Three Essays on the Analysis of Firms' Behaviors Under Staggered Treatment Adoption

Nazanin Sedaghatkish

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Economics and Science

Niloy Bose, Chair
Richard Ashley
Suqin Ge
Xu Lin

July 20, 2023
Blacksburg, Virginia

Keywords: Small U.S. Entrepreneurs, Loan Effects, Personal Finance, Collective Bargaining, Unionization, Financing Decisions, Innovation activities

Three Essays on the Analysis of Firms' Behaviors Under Staggered Treatment Adoption

Nazanin Sedaghatkish

ABSTRACT

This dissertation consists of three essays on firms' behaviors under staggered treatment adoption. The first essay draws information from a micro-lender and a credit bureau to identify the causal effects of small loans on the financial health of a group of small U.S. business owners. To achieve this, we exploit temporal variations in the loan disbursements and use an estimation strategy that controls for potential biases due to treatment effect heterogeneity. The results suggest that even small loans are effective in generating lasting positive impacts on widely accepted financial health indicators, such as *Vantage Score* (Credit Score), *Debt-to-Income Ratio*, and *Credit Utilization Ratio*. We obtain similar robust results for subprime and startup borrowers, who are known to face difficulties in securing credit.

The second essay combines unionization data from the National Labor Relations Board and financial data from Compustat to examine the causal effects of unionization on the financing decisions of publicly traded firms in the United States. In this essay, I exploit temporal variations in the election date of unionization across firms and use a dynamic difference-in-difference estimation strategy to identify the effects of unionization on a range of financial indicators, including the *Debt-to-Equity ratio, market leverage, book leverage, long-term book leverage, net leverage and cash to asset ratio*. I find that unionization negatively affect firms' financing decisions. For example, after unionization, firms rely less on leverage to raise capital. At the same time, unionization offers incentive to firms to hold more cash in hand. My analysis also suggests that the effects of unionization vary according to the political and

institutional structure of the states in which firms operate. For instance, the impacts on the outcome variables are more pronounced for the firms in democrat-led states and for firms which operate in states without right-to-work laws. The effects of unionization are also more noticeable for multi-establishment firms versus one-establishment firms. In addition, we find that the effects vary according to the margin of support for unionization within a firm.

The third essay examines the causal effects of unionization on innovation activities of publicly traded firms in the United States. As in the case of chapters 1 and 2, the analysis uses a dynamic difference-in-difference estimation strategy on a dataset that is compiled using information on unionization data from the National Labor Relations Board, financial data from Compustat and KPSS patent data. My analysis encompasses a wide range of innovation indicators, including the *number of patents, number of forward citations, market value of patents, average citations, number of patents to R&D expenditures ratio, number of citations to R&D expenditures ratio, number of patents per 1000 employees, capital expenditures to sales ratio and R&D expenditures to sales ratio.* The findings suggest a small positive impact of unionization on most of these innovation indicators, with the exception of *market value of patents and number of patents to R&D expenditures ratio.* I also find that the effects of unionization vary according to political orientations of states, industry type, firm size and firm age. The results demonstrate that the effects on innovation are more pronounced for smaller and younger firms and for firms operating in democrat-led states as well as manufacturing firms.

Three Essays on the Analysis of Firms' Behaviors Under Staggered Treatment Adoption

Nazanin Sedaghatkish

GENERAL AUDIENCE ABSTRACT

This thesis is a collection of three self-contained essays that examine the firms' behaviors in contexts where not all the units received the treatment at the same point in time.

In the first essay, we investigate how small loans affect the financial health of small business owners. By analyzing data from a lender and credit bureau, we identify the causal effects of receiving loans on the financial health of borrowers. The results indicate that even small loans have a positive and lasting impact on *credit scores, debt-to-income ratios, and credit utilization ratios*. This research also sheds light on the effects of loans on borrowers with less favorable credit status or those starting a new business, who often face challenges in accessing credit.

In the second essay, the focus shifts to the impact of unionization on the financing decisions of publicly traded firms in the United States. We examine the causal effects of unionization on various financial indicators. The findings reveal a negative effect of unionization on metrics such as *debt-to-equity ratio, market leverage, and book leverage*. However, *cash holdings* experience an increase. Furthermore, the effects of unionization vary based on the political and institutional structure of the states where firms operate, as well as the margin of support for unionization within a firm. The impact of unionization is more pronounced in democrat-led/without right-to-work law states, multi-establishment firms and when the support for unionization is stronger among employees.

In the third essay, we investigate the effects of unionization on innovation activities within publicly traded firms in the United States. By analyzing unionization data, financial data, and patent data, the study examines the causal effects of unionization on various innovation indicators. The results reveal a small positive impact of unionization on most innovation indicators, such as the *number of citations, number of patents per 1000 employees as well as ratio of number of citations to R&D expenditures*. However, the effects on *market value of patents and number of patent-to-R&D expenditure ratios* are not statistically significant. Moreover, the analysis considers factors like political orientations of states in which the firms operate, industry type, firm size and firm age. The findings indicate that the effects on innovation outcomes are more pronounced for smaller firms, younger firms, firms operating in democrat-led states and manufacturing firms.

Dedicated to my father and the memory of my mother.

Acknowledgments

I would like to take this opportunity to express my sincere gratitude to all of those who have played a significant role in the completion of this thesis. Their support, guidance, and encouragement have been instrumental in making this research possible.

First and foremost, my deepest appreciation goes out to my supervisor, Dr. Niloy Bose, for his unwavering support, kindness, expertise, and valuable insights. Beyond his role as a supervisor, he has proven to be an exceptional mentor who truly understands my concerns and consistently offers thoughtful solutions. His constructive feedback, and dedication to my academic growth have been pivotal throughout this journey. I am truly grateful for his support, as he always made time to provide insightful feedback, discuss ideas, and offer constructive suggestions to enhance the depth and clarity of my thesis. His patience and encouragement have been a constant source of motivation, especially during challenging times. His mentorship extended beyond the technical aspects of research, as he also provided guidance on professional development, critical thinking, and scholarly writing. I cannot express my gratitude enough for the care he has shown me at every stage of my academic progress. He invested his time and effort in helping me develop vital skills that will undoubtedly benefit me throughout my academic and professional career. I am truly fortunate to have had the privilege of working under his supervison, and his contributions have shaped the trajectory of my academic pursuits.

I extend my appreciation to the members of my thesis committee, Dr. Richard Ashley, Dr. Suqin Ge and Dr. Xu Lin, for their constructive suggestions and time dedicated to reviewing and evaluating my work. Their scholarly contributions have greatly enriched the quality of this research. I would also like to sincerely thank Dr. Sudipta Sarangi, our

Department Head, for his unwavering support and dedication to the Ph.D. students in our department. His commitment to our success and his efforts in creating a nurturing and inclusive environment have been instrumental in shaping our academic journey. I also owe thanks to the Department of Economics at Virginia Tech University for providing me with the invaluable opportunity to pursue my academic and research aspirations.

Additionally, I would like to express my gratitude to Dr. Sergio Barrera for organizing the Applied Reading Group sessions during the past year, both in the fall and spring semesters. Participating in these sessions was an incredible opportunity for me, allowing me to present my work twice and gain invaluable insights. I am also grateful to Dr. Djavad Salehi-Isfahani, Dr. Suqin Ge and Dr. Melinda Miller for their attendance and valuable suggestions during these sessions, which significantly contributed to the improvement of my research.

Furthermore, I would like to express my appreciation to Dr. Jaideep Chowdhury, Associate Professor in the Finance Department at James Madison University. His generous contribution of time and expertise has played a vital role in the analysis of the data utilized in my second and third chapters. His insights and guidance have been instrumental in enhancing the quality and rigor of my research.

I am deeply grateful to my dear friends from Sharif University of Technology, Sorena Rahi, Vahid Moghani and Zahra Hashemzadeh, for their unwavering moral support, stimulating discussions, and the invaluable sense of unity they brought to this research journey. Their encouragement, feedback, and shared experiences have been invaluable to me. Additionally, I would like to express my heartfelt appreciation to Asal Pilehvari, Ghadir Asadi, Brian D'Orazio, a number of former Ph.D. students of the department, and Amin Hosseinian from the Finance department, who have offered me more support and crucial advice than I could have ever hoped for. Their willingness to help me in any way possible has been a tremendous source of comfort and relief. I would also like to extend my gratitude to my fellow classmates,

Azat, Tommy, Yuan and Long as well as other friends in the department, including Arin, Byungki, Muyu, Nandini, Priya, Jamshid, Jonathan, Reza, and Martin for the unforgettable memories we have created together. Furthermore, I consider myself fortunate to have had the presence of amazing friends in Blacksburg, including Zahra, Morva, Shaghayegh, Simin, Mahshid, Elahe, Angineh, Nasim and Nila who have consistently been a source of support. While it is not possible to mention everyone individually, I want to express my heartfelt appreciation to all those who have supported me throughout this journey.

I would like to extend my gratitude to my family for their unwavering love, understanding, and encouragement throughout this academic pursuit. Their continuous support, belief in my abilities, and sacrifices have been a constant source of inspiration and motivation. I would like to extend my heartfelt gratitude to my father for his unwavering support and love. Talking to him was always a great source of comfort and joy. Despite the physical distance, his words of encouragement and love transcended any barriers. His ability to empathize with my challenges has been a tremendous source of solace and encouragement during my Ph.D. journey. My father's support, combined with the memory of my mother's unwavering belief in me, has propelled me forward during the most challenging times. I would like to thank my younger brother, Ali, for his constant support and encouragement. I am immensely thankful to my cousin, Negin, who has been by my side throughout the Ph.D. journey, sharing similar experiences and providing unwavering support. Your presence and understanding have been invaluable to me. Finally, I am beyond thankful for my husband, Armin, whose unwavering support and caring nature have been instrumental in my journey. His constant encouragement, understanding, and belief in my abilities have given me the confidence to pursue my dreams. His presence has provided me with the strength and motivation to overcome obstacles and reach new heights. I am forever grateful for his unwavering love and support, and I am blessed to have him by my side as we navigate life's adventures together.

Contents

[1] This is a joint work with Niloy Bose and Mohammad Mahdi Banasaz.

List of Figures

List of Tables

Chapter 1

Identification of Loan Effects on Personal Finance: A Case for Small U.S. Entrepreneurs [1]

1.1 Introduction

According to the U.S. Small Business Administration[2], small businesses—firms with less than 500 employees—account for 99.9 percent of all businesses, of which 60% of firms employ only 1-4 employees. Small businesses account for 47 percent of private-sector employment, 32 percent of export, and 60 percent of new job creation in the United States. Despite their importance, small businesses and small startups are known to face difficulties raising capital using traditional channels and experience low survival rates[3]. As a remedial step, there has been a rapid growth of a network of microcredit and microfinance programs in the U.S. Examples of private initiatives include Accion USA, Grameen America, and Kiva. The U.S. Small Business Administration (SBA) has also rolled out programs[4] where SBA brings

[1] This is a joint work with Niloy Bose and Mohammad Mahdi Banasaz.

[2] https://advocacy.sba.gov/wp-content/uploads/2021/12/Small-Business-FAQ-Revised-December-2021.pdf

[3] Between 1994-2019, on the average, only 67.6% of new small business establishments survived beyond two years.

[4] In addition to offering easier access to credit, these programs also offer other valuable services such as management training, technical assistance, and contracting opportunities.

together small entrepreneurs and potential lenders and acts as a guarantor for small business loans. But, how effective are small loans in improving financial health and creditworthiness of small U.S. business owners? We seek an answer to this question using a rich data set obtained directly from a U.S. based microfinance institution and a U.S. credit bureau.

The network of U.S. programs that we see today draws inspiration from the experimentation with microcredit and microfinance programs in the low and middle-income countries. The emerging evidence seems to suggest that the impacts of micro-loans on economic and social measures vary widely according to borrowers' characteristics, loan terms, and the social and institutional environment in which borrowers operate (Banerjee et al. [11], [32], Banerjee et al. [10], Banerjee et al. [13] and Banerjee [14]). Evidently, the environment in which U.S. micro-enterprises operate is different. Most small businesses in the U.S. operate in the formal sector with greater access to infrastructure, formal contracts, markets, and other facilities. Small U.S. enterprises face softer credit terms relative to their counterparts in the low and middle-income world, and if needed, some small business owners in the U.S. can even access alternative credit sources (e.g., credit cards, home equity credit lines, etc.) to supplement borrowed funds (Bernanke [17]). It also may be the case that shared norms, personal trust, and reciprocity play a lesser role for U.S. small businesses. When such differences are present, it is imprudent to conclude the impacts of small credit programs in the U.S. using lessons learned in the different part of the world.

Surprisingly, the existing literature offers little guidance in understanding the impacts of small loans on U.S. micro-enterprises. When it exists, the evidence is largely based on anecdotes, case studies, descriptive statistics, or an empirical strategy that fails to disentangle correlation from causation. The *Longitudinal Impact study by Accion and Opportunity Fund Small Business Lending (2018)* reports a summary of survey responses from a nationwide cohort of 350 borrowers, some of whom were tracked for a post-loan period of three years.

The report suggests a strong positive impact of loans on business indicators such as cash flow, employee hiring, equipment purchases, and business owners' personal and household well-being. Similarly, *Aspen Institute*[5] offer summaries of survey responses from small enterprises highlighting the short and long-run impact of microloan and services. There are a few studies (Young et al. [95], Craig et al. [31] and Lee [65]) that rely on the data aggregated at the MSA level for evaluating the impacts of the SBA loans on employment, startups, and income growth. We differ from these approaches on two grounds. First, neither we rely on the survey data, which are prone to inaccurate measurements and response biases, nor on aggregated data, making it difficult to map intervention to its effects with precision. Our data is based on actual and measurable outcomes where we can track each borrower for a period of 10 years including the pre and post-loan periods. The panel also helps us to control for individual-level unobserved heterogeneity. Second, to capture causal effects, we exploit temporal variation in loan disbursements and use recent advances in the causal inference literature to minimize potential biases arising due to treatment effect heterogeneity.

A part of our data comes directly from the Wisconsin Women's Business Initiative Corporation's (WWBIC) in-house data repository. WWBIC is the largest non-profit microfinance and economic development institution in Wisconsin, assisting the Wisconsin small business community since 1987. We furnish more details about WWBIC in the data section. This data contain demographic and loan information on 737 borrowers who received business loans from WWBIC during 2007-2016, but not all at the same time. The data come to us in an anonymous format where each borrower is assigned a unique identifier. To this file, WWBIC has added borrowers' credit file information for the period 2007-2016, obtained from one of the three credit bureaus.

The borrowers' business information—such as sales, profit, business credit score, etc.—are

[5] https://www.aspeninstitute.org/programs/business-ownership-initiative/data/

missing in the credit file. But the file captures rich information on each borrower's *Vantage Score (Credit Score), debt to income ratio,* and *credit utilization ratio* over ten years that nests pre- and post-loan periods. For two reasons, finding the causal effects of microloans on these variables is meaningful to understanding business outcomes. First, these key variables reflect a borrower's personal financial status, which is arguably correlated with the business's health. Second, personal credit matters directly for business credit access and growth. This is particularly true for sole small proprietors whose personal credit information is often reviewed for business lending. The personal credit status is also reviewed for those borrowers whose business history is lacking. Such borrowers represent at least 45% of borrowers in our sample. We lean on these facts to argue that the outcome variables we consider are informative about the personal as well as business health of borrowers.[7]

The staggered treatment adoption across borrowers naturally lend the analysis to a standard linear two-way fixed effect model or an event study design. However, the recent advances in the econometric literature (Goodman-Bacon [46], Sun and Abraham [88], Athey and Imbens [8], Callaway and Sant'Anna [29], Borusyak and Jaravel [21], Gardner [45] and De Chaisemartin and d'Haultfoeuille [34]) offer reasons to be skeptical about causal estimates from a standard staggered DID framework. The consensus is that, even when the parallel trends assumption is satisfied and treatment assignments are random, the standard approach can lead to misleading estimates in the presence of treatment effects heterogeneity across units or over time. We take this consensus seriously and use a solution proposed by Callaway and Sant'Anna [29], which, in addition to reducing potential biases, offers easily interpretable estimates akin to an event study for capturing both the short and long-run effects of loans. In addition, we subject our estimates to several robustness checks to add confidence to our

[7]For more details, please refer to the publication (https://www.nfcc.org/blog/why-personal-credit-matters-when-applying-for-business-loan/) by National Foundation for Credit Counseling; Washington D.C. Under some circumstances, the Fair Credit Reporting Act does allow lenders to review personal credit history of sole proprietors for business lending.

conclusions.

Our results offer valuable insights. We find a significant positive impact of small loans on the financial health of borrowers in our full sample. For example, within 2-4 (4-6) years of receiving loans, the average *Vantage Score* improves by 68 (92) points. We witness similar improvements in credit utilization and debt-income ratios, both of which play significant roles in lenders' decisions. How impactful are these improvements? At the time of securing loans, 46% of the borrowers in our sample were subprime borrowers (with *Vantage Scores* below 650). If we apply the estimate on *Vantage Score* to this group of these borrowers, then 56% (66%) of subprime borrowers come out of this category within 2-4 (4-6) years of receiving loans. We also find that loan sizes matter, but the impacts of small loans are no less visible. For example, within 2-4 (4-6) years of receiving treatments, the ATT on *Vantage Score* registers 52 (69) point improvements in a sub-sample of 433 borrowers with a mean and the median loan size of only $24,000 and $20,000, respectively. The WWBIC data allows us to identify borrowers seeking loans to start a business, and 45% of borrowers belong to this category. We find that the loan effects on this group of borrowers are particularly robust.

The remainder of this paper is organized as follows. The data and its sources are described in Section 1.2. Section 1.3 offers a detailed overview of the methodology. Section 1.4 discusses the results and includes a series of robustness checks. Finally, Section 1.5 concludes with some comments. Figures and tables are reported in section 1.6. Finally, section A provides additional details on robustness checks in the appendix.

1.2 Background and Data

Wisconsin Women's Business Initiative Corporation[9] (WWBIC) is a non-profit statewide economic development corporation assisting the small business community in Wisconsin since 1987 by lending fair capital for business startups and expansions[10]. It is currently the state's largest micro-lender which operates with support from federal, state, and private donors. Since its inception, the WWBIC has disbursed $ 90 million in loans[11] to small businesses with the mission of helping borrowers who typically face difficulty in accessing capital using traditional channels. Such borrowers include sub-prime, low-income, low-wealth, minority, and women borrowers. For example, across all the categories of assistance (including training and consultation), 70% of WWBIC's clients are female, 61% are people of color, and 63% come from low to moderate-income households.

Our data set consists of **737** individuals who have received business loans from WWBIC during **2007-2016**. The data comes to us in an anonymous format where each borrower is assigned a unique identifier. The data is compiled using information from three sources. The first source is the WWBIC client information file. Individuals who wish to access loans and/or services are required to register with WWBIC as clients. The registration process requires individuals to fill in a form and share various financial and personal details with the WWBIC. This information is saved electronically as the client information file, which contains information on each borrower's age, gender, ethnicity, education, marital and minority status. In our sample, 61% of borrowers are female, 56% of borrowers are married, and 77% of borrowers have a college degree or equivalent. The age distribution of borrowers is skewed toward younger borrowers with a mean and median age of 41 and 39

[9] www.wwbic.com

[10] The WWBIC also offers business and financial education to its members in areas that include, but are not limited to, business planning, business accounting, access to capital, marketing strategy, personal budgeting, debt reduction, saving plans, and legal help.

[11] https://www.wwbic.com/wp-content/uploads/2022/04/2022-Data-Sheet.pdf

years, respectively, and only 17% of borrowers are with minority status[12].

The second source of information is the WWBIC's loan file which contains details on each loan, including loan amounts and closing dates. All individuals in our sample have received at least one loan, while 17% of borrowers received a second loan during the period of analysis. The closing date of loans varies across the borrowers, and the frequency of loan disbursement is slightly skewed toward the later years: half of the borrowers in our sample received their first (second) loan before 2013 (2011). Although WWBIC normally caps the loan sizes at $250,000, the actual loan sizes are much smaller. For example, the mean and the median size of the first loan in our sample is $40,820 and $25,700, respectively. The loan sizes are even smaller for the second loans[14]. According to WWBIC, loan rates generally range from 8-10%. In table 1.1, we summarize the details on loan disbursements. The loan file also distinguishes startup borrowers from borrowers with running businesses. In our sample, 45% of the borrowers have received a loan to start a new business.

Our third and final source of information is the credit files on the borrowers, which the WWBIC has secured from a credit bureau. The credit file is also in an anonymous format, that maintains the same borrower-specific identifiers used in the other files. Thus, we can track each borrower's financial status for a period of 10 years, which nests both the pre- and post-loan periods for borrowers. Our main outcome variable is the *Vantage (Credit) Score*, which is a good proxy for an individual's near past and current financial health and determines a borrower's access to personal and business credit. We also consider the credit utilization and debt-income ratios as supplementary outcome variables. The credit utilization ratio—the ratio of the amount of revolving credit used and the amount of revolving credit available—reflects the extent to which a borrower is credit constrained. This variable receives a significant weight (30%) in the construction of the *Vantage Score*. The debt-income ratio—

[12]According to WWBIC, minority borrowers include African American and Hispanic.
[14]The mean and the median of the second loan are $31,000 and $22,000, respectively.

the ratio of total recurring monthly debt (including credit payments, mortgage, and auto loan) and the gross monthly income— serves as a proxy for a borrower's ability to service additional debt. This ratio does not play a direct role in the construction of the *Vantage Score*, but lenders take this information seriously in the loan approval process. We observe these variables at a two-year frequency starting December 2008. The additional details on these outcome variables are furnished in table 1.2.

1.3 Empirical Methodology

For our analysis, receiving loans is synonymous with receiving treatments, which all individuals in our sample have received, but not all at once. This staggered adoption setting is suitable for using a static linear two-way fixed effect (TWFE) model or an event study design.[15] However, the recent advances in the literature offer valuable insights which put both designs under scrutiny. For example, we now know that the causal parameter in a TWFE design is a weighted average[16] of all possible 2x2 DID estimators that compare timing groups to each other. We also know that in a TWFE model, already treated units act as comparison units. Thus, even when the parallel trend assumption holds and the assignment of treatments is random, the causal estimates in a TWFE model can be misleading due to the presence of treatment effects heterogeneity.[17] As in a static specification, the dynamic TWFE models also fail to yield sensible estimates of dynamic causal effects under

[15] According to De Chaisemartin and d'Haultfoeuille [34], 20% of the empirical papers published in the *American Economic Review* between 2010-2012 are based on the TWFE model.

[16] The weight assigned to each 2x2 DID is sensitive to the panel length, the groups' timing of the treatments, and the relative size of the treatment and control group in the sample.

[17] Since the estimate partly depends on the difference between the changes in the outcomes of (already treated) control units and the changes in outcomes for units that are treated later, the possibility of contamination arises in the presence of time-varying treatment effects.

heterogeneity across cohorts.[18] (Sun and Abraham [88]).

The recent advances in the econometric literature (e.g., Goodman-Bacon [46], Sun and Abraham [88], Athey and Imbens [8], Callaway and Sant'Anna [29], Borusyak and Jaravel [21], Imai and Kim [56], Gardner [45] and De Chaisemartin and d'Haultfoeuille [34]) offer solutions to eliminate potential bias arising due to the heterogeneity in the treatment effects. A direct approach to eliminate bias would be to modify the set of effective comparison units so that units receiving treatments are not compared to previously treated units. Thus, only the 'never treated' units are allowed to act as controls. However, given a relatively small sample size, such a restriction is costly for us. As a solution, we turn to the methodology proposed by Callaway and Sant'Anna [29] (henceforth CS 2021), which allows 'not yet treated' units to act as controls. In addition to providing sensible estimates, this methodology permits us to use our data to its fullest extent and report treatment effects over an extended period while maintaining the balance between the treated and control units. Below, we briefly describe the CS 2021 methodology used in this paper.

The CS 2021 takes a ground-up approach, using group-time specific average treatment effects on the treated, $ATT(g,t)$, as the building blocks. The groups are created according to when the units received (absorbing) treatments in the sample. The $ATT(g,t)$ measures the average treatment effect at time t for the group first treated in time $g < t$, and is defined as

$$ATT(g,t) = \mathbb{E}\left[Y_t(g) - Y_t(0) \mid G_g = 1\right] \tag{1.1}$$

where G_g is a dummy variable equal to one if the unit is in treatment time group g. $Y_t(g)$ is the outcome variable at time t for treated units if they were to first become treated in

[18] In addition, Sun and Abraham [88] noted that the evaluation of pre-trends based on these coefficients could also be misleading. The treatment lead coefficients are not guaranteed to be zero even when parallel trends are satisfied in all periods.

time period g, and $Y_t(0)$ is the potential outcome for the treated units had they not been treated. In the absence of data on $Y_t(0)$, the identification strategy relies on control groups consisting of only those units which have not received treatment up to the time g.

According to CS 2021 and Sant'Anna and Zhao [81], $ATT(g,t)$ can be semi-parametrically estimated using a doubly robust approach that combines the outcome regression (OR) approach of Heckman et al. [51] and the inverse probability (IPW) approach of Abadie [1], and is given by

$$ATT(g,t) = \mathbb{E}\left[\overbrace{\left(\frac{G_g}{\mathbb{E}\left[G_g\right]} - \frac{\frac{p_{g,t}(X)(1-D_t)(1-G_g)}{1-p_{g,t}(X)}}{\mathbb{E}\left[\frac{p_{g,t}(X)(1-D_t)(1-G_g)}{1-p_{g,t}(X)}\right]}\right)}^{\text{Inverse Probability Weight}} \overbrace{\left(Y_t - Y_{g-1} - m_{g,t}(X)\right)}^{\text{Outcome Regression}}\right] \tag{1.2}$$

As before, the term G_g is a binary variable that is equal to 1 when an individual is first treated in period g. The set of covariates is given by X, and the term $p_{g,t}(X)$ represents the propensity score prediction of receiving treatment (loans) at $t = g$. This prediction is based on the group of individuals receiving loans at $t = g$, and those groups who have received loans no earlier than the period $t+1$. The first expression represents the Inverse Probability Weighting (IPW) term (Abadie [1]), and the term $m_{g,t}(X) = \mathbb{E}\left[Y_t - Y_{g-1} \mid X, D_t = 0, G_g = 0\right]$ is the estimated conditional expectation function from the outcome regression (OR) approach of Heckman et al. [51] measuring change from $g-1$ to t in outcome Y_t for the group receiving loan no earlier than the period $t+1$, conditional on covariates X. Thus, the $ATT(g,t)$ in equation (1.2) is the weighted average of the difference in the evolution of outcomes between the treated and the control groups, where higher weights are assigned to the difference in outcomes between the treated and control group when the control group shares similar characteristics to those found in the treated group. This weighting procedure guarantees that the covariates of the group g and the control group are balanced. The OR approach requires

researchers to correctly model the outcome evolution of the comparison group. The IPW approach, on the other hand, requires one to correctly model the conditional probability of unit i being in group g given X. According to CS 2021, the above $ATT(g,t)$ estimator is doubly robust in the sense that for the estimate to be valid, one is required to correctly specify either the outcome evolution for the comparison group or the propensity score model, but not necessarily the both.

We use the methodology outlined above to divide the sample period into five sub-intervals: 2007-2008, 2009-2010, 2011-2012, 2013-2014, and 2015-2016, and assign a group identifier to a borrower according to which sub-interval includes the borrower's first loan date. For example, we assign borrowers to group 2 if these borrowers received their first loan during the two-year interval 2009-2010. Similarly, we assign borrowers to group 5 if their loan dates belong to the interval 2015-2016. These sub-intervals and groups is not arbitrary. They are designed to utilize the maximum amount of information on key outcome variables and to capture the loan effects over a longer horizon.

As a first step, we seek to use equation (1.2) to estimate $ATT(g, t = g+e)$: the average loan effects on borrowers belonging to group g as a function of time e relative to the treatment period. As noted earlier, we observe outcome variables at a two-year frequency starting December 2008. These reporting dates do not necessarily align with the borrowers' loan dates. For example, borrowers with loan closing dates in 2009-2010 belong to group 2, for whom we observe the earliest post-treatment outcomes on December 2010. At this date, the length of exposure to treatment for group 2 borrowers could vary between 0 - 2 years. We face a similar issue for borrowers in the other groups, requiring us to interpret $ATT(g, t = g + e(= 0))$ as the average loan effect materializing for group g within the 0 - 2 years of receiving loans. Similarly, $ATT(g, t = g + e(= 2))$ needs to be interpreted as the average loan effect for the group g in the 2 - 4 year interval, and so on.

With group $ATTs$ in hand, our next goal is to construct estimates that are well suited to capture instantaneous as well as time-varying loan effects. For each value of e, we construct a weighted average of the group $ATTs$, assigning the weights by the group sizes. We repeat this exercise for every relative period e, including the pre-treatment ($e \leq 0$) periods, to present our results in a standard event study format. Besides capturing the evolution of loan effects, the event study format presents an opportunity to conduct an informative test of the parallel trend assumption using formal inference based on CS 2021 recommended bootstrapping procedure which reports simultaneous confidence bands robust to multiple hypothesis testing and individual cluster errors. To create a single, overall point estimate, we take the average of these aggregated relative time estimates with $t \geq g$. Our analysis takes advantage of the `R` code, which has been made publicly available[21] by Callaway and Sant'Anna [29] as a supplement to their research.

Vantage (Credit) Score is our primary outcome variable, which we observe at a two-year frequency starting December 2008. As a result, pre-treatment information on *Vantage Score* on 50 borrowers in group 1 (who received loans during 2007-2008) is missing in our data. We exclude these borrowers from the analysis. Our methodology also prohibits us from using already treated units as controls. Thus, the last treated borrowers (group 5) lack controls of their own and can only serve as comparison units in earlier periods. Therefore, the loan effects that we report below are based on $ATTs$ that materialized for the remaining three groups.

[21] Callaway and Sant'Anna's `R` Package, `DiD`, version 2.0.1.908. See https://bcallaway11.github.io/did/ for more information on this package.

1.4 Loan Effects

Table 1.3 summarizes loan effects on the *Vantage Score*. The single estimate in the first row represents a weighted average of all group-time $ATTs$. The next row reports the weighted average of the groups' treatment effects by the length of exposure to the treatment. The results suggest that the group-average treatment effect within the first two years of receiving loans ($e = 0$) is marginally negative and indistinguishable from zero. The effects, however, become visible with time. In the 2-4 years interval ($e = 2$), the group-average ATT on the *Vantage Score* records an 68 points improvement. We experience a loss of (not-yet-treated) comparison units as we move forward in time. Despite this, the loan effects on the *Vantage Score* in the 4-6 year intervals ($e = 4$) remain statistically significant and register an increase of 92 points. How significant are these effects? At the time of receiving loans, nearly 46% of borrowers in groups 2, 3, and 4 were sub-prime borrowers (with a *Vantage Score* below 650). Within the 2-4 years (4-6 years) post-treatment, the improvements in the *Vantage Score* is large enough to move 56% (66%) of these borrowers out of the subprime category with a prospect of better financial health and better access to credit. Improvements of this magnitude clearly hold the potential for transformational impacts in the lives of borrowers.

In Figure 1.1, we supplement the results with an event study plot for pre- and post-treatment estimates with 95% simultaneous confidence bands. The pre-treatment estimates offer information about the validity of the (conditional) parallel trend assumption, requiring that in the absence of the treatment, the outcome variable follows the same trend across borrowers with similar characteristics. The individual characteristics that we consider are borrowers' age, gender, minority status, level of education, marital status, and business startup status that we observe at the time of securing loans. We use these characteristics as conditional variables for the full sample. However, for some sub-samples, insufficient variation in the data prevents us from executing a meaningful outcome regression component of the equation

(1.2). For such cases, we use a subset of the conditioning variables. In Figure 1.1 as well as table 1.4, the *ATTs* prior to treatments are indistinguishable from 0, rendering support to the (conditional) parallel trend assumption.

There is evidence in the recent literature [10, 12] to suggest that borrowers' characteristics are key to the success of microcredit programs. Could it be the case that WWBIC only targets a specific group of borrowers? At the time of loan closing, the borrowers vary widely in their credit scores, and nearly half of the borrowers appear in our sample as subprime borrowers. The borrowers' credit files also offer pre-loan income information and includes an income insight variable built on the credit bureau's income estimation model. Although this variable does not necessarily capture the actual income of borrowers, the variable offers additional insight into the borrowers' composition. When receiving loans, borrowers' scores range from 10 to 206, with higher values representing higher income. The mean and the standard deviation of the income insight variable are 52 and 25, respectively, with 50% of the borrowers having a score less than 47. We take this diversity among borrowers at the time of getting loans as support against sample biases. The validity of our estimates also depends on the 'no anticipation of the treatments'. WWBIC returns its loan approval decisions on a rolling basis within a short period of receiving the applications, where it is unlikely that borrowers are able to choose either the treatment status or have prior knowledge about the treatment path. Also, the treatment effects we capture are not immediate, but span many years. We view these features mitigating the possibility of anticipation effects contaminating our estimates.

The *Vantage Score* depends on several factors, such as an individual's credit utilization ratio, payment history, the number of new credit inquiries, etc. Information on these variables is available in the credit file. But, not all variables—such as payment history and the number of new credit inquiries—are well suited for our analysis due to the format in which they

appear in the data making it difficult for us to draw a clear distinction between the pre and post-treatment observations for borrowers belonging to the same group.[23] The data on the credit utilization ratio is, however, suitable for our analysis. As in the case of the *Vantage Score*, we observe this variable at a two-year frequency for each individual, and we can clearly separate pre-treatment observations from the post-treatment observations. We also consider the debt-to-income ratio as an outcome variable. Though not used for the *Vantage Score*, lenders pay close attention to this variable to assess a borrower's ability to service new debt obligations and prefer to see this ratio not exceeding 0.36 at the time of loan approval. Information on this variable is also available at a two-year frequency starting December 2008.

We report event study plots for these two additional variables in Figures 1.2a and 1.2b. The *ATTs* on the utilization ratio become visible in the 2-4 year interval, and compared to the control group, the ratio registers a 24% (57%) drop by the 2-4 (4-6) years post treatment. In contrast, the debt-income ratio is immediately impacted by loans and experiences an increase in the 0-2 year interval. This increase is transient and is followed by a steady decline, though not always statistically significant. Only in the 4-6 years interval, the ratio registers a 16% drop (with a borderline statistical significance). At the time of loan closing, 37% of borrowers in our sample did not meet the 0.36 debt-income threshold. By 4-6 years post-treatment, the size of the ATT treatment effect is large enough to improve the creditworthiness of half of these borrowers by lowering their debt-income ratios below 0.36.

Only 17% of the borrowers in our sample have received multiple loans. It is, however, possible that this small group of borrowers is driving our results. To examine this, we

[23]The data on the payment history and inquiries come with a time reference—such as the number of new inquiries or delinquencies in the last 12 months. Whereas, we associate group identifiers based on loan dates that lie in two-year intervals. In these intervals, the data that we observe on these two variables show up as pre-treatment data for some borrowers and post-treatment data for others, even when both sets of borrowers share the same group identifier.

restrict our sample to 613 borrowers with a single loan. The event study plots in Figure 1.3 validate our full-sample results and even present a more convincing case for all three outcome variables. It also may be the case that large average treatment effects are primarily driven by borrowers with larger loans. To examine this, we re-do the analysis after excluding the top 30% single loan recipients by loan size, leaving us with 433 borrowers with loan sizes less than or equal to $50,000. The mean and median loan sizes of this sub-sample are $24,000 and $20,000, respectively. The results (in Figure 1.4) suggest that while smaller loans diminish the treatment effects, their effects remain quite noticeable: in the 2-4 and 4-6 year intervals, the *Vantage Score* increases by 52 and 69 points. The credit utilization and debt-income ratios follow trends similar to those observed for the full sample.[25] Though in some cases—as in the case of the debt-income ratio—the *ATTs* lose statistical significance, the results offer evidence to suggest that even small loans are effective for positive impacts on entrepreneurs' financial health and creditworthiness.

The startup borrowers are known to face difficulties accessing credit and represent 45% of our sample. To understand the loan effects on this group, we restrict our sample to only startup borrowers and estimate the average treatment effects while drawing both treatment and comparison units from this sub-sample. In other words, the estimates in Figure 1.5 capture the loan effects within the group of startup borrowers. Despite having a smaller sample, the loan effects in the 'startup' group show up more robustly than in the full sample. The estimates are also statistically significant across all categories of indicators. The sub-prime borrowers are also visible in our sample, with a share of 46%. Since 50% of the borrowers in the startup sample are also sub-prime borrowers, the results on 'startups' partly inform about the loan effects on the sub-prime borrowers. To be sure, we construct a sub-sample consisting only of subprime borrowers. As in the case of 'startups', we draw both treated

[25] We repeat but do not report the exercises where we exclude the top 10% and 20% of loans by size. The positive correlation between the loan sizes and loan effects is also transparent in these exercises.

and comparison groups from the sub-sample. The results in Figure 1.6 suggest that effects are mixed and unlike the 'startups', loan effects are less robust. For example, in the 2-4 year interval, the *Vantage Score* improves by 50 points, but improvements are not statistically significant. The robust improvements show up only in the 4-6 years post-treatment interval. Though the loan effects are also missing in the case of the debt-income ratio, the results for the credit utilization ratio appear robust and within four years of receiving loans, the ratio registers more than a 20% decrease compared to the control group.

