] While we do not have reliable data on the extent of the gray market provision of credit in Sweden, Englund (2015) suggests that "the institutions that had been most directly hit by the regulations now expanded most rapidly, banks by 174 per cent and mortgage institutions by 167 per cent between 1985 and 1990."

with removals of interest rates restrictions, which were the true cause of the credit booms we observe in the data. If removals of these interest rate restrictions also generated exogenous credit supply shocks, it would not change our conclusions on the aftermath of credit booms, even if we misattributed the ultimate source of these credit shocks.)

The second reason is to gain an understanding of why credit booms emerge in the first place, which types of policies may give rise to them, and which types of macroprudential policies could be used to prevent them. Thus, it is useful to contrast the effects of credit ceiling removals with the effects of other types of financial deregulations.

In this analysis, we narrow the set of countries in our ensuing analysis to the set of 13 credit ceiling countries, in order to focus on compiling detailed accounts of the regulatory environments in this subset of countries. We then construct detailed timelines of financial reforms across our 13 credit ceiling countries focusing on several broad classes of policies that could plausibly affect credit growth or financial stability. Our major policy categories are: interest rate restrictions, branching restrictions, barriers to entry of new banks, specialization requirements (e.g. rules prohibiting banks from underwriting or trading securities), international capital controls, and controls in non-bank credit markets (e.g. bond or commercial paper markets). In some cases, there are slight disagreements between different sources regarding the years in which policy changes are enacted, while in a few other cases the distinctions between different types of financial reforms are blurry. Nonetheless, it does not appear that these sources of uncertainty are sufficient to meaningfully change our results. We describe the set of specific financial reforms and the set of sources from which we determine the dates these reforms were enacted in the appendix.

As a preliminary visual indicator of our results, Figure 2.9 in Appendix II plots the deregulatory timelines associated with each country by overlaying these reforms on the progression of credit to GDP ratios in each country. The timing of credit ceiling removals does not follow any immediately obvious pattern relative to the removals of other types of restrictions. In general, most of these countries begin a gradual process of financial market deregulation beginning in the 1970s, and extending through the 1980s and into the early

1990s, in some cases. Credit ceiling removals often appear to be roughly in the middle of a country's broader history of reforms, though perhaps somewhat closer to the end than the beginning. In a number of cases, it appears that credit ceiling removals come right at the beginning of an inflection point in the progression of credit. For example, in Sweden and Australia, credit ceilings come at the end of long periods of relatively static credit to GDP ratios, and at the beginning of steep and sustained increases.

After compiling our regulatory timelines, we create a series of indicator variables, with one variable for each of the six policy varieties mentioned above, that take a value of one in year t and in country i, if country i *removes* a restrictive policy of that type in year t.

Table 2.9 in Appendix II reports correlations between these policy indicator variables for each pair of policies, which are close to zero in all cases, indicating that these policy reforms rarely coincided in time. The type of reform with which credit ceiling removals coincide most closely is the removal of interest rate restrictions, as the two indicator variables corresponding to these policies have a correlation coefficient close to .2, corresponding to three instances (in New Zealand, Sweden, and the UK) in which credit ceilings and interest rate controls are removed at the same time.

To compare credit ceiling removals to other types of financial deregulations, we first ask how well each type of policy predicts credit growth. In particular, we ask whether the estimated effects of credit ceiling removals survive controlling for other types of deregulatory policies. We thus estimate regressions of the following form:

$$\Delta_h Credit - to - GDP_{i,t} = \alpha^h + \beta^h * Policy_{i,t} + \gamma^h * \Delta_1 X_{i,t-1} + \varepsilon_{i,t+h} \tag{4}$$

For h=3.

The *Policy* variable can be a vector in some specifications (that is, we will look at specifications with a single policy variable as well as those with multiple policy variables). The vector X contains the control variables which consist of lagged GDP growth, and the lagged investment to GDP ratio.

Table 2.6 panels A and B display the results of this analysis.[88] Panel A shows results from specifications where we look at each policy-variable individually, for the sake of comparing the raw predictive power of each of the types of regulatory reforms (indicated in the first column of the table), without considering the collinearity between different types of policies. These results indicate that credit ceilings are the only variable that strongly predicts credit growth over a three-year period, in settings with a single policy-variable. Credit ceilings, though assessed in a simpler regression format than our earlier analyses, nonetheless appear to predict three-year credit growth of 7.6%, which is similar in magnitude to our earlier estimates. Among the other policy variables, none predicts credit growth over a three-year period anywhere close to as large as credit ceiling removals, and none of the others are statistically significant. In magnitude, the closest competitor to credit ceiling removals is the removal of barriers to entry of foreign banks, which forecast 3.3% growth to a country's credit to GDP ratio over the three years that follow removal.

Panel B of Table 2.6 displays results for regression specifications that include multiple policy variables at the same time. The first six specification pair credit ceiling removals with each individual other deregulation to ask whether any of the policy variables other than credit ceilings can significantly diminish the estimated effect that credit ceilings have on credit to GDP growth. In the final column of panel B we show the results of the "horse race" regression where we include all policy variables together. The results in panel B suggest that the additions of multiple policy variables do little to affect the ability of credit ceiling removals to predict credit growth. The magnitude of the coefficient on the credit ceiling policy-variable is remarkably stable across the seven specifications, ranging from a value of roughly 7.4% to a maximum of 7.7%. The only other policy variable with a coefficient that attains a marginal level of statistical significance is the "Fin. Reform" variable, which corresponds to removals of restrictions in the bond market (and other non-bank credit markets).

In a final piece of the analysis on alternate policies, we look to assess which types of

[88] All these results are broadly similar in specifications that look at a five-year credit growth horizon rather than at three-year windows.

financial reforms best forecasts the *beginnings* of credit booms in terms of timing. For any of the policies we analyze, it may be that their removals simply coincide with credit booms that were already underway at the time they were removed and that they forecast credit growth associated with the tail-end of a boom that began earlier. Given these considerations, we estimate the ability of each type of policy reform to predict a credit boom in the context of a probability model. Specifically, we estimate logistic regressions, where we assume that the probability of observing the beginning of a credit boom, in any country i, and any year t, takes the form:

$$Log\left(\frac{p_t^{boom}}{1-p_t^{boom}}\right) = \alpha + \beta * Policy_t + \Delta_1 X_{t-1} \tag{5}$$

Where the variable p_t^{boom} represents the probability that a credit boom will begin in year t. The policy variables and the set of controls take the same form as our linear regression specifications in equation (4). In particular, we estimate versions of the model of equation (5) that include a single policy variable, as well as versions that include multiple policy variables. We estimate model (5) using maximum likelihood methods. In order to estimate (5), we need to define a binary variable equal to one in the event that a credit boom begins at time t, and zero otherwise. To do this, we separate credit booms in our sample from periods of stasis or credit declines. A credit boom can begin, in our algorithm, in the first year in which a country experiences at least two-consecutive years of credit to GDP growth exceeding the median growth rate of the previous five-year window, or when the country experiences a single year of credit to GDP growth exceeding 1%. A credit boom may end if a country previously experiencing a credit boom subsequently sees three or more consecutive years of credit to GDP growth below the median growth rate experienced during the prior five year period. We adopt the restriction that a credit boom can only begin if it follows a prior period of stasis or credit contractions, that is, a new credit boom cannot arise as the result of credit growth that becomes even stronger once a country is already in a boom.

Our methodology reflects the intuition that a credit boom should be an inflection point

where credit growth suddenly rises relative to its recent past. In the appendix, we further discuss the intuition behind our credit boom classification and discuss results that arise from plausible alternative classification approaches. We believe that as compared to estimating models such as equation (4) or the Jorda (2005) projections that preceded them, the logit estimation in (5) is less sensitive to the specific parametric assumptions of a linear regression model, such as the correct lag structure for auto-regressive terms, the assumption that lagged values of credit growth affect credit linearly and with a constant relationship over time, and so on.

Table 2.7 displays the results of our logit estimation. Panels A and B display results of univariate and multivariate tests, in an analogous fashion to Table 2.6. Each column of Table 2.7 (for any of the panels) represents a separate regression specification and each row corresponds to a coefficient representing one of the types of policy reforms.

The results in panel A show a coefficient of 2.041 for the credit ceiling removal variable, when estimating (5) in a single policy-variable setting. This suggests that removal of credit ceilings, in an arbitrary year, is estimated to increase the log odds ratio corresponding to the probability of a credit boom, by 2.041. Straightforward calculations show that this translates to a credit ceiling removal being associated with an increase the probability that a credit boom will begin in that year from 10% to 46.1%.[89] In single policy-variable specifications, none of the other policy liberalization variables forecast the beginning of credit booms with any kind of statistical significance.

The results in panel B show that credit ceiling removals continue to strongly forecast credit booms in specifications with multiple policy variables. In bivariate specifications no other type of deregulatory policy substantially lowers the predictive power of credit ceilings, as the credit ceiling coefficient ranges from 2.023, in the specification in which it is paired with removals of branching restrictions, to 2.664 in the regression with all policy

[89] The proportion of years covered by the beginnings of Type I credit booms is roughly 10% in our sample. Suppose then that in an arbitrary year, we expect that the probability that a credit boom will begin is 10%. The associated log odds ratio, is then log(.1/(1-.1)). = -2.197. If we were to be told, in addition, that credit ceiling policies would be removed in that same year, we would revise our log odds ratio upwards by 2.041, to -0.156. This odds ratio corresponds to a credit boom probability of exp(-.156)/(1+exp(-.156)) = 0.602.

variables in tandem. These coefficients retain statistical significance at the 1% level across all specifications. Moreover, no other policy variable appears to be strongly associated with the beginnings of credit booms. These results suggest that it is unlikely that credit ceiling removals coincide with other types of policy changes that accounted for both the credit growth and financial instability that we attribute to credit ceiling removals in our sample. In the appendix, we discuss the results of additional tests of the robustness of our results, including tests which assess the predictive power of our policy variables in forecasting credit booms at different time horizons, and additional regression specifications where we include all possible permutations of policy variables. The results of these additional tests do little to suggest that other policy variables may be responsible for driving our results.

7. Conclusions

Recent research has documented a strong link between credit growth and subsequent economic and financial turmoil. Mian, Sufi, and Verner (2019) showed, using a natural experiment, that credit growth can substantially amplify business cycle downturns.. We have performed a similar analysis in an international context and have attained results suggesting that credit expansion may not merely amplify economic downturns, but may generate them on their own. While financial crises are relatively rare episodes, we found that 11 of 13 countries that removed credit ceilings experienced crises in the 10 years that followed these reforms. Credit booms were also associated with booms and busts in real estate markets and precipitous drops in bank stocks, suggesting that the banking sector was at the epicenter of the downturns that followed deregulation.

Our results do not identify exactly why these credit booms had such negative impacts on the broader economy. We take our findings as being consistent with the view of Minsky (1977) that periods of strong growth in credit markets can generate recklessness among investors. However, we cannot rule out the interpretation of credit booms by Krishnamurthy and Muir (2018) and others who argue that high credit growth and leverage serve to make

the financial sector more vulnerable to shocks that emanate elsewhere in the economy, since we do not have bank-level data that allows us to directly show that lending standards were loosened post-liberalization. We nonetheless have sympathy for the market sentiment view of credit booms, for two reasons.

First, if banker expectations are formed, in part, by erroneously extrapolating past rates of loan default into the future, as in Greenwood, Hanson, and Jin (2019), credit ceiling removals would be prime candidates for the emergence of overly optimistic sentiment. This is because these episodes involved the transition from tightly-rationed banking regimes, where central banks and other government agencies explicitly sought to minimize default rates, to regimes where banks were free to extend credit as they pleased. For example, Moe, Solheim, and Vale (2004) note that in Norway "[b]anks had been exposed to little credit risk during the regulatory regime that had more or less been in place between 1945 and 1984... because the regulatory regime did not allow any bank to expand its lending rapidly. Furthermore, the regime implied a rationing of credit that allowed banks to pick mainly the best credit risk among the queue of unsatisfied credit demand. Thus, when the quantitative regulation was lifted, banks had hardly any experience in how to operate in this new much more competitive environment." Similarly, in Japan Similarly, in Japan, Rhodes and Yoshino (2007) note that prior to deregulation, "the [Ministry of Finance] provided an implicit guarantee that no bank would be allowed to fail and no bank did." It is thus easy to imagine that low default rates pre-liberalization might lead to biased expectations, if bankers did not appreciate the true extent of the regulatory regime change.

Secondly, we view the three empirical phenomena we identify in section 5 as most consistent with a sentiment-driven view of credit cycles. The calm before the storm phenomenon of Greenwood, Hanson, and Jin (2019) is hard to square with most rational models where business cycles and credit cycles tend to be synchronized. Our evidence on the irreversibility of credit booms suggests that even when policymakers have put measures in place to control credit after the onset of the boom, such measures do not appear to prevent financial collapse. This result held even in situations in which levels of credit to GDP began to go

back down after an initial boom, suggesting that the level of credit to GDP may not be as important as the initial surge in credit. Meanwhile, the fragility narrative of Krishnamurthy and Muir (2018) relies on the notion that high ¬levels of credit growth make the economy more vulnerable to adverse shocks.

Our results may also speak, to some degree, to the potential usefulness of various forms of macroprudential policy. Our results certainly do not suggest that credit booms are the only factor that can precipitate financial crises, and thus we cannot say that policies unrelated to the supply of credit do not have some bearing on financial stability nonetheless. However, we can say that a wide variety of financial reforms, other than credit ceilings, appear to have little effect on the quantity of credit when they are removed. Thus, if we believe that a central goal of macroprudential policymakers should be to constrain uncontrolled growth in credit, our analysis suggests that the reinstitution of Glass-Steagall era specialization restrictions, or the reinstatement of branching restrictions may be of little use. In contrast, while draconian controls on bank lending may be impractical in the current regulatory environment, our analysis suggests that the adoption of policies meant to restrain rapid credit growth may prove useful as part of the regulatory toolkit.

References

Aikman, D., O. Bush, and A. M. Taylor, "Monetary versus macroprudential policies: causal impacts of interest rates and credit controls in the era of the UK Radcliffe report," NBER Working Papers, No. 22380.

Ballantyne, A., J. Hambur, I. Roberts, M. Wright, (2014). "Financial Reform in Australia and China," Research Discussion Paper No. 2014-10, Reserve Bank of Australia Research Discussion Paper Series.

Baron, M., E. Verner, and W. Xiong (2020). "Banking crises without panics," *Quarterly Journal of Economics*, vol. 136(1), pages 51-113.

Baron, M., and W. Xiong, (2017). "Credit expansion and neglected crash risk," *Quarterly Journal of Economics*, vol. 132(2), pages 713-764.

Barradas, R., S. Lagoa, E. Leão, and R. P. Mamede (2013). "Report on the financial system in Portugal," *FESSUD Studies*, Financialization, Economy, Society & Sustainable Development (FESSUD) Project.

Borio, C., and P. Lowe, (2002). "Asset prices, financial and monetary stability: exploring the nexus," BIS Working Papers, No. 114.

Englund, P. (2015). "The Swedish 1990s banking crisis: a revisit in the light of recent experience," Riksbank Macroprudential Conference, June 2015.

Farhi, E., and I. Werning, (2016). "A theory of macroprudential policies in the presence of nominal rigidities," *Econometrica*, vol. 84(5), pages 1645-1704.

Glocker, C., and P. Towbin (2015). "Reserve requirements as a macroprudential instrument - evidence from Brazil," *Journal of Macroeconomics*, vol. 44(c), pages 158-176.

Gomes, J., M. Grotteria, and J. Wachter, (2018). "Foreseen risks," NBER Working Papers No. 25277, National Bureau of Economic Research.

Gomes, J., M. Grotteria, and J. Wachter, (2019). "Cyclical dispersion in expected defaults," *Review of Financial Studies*, vol. 32(4), pages 1275-1308.

Goodhart, C.A.E. (2014). "Competition and credit control," London School of Economics FMG Special Papers sp229, Financial Markets Group, London School of Economics.

Goodman, J. B. (1992). "Monetary policy and financial deregulation in France," *French Politics and Society*, vol. 10(4), pages 31-40.

Greenwood, R. and S. G. Hanson (2013). "Issuer quality and corporate bond returns," *Review of Financial Studies*, vol. 26(6), pages 1483-1525.

Greenwood, R., S. G. Hanson, and L. J. Jin, (2019). "Reflexivity in credit markets," NBER Working Papers No. 25747, National Bureau of Economic Research.

Hall, M., (1987). "Financial deregulation: a comparative study of Australia and the United Kingdom," London, England: Palgrave Macmillan.

Hamilton, J. D. (2018). "Why you should never use the Hodrick-Prescott filter," *Review of Economics and Statistics*, vol. 100(5), pages 831-843.

He, Z., and A. Krishnamurthy (2013). "Intermediary asset pricing," *American Economic Review*, vol. 103(2), pages 732-770.

Hodgman, D. R. (1976). "Selective credit controls in Western Europe: A Study Prepared for the Trustees of the Banking Research Fund, Association of Reserve City Bankers," Chicago, Illinois: The Association.

Holmstrom, B., and J. Tirole (1997). "Financial intermediation, loanable funds, and the real sector," *Quarterly Journal of Economics*, vol. 112(3), pages 663-691.

Hoshi, T., D. Scharfstein, and K. J. Singleton (1993). "Japanese corporate investment and Bank of Japan guidance of commercial bank lending," NBER Chapters, in: Japanese Monetary Policy, pages 63-94, National Bureau of Economic Research.

Jordà, O. (2005). "Estimation and inference of impulse responses by local projections," *American Economic Review*, vol. 91(1), pages 161-182.

Jordà, O., M. Schularick, and A. M. Taylor (2015). "Betting the house," *Journal of International Economics*, vol. 96(s1), pages 2-18.

Jordà, O., M. Schularick, and A. M. Taylor (2017). "Macrofinancial History and the New Business Cycle Facts," NBER Macroeconomics Annual 2016, volume 31, edited by Martin Eichenbaum and Jonathan A. Parker. Chicago: University of Chicago Press.

Jordà, O., M. Schularick, and A. M. Taylor, (2020). "The effects of quasi-random monetary experiments," *Journal of Monetary Economics*, vol. 112, pages 22-40.

Kaufman, G. G., (1992). "Banking structures in major countries," Alphen aan den Rijn, Netherlands: Kluwer Academic Publishers.

Kleivset, C., B. Klunde, R. Troite, C. Venneslan (2011). "Independence within government: a comparative perspective on central banking in Norway 1945-1970," Norges Bank Working Paper.

Korinek, A., and A. Simsek, (2016). "Liquidity trap and excessive leverage," *American Economic Review*, vol. 106(3), pages 699-738.

Krishnamurthy, A. and T. Muir (2018). "How credit cycles across a financial crisis," NBER Working Papers 23850, National Bureau of Economic Research.

Liebscher, K., J. Christl, P. Mooslechner, D. Ritzberger-Grünwald, (2007). "Financial development, integration, and stability: evidence from Central, Eastern, and South-Eastern Europe," Cheltenham, England: Edward Elgar Publishing.

López-Salido, D., J. Stein, and E. Zakrajšek, (2017). "Credit-market sentiment and the business cycle," *Quarterly Journal of Economics*, vol. 132(3), pages 1373-1426.

Mahar, M. and, J. Williamson (1998). "A survey of financial liberalization," Princeton, New Jersey: International Finance Section, Department of Economics, Princeton University, Essays in International Finance No. 211.

Melitz, J., (1990). "Financial deregulation in France," *European Economic Review*, vol. 34(2-3), pages 394-402.

Mian, A., A. Sufi, and E. Verner (2017). "Household debt and business cycles worldwide," *Quarterly Journal of Economics*, vol. 132(4), pages 1755-1817.

Mian, A., A. Sufi, and E. Verner, (2020). "How does credit supply expansion affect the real economy? The productive capacity and household demand channels," *Journal of Finance*, vol. 75(2), pages 949-994.

Moe, T. G., J.A. Solheim, and B. Vale, "The Norwegian banking crisis", Norges Bank Occasional Papers Series 33/2004, Norges Bank, pages 1-33.

Monnet, E., (2014). "Monetary policy without interest rates: evidence from France's Golden Age (1948-1973) using a narrative approach," *American Economic Journal: Macroeconomics*, vol. 6(4), pages 137-169.

Moreira, A., and A. Savov, (2017). "The macroeconomics of shadow banking," *Journal of Finance*, vol. 72(6), pages 2381-2432.

Perez, S., and J. Westrup (2008). "Finance and the macro-economy: the politics of regulatory reform in Europe", Center for European Studies Working Paper Series #156.

Yoshino, N., and J. R. Rhodes (2007). "Japan's monetary policy transition, 1955-2004," GRIPS Policy Information Center, Discussion paper 07-04.

Rojas-Suarez, L., and S. R.Weisbrod (1995). "Financial fragilities in Latin America: the 1980s and 1990s," International Monetary Fund, Occasional Papers Series 1995/012.

Romer, C., and D. Romer, (1993). "Credit channel or credit actions? An interpretation of the postwar transmission mechanism," NBER Working Papers No. 4485, National Bureau of Economic Research.

Santos, T., and P. Veronesi, (2021). "Leverage," *Journal of Financial Economics*, forthcoming.

Schmitt-Grohé, S., and M. Uribe, (2016). "Downward nominal wage rigidity, currency pegs, and involuntary unemployment," *Journal of Political Economy*, vol. 124(5), pages 1466-1514

Schularick, M. and A. M. Taylor (2012). "Credit booms gone bust: monetary policy, leverage cycles, and financial crises 1870-2008," *American Economic Review*, vol. 102(2), pages 1029-1061.

Shigehara, K. (1991). "Japan's experience with use of monetary policy and the process of liberalization," *BOJ Monetary and Economic Studies*, vol. 9(1)

Sonoda, K., and N. Sudo, (2016). "Is macroprudential policy instrument blunt?" BIS Working Paper No. 536, Bank for International Settlements.

Suzuki, Y. (1987). "The Japanese financial system," Oxford, UK: Oxford University Press.

Verheirstraeten, A. (1981). "Competition and regulation in financial markets," London, England: Palgrave Macmillan.

Walsh C.E. (1988). "Financial deregulation and monetary policy in New Zealand," In: Cheng HS. (eds) Monetary Policy in Pacific Basin Countries, Dordrecht, Netherlands: Springer Books.

Tables and Figures

Figure 2.1

In this figure, we display impulse responses of various dependent variables to the removal of credit ceilings. Our impulse responses are estimated via a local projections approach, as described in Jordà (2005), and detailed in equation (1). Impulse responses (shown as red lines below) plot coefficients on the *Liberalize* variable in equation (1) over successive time windows. We plot impulse responses of credit to GDP, GDP, bank stock index values, residential real estate prices, and the level of residential real estate investment, in panels A-E, respectively. All variables are expressed in real terms, deflated, where necessary, by a country's CPI. 95% Confidence intervals, with standard errors adjusted for possible clustering at the country level, are displayed via blue dashed lines.

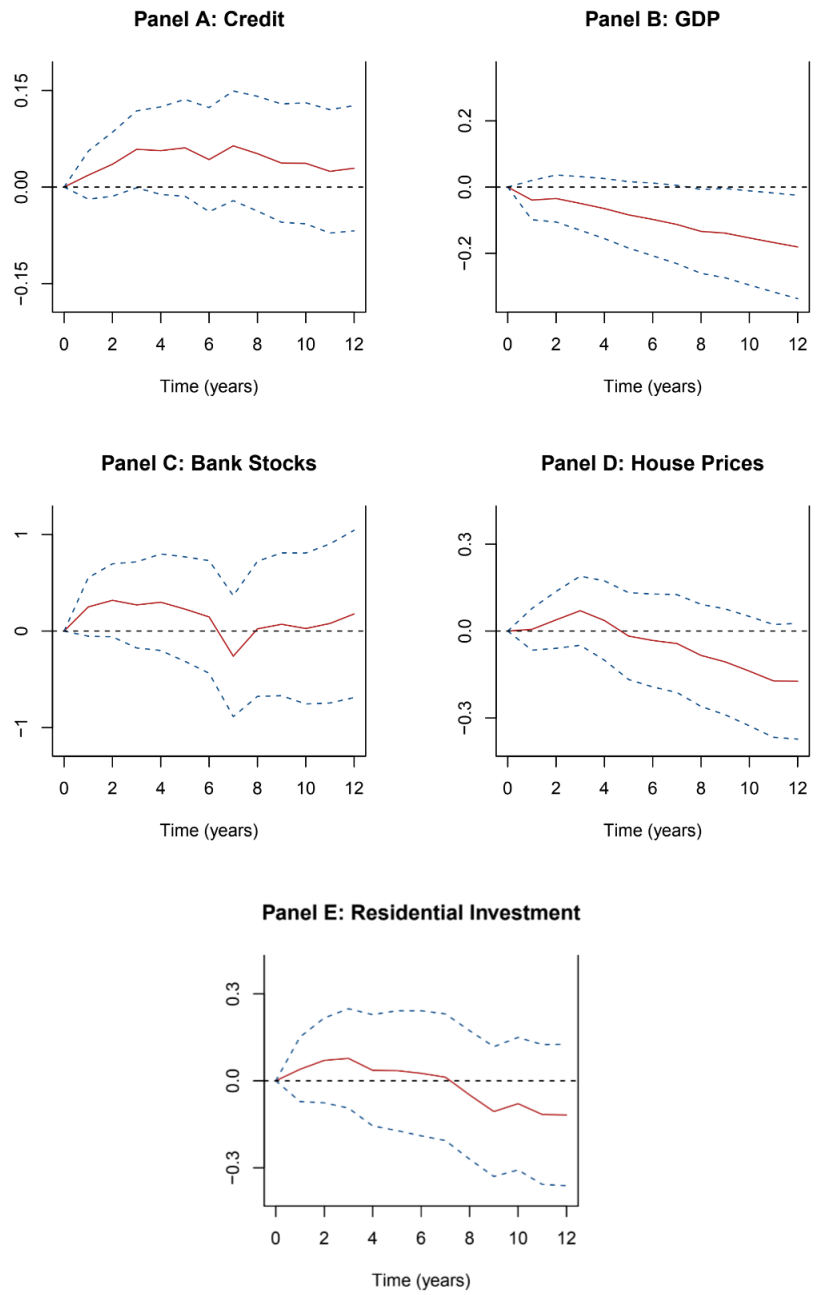


Figure 2.2

This figure plots impulse responses of various variables to an exogenous upward shock to a country's bank credit relative to GDP. Impulse responses are estimated via a linear projection with instrumental variables (LPIV) approach of Jordà, Schularick, and Taylor (2020). The figure plots estimated coefficients on the *Credit-to-GDP* variable generated from the second stage regression equation (3). Analogous results are displayed in tabular form in Table 2.4. Vertical bars in each panel represent the end of the "boom" period of our analysis (i.e. an assumed credit boom from time 0-3) and the beginning of the "bust" stage of the credit cycle. Left-hand side variables for the impulse responses are analogous to Figure 1. 95% confidence intervals for the impulse responses are drawn via the blue dashed lines in each panel. Values, in decimal form, are the growth of each variable in response to a 1% shock to bank credit to GDP growth.