We allocate the rest of this section to additional robustness checks. For our methodology, the composition of the treated groups changes with $e(= 0, 2, 4)$. For example, the event study estimate for the $0 - 2$ year interval ($e = 0$), uses treatment effect information on all three groups. In contrast, for the 4-6 year interval ($e = 4$), our data only permits the dynamic coefficient to use information on group 2. Such issues are not a concern if treatment effects are common across the groups, which may not hold in our case since groups are treated on different dates and may have encountered different environments. To address this concern, we restrict the analysis only to $e = 0$ and $e = 2$ for which the composition of the treated groups remains unchanged. The Event Study (Balanced) row in Table 1.3, reports the estimates using the information on group 2 and group 3 for both periods. By construction, the Event Study and the Event Study (Balanced) use the same treatment and comparison groups for $e = 2$, and return the same treatment effect. The difference, however, appears in the case of $e = 0$ where the initial decrease in the *Vantage Score* is now more pronounced and is statistically different from zero.

In the Online Appendix [Figure A.1], we present results from an experiment where we randomly assign borrowers to placebo treatment groups without paying attention to their true loan closing dates. We consider 80 such treatment assignments and estimate the dynamic treatment effects for each assignment using the same set of covariates as in our main analysis.

Figure A.1, presents the distribution of the treatment effects on the *Vantage Score* in 4-6 year interval, post-treatment. The results suggest that *ATT* placebo estimates are centered approximately around 0 with an average of 2.2, and the upper bound of the distribution is significantly smaller than our original estimate. We take these results as support for our main conclusions.

Most of the results support that the average outcomes for the treated and control groups follow parallel paths in the absence of treatment. However, the presence of parallel trends during pre-treatment does not guarantee that the time trend for untreated units is comparable to the counterfactual time trend for treated units during post-treatment periods. Some time-varying confounds or some other shocks can add biases to the estimates by causing the post-treatment trends to differ, even if pre-treatment trends were the same. In addition, the pre-tests can be under powered and may fail to detect violation of parallel trends—particularly for a small sample. To address these concerns, we turn to Rambachan and Roth [78] which offers a formal sensitivity analysis that relates the magnitude of violations of parallel trends to the robustness of treatment estimates in post-treatment periods. We use the updated version of the `DID`[27] package to generate results. The results suggest that the loan effects on the *Vantage Score* remain significant and statistically different from 0 even when we allow for some violation of PTA. In the case of 2-4 (4-6) year loan effects, our conclusions are robust to the violation of PTA up to 1.1 (0.9) times the largest violation in the pre-treatment periods. We obtain similar thresholds in the cases of utilization ratio. We include a more detailed analysis in the online appendix (Figures A.2, A.3, A.4 and A.5).

[27] https://github.com/bcallaway11/did

1.5 Conclusion

Despite the rapid growth of microcredit network in the U.S., there remains little systematic evidence to evaluate the impacts of small loans on small U.S. business owners. This paper seeks to fill in this gap by focussing on the effects of loans on three key outcome variables—*Vantage (Credit) Score, debt-income ratio,* and *credit utilization ratio.* These variables not only inform about the effects of loans on borrowers' personal financial status, but also inform about how small loans shape the prospect of business credit and business success. We find positive and statistically significant impacts of loans on these variables, for borrowers including startups and subprime borrowers. Though loan sizes matter, smaller loans also hold the potential for transformational impacts. The development experts are divided about the effectiveness of the microcredit programs in developing countries. The results that we obtain in the context of U.S. are too strong to ignore and probably suggest that it is not only the access to credit, but its interaction with the institutional and social factors, that hold the key to the success of small-loan programs.

1.6 Figures and Tables

Table 1.1: Distributions of Loan Timings and Amounts

Panel A: Frequency of Loan Disbursements										
Year	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Number of borrowers	28	22	37	56	89	76	96	113	103	117

	count	mean	sd	min	max	p25	p50	p75	**Source**	Frequency
Panel B: Loan Details										
1st Loan Amount (Whole Sample)	737	40820.47	37053.17	1000	312500	15000	25661.41	50000	WWBIC	Fixed
1st Loan Amount (Single Loan Group)	613	42339.73	38493.84	2000	312500	15000	30000	55000	WWBIC	Fixed
2nd Loan Amount	124	31177.78	28797.9	1500	150000	9969.56	21781.14	43061.84	WWBIC	Fixed

[a] Std. Errors are in parentheses and * indicates p-value < 0.05.

Table 1.2: Distribution of Pre-treatment (one period before getting the loan) Outcome Variables

	count[a]	mean	sd	min	max	p25	p50	p75	**Source**	Frequency
Panel A: Outcome Variables										
Vantage Score	683	653.5608	88.22721	341	824	595	656	722	Credit Bureau	Biyearly
Debt to Income Ratio	644	30.23292	23.46053	0	101	11	27	45	Credit Bureau	Biyearly
Utilization Ratio	562	.4441608	.3453867	0	2.85125	.1512403	.4023014	.7081652	Credit Bureau	Biyearly

[a] Notes: Number of borrowers (count) are less than 737 because we don't have information on the pre-treatment status of Group 1. Therefore, they are not included in the analysis of the summary statistics of outcome variables based on the status prior to receiving loans.

Table 1.3: Loan Effect Estimates on Vantage Score

	Partially Aggregated				
	(1)	(2)	(3)	(4)	Single Parameters (5)
Conditional Parallel Trends					
Simple Weighted Average					31.9095 *
					(13.4738)
	e=0	e=2	e=4		
Event Study	-0.6837	67.9937 *	91.9384*		53.0828 *
	(10.2447)	(20.2825)	(28.759)		(17.190)
	e=0	e=2			
Event Study (Balanced Groups)	-20.2912 *	67.9937 *			23.8512 *
	(5.9970)	(20.0118)			(11.0867)

Notes: The table reports aggregated treatment effect parameters under conditional parallel trends assumptions and with clustering at the individual level. The row 'Simple Weighted Average' reports the weighted average (by group size) of all available group-time average treatment effects. The row 'Event Study' reports average treatment effects by the length of exposure to the loan; For example, estimates corresponding to $e = 0$, represents weighted average of group-time $ATTs$ in the 0-2 year interval from receiving loans. Similarly, $e = 2$, represents weighted average of group-time $ATTs$ in the 2-4 year interval from receiving loans and so on. The row 'Event Study (Balanced Groups) reports average treatment effects by length of exposure while making sure that the composition of the treatment group does not change with e. The last two elements in the 'Single Parameters' column are the average of dynamic effects reported in the Event Study and Event Study (Balanced Groups) rows. '*' indicates that the results are statistically significant (p-value < 0.05).

Table 1.4: Dynamic Effect of the Loan Receipt on Outcome Variables[a]

	(1)	(2)	(3)
e	Vantage Score	DTI	Utilization Ratio (18)
-4	10.7320	-1.1812	0.0448
	(8.6211)	(2.0704)	(0.0403)
-2[b]	4.2334	-0.9328	-0.1142
	(10.6771)	(4.1376)	(0.0546)
0[c]	-0.6837	9.3118*	0.0130
	(10.1093)	(3.2789)	(0.0389)
2[d]	67.9937*	-3.7454	-0.2358*
	(22.5877)	(8.8020)	(0.0921)
4	91.9384*	-15.9220	-0.5656*
	(28.8007)	(8.2614)	(0.1360)
Overall	53.0828*	-3.4519	-0.2628*
	(16.0519)	(6.5312)	(0.0694)

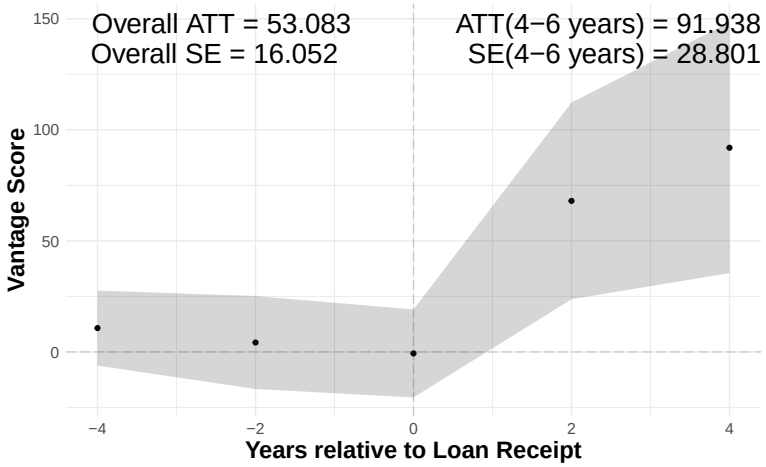
[a] Std. Errors are in parentheses and * indicates p-value < 0.05.

[b] $e = -2$ indicates the average loan effect within the interval 2-4 years prior to getting the loans

[c] $e = 0$ means the instantaneous effect of the loans (the average loan effects that materialized within the 0-2 years after receiving the treatments.)

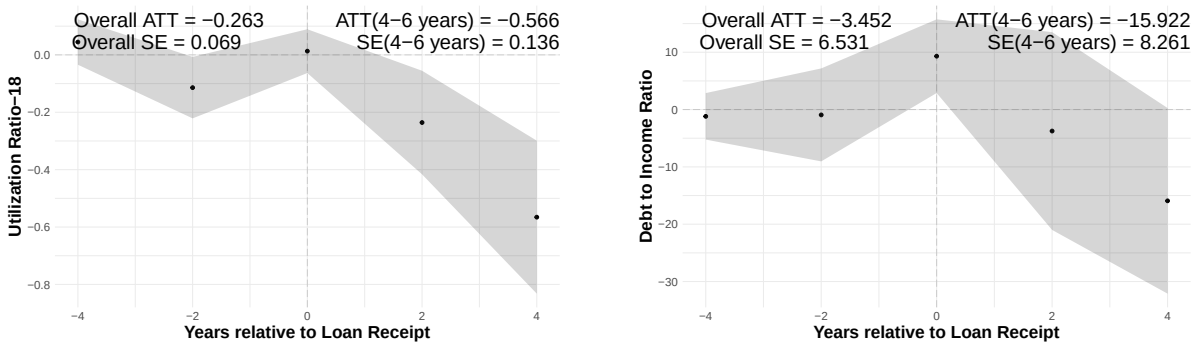
[d] $e = 2$ indicates the effect of the loan after 2-4 years

Figure 1.1: Event Study plot for Vantage Score



Notes: The figure presents dynamic difference-in-difference estimates under conditional parallel trends assumptions and simultaneous 95% confidence bands. Standard errors are clustered at the individual level. The outcome variable is Vantage Score and covariates include age, gender, married, education, start-up and minority. Years relative to loan receipt capture exposure to the treatment (e); for example, estimates corresponding to $e = 0$, represents weighted average of group-time $ATTs$ in the 0-2 year interval from receiving loans. Similarly, $e = 2$, represents weighted average of group-time $ATTs$ in the 2-4 year interval from receiving loans and so on. For each e, ATT is calculated by averaging dynamic treatment effects across all groups. The Overall ATT is calculated by averaging dynamic treatment effects across all event times. The average treatment effects are not significant when confidence band include 0.

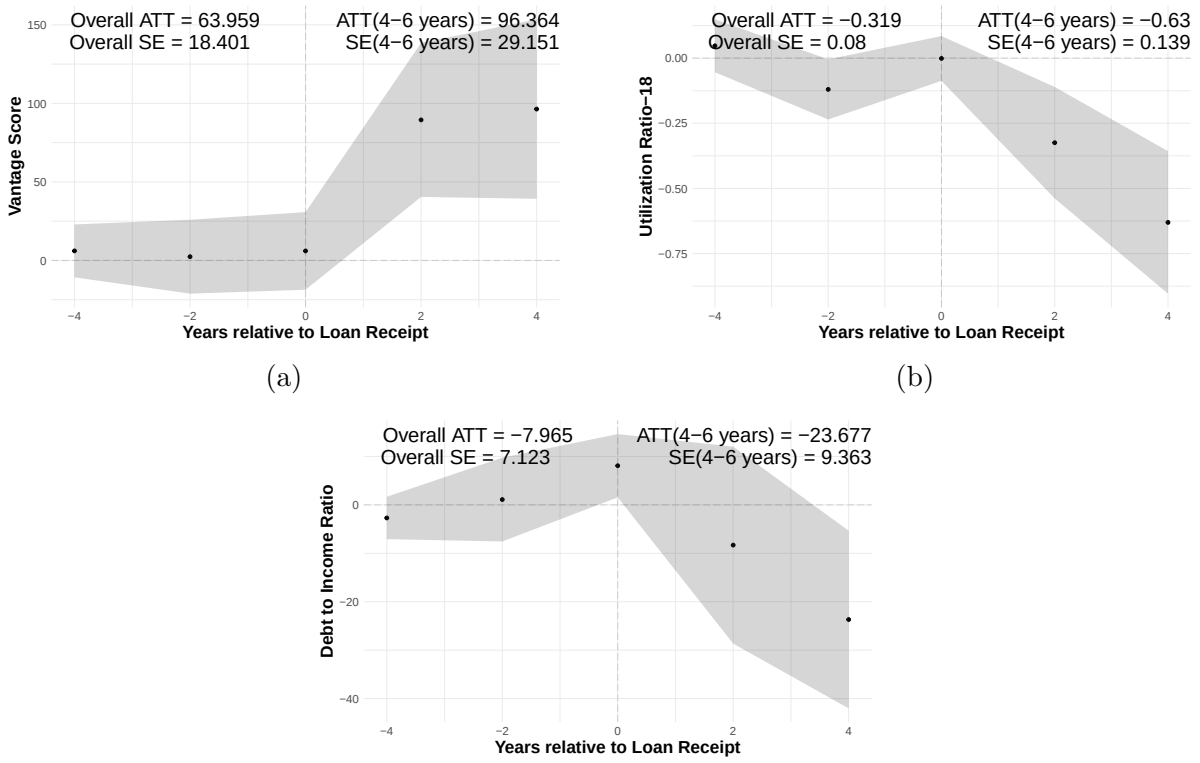
Figure 1.2: Event Study plots for Utilization Ratio and Debt to Income Ratio



(a) (b)

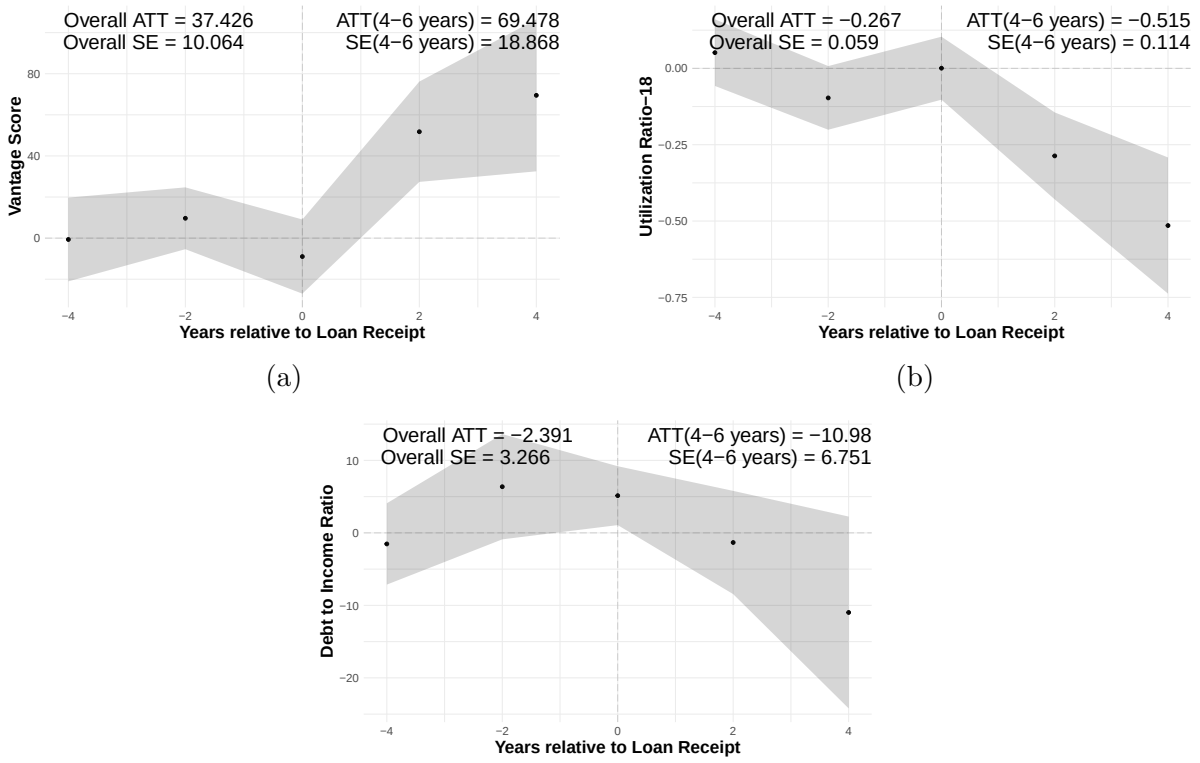
Notes: The figures present dynamic difference-in-difference estimates under conditional parallel trends assumptions and simultaneous 95% confidence bands. Standard errors are clustered at the individual level. The outcome variables include Utilization Ratio and Debt to Income Ratio and covariates include age, gender, married, education, start-up and minority. "Years relative to loan receipt" capture exposure to the treatment (e); for example, estimates corresponding to $e = 0$, represents weighted average of group-time $ATTs$ in the 0-2 year interval from receiving loans. Similarly, $e = 2$, represents weighted average of group-time $ATTs$ in the 2-4 year interval from receiving loans and so on. For each e, ATT is calculated by averaging dynamic treatment effects across all groups. Overall ATT is calculated by averaging dynamic treatment effects across all event times. The Credit Bureau reports DTI as a multiple of 100, for example, DTI = 50 refers to 0.5. For our analysis, we keep the Credit Bureau data format. Therefore, the reported ATTs for DTI in the Event Study plots, needs to be divided by 100 to express it in the fraction format. The average treatment effects are not significant when confidence band include 0.

Figure 1.3: Single Loan Event Study Plots



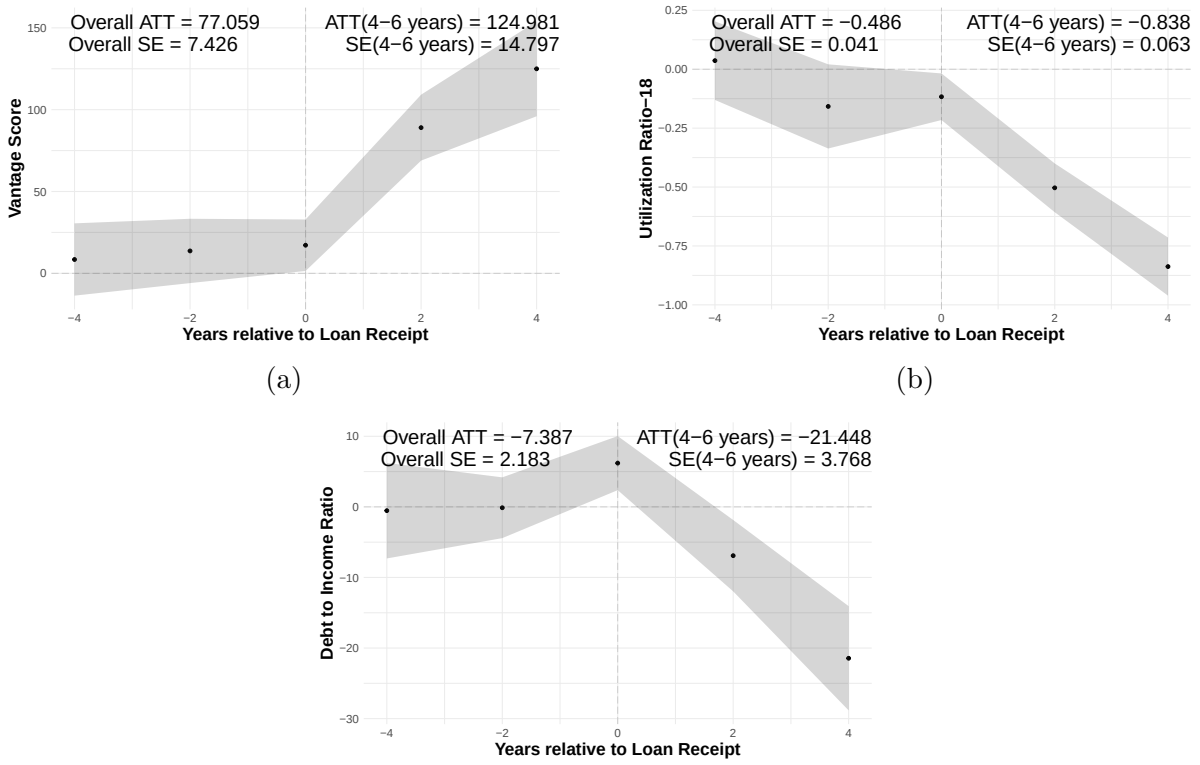
Notes: The figures present dynamic difference-in-difference estimates for the single loan group under conditional parallel trends assumptions and simultaneous 95% confidence bands. Standard errors are clustered at the individual level. The outcome variables include Vantage Score, Utilization Ratio and Debt to Income Ratio and covariates include age, gender, married, education, start-up and minority. "Years relative to loan receipt" capture exposure to the treatment (e); for example, estimates corresponding to $e = 0$, represents weighted average of group-time $ATTs$ in the 0-2 year interval from receiving loans. Similarly, $e = 2$, represents weighted average of group-time $ATTs$ in the 2-4 year interval from receiving loans and so on. For each e, ATT is calculated by averaging dynamic treatment effects across all groups. Overall ATT is calculated by averaging dynamic treatment effects across all event times. The Credit Bureau reports DTI as a multiple of 100, for example, DTI = 50 refers to 0.5. For our analysis, we keep the Credit Bureau data format. Therefore, the reported ATTs for DTI in the Event Study plots, needs to be divided by 100 to express it in the fraction format. The average treatment effects are not significant when confidence band include 0.

Figure 1.4: Event Study Plots for Single Loans less than $ 50,000



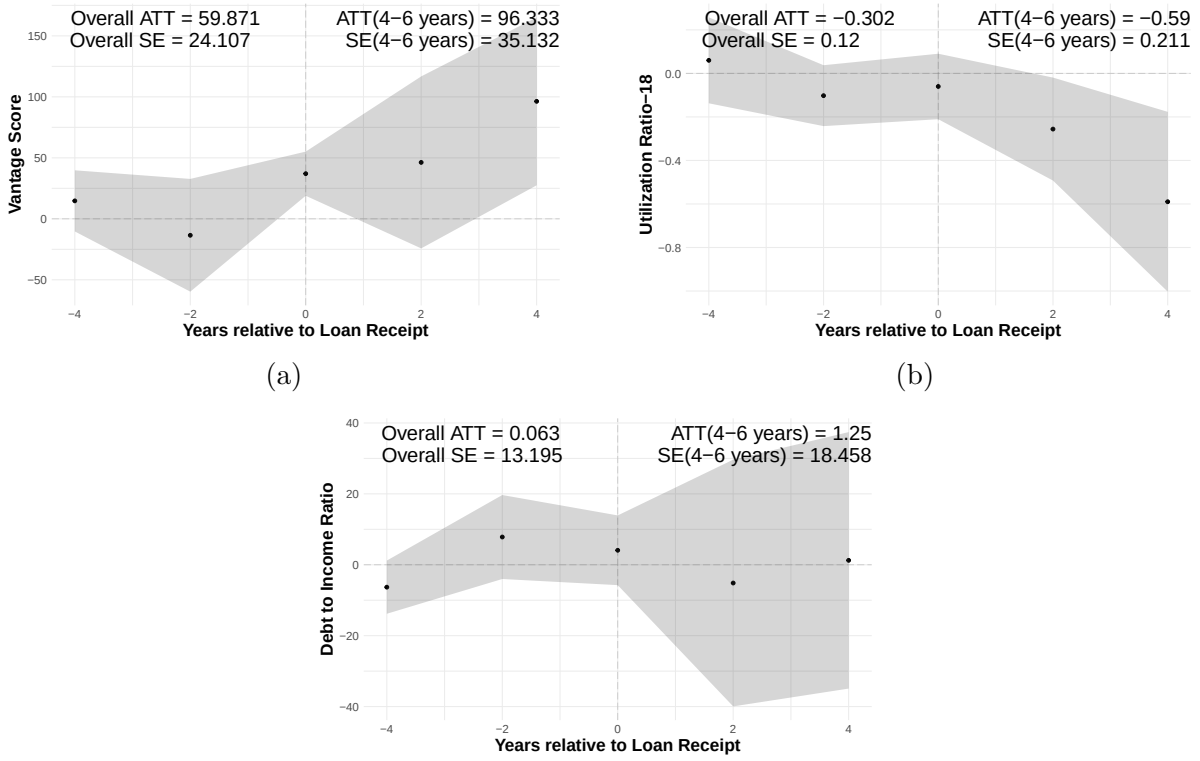
Notes: The figures present dynamic difference-in-difference estimates for the group of single loans less than $ 50,000 under conditional parallel trends assumptions and simultaneous 95% confidence bands. Standard errors are clustered at the individual level. The outcome variables include Vantage Score, Utilization Ratio and Debt to Income Ratio and covariates include age and start-up. "Years relative to loan receipt" capture exposure to the treatment (e); for example, estimates corresponding to $e = 0$, represents weighted average of group-time $ATTs$ in the 0-2 year interval from receiving loans. Similarly, $e = 2$, represents weighted average of group-time $ATTs$ in the 2-4 year interval from receiving loans and so on. For each e, ATT is calculated by averaging dynamic treatment effects across all groups. Overall ATT is calculated by averaging dynamic treatment effects across all event times. The Credit Bureau reports DTI as a multiple of 100, for example, DTI = 50 refers to 0.5. For our analysis, we keep the Credit Bureau data format. Therefore, the reported ATTs for DTI in the Event Study plots, needs to be divided by 100 to express it in the fraction format. The average treatment effects are not significant when confidence band include 0.

Figure 1.5: Start-up Event Study Plots



Notes: The figures present dynamic difference-in-difference estimates for the group of startups under unconditional parallel trends assumptions and simultaneous 95% confidence bands. Standard errors are clustered at the individual level. The outcome variables include Vantage Score, Utilization Ratio and Debt to Income Ratio. "Years relative to loan receipt" capture exposure to the treatment (e); for example, estimates corresponding to $e = 0$, represents weighted average of group-time $ATTs$ in the 0-2 year interval from receiving loans. Similarly, $e = 2$, represents weighted average of group-time $ATTs$ in the 2-4 year interval from receiving loans and so on. For each e, ATT is calculated by averaging dynamic treatment effects across all groups. Overall ATT is calculated by averaging dynamic treatment effects across all event times. The Credit Bureau reports DTI as a multiple of 100, for example, DTI = 50 refers to 0.5. For our analysis, we keep the Credit Bureau data format. Therefore, the reported ATTs for DTI in the Event Study plots, needs to be divided by 100 to express it in the fraction format. The average treatment effects are not significant when confidence band include 0.

Figure 1.6: Subprime Event Study Plots



Notes: The figures present dynamic difference-in-difference estimates for the group of subprime borrowers under conditional parallel trends assumptions and simultaneous 95% confidence bands. Standard errors are clustered at the individual level. The outcome variables include Vantage Score, Utilization Ratio and Debt to Income Ratio and covariates include age and gender. "Years relative to loan receipt" capture exposure to the treatment (e); for example, estimates corresponding to $e = 0$, represents weighted average of group-time $ATTs$ in the 0-2 year interval from receiving loans. Similarly, $e = 2$, represents weighted average of group-time $ATTs$ in the 2-4 year interval from receiving loans and so on. For each e, ATT is calculated by averaging dynamic treatment effects across all groups. Overall ATT is calculated by averaging dynamic treatment effects across all event times. The Credit Bureau reports DTI as a multiple of 100, for example, DTI = 50 refers to 0.5. For our analysis, we keep the Credit Bureau data format. Therefore, the reported ATTs for DTI in the Event Study plots, needs to be divided by 100 to express it in the fraction format. The average treatment effects are not significant when confidence band include 0.

Chapter 2

The Effects of Unionization on Firms' Financing Decisions: Evidence from U.S. Firms

2.1 Introduction

Labor unions typically bargain for higher wages, enhanced job security, and better benefits. Notably, the United States has witnessed a decline in trade unionism over the past few decades, with percentage of workers under union contract decreasing from 21% in 1984 to approximately 10% in 2021. However, the American labor movement experienced substantial growth since the beginning of the pandemic. In 2022, more than 16 million workers in the United States were represented by a union, marking an increase of 200,000 from the previous year. Furthermore, there have also been notable unionization efforts by workers at major corporations such as Starbucks and Amazon[1]. The recent unionization efforts have rekindled debates surrounding the efficacy and desirability of unionization in today's economy.

There already exists a large literature that explores the impact of unions on wages (Bryson

[1]In November 2021, over 400 Starbucks workers in Buffalo, New York filed for a union election, seeking to form the first union at a company-owned Starbucks in the United States. Similarly, in February 2021, workers at an Amazon warehouse in Bessemer, Alabama voted on whether to form a union, which garnered national attention and sparked a broader discussion on the role of unions in modern workplaces.

[25]), job security (Farber and Saks [40]), and productivity (Freeman and Medoff [44], Doucouliagos and Laroche [36] and Hirsch [53]). In contrast, the literature offers little to understand the effects of unionization on firms' decisions about raising capital. Existing literature offers two diverging views about how unionization can affect firms' decisions to raise capital. First, *bargaining device* hypothesis, suggests that unionization incentivize firms to shift the mode of financing in favor of leverage. In contrast, the *crowding-out* hypothesis (Woods et al. [94]), suggests that unionization makes it difficult for firms to raise capital using leverage. The main arguments underlying these two hypothesis are outlined below. The primary goal of this chapter is to revisit the relationship between unionization and firms' financing decisions, using a more carefully constructed datasets and a more up to date empirical methodology.

The *bargaining device* hypothesis (Matsa [70] and Bronars and Deere [23]) proposes that unionization leads firms to rely more on firm leverage as a possible means of raising capital. According to this view, financial leverage serves as a strategic tool to limit the appropriation of rents by workers. The basic idea is that unions and managers engage in negotiations to determine the distribution of future cash flows. Prior to engaging in the bargaining process, managers can increase the firm's leverage and use this to return capital to shareholders (e.g., through a leveraged buyback or special dividend). Such actions enable firms to pay the shareholder larger present dividends against future cash flows. On the other hand, an increase in firms' present debt limits the union's scope to claim a larger share of future profit. Such strategies are also effective for non-unionized firms facing the threat of unionization, as workers can easily organize themselves into a union (Bronars and Deere [23]). Therefore, according to this hypothesis, unionization will increase the firm leverage.

The *crowding-out* hypothesis presents a contrasting view. This perspective starts with the observation that unions impose significant economic costs on firms. These costs include the

union wage premium, which is estimated to be historically around 15% in the United States (Aidt and Tzannatos [5]), as well as costly fringe benefits like severance pay, paid holidays, paid sick leave, and pension plans. In addition, unions limit the manager's discretion in hiring and firing decisions, leading to operating inflexibility. This limits the firms' ability to reduce the negotiated labor costs if needed (Kuzmina et al. [61], Simintzi et al. [85] and Woods et al. [94]). Taken together, the negotiated labor costs with unions act as a super-senior "debt-like" contract, and therefore crowds out firm capacity to service debt. In other words, with unionization, the ability of firms to utilize financial leverage to finance their operations decreases. As a result, according to this hypothesis, unionization lead to a decrease in firm leverage. While both *bargaining device* and *crowding-out* hypotheses appear reasonable, there is a lack of consensus in the existing empirical literature about which of the two hypothesis is relevant in today's economy. The primary goal of this chapter is to seek answers to these questions.

The lack of consensus in the literature is not the only reason for me to revisit the effect of unionization on firms' financing decisions. The seminal paper by Matsa [70] was based on the historical data for the period 1950 to 1970. Since then the U.S labor market has experienced a significant change in the regulatory framework, which hold the potential of moderating the firm's reactions to unionization. The example of such regulations include increasing participation in *Right-to-Work* laws, introduction of chapter 11 bankruptcy and plant-level collective bargaining. In states that have right-to-work laws, employees usually have the freedom to choose whether or not to join a union or financially support the union. Consequently, labor unions would likely have a stronger presence in companies located in states that do not enforce right-to-work laws. Chapter 11 bankruptcy, introduced in 1978, changed the dynamics of labor-management relations by limiting union power in bankruptcy proceedings and reducing the effectiveness of using debt as a bargaining tool in labor negoti-

ations. Similarly, plant-level collective bargaining, coupled with increased capital mobility, has made strategic debt less necessary. Instead, firms can use the threat of plant closures or relocations to extract concessions from unions without resorting to financial debt. In this chapter, I used the most up to date data which holds the potential of including the effect of these changes on how firms react to unionization.

Our study employs a dataset that is compiled from two sources. We obtain union election data for the period 1999 to 2021 from the National Labor Relations Board (NLRB) website. This data provides detailed information on name of firms that held union elections, filing date, election date, closing date, total number of votes, and the number of votes in favor of forming a union. We obtain financial information from Compustat database. More details regarding Compustat data is included in section B.1.2 in the appendix. Compustat data include several firm identifiers, including *permco*[2], *permno*[3] and *gvkey*[4]. Unfortunately, a common identifier is not available in the NLRB dataset. In order to do the analysis, we need to merge the election data from NLRB to the financial information of firms in Compustat. But, we only have access to the company names that may not entered the same way in the two datasets. Therefore, we start by standardizing and cleaning the name strings[5]. Then, we rely on a new matching criteria to link the data on unionized firms and firms held the election to Compustat data using the company names which gives us a final sample of 430 firms. We further explain the details of this approach in section B.1.4 in the appendix.

In our study, we observe that the process of unionization does not happen simultaneously for all firms. Instead, it can be characterized as a staggered treatment, with each firm receiving unionization at a specific point in time. This naturally leads us to consider using a

[2] *permco* is a unique company level identifier that remains unchanged throughout the whole term of company's existence even if the company changed

[3] *permno* is a unique stock (share class) level identifier

[4] (Global Company Key) is a unique number assigned to each company in the Compustat-Capital IQ database.

[5] Our cleaning procedure builds on the `stnd_compname` Stata name standardization programs [93].

standard linear two-way fixed effect model or an event study design as many existing studies in the literature have relied on these methods or regression discontinuity (RD) design to investigate the relationship between unionization and financial leverage. However, recent advancements in the econometrics literature (Goodman-Bacon [46], Sun and Abraham [88], Athey and Imbens [8], Callaway and Sant'Anna [29], Borusyak and Jaravel [21], Gardner [45] and De Chaisemartin and d'Haultfoeuille [34]) have raised concerns about the reliability of causal estimates obtained from a standard staggered difference-in-differences (DID) framework. The consensus among researchers is that the standard approach can yield misleading estimates, even when the parallel trends assumption is met and treatment assignments are random. We take this consensus seriously and address this issue by adopting a solution proposed by Callaway and Sant'Anna [29]. Their approach not only helps reduce potential biases but also provides easily interpretable estimates similar to an event study, allowing us to capture both short-term and long-term effects of unionization. By employing this recent methodology, we aim to estimate the causal effect of unionization on a range of financial performance indicators, including *Debt to Equity ratio, market leverage, book leverage, long-term book leverage, net leverage and cash to asset ratio*, thereby ensuring that our results are not biased under heterogeneous treatment effects and staggered treatment timing.

The main findings of my analysis are summarized below:

1. We find robust support for *crowding out* hypothesis. All the outcome variables such as *Debt to Equity ratio, market leverage, book leverage, long-term book leverage and net leverage* decrease after 6-7 years since getting unionized. This trend continues throughout the period of our analysis.

2. We extend our analysis by examining the effect of labor unions on corporate cash-holding decisions. As in the case of the leverage, there are two competing views about how unionization can affect cash holdings. One view suggest that unionization lead to

higher labor adjustment costs. As a result, the firms increase the cash holdings to meet the precautionary demand for funds in unfavorable states. The other argument suggest that by lowering cash holdings, firms can reduce the pie over which the bargaining takes place between unions and firms[6]. Our results suggest that unionization has a positive effect on cash holdings. Thus, as in the case of leverage, the results run counter to the argument that is based on the *bargaining device* hypothesis.

3. The NLRB dataset include details about the election results. This offers an opportunity to measure variation in the support in favor of forming a union within an establishment. Our findings suggest that firms react strongly and reduce leverage and increase cash holdings more when unions won with a larger margin of support.

4. By examining the effects of unionization on financing decisions separately for one-establishment and multi-establishment firms, we aim to uncover any differential impact based on the organizational structure of the firms. Given that half of our sample includes multi-establishment firms, we are interested to know to what extent the initiative for unionization in one of the franchises can affect the financing decisions of the firms' headquarter. We find stronger results in the case of the firms with multiple

[6]There are two competing hypotheses drawing on the literature that evaluates the effects of labor market institutions, such as unionization on corporate decisions. On the one hand, strong unions are likely to increase labor adjustment costs for firms by making the firing and hiring more difficult and less timely (Autor et al. [9], Millán et al. [72]; Berglund and Furåker [16]). This higher labor adjustment cost implies a greater burden of fixed wage claims for a firm to service even in an unfavorable state, and in turn brings about a higher level of operating leverage and the distress risk (Alimov [6]). As a result, firms facing stringent unions may exhibit a greater precautionary demand for cash to counteract the distress risk associated with labor adjustment costs. We refer to this effect as the "*labor adjustment cost effect*." On the other hand, one might argue that an increase in the stringency of *employment protection laws* (EPLs) could enhance workers' bargaining power within a firm. The increased bargaining power for a firm's employees can then lead to the firm' strategic use of financial policies. For example, prior research documents that firms with unionized workers are more likely to increase their leverage ratio (Bronars and Deere [23]; Matsa [70]) and reduce cash holdings (Klasa et al. [59]; Schmalz [82]) in order to improve their bargaining power over unionized workers. The rationale behind this alternative prediction is that if a firm has a low leverage ratio and a large cash cushion, it may face greater difficulty in refusing the labor union's demands for wage increases. This line of reasoning then yields an alternative prediction that stringent EPLs might encourage firms to lower their cash reserves ("*bargaining power effect*").

franchises.