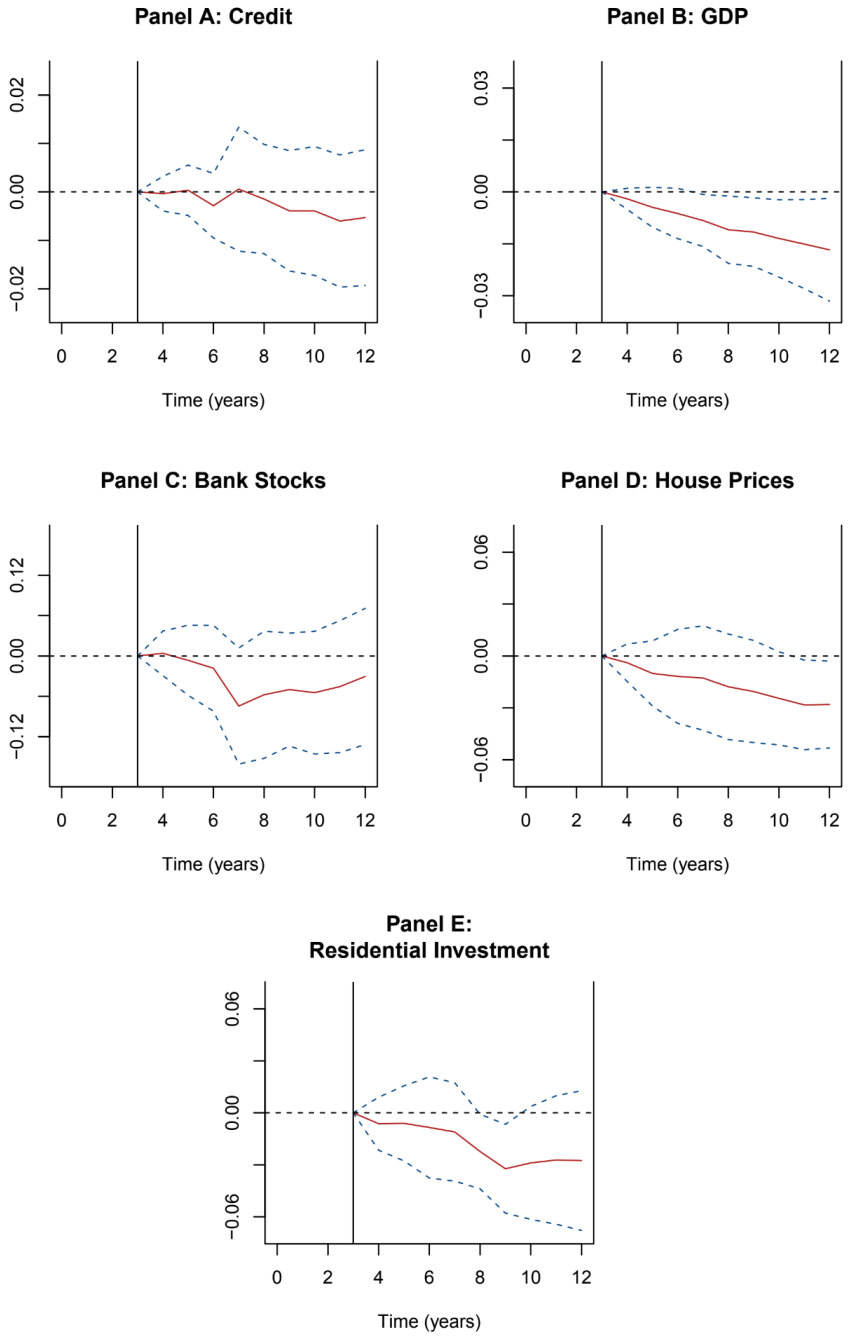


Figure 2.3

In this figure, we use data at the individual bank level (in the case of Sweden) and across classes of institutions (in Norway and Australia) to provide additional evidence suggesting that the credit booms we observe were the result of credit ceiling removals. In panels A and B we display the growth of credit in Norway by private banks and special government controlled state banks. Panel A shows the new net credit supplied by these institutions (i.e. total loans net of maturing debt) from 1977-1991. Credit ceilings were removed at the beginning of 1984 (announced in late 1983) for the private banks but not the state banks. Panel B shows how the evolution of credit for these types of banks looks from the perspective of total outstanding credit from these types of banks, with the year of ceiling removals set to an index value of 100. Panel C compares total outstanding credit of banks and non-bank financial institutions in Australia from 1975-1995. Outstanding credit in 1982 (the year of ceiling removals) assumes an index value of 100. NBFIs were not administratively controlled under credit ceilings, while banks were tightly regulated. Panel D shows the value of outstanding bank loans across large and small banks in Sweden from 1980-1989. Both large and small banks were subject to credit ceilings, but large banks were afforded more lenient treatment by the bank of Sweden, as they were permitted higher credit expansion, and tended to be favored in other ways by government policy.

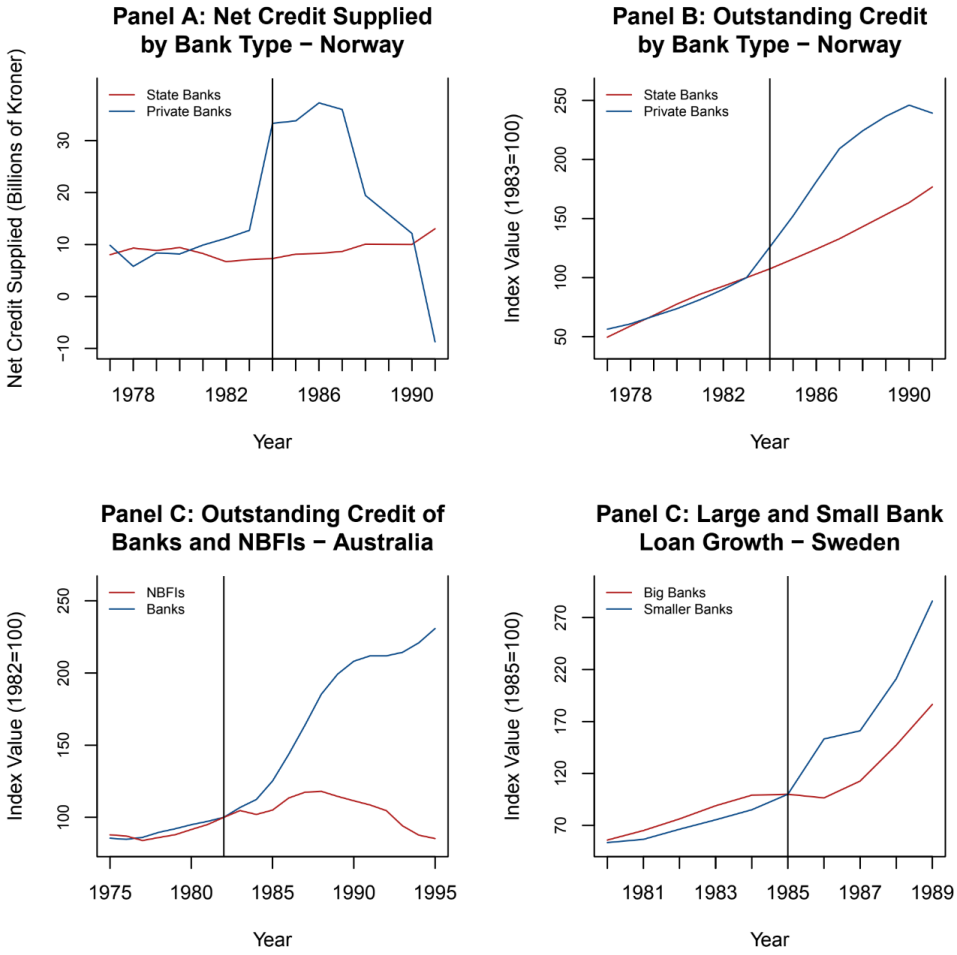


Table 2.1

This table displays each of the 13 credit ceiling countries and the year in which each country removed its credit ceilings. The credit ceiling countries appear in the order in which ceilings were removed. Dates for credit ceiling removals come from a number of academic and government sources. A full list of sources, and a more thorough discussion of the policies that we classify as credit ceilings for the purposes of this study can be found in the appendix. In some cases, there is a certain degree of ambiguity surrounding the "correct" date of credit ceiling removal. For instance, France announced it would gradually remove its credit ceilings in 1984, but did not complete the process of their removal until 1987; in other cases, countries reimposed restrictions after their initial removal prompted sharp rises in credit growth. See section 6 and the appendix for further details.

Country	Year of Liberalization
United Kingdom	1971
Chile	1975
Argentina	1977
South Africa	1980
Austria	1981
Australia	1982
Italy	1983
Japan	1983
Norway	1984
New Zealand	1984
Sweden	1985
France	1987
Portugal	1990

Table 2.2

This table shows summary statistics for the main variables of interest in the study. Observations in our data set are expressed at the country-year level. The first five rows of the table express summary statistics for the annual growth rates of each variables. Real GDP, bank stock prices, real house prices, and real residential investment are expressed in year-over-year percent changes while credit-to-GDP is expressed as a simple year-over-year difference. CPI inflation is a year-over-year percent change in the consumer price index. The final row of the table expresses summary statistics in levels for credit-to-GDP. The first column in the table displays the number of unique countries for which we have data for a given variable. The second column displays the number of non-missing observations available for each variable, prior to taking differences. The remaining columns depict the mean, median, inter-quartile range, and standard deviation for each variable. See section 4 and the appendix for a full list of data sources.

Variable	No. Countries	No. Obs.	Mean	Median	25th Pct.	75th Pct.	St. Dev
Real GDP	39	3394	0.0362	0.0357	0.0162	0.0577	0.0444
Credit-to-GDP	39	3394	0.0088	0.0081	-0.0075	0.0269	0.0485
Bank Stock Index	39	2899	0.0902	0.0562	-0.0810	0.1942	0.3821
House Prices	18	1049	0.0272	0.0236	-0.0202	0.0647	0.0824
Res. Investment	26	962	0.0279	0.0270	-0.0263	0.0884	0.1235
CPI Inflation	39	3394	0.0691	0.0319	0.0108	0.0760	0.1424
Credit/GDP Level	39	3394	0.5460	0.4603	0.2520	0.7667	0.3759

Table 2.3

This table compares growth rates of credit to GDP ratios in the years following credit ceiling removals to credit growth during all other time periods. For each of the 13 credit ceiling countries, we divide the sample into three-year windows and find the total growth of a country's credit to GDP ratio during each of those windows, comparing credit growth in the three years after credit ceilings are removed to the average growth rate during all other three-year windows, on a country by country basis. The first two columns show average credit growth in credit ceiling years and other years, respectively. The third column shows the difference between these values, and displays the t-statistic and p-value that arise from a test of the null hypothesis that credit growth tends to be equal across both liberalization years and non-liberalization years. The final column displays the proportion of countries for which liberalization-period credit growth exceeds the average credit growth rate during all other time periods, and gives a p-value under the null hypothesis that credit growth is equally likely to be above or below its unconditional mean in any 3-year window. Credit growth across countries is assumed independent for this test.

Comparing 3-year Credit Growth in Liberalization and Non-Liberalization Years				
	Lib. Years	Non-Lib. Years	Difference	Prop. Positive
Average 3-yr. Credit Growth	0.109	0.037	0.072***	.923***
t-Statistic			3.370	
p-Value			.005	.002

Table 2.4

This table shows results generated from estimating equations (2) and (3) via a two-stage linear projections with instrumental variables (LPIV) approach. The tables display coefficients on the credit to GDP variable estimated via these instrumental variables regressions. That is, we display $\beta^{s2,h}$ from regression equation (3) (denoted here as β^{Cred}_{t+3} for clarity) for values of h=4, 5,...,9. Each panel displays results of regressions with different set of outcome variables. Panel A shows regressions where the dependent variables are GDP growth; the Panel B outcome variables are real bank stock prices; Panel C outcome variables are real residential real estate prices; Panel C outcome variables are real residential investment. In panels E and F, credit growth is used to predict the incidence of financial crises, with crisis dates defined by Reinhart and Rogoff, and by Baron, Verner and Xiong (2020), in respectively. Crisis variables are defined as the number of financial crises that occur within the window denoted by the parameter h in each regression. For each coefficient in panels A-E, standard errors are given in parentheses. *, **, and *** denote statistical significance at teh 10%, 5%, and 1% levels, respectively.

Panel A: GDP						
	Dependent Var: $\Delta_h y_{i,t+3},\ h = 4, 5, ...9$					
	(h=4)	(h=5)	(h=6)	(h=7)	(h=8)	(h=9)
β^{Cred}_{t+3}	-0.827**	-1.093**	-1.160 **	-1.345**	-1.508**	-1.673**
	(0.345)	(0.447)	(0.455)	(0.514)	(0.591)	(0.683)
Observations	2066	2027	1988	1949	1910	1871

Panel B: Real Bank Stock Prices						
	Dependent Var: $\Delta_h y_{i,t+3},\ h = 4, 5, ...9$					
	(h=4)	(h=5)	(h=6)	(h=7)	(h=8)	(h=9)
β^{Cred}_{t+3}	-7.445*	-5.767	-4.996	-5.455	-4.554	-3.052
	(3.887)	(4.194)	(3.727)	(4.048)	(4.353)	(4.494)
Observations	1676	1631	1587	1544	1501	1459

Panel C: Real House Prices						
	Dependent Var: $\Delta_h y_{i,t+3},\ h = 4, 5, ...9$					
	(h=4)	(h=5)	(h=6)	(h=7)	(h=8)	(h=9)
β^{Cred}_{t+3}	-1.274	-1.780	-2.055	-2.449*	-2.828**	-2.807**
	(1.314)	(1.329)	(1.285)	(1.172)	(1.128)	(1.095)
Observations	868	850	832	814	796	778

Panel D: Real Residential Investment						
	Dependent Var: $\Delta_h y_{i,t+3},\ h = 4, 5, ...9$					
	(h=4)	(h=5)	(h=6)	(h=7)	(h=8)	(h=9)
β^{Cred}_{t+3}	-1.101	-2.232**	-3.226**	-2.890*	-2.720	-2.757
	(1.260)	(0.955)	(1.138)	(1.446)	(1.616)	(1.757)
Observations	701	675	649	623	597	571

Panel E: Reinhart and Rogoff Crises						
	Dependent Var: $\Delta_h y_{i,t+3},\ h = 4, 5, ...9$					
	(h=4)	(h=5)	(h=6)	(h=7)	(h=8)	(h=9)
β^{Cred}_{t+3}	4.971**	5.312**	7.550***	6.835***	6.095***	6.320**
	(1.851)	(2.009)	(2.162)	(2.022)	(1.886)	(2.162)
Observations	2066	2027	1988	1949	1910	1871

Panel F: Baron, Verner, and Xiong Crises						
	Dependent Var: $\Delta_h y_{i,t+3},\ h = 4, 5, ...9$					
	(h=4)	(h=5)	(h=6)	(h=7)	(h=8)	(h=9)
β^{Cred}_{t+3}	5.845**	7.163**	8.365***	7.620**	6.835**	7.039**
	(2.062)	(1.971)	(2.715)	(2.576)	(2.438)	(2.625)
Observations	2066	2027	1988	1949	1910	1871

Table 2.5

The following table illustrates three prominent phenomena associated with credit booms. In panel A we illustrate the "calm before the storm" phenomenon, investigating the relative timing of the peaks of the business cycle and credit cycle. In the residuals-based method, we define the beginning of a downturn as the first of two consecutive years where GDP or credit see negative forecast errors when regressing these quantities on their four most recent lagged values, and we define peaks as the last year prior to a downturn. In the first column, we show the average time between peaks, and the proportion of observations for which GDP sees a cyclical peak prior to credit. In the second column, we show analogous results using our crisis window methodology. Using this approach, we define a cyclical peak as the highest value GDP or credit attains in the years prior to the onset of a financial crisis, with crisis dates defined by Baron, Verner, and Xiong (2020). In panel B, we document the phenomenon of successive bubbles in the real estate market, whereby an asset price boom in one sector of the real estate market (either residential or commercial) is often followed by a smaller price boom in the other sector. The first column shows the time between the first year of each price boom, while the second column documents the average time between peaks. In panel C, we document the irreversibility of credit booms in the six countries that partially reinstated credit controls after liberalization. The first column shows the number of years that credit continued to grow after the enactment of the new restrictions. The second column asks whether a country experiences a financial crisis within the five years after new credit restrictions.

Panel A: The Calm Before the Storm – Business Cycle Peaks Before Credit Cycle		
	Residuals-based Method	Crisis-window Method
Average time between peaks (years)	1.30	1.82***
t-Statistic	1.109	3.390
Observations	13	11
Pos/(Pos+Neg)	.636	.888**
Bernouilli p-value	.274	.020

Panel B: Successive Bubbles – Large Asset Price Boom is Followed by a Smaller One		
	Time between start dates	Time between peaks
Average Time Difference (years)	2.75**	.75***
Observations	8	8
t-Statistic	2.433	3.00
Panel C: Similar Results for Real Estate Investment		
Average Time Difference (years)	.667*	1.0***
Observations	9	9
t-Statistic	2.309	3.0
Bernouilli p-value		

Panel D: Irreversibility of Credit Booms			
Country	Year of Regulation	Years of Cred. Growth	Crisis within 5-years?
Argentina	1982	0	yes
Italy	1986	7	yes
Japan	1989	1	yes
Norway	1986	4	yes
South Africa	1985	0	yes
UK	1973	1	yes

Table 2.6

This table depicts results of regressions which assess the ability of various policy reforms to predict credit growth. Credit growth over a 3-year and window was regressed on policy liberalization indicator variables. Panel A shows coefficients on the various policy liberalization variables in specifications where credit growth is regressed on a single policy variable and controls. Panel B shows policy variable coefficients in specifications with two or more policy variables. In the tables below, the abbreviation "Int. Rates" stands for any removal of interest rate restrictions (on deposits, loans, or other forms of bank credit). "Barriers" refers to the removal of barriers to entry of new banks. "Branching" refers to removal of restrictions on the ability of banks to open new branches. "Specialization" refers to removal of restrictions on the types of services banks are allowed to offer or the activities they may engage in (e.g. bank restrictions on securities trading, or on raising long-term debt in credit markets). "Cap. Flows" refers to the romval of restrictions on international capital flows. "Fin. Reform" refers to financial reform in the bond market or money markets. Standard errors are listed in parentheses beneath each coefficient and *, **, and *** labels denote statistical significance at the 10%, 5%, and 1% levels respectively.

Panel A: Association Between Deregulation and 3-yr. Credit Growth							
	Single Policy-Variable Specifications						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ceilings	7.638***						
	(2.329)						
Int. Rates		1.890					
		(1.641)					
Barriers			3.336				
			(2.498)				
Branching				-0.025			
				(3.525)			
Specialization					2.862		
					(2.496)		
Cap. Flows						1.600	
						(1.532)	
Fin. Reform							2.667
							(1.955)
Controls?	✓	✓	✓	✓	✓	✓	✓
R^2	.045	.024	.025	.021	.024	.023	.025
Observations	423	423	423	423	423	423	423

Panel B: Association Between Deregulation and 3-yr. Credit Growth							
	Two Policy-Variable Specifications						All Vars
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ceilings	7.392***	7.396***	7.639***	7.695***	7.553***	7.695***	7.524***
	(2.382)	(2.341)	(2.332)	(2.328)	(2.331)	(2.327)	(2.294)
Int. Rates	0.832						-0.092
	(1.660)						(1.724)
Barriers		2.546					2.545
		(2.485)					(2.507)
Branching			0.131				0.202
			(3.485)				(3.500)
Specialization				3.027			2.415
				(2.467)			(2.533)
Cap. Flows					1.409		0.834
					(1.516)		(1.599)
Fin. Reform						2.778	3.990*
						(1.933)	(2.162)
Controls?	✓	✓	✓	✓	✓	✓	✓
R^2	.046	.048	.045	.049	.047	.050	.060
Observations	423	423	423	423	423	423	423

Table 2.7

This table shows the results of logistic regressions where we assess the ability of various policy reforms to forecast the *beginnings* of credit booms. We define credit booms as beginning in the first of two consecutive years where credit to GDP growth exceeds the median growth rate of the previous five-year period, or when credit to GDP growth exceeds 1% in a single year. A credit boom ends after two consecutive years of negative credit to GDP growth, or after a single year with a credit contraction greater than 1%. A credit boom may also end after *three* consecutive years of static credit growth (i.e. growth below the median growth rate of the prior five-year period). We adopt the restriction that a new credit boom cannot begin while a previously existing boom is already underway. Panel A shows policy variable coefficients in specifications where a single policy variable is used to predict the onset of a credit boom. Panel B shows the results of regressions with multiple policy variables. The tables also report the Akaike Information Criterion (labeled AIC at the bottom of each table) for each of the regressions. Standard errors are listed in parentheses below each coefficient and statistical significance at the 10%, 5%, and 1% levels are denoted by *, **, and ***, respectively.

Panel A: Predicting 1-year Ahead Credit Booms with Policy Variables							
	Single Policy Variable Logit Specifications						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ceilings	2.041***						
	(0.747)						
Int. Rates		-0.011					
		(0.661)					
Barriers			-15.128				
			(905.661)				
Branching				-14.090			
				(727.532)			
Specialization					0.342		
					(0.830)		
Cap. Flows						0.425	
						(0.545)	
Fin. Reform							0.082
							(0.807)
Controls?	✓	✓	✓	✓	✓	✓	✓
AIC	323.380	330.941	327.998	329.310	330.780	330.368	330.931
Observations	328	328	328	328	328	328	328

Panel B: Predicting 1-year Ahead Credit Booms with Policy Variables							
Multiple Policy Variable Logit Specifications (Coefficient Averages)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ceilings	2.118***	2.450***	2.023***	2.053***	2.082***	2.048***	2.664***
	(0.767)	(0.853)	(0.747)	(0.747)	(0.748)	(0.747)	(0.897)
Int. Rates	-0.352						-0.616
	(0.721)						(0.780)
Barriers		-15.869					-15.995
		(820.30)					(807.68)
Branching			-14.002				-15.176
			(727.45)				(1186.99)
Specialization				0.416			0.165
				(0.831)			(0.870)
Cap. Flows					0.515		0.696
					(0.547)		(0.592)
Fin. Reform						0.169	0.203
						(0.808)	(0.820)
Controls?	✓	✓	✓	✓	✓	✓	✓
AIC	325.126	320.664	323.873	325.145	324.554	325.337	327.233
Observations	328	328	328	328	328	328	328

Appendix I: Documentation and Sources for Credit Ceiling Policies and Deregulation Dates

In this appendix we provide further information on the financial systems in the credit ceiling countries in our sample, and we provide documentation on the sources through which we compiled this information. For each credit ceiling country in our sample, we first list the sources through which we identified the years that credit ceilings were active (and in particular, the year credit ceilings were removed). Next, where available, we provide a brief description of how credit ceilings operated in the country in question, including an account of which institutions and types of credit were targeted by the ceilings and a description of the policy goals underlying the implementation and removals of credit ceilings. In some cases, we also provide a brief account of the other credit policies that were implemented and removed alongside credit ceilings.

Argentina

Removal of credit ceilings in Argentina occurred either in 1977 or 1978, depending on the source. Mahar and Williamson (1998) give 1977 as the date for removal of credit ceilings, though the authors also note that quantitative controls were partially reimposed in 1982. Rojas-Suarez and Weisbrod (1995) give the deregulation date for credit ceilings, as well as interest rate ceilings and other directed credit, as 1978.

Australia

Evidence for the removal date of credit ceilings is found in Hall (1987), page 30: "With effect from the end of June [1982] the request, made on 15 September 1981 to trading banks to keep growth in their advances to not more than 12 per cent per annum was withdrawn. This represented the abolition of quantitative controls on trading-bank lending (in force since 1975)."

Corroborations of 1982 as the key date for removal of quantitative lending controls can be found in Ballantyne, et al. (2014) and in Mahar and Williamson (1998).

In Australia, credit regulations generally applied, with different levels of stringency, to all of the federally chartered banks. The major deposit taking entities were the trading banks and savings banks. Trading banks were analogous to merchant banks, and primarily made loans and transacted with businesses. Savings banks transacted with individual savers and mostly were restricted to providing mortgage loans. There were also various types of non-bank financial institutions, such as permanent building societies, which fell outside the scope of banking regulation. Quantitative lending restrictions applied to both the trading and savings banks, however the relative strictness with which lending guidelines were applied to these types of institutions is not always clear during our sample period.

The Bank of Australia was given statutory authority, beginning with the 1959 Banking Act, to set lending guidelines for trading and savings banks, both in terms of prices (i.e. rates) and volumes (see Hall, 1987, for further details).

Since credit ceilings and other controls applied to the trading and savings banks, but not to finance companies, building societies, and other non-bank financial institutions, removal of these ceilings would be likely to jumpstart bank credit but have little effect on financial institutions. In Figure 2.3, we show that this is indeed what happened in Australia. Anecdotal evidence from Hall (1987) further suggests that these changes in credit composition were likely the result of changes to regulation. For example, the author argues: "given that trading banks, especially those subject to Federal regulation, suffered the most under the old regulatory regime...their share of total assets of all financial intermediaries fell from 32 per cent in 1953 to 23 per cent in 1963, a level maintained ever since—they are likely to be the main beneficiaries from financial deregulation. Removal on quantitative controls on lending and of maturity controls and restrictions on the payment on interest on checking accounts and call money will take business away from money-market corporations, authorized money-market dealers and other NBFI deposit-takers...Savings banks are another group clearly likely to benefit from deregulation...Largely as a result of regulatory 'straightjackets', savings banks' share of total financing declined markedly between 1953 and 1982 from 20 per cent of total assets of all financial institutions to 13 per cent. This performance

contrasts sharply with that of their major competitors, the permanent building societies and credit unions... The largest group of NBFIs subject to the Financial Corporations Act are the finance companies (and general financiers). Their share of total assets peaked at 14 per cent in 1982... Although [their] business improved in the latter half of 1983 and early 1984 their future remains uncertain, as the 'unshackling' of the banks is likely to induce a continuing relative decline of those which are non-bank-owned and the possible absorption of those affiliated with trading banks."

Austria

In Austria credit ceilings (referred to as the system of "Limes") were removed in 1981, as stated by Liebscher, et al. (2007): "In 1973 this passive credit control mechanism was... complemented by an active credit control mechanism. This new instrument foresaw an upper limit for the monthly growth rate of loans, the so-called Limes. Complying with the Limes was a pre-condition for access to the central bank's refinancing instruments... With the exception of a temporary weakening in 1975 and 1976, these rules were kept in place until 1981, when active and passive credit controls were lifted. Interestingly, they were lifted separately and with different arguments: most of the active measures were abandoned in February 1981, as the central bank blamed them for competitive distortions between credit institutions... The passive measures were abandoned in June 1981, when the ministry of finance, the central bank and the credit institutions were not able to agree how these measures might go along with the new Banking Act of 1919. In May 1982 the very few remaining active measures which restricted credit growth for households (Limes on consumer credit) were abolished."

Chile

Our main sources of information on key deregulation dates for Chile are Mahar and Williamson (1998) and Velasco (1988). Both of our key sources list 1975 as the initial date for removal of quantitative credit controls.

There was a broad range of financial reforms from 1974-1976, which took place for ide-

ological reasons. After the wave of bank nationalizations in 1970-1973, initiated by the administration of Salvador Allende, the military government that took control after the 1973 coup d'état, was significantly more market-oriented than the prior administration, and sought to make the banking system more competitive, gradually privatizing the nationalized banks, and removing administrative controls that previously existed. Velasco (1988) specifically notes that credit ceilings (and other bank portfolio restrictions) were removed in the hope that liberalization would more efficiently channel capital to its most productive uses, noting that: "another much-criticized feature of the previous system [pre-reform]---an array of quantitative controls on credit—was also eliminated. Selective credit controls had been a widely used tool of economic policy until then. After 1975 it was expected that the market mechanism alone would allocate credit, with the presumption that this would improve the allocation of resources."