5. We explore the differential impacts of unionization with respect to the presence of Right-to-Work legislation in the states in which firms operate and the political orientation of states (Democrat-led versus Republic-led according to 2020 presidential election). We know that *right to work* laws can hurt the unionization effects by reducing union funding and weakening its bargaining power. These factors collectively diminish the ability of unions to negotiate favorable terms and conditions for their members. On the other hand, Democratic states, may have a higher proportion of unionized workers and more favorable attitudes toward unions, whereas it is not the case for Republican states. we can better understand how political orientation of the states in which firms are operating can intervene with unionization effects if we examine the results in Democrat-led versus Republic-led states separately. Unionization effects on the outcome variables are more pronounced for the firms in democrat-led/without RTW law states.

6. In the above results, unionization is defined on the basis of exploiting the variation in the timing for unionization among firms which won the first election and whose franchises did not apply for "*decertification*" after the initial win result. However, in our sample nearly 57% of elections have ended with a loss representing the decision of workers against forming a union. Still, firms can read the losses as threat to unionization in the future and alter their current behavior. To examine such possibilities, we make two modifications to our analysis. As the first modification, I take the election date as treatment, regardless of the result of that election. As the second modification, subset of the firms with loss results is considered. In neither of these cases, the results are robust, implying that *threat to unionization* does not play a significant role in firms' financing decisions.

The remainder of this paper is organized as follows. Section 2.2 explains the unionization process and the role of the National Labor Relations Board (NLRB). The data and its sources are described in section 2.3. Section 2.4 offers a detailed overview of the methodology. Section 2.5 discusses the results and includes a series of robustness checks. Section 2.6 concludes with some comments. Figures and tables are reported in section 2.7. Finally, section B provides additional details on data matching and supplementary analyses in the appendix.

2.2 Process of Unionization through NLRB Elections

Most workers in the United States have the right to collective bargaining under the National Labor Relations Act (NLRA). When a union represents a group of workers, the employer is compelled by the NLRA to bargain with the union on working conditions[7]. This often occurs at the establishment level (Traxler [90]). During negotiations, the union may go on strike, or the employer may "lockout" employees in order to put pressure on the opposing party. The National Labor Relations Board (NLRB)[8] is an independent federal agency in the United States that oversees and enforces labor laws related to collective bargaining, union representation, and unfair labor practices. Its primary role in union certification is to conduct representation elections and determine the appropriate bargaining unit for a group of employees seeking union representation. Much of the current policy debate in the United States over organized labor focuses on boosting representation in nonunionized establishments[9] (Young et al. [96]). Our findings speak directly to the potential effects of

[7]During the collective bargaining process, the ultimate objective is to reach an agreement on a contract, which typically outlines the wage and non-wage benefits associated with each job title, procedures for handling grievances, policies for carrying out layoffs, as well as guidelines for promotions Slichter et al. [86].

[8]https://www.nlrb.gov/

[9]For example, the debated *Protecting the Right to Organize (PRO) Act of 2021* aims to enhance the existing Federal laws that safeguard workers' ability to establish a union and collectively bargain for better wages, benefits, and working conditions.

these unionization attempts.

A secret-ballot NLRB election is the principal method for private-sector workers to obtain union representation. Therefore, workers at the establishment start the organizing push, either on their own initiative or as a result of union contact. The first step is to collect cards indicating union support signed by workers in the proposed "bargaining unit" (i.e., the workers the union would represent). In most cases, the bargaining unit consists of workers from a single establishment. It can range from employees in a single occupation (for example, delivery truck drivers) to all non-managerial employees. The process of initiating an election to determine union representation requires the submission of a petition to the nearest NLRB Regional Office, which must demonstrate that at least 30% of employees are interested in either forming or removing/decertifying a union. After the union files an election petition with the NLRB and gathers enough signatures from the bargaining unit, the NLRB checks the cards to ensure there is enough support for the union and settles any disputes regarding the composition of the bargaining unit. Also, NLRB agents conduct an investigation to ensure that the Board has jurisdiction over the matter and that no current labor contracts would preclude the holding of an election. The NLRB then schedules the election. However, employers often try to delay the process in an attempt to decrease union support, such as by contesting the composition of the bargaining unit (Levitt et al. [67]).

Prior to an election, both unions and employers engage in active campaigning either for or against union representation. Pro-union workers and organizers can campaign by conversing with coworkers on the job or during "house calls", showing solidarity publicly (such as through rallies or wearing pro-union attire), and seeking the support of community groups (Bronfenbrenner and Juravich [24]). Employers also have various campaign tools at their disposal, including one-on-one meetings with supervisors, captive audience meetings that require employee attendance, and the hiring of union avoidance consultants and law firms

(Logan [68]). Finally, despite legal restrictions on firing pro-union workers and threatening to close establishments, these tactics still occur (Schmitt et al. [83]). If a majority of workers vote for the union, the NLRB certifies the union to represent the bargaining unit. Following certification, the employer is obligated to bargain in good faith with the union, but they are not required to come to an agreement. If a contract is not reached one year after certification, employees can hold a *decertification* election to vote out the union.

2.3 Background and Data

Our study employs a dataset that is constructed from various sources. The union election data from 1999 to 2021 is retrieved from the National Labor Relations Board (NLRB) website. This data provides crucial details about the establishments that held union elections, such as the company name, filing date, election date, closing date, total number of votes, and the number of votes for the union. We determine one of our treatment times using the earliest election date that is available for each firm. We obtain financial information from Compustat database. More details regarding Compustat data is included in section B.1.2 in the appendix.

We utilize a new matching criteria to link the data on unionized firms and firms held the election to Compustat data using the company names which result in 430 firms. We further explain the details of this approach in section B.1.4 in the appendix. We only include union elections with available election results, and for firms with multiple election records, we retain the first one (as per Bradley et al. [22], Young et al. [96]). Besides, we eliminate contested elections, which are those in which multiple unions are on the ballot. These elections frequently include incumbent unions (e.g., "union raids") and may thus be unrelated to changes in union representation [80]. Furthermore, we eliminate elections with less than six

workers in the bargaining unit to ensure that the election can result in a significant increase in union representation. We only keep the *first* election for each establishment. This means that our estimates should be taken as the effects of winning the first union election at an establishment. Consistent with standard practices in corporate finance research, we exclude financial firms (SIC 6000-6999) and utility companies (SIC 4900-4999) due to their highly regulated industries. Additionally, we drop non-U.S. firms. Our final sample consists of 430 union elections conducted between 2000 and 2021.

The firm-level controls refer to a group of common characteristics found in the literature that affect a company's capital structure. These characteristics are based on various studies by Rajan and Zingales [77], Frank and Goyal [43], Lemmon et al. [66] and Öztekin [75]. The firm-level controls consist of four main factors: SIZE as a proxy for diversification and a firm's default risk; TANGIBILITY as a measure of assets in place and the extent of adverse-selection costs; PROFITABILITY as a proxy for the availability of internal funds, and GROWTH (Tobin's q) as a control for growth opportunities. More details regarding covariates are included in table 2.1.

We construct two variables related to union elections. Union votes (support) is the number of votes in favor of unionization divided by the total number of votes in an election. Passage is the election result dummy which equals 1 when the union votes is greater than 50% and 0 otherwise (Daniel et al. [33]).

We use the ratio of financial debt to book assets (book leverage) as our primary choice for outcome variable. The reason is we are interested in a firm's deliberate choice of leverage rather than changes in leverage caused by market reactions to unionization elections[10]. De-

[10]The detailed reasons for why we focus primarily on book leverage in our analysis are listed below: Firstly, research by Graham and Harvey [47] reveals that managers heavily rely on book values when making capital structure decisions. By focusing on book leverage, we align our analysis with the decision-making framework employed by managers themselves. Secondly, book leverage offers advantages in terms of adjustability compared to market leverage. Book leverage primarily reflects the assets in place rather than the firm's

spite this, the findings on market leverage are consistent with book leverage results. To avoid accounting changes impacting the level of leverage we measure, we use total liabilities as the numerator in most of our leverage calculations. Alternative definitions, on the other hand, have been tested, and the results remain robust to alternative definitions. More details regarding our outcome variables are included in table 2.1.

In our initial analysis, we utilize an alternative measure of leverage similar to Heider and Ljungqvist [52] and Simintzi et al. [85]. We specifically employ the long-term book leverage (long-term debt to total assets). By specifically examining the proportion of debt that is more likely to be sensitive to variations in employment protection laws (EPL)/unionization, such as long-term debt, we can focus on the component of leverage that is most likely to respond to changes in labor-related factors rather than being influenced by factors like working capital requirements. Incorporating long-term book leverage provides us with a more targeted perspective on the relationship between labor power and the specific debt structure of firms.

In table 2.2, we can see the distribution of firms according to the time they held an election for the purpose of unionization. In total, we have 430 firms. Next, in table 2.3, we will examine the distribution of the firms certified as unionized after holding the election. We have 168 such firms in our data.

Tables 2.4 and 2.5, report the summary statistics of the main variables used in our analysis. It is worth noting that the passage rate for union elections is approximately 43%, indicating that approximately half of the union elections in our sample result in union formation.

growth opportunities, as highlighted by Myers [74]. This distinction is crucial because market leverage can be influenced by equity value changes driven by stock returns. Given that labor regulations strengthening labor power might be perceived as negative news by the capital market, leading to a decline in equity value (as observed in studies by Besley and Burgess [19] and Lee and Mas [64]), using market leverage could confound the effects of legal changes on firms' debt usage. In contrast, book leverage remains unaffected by stock price fluctuations, making it a more suitable measure for isolating the impact of labor-related factors on debt decisions.

Furthermore, among the 184 firms with successful union elections (43% of firms), only 168 firms have shown no initiative to "decertify" the union after its formation, which will form one of our treatments. Moreover, 69% of firms are operating in democrat-led states and 36% of firms are located in states with right-to-work laws. Also, 53% (228) of firms only have one establishment whereas the remaining 47% (202) of firms include multi-establishments.

2.4 Methodology

In our analysis, we determine the treatment timing based on the date when the election for unionization were held in the firms, which was the case for all firms in our sample but not simultaneously. This staggered adoption setting is appropriate for employing either a static linear two-way fixed effect (TWFE) model or an event study design. However, recent advancements in the literature have highlighted concerns and provided valuable insights regarding these designs. For instance, it is now understood that the causal parameter in a TWFE design represents a weighted average of all possible 2x2 Difference-in-Differences (DID) estimators comparing different timing groups. Additionally, in a TWFE model, previously treated units serve as comparison units. Consequently, even when the parallel trend assumption holds and treatment assignment is random, the causal estimates in a TWFE model can be misleading due to the presence of treatment effects heterogeneity. Furthermore, both static and dynamic TWFE models fail to provide reliable estimates of dynamic causal effects when there is heterogeneity across cohorts.

To address these issues, we adopt the methodology proposed by Callaway and Sant'Anna [29] (CS 2021). This methodology allows us to use "not yet treated" units as controls, instead of only "never treated" units, which maximizes the use of our data and provides reliable estimates. We focus on the average treatment effects on the treated ($ATT(g,t)$), using group-time specific average treatment effects as building blocks.

The CS 2021 takes a ground-up approach, using group-time specific average treatment effects on the treated, $ATT(g,t)$, as the building blocks. The groups are created according to when the units received (absorbing) treatments in the sample. The $ATT(g,t)$ measures the average treatment effect at time t for the group first treated in time $g < t$, and is defined as

$$ATT(g,t) = \mathbb{E}\left[Y_t(g) - Y_t(0) \mid G_g = 1\right] \tag{2.1}$$

where G_g is a dummy variable equal to one if the unit is in treatment time group g. $Y_t(g)$ is the outcome variable at time t for treated units if they were to first become treated in time period g, and $Y_t(0)$ is the potential outcome for the treated units had they not been treated. In the absence of data on $Y_t(0)$, the identification strategy relies on control groups consisting of only those units which have not received treatment up to the time g. We assume no anticipation before the year of the election.

The CS 2021 methodology estimates $ATT(g,t)$ using a doubly robust approach that combines the outcome regression (OR) approach and the inverse probability weighting (IPW) approach. It involves weighting the difference in outcomes between treated and control groups by the propensity score and estimating the conditional expectation function for the comparison group. The $ATT(g,t)$ estimator is doubly robust, meaning that it remains valid as long as either the outcome evolution for the comparison group or the propensity score model is correctly specified.

$$ATT(g,t) = \mathbb{E}\left[\overbrace{\left(\frac{G_g}{\mathbb{E}\left[G_g\right]} - \frac{\frac{p_{g,t}(X)(1-D_t)(1-G_g)}{1-p_{g,t}(X)}}{\mathbb{E}\left[\frac{p_{g,t}(X)(1-D_t)(1-G_g)}{1-p_{g,t}(X)}\right]}\right)}^{\text{Inverse Probability Weight}} \overbrace{\left(Y_t - Y_{g-1} - m_{g,t}(X)\right)}^{\text{Outcome Regression}}\right] \tag{2.2}$$

To implement this methodology, we divide the sample period into sub-intervals from 2000

to 2021 and assign a group identifier to each firm based on its election/unionization year. We estimate $ATT(g, t = g + e)$ for different values of e to capture unionization effects at different time intervals relative to the treatment period. We then construct weighted averages of the group $ATTs$ for each value of e and aggregate them to create an overall point estimate. To test the parallel trend assumption, we use formal inference and CS 2021 recommended bootstrapping procedure to obtain simultaneous confidence bands robust to multiple hypothesis testing and individual cluster errors.

The details of the methodology can be found in Section 1.3 of the first chapter. We utilize the publicly available `R` code provided by Callaway and Sant'Anna [29] as a supplement to their research[11].

2.5 Results

The goal of this chapter is to investigate the impacts of workers' decision to become unionized on firms' financing and corporate cash holdings decisions. We also consider other types of treatments such as workers' initiative to hold elections which lead to win result and workers' lack of initiative to decertify a union after it is formed. The results pertaining to these different types of treatments collectively paint the picture about how firms' mode of financing is affected by unionization.

The primary dependent variable in the study is the ratio of book debt to the book value of total assets (*book leverage*), but we also provide estimates for four alternative measures of leverage, such as *long-term book leverage, net leverage, market leverage and debt to equity ratio*. In addition, we consider the effect of unionization on *cash to asset ratio*. Collectively,

[11]For more comprehensive information, please refer to Callaway and Sant'Anna [29] and the specified section for detailed explanations.

analysis of this set of financial leverage and corporate cash holdings outcome variables, will give us a better picture regarding the effect of unionization on firms' financing decisions. Table 2.1 summarizes details regarding all financial leverage proxies and the key covariates.

Since our empirical design compares the outcomes for elections in different time periods with similar baseline characteristics, we have reviewed the existing literature on selection into holding and winning elections. This information has helped us identify which baseline characteristics to condition on. For selection *into* elections, Dinlersoz et al. [35] find that establishments that are larger, more productive, and younger are more likely to hold elections. To address this potential selection bias, we have limited our comparisons to establishments that have already held elections.

Since the result of the election in favor or against unionization plays a significant role in addressing our question of interest, in section 2.5.1, we begin our analysis with the benchmark case in which we explore the results for the subset of firms hold elections with 'win' and 'loss' results separately. Next, within the subset of firms with 'win' result, we are interested to investigate whether firms which received a larger support for unionization will behave differently compared to their counterparts with smaller support for unionization. The reason for this analysis is the possibility of "vote-share manipulation" around the 50% threshold and reduced unions' bargaining power due to debates about contested votes that can result in long delays before bargaining with the firms in close elections. These factors can imply that the effects of elections with close number of votes in favor and against unionization might differ from those of higher vote-share elections. The results of this analysis is provided in section 2.5.2.

Since in our benchmark analysis, we fail to acknowledge the possibility that employees may lose their union representation through a *decertification* election at a later point in time, next, we extend our analysis by considering a subset of firms which hold the elections with

'win' result and there is lack of initiative to decertify a union after it is formed. Here, we consider the 'closing date' instead of 'election date', as the treatment year since this is the date on which the results of elections are announced (refer to section 2.5.3 for more details). We explore a similar analysis on the role of margin of support for unionization for this subset of firms as well.

We also conduct subgroup analyses to explore potential heterogeneity in the effects of unionization on financing decisions of the firms. By examining the effects of unionization separately for one-establishment and multi-establishment firms, we aim to uncover any differential impact based on the organizational structure of the firms. Given that half of our sample includes multi-establishment firms, we are interested to know to what extent the initiative for unionization in one of the franchises can affect the financing decisions of the firms' headquarter. We know that *right to work* laws can hurt the unionization effects by reducing union funding and weakening its bargaining power. These factors collectively diminish the ability of unions to negotiate favorable terms and conditions for their members. As a result, we estimate the effects of unionization on financial leverage indicators for firms which are located in states that have/have not Right-to-Work legislation in place separately. Furthermore, since Democratic states, may have a higher proportion of unionized workers and more favorable attitudes toward unions, whereas it is not the case for Republican states, we can better understand how political orientation of the states in which firms are operating interact with unionization if we examine the effect of unionization in each set of states, i.e., Democrat-led versus Republic-led according to 2020 presidential election, separately. The results for our subgroup analysis are included in section 2.5.4.

2.5.1 Effects of Forming Unions

We start our analysis with the benchmark case where the treatment is synonymous with a win election result (vote share in favor of unionization, of more than 50%). In practice, however, a union can be *decertified* at a later date. We consider such possibilities later in our analysis. Furthermore, unions may hold multiple elections in different franchise of a firm, but the analysis only takes into account the first election of a firm. After winning the election, the union and firms need to bargain in good faith. Consequently, here we explore the dynamic effects of winning the first union election, which may not necessarily correspond directly to being represented by a union.

Figure 2.2, plots the estimates for all the financial leverage variables in addition to log cash to asset ratio. This event study plot include pre- and post-treatment estimates with 95% simultaneous confidence bands. The pre-treatment estimates offer information about the validity of the (conditional) parallel trend assumption. In Figure 2.2, the $ATTs$ prior to treatments are indistinguishable from 0, rendering support to the (conditional) parallel trend assumption. Therefore, the estimates presented in Figure 2.2 indicates that firms that had successful union elections had similar conditional pre-election log of financial leverage and cash to asset ratio compared to those who are yet to be treated. Our analysis controls for several firm specific variables such as, size, tobin's q, tangibility and profitability that we observe at the time of treatment. We use these characteristics as covariates for the full sample. However, for some sub-samples, insufficient variation in the data prevents us from executing a meaningful outcome regression component of the equation (2.2). For these cases, we use a subset of the conditioning variables and we rely on inverse probability weighting (IPW) approach, instead of doubly robust (DR) approach.

The effects of holding an election with 'win' result become visible with time. After 6-7 years from holding the election, the effects start to become visible and firms experience significant

decline in their financial leverage outcome variables, such as *net leverage, debt to equity ratio, market leverage* and noticeable increase in their cash to asset ratio on average. We experience a loss of (not-yet-treated) comparison units as we move forward in time. Despite this, when the firms experience successful election, their financial leverage outcome variables continue to decline and for the case of the cash to assets ratio, it continues to increase throughout the period of our analysis. As an example, nine years after a successful union election, the log of net leverage is around 0.2 lower, than firms which are not yet hold an election (consistent with 19% decrease in net leverage). The estimates on cash to assets ratio show that five years after an election, establishments with successful elections are four percent more likely to increase their cash to asset ratio, and this effect increases slightly to ten percent after fifteen years. The findings of the study hold true regardless of the measure of leverage used. We find strong negative and significant effect of unionization on financial leverage variables. We show that baseline results hold across different operational definitions of leverage, such as market leverage, debt to equity ratio and net leverage.

Since the result of the election in favor or against unionization plays a significant role in addressing our question of interest; as a robustness check, we redo the analysis with the subgroup of firms which hold an election with 'loss' result (vote share in favor of unionization, of less than 50%). According to the results, as depicted in Figure 2.3, the effects of unionization is not evident in the subgroup of elections with "Loss" results, which highlights the point that *threat to unionization* does not appear to have a significant effect on the financing decisions of firms and further provide support to our previous results.

2.5.2 Effects of Forming Unions with Large Margin of Support (Successful Elections)

In this section, we want to explore the effects of unionization for firms that have received a large support for unionization compared to their counterparts with relatively low support in favor of unionization. We define the level of union votes *support* variable as the number of votes in favor of unionization divided by the total number of votes in an election. For the purpose of this analysis, we focus only on firms that held an election resulting in a "win" ($support > 0.5$). Within this subset of firms, the mean and median of the *support* variable are 0.70 and 0.65, respectively, with 25% of the firms receiving a support in favor of unionization of less than 0.57. We refer to firms that hold an election and received less than 0.57 of the total votes in favor of unionization as *Low Support* category. On the other hand, *High Support* subset, includes those firms that received support for unionization of greater than 0.57. The reasons for exploring the effects of Larger Margin-of-Support Elections are listed below (Young et al. [96]).

1. Generally, receiving a support of greater than 0.50 in favor of unionization in an election is enough to be recognized as a union. However, in this case, we consider the effects of unionization for a subset of firms that received a *support* level higher than 0.57 separately to guard our approach against the criticism that is usually directed to "Regression Discontinuity Design" when there is evidence of non-random sorting of elections around the 50% threshold, which is often referred to as "vote-share manipulation." Figure 2.1, depicts the vote-share distribution for the elections in our sample and highlights the previously documented manipulation around the 50% threshold by Frandsen [42].

2. Another motivation for our analysis in this section is to examine the potential depen-

dence of the treatment effect of unionization on the vote share in the election. Previous research, such as the study conducted by Lavetti et al. [62], has shown that the negative effects of unionization on the stock market are more pronounced in elections with higher margins of victory. One possible explanation for this variation is that close union elections are often followed by long delays before the bargaining process begins, which may involve debates and disputes over contested votes. These delays can potentially reduce the bargaining power of unions. Therefore, the effects of close elections may differ from those of elections with a higher vote share.

3. Moreover, there are other factors that contribute to the distinct effects of close elections compared to elections with higher margins of victory on firms. One key factor is that firms may anticipate the possibility of a future decertification election and respond to close elections by intentionally delaying the bargaining process. This strategic delay can weaken the bargaining power of unions, thereby leading to different outcomes. Consequently, establishments with higher margin-of-victory elections have a greater chance of successfully reaching their first contracts, which can lead to more substantial changes.

4. Finally, unions that achieve a higher level of support in elections have more bargaining power, as they can more credibly threaten to strike. For example, in the manufacturing industry, the probability of a post-election works stoppage increases with the election vote share. These results suggest that multiple indicators of unions' bargaining power, increase as the share of votes in favor of unionization rises, implying that the effects of unionization may differ based on the election vote share, highlighting the importance of this subgroup analysis.

Our results (Figures 2.4 and 2.5) indicate that the impact of unionization on corporate fi-

nancial decisions is more pronounced for firms with higher margins of support in favor of unionization. Specifically, we found a stronger and statistically significant effect of unionization on financial leverage and corporate cash holdings for the higher margin of support subgroup, compared to their counterparts with lower margin of support. This suggests that the bargaining power of labor resulting from unionization has a more significant impact on the financing decisions of firms when the support for unionization is stronger among employees.

Similar to the benchmark case, the effects of a successful election with higher margin of support becomes evident over time, particularly around 6-7 years after the election takes place. Firms tend to experience a significant decline in various financial leverage indicators, such as net leverage, debt to equity ratio, and market leverage. Additionally, there is a significant increase in the cash to asset ratio on average. As we progress in time, it is important to note that we encounter a loss of comparison units that have not yet undergone treatment. However, despite this loss, when the firms experience a successful election with higher margin of support, their financial leverage indicators continue to decline throughout the period of our analysis. Conversely, the cash to assets ratio consistently increases over time. These findings suggest that the consequences of a successful election with high support in favor of unionization gradually manifest themselves in firms' financial leverage outcomes. The decline in financial leverage indicators indicates a potential reduction in firms' reliance on debt financing, while the increase in the cash to asset ratio suggests an accumulation of cash reserves. Furthermore, according to Figure 2.4, the $ATTs$ prior to treatments are indistinguishable from 0, rendering support to the (conditional) parallel trend assumption.

The implications of our findings are relevant for both labor relations and corporate finance. For labor relations, our results suggest that unions may have a more substantial impact on bargaining power and worker outcomes when there is a strong base of support among

employees. This highlights the importance of union organizing efforts and building strong relationships with workers to achieve successful collective bargaining outcomes. From a corporate finance perspective, our findings have implications for firms' capital structure decisions. The negative relationship between unionization and financial leverage suggests that firms facing unionization threats may choose to rely less on debt financing to maintain their financial flexibility. On the other hand, the positive relationship between unionization and corporate cash holdings indicates that unionized firms may hold more cash as a precautionary measure against potential labor disputes and strikes. These findings contribute to the growing literature on the effects of labor on corporate financial decisions and have important implications for policymakers, firms, and labor unions.

2.5.3 Unionization Not Followed by Decertification

Since, in our previous analysis on successful elections, the underlying assumption fails to acknowledge that employees may lose their union representation through a decertification election and the possibility of holding multiple elections in other franchises after holding the initial election, one could argue that these mechanisms have the potential to confound our results. To mitigate this concern, in this section, we have designated the first close date, which is when the union is officially certified, for a successful unionization petition[12] as the treatment year. Here, we define a stricter version of unionization and have only considered those firms as treated where there has been workers' lack of initiative to "decertify"[13] a union after it is formed.

When we include the elections with win results combined with no later removal of unionization from any of the franchises, according to Figure 2.6, we find that the effects of union-

[12]known as the "RC type"
[13]known as the "RD type"

ization on financial leverage and cash holdings are stronger compared to relying only on subgroup of firms in which the first election had a win result. This result suggests that the effect of unionization on financial decisions of firms is not only due to the initial shock of unionization, but also due to the ongoing presence of the union. The ongoing presence of the union may continue to influence the bargaining power of labor and the financial decisions of the firm.

Similar to the benchmark case, the effects of a successful election with lack of workers' initiative to decertify a union after it is formed, gradually become apparent in firms' financial leverage outcomes, particularly around 6-7 years following the unionization. After that, firms tend to experience a significant decrease in various financial leverage indicators, such as net leverage, debt to equity ratio, and market leverage. Additionally, there is a notable increase in the cash to asset ratio on average. As we move forward in time, it is important to acknowledge that we encounter a loss of comparison units that have not yet undergone treatment. Nevertheless, despite this loss, when the firms experience a successful election combined with lack of workers' initiative to decertify a union after it is formed, their financial leverage indicators continue to decline on average throughout the period of analysis. Conversely, their cash to asset ratio consistently rises over time. The decline in financial leverage indicators implies a potential reduction in firms' reliance on debt financing, while the increase in the cash to asset ratio indicates a buildup of cash reserves. Furthermore, as depicted in Figure 2.6, the pre-treatment average treatment effects (ATTs) are statistically indistinguishable from 0, supporting the (conditional) parallel trend assumption. This finding provides further support for the validity of the analysis, indicating that the groups being compared exhibited similar trends prior to the treatment.

Effects of Forming Unions with Large Margin of Support (Without Decertifying the Union)

In this section, our aim is to investigate the effects of unionization (successful election with lack of workers' initiative to decertify a union after it is formed) on firms that have obtained significant support for unionization, in contrast to firms with relatively low support. For our analysis, we specifically focus on firms that held an election resulting in a "win" combined with lack of workers' initiative to decertify a union after it is formed where the support for unionization exceeds 0.5. Within this subset of firms, we observe that the mean of *support* variable is 0.72, and the median of *support* variable is 0.68. Additionally, approximately 25% of the firms in this subset received a support in favor of unionization lower than 0.59. To categorize the firms based on their level of support, we distinguish between the "Low Support" category and the "High Support" subset. The "Low Support" category includes firms that held an election and received less than 0.59 of the total votes in favor of unionization. Conversely, the "High Support" subset comprises firms that obtained a support for unionization greater than 0.59. By dividing the firms into these distinct categories, we can examine the differential effects of unionization based on the level of support received.

Similar to our previous discussion, based on figures B.3 and B.4 in the appendix, we find stronger and more significant results for the subgroup of unionized firms with *higher margin of support*.

2.5.4 Heterogeneous Treatment Effects (Analysis of Subgroups)

Effects of Labor Unions in One-Establishment versus Multi-Establishment Firms

So far, in our sample, we did not differentiate between multi-establishment (or multi-unit, MU) firms vs single-establishment (or single-unit, SU) firms. However, there is a need to

differentiate this two groups and estimate whether the effects of unionization are different between multi-establishment that account for 48% of unionized firms (successful election combined with workers' lack of initiative to decertify a union after it is formed) and single-establishment firms. The rationale behind this analysis is that a couple of multi-establishment firms include a large number of franchises and the initiative in favor of unionization in one franchise may not significantly affect the financing decisions of the firms' headquarters. As a result, we are interested to investigate to what extent the unionization in one franchise can affect the financing decisions of the firms' headquarters. Our method entails defining "an election at a MU firm" depending on whether the firm that owns the establishment has at least one other establishment under its control during the period of our analysis. Figures 2.7 and 2.8, plot the estimates for all the financial leverage variables in addition to log cash to asset ratio for *Multi-Establishment* versus *One-Establishment* firms, respectively.

Based on Figure 2.7, the subgroup of *Multi-Establishment* firms, experienced significantly large relative declines in financial leverage levels and stronger increase in cash to assets ratio after the unionization compared to the subset of *One-Establishment* firms. Over time, for the subset of *Multi-Establishment* firms, the effects of a successful election combined with workers' lack of initiative to decertify a union gradually become evident in firms' financial leverage outcomes, particularly around 6-7 years after the unionization process. Despite the loss of comparison units as we move forward in time, after unionization, the financial leverage indicators continue to decrease and cash to asset ratio consistently rises throughout the period of our analysis. As an example, ten years after unionization, the log of Long-term book leverage is approximately 0.15 lower, than firms which are not yet unionized (consistent with 14% decrease in Long-term book leverage). The estimates on cash to assets ratio show a stronger positive and significant increasing pattern as well. On the other

hand, the corresponding effects of unionization are mostly insignificant for the subset of *One-Establishment* firms according to Figure 2.8. Also, there is no evidence of differential pre-unionization evolution of outcome variables for either group. Moreover, for the subset of *Multi-Establishment* firms, we observe stronger declines in *book leverage* and *long-term book leverage* as a result of unionization that start to show up earlier compared to the analysis for the full sample of unionized firms, that accordingly highlight the direct more aggressive choice of firms on how to rely less on debt. The decline in financial leverage indicators implies a potential reduction in firms' reliance on debt financing, while the increase in the cash to asset ratio indicates a buildup of cash reserves.

The analysis demonstrates that the financial leverage decreases and cash holdings increases caused by unionization are substantially greater at multi-establishment enterprises, albeit the results for cash holdings are still significant at single-employer firms. This conclusion shows *Multi-Establishment* may feel a greater need to respond strongly to unionization due to concerns that it would spread to their other locations. Multi-establishment firms have operations in multiple locations or regions.

So far, our findings reveal strong and statistically significant negative unionization effects on financial leverage variables and positive effect for the case of the cash to asset ratio, within the unionized firms (successful election and lack of workers' initiative to decertify a union after it is formed). Since this sample includes both one-establishment and multi-establishment firms, a valid concern arises regarding the extent to which unionization in one establishment can impact the overall performance of a firm's headquarters, particularly when the firm has a high number of franchises. Within the subset of *multi-establishment* unionized firms, the mean and median of the *Number of franchises* are 7.7 and 3, respectively, with 75% of the firms having less than or equal to 6 number of franchise. We refer to firms that have greater than 6 number of franchises as "High number of franchises" group. As a

robustness check, we have excluded this group of high number of franchises, resulting in our final sample consists of unionized firms with fewer than 7 number of franchises. This subset includes both *one-establishment* and *multi-establishment* firms with less than 7 franchises.

Figure 2.9 presents the estimates for all the financial leverage variables, as well as the logarithm of the cash-to-asset ratio, focusing solely on the subset of unionized one-establishment firms and multi-establishment firms with less than seven franchises. It is important to note that while the effects observed in this subgroup are relatively smaller compared to the broader category of unionized firms, they remain both significant and robust. Despite the diminished magnitude of the effects, our results confirm the importance of exploring the subset of multi-establishment firms versus one-establishment firms separately.

Effects of Labor Unions in States With or Without Right-To-Work Laws

Up to this point, we have not distinguished between firms operating in states with and without right-to-work (RTW) laws. Nonetheless, it is crucial to examine the impact of unionization on the financing decisions of firms within each group of states separately. Right-to-work (RTW) legislation, permits employees to benefit from the actions of a union without being contractually compelled to join or financially support it (Bradley et al. [22]). As a result, RTW creates a "free-rider" issue that reduces the funding available per worker covered by the union, limiting the union's ability to contract high-quality experts to assist in negotiations. Therefore, the bargaining power of union is reduced (Holmes [55], Matsa [70]). However, this legislation also reduces the bargaining power of non-organized labor by reducing the *threat* of unionization, allowing firms to pay their non-organized workers lower wages. The passage of RTW is often seen as a weakening of both union power and state labor rights. We obtain information on the year that each state adopted the RTW legislation from the Department of Labor.

Therefore, there is a need to explore the effects of unionization on financial leverage variables and cash to assets ratio for firms that operate in states with and without right-to-work laws separately. Figures 2.11 and 2.12, plot the estimates for all the financial leverage variables in addition to log cash to asset ratio for the subset of firms that operate in states without RTW and with RTW, respectively. The estimates presented in Figure 2.11 for firms that operate in states without RTW laws, indicate that firms that had successfully became unionized, had similar conditional pre-election log of financial leverage and cash to asset ratio compared to those who are yet to be treated. However, they experienced large significant declines in financial leverage levels and increase in cash to assets ratio after the unionization compared to the subgroup of firms that operate in states with RTW laws.

The estimates presented in Figure 2.11 for firms that operate in states without RTW laws, indicate that for firms that had successfully became unionized, the effects, become visible with time, particularly around 6-7 years following the unionization. After that, firms tend to experience a significant decrease in financial leverage indicators and increase in the cash to asset ratio that are in place for the period of our analysis. As an example, seven years after unionization, the log of long-term book leverage is approximately 0.11 lower, than firms which are not yet unionized (consistent with 11% decrease in long-term book leverage). For the case of firms that operate in states without RTW laws, the decline in financial leverage indicators implies a potential reduction in firms' reliance on debt financing, while the increase in the cash to asset ratio indicates a buildup of cash reserves. Furthermore, as depicted in Figure 2.11, the pre-treatment average treatment effects (ATTs) are statistically indistinguishable from 0, supporting the (conditional) parallel trend assumption. This finding provides further support for the validity of the analysis, indicating that the groups being compared exhibited similar trends prior to the treatment. On the other hand, the corresponding effects of unionization did not show up that much for the subgroup of firms in states with RTW laws

according to Figure 2.12. As a summary, in states without Right-to-Work laws, unions have a stronger position and can exert greater pressure on firms to provide higher wages and benefits. This increased bargaining power may lead to a more pronounced crowding-out effect on financial leverage.

Effects of Labor Unions in States With Different Political Orientation (Democrat Versus Republic)

So far, we did not differentiate the firms that operate in democrat-led versus republic-led states. However, it is important to explore the effect of unionization on financing decisions of firms in each set of states, separately. Unionization rates and attitudes toward unions may differ between Democratic and Republican states. Democratic states, for example, may have a higher proportion of unionized workers and more favorable attitudes toward unions, whereas Republican states may have a smaller number of unionized workers and unfavorable attitudes toward unions. Besides, when it comes to policy differences, Democrat and Republican states may have differing union and labor rights rules and regulations. We can better understand how political orientation of the states interact with unionization if we examine the effect of unionization for firms located in each set of states based on political orientation separately. According to Figure 2.13, we divide states to "Republic" versus "Democrat" using the result of 2020 presidential election.

There is a high overlap between Republic states and states with *right to work* laws. According to Feigenbaum et al. [41], right-to-work laws cut Democratic Presidential vote shares by 3.5 percentage points. They investigate the mechanics and discover that right-to-work legislation reduce organized labor contributions to Democrats and make potential Democratic voters less likely to be approached to vote.

Therefore, we estimate the effects of unionization on financial leverage variables and cash to assets ratio for democrat versus republic states separately. Figures 2.14 and 2.15, plot the estimates for all the financial leverage variables in addition to log cash to asset ratio for the subsample of democrat and republic states, respectively. The estimates presented in Figure 2.14 indicates that for firms that had successfully became unionized, the effects, become visible with time, particularly around 6-7 years following the unionization. After that, firms tend to experience a significant decrease in financial leverage indicators and increase in the cash to asset ratio that are in place for the period of our analysis. As an example, nine years after unionization, the log of debt/equity is approximately 0.5 lower, than firms which are not yet unionized (consistent with 39% decrease in debt/equity ratio). For the case of firms that operate in democrat-led states, the decline in financial leverage indicators implies a potential reduction in firms' reliance on debt financing, while the increase in the cash to asset ratio indicates a buildup of cash reserves. Furthermore, as depicted in Figure 2.14, the pre-treatment average treatment effects (ATTs) are statistically indistinguishable from 0, supporting the (conditional) parallel trend assumption. This finding provides further support for the validity of the analysis, indicating that the groups being compared exhibited similar trends prior to the treatment. On the other hand, the corresponding effects of unionization did not show up for the subsample of republic states according to Figure 2.15. Since Democrat-led states generally have a more favorable stance towards labor unions and workers' rights compared to Republican-led states, these states may have stronger labor protection laws, which could enhance the bargaining power of unions and their ability to negotiate favorable labor contracts. As a result, the crowding-out effect on financial leverage may be more significant in these states due to the higher level of unionization and labor influence.

2.6 Conclusion

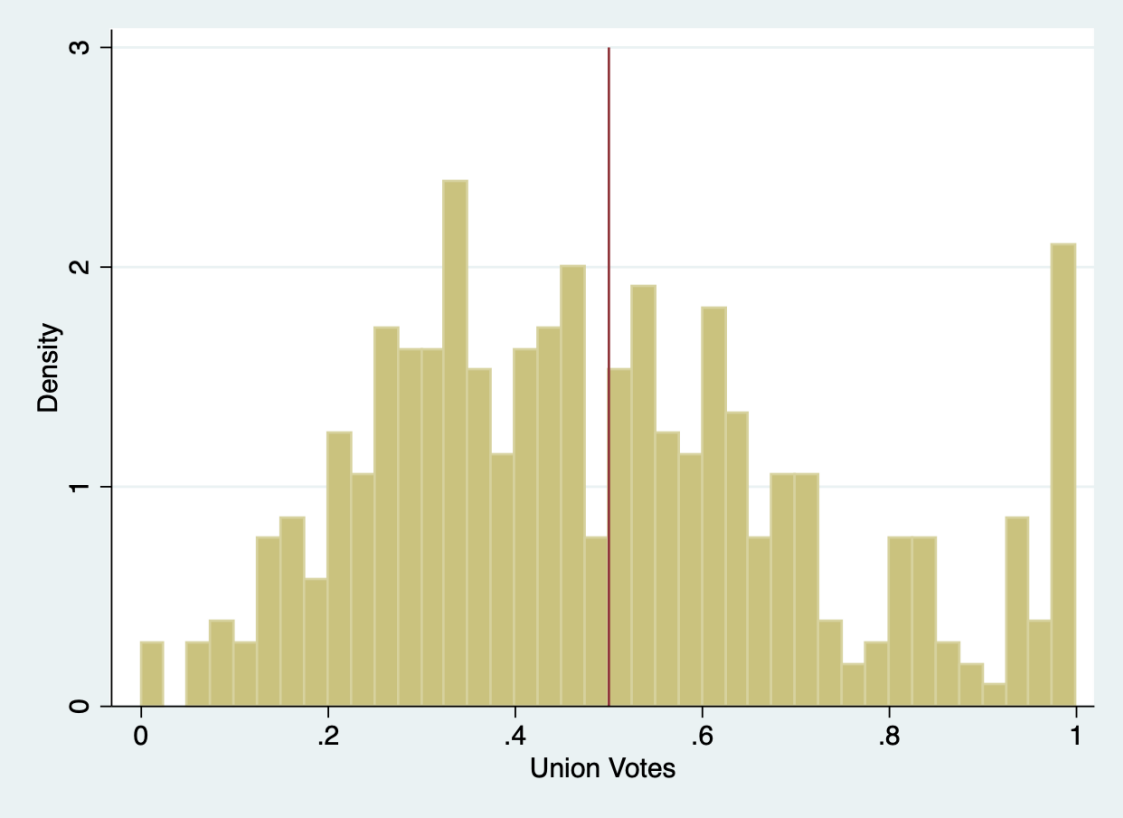
Despite the decrease in trade union membership in the United States, labor unions still have a potentially significant influence on financial policies. This chapter offers new evidence on the impact of unionization on the financing decisions of firms in the United States. Our identification strategy involves exploiting exogenous variation in the timing of unionization elections. We find robust support for *crowding out* hypothesis. All the outcome variables such as *Debt to Equity ratio, market leverage, book leverage, long-term book leverage and net leverage* decrease after 6-7 years since getting unionized. Also, we find that the effects of unions on cash holdings are positive. Moreover, our analysis shows the potential heterogeneity in the effects of unionization by exploring the effects of unionization on financial leverage and cash holdings within different subgroups, such as one-establishment versus multi-establishment firms, and Democrat-led versus Republican-led states. Collectively, effects on the outcome variables are more pronounced for the firms in democrat-led/without RTW law states, multi-establishment firms and when the support in favor of unionization is stronger among employees. Overall, our study contributes to the ongoing discussion on the role of unions in the labor market and their impact on corporate finance decisions.

2.7 Figures and Tables

Table 2.1: Variable Definitions (Financial Leverage Outcome Variables, Cash to Assets Ratio and Covariates)

Firm-level Variables	**(Compustat item names in the parentheses where applicable)**
Outcome Variables	
Book leverage	Long term debt (dltt) plus debt in current liabilities (dlc), divided by total assets (at)
Long-run book leverage	Long term debt (dltt), divided by total assets (at)
Net Leverage	Long-term debt (dltt) plus debt in current liabilities (dlc) minus cash (che), divided by total assets (at)
Debt to Equity	Long term debt (dltt), divided by the market value of common equity (csho*prcc_f)
Market leverage	Total liabilities (lt), divided by the market value of common equity (csho*prcc_f)
Cash/Assets	Cash and Short-Term Investments (che) divided by total assets
Covariates	
Size	Natural logarithm of total assets.
Tobin's q	Book value of assets (at) minus book value of common equity (ceq) plus the market value of common equity (csho*prcc_f) scaled by total assets (at)
Fixed Assets/Assets (Tangibility)	Fixed assets (ppent) divided by total assets (at)
ROA (Return on Assets)	Earnings before interest and tax (ebit) divided by total assets (at)

Figure 2.1: Distribution of Votes in Favor of Unionization



Notes: The figure presents distribution of vote-shares for the entire sample. It has 40 equally spaced bins for vote-shares. In order for the regression discontinuity (RD) design to be valid, there must be no deliberate manipulation of election outcomes right around the 50% vote-share threshold. In other words, there must not be systematic sorting of firms, within close proximity to that threshold. Such sorting would be visible from a discontinuity in the vote-share distribution at the 50% vote-share threshold.

Table 2.2: Distribution of Election Timings (Based on Merged Data from NLRB and Compustat)

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Number of Firms	57	50	42	34	32	21	19	18	11	10	15
Year	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
Number of Firms	11	12	13	13	14	15	13	5	12	5	8

Table 2.3: Distribution of Unionization Timings (Based on Merged Data from NLRB and Compustat)

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Number of Firms	23	15	14	14	9	7	5	8	2	1	7
Year	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
Number of Firms	6	3	5	10	10	10	5	2	7	2	3

Table 2.4: Summary Statistics Using the NLRB Information (Based on Merged Data from NLRB and Compustat)

	count	mean	sd	min	max	p25	p50	p75
closing_method	430	4.583721	.4935152	4	5	4	5	5
num_elig_employees	429	115.4872	256.0753	6	4064	19	43	116
N_franchise	430	2.927907	6.594616	1	119	1	1	3
support	426	.4895376	.2371151	0	1	.3103448	.4571429	.625
democrat_indicator	299	.6856187	.4650473	0	1	0	1	1
rtw_indicator	299	.3578595	.4801742	0	1	0	0	1
N	430							

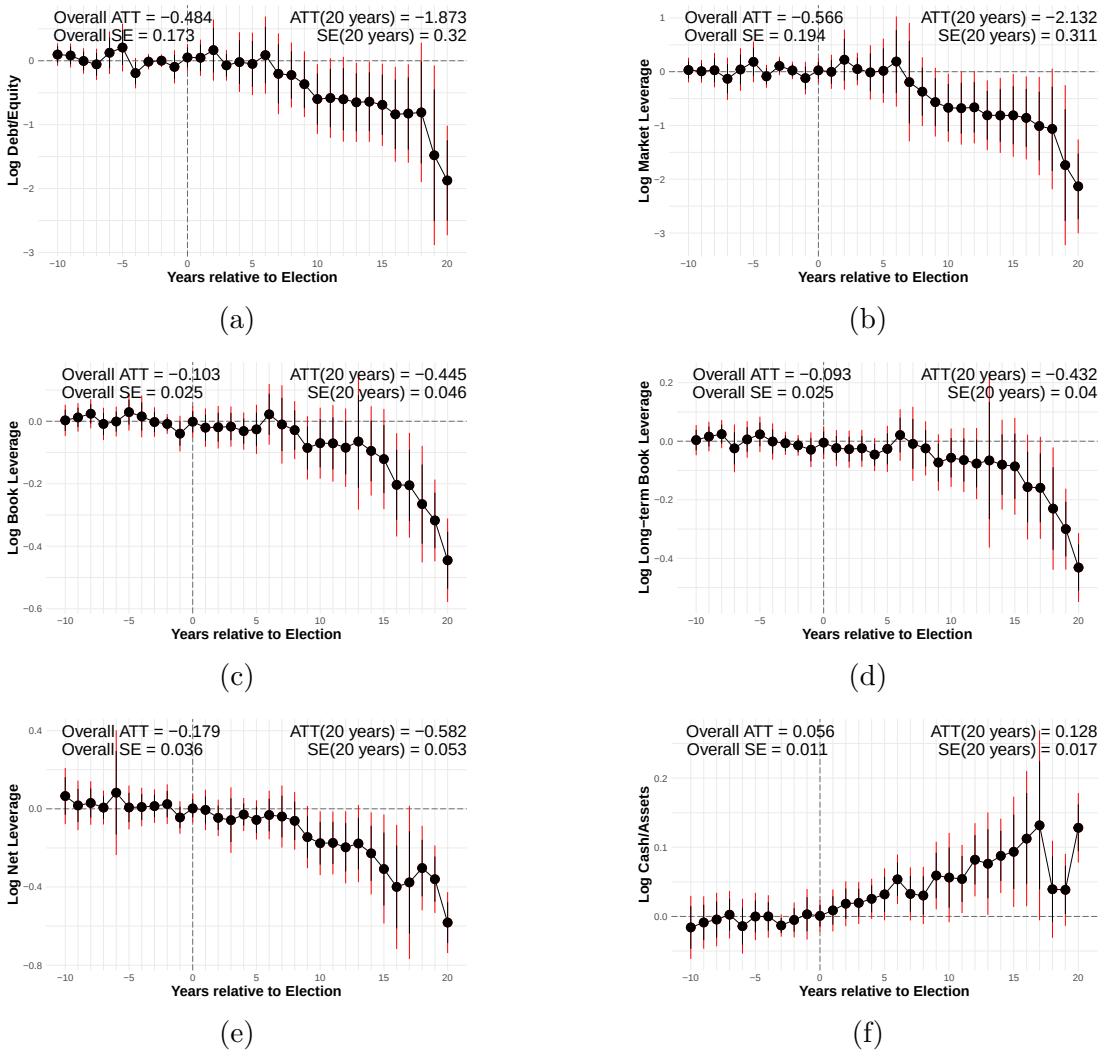
Notes: This table presents the summary statistics for several variables obtained from NLRB (National Labor Relations Board) data, along with the fraction of firms operating in states led by Democrats and states with "right to work" laws. This table includes the summary statistics for a couple of variables from NLRB data and fraction of firms operating in democrat states as well as states with *right to work* laws. *closing_method* captures the outcome of the unionization election. A value of "4" indicates a win or certification of a representative, while "5" represents a loss or certification of the results. *num_elig_employees* indicates the number of eligible voters in the unionization election. It reflects the total count of employees who had the right to participate in the voting process. *N_franchise* represents the number of franchises associated with each firm. *support* is calculated by dividing the number of votes in favor of unionization by the total number of eligible voters. It provides a measure of the level of support for unionization among the employees. *democrat_indicator* is a binary variable that takes a value of one if the state in which the firm operates is led by Democrats. It is equal to zero if the state is led by Republicans, according to the 2020 presidential elections. Finally, *rtw_indicator* is a dummy variable that equals to 1 if the state in which the firm is located, has right to work laws and 0 otherwise.

Table 2.5: Summary Statistics: Status of Financial Leverage Outcome Variables and Covariates in 2000

	count	mean	sd	min	max	p25	p50	p75
size	329	6.558091	1.914808	.7026024	12.9877	5.356931	6.615038	7.765768
tobin_q	284	1.784385	1.26307	.4978638	7.820542	1.00125	1.340579	2.018938
ebit_at	327	.07921	.1400456	-1.435022	.4173743	.0491357	.0890076	.1303525
ppent_at	329	.3486903	.2142087	0	.9594127	.1738089	.3259941	.4848818
book_l	329	.3692364	.2573526	0	1.316005	.1782781	.3330426	.5209277
book_l_l	329	.3144259	.2523319	0	1.311523	.1279241	.2734874	.4463608
debt_equity	284	283.344	4715.507	0	79468.22	.0743213	.3211928	1.012153
market_leverage	284	452.388	7523.617	.0077681	126793.7	.3492301	.913489	2.357301
net_leverage	329	.3005065	.3190311	-.8447217	1.305013	.1303637	.309367	.4958082
cash_at	329	.0687299	.1288072	0	.8830084	.0091555	.0201083	.0622495
N	329							

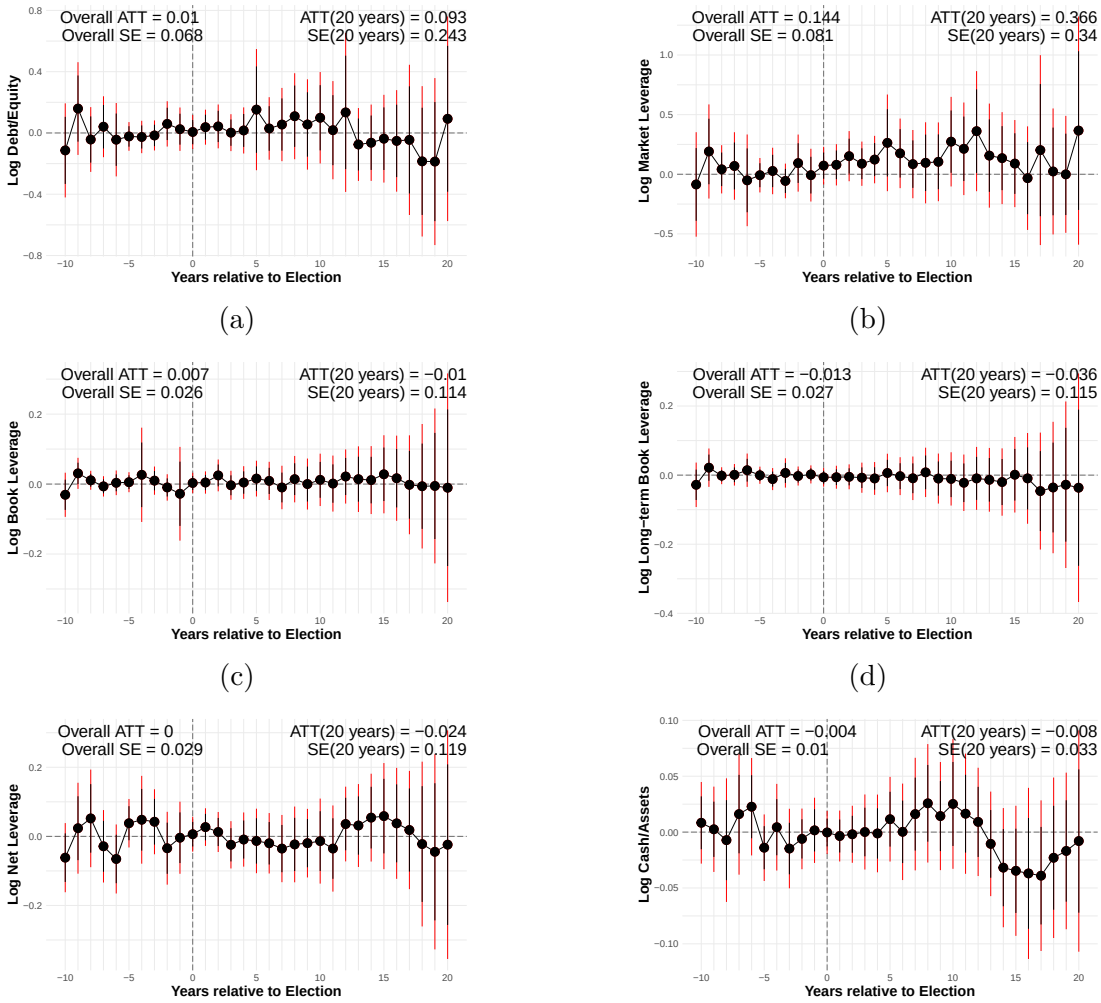
Notes: This table summarizes the status of financial leverage outcome variables in addition to cash to assets ratio and covariates in the year 2000. Among the covariates, *size* represents the logarithm of total assets, indicating the size of the firm. *tobin_q* measures the difference between the book value of assets and the book value of common equity, plus the market value of common equity, scaled by total assets. It serves as an indicator of the firm's investment opportunities. *ebit_at* is the return on assets (ROA), calculated by dividing earnings before interest and tax (EBIT) by total assets. It provides insight into the firm's profitability. *ppent_at* represents tangibility, specifically the ratio of fixed assets to total assets. Moving on to the outcome variables, *book_l* denotes book leverage, which is calculated by dividing the sum of long-term debt and debt in current liabilities by total assets. *book_l_l* represents long-run book leverage, specifically long-term debt divided by total assets. *debt_equity* is the ratio of long-term debt to the market value of common equity, providing a measure of the firm's debt relative to its equity value. *market_leverage* is calculated by dividing total liabilities by the market value of common equity, offering insight into the firm's leverage in relation to its market equity. *net_leverage* captures the ratio of long-term debt plus debt in current liabilities minus cash to total assets, reflecting the firm's leverage after accounting for available cash. *cash_at* represents the cash-to-assets ratio, obtained by dividing cash and short-term investments by total assets. It highlights the proportion of the firm's assets held in cash. For further information on the definition of these variables, please refer to the table 2.1 in this chapter.

Figure 2.2: Event Study Plots for Financial Leverage Outcome Variables and Cash to Assets Ratio (Subset of Successful Elections)



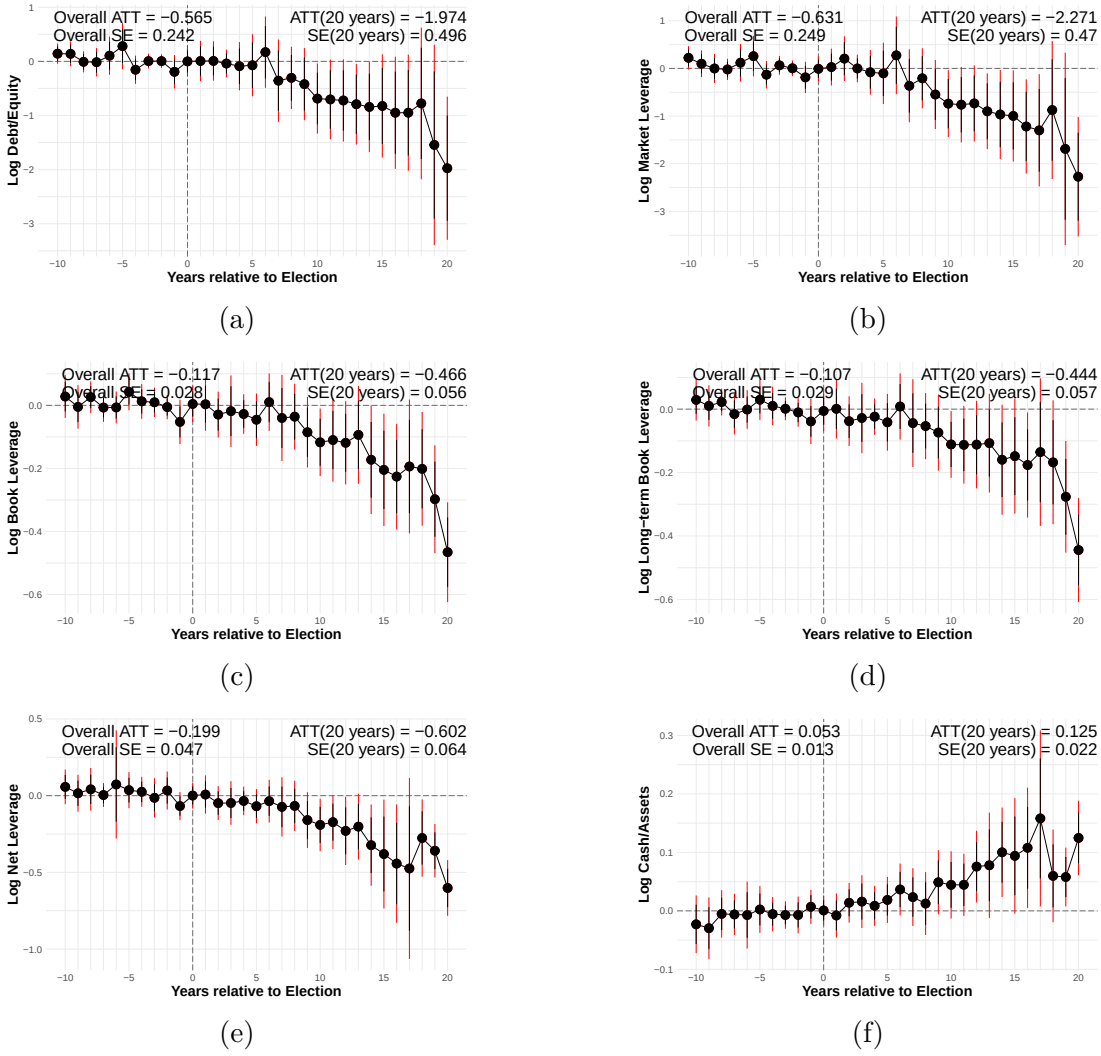
Notes: The figures present dynamic difference-in-difference estimates for the subset of elections resulting in a "win", under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of D/E, market leverage, book leverage, long-term book leverage, net leverage, and cash/asset ratio. The covariates are size, tangibility, profitability, and tobin's q. "Years relative to Election" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after unionization, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 2.3: Event Study Plots for Financial Leverage Outcome Variables and Cash to Assets Ratio (Subset of Unsuccessful Elections)



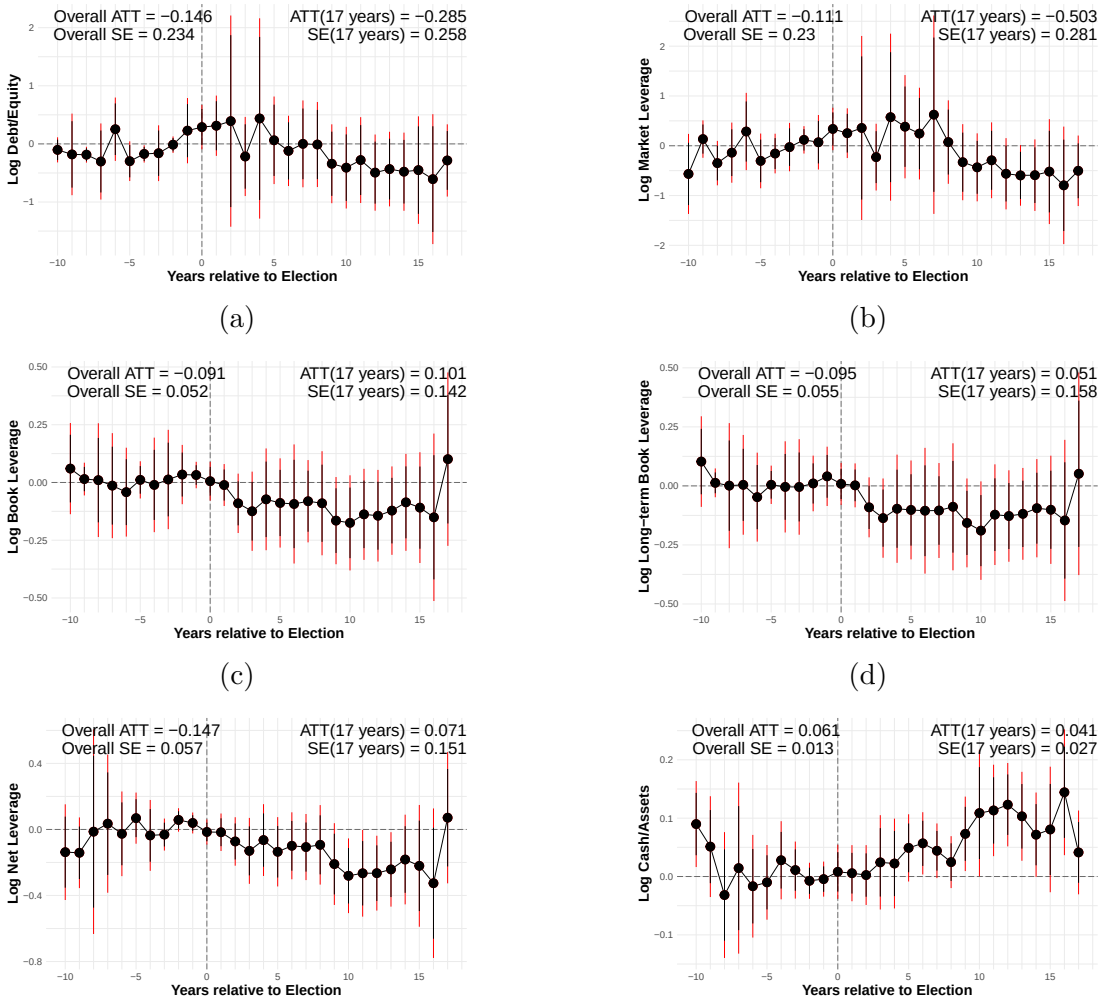
Notes: The figures present dynamic difference-in-difference estimates for the subset of elections resulting in a "loss", under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of D/E, market leverage, book leverage, long-term book leverage, net leverage, and cash/asset ratio. The covariates are size, tangibility, profitability, and tobin's q. "Years relative to Election" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after holding the election, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 2.4: Event Study Plots for Subset of Elections with Large Margin of Support (Successful Elections)



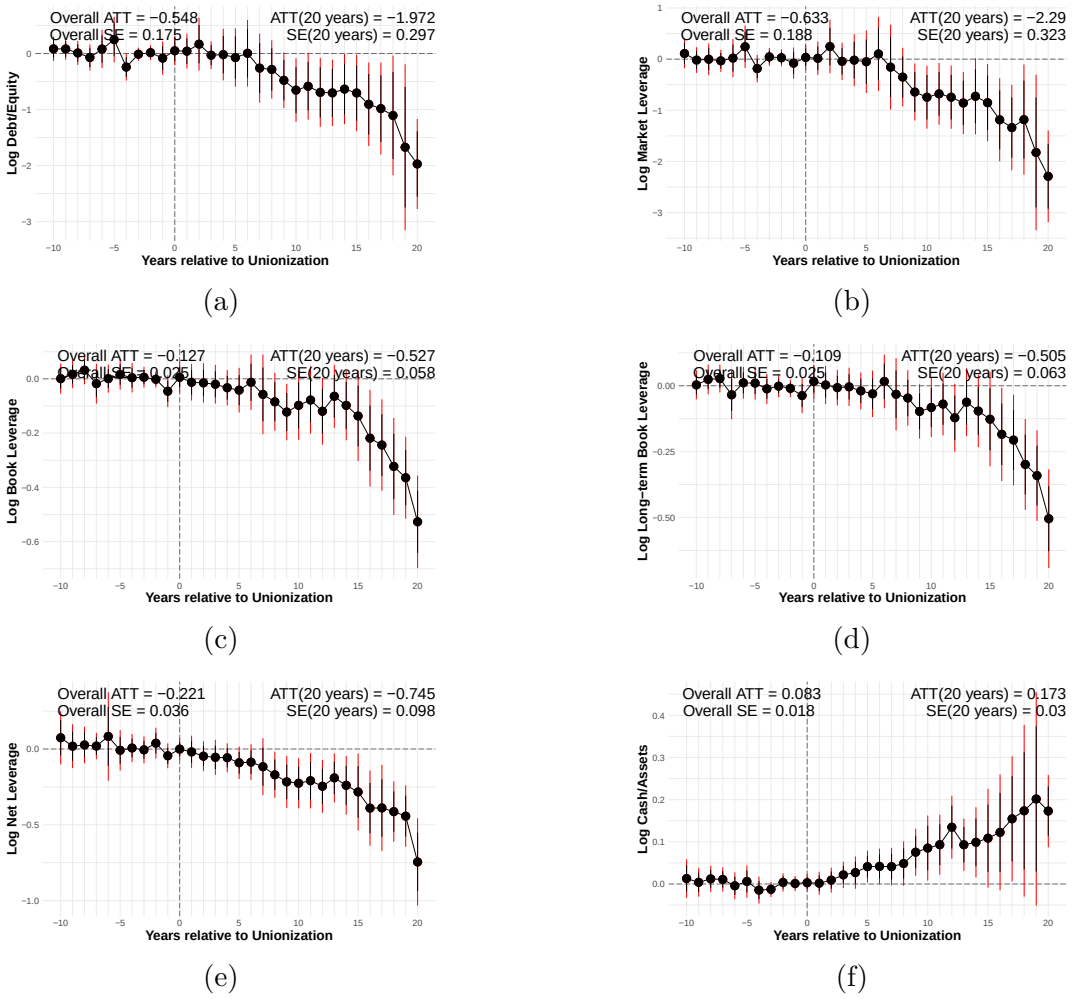
Notes: The figures present dynamic difference-in-difference estimates for the subset of elections resulting in a "win" and with "high" support, under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of D/E, market leverage, book leverage, long-term book leverage, net leverage, and cash/asset ratio. The covariates are size, tangibility, profitability, and tobin's q. "Years relative to Election" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after unionization, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 2.5: Event Study Plots for Subset of Elections with Small Margin of Support (Successful Elections)



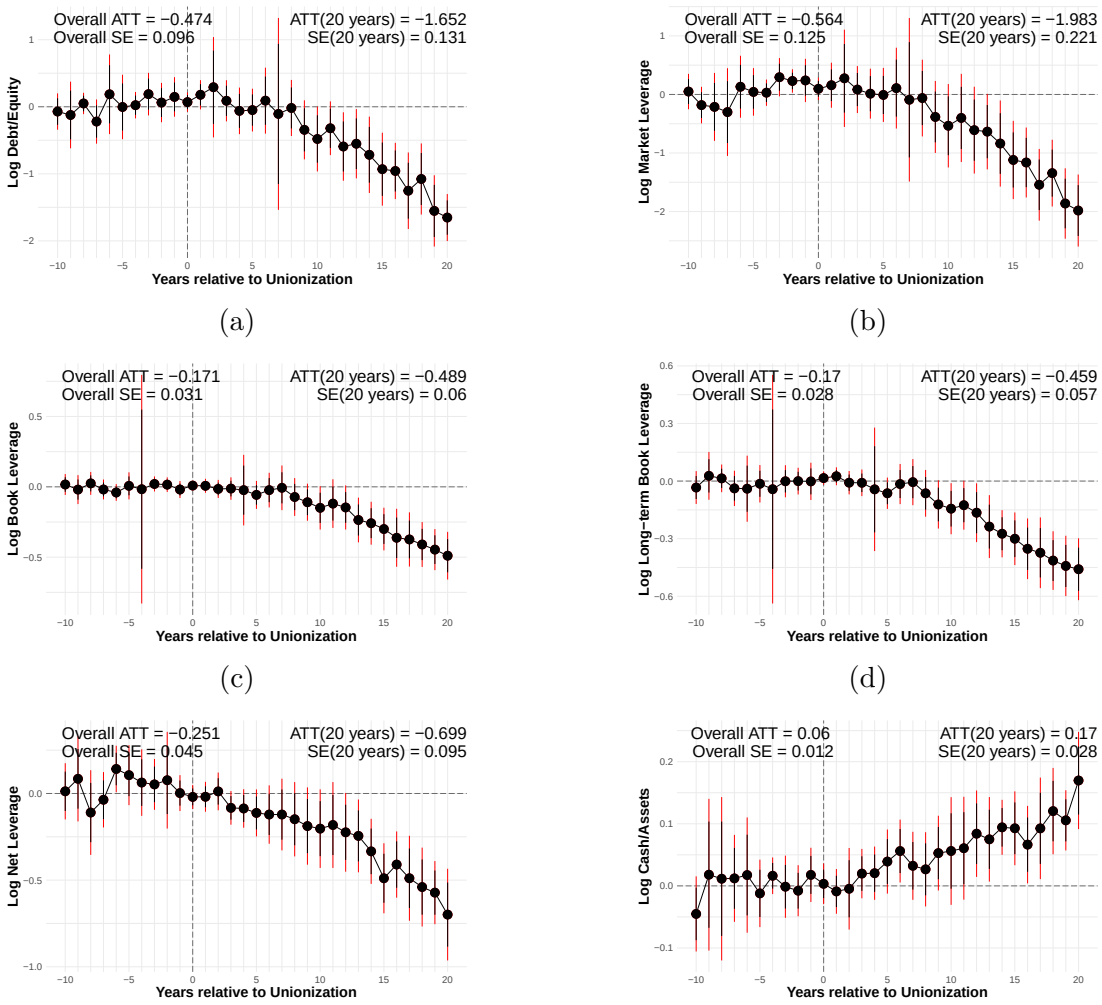
Notes: The figures present dynamic difference-in-difference estimates for the subset of elections resulting in a "win" but with "low" support, under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of D/E, market leverage, book leverage, long-term book leverage, net leverage, and cash/asset ratio. The covariates are size, tangibility, profitability, and tobin's q. "Years relative to Election" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after unionization, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 2.6: Event Study Plots for Subset of Unionized Firms Not Followed by Decertification



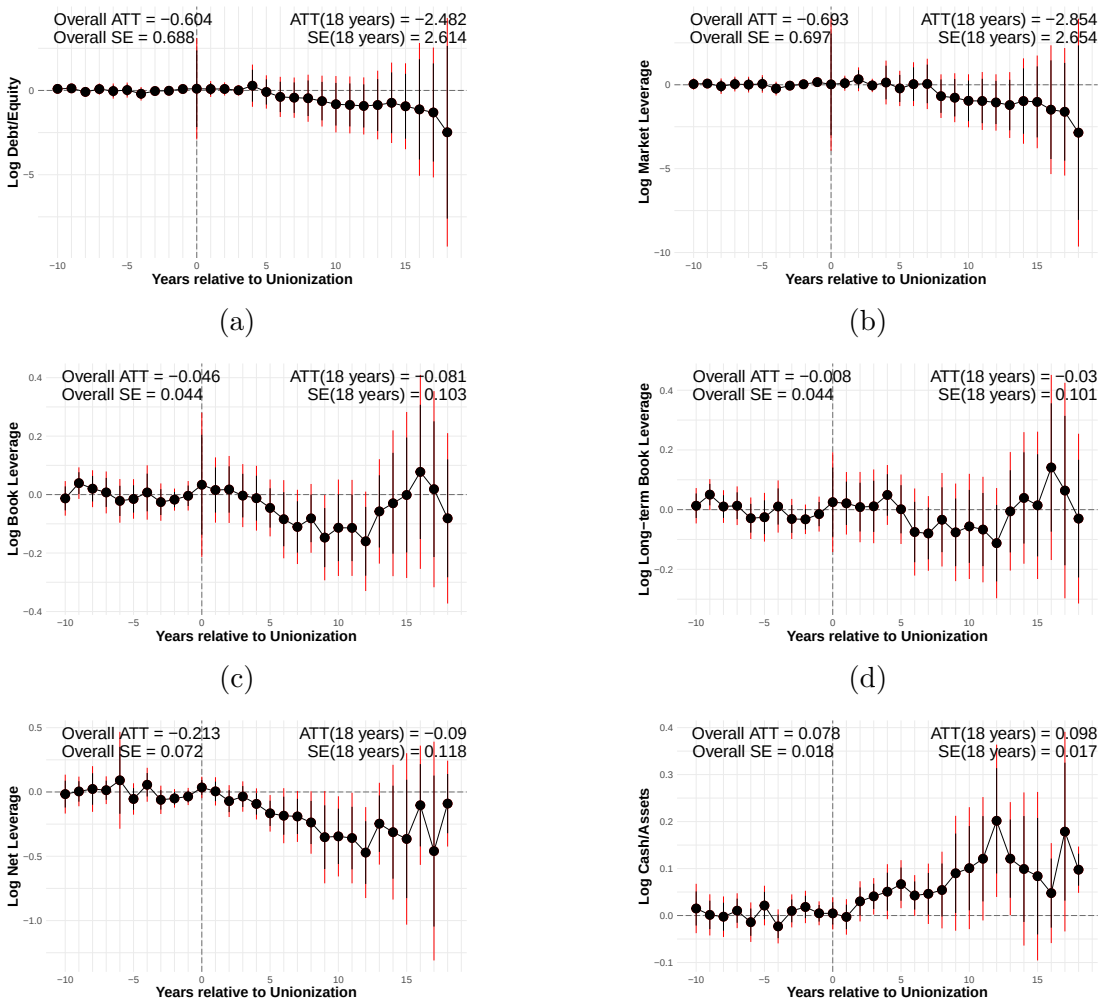
Notes: The figures present dynamic difference-in-difference estimates for the subset of truly unionized firms (first petition RC (win) and no RD later), under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of D/E, market leverage, book leverage, long-term book leverage, net leverage, and cash/asset ratio. The covariates are size, tangibility, profitability, and tobin's q. "Years relative to Unionization" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after unionization, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 2.7: Event Study Plots for Subset of Unionized Multi-establishment firms