Credit ceilings primarily targeted the commercial banks, rather than the savings and loans, which largely focused on residential mortgage lending. Alongside credit ceiling liberalization, the commercial banks were also given greater latitude to make residential mortgage loans (see Velasco, 1988, page 8). The other major policy reforms adopted alongside the removal of credit ceilings and bank privatization included removal of interest rate controls on deposits (also in 1975) and the removal of bank specialization requirements that separated commercial banks, investment banks, and mortgage lenders (savings and loans).

France

Our main sources for verifying the dates that credit ceilings were implemented and removed in France include Hodgman (1985), Monnet (2012), Goodman (1992), and Melitz (1990). Hodgman (1985) and Monnet (2012) describe France as gaining statutory authority to implement credit ceilings in the early 1970's (previously relying on informal guidance and moral suasion), with the latter of these sources suggesting ceilings were implemented as a relatively permanent fixture of French monetary policy by 1973 (see Monnet, 2012, page 210). France is an example of a country that gradually weakened its credit ceiling policies

over time.

Goodman (1992) asserts that France began its process of removing credit ceilings by first announcing, in November of 1984, its plan to gradually remove its ceilings, beginning with a sharp reduction of the penalty that a bank would have to pay for exceeding its individual ceilings. This policy took effect at the beginning of 1985, and credit ceilings were removed in their entirety in 1987, according to the author. This timing is echoed by Melitz (1990) who, when discussing financial reforms in the 1980's, mentions the new Banking Act of 1984 and notes that "this Act, replacing 1941 and 1945 legislation, removes old divisions between investment and commercial banks and introduces a new, uniform set of prudential rules for all financial institutions. The deregulation itself may be dated with the November 1984 announcement of the cessation of the 'encadrement' [i.e. credit ceilings] on New Year's day [of 1985]." A reading of primary source government documents from France's National Credit Council confirms that the credit ceilings ceased to exist, in any form, by 1987.

French credit ceilings operated through a system of supplementary reserve requirements. If a bank exceeded its permitted rate of lending growth, it would be required to hold additional non-interest bearing reserves. The rate at which it would have to make additional reserve commitments increased exponentially as the threshold was exceeded. According to Melitz (1982) "As currently operated, the 'encadrement' combines ceilings (norms) on credit growth over the next (calendar) semester with penalties for transgressions consisting of legal reserve requirements. The legal reserve requirements are a percentage of the total credit distribution. Further, this percentage rises exponentially with the level of the transgressions." As an example of how this system worked, in 1982, the formula for supplemental reserve penalties took the form $T=(.3+.15X)X$ where X is the the number of percentage points by which a bank exceeds the credit ceiling, while T is the ratio of the amount of new non-interest-bearing reserves that need to be provisioned to the quantity of credit supplied. Thus, if the ceiling is exceeded by 2%, the bank will have to place $(.3+.15*2)*2=120\%$ of the value by which lending has exceeded the ceiling into non-interest bearing reserves.

The French system of credit ceilings mainly targeted the commercial banks, though the

credit ceilings maintained rigid control over aggregate credit, because bond markets were tightly controlled, and other types of lenders (e.g. the savings banks) were also directly controlled by the government. Credit ceilings did affect some types of loans more than others. After the inception of its credit ceilings, France also carved out various exceptions for types of credit that were given priority status under the ceilings, and were thus either entitled to grow more rapidly than non-priority credit, or were exempted from ceilings entirely. According to Aftalion (1981) five sectors that received priority status under the credit ceilings were agriculture, housing, exports, industrial equipment, and local government loans. Aftalion (1981) notes that "[s]everal specific sectors of the economy (housing, exports, industrial development, agriculture, local entities) receive support from the government, in particular through subsidized loans (credits aides). Restricting the volume of these loans would defeat their purpose. Therefore, their expansion is less severely limited than that for 'ordinary' loans. In the latter case, monthly 'ceilings' are fixed for each bank."

The purpose of the French credit ceilings was very clearly stated in the case of France, and had to do with the government's desire to control the money supply and regulate inflation. For example, Aftalion (1981) notes: "during the Raymond Barre years, at the top executive level the main official objectives of monetary policy were the reduction of the inflation rate and the stabilization of the foreign exchange rate. Another objective was to keep short-term interest rates above the inflation rate. The instruments used to achieve these objectives were respectively: control of the money supply and control of interest rates. The technique used for monetary control was, and still is, control of the volume of 'ordinary' loans extended by banks."

Credit ceilings were removed alongside other efforts to modernize monetary policy. Melitz (1990) mentions: "As of January 1985, the French monetary authorities began a sharp transition toward exclusive reliance on interest rates and legal reserve requirements for monetary management in which the reserve requirements played a major part. A series of impressive changes took place in the next fifteen months or so. The authorities set up a new market for debt instruments of all maturities of up to seven years in which any one

could participate ('the money market')... Moreover, banks were allowed to issue certificates of deposits; firms were permitted to float commercial paper (heretofore unknown); and the purchase of Treasury bills was opened to everyone. The so-called 'petit marche', where modest bond issues were possible without waiting one's turn on the administered queue, became available for issues five times the previous size."

There was also a sense that the various sectors of credit which were exempt from the ceilings made these ceilings insufficiently precise in controlling the money supply. Melitz (1990) notes, for example, that "the administrative complexities of reconciling the money growth targets with the subsidized credits grew incessantly."

Italy

Szego and Szego, writing a chapter on the Italian banking system in Kaufman (1992), write that: "From 1973 to 1988, the Bank of Italy adapted direct measures to restrict credit and at the same time to achieve its target of placing Treasury debt instruments and facilitating the financing of special credit institutions. In 1973 a *portfolio constraint* was introduced. This measure compelled banks to invest part of their resources in bonds issued by special credit institutions... Besides this portfolio constraint, the Bank of Italy also used another direct tool to restrict credit: the *ceiling on lending growth*, which was in force almost continuously from 1973 to 1982 and was renewed for short periods in 1986 and 1987. This ceiling consisted in the establishment of a maximum rate of expansion of bank lending." The authors note that ceilings are removed in the following year, which also is echoed by the timeline in Mahar and Williamson (1998).

The ceilings were directed toward all banks, but not all loans were subject to ceilings. Verheirstraeten (1981) notes that: "loans not subject to ceiling can be expanded freely. These loans are extended to small customers: the freedom of action left to banks in granting them was officially justified with the necessity of avoiding a too large incidence of credit restrictions on smaller enterprises."

Credit ceilings in Italy were a part of the monetary policy regime in Italy, and were

implemented alongside portfolio requirements, which helped allocate credit to priority sectors. Ceilings were put in place in order to regulate inflation. For example, Verheirstraeten (1981) notes that "Monetary authorities were compelled in June 1973 to introduce a scheme of direct controls on bank credit which... put limits on the rate of growth of bank loans and forced banks to acquire medium- and long-term bonds. The controls, which in their original version extended only to a short period of time, were justified on the grounds of short-run considerations. The objectives which the controls intended to achieve, according to statements by the monetary authorities, were those of external equilibrium and of a reduction in the rate of inflation."

There is also some evidence suggesting that credit ceilings were put in place, in part, because of the Bank of Italy's desire to induce banks to purchase government debt, which they did when credit ceilings prevented them from making additional loans, despite possessing ample reserves. Verheirstraeten (1981) suggests that "[i]n an indirect way, banks have also often been forced by the controls system to buy government bonds, which are not included in the portfolio requirement. Banks have been pushed to this kind of behavior by the existence of the ceiling which, preventing them from expanding their loans, has often induced them to buy government securities, besides the bonds of special credit institutions which they are forced to buy." The bank of Italy was required, by law, during much of this period to purchase all government debt that was not sold on open markets at the yield at which these bonds were offered. Kaufman (1992) notes that the "Bank of Italy was bound to purchase all debt papers issued by the Treasury and not sold to the public. This formal relationship lasted till 1982 and ended with the so-called divorce between the Bank of Italy and the Treasury." The bank of Italy thus was restrained from using open market operations (i.e. the purchase of government bonds) as an independent tool for conducting monetary policy and thus needed some other policy tool.

The Bank of Italy may also have found credit ceilings, in tandem with portfolio requirements, as desirable tools for inducing banks to purchase government bonds to reduce the residual amount that they themselves would have had to purchase. After 1982, when the

Bank of Italy was released from this government bond purchasing requirement, the need for credit ceilings as an indirect tool to induce banks to purchase government debt may have lessened.

Japan

In Japan, credit ceilings were implemented by the Bank of Japan, via a semi-formal process referred to as "Window Guidance." At various times, these policies targeted only the largest banks (the so-called "city banks" that operated in Japan's largest urban areas) while at other times these policies were broadened to cover larger sets of lenders.

As written in Suzuki (1987): "During its day-to-day contact through deposit and lending transactions with client financial institutions, the Bank of Japan receives various types of information concerning the plans and actual developments of the funds and lending positions of these institutions. Using this information, the Bank of Japan provides guidance to the financial institutions to keep the increase in their lending to clients within limits that the Bank of Japan feels to be appropriate. Such guidance is particularly important in times of tight money and for lending by major institutions such as city banks. This type of guidance is in general called 'window guidance', and the type of guidance is changed from time to time according to financial conditions. Reflecting these conditions, the guidance sometimes takes the form of regulation of increases in loans and sometimes the form of guidance of overall positions. For example, during each of the tight money periods between 1957 and around 1963, the Bank of Japan gave directives to keep lending by financial institutions within specified limits on a monthly basis, in order to prevent borrowings from the Bank of Japan from rising too much. This guidance was carried out primarily vis-à-vis city banks that had high levels of Bank of Japan borrowings. In the tight money period of 1964 there were directives to control the increases in lending on a quarterly basis, and the objects of such guidance were expanded from city banks and long-term credit banks to include trust banks (the banking accounts thereof) and regional banks. In the tight money period of 1967, the institutions subject to guidance were again expanded to include the larger sogo banks.

In the tight money period following January 1973, the extent of guidance was expanded further to include the larger shinkin banks and the larger foreign banks in Japan, that is, almost all the Bank of Japan's client financial institutions. In addition, there was guidance concerning not only the restraint of overall lending but on lending to trading companies, and also guidance concerning restraint of securities investment. After 1975, the period of tight money was gradually relaxed, and from July 1977 a new formula was introduced under which the voluntary lending plans of the various financial institutions were essentially accepted within the window guidance framework; the framework itself, however, was maintained in order to continue appropriate management of the money supply. In 1979, a strict guidance was implemented that was similar to that in previous periods of tightening; but in the second half of 1980 this guidance was gradually eased along with the change of policies. Since early 1982, the lending programmes of the individual financial institutions have been accepted completely."

Hoshi, Scharfstein, and Singleton (1993) list the date of deregulation as 1983 and write "After 1982 and until 1989, window guidance played an insignificant role in the conduct of monetary policy, as the lending programs of financial institutions were accepted completely."

As suggested above, it is well documented that credit ceilings in Japan were implemented as a monetary policy tool, with the goal of controlling money growth and inflation. The notion that credit ceilings were primarily instituted to maintain control of the price level is further supported, for example, by Hoshi, Scharfstein, and Singleton (1993) who note that "during 1990, the BOJ once again relied on window guidance in an effort to control inflation in Japan. For instance, in the last quarter of 1990, the BOJ reduced the net lending of the twelve city banks by more than 30% from a year before."

The use of credit controls, rather than interest rate policy, as a key instrument of monetary control, were implemented for somewhat idiosyncratic reasons. Under a 1947 law, the Temporary Interest Rate Adjustment Law (TIRAL), the Ministry of Finance (MOF) set interest rates across a number of bank products in an effort to regulate competition between banks. According to Suzuki (1987) "the purpose of this law was to prevent interest

rate competition that was destructive to the profitability of financial institutions... Under the TIRAL, the Minister of Finance would determine whether interest rate regulation was necessary in the light of general economic conditions." The Bank of Japan, while given some de facto authority to recommend changes to interest rate controls given economic conditions, could not reliably adjust rates at will to control inflation. Thus, window guidance was adopted as a tool for preventing the economy from overheating. As stated by Rhodes and Yoshino (2007), "since indirect control of lending and investment was not possible under the interest rate control regime enforced by the MOF, the BOJ relied on quantitative controls measures."

Credit ceilings were discontinued as part of a general push toward financial liberalization in the 1980's. The regime of administratively controlled interest rates was gradually dismantled in the late 1970's and early 1980's, giving the Bank of Japan greater flexibility to set interest rates. Liberalization of capital markets allowed the BOJ greater ability to perform open market operations to flexibly control the money supply. Thus, strict credit controls were likely no longer seen as necessary. There was also a general concern that credit ceilings hampered competition among financial institutions. Suzuki (1987) notes that "if strong window advice is continued for long periods, there emerge disequilibria among financial institutions between those that are subject to controls and those that are not. In addition, the lending shares within one type of financial institution tend to become fixed."

In the 1980's, it was thought that the liberalization of credit markets, including the expansion of corporate debt and increased availability of short-term borrowing via capital markets rendered credit ceilings less effective, and thus obsolete. For example, Shigehara (1987) argued: "it is true that in the past when the Japanese financial system was highly regulated and international capital flows were tightly controlled, the Bank of Japan relied on direct quantitative control on commercial bank lending to non-bank sectors as an important supplementary instrument to adjust the volume of total credit to these sectors. Then, the development of the money supply was largely in parallel with domestic credit expansion, which, in turn, followed more or less the same course as total bank lending. However,

direct quantitative control on commercial bank lending is no longer in use... In any event, given a significant progress in the de-regulation of financial transactions both domestic and international, Japanese non-bank borrowers' access to alternative sources of finance and innovations in financial engineering could negate the effect of compulsory control on the volume of domestic bank lending, even if it were reintroduced."

New Zealand

Our main source of information on the implementation and removal of credit ceilings is Walsh (1988). The author writes that credit ceilings were removed alongside various other financial market reforms in 1984. The author notes that "New Zealand, prior to 1984, had a highly segmented financial system which was reinforced by regulation which sought to channel funds to priority sectors of the economy, while controlling aggregate credit." The credit ceilings that were in place prior to removal in 1984 usually required banks to keep the growth in their loans outstanding to no more than 1% per month. In New Zealand credit controls were abolished concurrently with the opening of a true central bank discount window, the removal of interest rate controls, and the relaxation of restrictions on private overseas borrowing.

Norway

Various sources mention 1984 as the key date for the removal of the supplementary reserve requirements (i.e. credit ceilings), including Drees and Pazarbasioglu (1984), and Moe, Solheim, and Vale (2004).

In Norway, credit ceilings were operated via a few mechanisms. For the government-controlled "state banks," direct controls were set in so-called "credit budgets," declared annually by the Ministry of Finance. Moe, Solheim, and Vale (2004) note that "during the 1960s and 1970s, the government developed a 'credit budget' framework for macroeconomic planning, involving special government lending institutions ('state banks') responsible for different sectors like the housing sector, manufacturing, agriculture and fisheries. The idea was both to control aggegate demand (jointly with fiscal policy), and sectoral investment spending by means of a housebuilding permit system, regulation of the bond market and

credit flows from private and public financial institutions, and regulation of foreign exchange and cross-border capital movements."

The main system of credit ceilings for the privately operated banks was operated through the Norges Bank's so-called "supplementary reserve requirements." In this system, the central bank would set a threshold rate of credit expansion for banks. Loans above this threshold would have to be accompanied by an additional set-aside of reserves in non-interest bearing accounts at the Norges Bank. Since nominal rates and inflation tended to be high during this period, it was generally quite costly to lend above these thresholds, and sources tend to suggest that this system effectively created a firm ceiling for credit growth among the privately owned banks. Drees and Pazarbasioglu (1995) describe this system of supplementary reserve requirements: "to restrict bank lending, a supplementary reserve requirement was imposed on Norwegian banks in the early 1980s that mandated that a prohibitively large part of the lending increase that exceeded the credit ceiling had to be deposited in non-interest-bearing accounts at Norges Bank. In addition, banks were not permitted to pass the extra costs on to borrowers. This requirement, which was in effect from 1981 to 1983, was sufficiently high to restrict bank lending." In addition to supplementary reserve requirements, direct controls were also occasionally instituted for non-bank lenders (e.g. insurance companies and finance companies), mainly as a tool to supplement its more permanent system of credit controls on the banking sector in the wake of lending increasingly gravitating toward these non-bank sources in periods of high lending demand.

Credit ceilings began informally, beginning in the early 1950's, as Kleivset, et al. (2011) recount that the Norges Bank developed "qualitative credit guidelines concerning lending volumes, credit allocations and liquidity positions with private banks on a voluntary basis." The authors note that these guidelines on lending volume were formalized 1965 with the inception of supplementary reserve requirements for banks. Credit ceilings were implemented alongside a number of other policies for regulating credit and managing the macroeconomy, such as the aforementioned state-bank system, and the "credit budget" approach that governed this system. It is well documented that Norway's system of credit controls was

motivated by the desire for monetary control, and did not have macroprudential purposes. For example, Moe, Solheim, and Vale (2004) note "the bank regulation [i.e. credit control system] that was in place between the end of World War II and 1984-1985 did not have prudential purposes. The main purpose of the quantitative regulations of banks was to control aggregate credit supply as a substitute for a market-based monetary policy."

Credit ceilings in Norway were dismantled alongside a gradual wave of deregulations in the late 1970's and early 1980's. Interest rate restrictions were eased (though not abolished) in 1980. Interest rate restrictions were lifted altogether and portfolio requirements for banks were dismantled in 1985. Moe, Solheim, and Vale (2004) believe that along with the ideological desire to move toward more market-oriented monetary policy, and make the banking sector more competitive, credit ceilings and other regulations were also made ineffective by the rise of non-banks and financial innovation. The authors note, for example, that "[t]he interest rate regulation policy also generated powerful incentives to channel credit outside the regulated credit market by numerous shadow market operations. Over time, new innovative ways of circumventing the regulations triggered new regulatory measures... In the beginning of the 1980's, the growth of the eurokrone market, financial innovations and increasing flexibility of the shadow credit market made it much more difficult for the government to constrain the underlying market forces by credit regulations... This problem appears to be the main reason why the government decided to move away from credit regulations in the fall of 1983."

Portugal

Our main source of information on credit ceilings in Portugal is Barradas, Lagoa, Leao, and Mamede (2013). According to the authors, the implementation of credit ceilings began in the early 1980's, though credit was effectively regulated prior to this, more informally, given that most banks were nationalized in the late 1970's and subject to various forms of government control. Deregulation began in 1983 with the opening of the business of banking to private firms (previously only nationalized banks operated in the country, with

the exception of smaller mutual and cooperative banking organizations). The authors note that "the third state of [credit market liberalization] is related to the elimination of interest rate controls (controls on lending interest rates were removed in 1988, while controls on deposits interest rates were eliminated only in 1992), the abolishment of credit ceilings (in 1990) and other State controls on the number of branches and new products and services."

South Africa

Mahar and Williamson (1998) note that South Africa implements credit ceilings in 1965, and removes them for good in 1980.

Sweden

Drees and Pazarbasioglu (1995) state that "the system of liquidity ratios for banks was abandoned in 1983 and the ceilings for commercial bank lending were removed in 1985. At the same time, restrictions on lending rates were lifted, and by 1989 all remaining foreign exchange restrictions had been removed." Several other sources corroborate that 1985 was the date of credit ceiling removal, including Englund (2015).

Credit ceilings in Sweden were a part of a broader agenda of state controls on credit, including directed credit to the housing sector, agriculture and other priority sectors. Banks were assigned individual quotas for new housing construction loans, and had portfolio requirements mandating minimum investments in medium-term housing bonds. Portfolio, liquidity, and cash ratio requirements were set heterogeneously across various types of institutions, including commercial banks, savings banks, and agricultural credit associations (see Hodgman, 1976). Credit ceilings were initially informal, but later gained the force of law, as stated by Englund (2015): "bank actions were continuously scrutinized by the Riksbank and views on proper bank behavior were communicated in weekly meetings between the governor and representatives of the major banks. As one result of these meetings, the banks would commit to keep their lending within certain limits. It was only in 1974, however, that a law was passed giving the Riksbank the right to impose legally binding regulations."

Credit ceilings were set at different rates across institution types of institutions, with

larger credit expansion often permitted for the large banks. For example, Hodgman (1986) notes that "the Riksbank imposed a ceiling on the rate of expansion of bank loans for purposes other than house building...In 1974 the ceiling rate on loan expansion over a 12 month period was 18 per cent for commercial banks, 10 per cent for [smaller] savings banks and 13 per cent for cooperative banks. Other banking institutions were requested not to expand their lending at the expense of commercial banks." Credit was also tightly controlled in Sweden because bond markets were virtually non-existent for entities other than those with strong ties to the state (e.g. government bonds and bonds for housing construction). As in other countries, there is evidence that the process of deregulation, the growth of financial markets, and regulatory arbitrage, had made credit ceilings less effective, leading in part to the decision by central bankers to get rid of credit ceilings. For example, Drees and Pazarbasioglu (1995) note that "banks attempted to bypass the interest rate regulations by establishing their own finance companies—which formed an important part of the grey credit market in Sweden."

United Kingdom

Tew (1978) notes that credit ceilings were dismantled in 1971, alongside a number of other reforms, including the dismantling of interest rate controls (the so-called interest rate "cartel" set by the major London clearing banks). Similarly, Perez and Westrup (2008) mention that "in Britain, direct credit controls were dismantled in 1971 in favor of 'competition and credit controls' (the lifting of quantitative controls and the replacement of the regulated Bank Rate by minimum lending rates and reserve ratios deposited with the Bank of England."

As in other countries, credit ceilings in the UK were largely informal and applied via moral suasion. According to Tew (1978), "[f]rom 1950 onward, the banks had been accustomed to receive from time to time official advice on the composition, and subsequently also on the total amount, on their sterling lending to the private sector. In the 1950's the requests were addressed only to the London and Scottish clearing banks, but as from 1961 they were extended to other banks and to the larger finance houses."

Credit ceilings were implemented as a form of monetary control, to fight inflation. For example, a Bank of England Official, Charles Goodhart, wrote in a letter to the Governors at the Bank of England: "The attached paper considers the reasons for examining at this time various methods of direct controls over banks. The main reason is seen to be that of restraining monetary expansion without an unpalatable increase in interest rates. This can be achieved by preventing the banks, by controls, from carrying out their role as intermediaries, though this needs to be reinforced by an agreement with the clearing banks, at least, that they will not profiteer from such direct rationing by charging what the traffic will bear."

Credit ceilings were preferred, at times, to interest rate policy, because the government sought to control inflation without raising borrowing costs for the government. For example, Hodgman (1976) claims that "one objective of selective credit controls has been to shelter the gilt-edged [i.e. sovereign debt] market from the full force of tight credit policies and attendant higher costs of financing the large British national debt... For some years prior to the reform of 1971, they attempted to restrain upward pressure on gilt-edged yields by means of their control over interest rates in the narrowly-defined money market, by restricting bank advances to the private sector [i.e. credit ceilings], and ultimately by open market operations on the part of the Bank of England to support gilt-edged prices."

Hodgman (1976) argues that credit ceilings were removed for ideological reasons. It was initially believed, by earlier governments, which were less ideologically averse to heavy-handed credit regulations that "credit control techniques permitted more efficient control over interest rates, total credit, and aggregate demand than would have a monetary policy that relied upon market forces and control over monetary aggregates." Removal of credit ceilings was largely motivated by the desire to modernize monetary policy in a way that relied on market forces rather than government controls. Hodgman (1976) notes that "in 1971, however, after six years of heavy reliance on selective controls, the authorities became convinced that the growing disadvantages of this system had come to outweigh its advantages. In September 1971 most of the selective controls were removed in a broad reform that

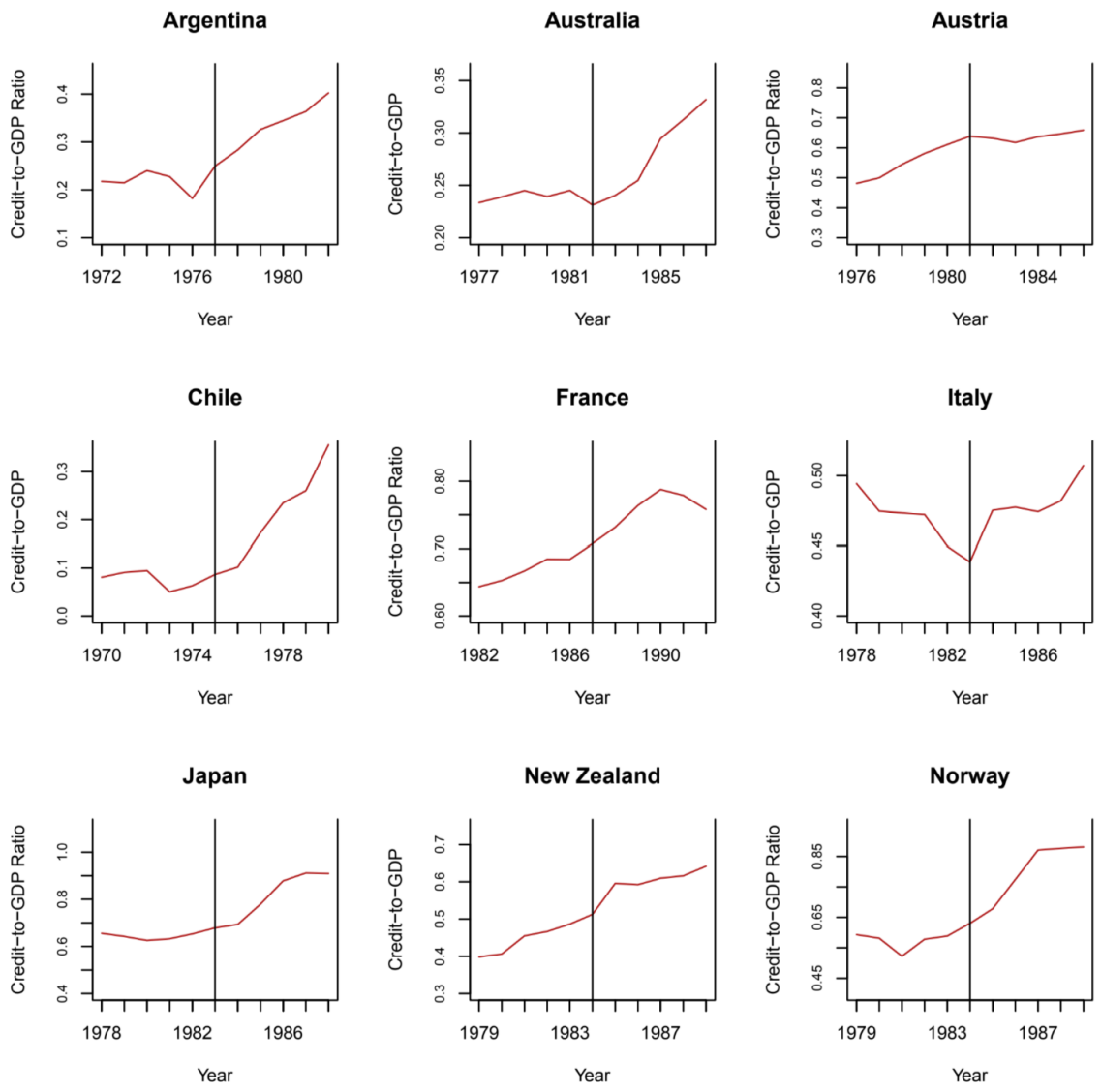
sought to enhance the role of competition and of market forces in transmitting the effects of official monetary and credit policies to the economy."

Contemporaneous accounts of central bank policy corroborates the notion that reformers in the Bank of England sought to remove credit controls to better stimulate competition within the banking sector and to modernize monetary policy. John S. Fforde, an official at the Bank of England, wrote in a letter to Her Majesty's Treasury: "prolongation of the present system is inconsistent with the Bank's fundamental and correct view that the shape of the banking industry, and its capability for change, should not be notably subordinated to the requirements of monetary policy. Banking, as a legitimate commercial activity, often inconveniences the Government of the day. There is accordingly a persistent temptation to convert the banks into mere slaves of official policy. We have always said, and rightly, that this is a temptation that must be resisted. The strength of our resistance to temptation can be judged by six years of ceilings... Our system of controls, by a delightful paradox, quite largely derives from a devoted adherence to the fundamental attitude mentioned above. It was fashioned according to the shape of the banking industry as we found it. It was intended to be entirely without prejudice to the development of that industry. But the credibility of this intention rested upon the presumption that the controls were only temporary, and their nature justified by the existence of an impermanent emergency. The idea that strict curbs on bank lending would be the rule rather than the exception finds no place in the official thinking of the early and middle 1960's upon which the present system of credit control was founded."

Appendix II: Additional Tables and Figures

Figure 2.4

This figure displays the growth of bank credit to GDP ratios for each of the 13 credit ceiling countries in our sample. For each country we plot bank credit relative to GDP in the 11 year period that extends from five years before credit ceilings are removed to five years after removal. Ratios are displayed in decimal form (i.e. credit/GDP) with the value in a given year determined by the height of the red line in each plot. The vertical bar in each plot shows the year (on the X-axis) in which credit ceilings are removed, according to sources listed in the next section of the appendix.



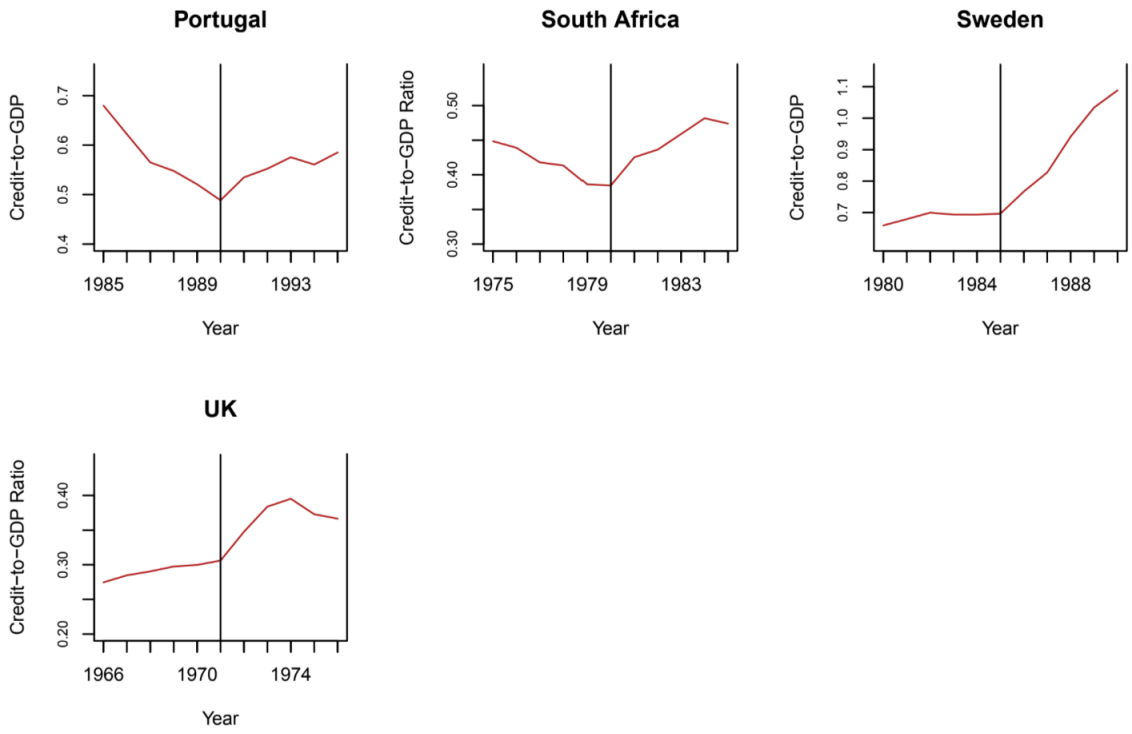
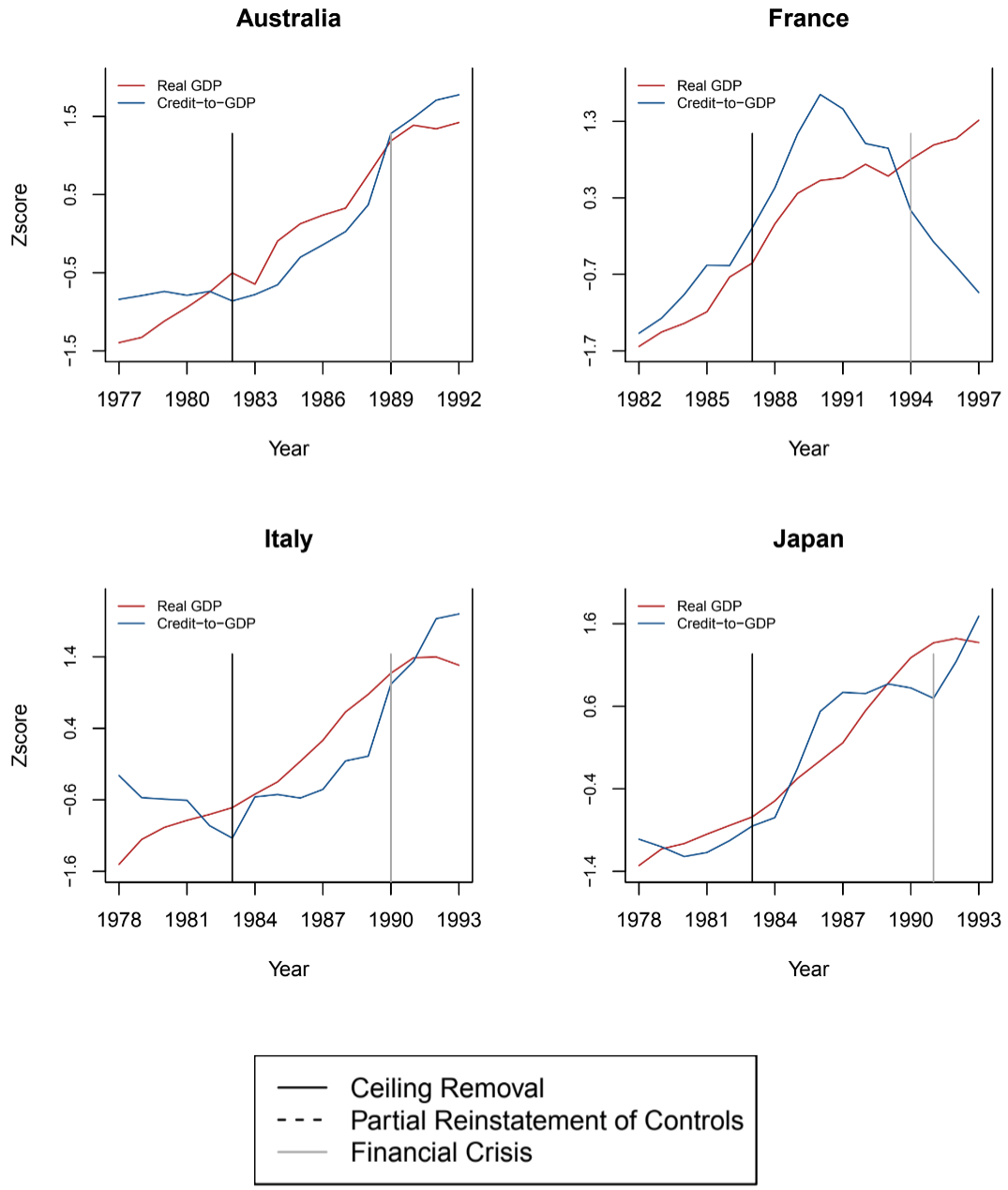
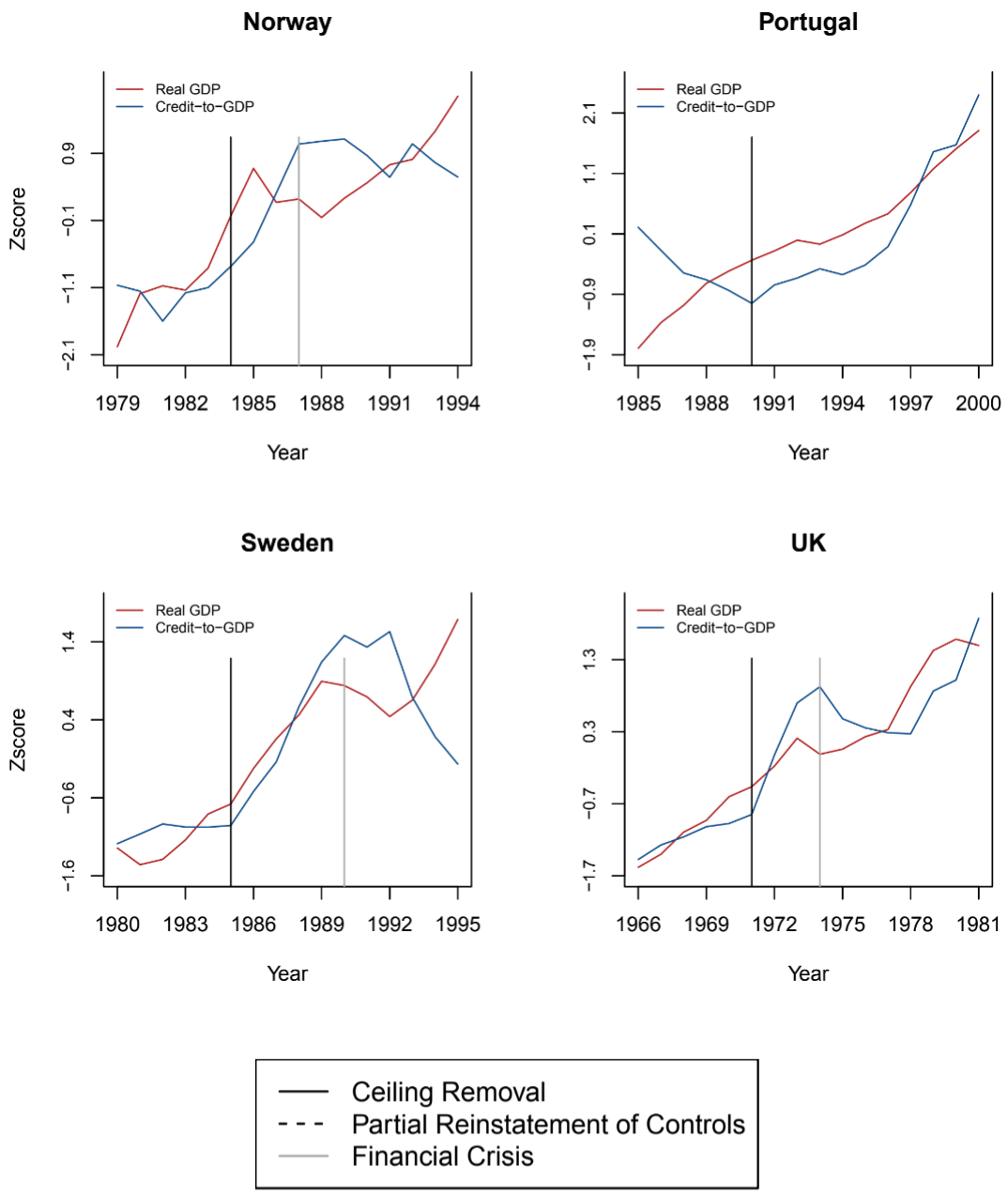

Portugal
Credit−to−GDP
0.4
0.5
0.6
0.7
1985
1989
1993
Year
South Africa
Credit−to−GDP Ratio
0.30
0.40
0.50
1975
1979
1983
Year
Sweden
Credit−to−GDP
0.7
0.8
0.9
1.0
1.1
1980
1984
1988
Year
UK
Credit−to−GDP Ratio
0.20
0.30
0.40
1966
1970
1974
Year

Figure 2.5

In this figure we display real GDP alongside bank credit to GDP ratios. Variables are standardized, with the y-axis displaying the number of standard deviations of each variable from its mean during the 16 year sample period shown in the plot. We display these quantities for the period beginning five years prior to credit ceiling removal and ending ten years after ceiling removal. Black vertical bars on each plot show the year that credit ceilings are removed, while gray bars indicate the year of a financial crisis, according to the Baron, Verner, and Xiong (2020) financial crisis list.




Norway
Real GDP
Credit−to−GDP
Zscore
−2.1
−1.1
−0.1
0.9
1979
1982
1985
1988
1991
1994
Year
Portugal
Real GDP
Credit−to−GDP
−1.9
−0.9
0.1
1.1
2.1
1985
1988
1991
1994
1997
2000
Year
Sweden
Real GDP
Credit−to−GDP
Zscore
−1.6
−0.6
0.4
1.4
1980
1983
1986
1989
1992
1995
Year
UK
Real GDP
Credit−to−GDP
−1.7
−0.7
0.3
1.3
1966
1969
1972
1975
1978
1981
Year
Ceiling Removal
Partial Reinstatement of Controls
Financial Crisis

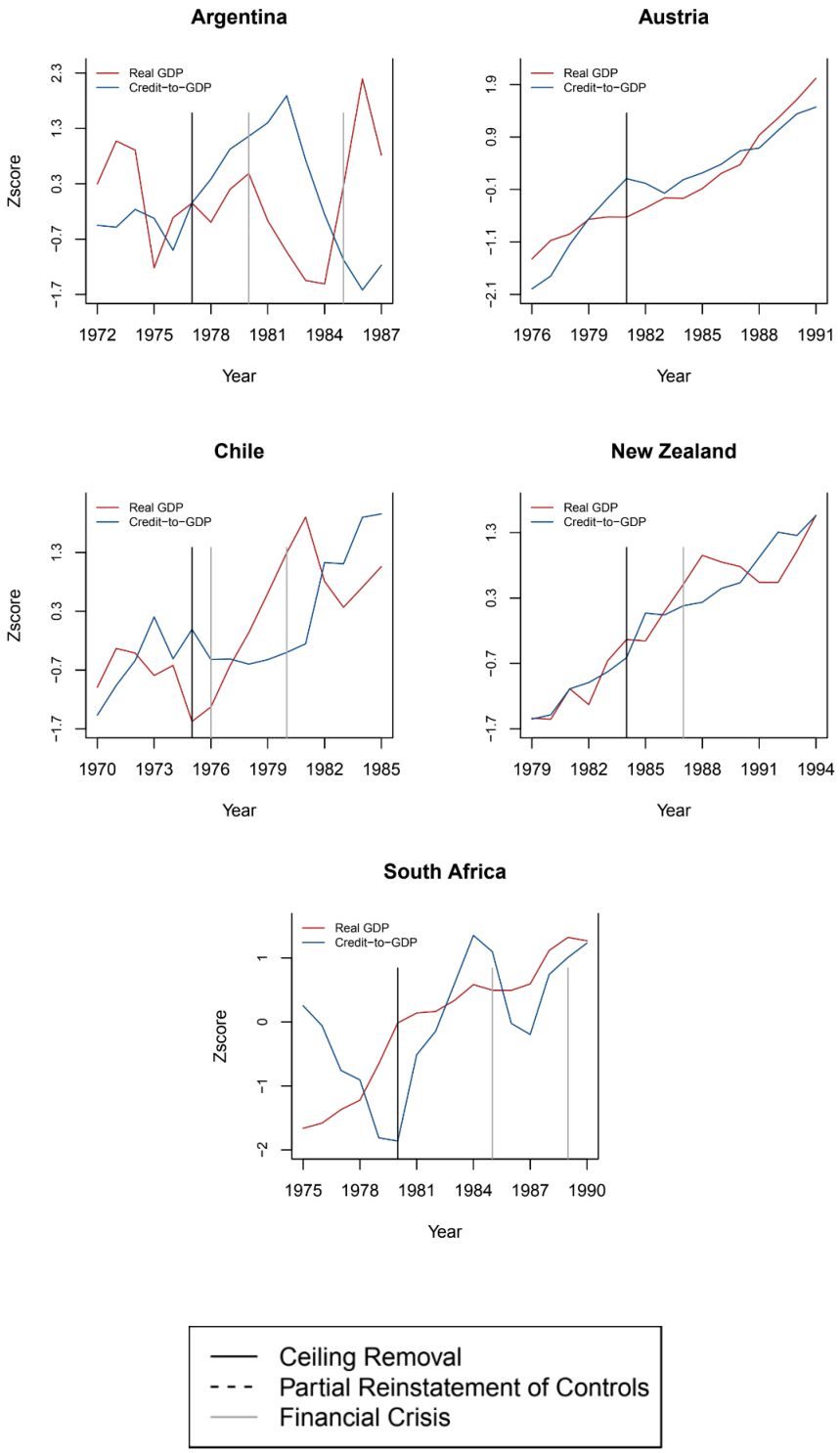
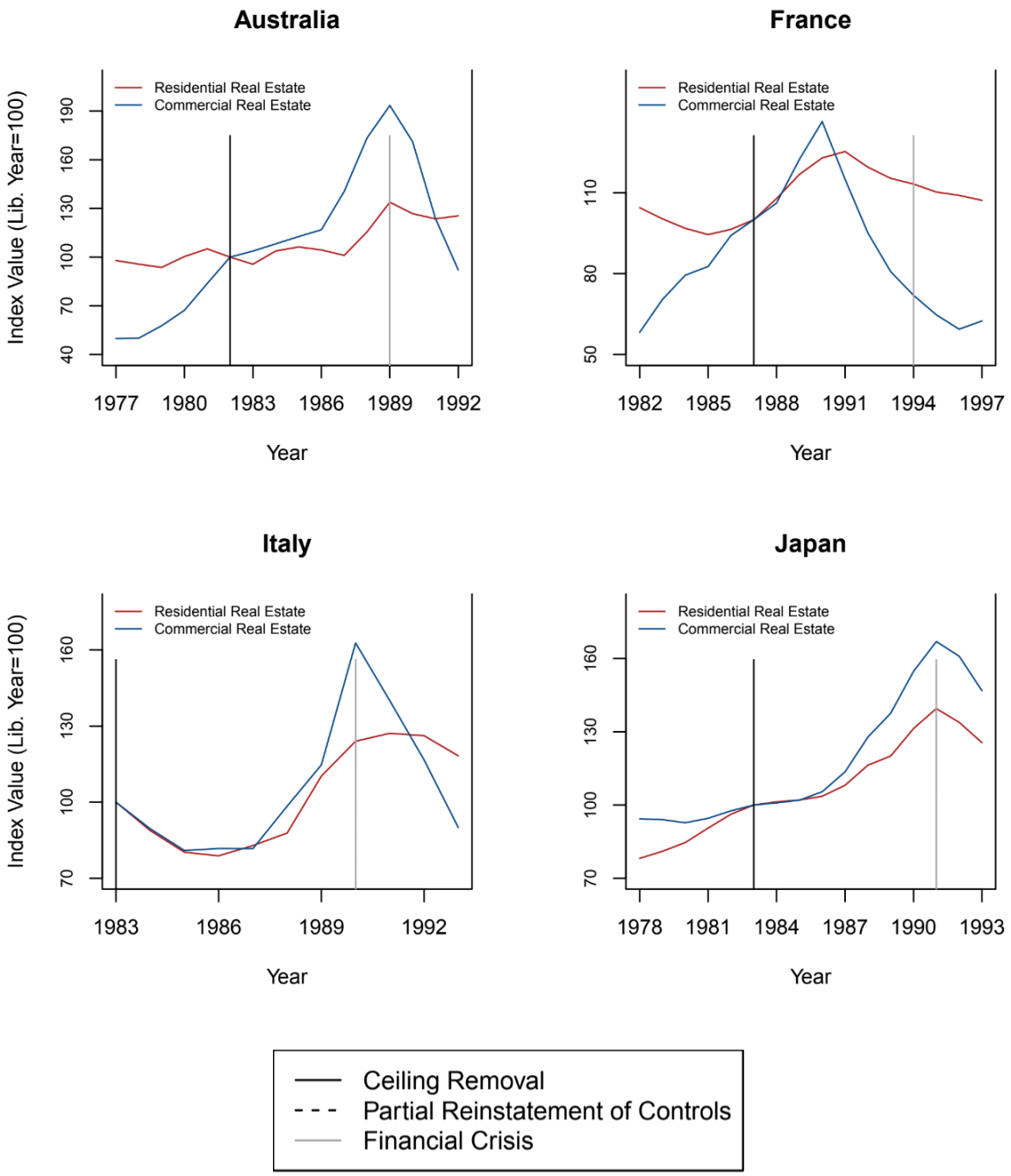

Argentina
Real GDP
Credit-to-GDP
Zscore
2.3
1.3
0.3
−0.7
−1.7
1972
1975
1978
1981
1984
1987
Year
Austria
Real GDP
Credit-to-GDP
1.9
0.9
−0.1
−1.1
−2.1
1976
1979
1982
1985
1988
1991
Year
Chile
Real GDP
Credit-to-GDP
Zscore
1.3
0.3
−0.7
−1.7
1970
1973
1976
1979
1982
1985
Year
New Zealand
Real GDP
Credit-to-GDP
1.3
0.3
−0.7
−1.7
1979
1982
1985
1988
1991
1994
Year
South Africa
Real GDP
Credit-to-GDP
Zscore
1
0
−1
−2
1975
1978
1981
1984
1987
1990
Year
Ceiling Removal
Partial Reinstatement of Controls
Financial Crisis

Figure 2.6

In this figure, we show the behavior of residential and commercial real estate prices, in a selection of credit ceiling countries (i.e. those for which sufficient commercial real estate data are available) in the period surrounding the removal of credit ceilings. We show these prices, when available, during the 16 year period extending five years before the removal of credit ceilings to ten years after removal (otherwise we begin the chart in the first year for which both of these data series are available). Residential real estate prices are shown in red and commercial real estate prices are given by the blue lines. Prices are shown as index values with each index set to a value of 100 in the year that credit ceilings are removed. Solid black vertical bars denote the year that credit ceilings are removed, while gray vertical bars denote financial crisis years.