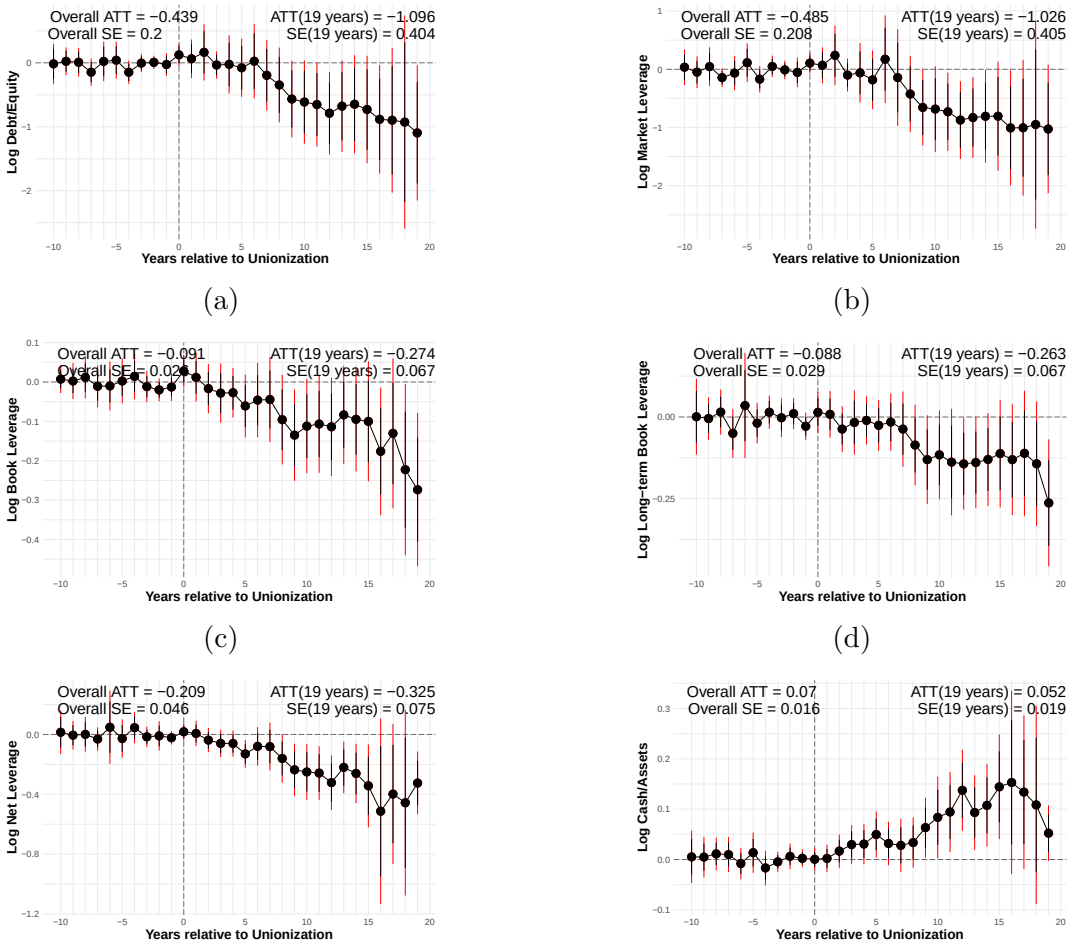
Notes: The figures present dynamic difference-in-difference estimates for the subset of unionized multi-establishment firms, under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of D/E, market leverage, book leverage, long-term book leverage, net leverage, and cash/asset ratio. The covariates are size, tangibility, profitability, and tobin's q. "Years relative to Unionization" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after unionization, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 2.8: Event Study Plots for Subset of Unionized One-establishment firms



Notes: The figures present dynamic difference-in-difference estimates for the subset of unionized one-establishment firms, under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of D/E, market leverage, book leverage, long-term book leverage, net leverage, and cash/asset ratio. The covariates are size, tangibility, profitability, and tobin's q. "Years relative to Unionization" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after unionization, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 2.9: Event Study Plots for Subset of Unionized firms with less than 7 number of franchises



Notes: The figures present dynamic difference-in-difference estimates for the subset of unionized firms with less than 7 number of franchises, under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of D/E, market leverage, book leverage, long-term book leverage, net leverage, and cash/asset ratio. The covariates are size, tangibility, profitability, and tobin's q. "Years relative to Unionization" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after unionization, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 2.10: States with right-to-work laws

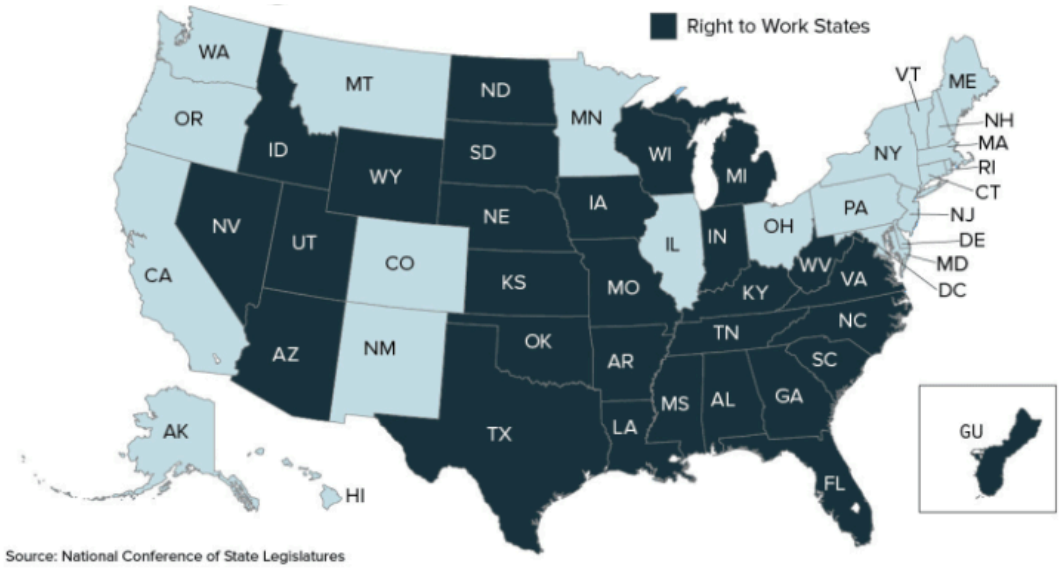
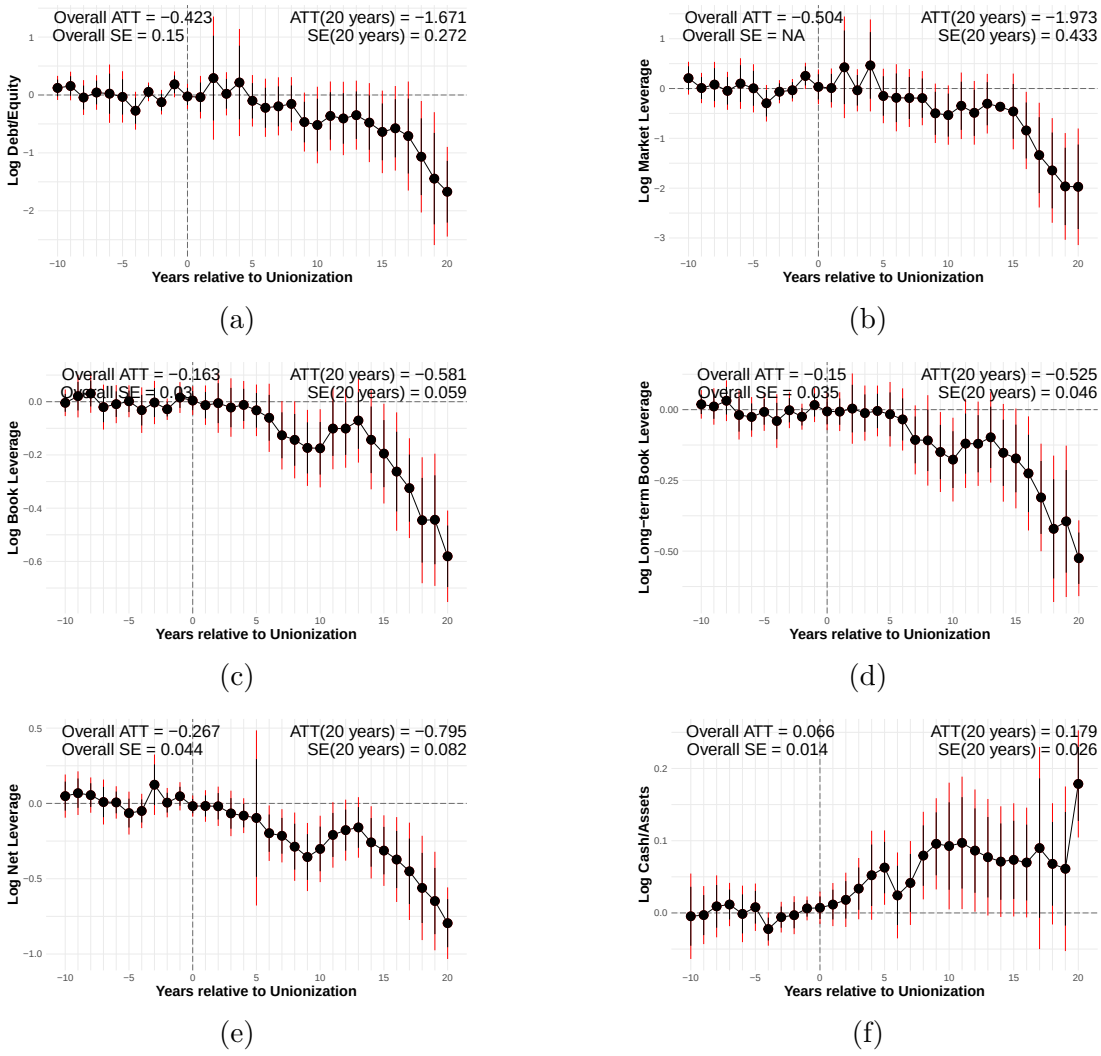
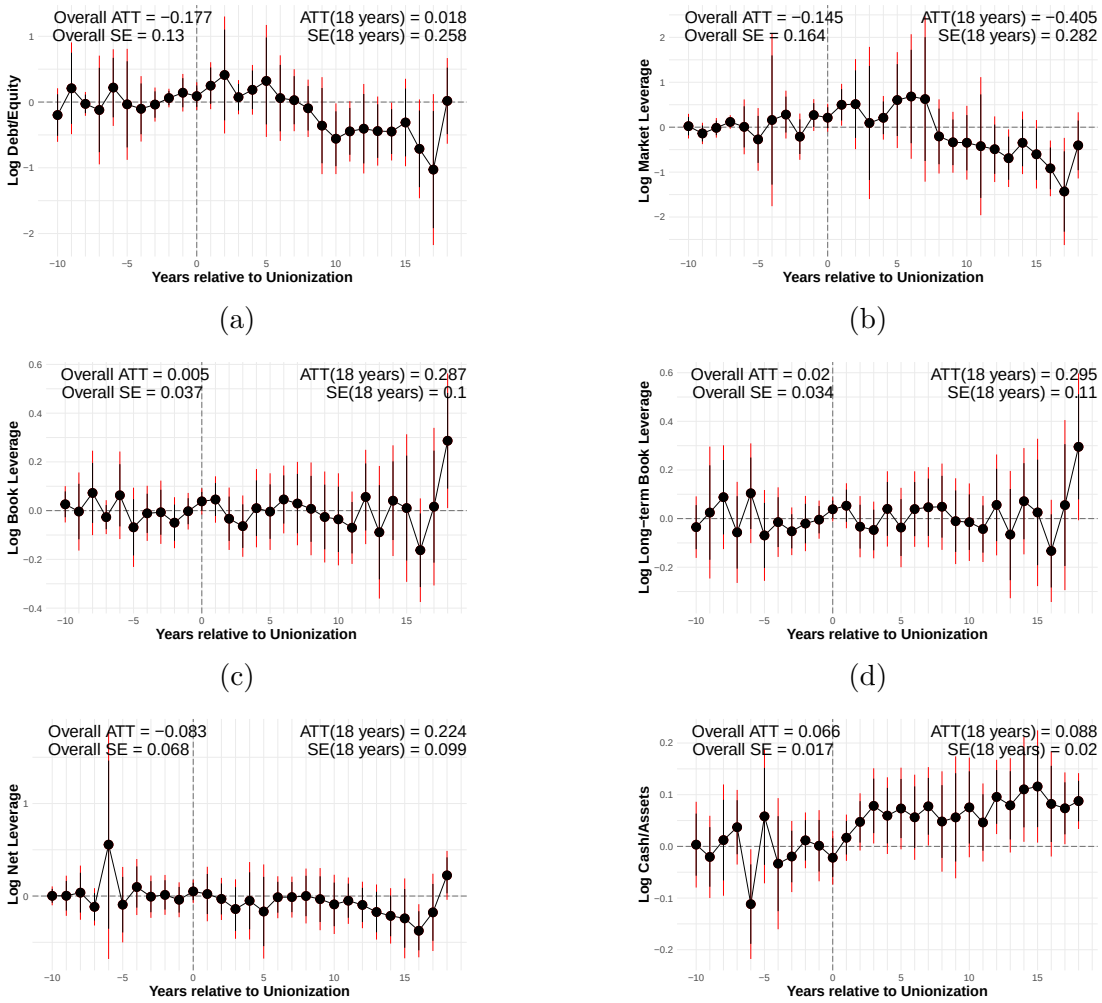


Figure 2.11: Event Study Plots for Subset of Unionized Firms in States without Right-to-Work Law



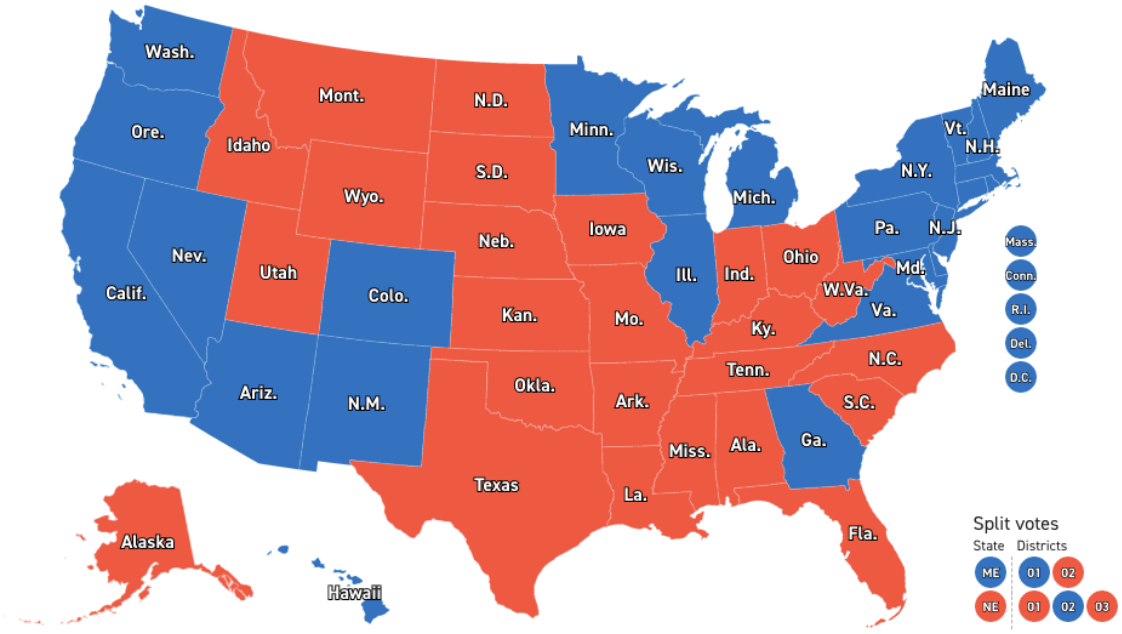
Notes: The figures present dynamic difference-in-difference estimates for the subset of unionized firms in states without RTW, under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of D/E, market leverage, book leverage, long-term book leverage, net leverage, and cash/asset ratio. The covariates are size, tangibility, profitability, and tobin's q. "Years relative to Unionization" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after unionization, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 2.12: Event Study Plots for Subset of Unionized Firms in States with Right-to-Work Law



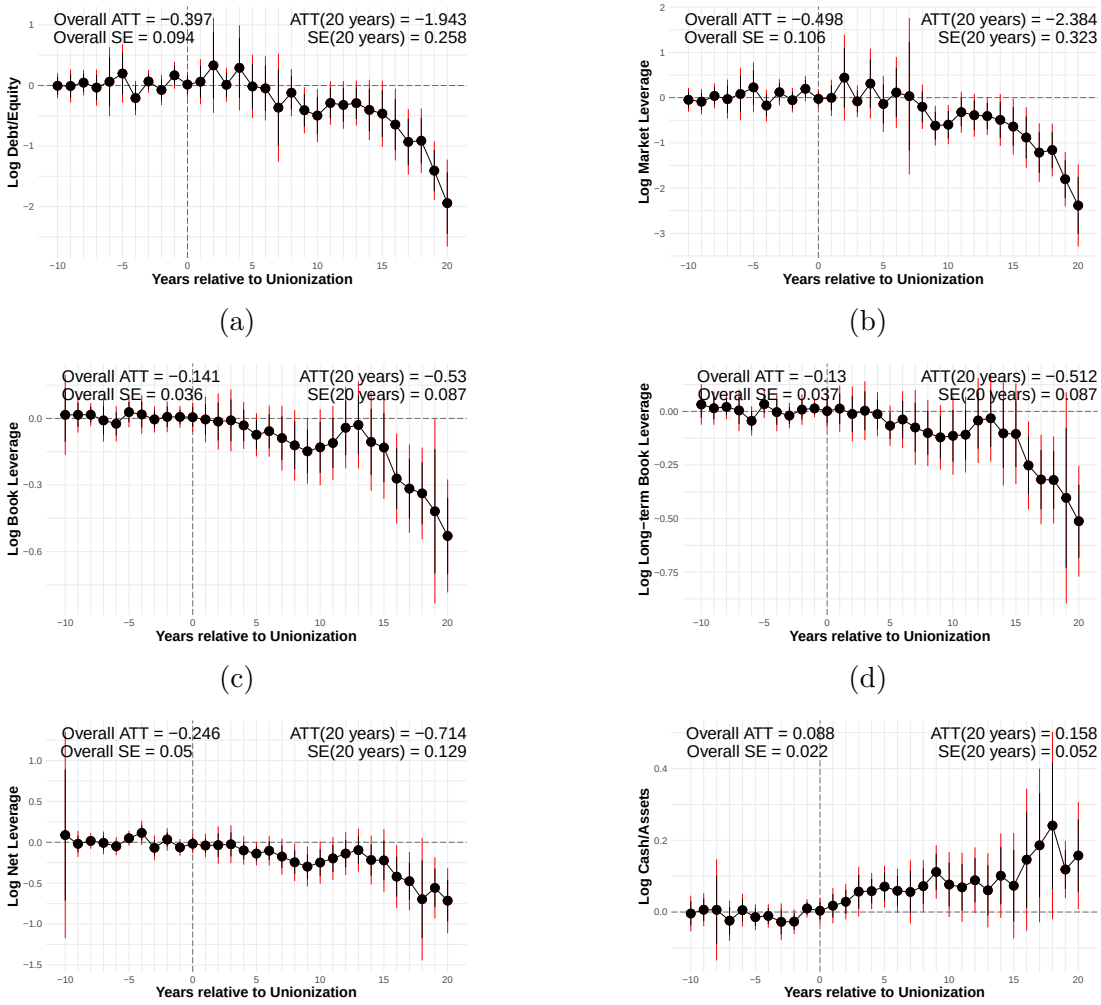
Notes: The figures present dynamic difference-in-difference estimates for the subset of unionized firms in states with RTW, under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of D/E, market leverage, book leverage, long-term book leverage, net leverage, and cash/asset ratio. The covariates are size, tangibility, profitability, and tobin's q. "Years relative to Unionization" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after unionization, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 2.13: Presidential Election Results (2020)



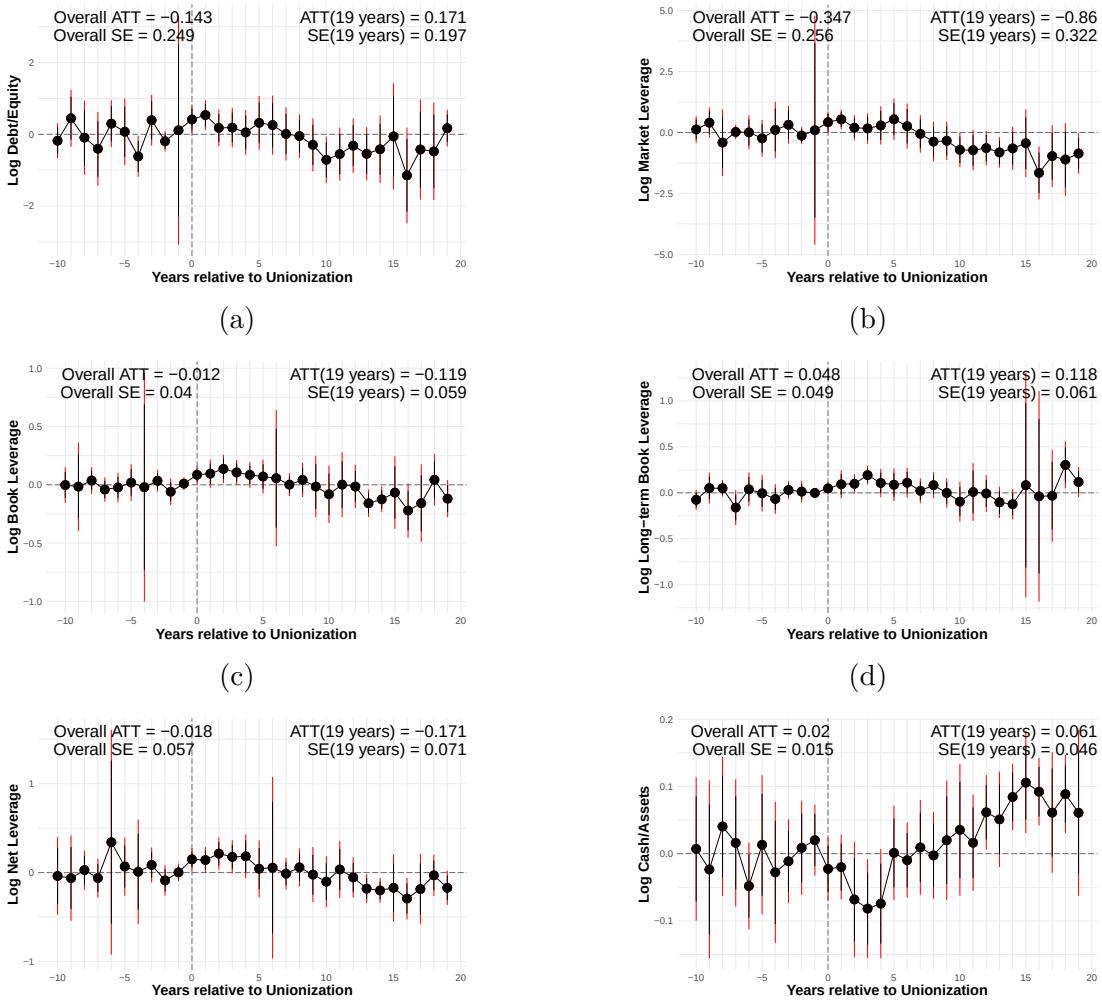
Source for the presidential election results: Politico

Figure 2.14: Event Study Plots for Subset of Unionized Firms in Democrat-led States



Notes: The figures present dynamic difference-in-difference estimates for the subset of unionized firms in Democrat-led States, under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of D/E, market leverage, book leverage, long-term book leverage, net leverage, and cash/asset ratio. The covariates are size, tangibility, profitability, and tobin's q. "Years relative to Unionization" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after unionization, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 2.15: Event Study Plots for Subset of Unionized Firms in Republic-led States



Notes: The figures present dynamic difference-in-difference estimates for the subset of unionized firms in Republic-led States, under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of D/E, market leverage, book leverage, long-term book leverage, net leverage, and cash/asset ratio. The covariates are size, tangibility, profitability, and tobin's q. "Years relative to Unionization" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after unionization, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Chapter 3

The Effects of Unionization on Firms' Innovative Capacities: Evidence from U.S. Firms

3.1 Introduction

Labor unions typically negotiate for higher wages, increased job security, and improved benefits for their members. However, union influence extends beyond these traditional areas and there is a growing recognition that unions can impact companies' innovative efforts. Innovation is a key driver of economic growth, productivity and competitiveness in the modern global economy. As firms seek a competitive advantage, they invest in research and development (R&D) activities to create new products, processes and technologies. These innovations not only enhance corporate performance, but also contribute to overall societal progress. Existing literature offers two contrasting views on how unionization can affect firms' innovative capacities. The main arguments underlying these two views are outlined below. The primary goal of this chapter is to revisit the relationship between unionization and firms' innovative activities, using a more carefully constructed dataset and a more up to date empirical methodology.

The relationship between unionization and innovation is complex and multifaceted. On

the one hand, trade unions can pose challenges to the innovation activities of firms. Union presence might encourage shirking by reducing the likelihood of being fired, thereby reducing the negative consequences of underperforming (Tressel and Scarpetta [91]). Furthermore, unions are engaged in rent-seeking behavior and hinder the market outcomes. According to Grout [48], wage bargaining reduces the gains from innovation by imposing a kind of tax on sunk capital (Connolly et al. [30]), thus discouraging upfront R&D investment and innovations. Generally, unionized companies could face higher labor costs, stricter work rules and slower decision-making processes, which could potentially hamper their ability to innovate and reduce productivity due to employees' potential (and realization) to strike. Therefore, according to this perspective, unionization will decrease the firm innovation.

On the contrary, unions can play a role in fostering innovation within companies. Unions are considered as an institution of collective voice which improves labor productivity which in turn increases innovation through a variety of different channels. Freeman and Medoff [44] show that union formation can result in decrease in worker grievances leading to better labor productivity and innovation. This way, unions contribute to enhanced innovation by providing better job security and negotiating for higher wages, which, in turn, reduces employee turnover and boosts productivity. Moreover, formation of unions may reduce labor turnover. Also, according to one side of the *employment protection* view, providing workers with better employment protection can encourage innovation by ensuring tolerance for early failure. Furthermore, based on the *oligopolistic markets* view, when unions bargain to maintain high employment level, this forces the firms to invest more in innovation and R&D in order to protect their own market share. As a result, according to this perspective, unionization will increase the firm innovation. These widely differing results are related to the ambivalent role of unions, which can either behave as rent-seekers as originally posited by Grout [48] or reinforce voice mechanisms that enhance organizational performance (Freeman and Medoff

[44]). With this unsolved puzzle, it is not surprising that the impact of unionization on firm innovation is still an open issue.

While both views regarding the positive and negative effects appear reasonable, there is a lack of consensus in the existing empirical literature about which of the two perspectives is relevant in today's economy. The primary goal of this chapter is to seek answers to these questions. Innovation is essential to a country's economic progress (Solow [87], Romer [79]) and a firm's long-term success (Porter [76]). Due to the recent availability of high-quality patent and citation data that capture a country's and/or a firm's innovation output, there has been a fast-growing strand of literature that explores various determinants and consequences of corporate innovation. Therefore, this study aims to contribute to the existing literature by examining the causal effects of unionization on the innovation activities of publicly traded firms in the United States.

The lack of consensus in the literature is not the only reason for me to revisit the effect of unionization on firms' innovative activities. Previous studies have not comprehensively covered all dimensions of innovation simultaneously. For example, according to the study by Bradley et al. [22], only the *number of patents* and *number of citations* are used to determine the effect of unionization on innovation. To address this gap, our study takes into account a wide range of innovation indicators. We begin by analyzing the *Number of Patents*, that measures the innovation output quantity but does not capture the scientific importance or technological value of the patents. Then, we use the *Number of Citations* subsequently received by the patents filed by the firm in a year that measure the innovation output quality in terms of scientific value. We also consider the market value of the patent filed by the firm in a year as reported in the KPSS dataset (*Patent value*) as our third innovation output proxy[1]. In order to capture the other dimensions, we also add alternative measures

[1]We also scale the total patent value by the number of patents filed by the firm in a year to measure the average value of each patent (*Average patent value*). We build a similar ratio by dividing the total number of

of innovation output. We construct two innovation efficiency measures by using number of patents granted, scaled by R&D expenditures *(Innovation Efficiency (patents))*, and patent citations, scaled by total R&D expenditures *(Innovation Efficiency (citations))*. We also use innovation productivity as another alternative measure of innovation by considering the number of patents per 1,000 firm employees. Finally, as other outcome variables, we capture the firm investment intensity as total capital expenditure divided by total sales (*capital expenditure*) and total R&D expenditure divided by total sales (*R&D expenditure*), since they can capture the input to innovation process. Overall, our study aims to provide a comprehensive analysis of the relationship between unionization and firms' innovative activities by considering various dimensions of innovation and utilizing multiple indicators.

Our study utilizes a dataset that combines information from three different sources. The union election data for the period spanning from 1999 to 2021 is obtained from the National Labor Relations Board (NLRB) website. This dataset provides comprehensive details such as the names of establishments that conducted union elections, the filing date, election date, closing date, total number of votes, and the number of votes in favor of forming a union. On the other hand, financial information is sourced from the Compustat database. For a more thorough understanding of the Compustat data, please refer to section B.1.2 of the previous chapter in the appendix. Finally, for our third reference of data, we use the patent data compiled by Kogan et al. [60] (Hereafter KPSS). The KPSS patent data allows us to track the firm's patenting activities based on patents filed and eventually issued at the US Patent and Trademark Office (USPTO). Compustat dataset includes various identifiers for each firm, such as the unique company-level identifier *permco*[2], the unique stock-level identifier

citations by the number of patents filed by the firm in a year to measure the average citation of each patent (*Average patent citation*).

[2] *permco* remains unchanged throughout the whole term of company's existence even if the company changed

permno, and the Global Company Key *gvkey*[3]. Unfortunately, the NLRB dataset does not have a common identifier that matches the identifiers in the Compustat data. To conduct our analysis, we need to merge the election data from NLRB with the financial information of the firms in Compustat. However, we only have access to the company names, which may not be entered consistently in both datasets. To address this issue, we begin by standardizing and cleaning the name strings using a standardization procedure [4]. Subsequently, first we employ a new matching criteria that utilizes the company names to link the unionized firms and the firms that held elections to the Compustat data. Then, since the KPSS dataset, includes information on *permno*, we rely on this identifier to build our final dataset that includes information from all references. This process results in a final sample of 168 firms.

In our study, we know that the process of unionization does not occur simultaneously for all firms. Instead, it is considered a staggered treatment, with each firm experiencing unionization at a specific point in time. This naturally leads us to consider using a standard linear two-way fixed effect model or an event study design, as many existing studies in the literature have utilized these methods. Additionally, regression discontinuity (RD) design has been used to examine the relationship between unionization and innovation-related indicators. However, recent advancements in econometrics literature (Goodman-Bacon [46], Sun and Abraham [88], Athey and Imbens [8], Callaway and Sant'Anna [29], Borusyak and Jaravel [21], Gardner [45] and De Chaisemartin and d'Haultfoeuille [34]) have raised concerns regarding the reliability of causal estimates obtained from a standard staggered difference-in-differences (DID) framework. Researchers widely agree that the standard approach can yield misleading estimates, even when the parallel trends assumption holds and treatment assignments are random. Recognizing the significance of this issue, we address it by adopting a solution proposed by Callaway and Sant'Anna [29]. Their approach not only

[3]It is a unique number assigned to each company in the Compustat-Capital IQ database.
[4]Our cleaning procedure builds on the `stnd_compname` Stata name standardization programs [93].

helps mitigate potential biases but also provides easily interpretable estimates akin to an event study, allowing us to capture both short-term and long-term effects of unionization. By applying this recent methodology, we aim to estimate the causal effect of unionization on various innovation-related indicators. This ensures that our results are unbiased under heterogeneous treatment effects and staggered treatment timing.

Our study contributes to the ongoing discussion on the role of unions in the labor market and provides new evidence on the impact of unionization on the innovation activities of firms in the United States. The main findings of our analysis are summarized below:

1. We find modest effect of unionization on a couple of innovation-related indicators. Our analysis documents a small positive and significant effect on the *number of citations and average patent citations* for up to the initial five years after the unionization election. This contrasts the earlier finding by Bradley et al. [22] that reports negative effect of unionization on number of patents and citations three years after union elections. Furthermore, we observe a relatively small positive effect of union elections on variables such as the *number of citations to R&D expenditures ratio, the number of patents per 1000 employees, capital expenditures to sales ratio, and R&D expenditures to sales ratio* in the latter half of the years (10-20) following the union elections, on average supporting the *voice mechanisms*. Although there are some positive effects observed in certain variables following successful elections, the overall impact is relatively modest. This implies that the presence of a union, may have a limited influence on firm innovation beyond the initial years following the election.

2. By examining the effects of unionization on innovation activities of firms separately for manufacturing and non-manufacturing firms, we aim to uncover any differential impact based on the type of firm activities. Our results suggest that *manufacturing* firms which hold union elections exhibit a relatively small positive and significant

effect on the *number of citations and average patent citations* within the initial 5 years following the election. These findings align with the results observed in the full sample of elections, suggesting a relatively modest positive effect of union elections on other variables such as the *number of patents per 1000 employees, capital expenditures to sales ratio, and R&D expenditures to sales ratio* in the latter half of the years of our analysis, which refers to years 10-20 following the union elections.

3. We divide the firms according to the size and explore the results is each subset (below median or above median) separately. By considering smaller and larger firms separately to determine if the size of a firm influences the relationship between unionization and innovative choices, our results suggest the unionization effect are more pronounced for smaller firms. The outcome variables for which the results are evident for the subset of *small* firms, are similar to that of the *manufacturing* firms.

4. Similarly, we examine the effects of unionization on firms' innovative decisions separately based on the age of firms, comparing younger firms versus older ones. This subgroup analysis seeks to identify any variations in the impact of unionization on innovation based on the stage of a firm's development. The results of unionization are stronger for younger firms. Our analysis suggest that *young* firms, which undergo union elections, exhibit a relatively small yet positive and significant effect on several innovation variables. These include the *number of patents, the number of citations to R&D expenditures ratio, the number of patents to R&D expenditures ratio, the number of patents per 1000 employees, capital expenditures to sales ratio, and R&D expenditures to sales ratio.* These effects manifest in the latter half of the years following the union elections.

5. We examine the varying impacts of unionization based on political orientation of states. Democratic-led states may have a higher proportion of unionized workers and more

favorable attitudes towards unions compared to Republican-led states. By separately analyzing the results for these different contexts, we gain a better understanding of how the political orientation of the states in which firms operate can influence the effects of unionization. Specifically, we observe that the effects of unionization on the outcome variables are more pronounced in Democrat-led states.

6. Furthermore, we address the possibility that firms may respond to the threat of unionization, even if the actual election results in a loss for the union. To investigate this, we make two modifications to our analysis. Firstly, we consider the election date as the treatment, regardless of the election result. Secondly, we narrow our focus on a subset of firms that experienced election losses. By examining this group separately, we can isolate the impact of unsuccessful unionization attempts on firms' innovative activities. In the case of treating the election date as the treatment, we observe small positive and significant impact of holding the election on a couple of innovation outcome variables that hold the potential for the "threat to unionization" to significantly impact firms' innovative activities. However, in the scenario where we consider firms that experienced election losses, we do not observe significant effects. This suggests that the actual success of unionization attempts is a crucial driver of any noticeable impact on firms' innovative activities.

The remainder of this paper is structured as follows: Section 3.2 provides an overview of the relevant literature on the intersection of unions and innovation. Section 3.3 describes the data sources, variables and methods used in our analysis. Section 3.4 offers a detailed overview of the methodology. Section 3.5 presents the empirical results and their implications. Section 3.6 concludes the paper by summarizing the main findings and discussing opportunities for future research. Figures and tables are reported in section 3.7. Finally, section C provides additional details on supplementary analyses in the appendix.

3.2 Related literature

3.2.1 Theory

There is an established literature on the relationship between unionization and firm innovation. There are various strand of theoretical research regarding the relationship between unionization and firm innovation. First strand of research considers unions as an institution of collective voice which improves labor productivity which in turn increases innovation through a variety of different channels. Freeman and Medoff [44] show that union formation can result in decrease in worker grievances leading to better labor productivity and innovation. Moreover, formation of unions may reduce labor turnover. This allows the firms to accumulate firm-specific human capital resulting in greater labor productivity and higher innovations. Furthermore, by facilitating the introduction of organizational innovations, they can increase job satisfaction and improve learning outcomes, with positive spillovers on labor productivity and innovation. Models based on unions being collective voice predicts a positive relationship between unionization and firm innovation.