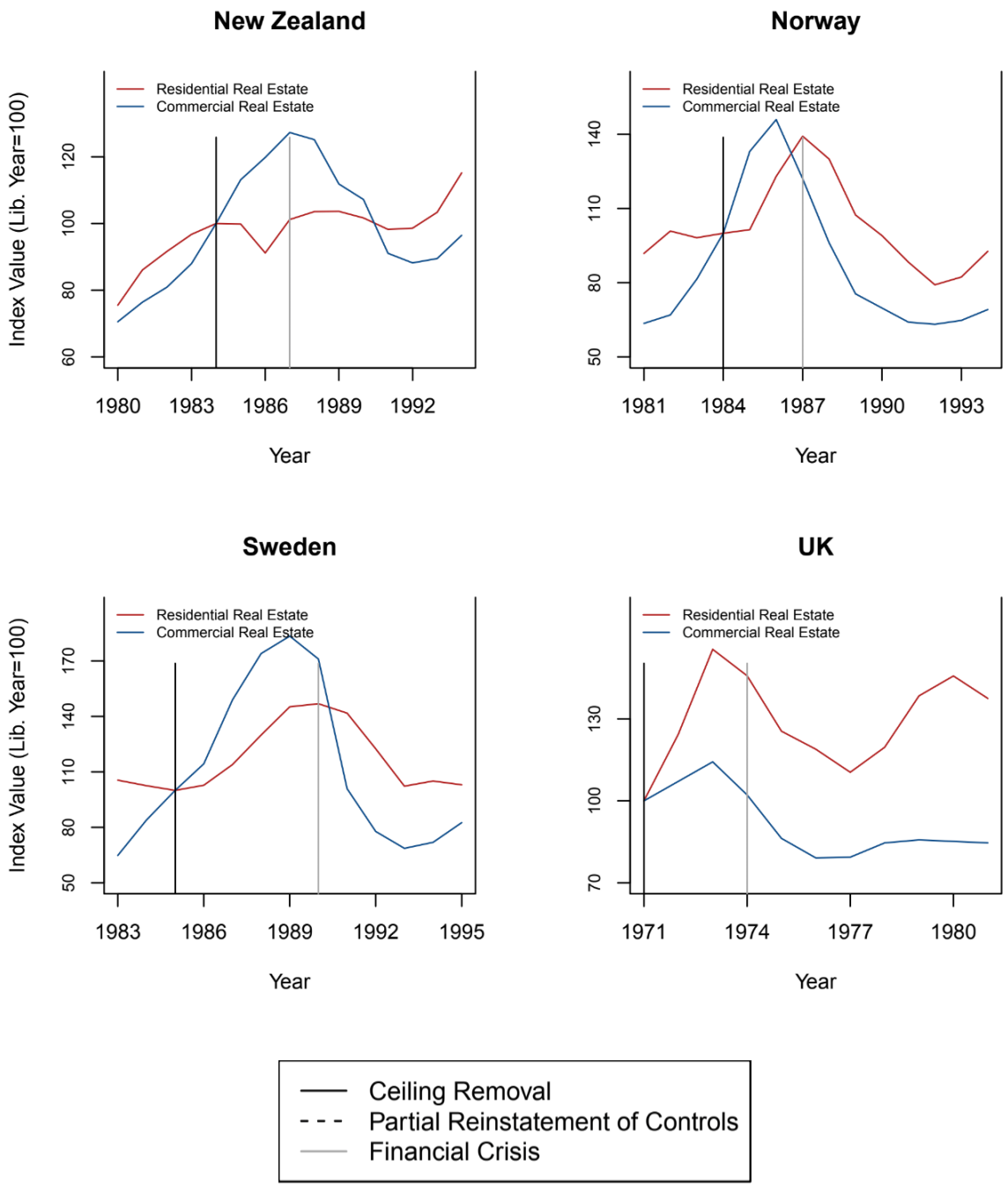
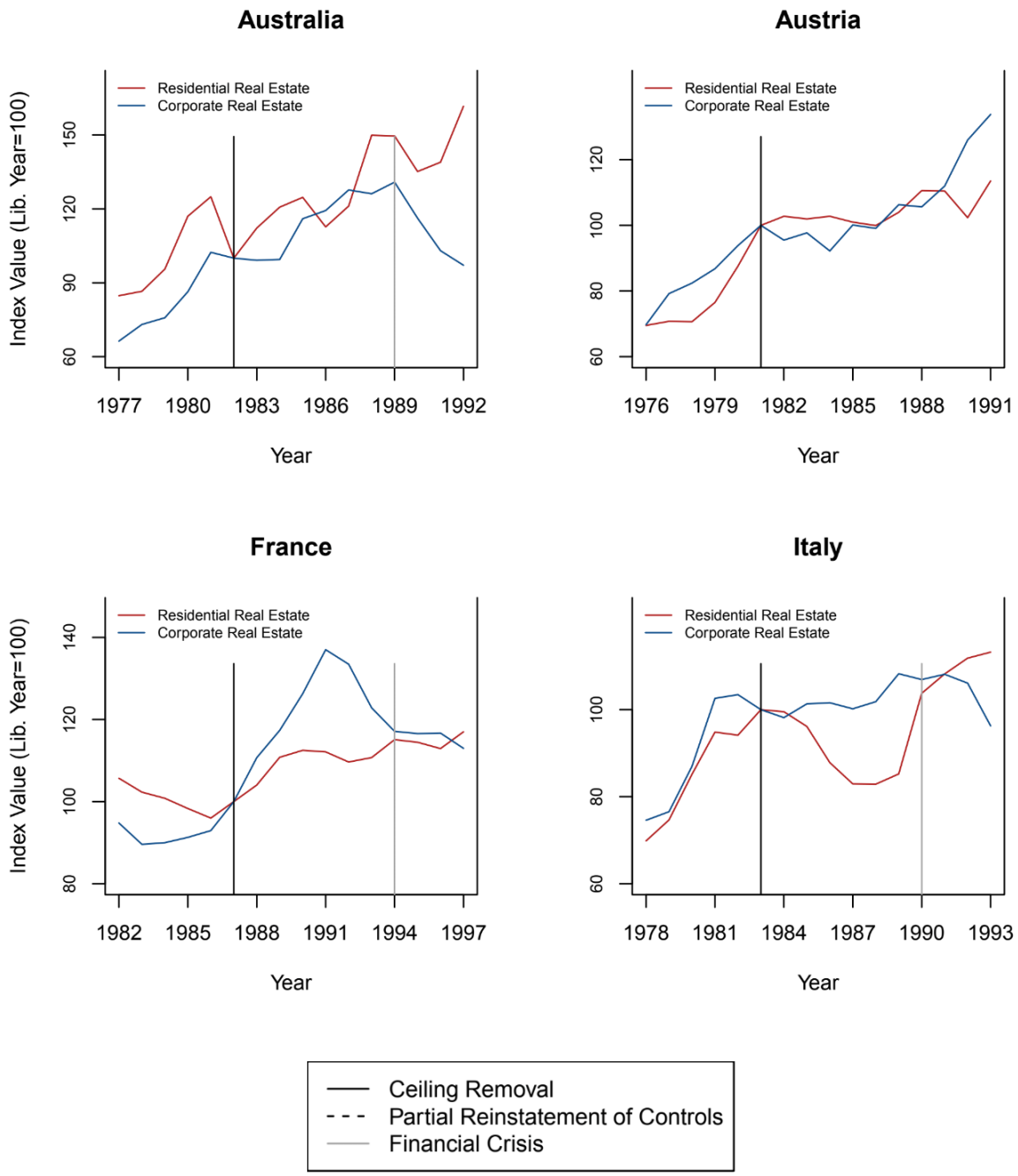
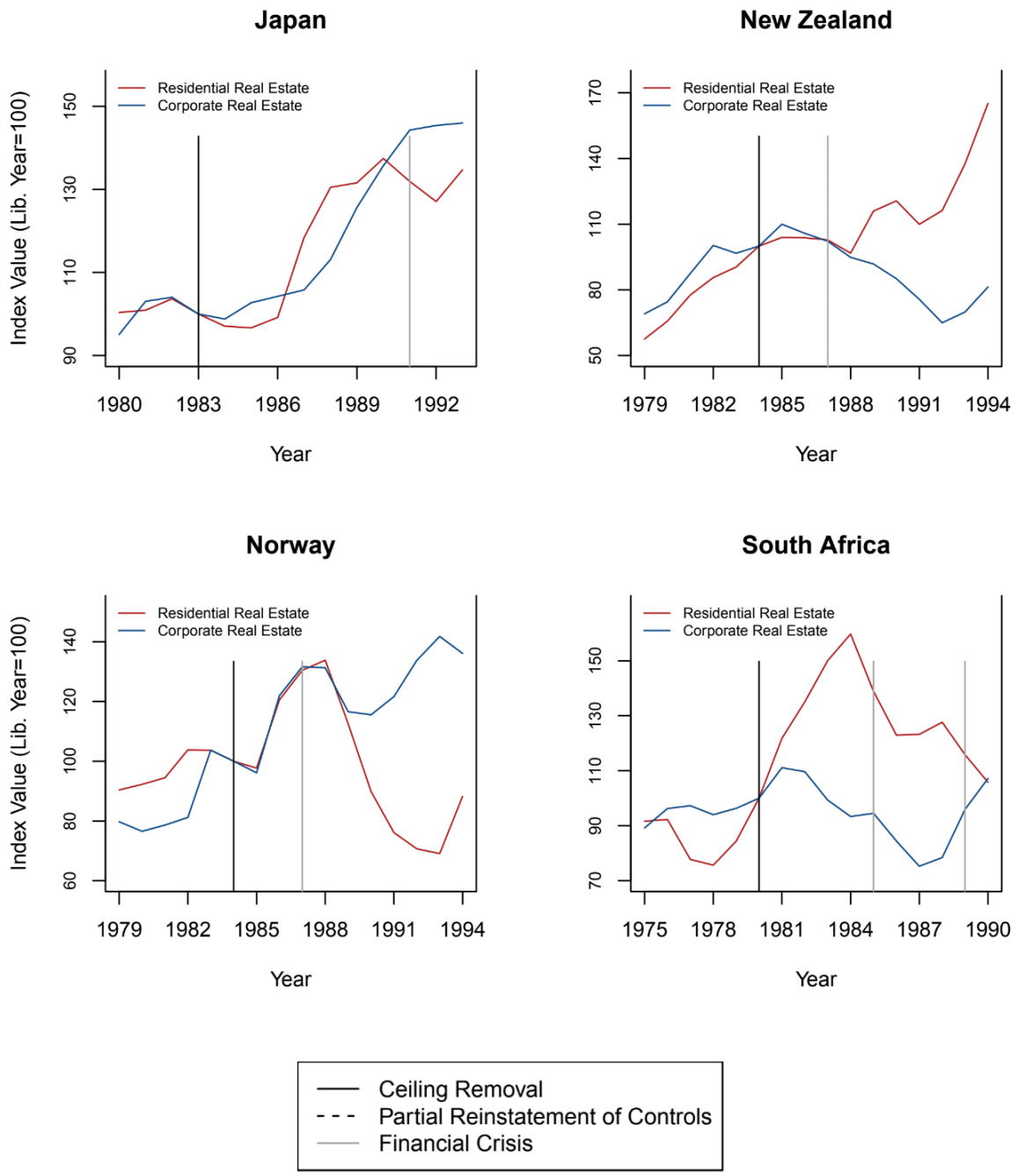

New Zealand
Norway
Sweden
UK
Residential Real Estate
Commercial Real Estate
Index Value (Lib. Year=100)
Year
Ceiling Removal
Partial Reinstatement of Controls
Financial Crisis

Figure 2.7

This chart shows analogous results to those of Figure 2.6 above, with investment quantities rather than prices. Due to the limitations of available data, "corporate real estate" investment is not necessarily calculated on the same structures for which commercial real estate prices are shown above. Commercial real estate prices are generally calculated from prices on real estate transactions on office buildings and other commercial spaces in a handful of major cities within a country. Investment quantities are corporate investments across a broader array of structures that include office buildings as well as factories and other classes of structures.




Japan
New Zealand
Norway
South Africa
Residential Real Estate
Corporate Real Estate
Index Value (Lib. Year=100)
Year
Ceiling Removal
Partial Reinstatement of Controls
Financial Crisis

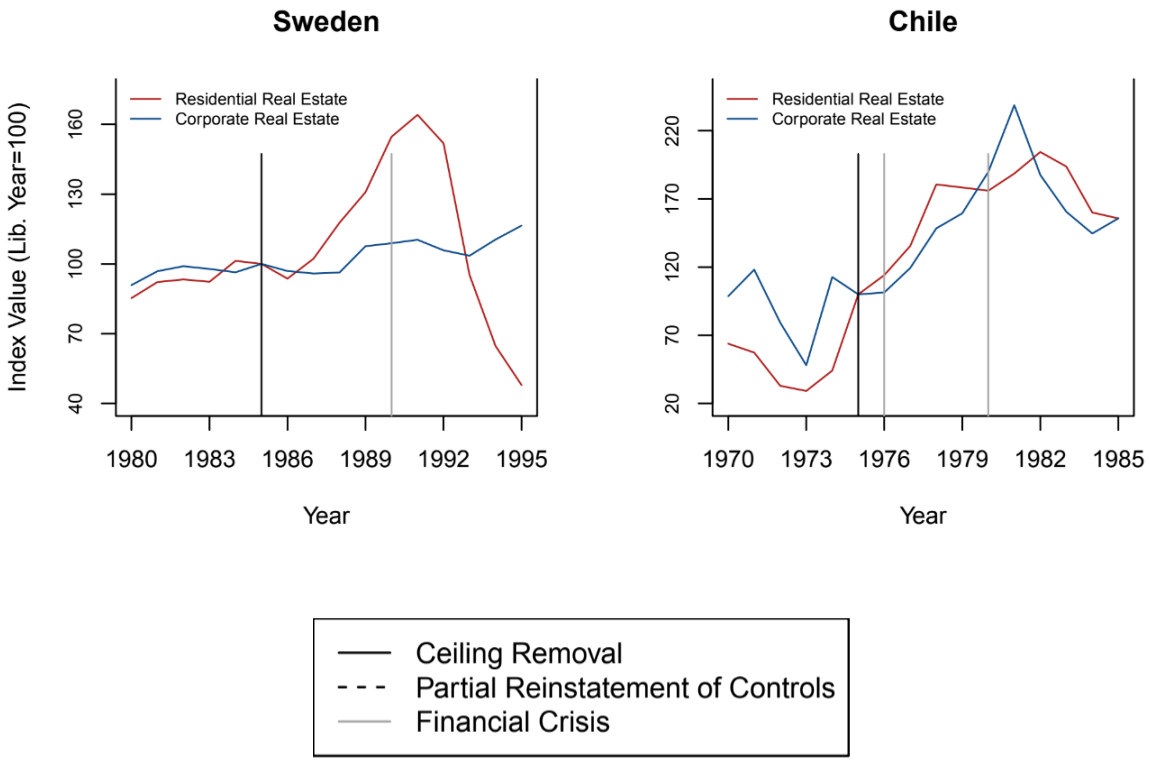
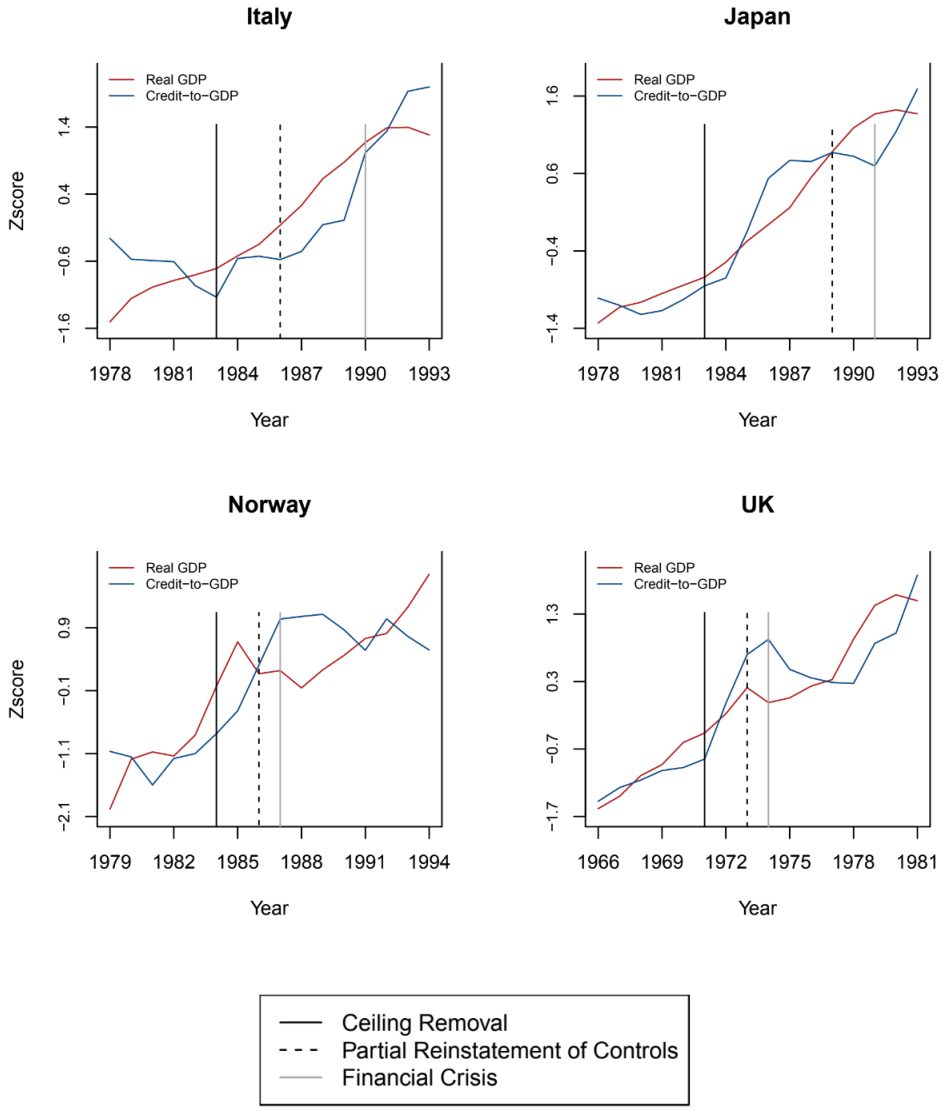

Sweden
Chile
Residential Real Estate
Corporate Real Estate
Index Value (Lib. Year=100)
40
70
100
130
160
20
70
120
170
220
1980
1983
1986
1989
1992
1995
1970
1973
1976
1979
1982
1985
Year
Ceiling Removal
Partial Reinstatement of Controls
Financial Crisis

Figure 2.8

In this figure we show evidence suggesting that credit booms have an irreversible quality. We show six countries for which new quantitative restrictions on credit were imposed in the years following credit ceiling removal. None of these countries fully reinstated the regimes that prevailed pre-liberalization. Italy, Japan, Norway, and Argentina installed new restrictions to credit growth which were generally weaker than pre-existing controls, and which were understood to be temporary (e.g. Italy explicitly announced new ceilings would last for a year, while Japan placed temporary growth targets only on a handful of large banks; Norway imposed weaker and temporary supplementary reserve requirements). The UK imposed new portfolio requirements on banks, and South Africa placed new controls on foreign capital. We show how bank credit to GDP and real GDP grow in the years surrounding credit ceiling removal. Here, we add a new dashed vertical bar to the charts in each year where partial ceilings are reimposed. We show the number of years that credit continues to grow; we also display, with gray vertical lines, the incidence of financial crises within short windows following the new controls on credit. We again focus on 16 year stretching from the 5 years prior to ceiling removal to 10 years after liberalization.



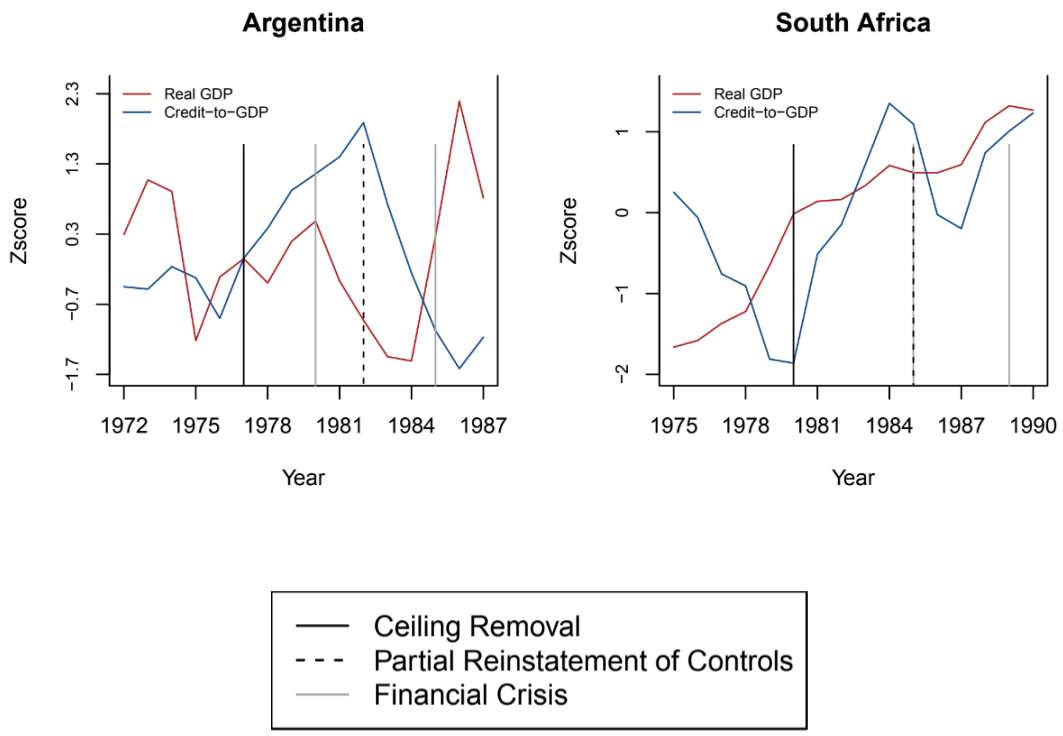
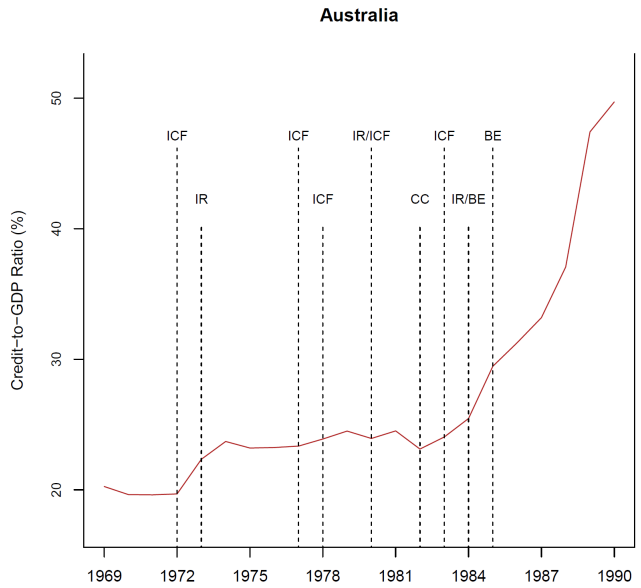
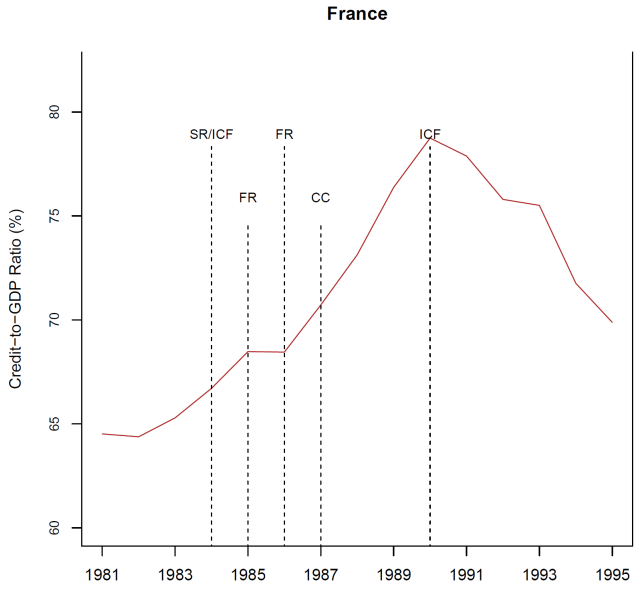
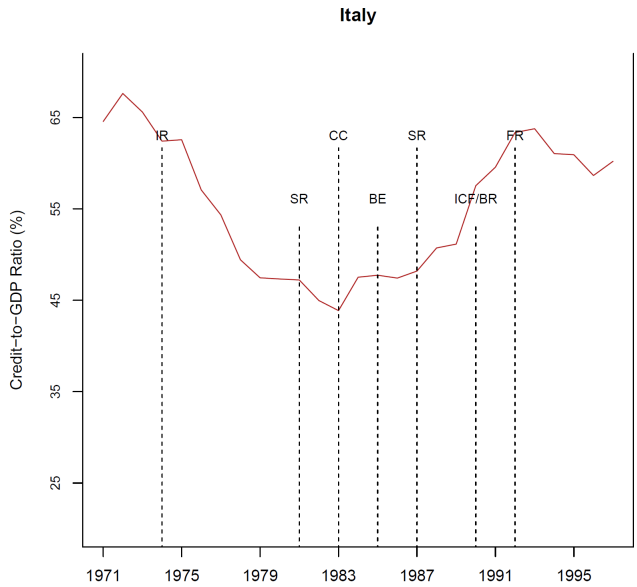

Argentina
Real GDP
Credit-to-GDP
Zscore
2.3
1.3
0.3
−0.7
−1.7
1972
1975
1978
1981
1984
1987
Year
South Africa
Real GDP
Credit-to-GDP
Zscore
1
0
−1
−2
1975
1978
1981
1984
1987
1990
Year
Ceiling Removal
Partial Reinstatement of Controls
Financial Crisis

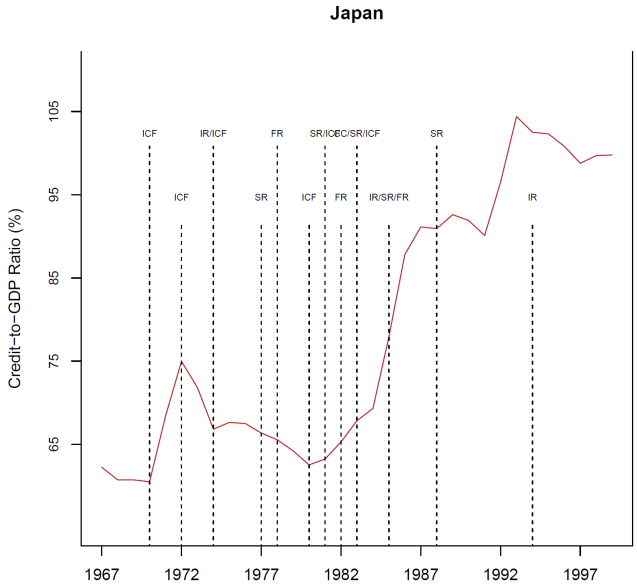
Figure 2.9

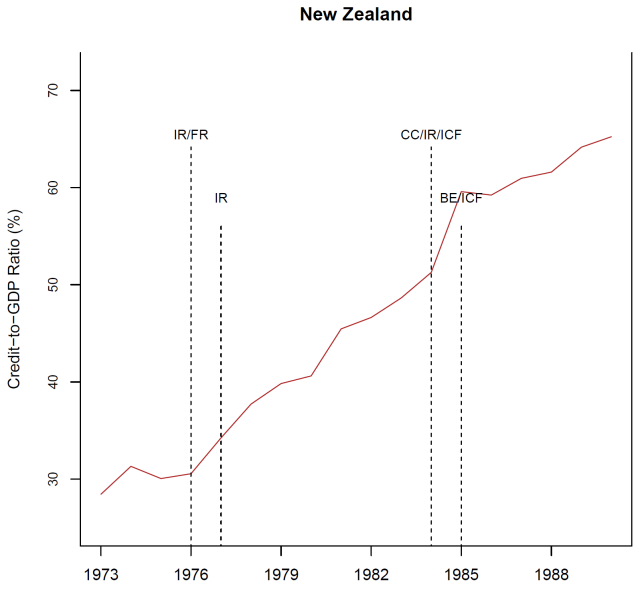
In this figure, we display timelines of major financial reforms in a number of credit ceiling countries. Vertical dashed bars denote the years of major financial reforms of seven types. These include removals of credit ceilings, removals of branching restrictions, removals of barriers to entry of foreign banks, removals of interest rate controls, abolitions of restrictions on international capital flows, financial reforms in capital markets (i.e. removals of restrictions in bond markets, or short-term credit markets/money markets), and removals of bank specialization restrictions (i.e. prohibitions on banks trading securities, offering mortgages, etc.). The abbreviations for each of these policy types are displayed above the vertical dashed bars, with the legend below further explaining each abbreviation. In some years, multiple policies are liberalized, in which case policy types are separated with a "/" above (e.g. "IR/ICF" would means both interest rate controls and international capital flow restrictions are removed in the same year). In addition to these deregulatory timelines, we also display each country's credit to GDP ratio across the years covered by the plot. This is to give a sense of the correlations between credit growth and deregulation of various types.

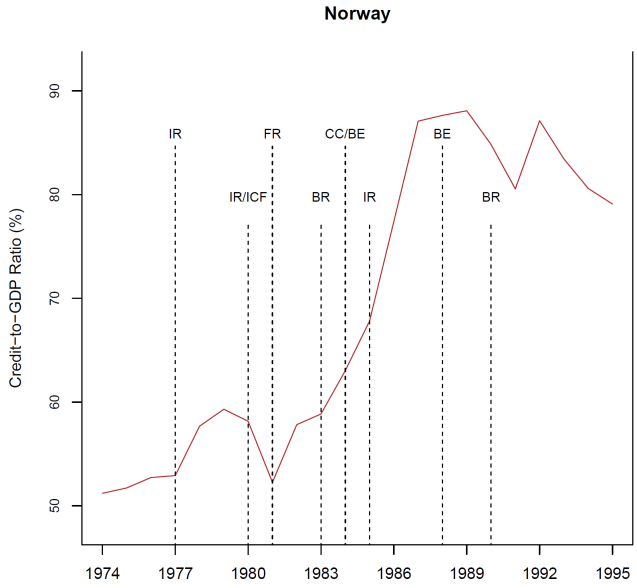


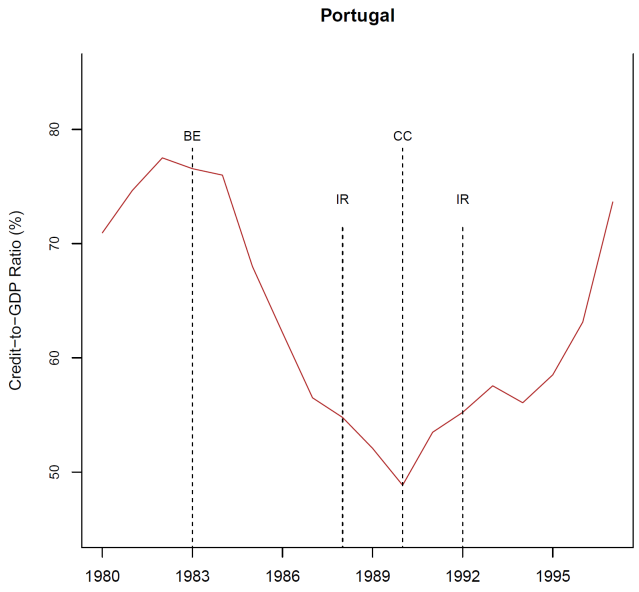


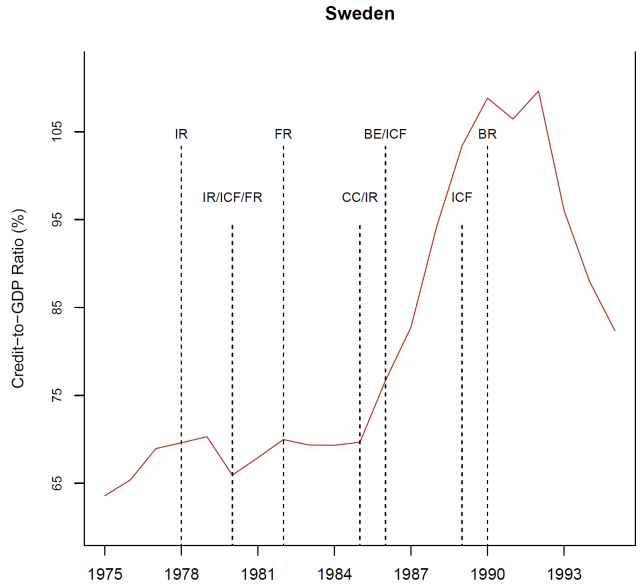

Italy
Credit−to−GDP Ratio (%)
25
35
45
55
65
IR
SR
CC
BE
SR
ICF/BR
FR
1971
1975
1979
1983
1987
1991
1995


Japan
Credit−to−GDP Ratio (%)
65
75
85
95
105
ICF
IR/ICF
FR
SR/ICC/SR/ICF
SR
ICF
SR
ICF
FR
IR/SR/FR
IR
1967
1972
1977
1982
1987
1992
1997


New Zealand
Credit−to−GDP Ratio (%)
30
40
50
60
70
IR/FR
IR
CC/IR/ICF
BE/ICF
1973
1976
1979
1982
1985
1988


Norway
Credit−to−GDP Ratio (%)
50
60
70
80
90
IR
IR/ICF
FR
BR
CC/BE
IR
BE
BR
1974
1977
1980
1983
1986
1989
1992
1995


Portugal
Credit−to−GDP Ratio (%)
50
60
70
80
BE
IR
CC
IR
1980
1983
1986
1989
1992
1995


Sweden
Credit−to−GDP Ratio (%)
65
75
85
95
105
IR
IR/ICF/FR
FR
CC/IR
BE/ICF
ICF
BR
1975
1978
1981
1984
1987
1990
1993

BE=Barriers to Entry
BR=Branching
CC=Credit Ceilings
FR=Fin. Market Reforms
ICF=Capital Flows
IR=Int. Rate Restrictions
SR=Specialization

Table 2.8

This table compares credit growth in the three years following credit ceiling removal, to credit growth in across all other three-year windows. To calculate credit growth in the years after liberalization, we subtract a country's credit to gdp ratio in year t from its credit to GDP ratio in year t+3, where year t denotes the year credit ceilings were removed. We perform this comparison country-by-country and show the difference between credit growth after liberalization and average credit growth during other three-year windows. This difference is shown in the final column.