The second strand of research is based on employment protection. Models based on employment protection can often generate contradictory predictions. On the one hand, providing workers with better employment protection can encourage innovation by ensuring tolerance for early failure. On the other hand, union presence might encourage shirking by reducing the likelihood of being fired, thereby reducing the negative consequences of underperforming. Tressel and Scarpetta [91] develop a theoretical model where unionization decrease the probability of getting fired. The employees shirk in their jobs resulting in lower innovations. The model predicts a negative relationship between unionization and firm innovation. In contrast, Acharya et al. [3] develop a theoretical model based on wrongful discharge laws. These laws prohibit the firms to terminate the employees in bad faith. The model predicts

that a positive relationship between employee protectionism and firm innovation.

The third strand of the theoretical literature addressing the effect of unionization on firm innovation is based on oligopolistic markets (Beath et al. [15] and Ulph and Ulph [92]). In this framework, when unions bargain to maintain high employment level, this forces the firms to invest more in innovation and R&D in order to protect their own market share. Haucap and Wey [50] propose a Cournot duopoly model which examines three different unionization regimes, namely, centralized, coordinated and decentralized. The model predicts that the relationship between unionization and firm innovation is non-monotonic with respect to the degree of wage centralization. Different unionization regimes result in different innovation incentives. Bryson and Dale-Olsen [27] extend the Cournot duopoly model of Haucap and Wey [50] by including a voice effect and relax the assumption that unions have no effects upon productivity. They find that the relationship between unionization and firm innovation can either be positive or negative. Berton et al. [18] improve upon the model of Bryson and Dale-Olsen [27] to differentiate between product and process innovation. Labor productivity are different across unionized and non-unionized sectors because of collective voice mechanisms. They find that in case of product innovation, unionization will result in greater innovation incentives compared to a sector where wages are competitive, if the productivity gains due to union voice mechanisms are large. In case of process innovation, the effect of union voice on firm innovation is ambiguous.

The fourth strand of the literature believes that the unions are engaged in rent-seeking behavior and hinders the market outcomes. According to Grout [48], wage bargaining reduces the gains from innovation by imposing a kind of tax on sunk capital (Connolly et al. [30]), thus discouraging upfront R&D investment and innovations. However, this negative relationship between unions and firm innovation may be weaken or even vanish based on efficient bargaining between union and firms (J OSWALD' and J TURNBULL [58]; Layard et al. [63]

and Booth [20], Mukherjee and Pennings [73]). The details of these different mechanisms are summarized in table 3.1 as included in Berton et al. [18].

Some other papers focus on how innovation is shaped by firm-level corporate culture. In a seminal theoretical paper, Manso [69] shows that the optimal incentive scheme to motivate innovation needs to tolerate early failure and reward for long-term success. In order to create such an innovative corporate culture, firms should adopt long-term compensation plans, offer job security, and provide timely feedback on performance to managers and key R&D employees.

3.2.2 Empirical Evidence

Empirical evidence about the effect of unionization on firm innovation have been mixed. Addison et al. [4] and Bradley et al. [22] show respectively, that there is no evidence of a negative relationship between unions and innovations in Germany while the relationship is negative and significant in the US. Ederer and Manso [38] report that firms and labor unions enter into contracts which motivate innovation by tolerating early failures and rewarding long-term success suggesting a positive relationship between unionization and innovation. Acharya et al. [3] reports that the US states which adopted wrongful discharge laws exhibit an increase in the annual number of patents and citations by 12.2% and 18.8% respectively. Their findings support the employment protection hypothesis which states that stringent employment protection laws foster innovation. Acharya et al. [2] uses data from three European countries, namely, Germany, France and United Kingdom and find evidence in support of employment protection hypothesis. To the extent that the passage of stringent employment protection laws are results of union activism, the results suggest a positive relationship between unionization and innovation. On the contrary, Bradley et al. [22] report that winning

union elections lead to a 8.7 % decline in patent quantity and quality and reduction in R&D expenditure three years after union elections. Further, unionization results in firms moving away from states with stronger labor laws.

Doucouliagos and Laroche [37] conducted a meta-regression analysis on 27 studies from four countries (UK, US, Canada, Germany) and discovered a negative correlation between unionization and innovation in all four countries. However, the effect is declining over time and increasing in the flexibilization of labor markets. This is similar to the findings of Menezes-Filho and Van Reenen [71].

Bryson et al. [26] found that union presence can mitigate job-related anxiety due to the introduction of process innovations, indicating that implementing innovations may be more costly for non-unionized firms. Holman and Rafferty [54] found that more unionized systems of industrial relations have greater organizational innovations. Antonioli et al. [7] identified a positive relationship between on-the-job well-being and organizational innovation and cooperative industrial relations. Finally, Bryson et al. [28] use British data on the introduction of high-involvement management (HIM) practices (also referred to as "highperformance work systems") at the firm-level and found that the positive effect of organizational changes on labor productivity is limited to unionized firms. These pieces of evidence suggest that unions may act as collective voice institutions.

Acharya et al. [3] tested the "employee protectionism" hypothesis using US data and found that wrongful discharge laws can increase the number of patents and citations. They also extended their analysis to Germany, UK, and France and found that more stringent dismissal laws foster innovation, particularly in knowledge-intensive industries. Ederer and Manso [38] supported Manso [69] principal-agent model by showing that innovation-motivating contracts ensure rewards for long-term success and tolerance for early failures.

There are a couple of other papers that address the role of monopoly power. Schnabel and Wagner [84], find that union density has a positive impact on R&D only when union monopoly power is not too high, while Fang and Ge [39] point out that the positive association between union presence and innovation in China can be explained by the low bargaining power of Chinese unions. Menezes-Filho and Van Reenen [71], in turn, justify the existence of non-linearities in the relationship between unions and innovation in Europe by claiming that unions have a positive impact on innovation when their bargaining power is low and a negative impact on innovation when their bargaining power is high.

3.3 Background and Data

Our study employs a dataset that is constructed from various sources. The union election data from 1999 to 2022 is retrieved from the National Labor Relations Board (NLRB) website. This data provides crucial details about the companies that held union elections, such as the company name, filing date, election date, closing date, total number of votes, and the number of votes for the union. We determine one of our treatment times using the filing date, which is the earliest date recorded for each election. We obtain financial information from Compustat database. More details regarding Compustat data is discussed in section B.1.2 of previous chapter.

We carefully link the data on unionized firms and firms held the election to Compustat data using the company names (as outlined in section B.1.4 of previous chapter). We only include union elections with at least 6 voters and with available election results, and for firms with multiple election records, we retain the first one (as per Bradley et al. [22], Young et al. [96]). We only keep the *first* election for each establishment. This means that our estimates should be taken as the effects of winning the first union election at an establishment. Consistent with

standard practices in corporate finance research, we exclude financial firms (SIC 6000-6999) and utility companies (SIC 4900-4999) due to their highly regulated industries. Additionally, we drop non-U.S. firms.

To measure firms' innovation output, for our third reference of data, we use the patent data compiled by Kogan et al. [60] (Hereafter KPSS). The KPSS patent data allows us to track the firm's patenting activities based on patents filed and eventually issued at the US Patent and Trademark Office (USPTO). The dataset provides the number of the patents filed by the firm, the number of citations received by the patent, and the estimated market value of the patents filed by the firm with the USPTO[5]. Recently, they have updated the data until the end of 2022. In addition to the patent and citation count measures, the economic value of patents is an important and useful innovation outcome metric as it is denoted in dollar amount and thus comparable over time and across industries. We follow the innovation literature and use the patent filing date instead of grant date since patent filing dates are closer to the firm innovation activities. KPSS includes the patents that have been granted, it generally takes about two years from the patent filing to patent grant, thus, we do not observe the patents filed after 2020 (Hall et al. [49], Sunder et al. [89], Islam and Zein [57]). Our final sample consists of 168 union elections conducted between 1999 and 2022 for which we have information from Compustat and KPSS as well. We can see the distribution of firms according to the time they held an election for the purpose of unionization in table 3.3. Next, we will examine the distribution of firms that certified as unionized after holding the election (first petition result in win and no *decertification* in any of the franchises). According to table in table 3.4, we have 63 such firms in our data.

[5]In particular, the patent value is estimated on the three-day window following the patent issue date and deflated to 1982 (million) dollar using the CPI. Three-day return is computed as the cumulative market adjusted return over the three-day period from Tuesday to Thursday (Patents are always issued on Tuesdays).

3.3.1 Measuring Firm Innovation

We use several measures to proxy firm-level innovation outcomes in this study. The first metric is the number of patent applications that are filed (and eventually granted) for each firm-year (*Number of Patents*). The patent number measures the innovation output quantity but does not capture the scientific importance or technological value of the patents. Our second proxy is the number of citations subsequently received by the patents filed by the firm in a year (*Number of Citation*). The citation received by the patents gauges the innovation output quality in terms of scientific value. Another important measure of firm innovation quality is the patent economic value, we then use the market value of the patent filed by the firm in a year as reported in the KPSS dataset (*Patent value*) as our third innovation output proxy. The economic value of patents is estimated using the forward-looking stock market data and thus captures the private value to the patent holders (e.g., gaining competitive advantage over rivalries), which do not necessarily coincide with the scientific value of patents. More importantly, the economic value of the patents is in terms of dollars, facilitating its comparison over time and across industries. We also scale the total patent value by the number of patents filed by the firm in a year to measure the average value of each patent (*Average patent value*). We build a similar ratio by dividing the total number of citations by the number of patents filed by the firm in a year to measure the average citation of each patent (*Average patent citation*).

We add alternative measures of innovation output. First, we construct two innovation efficiency measures and examine whether unionization leads to lower or higher innovation efficiency. To measure innovation efficiency, we use patents granted, scaled by R&D expenditures *(Innovation Efficiency (patents))*, and patent citations, scaled by total R&D expense *(Innovation Efficiency (citations)).* We also use innovation productivity as another alternative measure of innovation. If a firm has more incentive to innovate due to unionization, it

should have higher innovation productivity. To test this, we follow the approach by Acharya et al. [3] and estimate the number of patents per 1,000 firm employees (*log(1 + Patents/Employees)*). Finally, we calculate the firm investment intensity as total capital expenditure divided by total sales (*capital expenditure*) and total R&D expenditure divided by total sales (*R&D expenditure*), since they can capture the input to innovation process. We use the natural logarithm of one plus each innovation measure in our analysis due to the skewness of the variables. More details regarding our outcome variables are included in table 3.2.

3.3.2 Covariates

We control for standard covariates that can explain corporate innovation following the innovation literature (e.g., Islam and Zein [57], Sunder et al. [89]). At the firm level, we use natural logarithm of book value of total assets to proxy for firm *size* to reduce the skewness in total assets. We use *tobin's q* to proxy the firm growth opportunities. Furthermore, we add *tangibility*[6] and *profitability* to control for firm investment intensity and availability of internal funds, respectively. Finally, we consider *age of the firm* as another covariate that can affect the innovation activities of the firm. More details regarding covariates are included in table 3.2. Tables 3.5 and 3.6, report the summary statistics of the main variables used in our analysis.

3.4 Methodology

For our analysis, we focus on the impact of unionization on firms' innovative capacities. To determine the treatment timing, which refers to when firms held the unionization election, we use the date when the election petition was filed. However, since not all firms in our

[6]Fixed assets divided by total assets

sample held the election at the same time, we have a staggered adoption setting. This setting allows us to employ a static linear two-way fixed effect (TWFE) model or an event study design. However, recent advancements in the literature (e.g., Goodman-Bacon [46], Sun and Abraham [88], Athey and Imbens [8], Callaway and Sant'Anna [29], Borusyak and Jaravel [21], Imai and Kim [56], Gardner [45] and De Chaisemartin and d'Haultfoeuille [34]) have shed light on the limitations of both the TWFE model and the event study design. The TWFE model estimates the causal parameter as a weighted average of all possible 2x2 difference-in-differences (DID) estimators, comparing different timing groups. We also know that in a TWFE model, already treated units act as control units, but this can lead to misleading estimates due to heterogeneity in treatment effects.

To address the potential bias arising from heterogeneity in treatment effects, recent econometric literature provides solutions. One approach is to modify the set of comparison units by excluding previously treated units, allowing only "never treated" units to act as controls. However, this restriction can be costly, especially with a relatively small sample size. To overcome this issue, we adopt the methodology proposed by Callaway and Sant'Anna [29] (henceforth CS 2021), which allows "not yet treated" units to serve as controls. This methodology provides sensible estimates and allows us to leverage our data fully, reporting treatment effects over an extended period while maintaining a balance between treated and control units.

The CS 2021 methodology takes a ground-up approach and uses group-time specific average treatment effects on the treated, denoted as $ATT(g,t)$, as the building blocks. We create groups based on the timing of treatments in our sample. The $ATT(g,t)$ measures the average treatment effect at time t for the group first treated in time $g < t$, and is defined as

$$ATT(g,t) = \mathbb{E}\left[Y_t(g) - Y_t(0) \mid G_g = 1\right] \tag{3.1}$$

where G_g is a dummy variable equal to one if the unit is in treatment time group g. It is defined as the expected difference between the outcome variable Y_t for treated units at time t if they were first treated at time g, and the potential outcome $Y_t(0)$ for the same treated units if they had not been treated. To identify the treatment effects, we rely on control groups consisting only of units that have not received treatment up to time g.

To estimate $ATT(g,t)$, we adopt a doubly robust approach that combines the outcome regression (OR) approach of Heckman et al. [51] and the inverse probability weighting (IPW) approach of Abadie [1], and is given by

$$ATT(g,t) = \mathbb{E}\left[\overbrace{\left(\frac{G_g}{\mathbb{E}\left[G_g\right]} - \frac{\frac{p_{g,t}(X)(1-D_t)(1-G_g)}{1-p_{g,t}(X)}}{\mathbb{E}\left[\frac{p_{g,t}(X)(1-D_t)(1-G_g)}{1-p_{g,t}(X)}\right]}\right)}^{\text{Inverse Probability Weight}} \overbrace{\left(Y_t - Y_{g-1} - m_{g,t}(X)\right)}^{\text{Outcome Regression}}\right] \tag{3.2}$$

This approach allows us to account for both the outcome evolution in the control group and the propensity of units to be in group g given covariates. The $ATT(g,t)$ estimator is the weighted average of the difference in outcome evolution between the treated and control groups, where the weights reflect the similarity of characteristics between the groups. CS 2021 highlights that the doubly robust estimator is valid as long as either the outcome evolution or the propensity score model is correctly specified. It is important to note that more detailed information on the estimation procedure and additional findings can be found in section 1.3.

The estimation results provide evidence of the unionization effects over a longer time horizon, as we divide the sample period into twenty-three sub-intervals from 2000 to 2022. Each sub-interval corresponds to a specific group based on the year of the firm's election/unionization. For example, we assign firms to group 2 if these firms held their election for unionization

during the year 2001. Similarly, we assign firms to group 23 if their election date years belong to the 2022. The choice of these sub-intervals and groups is not arbitrary. They are designed to utilize the maximum amount of information on key outcome variables and to capture the unionization effects over a longer horizon. By considering different time periods relative to the treatment, we capture both instantaneous and time-varying effects.

To present our results in a standard event study format and assess the parallel trend assumption, we construct weighted averages of the group $ATTs$ for each relative time period. These aggregated estimates, weighted by group sizes, allow us to analyze the evolution of the unionization effects. We conduct formal inference using CS 2021's recommended bootstrapping procedure, which provides simultaneous confidence bands robust to multiple hypothesis testing.

We can observe the outcome variables starting from 1999. As a result, pre-treatment information on firms in group 1 who filed an election in 2000 are also available. Therefore, we don't need to exclude this group from the analysis. On the other hand, our methodology prohibits us from using already treated units as controls. Thus, the last treated firms (group 23) lack controls of their own and can only serve as comparison units in earlier periods. Therefore, the unionization effects that we report below are based on $ATTs$ that materialized for the remaining groups.

3.5 Results

The goal of this chapter is to investigate the impacts of workers' decision to become unionized on firms' innovation activities. We also consider other types of treatments such as workers' initiative to hold elections which lead to win result and workers' lack of initiative to decertify a union after it is formed. The results pertaining to these different types of

treatments collectively paint the picture about how firms' innovative activities are affected by unionization.

To measure the impact of unionization on innovation, we primarily focus on innovation-related variables from the KPSS dataset, including *number of patents, number of citations, market value of patents, average patent value and average patent citations*. Furthermore, we provide results for several other alternative measures of innovation such as *the number of patents to R&D expenditures ratio and the number of citations to R&D expenditures ratio*). Also, we consider other indicators of innovation including *number of patents per 1000 employees, capital expenditures to sales ratio and R&D expenditures to sales ratio*. We are interested to see if the choice of innovation measure may affect the results, and we present a comprehensive summary of all innovation proxies and key covariates in Table 3.2.

We begin our analysis with the benchmark case, where the treatment is defined as holding an election, irrespective of the outcome. For more detailed information, please refer to section 3.5.1. To address our research question effectively, we recognize the importance of considering the outcome of unionization elections. In section 3.5.2, we continue our analysis by examining the results separately for firms that experienced election victories ("win" results) and firms that experienced election losses ("loss" results). This approach allows us to explore the distinct effects based on the election outcome. However, it is essential to acknowledge that union representation can be revoked through a *decertification* election at a later point in time. To account for this possibility, we further extend our analysis in section 3.5.3 by focusing on a subset of firms that hold the elections with "win" result and did not initiate any decertification process after forming a union. In this case, we consider the "closing date" as the treatment year, as it represents the announcement date of the election results. This modification allows us to provide a more comprehensive understanding of the effects of unionization on firms' innovative activities.

We also conduct subgroup analyses to explore potential variations in the effects of unionization on firms' innovative activities. Considering the political orientation of the states in which firms operate can provide valuable insights into the effects of unionization on innovation. Democratic-led states typically have a higher proportion of unionized workers and more favorable attitudes toward unions, while Republican-led states often have different perspectives. By separately analyzing the effects of unionization in Democrat-led states and Republican-led states, as determined by the 2020 presidential election, we can gain a better understanding of how the political context interacts with unionization and influences firms' innovative activities.

In addition, we investigate the heterogeneous treatment effects in subgroups based on the manufacturing versus non-manufacturing sectors to explore any potential variations in the effects of unionization on firms' innovative decisions within these sectors. This analysis aims to determine if the impact of unionization on innovation differs depending on industry type. Furthermore, we intend to examine the effects of unionization on innovative decisions based on the age of firms separately, comparing younger firms versus older ones. This subgroup analysis seeks to identify any variations in the impact of unionization on innovation based on the stage of a firm's development. Moreover, we plan to analyze the effects of unionization on innovative decisions among smaller and larger firms separately to determine if the size of a firm influences the relationship between unionization and innovative choices. These findings contribute to a better understanding of the mechanisms through which unionization influences innovation outcomes. The detailed results of these subgroup analyses will be presented in section 3.5.4.

3.5.1 Effects of Holding Union Elections

We begin our analysis with the benchmark case, where the treatment corresponds to holding an election regardless of the results. Therefore, we ignore the vote share in favor of unionization at this stage. Also, it is important to acknowledge that in practice, a union can be *decertified* at a later date, and we take this possibilities into account later in our analysis. Furthermore, unions may conduct multiple elections in different franchises of a firm. However, for our analysis, we focus solely on the first election that occurs within a firm. In this section, we examine the effects of union elections on various innovation-related variables.

Figures 3.2 and 3.3, plot the estimated effects for the indicated innovation variables under conditional parallel trend assumption[7], for the subset of firms that hold the unionization election. This event study plot include pre- and post-treatment estimates with 95% simultaneous confidence bands. The pre-treatment estimates provide valuable insights into the validity of the (conditional) parallel trend assumption. In Figures 3.2 and 3.3, the average treatment effects (ATTs) prior to the treatment are indistinguishable from 0, supporting the (conditional) parallel trend assumption. Our analysis controls for various firm-specific variables, such as size, Tobin's q, tangibility, profitability, and age which we observe at the time of the treatment. These variables serve as covariates for the entire sample. However, for certain sub-samples, there is insufficient variation in the data, making it challenging to conduct a meaningful outcome regression component of Equation (3.2). In these cases, we utilize a subset of conditioning variables and employ the inverse probability weighting (IPW) approach instead of the doubly robust (DR) approach. This allows us to account for potential biases and obtain reliable estimates despite the limited variation in the data.

According to the results, certain outcome variables exhibited significant effects, while others

[7]The plots under unconditional parallel trends assumption are included in the appendix (Figures C.1, C.2).

showed relatively small positive effects of unionization in the latter half of the years, which refers to years 10-20 following the union elections. When it comes to the variables of *the number of citations and average patent citations*, our analysis demonstrates that the effects of holding an election become visible with time. After 2 years from holding the election, the effects start to become visible and firms experience significant increase in their number of forward citations. This suggests that firms that held union elections, on average, experienced higher levels of future citations for their patents (their patents received greater citation counts). This finding indicates a positive impact of unionization on the visibility and influence of firms' patented innovations. We experience a loss of (not-yet-treated) comparison units as we move forward in time. Despite this, when the firms hold elections, their number of citations continues to increase throughout the period of our analysis.

On the other hand, when examining the variables such as *the number of citations to R&D expenditures ratio, the number of patents per 1000 employees, capital expenditures to sales ratio, and R&D expenditures to sales ratio*, we observe a relatively small positive effect of union election in the latter half of the years of our analysis, which refers to years 10-20 following the union elections. This suggests that unionization may have a modest but discernible influence on these particular aspects of innovation. For example, firms experiencing union elections may allocate a slightly higher proportion of their resources towards research and development in the long-run, resulting in a slightly higher number of patents per employee or an increased investment in capital expenditures relative to their sales.

However, it is important to note that we did not observe any significant effect of unionization on *the number of patents, market value of patents, average patent value, and the number of patents to R&D expenditures ratio* variables. This indicates that holding elections did not have a substantial impact on the overall quantity or market value of patents held by the firms, nor did it significantly affect the efficiency of converting R&D investments

into quantity of patentable innovations. Overall, these results contribute to the literature on the relationship between unionization and innovation, offering insights into the specific dimensions of innovation that may be influenced by union elections.

3.5.2 Effects of Forming Unions

So far, we have not incorporated the outcome of the election in our analysis. However, considering that the election result can have an impact on the results, we extend our analysis by defining the treatment as a successful election outcome with a vote share in favor of unionization exceeding 50%. It is important to note that this definition assumes that the treatment is absorbing and does not take into account potential *decertification* elections. Additionally, unions may hold multiple elections in different franchises of a firm, but our analysis solely focuses on the first election within a firm. After winning the election, both the union and the firms are obligated to engage in good faith bargaining. Consequently, we explore the dynamic effects of winning the first union election, which may not necessarily align directly with being represented by a union. Furthermore, we report the results for the subset of unsuccessful union elections as well.

First, we estimate the effects of successful union elections on firm innovation indicators. Figures 3.4 and 3.5 present the estimates for all the innovation-related variables. The results depicted in Figure 3.4 indicate that firms with successful union elections show a small positive and significant effect on the *number of citations and average patent citations* for up to the initial five years after the election. Similar to the findings in the full sample of elections, we observe a relatively small positive effect of successful union election on variables such as the *number of citations to R&D expenditures ratio, the number of patents per 1000 employees, capital expenditures to sales ratio, and R&D expenditures to sales ratio* in the latter half of

the years of our analysis, which refers to years 10-20 following the union elections. However, the results for the remaining outcome variables are not significant.

Since the result of the election in favor or against unionization plays a significant role in addressing our question of interest; as a robustness check, we redo the analysis with the subgroup of firms which hold an election with "loss" outcome, where the vote share in favor of unionization was less than 50%. The results of this analysis, illustrated in Figures 3.6 and 3.7, indicate that the effects of unionization are not evident within the subgroup of elections with "loss" results. This highlights the point that the "threat to unionization" does not appear to have a significant effect on the innovation decisions of firms. providing further support for our previous results.

These findings shed light on the differential effects of successful versus unsuccessful union elections on firm innovation indicators. While there are positive effects of successful elections on certain variables, the overall impact is relatively small. It suggests that the presence of a union, as indicated by a successful election, may have limited influence on firm innovation beyond the initial years following the election.

3.5.3 Unionization Not Followed by Decertification

Since, in our previous analysis on first win elections, the underlying assumption fails to acknowledge that employees may lose their union representation through a decertification election and the possibility of holding multiple elections in other franchises after holding the initial election, one could argue that these mechanisms have the potential to confound our results. To mitigate this concern, in this section, we have designated the first close date, which is when the union is officially certified, for a successful unionization petition[8] as the

[8]known as the "RC type" that is a petition filed by an employee or group of employees, or any individual or labor organization acting in their behalf, for an election to certify a representative for collective bargaining

treatment year. Here, we define a stricter version of unionization and have only considered those firms as treated where there has been workers' lack of initiative to "decertify"[9] a union after it is formed.

Similar to the previous section, the findings depicted in Figure 3.8 reveal that a successful election combined with a lack of workers' initiative to decertify a union after its formation has a small positive and significant effect on the *number of citations* for up to the initial five years following the election. However, we observe stronger effects of unionization on certain innovation variables (*number of citations, number of citations to R&D expenditures ratio and capital expenditures to sales ratio*) when compared to solely relying on the subgroup of firms where the first election resulted in a win. On the other hand, the results for the *number of citations to R&D expenditures ratio and capital expenditures to sales ratio* become evident after two years of unionization, which was not observed in the previous analyses. As we progress further in time, there is a loss of comparison units that have not yet undergone the treatment. Nonetheless, even with this loss, firms that experience a successful election without subsequent decertification efforts witness a continued increase in their *number of citations to R&D expenditures ratio and capital expenditures to sales ratio* throughout the analysis period. The increase in certain innovation indicators suggests a potential rise in firms' reliance on innovative activities. Furthermore, Figures 3.8 and 3.9 illustrate that the average treatment effects (ATTs) before the treatment are statistically indistinguishable from zero, providing support for the (conditional) parallel trend assumption. This finding further validates the analysis, indicating that the compared groups exhibited similar trends prior to the treatment.

in an appropriate unit (*Source:* NLRB Election Reports).

[9]known as the "RD type" that is a petition filed by employees alleging that the union previously certified or currently recognized by the employer as their collective - bargaining representative no longer represents a majority of the employees in the appropriate unit and seeking an election to determine this (*Source:* NLRB Election Reports).

These results provide valuable insights into the dynamic relationship between unionization and firm innovation. It indicates that the effects of unionization extend beyond the initial unionization event and highlight the importance of considering the ongoing presence of the union when examining its impact on innovation outcomes. The continued influence of the union on bargaining dynamics and labor-management relations may play a significant role in shaping the innovation strategies and decisions of firms.

3.5.4 Heterogeneous Treatment Effects (Analysis of Subgroups)

Effects of Labor Unions in States With Different Political Orientation (Democrat Versus Republic)

So far, our analysis did not distinguish between firms operating in Democrat-led states versus Republican-led states. However, it is crucial to investigate the effect of unionization on firms' innovative capacities within each set of states separately. This differentiation allows us to account for potential variations in unionization rates and attitudes toward unions between Democratic and Republican states. Democratic states tend to have a higher proportion of unionized workers and more favorable attitudes toward unions. This supportive environment for unions can influence the bargaining power and effectiveness of unions in these states. On the other hand, Republican states may have a smaller number of unionized workers and less favorable attitudes toward unions. Additionally, there may be differences in union and labor rights regulations and policies between Democrat-led and Republican-led states. Democrat-led states often implement policies that support labor rights and create a more supportive environment for unions. These policies may enhance the bargaining power of unions, leading to better working conditions and potentially fostering an environment conducive to innovation. By analyzing the effect of unionization independently in each type

of state based on political orientation, we can gain a deeper understanding of how political factors interact with unionization.

According to Figure 2.13 in the previous chapter, we divide states to "Republic" versus "Democrat" using the result of 2020 presidential election. From another viewpoint, there is a high overlap between Republic-led states and states with *right to work* laws. According to Feigenbaum et al. [41], right-to-work laws cut Democratic Presidential vote shares by 3.5 percentage points. They investigate the mechanics and discover that right-to-work legislation reduce organized labor contributions to Democrats and make potential Democratic voters less likely to be approached to vote. Right-to-work laws have been shown to weaken labor unions resulting in diminished support for unionization. These laws may inhibit the ability of unions to effectively advocate for worker rights and may limit their impact on innovation decisions within Republican-led states.

Therefore, we proceed to estimate the effects of union elections on innovation-related outcome variables for firms located in democrat-led versus republic-led states separately. Figures 3.10, 3.11, 3.12 and 3.13, plot the estimates for all the innovation-related variables for the subset of firms in democrat and republic states, respectively. The estimates presented in Figure 3.10 and 3.11, suggest that firms which hold union elections exhibit a small positive and significant effect on the *average patent value and average patent citations* for up to the initial 5 years after the election in democrat-led states. For the case of *average patent value* the positive effects of holding elections are transient and it is followed by a steady decline. Consistent with the findings in the full sample of elections, we observe a relatively small positive effect of union election on variables such as the *number of patents per 1000 employees, capital expenditures to sales ratio, and R&D expenditures to sales ratio* in the latter half of the years in our analysis (years 10-20 following the union elections). However, the results suggest that the effects for the *number of citations to R&D expenditures* become visible with time and

firms experience significant increase in this ratio after ten years following the election. It continues to increase throughout the period of our analysis as well. On the other hand, the corresponding effects of union election did not show up for the subsample of republic-led states according to Figure 3.12 and 3.13. This highlights the point that the threat to unionization does not seem to have a significant impact on the innovation decisions of firms in states led by republicans.

Overall, the divergent effects of unionization on innovation outcomes in Democrat-led and Republican-led states can be attributed to a combination of factors, including differences in unionization rates, attitudes towards unions, policy variations, and the influence of right-to-work laws. Understanding these underlying mechanisms is crucial for policymakers to develop context-specific strategies that promote unionization and protect labor rights, while also fostering an environment conducive to innovation in different political and policy contexts.

Effects of Labor Unions in manufacturing versus non-manufacturing Firms

In this section, we delve into the role of industry type on the effects of unionization, specifically exploring whether the impact of union elections differs between manufacturing and non-manufacturing firms. To categorize the firms, we utilize the NAICS code[10]. Figures 3.14, 3.15, 3.16 and 3.17, represent the estimates for all the innovation-related outcome variables for *manufacturing* and *non-manufacturing* firms, respectively. Our analysis reveals substantial heterogeneity in the effects of unionization across these two broad industry groups. Specifically, within the non-manufacturing sector, which includes the service sector and has witnessed a majority of recent union organizing efforts, we find relatively smaller and mostly insignificant effects of unionization.

[10]We define manufacturing as NAICS sectors 31-33 and services as NAICS 51-81 and NAICS 44-45 and the "Others" category include the remaining sectors. "Non Manufacturing" firms include both "services" and "others" categories.

The estimates presented in Figures 3.14, and 3.15, indicate that *manufacturing* firms which hold union elections exhibit a relatively small positive and significant effect on the *number of citations and average patent citations* within the initial 5 years following the election. These findings align with the results observed in the full sample of elections, suggesting a relatively modest positive effect of union elections on other variables such as the *number of patents per 1000 employees, capital expenditures to sales ratio, and R&D expenditures to sales ratio* in the latter half of the years of our analysis, which refers to years 10-20 following the union elections. Furthermore, Figures 3.14 and 3.15 illustrate that the average treatment effects (ATTs) before the treatment are statistically indistinguishable from zero, providing support for the (conditional) parallel trend assumption. This finding further validates the analysis, indicating that the compared groups exhibited similar trends prior to the treatment.

In contrast, the corresponding effects of union elections in the non-manufacturing firms, as depicted in Figures 3.16 and 3.17, are generally insignificant, except for a minor positive effect on *Average Patent Citation*, which is comparatively stronger than that of the manufacturing firms. This highlights that the threat to unionization does not seem to have a substantial impact on the innovation decisions of non-manufacturing firms.

Understanding the mechanisms behind these differences is crucial. Manufacturing firms often operate in industries with significant research and development activities, capital-intensive operations, and a higher reliance on intellectual property. In such contexts, the introduction of unions through union elections may lead to a relatively small yet noticeable positive impact on innovation outcomes. The collective bargaining power of unions can facilitate negotiations for better working conditions, increased investment in research and development, and the fostering of a collaborative and inclusive work environment, all of which can contribute to enhanced innovation.

On the other hand, the limited effects observed in non-manufacturing firms indicate that the

threat of unionization may not significantly influence their innovation decisions. The nature of the service sector, characterized by different organizational structures and lower levels of capital intensity, may render the impact of unionization less pronounced. Moreover, the dynamics of the service sector, including the higher prevalence of non-standard employment arrangements, variable job roles, and decentralized decision-making, may affect the direct influence of unions on innovation outcomes in these firms. Overall, our findings underscore the importance of considering industry-specific factors when assessing the effects of union elections on innovation.

Effects of Labor Unions with respect to *Size*

In this section, our aim is to investigate the potential impact of firm size on the relationship between union elections and innovative decisions. Specifically, we examine how this relationship differs between firms with below-median size and firms with above-median size. To determine the median size, we analyze the available data from the year 2002 among firms that held a union election and find out the median of firm sizes[11] to be 7.8, which serves as the threshold for categorizing firms into two groups. The first group, referred to as the *small size* category, includes firms that had a size smaller than 7.8 in 2002. The second group, known as the *large size* category, comprises firms with a size greater than 7.8. By examining the effects of union elections on various innovation-related variables separately for small size and large size firms, as depicted in Figures 3.18, 3.19, 3.20, and 3.21, we gain insights into how firm size influences the outcomes of unionization on innovative decisions.

Figure 3.20 and 3.21 reveal that *small* size firms, which undergo union elections, experience a small yet significant positive effect on the *number of citations and average patent citations* within the initial 5 years following the election. This finding is consistent with the results

[11]It is calculated based on logarithm of total assets from Compustat data.

observed in the full sample of elections, indicating a relatively modest positive effect of union elections on other variables such as the *number of patents, the number of citations to R&D expenditures ratio, the number of patents per 1000 employees, capital expenditures to sales ratio, and R&D expenditures to sales ratio* in the latter half of the years of our analysis after the union elections.

To gain a deeper understanding of these results, we can explore the underlying mechanisms behind the observed patterns. In the context of *small* size firms, union elections may lead to several factors that contribute to the positive impact on innovation outcomes. Firstly, the presence of unions can enhance employee representation and provide a platform for addressing workplace concerns and negotiating better working conditions. This increased employee voice and involvement in decision-making processes may foster a sense of ownership and engagement, ultimately leading to a more innovative work environment. Secondly, the presence of unions in *small* size firms may create a collaborative atmosphere where knowledge sharing, information exchange, and cooperation among employees are encouraged. This collaborative culture can stimulate creativity, facilitate the flow of ideas, and enhance problem-solving capabilities, all of which are essential drivers of innovation.

On the other hand, the lack of significant effects observed in the subsample of *large* size firms, as depicted in Figures 3.18 and 3.19, suggests that the threat of unionization may not substantially impact the innovation decisions of firms with relatively larger size. One possible explanation is that larger firms often have established internal structures, processes, and resources in place, which are less susceptible to the influence of unionization. These firms may already have robust mechanisms for addressing employee concerns, ensuring fair labor practices, and promoting innovation without the explicit need for union representation.

Overall, the differential effects of union elections on innovation outcomes between *small* size and *large* size firms can be attributed to a combination of factors, including the enhanced

employee representation and collaborative environment fostered by unions in *small* size firms, as well as the pre-existing structures and resources within larger firms that may mitigate the impact of unionization on innovation decisions. Also, holding elections in *small* firms can contribute to increased visibility and recognition, which can potentially lead to more citations for their patents.

Effects of Labor Unions with respect to *Age*

In this section, our objective is to explore the potential influence of firm age on the relationship between union elections and innovative decisions. Specifically, we investigate whether this relationship differs for firms below the median age and firms above the median age. To determine the median age, we analyze the available data from the year 2002 among firms that underwent a union election and find the median age threshold to be 34, based on number of years listed on Compustat. Using this threshold, we categorize firms into two groups: the *Young* category, comprising firms with a age less than 34 in 2002, and the *Old* category, consisting of firms with a age greater than 34. By separately examining the effects of union elections on various innovation-related variables for young and old firms, as depicted in Figures 3.22, 3.23, 3.24 and 3.25, we gain insights into the influence of firm age on the outcomes of unionization in relation to innovative decisions.

The estimates presented in Figures 3.24 and 3.25 suggest that *young* firms, which undergo union elections, exhibit a relatively small yet positive and significant effect on several innovation variables. These include the *number of patents, the number of citations to R&D expenditures ratio, the number of patents to R&D expenditures ratio, the number of patents per 1000 employees, capital expenditures to sales ratio, and R&D expenditures to sales ratio.* These effects manifest in the latter half of the years following the union elections, on average.

To understand the underlying mechanisms driving these results, it is crucial to explore the

dynamics at play in *young* firms. Union elections in *young* firms can introduce various factors that contribute to the observed positive impact on innovation outcomes. Firstly, the presence of unions can amplify the collective voice of employees, empowering them to voice concerns, negotiate better working conditions, and advocate for a supportive work environment conducive to innovation. Secondly, *young* firms often operate in dynamic and competitive industries, where innovation is crucial for survival and growth. Unionization can facilitate an environment of collaboration, information sharing, and employee involvement in decision-making processes. This collaborative culture fosters the exchange of ideas, stimulates creativity, and encourages problem-solving, thereby enhancing innovation within *young* firms.