	3-Year Avg. Credit Growth		
	Non-liberalization Years	Liberalization Years	Difference
Argentina	-0.013	0.095	0.108
Australia	0.063	0.063	0.000
Austria	0.045	-0.001	-0.047
Chile	0.031	0.148	0.117
France	0.029	0.047	0.017
Italy	0.035	0.036	0.001
Japan	0.023	0.200	0.176
New Zealand	0.069	0.097	0.028
Norway	0.034	0.241	0.207
Portugal	0.060	0.087	0.027
South Africa	0.023	0.074	0.051
Sweden	0.035	0.245	0.210
UK	0.057	0.089	0.032
Total Avg.	0.037	0.109	0.072

Table 2.9

This table shows the correlation coefficients associated with each pair of policy variables. Policy indicator variables are constructed by finding the dates of policy reform or removal for each type of policy, across the 13 credit ceiling countries. For each country, an indicator variable for a particular policy assumes a value of one in any year in which there was a meaningful liberalization of a policy of that particular type. The variable takes a value of zero in all other years. In the columns and rows below, the policy types are labelled similarly to tables 2.6 and 2.7 of the main exhibits. Refer to those tables, and the text, for a further description of the types of policies reforms included in each variable.

Correlations Between Binary Policy Variables							
	Ceilings	Rates	Barriers	Branch.	Spec.	Cap. Flows	Fin. Reform
Ceilings	1						
Int. Rates	0.205	1					
Barriers	0.100	0.054	1				
Branching	-0.014	-0.021	-0.014	1			
Specialization	-0.020	0.054	-0.019	-0.014	1		
Cap. Flows	0.039	0.206	0.124	0.087	0.202	1	
Fin. Reform	-0.018	0.170	-0.017	-0.012	0.083	0.033	1

Chapter 3: Elections Have Consequences: Federal Spending and Its Link to Presidential Politics

1. Introduction

In the United States, as in most developed economies, government spending accounts for a substantial proportion of GDP. In the 2021 fiscal year, federal spending alone accounted for roughly 30% of the $23 trillion American economy.[90] Given the vast sums involved, it is important to understand the determinants of federal spending, and to analyze the incentives facing the policymakers who allocate federal funds. Anecdotal evidence and empirical research suggest that federal electoral politics may play an important role in the provision of government funds. For example, Creary, Greer, Singer, and Willison (2019) show that during the 2017 hurricane season, federal hurricane relief to victims of Hurricane Maria in Puerto Rico was slower and less generous than the federal responses to hurricanes Harvey and Irma in Texas and Florida, respectively. Members of the media and the public have interpreted these responses as reflective of the large constituencies of Republican voters in Texas and Florida (during a Republican presidency)[91] and the relative lack of importance of Puerto Rico in federal presidential politics.[92] More systematic research on federal expenditures (e.g. Levitt and Snyder, 1995; Carsey and Rundquist, 1999; Bickers and Stein, 2000; Cohen, Coval, and Malloy, 2011) have shown that powerful members of Congress are able to exert disproportionate influence in directing funds to their districts, and that the ideological composition of influential congressional committees can be an important driver in the allocation of funds. A smaller, but growing, body of work has recognized the importance of the president, and of presidential politics, as important drivers of the disbursement of public funds.

[90] https://datalab.usaspending.gov

[91] See, e.g. a transcript of National Public Radio (NPR) commentary; https://www.npr.org/2018/03/29/598053556/a-comparison-of-how-the-government-responded-to-hurricanes-harvey-and-maria

[92] While Puerto Rico is a self-governing US territory, where nationals enjoy the rights of US citizenship, Puerto Ricans living in Puerto Rico do not have a vote for the US President, as presidential votes are given only to residents of the 50 states and the District of Columbia.

In this paper, we use 25 years of county-level data on government expenditures, in the form federal grant payments (largely federal aid for state and local projects) and government procurement contracts, to evaluate how this spending responds to the electoral incentives of the U.S. president and his or or her political party. We focus our attention on the presidency, both because the role of the president in the allocation of federal funds is relatively less well studied, and because we believe that the importance of national (as opposed to state-level or local) politics has likely become more important over time. While many studies have focused on the role of the legislative branch in allocating funds, because of the authority granted to that branch in the United States constitution to control taxation and spending, the president is in charge of proposing a federal budget, and wields veto power over the subsequent appropriations bills passed by Congress after making budget recommendations. Moreover, the president is in charge of various federal agencies which allocate funds appropriated by Congress. In practice, agencies are often given substantial discretion by Congress to allocate funds appropriated to them. Thus, the president, through various channels, wields substantial power in allocating federal funds, despite the formal vesting of taxation and spending power with the federal legislature.

We use our data on federal expenditures to test the two main empirical predictions that have arisen from theoretical models of U.S. budget process and political system. The first of these theories suggests that the president might seek to direct federal funds in a manner that rewards his most reliable supporters, and thus that federal expenditures will flow into jurisdictions where the president has garnered substantial support from voters. This hypothesis is supported by theoretical models presented by McCubbins and Cox (1986) and Mebane and Wawro (2002). The authors of the former study suggest that voters with strong preferences on particular issues form factions, or voting blocs, with like-minded voters, and candidates compete for the support of these factions by adopting policies, including spending proposals, which are favorable to them. In order to ensure the long-term viability of their political parties (i.e. the set of factions that make up their key governing coalitions), political candidates must adopt spending policies favored by their core supporters.

The second prominent theory of the president's role in federal spending proposes that the president will seek to reward voters in closely contested "battleground states," which are particularly crucial to a presidential candidate's ability to win an upcoming election. Given the series of state-by-state popular votes which typify the presidential election system in the United States, and the fact that election results across a substantial number of states are essentially a foregone conclusion prior to the beginning of a presidential campaign, the outcome of a presidential contest often boils down to the support of both candidates in this relatively small number of undecided electoral battlegrounds. Given the importance of these states to the overall result of an election, it is plausible that the president in office might seek to direct funds to these states in order to improve local economic conditions, and thus enhance their favorability among voters in these states.

In order to test these hypotheses, we divide states into three sets, based on the results of the most recent presidential election. These sets include states that strongly supported the winning presidential candidate in the most recent election (i.e. the current president), closely contested battleground states, and states which voted for the losing candidate by a substantial margin. We sort states into these three sets by sorting all states based on their margins of victory, and finding the so-called "tipping point state" or the state which, based on its margin of victory, put the winning candidate over the electoral vote threshold needed to win the election. We then label the ten states with margins of victory closest to this tipping point state as the set of battleground states. We then categorize states outside of this set of battleground states, as belonging solidly to one or the other political party.

Our method for sorting states is novel within the political spending literature, but we believe it is superior to the alternative approaches of defining battlegrounds as states with close absolute margins of victory or as states which have switched support between major candidates in recent elections. The first of these approaches is inappropriate for labeling battlegrounds in years, like 1996 or 2008, with relatively one sided election results, since states with close absolute margins of victory in those years would likely be a solid part of the opposing party's coalition in a year in which the overall election result was relatively close.

The latter approach, of defining battlegrounds as states which have switched affiliations in recent elections, is less forward-looking and would omit certain states which were known at the time to be presidential battlegrounds. For example, prior to 2016, Michigan and Wisconsin had voted for the Democratic party's candidate for president in every election going back to 1992, and would thus not be coded as battleground states under this alternative approach, despite having close margins of victory that made it clear that they would be heavily contested by both major candidates in 2016.

After sorting states into these groups, we test our hypotheses on the impact of presidential elections on federal spending using a "state borders" identification scheme. Rather than comparing spending outcomes across all fifty states, our identification approach focuses on comparisons of county-level spending outcomes across neighboring counties located on opposite sides of adjacent state borders in pairs of states which experienced different outcomes in recent presidential elections.

As an example of this approach, consider a test of the hypothesis that the president channels funds to politically important battleground states. In our empirical tests of this hypothesis, we focus on identifying pairs of bordering states, such as Pennsylvania and New York in the wake of the 2016 presidential election, where one state is a closely-contested battleground state (Pennsylvania) and the other state strongly supports one of the two major political parties (New York). We then compare counties located close to the borders of these adjacent states and examine whether we see sharply higher spending in counties located within the battleground state, as compared to the state with a relatively lopsided election outcome. The central idea behind these comparisons is that counties located nearby one-another tend to be far more similar than geographically distant counties and states. Thus, the influence of unobserved factors influencing a county's propensity to attract federal funds should be severely diminished when comparing sets of counties located close together. Our methodology in conducting these tests is similar to the earlier borders-based identification literature of Holmes (1998), which compares manufacturing outcomes across right-to-work and union-friendly states, and Pence (2006) which compares housing markets across states

with judicial foreclosure laws and those without such policies. The approach has also been used frequently in studies of local labor markets (see, for example, Moretti and Wilson, 2017; Ljunquist and Smolyanski, 2018; Campello, Gao, and Xu, 2019).

Our empirical tests uncover preliminary support for both of the main hypotheses on the effects of presidential politics on federal spending. We find that as compared to counties in a bordering state which voted strongly in favor of the losing candidate in the most recent presidential election, counties in states that strongly support the winning presidential candidate see between 18% and 29% stronger spending in the form of federal government procurement contracts and between 16% and 25% stronger federal grant spending, in our preferred specifications. We also find support for the notion that battleground states receive relatively strong federal funds allotments. We show that on average, as compared to a county located in a state with a relatively lopsided recent presidential election result, counties in neighboring battleground states see between 5% and 16% stronger federal procurement expenditures and between 4% and 16% stronger federal grant spending. While counties across neighboring state borders should be substantially similar, there can, in practice, be substantial differences across neighboring states, particularly if state-level government policies produce unique economic situations (e.g. a stronger manufacturing environment, or poorly-funded public schools) which tend to attract stronger federal spending to one state but not its neighbor. To deal with this possibility, we take advantage of our relatively long sample and look at state pairs which shift their political allegiances over time. Our regressions utilize state fixed-effects and a difference-in-differences setting to show that spending differentials between neighboring states tend to shift in favor of closely contested states in the years after an election in which they become electoral battlegrounds for the first time.

In the subsequent part of our analysis, we seek to more directly address the question of whether battleground states or states loyal to the sitting president more strongly attract federal outlays. In our initial tests of the hypothesis that battleground states attract strong federal spending, we do not differentiate within our set of paired control-states (i.e. states with lopsided election margins bordering battlegrounds) between states which vote

by large margins for the winning candidate and those which favored the losing candidate. Recognizing that our earlier tests, comparing states which strongly support the winning vs. losing candidates, imply heterogeneity between these sets of states, we face off battleground states and states with strong support for the winning candidate. We generally find that the tendency to reward "loyal" states is stronger than the tendency to direct spending to battlegrounds. We estimate that federal grant spending is roughly 10% higher in states that strongly supported the winning presidential candidate in the most recent election as compared to battleground states, though both see stronger spending than states that strongly supported the losing candidate.

Lastly, we show evidence of time variation in the relative strength of the effects we estimate. In particular, we show evidence suggesting that the tendency of the federal government to direct federal outlays toward battleground states grows stronger in presidential election years. While our results suggest that in a typical year outside of a presidential election, the value of federal grants directed toward states that strongly support the current president exceeds grant spending in battleground states by 10% or more, our results suggest that in an election year, this gap shrinks substantially. In a number of specifications, our results suggest that in an election year, aggregate spending in counties located within battleground states actually exceeds spending in neighboring states that strongly support the president. In the aggregate, we find that the quantity of federal grants directed to battleground states increases by between 3% and 9% in election years, depending on the specification. We interpret these results as suggesting that while the president might generally prefer to adopt policies that ensure the stability and loyalty of his core group of voters, in the lead-up to a presidential election, it becomes more valuable to focus on the support of voters in closely contested states.

Our paper contributes to a growing literature on the interplay between political campaigns, election outcomes, and government spending. A number of recent studies (e.g. Brogaard, Denes, and Duchin, 2015; Carvalho, 2014; and Finan and Mazzocco, 2016) analyze how political incentives, and the needs of politicians to raise campaign funds, or to

garner electoral majorities, affect the way discretionary funds are allocated. Relative to these studies, our paper focuses on the impact of presidential elections in the United States, the importance of individual states to election outcomes, and the spending patterns that arise from the unique features of the Electoral College system in U.S. presidential elections.

A few other studies exist which, like ours, focus on the importance of the presidency to spending patterns. Within this set of papers, empirical evidence is generally mixed with respect to how presidential politics affect federal spending. Wallis (1987) suggests that states which do not reliably support either political party from one election to another tend to attract higher federal spending. On the other hand, Anderson and Tollison (1991) and Larcinese, Rizzo, and Testa (2006) find evidence suggesting that regions containing larger shares of voters supportive of the sitting president tend to receive higher federal outlays.

We seek to advance this literature in a number of ways. First, we analyze a relatively long sample, comprised of federal expenditures in the form of grants and procurement contracts, covering the period from 1996-2020, which covers parts of six presidential election campaigns. A number of existing studies (e.g. Wallis, 1987; Mebane and Wawro, 2002) analyze relatively short periods that cover only one or two federal election cycles, making it difficult, given the lack of time-variation in state-level political outcomes, to rule out the possibility that states which attract high levels of federal funds do so for reasons other than their political affiliations.[93] Our longer sample covers a period over which relatively slow-moving regional political affiliations go through periods of change, allowing us to make comparisons across many different groups of states, and to include state fixed-effects in a number of specifications to observe how spending in individual states change following a change in that state's political allegiances.

Among existing empirical studies, Larcinese, Rizzo, and Testa (2006) also includes a relatively long sample and employs state fixed-effects. Relative to this paper, we make additional methodological contributions to more precisely test prevailing theories on presidential political incentives and federal outlays. Specifically, we conduct our tests using more

[93] The model in Mebane and Wawro (2002) has other, more elaborate, empirical predictions which are arguably addressed more easily by their mainly cross-sectional data.

granular county-level data, rather than the state-level spending data employed by Larcinese, Rizzo, and Testa. The county-level data allows us to control for a number of observable features which are likely to affect the allocation of federal funds within states. We then utilize this county-level data to employ the "state borders" identification scheme across a variety of empirical specifications, as previously discussed. Our empirical approach makes greater use of the geography of various states by focusing on comparisons between pairs of bordering states, rather than utilizing the full cross-section of states. Doing so allows us to better match our sample and identify sets of counties which, we argue, are substantially similar, and would be likely to attract similar levels of federal funds in the absence of differences in their political importance to presidential elections.

Existing resesarch on politically motivated spending has often used aggregate data across all categories of federal expenditures within a state (e.g. Larcinese, Rizzo, and Testa, 2006) or used very narrow classes of expenditures, such as defense contracts, to show, for example, how the composition of important senate committees contributes to the location of these outlays. We balance these two approaches by separately considering two types of federal expenditures: federal procurement contracts, which capture payments by the federal government (or specific federal agencies) to private firms for the provision of goods and services, and federal grants, which consist of awards by the federal government to communities to pursue a specific project or implement and federal program. We believe that analyzing these two classes of spending allows us to capture a substantial quantity of federal spending, while largely removing spending through programs like Social Security, over which the president has little short-term discretion.

Finally, our research is the first, to our knowledge, to examine the time variation of the two effects that we uncover and to show that federal spending appears to more strongly target presidential battleground states in the year of a presidential election. We interpret these results in light of existing theories of political spending, which suggest that targeting marginal voters is likely to be a short-term strategy which is most valuable in the lead-up to an election.

The rest of this article is organized as follows. In the next section, we give a brief background of presidential elections in the United States, and discuss our hypotheses and methodology in light of this institutional background. In section three, we discuss our data and present summary statistics. In section four, we test our two hypotheses on how federal spending responds to the incentives created by presidential elections. We compare the effects that we uncover in support of these hypotheses and discussing their relative magnitudes and timing. We discuss the implications of our results and present some concluding thoughts in section five.

2. Institutional Details and Hypotheses

Our hypotheses and empirical design are largely centered on the unique way in which presidential elections are decided in the United States. Unlike many other democracies, United States presidential elections are not decided on the basis of a simple nation-wide majority vote. Instead, the president is selected via a body called the "Electoral College," a group of individuals comprised of representatives (called "electors") from each of the fifty states, which meets after each presidential election to determine the president. In a presidential election, each state gets to nominate a collection of electors to represent their state and cast their votes for the presidency. A presidential candidate who gains the support of a majority of electors then becomes the president.

Each state is entitled to nominate a number of electors which depends on that state's population. In particular, the number of electors for a state is the total of that states representation in Congress: the sum of that state's senators and delegates in the House of Representatives. Each state has two senators, and a number of representatives that is allocated based on a state's population (but bounded below at one representative), so that larger states have more presidential electors than smaller states.[94] The total number of

[94]Since every state has two senators and at least one representative, large states do not have proportionally more representatives than smaller states. That is, the population per elector tends to be higher for more populous states.

electors apportioned through each presidential election is 538[95] and thus a winning candidate must receive 270 votes in the Electoral College.[96]

States have broad discretion to choose how to select their electors, however, in practice, all but two states allocate all of their electors to the winner of the state-wide popular vote.[97]Thus, the US presidential election can be thought of as a series of popular votes held in each of the fifty states. The reason that this is not equivalent to a simple nationwide popular vote is that a candidate achieving a substantial majority of votes in a particular state is no better-off than one who receives a narrow majority in that same state. Thus, it is feasible, both in theory and in practice, for a candidate who attains the support of a majority of voters nationwide to nonetheless lose the election.[98]

This procedure for selecting a president induces unique strategic considerations for presidential candidates. Since electoral victory is accomplished by winning majority of votes in a sufficient collection of states, candidates have the incentive to focus their efforts only on the states in which they have a reasonable chance of winning.The United States has strong and predictable regional voting patterns, based on the ideological and demographic composition of each state, and two-party system has tended to become more strongly entrenched in recent decades. Because of regional party allegiances, presidential election outcomes in a number of states are effectively foregone conclusions prior to the beginning of an election campaign, irrespective of the unique traits associated with the individual candidates.

Often, the outcome of a presidential election hinges on the results of winner-take-all elections in a small subset of states, generally referred to as "battleground" or "swing" states. In recent election cycles, the most hotly contested states have included a collection of states in the Midwest (such as Ohio, Michigan, Wisconsin, Iowa, and Pennsylvania), a few states in the Mountain West (such as Nevada, Colorado, and more recently, Arizona), and a few states in the Southeast (such as Florida, Virginia, North Carolina, and more

[95]The senate has 100 members (two elected officials from each of the 50 states) and the House of Representatives has 435 members, so the 50 states allocate a total of 535 electors. Residents of Washington, DC, are not residents of any of the 50 states. Washington, DC is granted 3 electors, despite not having voting representatives in the House or Senate.

[96]The even number of electors means there can be a tie between two candidates. There can also be cases in which no candidate garners 270 or more electoral votes. In cases where no candidate receives a majority, the election goes to the House of Representatives, which votes (in coalitions made up of each individual state delegation) for the President.

[97]As of the writing of this paper, Maine and Nebraska are the only two states which do not allocate all electoral votes to the winner of the statewide popular vote. These states, instead, allocate electors based on the winner of majority votes in each of their congressional districts.

[98]Recent elections in which the candidate who lost the nationwide popular vote nonetheless won the presidency include 2000 and 2016.

recently, Georgia). States located in the Northeast and on the West Coast tend to reliably vote for the Democratic Party candidate, while states in the Deep South, the prairie states (e.g. Kansas and the Dakotas) and a number of the inland western states (e.g. Texas, Oklahoma, Wyoming, and Idaho), tend to reliably vote for Republicans.

The incentives created by this electoral system lead, at least anecdotally, to various observable patterns. For example, candidates spend the majority of their time campaigning in these battleground states. Moreover, the majority of expenditures on local advertisement, both by the campaigns themselves and by other politically interested groups target these closely contested states.[99] For example, in the leadup to the 2020 presidential election, according to NPR, combined advertisement spending by the two major candidates was highest in Florida, Pennsylvania, and Michigan, where the two campaigns spent $258 million, $196 million, and $120 million respectively. Each of these three states was considered a battleground state, and each of the top 12 states with the highest campaign advertisement spending was considered a battleground state by at least some observers. Outside of these battleground states, no state attracted campaign ad spending exceeding $5 million.[100] Given these patterns, it seems plausible that a sitting president may attempt to curry favor with voters in these battleground states by directing federal funds to these areas. Since voters tend to support an incumbent politician when they feel that the economy is doing well, if targeted federal spending improves local economic outcomes, it is likely not crucial that voters are able to directly trace their economic welfare to government spening policies.

While the notion that federal spending may flow to politically important regions seems intuitive given the structure of United States presidential elections, this is not the only way that federal spending could follow the political incentives of the president. It may also be the case that federal expenditures are used to reward a president's most loyal supporters.

For example, McCubbins and Cox (1986) propose that the spending priorities of political

[99] Examples of groups outside of the individual campaigns which spend heavily on elections include the political party apparatuses (e.g. the Democratic Party or Republican Party), and outside interest groups like Political Action Committees (PACs and so-called Super-PACs).

[100] The states with the highest levels of ad spending include Florida, Pennsylvania, Michigan, North Carolina, Wisconsin, Arizona, Georgia, Nevada, Minnesota, Ohio, Iowa, and New Hampshire.

parties may be crucial to the way that these parties win supporters and form coalitions in the first place. In their model, groups of people with strongly held beliefs on particular issues form coalitions with likeminded individuals. In practice, we can think of these coalitions as groups like rural voters, labor groups, and voters with shared religious convictions, that share are strongly united on a subset of issues. Political candidates and parties then compete for the support of various factions by proposing policies, including spending patterns, which are favored by these factions. In this framework, in order for presidents major political parties to maintain the stability of their core groups of voters, they must adopt spending policies that raise the utility of voters within their coalition. While in practice, this does not require spending to be directed to the places these voters live, and could involve the pursuit of national projects that these groups prioritize, it is easy to imagine that various groups would favor policies which direct expenditures to their communities. An anecdotal example of a policy meant to shore-up the long-term stability of a president's coalition might include, for example, the high payouts of subsidies to farmers that took place during the Trump administration's trade war with China in 2019.[101] Such an action could be interpreted through the lens targeting rural voters who form part of the base of the Republican party. Much like the adoption of other types of policies (e.g. immigration or social policies) popular with a party's most loyal supporters, directing federal outlays toward a party's base can be seen as a tool for generating enthusiasm among a party's vocal proponents, and building support among those most likely to donate, volunteer, or fundraise on behalf of a candidate's campaign.

As compared to the battleground state hypothesis, the notion that presidents would want to reward their most fervent supporters may not require the same degree of geographic targeting by the president's party. For example, if targeted spending is primarily a tool for attracting small donations from loyal supporters, there would appear to be little additional incentive to target supporters in a state in which the incumbent president is favored or highly competitive as compared to a state in which a president is less competitive.

[101] See, for example: https://www.politico.com/news/2020/07/14/donald-trump-coronavirus-farmer-bailouts-359932

Nonetheless, there are a number of reasons why a presidential candidate seeking to reward loyal voters would seek to send funds, in particular, to states in which they garner significant support. First, to the extent that a sitting president is seeking to form or maintain a viable winning coalition for his party, such a coalition must comprise of voters that could reasonably be expected to generate electoral votes for the president's party. If a particular state does not contain a large enough base of supporters of the president's party for that state to realistically vote for that party in a presidential election, then targeting the president's supporters in that state is likely a futile exercise.

Next, federal funds are often granted at the state level, to agencies within states, or to local communities. As such, once funds are granted, state and local officials often have a degree of discretion with respect to the use of funds. It stands to reason then, that in order to ensure that funds are spent in accordance with the president's priorities, it is likely a better strategy to direct funds to states where voters and state officials share the president's ideological and party affiliations.

3. Data and Empirical Approach

3.1 Data and Summary Statistics

In our analysis, we focus on the two types of federal spending: government contracts (procurements) and federal grants. For both types of federal spending, we obtain information at an annual frequency, broken out by the geographic region in which spending takes place. The most granular geographic units of observation for which we can obtain reliable spending estimates for a relatively long time-series, across both categories of spending, are U.S. counties. As such, our empirical tests are conducted at the county level.

We consider procurements and grants separately in our analysis, since these forms of spending are appropriated for very different purposes. Procurements are federal government contracts with private firms. Procurement contracts are made for the provision of goods

and services to the federal government and can include items like defense equipment or medical devices for Veterans Administration (VA) hospitals, among a large list of other items. We obtain procurement data from USA Spending, the official US government-run database tracking federal outlays across a number of categories. This database presents information on a variety of characteristics of government contracts awarded to firms. Our procurement data covers the period from 2000-2020.[102] The more recent procurement data (beginning in 2010) are expressed at a town/city level and are linked to individual firms for each procurement contract. We use a comprehensive publicly available file, which lists the names of U.S. cities and their corresponding states and counties, to generate county-level procurement data.[103] We aggregate award amounts at the county level across firms to get the total value of procurement awards for each county. We consider only awards above a threshold amount of $100,000 to ensure the reliability of the data.

Federal grants consist of awards appropriated by Congress and allocated to specific projects. Grants are often awarded by federal agencies for projects that fall under the scope of their authorities. For example, the department of Housing and Urban Development may provide grants to local government for the construction of affordable housing or local housing assistance. The U.S. Department of Agriculture (USDA) funds programs for family farmers, and the Environmental Protection Agency may grant funds to local governments for local clean-up projects, or removals of lead pipes and clean water programs. Federal grants often go to state governments, local authorities, or state-level agencies nested within a state government.

The data on federal grants is obtained via two sources. For earlier years, we obtain county level federal grant spending from the Consolidated Federal Funding Reports (CFFR), published annually by the U.S. Census Bureau. We scrape text from the scanned PDFs of these files on the Census Bureau website. CFFR reports are available to us for the period from 1996-2010, after which the CFFR was discontinued. We extend our sample through

[102] It should be noted that the data up to 2008 follows a slightly different format from the later data, as it was retrieved under a previous version of the federal spending website, which is no longer available to the public.

[103] This file can be found, and accessed by the public, at: https://simplemaps.com/data/us-cities, which the authors last accessed in March of 2019.

more recent years by using grants data from USA Spending. We are thus able to generate a sample of federal grants covering the period from 1992-2020. The projects that qualify as federal grants do not necessarily line up across our two data sources. However, our tests rely on comparisons of counties made within a single year, so the effects of these time inconsistencies on our ultimate results are minimized. We also carry out tests on separate subsamples (i.e., 1992-2010, and 2010-2020) within the larger sample of federal grants to ensure the robustness of our results.

We combine our information on county-level spending with county-level demographic and economic information from the U.S. Census and other sources. We obtain time-series information on county populations from the intercensal population estimates from the NBER (prior to 2010) and from the U.S. Census Bureau. We also obtain the land area (in square miles) of U.S. counties, also from the U.S. Census, which we use to calculate population densities of each county (a proxy for the urban-rural makeup and remoteness of a county). We calculate average wages of employed residents using data from the County Business Patterns (CBP) database of the Census Bureau, and also used this database to generate estimates of the total number of employed workers in each county. We use information on the racial and ethnic composition of counties from the U.S. Census and the American Community Survey (ACS). While we estimate a number of regressions using annually updated information on county race/ethnic composition, for smaller counties we are only able to obtain information updated on a decennial basis. Table 3.1 displays summary statistics, at the county level, of the main variables that we include in our analyses, with county-level grant and procurements displayed in the final rows of the table. Grants, procurements, population, population density, and average wages are displayed as logs, as this is the form in which these variables enter into our regressions. The number of non-missing observations is substantially smaller for procurements than for grants and other variables, both because of the shorter coverage of the sample, and because a number of smaller counties have very few procurement contracts, and are dropped from the analysis. In contrast, almost all counties have non-missing information on federal grants.

Our empirical tests require us to have knowledge of the margin of victory of presidential candidates on a state-by-state basis for each presidential election, and the number of Electoral College votes allocated to each state for each election covered by our sample. The U.S. Census provides a spreadsheet on historical apportionment data for the house of representatives, which we use to calculate electoral votes across each state, and we use a number of online sources to compile margins of victory by state.[104]

3.2 Empirical Design

To test our hypotheses empirically, we begin by assuming that the president's information set is refreshed after each election. That is, after an election, a president can observe the most recent state-level election returns and, based on that information, form new expectations about how states are likely to vote in the future. We assume that the president's best forecast of battleground states for the subsequent presidential election is the set of battleground states from the previous election. Thus, we ask whether, after a presidential election, battleground states or states with large vote shares for the president's party, are likely to see large federal expenditures, as compared to other states. It is likely that a president's information set is larger than simply the set of closely contested states from the most recent elections; subsequent opinion polls and midterm election results are also likely to further refine the set of tightly contested states. However, we view it as a reasonable assumption that president's set of assumed battleground states is likely to closely mirror the set of tightly contested states from the previous election.

Our empirical design requires that we identify states as belonging to a set of battleground states, a set of states that solidly supports the incumbent president, and a set of states that supports the main opposition political party (i.e. whichever party, among the Democrats or Republicans, the sitting president does not belong to). To identify the set of battleground states, we begin by identifying the so-called "tipping point state" in each election. To do this, we begin by ordering each state, and the District of Columbia, according to the winning

[104] Including 270towin.com and Wikipedia, verified by news sources (e.g. the New York Times).

party's margin of victory in the most recent presidential election. Beginning with the state with the president's largest margin of victory, we tally the cumulative Electoral College votes earned by the president across all states in which the president amassed electoral votes. We label the state which generates the 270th electoral vote (the minimum threshold needed for victory) for the winning candidate as the tipping point state. Intuitively, this state is the state that just pushes the president over the threshold needed to win the election in a given year. Assuming the relative ordering by state of both parties' support among voters remained constant across elections, winning this tipping point state would be crucial for the winning candidate of either party. Since this state is theoretically crucial to both parties' chances at earning a presidential victory, we thus use it as our anchor in defining our set of battleground states. We define the larger set of battleground states in any election as the tipping point, and ten states surrounding the tipping point state, when sorted by the winning candidate's margin of victory. We include the five states with margins of victory larger than the margin in the tipping point state (in favor of the winning candidate), and five states with margins of victory smaller than the tipping point state. In Table 3.2, we illustrate this algorithm in tabular form for the 2000 election won by George W. Bush. We show states sorted by the Bush margin of victory (with negative values in the margin of victory column for states won by the losing candidate, Al Gore) and depict the set of states coded as battlegrounds for the years following that presidential election.

The remainder of states outside of the battleground state are states that we assume are not closely contested by the candidates in a given election. Those states are divided into states solidly supporting the president's political party, and states outside of the sitting president's winning coalition. States won by the sitting president but outside of the band of closely contested battleground states will be used as a treatment for testing whether federal funds are directed toward the president's core supporters.

An alternative to using this tipping point state approach could consist of merely choosing close states, with relatively small margins of victory (in absolute value terms) as the set of battleground states. The reason why this approach is inferior to the tipping point state

approach is that in an election in which the winning president wins in a relative landslide, the set of states with small absolute margins of victory could consist of states that were not particularly important to the victorious president's ability to win an election. For example, in 2008, the winning candidate, Barack Obama, won the state of North Carolina by .3 percentage points, making it one of the closest states, as ranked by margin of victory. However, President Obama would have still won the election even if he lost North Carolina, and would have likewise won even if he lost five other states where he achieved stronger support than he did in North Carolina, suggesting that this state was not particularly crucial to his ultimate ability to win the presidency.

After labelling each state according to its political affiliation and margin of victory in a given election, we identify state-pairs to test our main hypotheses. State pairs consist of neighboring states with differing political affiliations. To test the hypothesis that federal funds are directed toward a president's core supporters, we look at bordering states where one state within the pair is labeled as a solidly Republican state, while the other state within the pair is labeled as solidly Democrat. We then compare spending (either via procurement contracts or federal grants) across the border counties of these state pairs. If funds are directed in a way that seeks to channel funds to "loyal" states within the president's base, then we should see sharply higher spending in counties located within the state that voted for the party affiliated with the winning candidate. In regressions where we test the hypothesis that federal funds are awarded to states that vote strongly in favor of the winning candidate, we label states with large margins of victory for the winner of the election as "treated" states, while paired states that vote for the other party are untreated, or "control" states.

Similarly, to test the hypothesis that federal spending targets battleground states with closely contested elections, we compile state pairs where one state within the pair is a battleground state and the other state is strongly affiliated with one of the political parties (i.e. sees relatively lopsided voting margins). We analyze spending in border counties within these state pairs and look for sharp increases in spending as we cross into the battleground state from its paired state. In regressions where we test this hypothesis, we label closely

contested battleground states as belonging to the treatment group.

In Figure 3.1 we show how our sets of state pairs look in the years following the 2000 election. We first show the election map color-coded according to whether each state is a battleground, or whether it belongs solidly to one or the other political party. States shaded in gray are the battleground states, while states shaded in red and blue are Republican-won and Democrat-won states, respectively, with large absolute margins of victory. In panels B and C, we show the sets of state pairs that emerge in tests of our main hypotheses. Panel B shows state pairs featuring bordering states with solidly Republican and solidly Democrat electorates, which we use to test the hypothesis that the federal funds award strong supporters of the winning candidate. Panel C shows state pairs consisting of a battleground state and state that voted for the Democratic party candidate by a wide margin. In our first set of tests we will also include state pairs featuring battleground states that border states with large majorities in favor of the winning candidate (i.e. the Republican party in this example), however, in later tests we will consider heterogeneous treatment effects in comparing battlegrounds with states that strongly support the winning candidate and those that support the losing candidate.

After forming our collections of state pairs, we test our hypotheses by estimating regressions in which we compare government spending outcomes across the pairs of states that appear in our sample. We estimate spending regressions at the county level, using counties located within one of the states that appears in the set of bordering state pairs. We refine our sample of counties by only including counties which are located near the borders shared between a set of paired states. We focus on these border counties because we believe that comparing counties located very close together reduces the potential for omitted variable bias relative to regressions which include the full set of counties across all states. This is because counties located next to each other in neighboring states tend to share similar demographic traits and are subject to similar regional economic conditions and common shocks. Thus, federal funds meant to target particular demographic groups (e.g. the elderly) or occupations (e.g. farmers) would be likely to be apportioned similarly

across neighboring counties to the extent that these funds were designated for purposes unrelated to the political tendencies of the states in which these counties reside. That is, an identifying assumption is that for collections of counties near borders, differences in the proclivities of these counties to receive funds are as good as random in the absence of the political motivations of those directing federal spending.

In Figure 3.1 panel D we show a graphical depiction of our sample of counties for a test of the hypothesis that federal funds are directed toward residents of battleground states. Our sample consists of the set of state pairs highlighted in panel B of Figure 3.1. We show the set of counties contained in a specification in which we select an 80 mile border distance as the cut-off for sample inclusion. That is, in the example of panel D, our sample includes only counties located within 80 miles of the borders shared between adjacent states.[105]

In our state borders-based empirical tests, we estimate regressions of the form:

$$y_{ijbt} = \alpha_t + \gamma_b + \beta * Treat_{jt} + \delta * X_{ijbt} + \epsilon_{ijbt} \tag{1}$$

where the dependent variable, y, is some form of government spending, including either federal grants, or the aggregate dollar volume of county-level procurement contracts. The indices i, j, b, and t denote the county, state, border (i.e. the set of paired states, whose border is being referenced), and year. The *Treat* variable is an indicator variable that takes a value of one if a state is a treated state relative to the hypothesis being tested. The *Treat* variable has only j and t subscripts because treatment is assigned at the state level for a given election year. A state can appear in both the treatment and control groups after different election years. The vector X represents the set of control variables in our regressions. The vector X includes county-level characteristics: population, population density, racial composition (% of Black residents and Hispanic residents), average income, and the employment/population ratio of a county. The key coefficient in these regressions is

[105] In one state pair in this sample, the Michigan-Wisconsin pair, part of the border between these two states is comprised of Lake Michigan, a natural water-induced boundary. For robustness, we will estimate versions of our main specifications in which we drop counties separated by water borders, and we will also estimate specifications in which we drop large urban counties from the sample in order to attain better sample balance.

the value of β, the coefficient on the treatment variable. A positive value of this coefficient suggests that residents of the treated state tend to receive more federal funds than residents of untreated states. Depending on the regression specification a positive value of β would suggest that federal funds either target battleground states, or states in which the winning candidate in the previous election received substantial support (i.e. a state that aligns solidly with the sitting president's party). The variable y enters the regressions in log form, suggesting that the value of β reveals the expected percentage-increase in spending that would be attained when moving from a county located right next to the border in an untreated state to a county located just across the border in a treated state.

Since our panel covers a fairly long time period, particularly in our estimation of regressions with federal grants as the outcome variable, we are able to generate a fairly large set of state pairs across the full sample. Nonetheless, one concern is that there may be several state pairs which enter the sample on a recurring basis and that there might be persistent state-level factors that lead to higher federal spending in one state over another, which are unrelated to the political affiliations of voters in these states. Holmes (1998) shows that there tends to be higher manufacturing activity in right-to-work states as compared with closed shop (stronger union) states and shows that manufacturers are more likely to locate plants in counties within right-to-work states, as compared to neighboring counties just across the border. Given these manufacturing patterns, we may be concerned that since our set of treated and untreated states maps neatly into state-level ideology, our measure of treatment might merely capture differentials in manufacturing and other business activity across states, rather than the political motives of policymakers. As such, we will also estimate additional regressions, in which we include state fixed-effects to absorb state-level policies and other persistent differences between counties that operate at the state level. We estimate regressions of the form:

$$y_{ijbt} = \alpha_t + \gamma_b + \delta_j + \beta * Treat_{jt} + \delta * X_{ijbt} + \epsilon_{ijbt} \tag{2}$$

where the key difference relative to equation (2) is the fixed-effect δ_j. When estimating equation (2), we take two different approaches. When testing the hypothesis that federal funds are directed toward states that loyally support the winning presidential candidate, we refine our sample by focusing on state pairs which switch treatment status over the course of the sample. This switching of treatment status can happen when we observe two neighboring states where one state reliably supports Republican candidates and the other reliably supports Democrats over an extended period. In such state pairs, when we transition from a Republican president to Democrat (or vice versa), the state in the pair that had previously been untreated will flip to being a treated state. Our fixed effects specification will then tell us whether the average differential in federal spending between these sets of states shifts, after a presidential election, in favor of whichever state supported the winning candidate.

In federal grants regressions, our approach will be slightly different. In general, we do not observe many state pairs which fully switch treatment status over the course of the sample. That is, we do not observe many cases in which, across a number of elections, a battleground state shifts to voting solidly in favor of a particular pary, while its neighboring state makes the reverse transition, from favoring a particular party to being a battleground.[106] Instead, we consider all cases in which a state transitions from being a battleground state to a state which strongly supports a particular party, or vice versa. We then maintain all pairs of bordering states that include a state that makes this transition at some point during the sample. We then estimate equation (2) on this sample. This specification can be thought of as a difference-in-differences setting in combination with a border counties approach. That is, we ask whether the differential in spending between counties in two bordering states grows larger after one of these states becomes an electoral battleground.

It should be noted, that our two hypotheses, that federal funds are directed toward battleground states, and that they are directed toward states which support the president, and not completely mutually exclusive. That is, there can be some rank ordering between

[106] Two exceptions are Arizona-New Mexico, and Ohio-Pennsylvania.

battleground states, states in which a president has a strong base of support, and states which tend to vote against the sitting president's party. It is plausible that a president prefers both battlegrounds, and states that align with his party, over states which vote for the other party. However, ultimately, it is of interest to determine which of these effects dominates. That is, do presidents care more about battlegrounds than they do about states in which they are heavily supported? To answer these questions, we will also estimate regressions where we consider all sets of state pairs that test either hypothesis within a single regression. We estimate regressions of the form:

$$y_{ijbt} = \alpha_t + \gamma_b + \delta_j + \beta_1 * TreatLoyal_{jt} + \beta_2 * TreatBattleground_{jt} + \delta * X_{ijbt} + \epsilon_{ijbt} \quad (3)$$

where the variable *TreatLoyal* takes a value of one if county i is located within a state where the winning president in the previous election won by a substanial margin, while *TreatBattleground* takes a value of one if county i is located within a presidential battleground state.

4. Results

4.1 Does Federal Spending Target States where the President is Strongly Supported?

We first look to test the hypothesis that federal funds are used, in part, to reward a sitting president's loyal supporters. To test this hypothesis, we begin by estimating county-level government spending regressions of the form shown in equation (1). We use a sample of paired neighboring states where each state pair consists of a treated state, where the current president won by a substantial margin in the previous election, and an untreated state where the president lost by a relatively large margin. We estimate these county-level regressions by considering only sets of nearby "border" counties in each state pair which are likely to have

similar geographic and demographic features which might lead them to attract relatively similar federal expenditures in the absence of a political motive on the part of policymakers.

Table 3.3 panel A shows the results of these regressions. The first four columns of results display coefficients from regressions where the dependent variable is federal procurement expenditures, while the last four columns display results of federal grants regressions. We estimate our model using four different bandwidths, or cutoffs for inclusion of a county in the sample. In the left-most column we show results of the least restrictive specification in which we include all counties located within 100 miles of the border shared by the treated and untreated state. We also show results from regressions at bandwidths of 80 miles, 50 miles, and 30 miles, with distances to borders calculated using population centroids of each county. The 30 mile bandwidth is quite restrictive, as a number of counties located right on the border of their state are excluded from the sample in this specification.[107] Results in the first four columns of panel A suggest a strong positive treatment effect for procurement contracts, consistent with the hypothesis that government contracts target states where a president has a substantial base of supporters. The treatment effect in the 100 mile bandwidth regression for procurements is .358, suggesting that a county located within a state where the sitting president enjoys strong support sees about 36% stronger federal procurement spending than a similar county in a state that tends to vote for the opposing party. The treatment effect tends to get stronger at smaller bandwidths. The estimated effect rises monotonically from the 100 mile regressions to the 30 mile regressions, peaking at a value of .613 in the 30 mile specification. This suggests that the results are driven by the counties located closest to their state border. Treatment effects in the procurement regressions are large and statistically significant at the 1% level in all specifications.

Treatment effects are also consistently positive and large in the federal grants regressions. In a number of ways results with the federal grants data may be a better test of the extent to which federal spending targets particular jurisdictions, as federal grants cover a much broader arrary of government spending. Across specifications, treatment coefficients in the

[107] This can happen when the majority of a county's population lives more than 30 miles away from the border, despite the fact that the county itself shares a border with a county in the adjacent state.

federal grants regressions range from .183-.309 suggesting that counties in states that support the winner of a presidential election see between 18% and 31% stronger grant spending in the four years that follow the election, as compared to counties in states that the president loses by a wide margin. These coefficients vary in magnitude across bandwidths, with the largest results seen in the 80 mile and 50 mile bandwidth regressions. These results suggest that federal grants may be an important way in which a winning presidential candidate rewards his core voters.

While these regression results provide evidence suggesting that federal grants are directed toward states with a political party affiliation that matches the affiliation of the president, it may also be the case that state-level policies or other factors are driving these results. This should be mitigated to some degree because many of the state pairs that appear in the sample switch their treatment status over the course of the period over which we estimate our regressions, as the party holding the presidency changes. For example, the state pair Illinois-Indiana shows up in our model in several elections. During the Bush presidency of 2001-2008, Indiana was a treated state because of the relatively sizable margins of victory it delivered to the Republican party in the 2000 and 2004 presidential elections, while Illinois would be an untreated state because it voted heavily for the Democratic party candidate in both elections. However, once a Democrat was elected to the White House, this treatment assignment changed. While Indiana was a battleground state in the 2008 election, by 2012 it was again a solidly Republican state, and in these years, with a Democrat in the White House, Illinois became a treated state, according to our algorithm, while Indiana was its untreated paired state.

This time-variation should help ameliorate the concerns that our results stem from state-level policies or other unobserved state-level effects. However, not all state pairs show up in the sample with both treatment statuses, and nothing guarantees that even those state pairs which appear in the sample with both sets of treatment statuses will appear an equal number of times under both assignments. For example, in the procurement sample, the Illinois-Indiana state pair appears four times with Indiana assigned to the treatment group

on three of those four occasions (after both Bush victories and the Trump presidency). Since the treatment effect variable captures only the average effect, if Indiana had persistently higher federal expenditures than Illinois for reasons having to do with aggregate state-level differences (but unrelated to the paresident's party affiliation) the fact that Indiana appears in the sample more frequently than Illinois would tend to bias our results in favor of finding a treatment effect. In order to deal with these issues, we focus on only the set of states that appears in the sample under both treatment arrangement (i.e. each state in the state pair is treated on at least one occasion) and estimate regressions with state fixed-effects according to equation (2).

The results of this approach are shown in panel B of Table 3.3. In both the procurements and grants regressions, treatment effects are somewhat smaller than in specifications without state fixed-effects. However, results remain fairly consistent across panels A and B. When state fixed-effects are added to the procurement regressions, the estimated treatment effect is roughly cut in half. The coefficient in the 100 mile regression falls from .358 to .179, suggesting that states which tend to be in the treatment sample more often tend to have time-invariant features which attract stronger federal procurement spending, irrespective of the party of the president. The drop in magnitude is similar across the other bandwidths. However, in each of these specifications, the estimated treatment effect remains positive and statistically significant. Moreover, the magnitudes of the coefficients continue to suggest economically sizable shifts in spending following changes in the presidency. In a state like Illinois, which votes reliably for the Democratic candidate, a favorable election result could mean an 18% increase in federal procurement spending for an average county in the state, suggesting that outcome of a presidential election is rather meaningful for companies with operations located within the state, and for workers employed by these firms.

As compared to the results of the procurement regressions, treatment effects in the federal grants specifications remain quite similar across specifications with and without state fixed-effects. Treatment coefficients across these regressions suggest that other factors equal, federal grant spending is between 16% and 25% stronger in counties located within states

that voted decisively for the winning candidate in the most recent election, as compared to a county located within a state that voted for the losing candidate.

4.2 Does Federal Spending Favor Battleground States?

We next seek to test the hypothesis that federal spending is directed to closely contested battleground states. A president seeking to use federal expenditures to improve the prospects of winning the subsequent election might seek to direct funds to battleground states and generate favorable economic conditions in these important states. We test this hypothesis by comparing federal spending outcomes in battleground states to spending in states which tend to have comparatively lopsided outcomes.

The set of battleground states is rather persistent across the sample. States like Florida, Pennsylvania, Ohio, and Michigan all appear in the sample of battleground states in almost all years. There are almost no instances of state pairs which have switched over time, in the sense that a treated state has part of the sample of control states (with relatively one-sided election margins) while its neighboring state has entered the set of treated states. Since persistent state-level effects are a concern, we proceed by estimating versions of equation (2) which follow a difference-in-differences approach described in section 3. That is, we consider all states which transition from being a battleground to attaining a relatively clear party allegiance over time, or vice versa. Such states include Virginia, Arizona, New Hampshire, Ohio, Georgia, Missouri, Iowa, Washington, and North Carolina, among others. We then consider all sets of bordering state pairs which include exactly one of these battleground states, and we ask whether the differential in federal spending between these sets of neighboring states becomes larger after one of these states becomes a closely contested battleground (or smaller after a state loses battleground status).

Table 3.4 shows the results of this analysis on our panels of federal government procurement contracts and federal grant spending. The first four columns show results of the procurement regressions. In these regressions, estimated treatment effects are positive

across all specifications, consistent with the hypothesis the procurements are disproportionately directed to battleground states. The treatment coefficient in the 100 mile bandwidth regression is .091. This suggests that in the years after an election in which a state newly becomes a presidential battleground, the differential in procurement spending between the battleground state and its non-battleground neighbors increases by roughly 9.1% (i.e. shifts by 9.1% in favor of the battleground state) for a typical county. The estimated treatment effect is largest in the regression conducted at a bandwidth of 80 miles, where it attains a value of .158, before shrinking in the 50 and 30 mile regressions.

The last four columns of Table 3.4 display results for the federal grants regressions. On the whole, the results for federal grants are quite similar to the results of the procurement regressions. Estimated treatment effects are again positive across all specifications, ranging from a value of .041 to .149. The estimated treatment effect is largest in the 30 mile bandwidth regression, which includes a small subset of counties whose population lives very close to state borders. In this specification, our treatment effect would suggest that a county located within a closely contested battleground state would see, on average, roughly 15% stronger federal grant spending than a county in a state that tends to strongly support one or the other political party.

We interpret the consistently positive value of our estimated treatment effects across dependent variables and across subsets of counties as indicative that the political importance of a state to upcoming presidential elections has some bearing on that state's ability to attract federal funds. While the set of battleground states tends to change only marginally from one election to the next, our relatively long sample for federal grants can ameliorate concerns that our results are driven by strong spending differentials across a relatively small number of state-pairs. With only a small number of exceptions (e.g. Montana, Delaware, and Oklahoma) almost all of the continental 48 states enter the sample (as either a treated state, untreated paired state, or both) at some point. Moreover, our fixed-effects specifications control for time invariant policy differences at the state level, though there may be some concern that in a relatively long sample our specifications may omit some

important policy changes that occur contemporaneously with shifts in presidential voting patterns.

4.3 Comparing Magnitudes and Timing

We have shown preliminary evidence in support of both of our main hypotheses. We have shown that federal spending appears to grow more strongly in states that supported the winning candidate in the most recent presidential election, as compared to states that voted for the losing candidate. We have also shown evidence suggesting that swing states, on average, tend to attract more federal spending than states with less closely contested elections. However, the results that we've shown thus far do not clarify which of these effects wins out. That is, when choosing to allocate a marginal dollar of federal spending to a particular project, we are interested in determining whether a battleground state, or a state that votes strongly in favor of the president's party is likely to see a greater portion of those marginal funds.

It is tempting to merely compare the magnitudes of the treatment effects from the previous two sections. However, the tests of these sections were designed to establish and cleanly identify the existence of an effect of a state's political significance on its tendency to receive federal funds, and not necessarily designed to generate magnitudes that are directly comparable. In particular, the control groups across these two sets of tests was somewhat different. In our first set of tests we compared states that voted strongly in favor of the president's party with those voting against that party, while our second set of tests compared battleground states to *all* states with relatively one-sided election results, irrespective of whether they voted with or against the president. We know then, that on average, swing states tend to attract more funding than comparable states with lopsided elections. But the results of the first section suggest that among states with lopsided elections, it is important to know whether a state voted for or against the current occupant of the White House.

In order to gauge the relative magnitudes of our two effects, we estimate regressions that

take the form of equation (3) where we include separate treatment variables for states that voted by a large margin in favor of the president (denoted ast *TreatLoyal* in equation (3)), and for battleground states (denoted as *TreatBattleground*). In our regressions, we include all states which at some point during our sample were either a battleground state, or a state voting strongly for the winning candidate, and which bordered a state with a different treatment status. This includes virtually all states, and we estimate our regressions across all sets of borders between states which, at some point in the sample, have differently-coded values of either of the two treatment variables.[108] State pairs like North and South Dakota, which have consistently been safe Republican states, and Vermont and Massachusetts, which have consistently been safe Democratic states, are excluded from the sample, however, the majority of bordering states have had a different assignment to treatment groups at some point during the sample period.

An alternative approach to comparing the size of the rewards accruing to battleground states as compared to states which loyally support the winning candidate would involve estimating our regressions only on pairs of states that include a battleground and a state which voted strongly in favor of the winning candidate in the previous election. We favor our approach however, since as previously mentioned, the set of battleground states changes only gradually over time and so we seek a sample with maximum variation in the assignment of states to treatment or control-state groups.

We display the results of this analysis in panel A of Table 3.5. As usual, we group procurement results in the first four columns, followed by results from the longer federal grants sample. Our results generally suggest that the tendency of federal spending to gravitate toward states which voted strongly in favor of the winner of the most recent presidential election is stronger than the tendency to for federal grants to target battleground states. In procurements specifications, the *TreatLoyal* coefficient ranges from a value of .188 to .449 and is larger than the corresponding *TreatBattleground* coefficient in all specifications. The

[108] That is, a border between two states can enter our sample if it connects a state pair that consists of any of these three groups at some point during the sample period: a battleground state and a state that favors the the sitting president by a large voting margin; a battleground state and a state that voted for the losing candidate by a large margin; a state favoring the winning candidate by a large maring and the losing candidate by a large margin.

difference in magnitudes between the *TreatLoyal* and *TreatBattleground* coefficients in the procurement regressions ranges from .088 at the 100 mile bandwidth to .346 at 30 miles, suggesting that states that loyally support the winning candidate in a presidential election tend to see 9%-35% higher federal procurement expenditures in a typical county. A similar pattern is evident in the federal grants regressions. Our battleground state treatment effect ranges from .041 to .153, while the *TreatLoyal* coefficient ranges from .134-.291. Of the specifications we estimate on federal grants, only one bandwidth (30 miles) implies a larger treatment effect for battleground states than for states which solidly support the winning candidate.