Conversely, the lack of significant effects observed in the subsample of *old* firms, as depicted in Figures 3.22, and 3.23, suggests that the threat of unionization may not substantially impact the innovation decisions of relatively older firms. Older firms, with more established structures and practices, may have already developed internal mechanisms to address employee concerns, foster innovation, and maintain competitiveness. Consequently, the presence or absence of unionization may have a limited additional impact on their innovation outcomes.

3.6 Conclusion

Despite the decrease in trade union membership in the United States, labor unions still has a potentially significant influence on innovation activities of firms. Our research offers new evidence regarding the causal effect of unionization on the innovation-related variables.

Collectively, our findings suggest that union elections have a modest positive effect on certain innovation-related variables, such as the number of citations, average patent citations, and various other innovation indicators. We also explore the effects of successful versus

unsuccessful union elections on firm innovation indicators separately. It suggests that the presence of a union, as indicated by a successful election, may have limited influence on firm innovation beyond the initial years following the election, whereas the effects of unionization are not evident in the subgroup of elections with "Loss" result. When we consider the effect of successful elections combined with workers' lack of initiative to "decertify" a union after it is formed, we observe stronger effects of unionization on innovation variables compared to solely relying on the subgroup of firms where the first election date resulted in a win. However, the magnitude and significance of effects of holding union elections vary across different subgroups. Firstly, we find small significant and positive effect of unionization on innovation-related indicators for firms operating in democrat-led states, possibly due to presence of a more supportive environment for unions. On the other hand, when considering firm size, we observe that smaller firms experience a small yet significant positive impact of union elections on innovation outcomes. In contrast, the effects are not significant for larger firms. This highlights the differential influence of unionization depending on the size of the firm, with smaller firms potentially benefiting more from the collective bargaining power of unions in fostering innovation. Moreover, our analysis reveals that the impact of union elections is more pronounced in younger firms compared to older firms. Younger firms show positive and significant effects on various innovation variables, whereas older firms do not exhibit substantial impacts. This suggests that the threat of unionization and the presence of unions may play a more influential role in shaping the innovation decisions of younger firms, possibly due to their greater flexibility, adaptability, and openness to change.

Furthermore, our study emphasizes the industry-specific effects of unionization. Manufacturing firms, characterized by their research and development activities and capital-intensive operations, demonstrate modest yet significant positive effects of union elections on innovation outcomes. In contrast, non-manufacturing firms, primarily in the service sector, show

smaller and mostly insignificant effects, highlighting that the influence of unionization on innovation decisions may be less pronounced in these contexts. The mechanism behind these results lies in the collective bargaining power of unions, which can lead to improved working conditions, increased investment in research and development, and the creation of a collaborative work environment.

3.7 Figures and Tables

Table 3.1: Summary of Theories for Effects of Unionization on Innovation

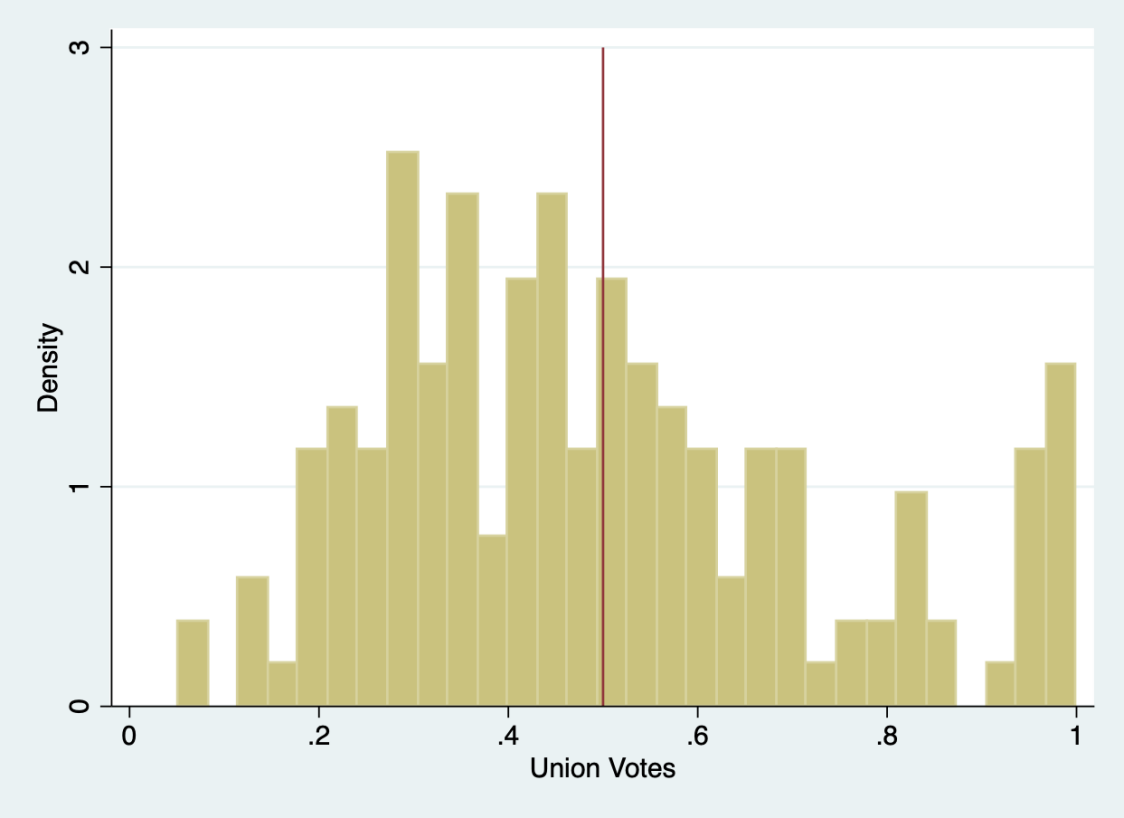
Theoretical View	Effect on Innovation	Mechanism
Monopoly Power	↓	↑ wages =↓ innovation gains (e.g., Grout [48])
Collective Voice	↑	↓ worker grievances/turnover =↑ productivity (e.g., Freeman and Medoff [44])
Strategic R&D	↑	↑ market share =↑ R&D investments (e.g., Beath et al. [15])
Employment Protection	↑	↑ long-term commitment =↑ greater worker effort (e.g., Acharya et al. [3])
Employment Protection	↓	↓ dismissal probability =↓ labor effort (e.g., Tressel and Scarpetta [91])

Source: Berton et al. [18], page 6

Table 3.2: Variable Definitions (Innovation-related Outcome Variables and Covariates)

Firm-level Variables	**(Compustat item names in the parentheses where applicable)**
Outcome Variables	
Number of Patents	Annual number of patents applied for and eventually granted, source: KPSS patent database
Number of citations	Annual number of citations per patent received, source: KPSS patent database
Dollar value	Annual dollar value of patents granted, source: KPSS patent database
Average citation	Scaling the total citations by the number of patents filed by the firm in a year, to measure the average citation of each patent source: KPSS patent database
Average value	Scaling the total patent value by the number of patents filed by the firm in a year, to measure the average value of each patent source: KPSS patent database
Innovation Efficiency (patents)	Innovation efficiency measure constructed by dividing the total number of patents filed by the firm in a year to R&D expenditures in the same year, source: KPSS patent database and Compustat
Innovation Efficiency (citations)	Innovation efficiency measure constructed by dividing the total number of citations received by the firm in a year to R&D expenditures in the same year, source: KPSS patent database and Compustat
Patents/Employees	Annual number of patents applied for and eventually granted per 1,000 firm employees (Innovation productivity), source: KPSS patent database and Compustat
Capital expenditures	Capital expenditure, scaled by sales, source: Compustat
R&D expenditure	R&D expense, scaled by sales, source: Compustat
Covariates	
Size	Natural logarithm of total assets.
Tobin's q	Book value of assets (at) minus book value of common equity (ceq) plus the market value of common equity (csho*prcc_f) scaled by total assets (at)
Fixed Assets/Assets (Tangibility)	Fixed assets (ppent) divided by total assets (at)
ROA (Return on Assets)	Earnings before interest and tax (ebit) divided by total assets (at)
Firm age	Number of years listed on Compustat, source: Compustat

Figure 3.1: Distribution of Union Votes



Notes: The figure presents distribution of vote-shares for the entire sample. It has 40 equally spaced bins for vote-shares. In order for the regression discontinuity (RD) design to be valid, there must be no deliberate manipulation of election outcomes right around the 50% vote-share threshold. In other words, there must not be systematic sorting of firms, within close proximity to that threshold. Such sorting would be visible from a discontinuity in the vote-share distribution at the 50% vote-share threshold.

Table 3.3: Distribution of Election Timings (Based on Merged Data from NLRB, Compustat and KPSS)

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	
Number of Firms	20	15	12	11	16	9	6	7	7	3	8	
Year	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Number of Firms	3	5	7	7	6	6	3	2	4	3	3	2

Table 3.4: Distribution of Unionization Timings (Based on Merged Data from NLRB, Compustat and KPSS)

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	
Number of Firms	8	5	4	3	2	3	3	0	2	1	5	
Year	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Number of Firms	1	2	0	6	4	3	2	1	3	2	1	2

Table 3.5: Summary Statistics Using the NLRB Information (Based on Merged Data from NLRB, Compustat and KPSS)

	count	mean	sd	min	max	p25	p50	p75
closing_method	165	4.593939	.4925911	4	5	4	5	5
num_elig_employees	165	143.1758	362.0207	6	4064	23	59	126
N_franchise	165	3.230303	9.613649	1	119	1	1	3
n_estab	165	.4424242	.4981859	0	1	0	0	1
democrat_indicator	121	.6446281	.480616	0	1	0	1	1
support	163	.4962829	.2312428	.0519481	1	.3227513	.4516129	.6382979
manufac_dummy	165	.6787879	.4683635	0	1	0	1	1
N	165							

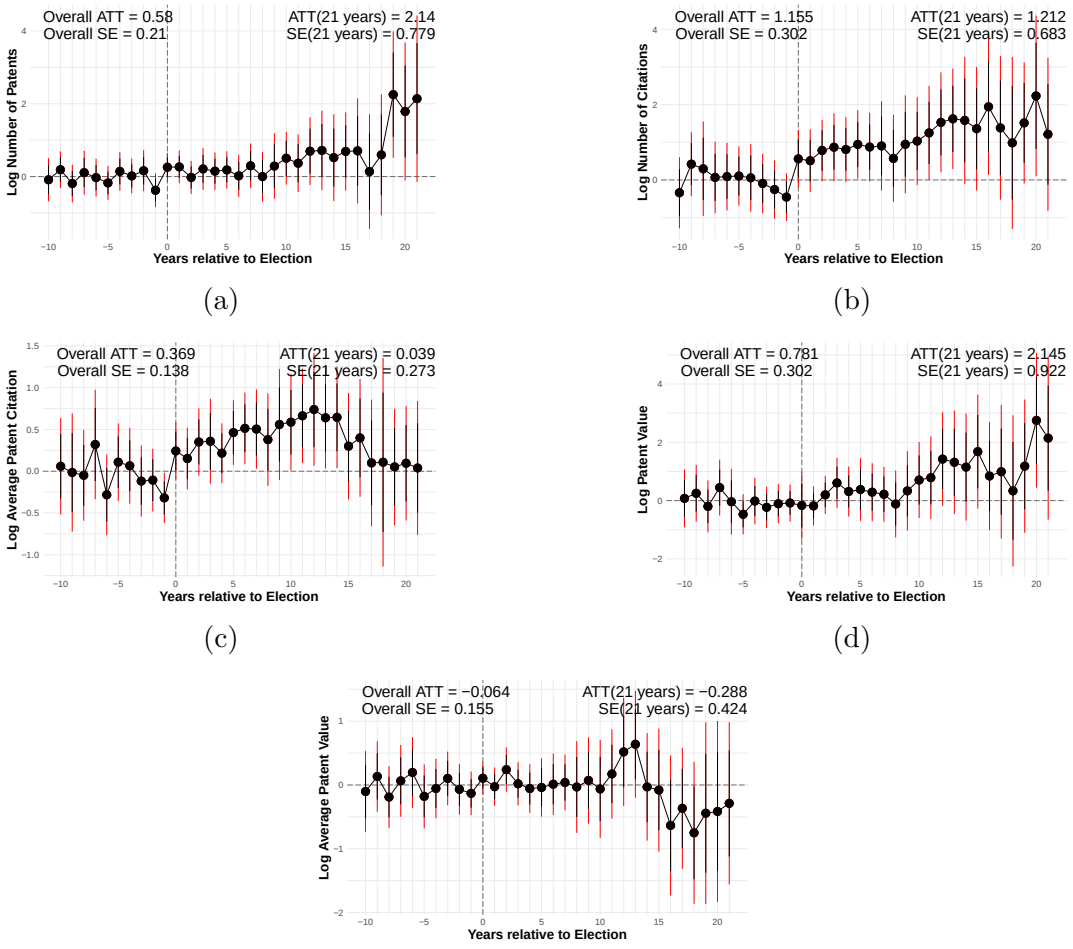
Notes: This table presents the summary statistics for several variables obtained from NLRB (National Labor Relations Board) data, along with the fraction of firms operating in states led by Democrats and states with "right to work" laws. This table includes the summary statistics for a couple of variables from NLRB data and fraction of firms operating in democrat states as well as states with *right to work* laws. *closing_method* captures the outcome of the unionization election. A value of "4" indicates a win or certification of a representative, while "5" represents a loss or certification of the results. *num_elig_employees* indicates the number of eligible voters in the unionization election. It reflects the total count of employees who had the right to participate in the voting process. *N_franchise* represents the number of franchises associated with each firm. *n_estab* is a binary indicator that takes a value of 1 if the firm operates multiple establishments, and a value of 0 if the firm operates only one establishment. *support* is calculated by dividing the number of votes in favor of unionization by the total number of eligible voters. It provides a measure of the level of support for unionization among the employees. *democrat_indicator* is a binary variable that takes a value of one if the state in which the firm operates is led by Democrats. It is equal to zero if the state is led by Republicans, according to the 2020 presidential elections. Finally, the variable *manufac_dummy* is a binary indicator that takes a value of 1 if the firm conducting the election belongs to the manufacturing industry, and a value of 0 if it does not.

Table 3.6: Summary Statistics: Status of Innovation-related Outcome Variables and Covariates in 2002

	count	mean	sd	min	max	p25	p50	p75
No_patents	93	56.34409	173.7821	1	1489	2	7	23
sum_cite	93	1018.204	3110.576	0	23021	32	122	414
sum_xi_real	93	858.2173	3763.64	.6113733	35231.95	7.50298	41.87725	359.8228
average_cite	93	19.39304	20.84304	0	101.25	8.2	12.88889	18.88889
average_value	93	14.04998	26.02105	.3056867	175.564	2.531859	5.990236	13.95908
no_patents_RandD	71	.7633167	1.021872	.018797	6.758742	.1997736	.4314534	.8846154
no_cites_RandD	71	17.62863	44.4198	.1030928	309.633	2.212145	6.005459	15.4669
no_patents_emp	89	5.475062	14.91927	.0068966	108.2353	.2962963	1.337409	3.75
capx_sale	93	.0446343	.0420932	.0006368	.3728505	.0243524	.036534	.0571073
R_and_D_sale	93	.0442964	.1502727	0	1.035935	.0014815	.0139688	.0298872
age_compustat	93	30.6129	18.71586	0	52	11	34	52
ppent_at	93	.3167578	.188376	.0072448	.8341225	.1694549	.2779251	.4461737
ebit_at	93	.0783387	.0822183	-.2115063	.2550609	.0353663	.0791205	.1312375
book_l	93	.2930231	.2213132	0	1.466121	.145668	.2683689	.4083494
tobin_q	88	1.694766	.9033234	.7061067	5.414176	1.172484	1.379999	1.91443
size	93	7.699661	1.769312	2.270268	13.26255	6.535444	7.800724	8.8578
N	93							

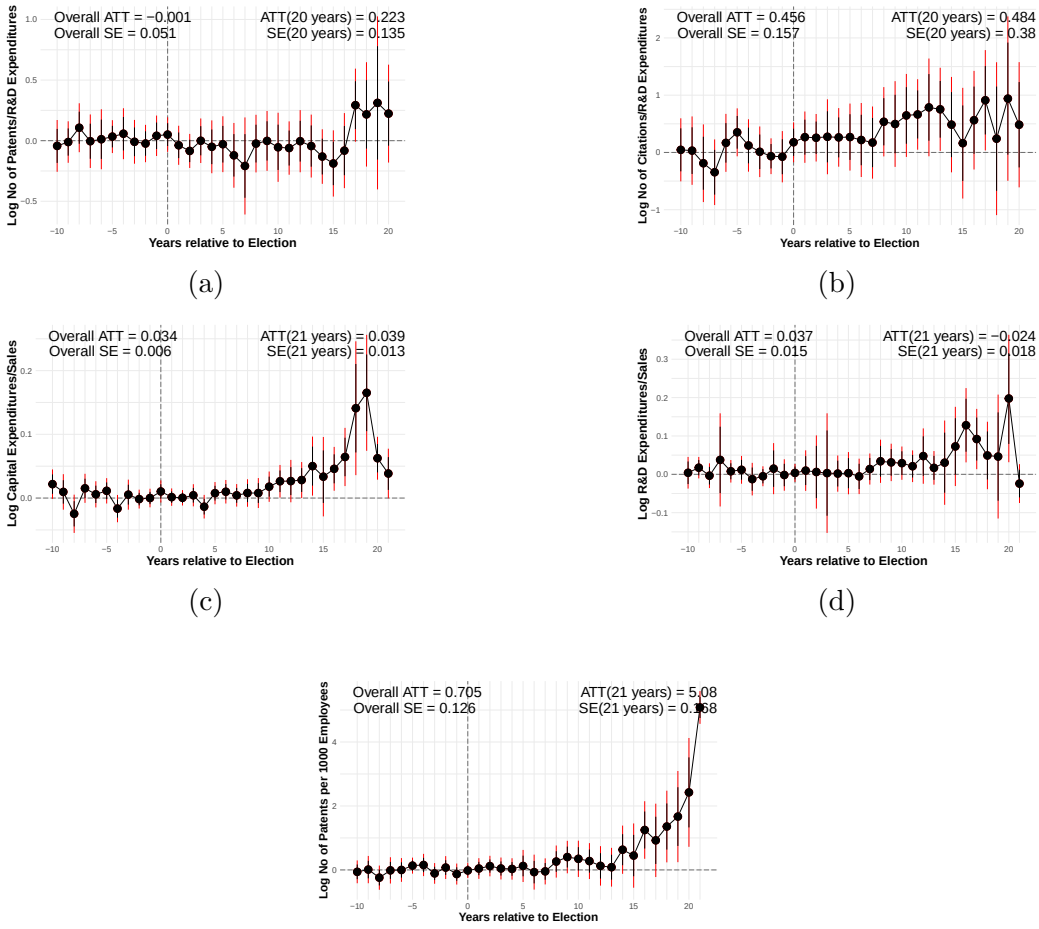
Notes: This table summarizes the status of innovation-related outcome variables and covariates in the year 2002. Among the covariates, the variable *size* represents the logarithm of total assets, indicating the size of the firm. *tobin_q* measures the difference between the book value of assets and the book value of common equity, plus the market value of common equity, scaled by total assets. It serves as an indicator of the firm's investment opportunities. *ebit_at* is the return on assets (ROA), calculated by dividing earnings before interest and tax (EBIT) by total assets. It provides insight into the firm's profitability. *ppent_at* represents tangibility, specifically the ratio of fixed assets to total assets. *age_compustat* denotes the number of years listed on Compustat for a firm and *book_l* denotes book leverage, which is calculated by dividing the sum of long-term debt and debt in current liabilities by total assets. Moving on to the outcome variables, the variable *No_patents* captures the number of patents. *sum_cite* denotes the summation of forward citations for all the patents filed by a firm in a year. *sum_xi_real* denotes the summation of the value of innovation of all patents filed by a firm deflated to 1982 (million) dollars using the CPI. *average_cite* is calculated by dividing the total number of citations by the number of patents filed by the firm in a year to measure the average citation of each patent. For *average_value* we also scale the total patent value by the number of patents filed by the firm in a year to measure the average value of each patent. We construct two innovation efficiency measures, *no_patents_RandD* is the number of patents granted, scaled by R&D expenditures and *no_cites_RandD* is equal to the number of patent citations, scaled by total R&D expenditures. As a proxy for innovation productivity, *no_patents_emp*, estimate the number of patents per 1,000 firm employees. To capture the input to the innovation process, *capx_sale*, is the total capital expenditure divided by total sales and *R_and_D_sale* represents the total R&D expenditure divided by total sales. For further information on the definition of these variables, please refer to the table 3.2 in this chapter.

Figure 3.2: Event Study Plots for Innovation Outcome Variables- Subset of Firms Holding Elections for Unions- Conditional Parallel Trends Assumption (part 1)



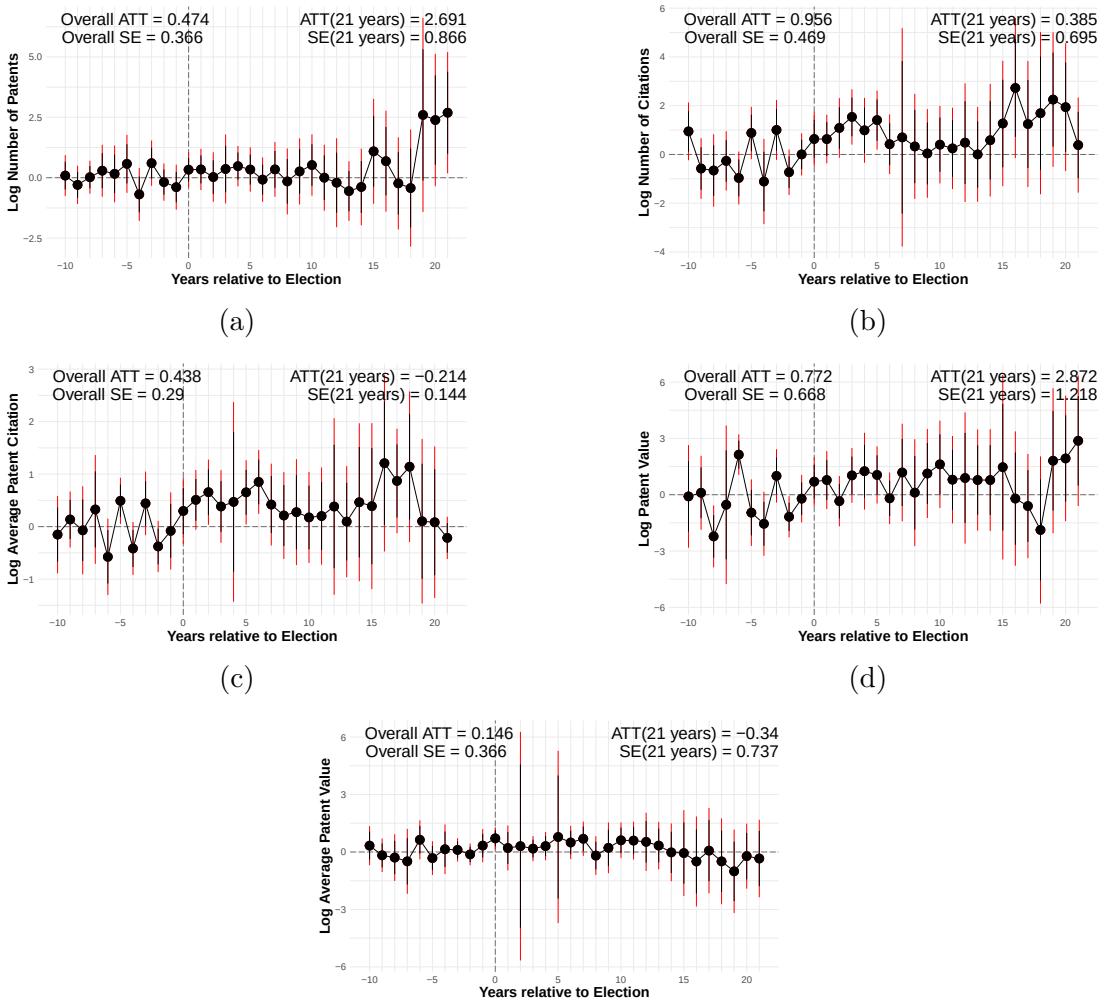
Notes: The figures present dynamic difference-in-difference estimates for the subset of firms holding elections for unionization, under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of *number of patents, number of citations, average citations, market value of patents and average value.* The covariates are size, tangibility, profitability, tobin's q and age of the firm. "Years relative to Election" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after holding the election, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 3.3: Event Study Plots for Innovation Outcome Variables- Subset of Firms Holding Elections for Unions- Conditional Parallel Trends Assumption (part 2)



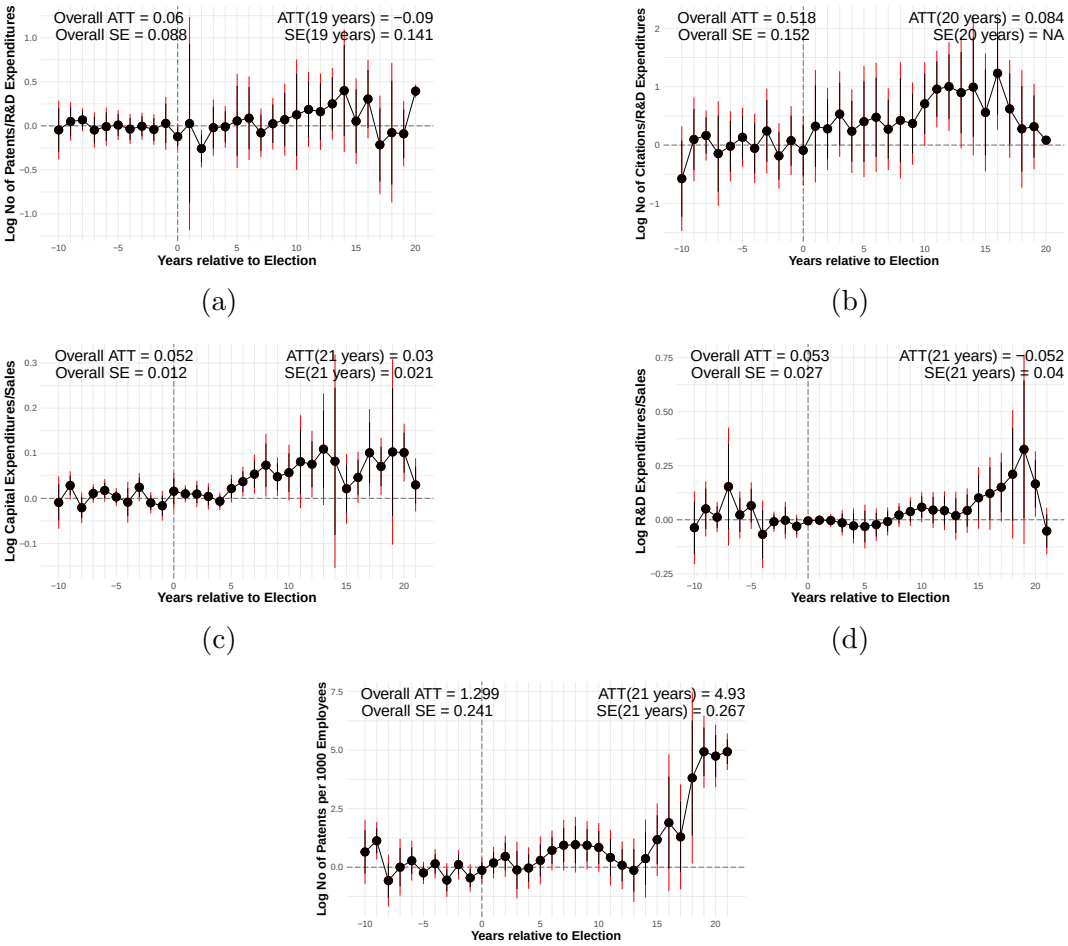
Notes: The figures present dynamic difference-in-difference estimates for the subset of firms holding elections for unionization, under unconditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of *number of patents to R&D expenditures ratio, number of citations to R&D expenditures ratio, capital expenditures to sales ratio, R&D expenditures to sales ratio and number of patents per 1000 employees.* The covariates are size, tangibility, profitability, tobin's q and age of the firm. "Years relative to Election" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after holding the election, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 3.4: Event Study Plots for Subset of Successful Elections (part 1)



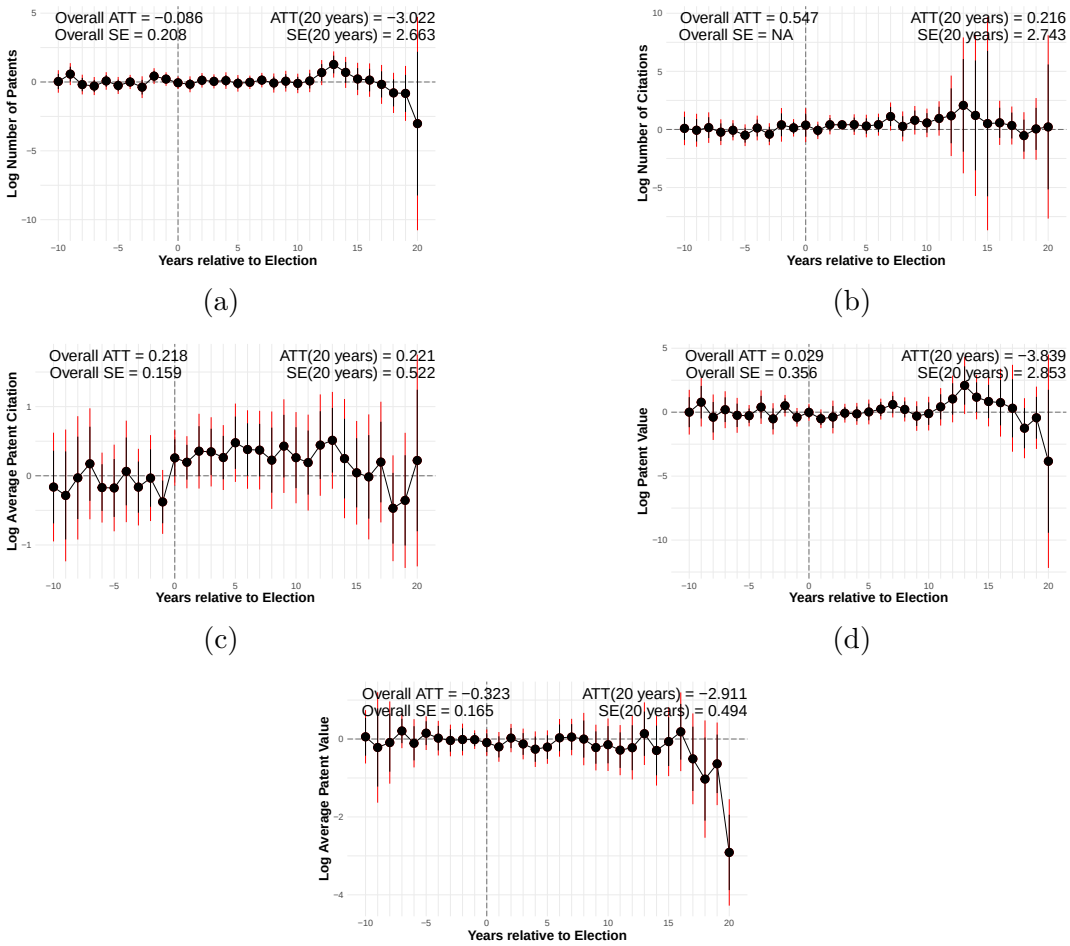
Notes: The figures present dynamic difference-in-difference estimates for the subset of elections resulting in a "win", under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of *number of patents, number of citations, average citations, market value of patents and average value.* The covariates are size, tangibility, profitability, tobin's q and age of the firm. "Years relative to Election" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after unionization, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 3.5: Event Study Plots for Subset of Successful Elections (part 2)



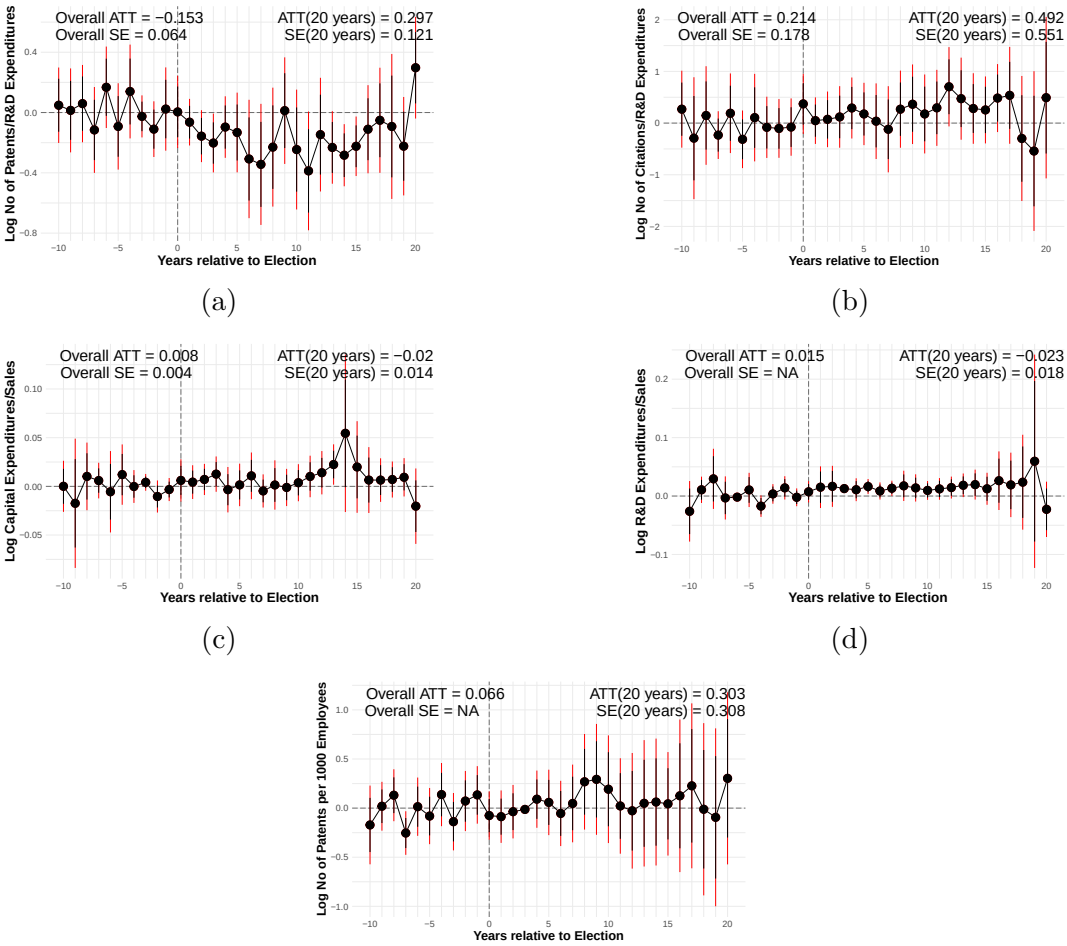
Notes: The figures present dynamic difference-in-difference estimates for the subset of elections resulting in a "win", under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of *number of patents to R&D expenditures ratio, number of citations to R&D expenditures ratio, capital expenditures to sales ratio, R&D expenditures to sales ratio and number of patents per 1000 employees.* The covariates are size, tangibility, profitability, tobin's q and age of the firm. "Years relative to Election" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after unionization, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 3.6: Event Study Plots for Subset of Unsuccessful Elections (part 1)



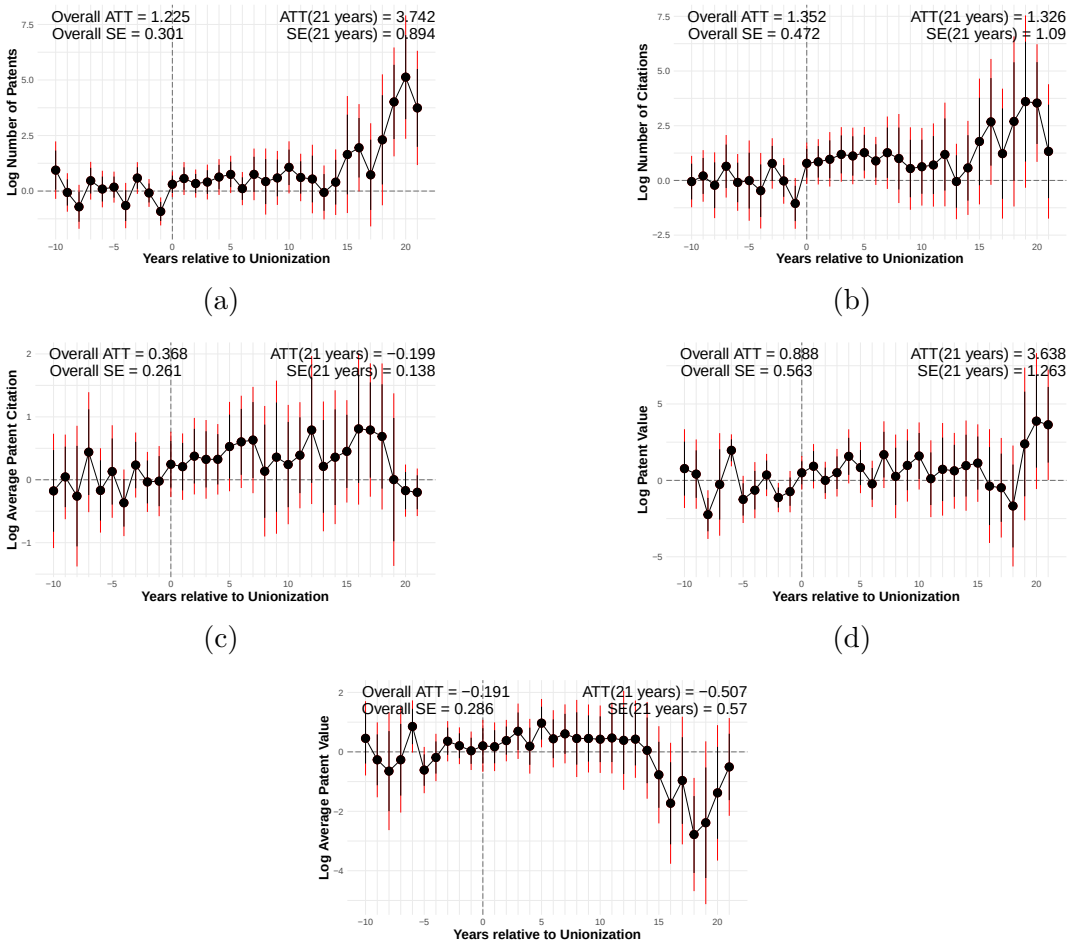
Notes: The figures present dynamic difference-in-difference estimates for the subset of elections resulting in a "loss", under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of *number of patents, number of citations, average citations, market value of patents and average value.* The covariates are size, tangibility, profitability, tobin's q and age of the firm. "Years relative to Election" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after holding the election, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 3.7: Event Study Plots for Subset of Unsuccessful Elections (part 2)