In a final piece of the analysis, we ask whether there is time variation in the strength of our treatment effects. In particular, we ask how the sizes of our treatment effects change in the year of a presidential election. We conjecture that it may be particularly important to channel funds to battleground states during an election year, since undecided voters often make up their minds on which candidate to support only in the months leading up to an election. In contrast, the theoretical intuition underpinning the hypothesis that presidents tend to reward states in which they garner substantial support does not, in our view, depend as crucially on the timing of spending relative to an election. While a presidents may need to reward their supporters in order to keep campaign promises and ensure the long-term viability of their party, it is likely that party loyalties are likely somewhat sticky in the short-run, suggesting that over a single year in the lead-up to an election, the president may be able to afford to direct some funds away from his most ardent supporters in favor of swing voters in battleground states. Moreover, if the president seeks to reward supporters in order to generate enthusiasm within his party and build networks of donors or campaign volunteers, such efforts likely take some time to come to fruition, and are thus likely to be undertaken in advance of an election.

To test whether the size of our treatment effects vary in election years as compared to other times, we estimate an additional version of equation (3) in which we incorporate terms in our regressions which interact our treatment variables with election year dummy variables

(i.e. indicator variables which assume a value of one in an election year, and zero otherwise). We can interpret the coefficients on these variables as indicating how much stronger our treatment effects of each type become in the lead-up to a presidential election. We display the results of these specifications in panel B of Table 3.5, and we display interaction terms beneath each of the original treatment coefficients (denoted by the **ElectionYr.* suffix). Our interaction terms are generally suggestive of the notion that the tendency to reward voters in battleground states grows stronger in the lead-up to an election. This result is stronger for grants than for procurements, though the estimated value of the battleground regression term generally remains positive in all of these regressions (with the exception of a single specification in the 30 mile bandwidth regression for procurements). In the federal grants regressions, our results suggest that spending channeled to counties in battleground states tends to rise by between 3-9.5 percentage points in the year of a presidential election. Across specifications, this implies roughly a doubling of the estimated battleground state effect on federal grant spending. In contrast, these interaction terms tend to be negative in federal grants specifications for the *TreatLoyal* interaction coefficient. It thus appears that the tendency to reward states that strongly support the incumbent president is not strongly time dependent and that this tendency does not grow stronger in the lead-up to elections. This falls in line with existing theory, since the tendency to reward loyal supporters is often described as a long-term approach geared toward ensuring the viability of a president's or party's core coalition, rather than a short-term approach of gleaning additional supporters in the service of marginally improving the odds of a favorable election outcome.

5. Conclusion

After taking office in 2009, President Obama famously quipped, in a discussion with congressional Republicans on economic policy, that "elections have consequences [and] I won." This saying has been repeated by subsequent presidents and other politicians, to suggest

that winning candidates have the right to implement their policy preferences without interferenence from the opposition. In this paper, we have shown that elections indeed have consequences for the federal budget. In particular, we find evidence consistent with the notion that a president seeks to reward his supporters in the wake of a presidential election by directing spending to states which strongly supported his candidacy. We also find evidence suggesting that as compared to states which voted against the winning candidate, federal funds are also directed toward key presidential battleground states. The tendency for funds to target these battleground states grows stronger in the lead-up to a presidential election, when gaining the support of these crucial voting blocks is likely to be most important.

Ours is not the first paper to examine the hypothesis that federal spending follows the incentives of the president's political party. However, our paper is novel in making the use of the unique mechnism governing presidential elections in the United States in testing these prevailing hypotheses. We note that to the extent that a president's main incentive is to get re-elected, or to help elect the next candidate within his political party, the Electoral College mechanism makes it so that the specific set of states in which the president directs these efforts is of critical importance. By focusing on pairs of neighboring states which have different voting tendencies, and zooming in on the border counties of those state pairs, we are able to better identify the effect of a president's political incentives on the allocation of federal outlays. Rather than comparing far-flung counties across states in very different regions, our border counties approach helps ensure that our main comparisons are conducted across relatively similar counties, with substantially similar demographics, and exposure to many of the same regional economic shocks. These comparisons help to diminish potential bias that could emerge if unobserved regional economic or demographic factors influence the propensity of a county to receive federal funds. We also use a long sample and make use, in a number of specifications, in time variation in the ideological loyalties of individual states, as well as the changing identity of the party that wins a presidential election. We are then able to use state fixed-effects, and a difference-in-differences setting, in combination with our borders approach to show that after a presidential election, the differential in spending

between two neighboring states with different political affiliations tends to shift in favor of whichever state supported the winning candidate in that election.

The magnitudes of the effects we find are far from trivial. In a number of specifications, we see treatment effects suggesting that a typical county might expect a roughly 20% increase in the dollar volume of federal grants it receives after supporting the winning candidate, as opposed to the losing candidate, in a particular election. While such a shift might represent only a few hundred thousand dollars in spending in the smallest counties, the largest counties may receive several billions of dollars of federal grants in a typical year.

It is difficult to assess the welfare implications of the spending patterns we uncover in this paper, as we do not directly observe the effect of federal spending on the utility of agents to whom it is directed. Nonetheless, if our results are to be believed, then the structure of the U.S. political system and the protocol for the conduct of presidential elections makes it so that very small changes in political affiliations and voting records can have dramatic effects on the way federal funds are allocated. In some sense, changes in the allocation of federal grants and procurement spending may be just a small example of the observably large policy shifts that can come to pass after a change in the party that holds the White House. Nonetheless, it is tempting to wonder whether, as the national electorate in the United States becomes more bifurcated, whether the economic costs associated with voting for the losing candidate in an election may also become more severe.

References

Bickers, K., and R. M. Stein (2000). "The Congressional pork barrel in a Republican era," *Journal of Politics*, vol. 62(4), pages 1070-1086.

Brogaard, J., M. Denes, and R. Duchin (2015). "Political Connections, Incentives, and Innovation: Evidence from Contract-level Data," Duke University Working Paper.

Campello, M., J. Gao, and Q. Xu, (2019). "Personal taxes and firm skill hiring: evidence from 27 million job postings," Kelley School of Business Research Paper No. 19-35.

Carsey, T. M., and B. Rundquist (1999). "Party and committee in distributive politics: evidence from defense spending," *Journal of Politics*, vol. 61(4), pages 1156-1169.

Carvalho, D. (2014). "The Real Effects of Government-Owned Banks: Evidence from an Emerging Market," *Journal of Finance*, vol. 69(2), pages 577-609.

Cohen, L., J. Coval, and C. Malloy (2011). "Do powerful politicians cause corporate downsizing?" *Journal of Political Economy*, vol. 119(6), pages 1015-1060.

Creary, M. S., S. L. Greer, P. M. Singer, and C. E. Willison (2019). "Quantifying inequities in U.S. federal response to hurricane disaster in Texas and Florida compared to Puerto Rico," *BMJ Global Health*, vol. 4(1), pages 1-6.

Holmes, T. J., (1998). "The effect of state policies on the location of manufacturing: evidence from state borders," *Journal of Political Economy*, vol. 106(4), pages 667-705.

Larcinese, V., L. Rizzo, and C. Testa (2006). "Allocating the U.S. federal budget to the states: the impact of the president," *Journal of Politics*, vol. 69(2), pages 447-456.

Levitt, D. S., and J. J. Snyder (1995). "Political parties and the distribution of federal outlays," *American Journal of Political Science*, vol. 39(4), pages 958-980.

Ljungqvist, A., and M. Smolyansky, (2018). "To cut or not to cut: on the impact of corporate taxes on employment and income," NBER Working Papers 20753.

McCubbins, M. D., and G. W. Cox (1986). "Electoral politics as a redistributive game," *Journal of Politics*, vol. 48(2), pages 370-389.

Moretti, E., and D. Wilson, (2017), "The effect of state taxes on the geographical location of top earners: evidence from star scientists," *American Economic Review*, vol. 107(7), pages 1858-1903.

Pence, K. M., (2006). "Foreclosing on opportunity: state laws and mortgage credit," *Review of Economics and Statistics*, vol. 88(1), pages 177-182.

Wallis, J. J. (1987). "Employment, politics and recovery in the Great Depression," *Review of Economics and Statistics*, vol. 69 (August), pages 516-520.

Tables and Figures

Table 3.1

In this table we present summary statistics of the variables used in our main panel investigating state political affiliations and federal spending using county-level data. We present information on our federal spending variables (federal grants and procurement contracts), in the last two rows of the table, as well as county-level demographic and economic information, which we control for in regressions. We include county-level population, population density, the percentage of Black and Hispanic residents, average wages of employed workers, and the employment/population ratio, among our observable county characteristics. Grant and procurement contract values, as well as population, population density, and wages, are expressed as natural logarithms. In the first column of the table, we express the number of non-missing observations for each variable. We express variable means, first quartile values, median values, third quartile values, 90th percentile thresholds, and standard deviations in the subsequent columns of the table.

Summary Statistics for Key Variables at the County Level							
Variable	Observations	Mean	Quartile 1	Median	Quartile 3	90th Pctl.	St Dev.
Population	75,414	10.24	9.30	10.13	11.08	12.16	1.45
Pop. Density	75,378	3.75	2.82	3.78	4.70	5.91	1.75
Pct. Black	75,356	0.095	0.007	0.026	0.111	0.307	0.147
Pct. Hispanic	75,356	0.083	0.016	0.033	0.082	0.214	0.132
Pct. White	75,356	0.858	0.811	0.928	0.968	0.981	0.164
Wage Avg.	77,281	10.26	10.05	10.27	10.47	10.65	0.312
Employment/Pop.	75,285	0.276	0.185	0.258	0.343	0.437	0.135
Procurements	19,119	10.82	8.65	10.30	12.43	14.70	2.81
Grants	74,343	8.00	6.86	7.95	9.02	10.30	2.04

Table 3.2

In this table, we give an example of the algorithm for classifying states into categories based on their political affiliations in a recent election. In this table we use the 2000 presidential election as an example. We first order each state (and the District of Columbia), in descending order, by the margin of victory of the election's winner in that state. The name of each state, in this order, is shown in the first column of the table. Next, we display the winning candidate's margin of victory (percentage vote share of the election's winning candidate minus the vote share of the losing candidate). In the third column we show the number of Electoral College votes designated to each state. Column four shows the cumulative electoral votes garnered by the winning candidate (Bush) by virtue of winning the state. That is, it shows the value of Electoral College delegates won by Bush from the state listed within the same row of the table, plus the sum of electoral votes won from all states in which he had a higher margin of victory (i.e. states in previous rows of the table). In the final column of the table, we list how each state is classified in our sample in the years that follow the 2000 presidential election (i.e. 2001-2004).

Illustration of Battlground State Coding: Vote Margins in 2000 Presidential Election				
State	Margin of Victory (Rep. - Dem.)	Electoral Votes	Cumulative Electoral Votes (Bush)	Battleground?
Utah	40.49	5	5	Solid Rep.
Wyoming	40.06	3	8	Solid Rep.
Idaho	39.53	4	12	Solid Rep.
Alaska	30.95	3	15	Solid Rep.
Nebraska	28.99	5	20	Solid Rep.
North Dakota	27.6	3	23	Solid Rep.
Montana	25.07	3	26	Solid Rep.
South Dakota	22.73	3	29	Solid Rep.
Oklahoma	21.88	8	37	Solid Rep.
Texas	21.32	32	69	Solid Rep.
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
Arizona	6.29	8	189	Solid Rep.
Arkansas	5.44	6	195	Solid Rep.
Tennessee	3.86	11	206	Battleground
Nevada	3.55	4	210	Battleground
Ohio	3.51	21	231	Battleground
Missouri	3.34	11	242	Battleground
New Hampshire	1.27	4	246	Battleground
Florida	0.01	25	271	Tipping Point
New Mexico	-0.06	5		Battleground
Wisconsin	-0.22	11		Battleground
Iowa	-0.31	7		Battleground
Oregon	-0.44	7		Battleground
Minnesota	-2.40	10		Battleground
Pennsylvania	-4.17	23		Solid Dem.
Maine	-5.11	4		Solid Dem.
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.

Table 3.3

In this table we present results of our tests of the hypothesis that federal spending is channeled to states where a president has a strong base of support. In panel A, we show specifications which follow the form of equation (1), while in panel B our regressions follow the form of equation (2), and we estimate specifications with state fixed-effects. We regress the log of two forms of federal spending on a treatment variable, which takes a value of one in counties located within a state that voted strongly in favor of the winning candidate in the most recent election and zero in states that vote for the losing candidate. We also include county-level demographic and economic characteristics as controls (coefficients on these variables are omitted). In the first four columns of the table, we display estimated treatment effects from regressions which include the log of aggregated county-level procurement contracts as the left-hand side variable. In the last four columns we display treatment effects from regressions where log federal grant spending is the dependent variable. We display treatment effects from regressions estimated at four different bandwidths (or distance cutoffs for inclusion of a county in the sample). We show distance cutoffs of 100, 80, 50, and 30 miles from a state border. Treatment coefficients are displayed in the first row with standard errors shown in parentheses beneath each coefficient. We also display R-squared values from each regression. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Baseline Specifications for Loyal Supporters Hypothesis								
	Procurements				Grants			
Bandwidth (Mi.):	100	80	50	30	100	80	50	30
Treatment Effect	.358***	.412***	.562***	.613***	.183***	.309***	.207***	.192***
	(.078)	(.084)	(.099)	(.109)	(.040)	(.043)	(.039)	(.044)
R-squared	.682	.687	.734	.765	.764	.771	.865	.903
Fixed-effects	Border	Border	Border	Border	Border	Border	Border	Border

Panel B: States which Switch Treatment Status (State Fixed-effects)								
	Procurements				Grants			
Bandwidth (Mi.):	100	80	50	30	100	80	50	30
Treatment Effect	.179***	.199***	.399***	.288***	.181***	.247***	.162***	.156***
	(.059)	(.062)	(.077)	(.088)	(.034)	(.035)	(.037)	(.046)
R-squared	.625	.641	.690	.736	.740	.751	.794	.824
Fixed-effects	Border + State	Border + State	Border + State	Border + State	Border + State	Border + State	Border + State	Border + State

Table 3.4

In this table we present results of our tests of the hypothesis that federal spending is channeled to key battleground states that are crucial to the results of a presidential election. We show specifications which follow the form of equation (2) and estimate regressions which include state fixed-effects. We regress the log of federal procurement spending and grant spending, at the county level, on a treatment variable that takes a value of one in counties located within closely contested battleground states (as determined using state-level voting patterns from the previous election) and zero in states that voted strongly in favor of the losing candidate. We also include county-level demographic and economic characteristics as controls. In the first four columns of the table, we display estimated treatment effects from regressions which include the log of aggregated county-level procurement contracts as the left-hand side variable. In the last four columns we display treatment effects from regressions where log federal grant spending is the dependent variable. We display treatment effects from regressions estimated at bandwidths of 100, 80, 50, and 30 miles from a state border. Treatment coefficients are displayed in the first row with standard errors shown in parentheses beneath each coefficient. We also display R-squared values from each regression. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

	Testing the Battleground States Hypothesis							
	Procurements				Grants			
Bandwidth (Mi.):	100	80	50	30	100	80	50	30
Treatment Effect	.091**	.158***	.083	.051	.128***	.041	.042	.149***
	(.031)	(.044)	(.054)	(.070)	(.024)	(.025)	(.030)	(.037)
R-squared	.620	.600	.641	.655	.788	.737	.762	.804
Fixed-effects	Border + State	Border + State	Border + State	Border + State	Border + State	Border + State	Border + State	Border + State

Table 3.5

In this table, we display results from regressions in which we jointly test the two major hypotheses regarding federal spending that we identify from the literature. We include all pairs of bordering states that include either: (a) state voting strongly in favor of the winning candidate's party in the most recent election bordering a state voting strongly in favor of the losing candidate, or (b) a closely contested battleground state bordering a state with a relatively lopsided margin of victory (for either party). Our estimating equations follow the form of equation (3), with both state and border fixed-effects. Panel A shows the results of these regressions for federal grants and procurement contracts without adding any additional interaction terms. The *Treatment (Loyal)* coefficient shows the additional spending directed toward counties in states voting strongly in favor of the winning candidate as compared to counties which voted heavily in favor of the losing candidate. The *Treatment (Battleground)* coefficient shows the additional spending directed toward counties in battleground states as compared to states with relatively large (in absolute value) margins of victory. In panel B we add additional terms where we interact the two treatment variables with indicators variables that take a value of one in an election year. We show bandwidths of 100, 80, 50 and 30 miles for both the procurement and grants regressions, as in previous exhibits. Standard errors are listed in parentheses beneath each coefficient. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

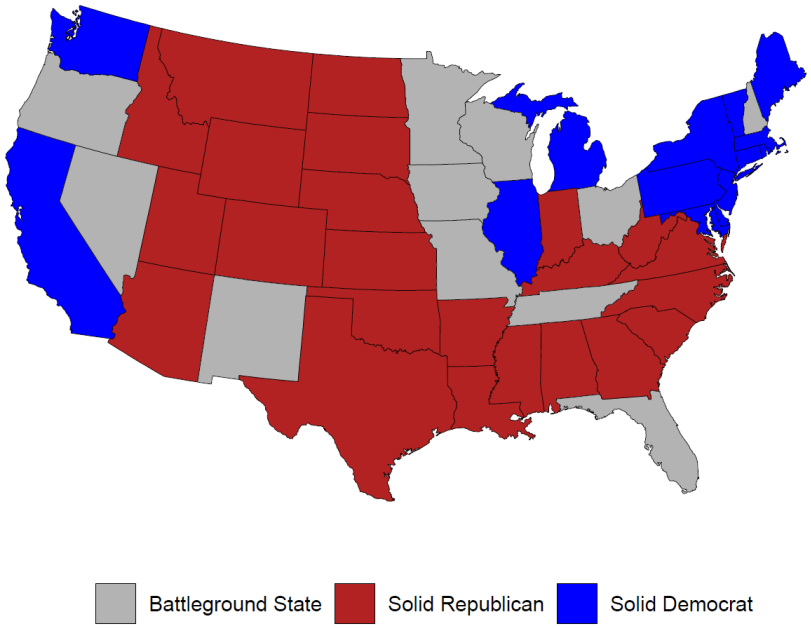
Panel A: Testing Both Hypotheses in Full Panel								
	Procurements				Grants			
Bandwidth (Mi.):	100	80	50	30	100	80	50	30
Treatment	.188***	.292***	.449***	.449***	.182***	.291***	.192***	.134***
(Loyal)	(.062)	(.067)	(.082)	(.094)	(.084)	(.035)	(.040)	(.046)
Treatment	.100***	.170***	.103*	.074	.129***	.043*	.041	.153***
(Battleground)	(.039)	(.044)	(.053)	(.070)	(.023)	(.025)	(.029)	(.037)
R-squared	.628	.614	.660	.681	.737	.746	.784	.822
Fixed-effects	Border + State	Border + State	Border + State	Border + State	Border + State	Border + State	Border + State	Border + State

Panel B: Adding Election-year Interaction Terms								
	Procurements				Grants			
Bandwidth (Mi.):	100	80	50	30	100	80	50	30
Treatment	.173***	.258***	.421***	.428***	.183***	.286***	.205***	.158***
(Loyal)	(.072)	(.077)	(.096)	(.111)	(.037)	(.039)	(.044)	(.050)
Treat (Loyal)	.056	.128	.097	.069	-.004	.030	-.070	-.139
*ElectionYr.	(.132)	(.143)	(.173)	(.197)	(.081)	(.084)	(.096)	(.111)
Treatment	.078*	.142***	.096	.082	.117***	.037	.022	.132***
(Battleground)	(.043)	(.048)	(.059)	(.077)	(.025)	(.027)	(.032)	(.040)
Treat (Battle)	.102	.129	.030	-.037	.057	.030	.095	.094
*ElectionYr.	(.084)	(.096)	.115	(.194)	(.048)	(.054)	(.065)	(.076)
R-squared	.628	.614	.660	.681	.737	.746	.784	.823
Fixed-effects	Border + State	Border + State	Border + State	Border + State	Border + State	Border + State	Border + State	Border + State

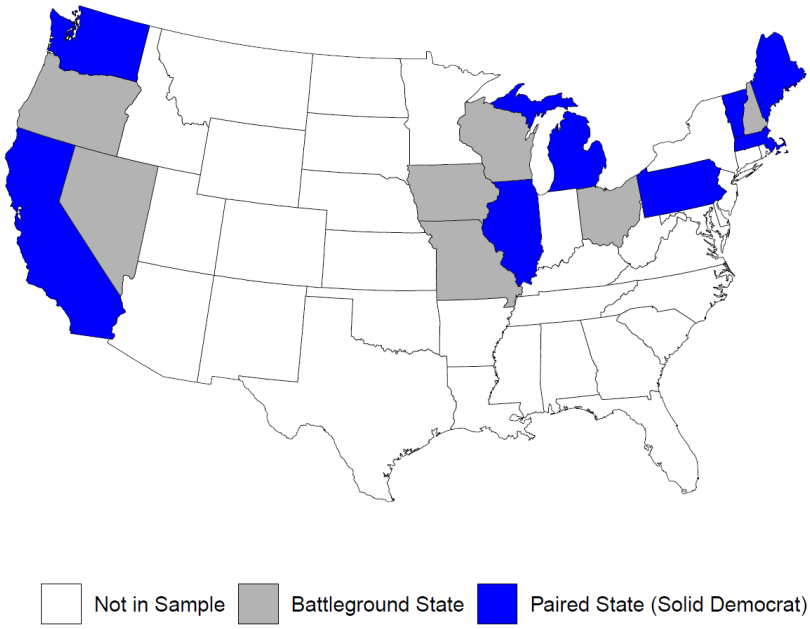
Figure 3.1

In this figure we illustrate how we form our sample of counties under various hypotheses using the election results from the 2000 election as an example. In panel A we display a map showing how all states in the continental U.S. would be coded based on their margins of victory in the 2000 presidential election. States shaded in gray make up the sample of battleground states, which includes the tipping point state and two sets of five states on either side of the tipping point state, as sorted by the winning candidate's margin of victory. States shaded red are coded as "solidly Republican," meaning that they were won by the Republican candidate and have a margin of victory larger than any battleground state. States shaded blue are coded as "solidly Democrat," meaning that they were won by the Democratic candidate and have lower margins of victory (i.e. margins more supportive of the Democratic candidate) than any battleground state. In panel B we display the set of treated and paired states that appear in the sample when testing the hypothesis that federal funds are directed towards battleground states. It highlights pairs of bordering battleground states and paired states that voted solidly for the losing candidate (which, in 2000, happened to be the Democratic party candidate). In panel C we display sets of states that appear in our test of the hypothesis that federal funds are directed toward states that vote heavily in favor of the winning candidate. We show pairs of states that vote solidly for the Republican candidate (treated states) and bordering states that vote heavily in favor of the Democrat (paired untreated states). Treated states (voting solidly for the winning candidate) are shaded in red, while paired states are shaded in blue. In panel D, we give an example of how our sample is formed at the county level using the battleground state hypothesis (i.e. states shown in panel B). We show sets of counties in treated battleground states (shaded in gray) and untreated states (in blue) that are located within 80 miles of the shared borders of these states.

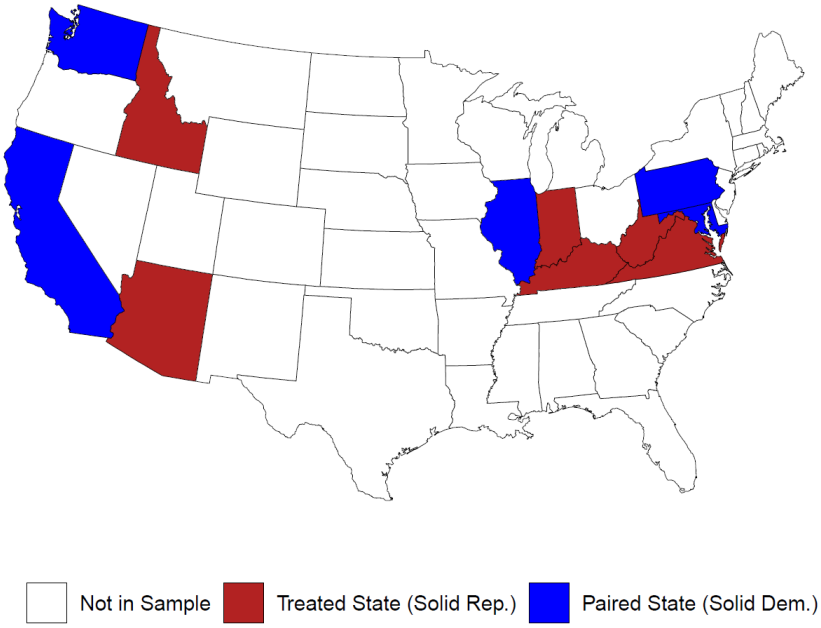
Panel A



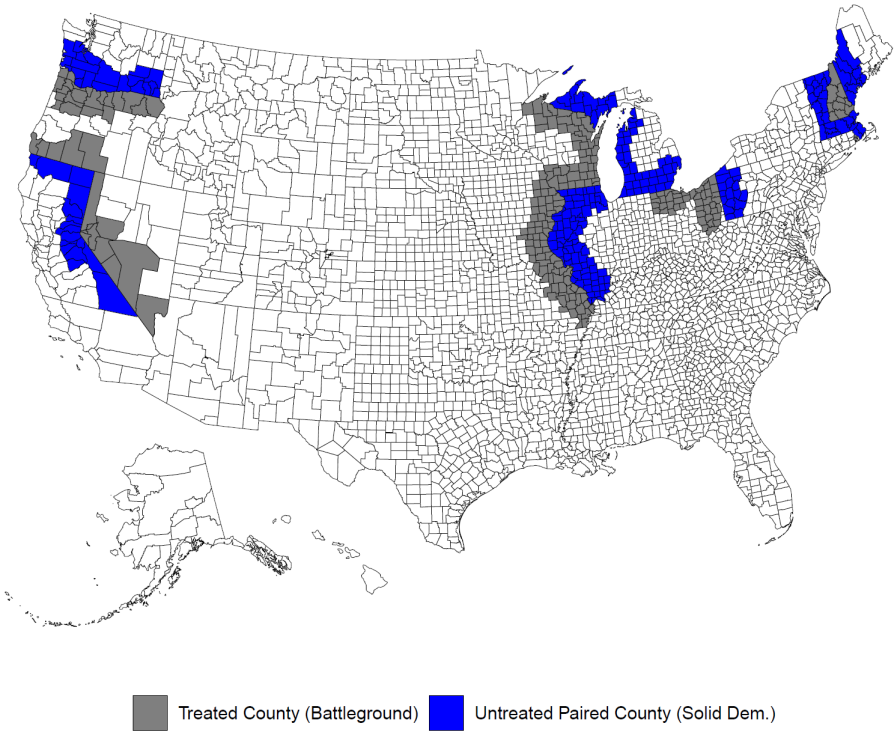
Panel B



Panel C



Panel D


Treated County (Battleground)
Untreated Paired County (Solid Dem.)