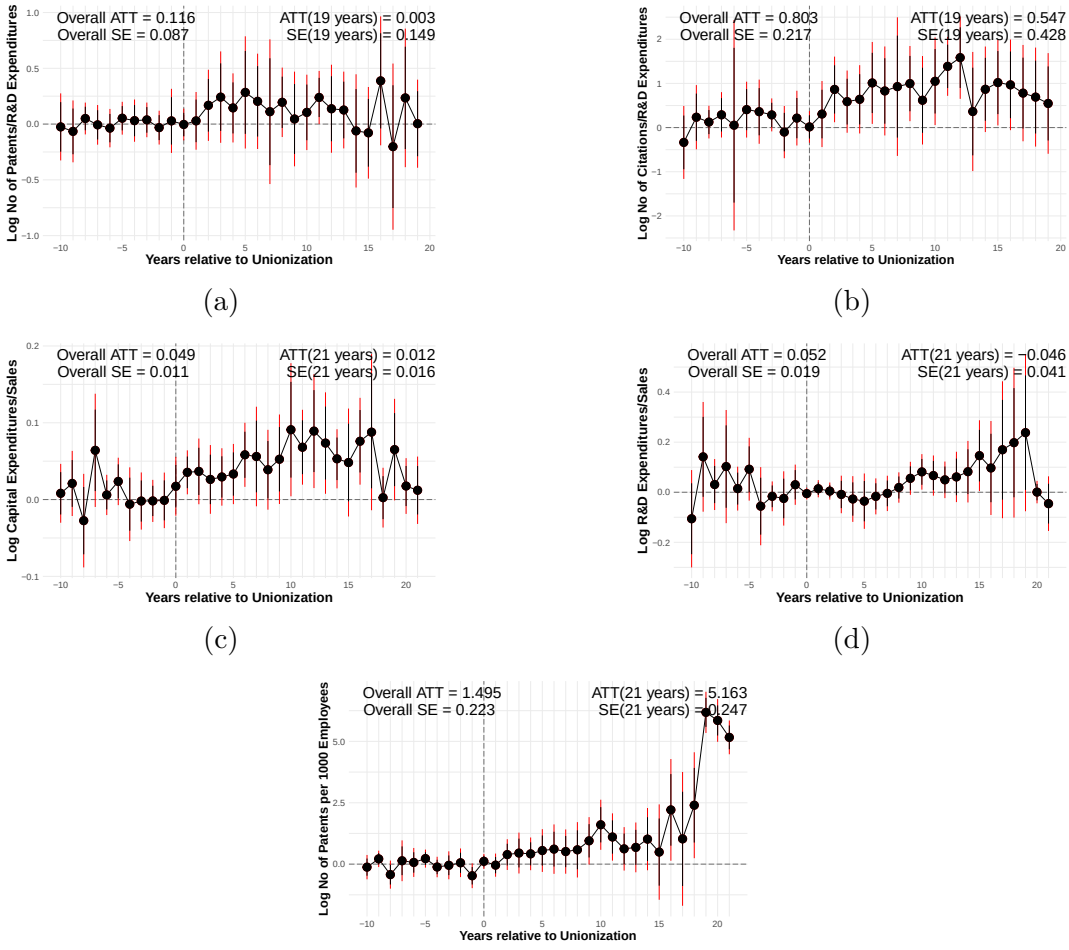
Notes: The figures present dynamic difference-in-difference estimates for the subset of elections resulting in a "loss", under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of *number of patents to R&D expenditures ratio, number of citations to R&D expenditures ratio, capital expenditures to sales ratio, R&D expenditures to sales ratio and number of patents per 1000 employees.* The covariates are size, tangibility, profitability, tobin's q and age of the firm. "Years relative to Election" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after holding the election, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 3.8: Event Study Plots for Subset of Unionized Firms Not Followed by Decertification (part 1)



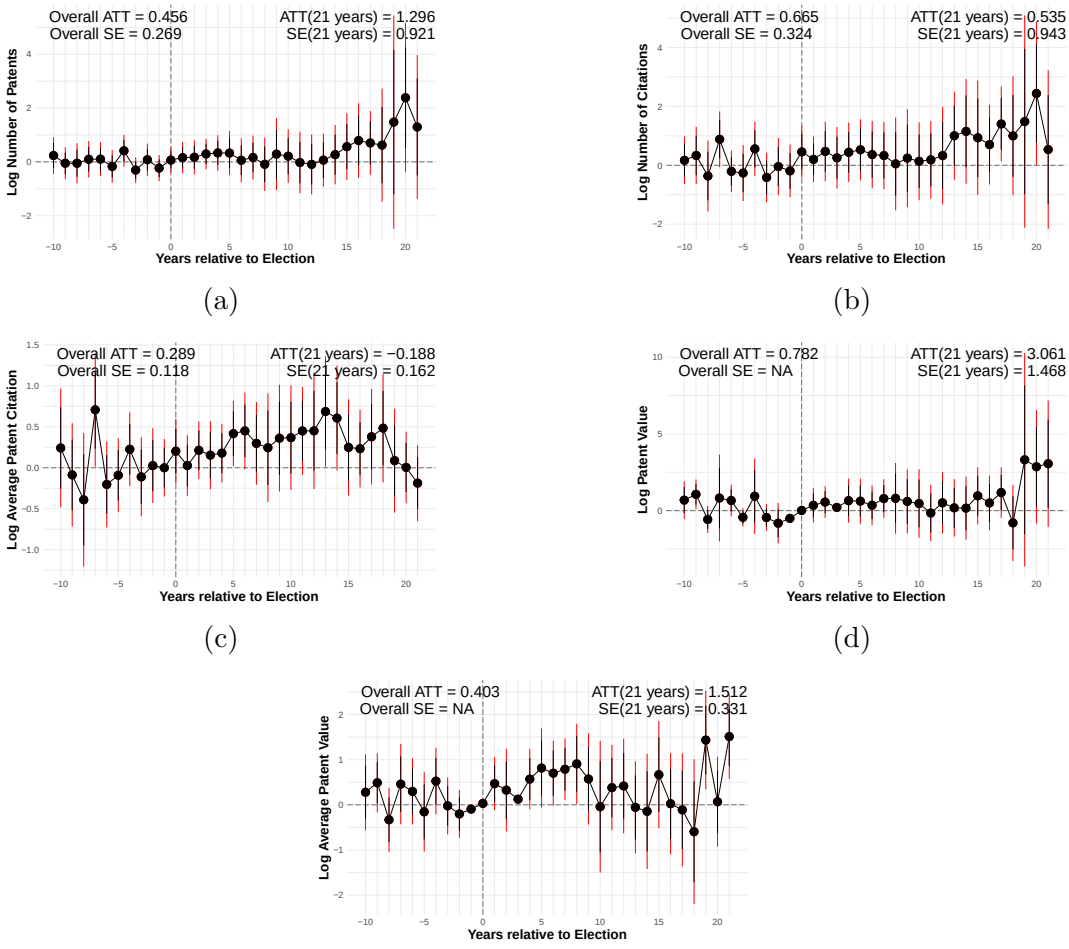
Notes: The figures present dynamic difference-in-difference estimates for the subset of truly unionized firms (first petition RC (win) and no RD later), under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of *number of patents, number of citations, average citations, market value of patents and average value.* The covariates are size, tangibility, profitability, tobin's q and age of the firm. "Years relative to Unionization" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of certifying the union. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after unionization, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 3.9: Event Study Plots for Subset of Unionized Firms Not Followed by Decertification (part 2)



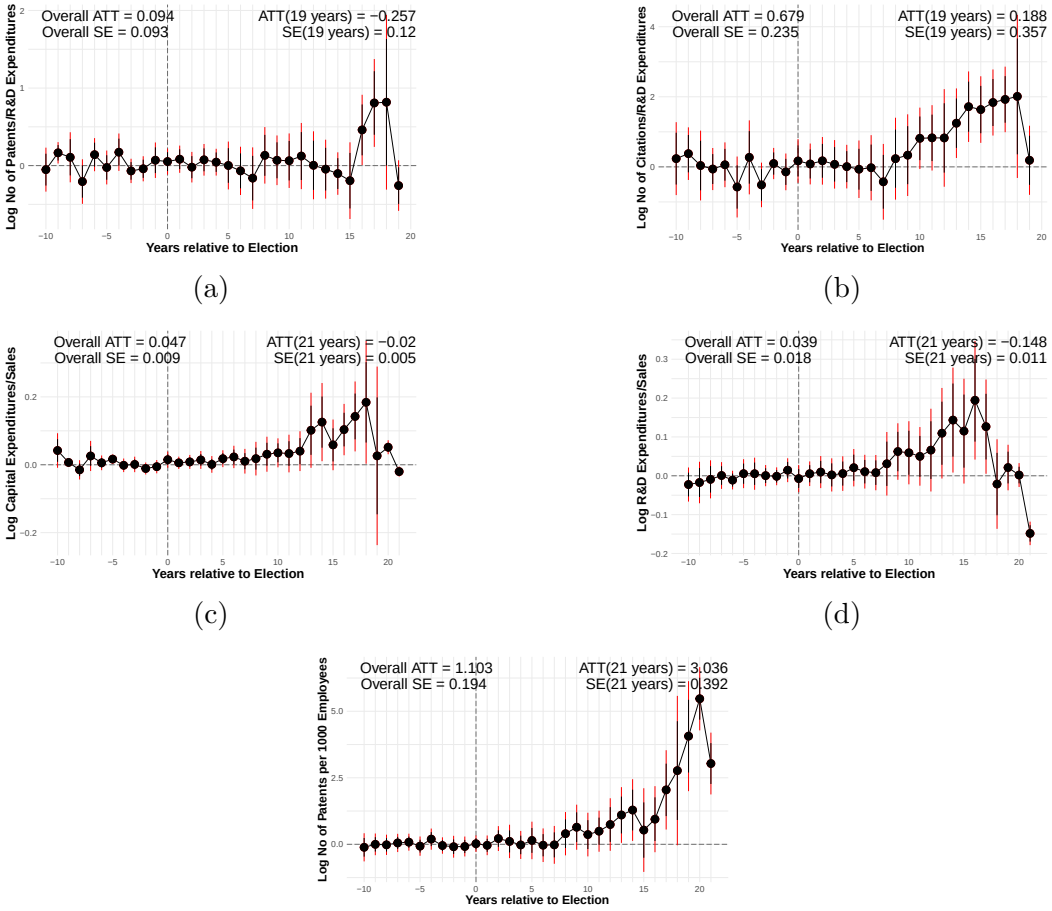
Notes: The figures present dynamic difference-in-difference estimates for the subset of truly unionized firms (first petition RC (win) and no RD later), under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of *number of patents to R&D expenditures ratio, number of citations to R&D expenditures ratio, capital expenditures to sales ratio, R&D expenditures to sales ratio and number of patents per 1000 employees.* The covariates are size, tangibility, profitability, tobin's q and age of the firm. "Years relative to Unionization" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of certifying the union. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after unionization, and so on. For each e, the *ATT* is calculated by averaging dynamic treatment effects across all groups. The overall *ATT* is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 3.10: Event Study Plots for Subset of Firms in Democrat-led States which hold elections (part 1)



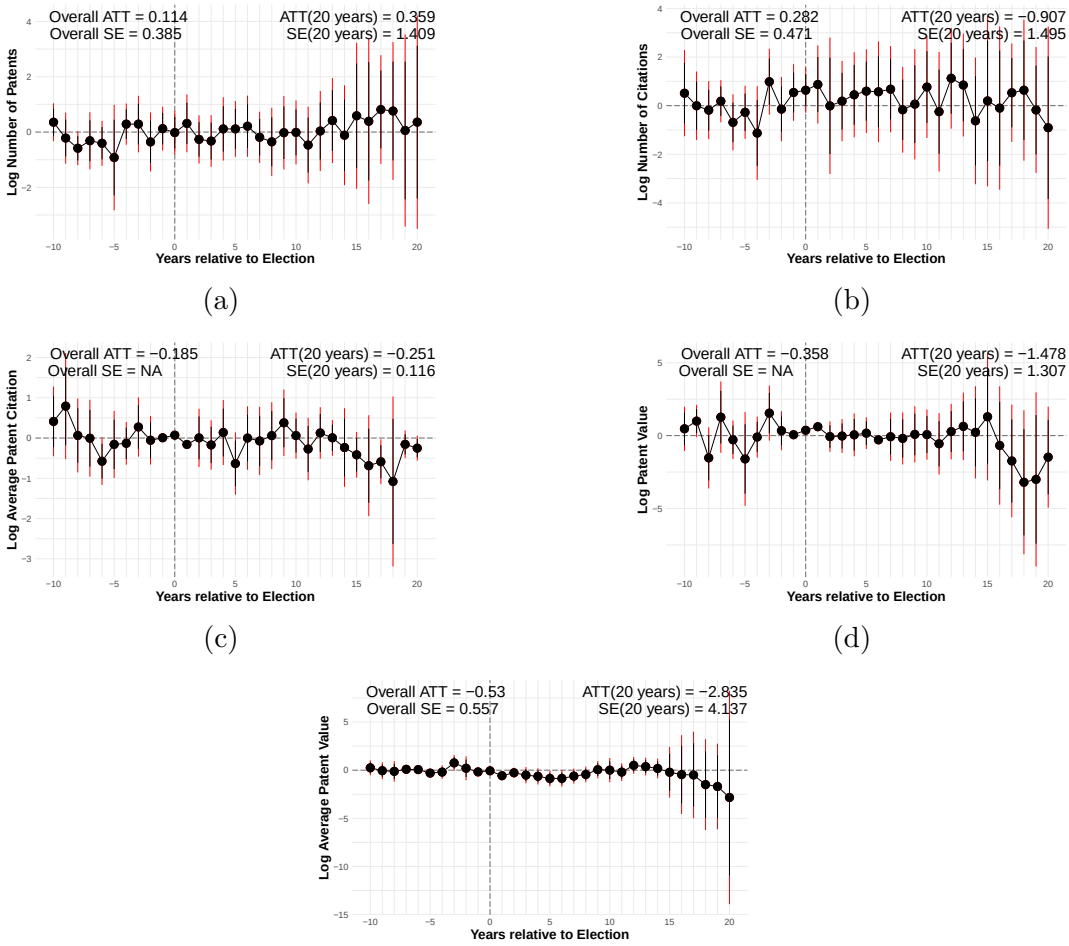
Notes: The figures present dynamic difference-in-difference estimates for the subset of firms in Democrat-led States which hold election, under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of *number of patents, number of citations, average citations, market value of patents and average value.* The covariates are size, tangibility, profitability, tobin's q and age of the firm. "Years relative to Election" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after holding the election, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 3.11: Event Study Plots for Subset of Firms in Democrat-led States which hold elections (part 2)



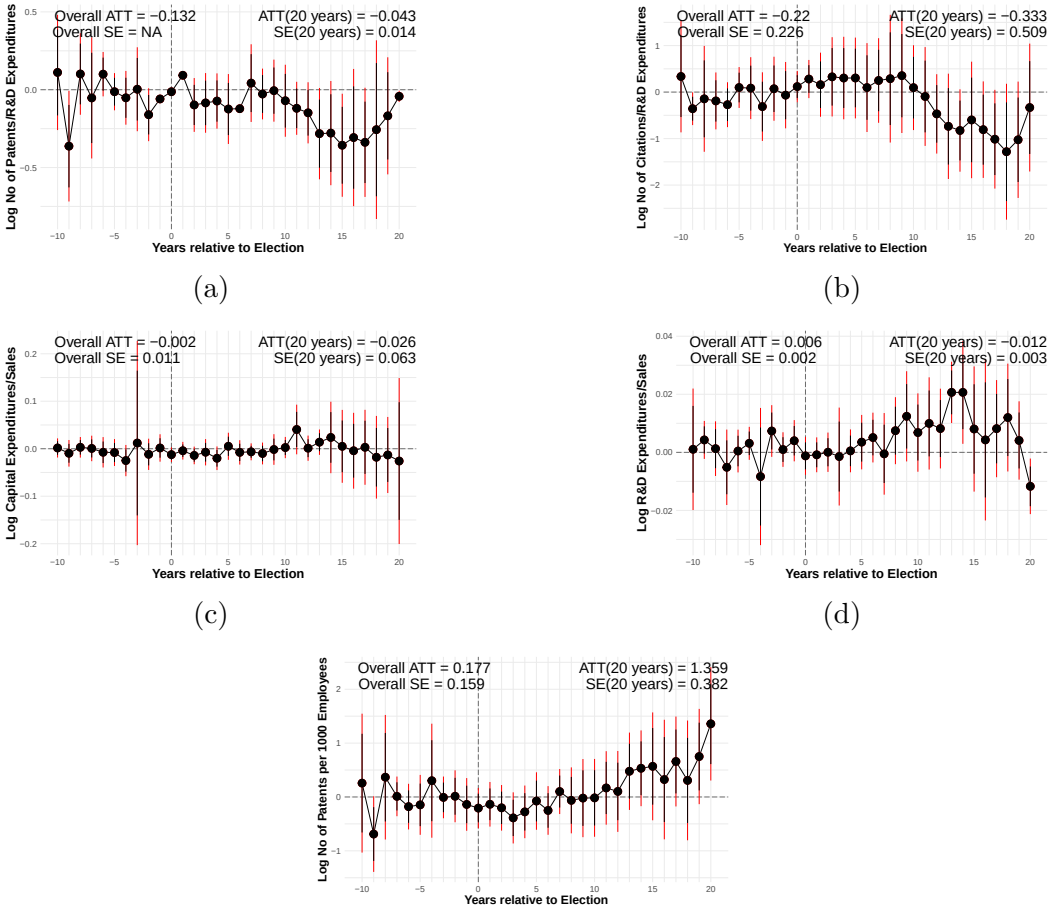
Notes: The figures present dynamic difference-in-difference estimates for the subset of firms in Democrat-led States which hold election, under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of *number of patents to R&D expenditures ratio, number of citations to R&D expenditures ratio, capital expenditures to sales ratio, R&D expenditures to sales ratio and number of patents per 1000 employees*. The covariates are size, tangibility, profitability, tobin's q and age of the firm. "Years relative to Election" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after holding the election, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 3.12: Event Study Plots for Subset of Firms in Republic-led States which hold elections (part 1)



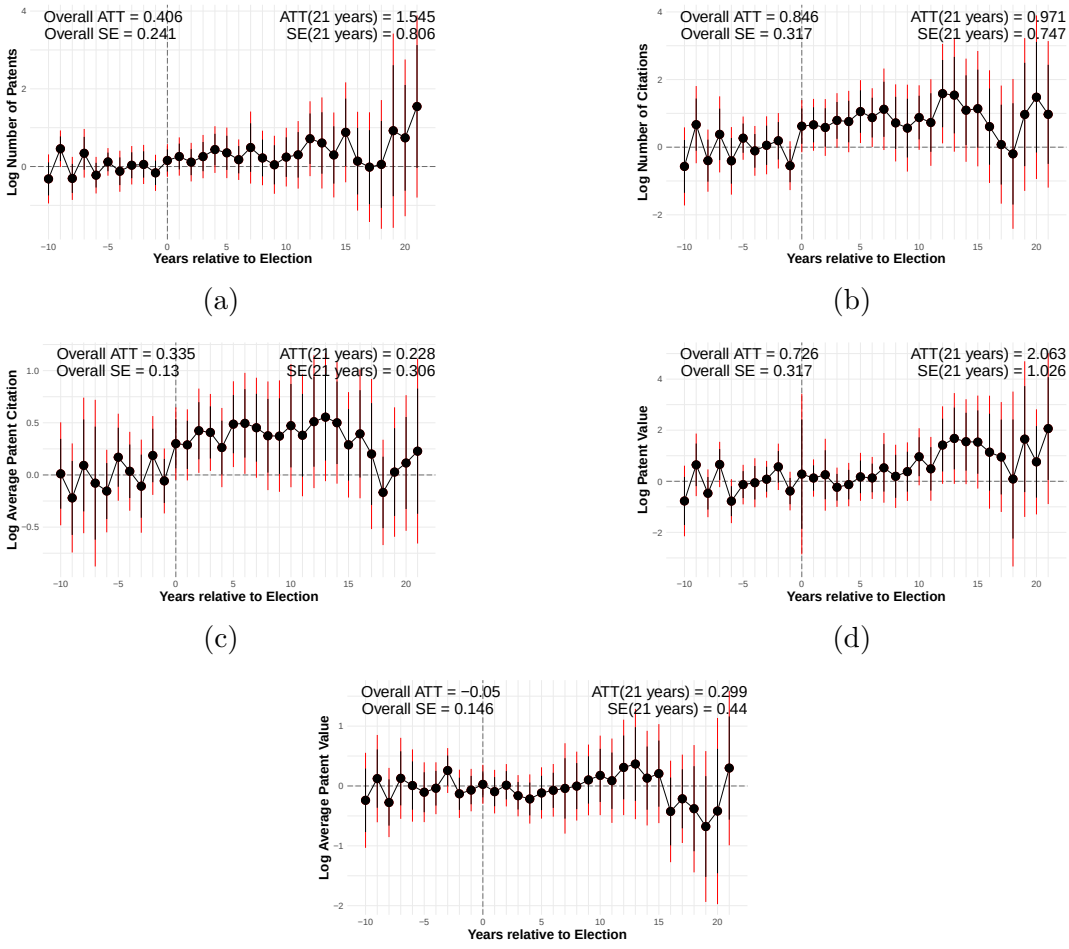
Notes: The figures present dynamic difference-in-difference estimates for the subset of firms in Republic-led States which hold election, under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of *number of patents, number of citations, average citations, market value of patents and average value.* The covariates are size, tangibility, profitability, tobin's q and age of the firm. "Years relative to Election" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after holding the election, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 3.13: Event Study Plots for Subset of Firms in Republic-led States which hold elections (part 2)



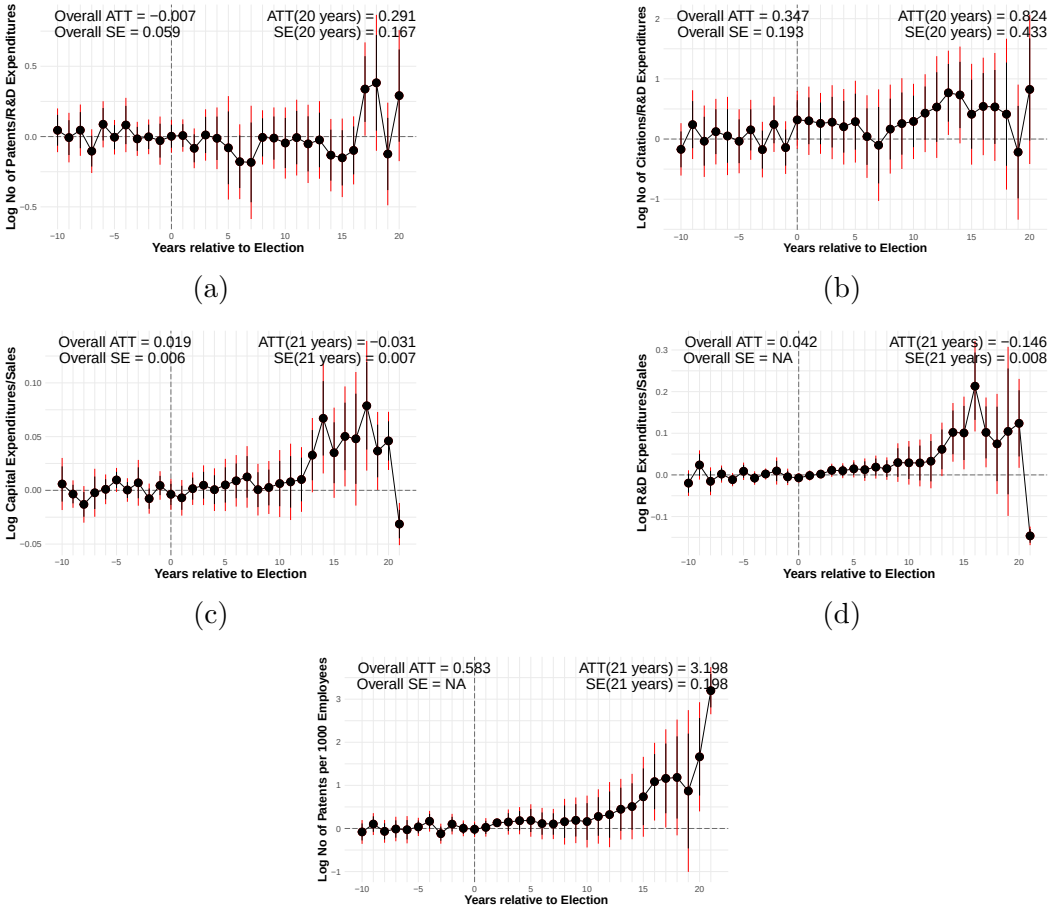
Notes: The figures present dynamic difference-in-difference estimates for the subset of firms in Republic-led States which hold election, under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of *number of patents to R&D expenditures ratio, number of citations to R&D expenditures ratio, capital expenditures to sales ratio, R&D expenditures to sales ratio and number of patents per 1000 employees*. The covariates are size, tangibility, profitability, tobin's q and age of the firm. "Years relative to Election" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after holding the election, and so on. For each e, the *ATT* is calculated by averaging dynamic treatment effects across all groups. The overall *ATT* is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 3.14: Event Study Plots for Subset of Manufacturing firms which hold elections (part 1)



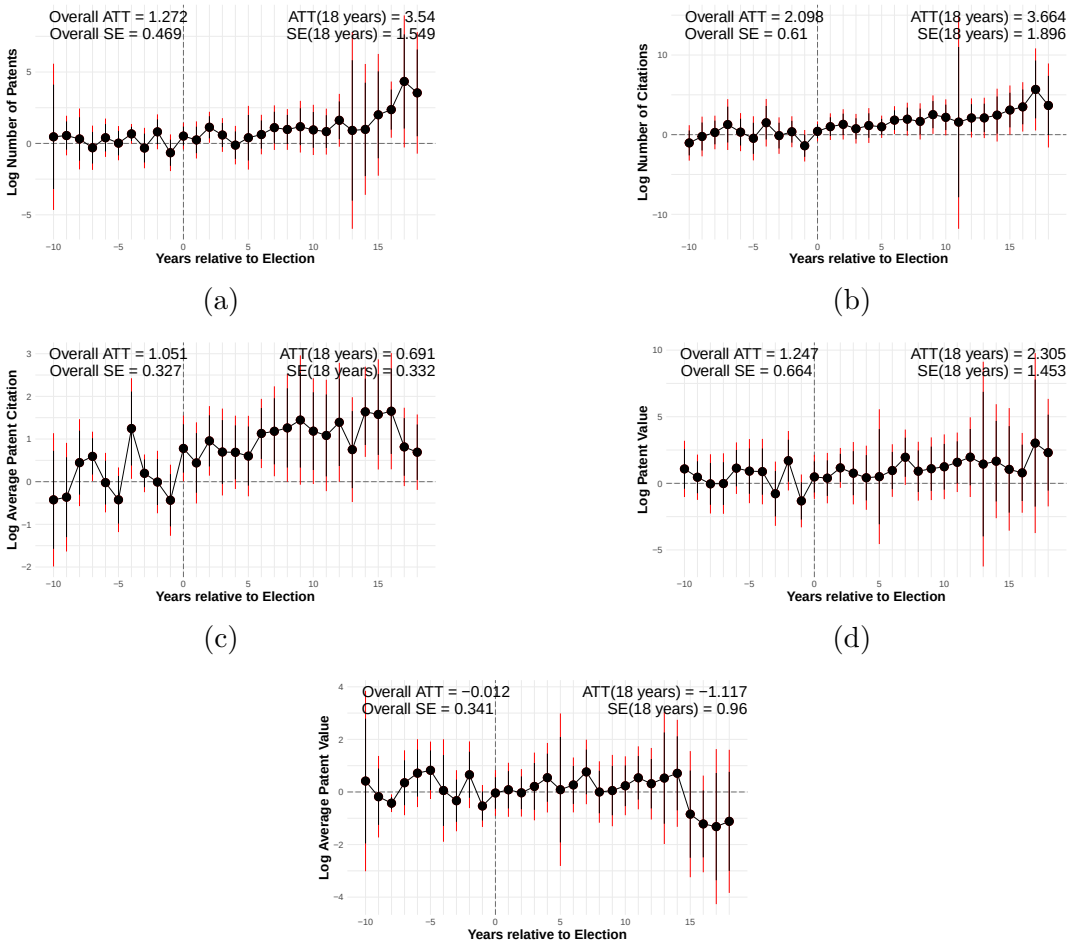
Notes: The figures present dynamic difference-in-difference estimates for the subset of manufacturing firms which hold election, under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of *number of patents, number of citations, average citations, market value of patents and average value.* The covariates are size, tangibility, profitability, tobin's q and age of the firm. "Years relative to Election" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after holding the election, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 3.15: Event Study Plots for Subset of Manufacturing firms which hold elections (part 2)



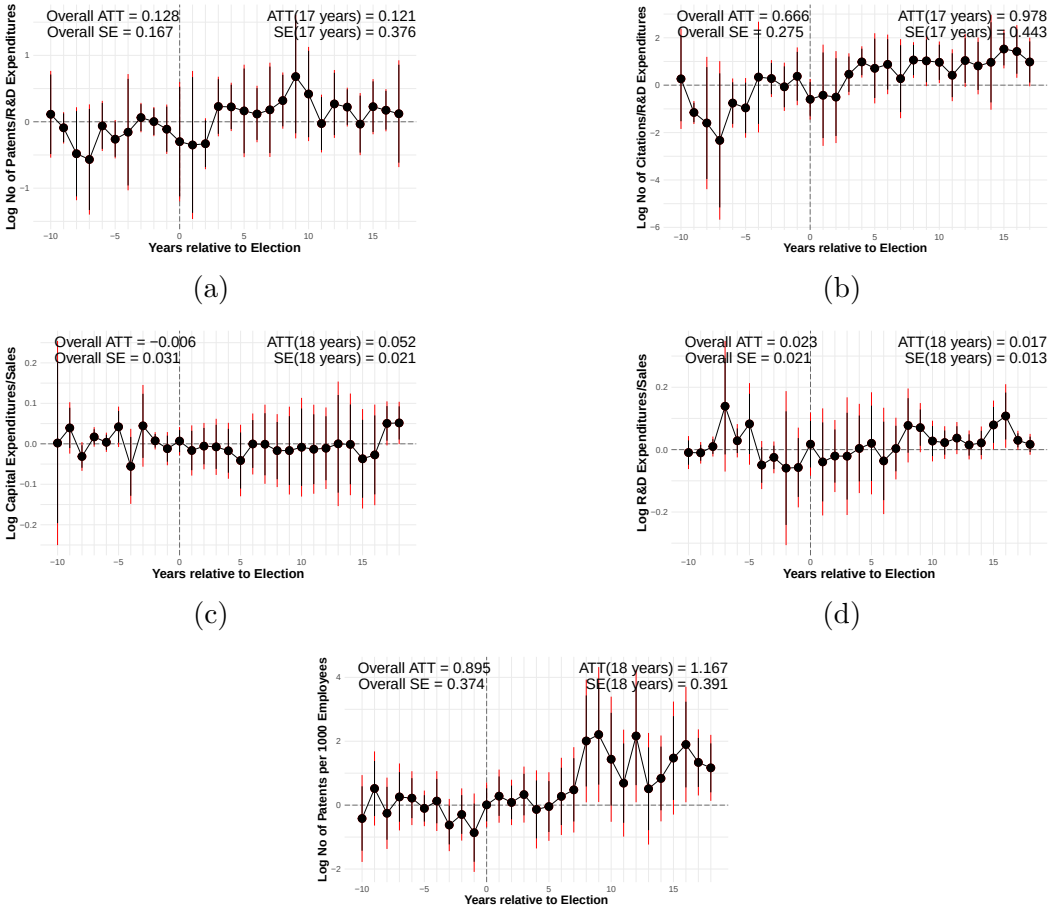
Notes: The figures present dynamic difference-in-difference estimates for the subset of manufacturing firms which hold election, under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of *number of patents to R&D expenditures ratio, number of citations to R&D expenditures ratio, capital expenditures to sales ratio, R&D expenditures to sales ratio and number of patents per 1000 employees.* The covariates are size, tangibility, profitability, tobin's q and age of the firm. "Years relative to Election" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after holding the election, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 3.16: Event Study Plots for Subset of Non-manufacturing firms which hold elections (part 1)



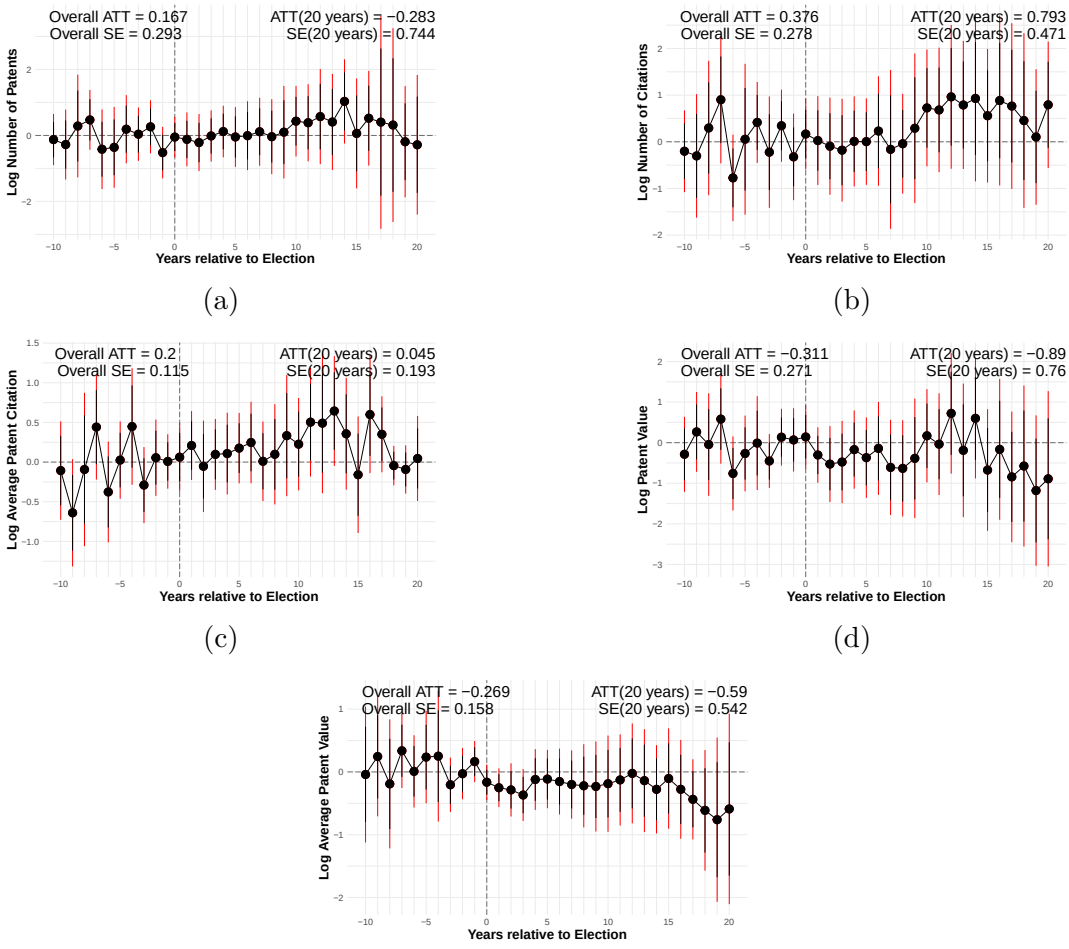
Notes: The figures present dynamic difference-in-difference estimates for the subset of non-manufacturing firms which hold election, under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of *number of patents, number of citations, average citations, market value of patents and average value.* The covariates are size, tangibility, profitability, tobin's q and age of the firm. "Years relative to Election" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after holding the election, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 3.17: Event Study Plots for Subset of Non-manufacturing firms which hold elections (part 2)



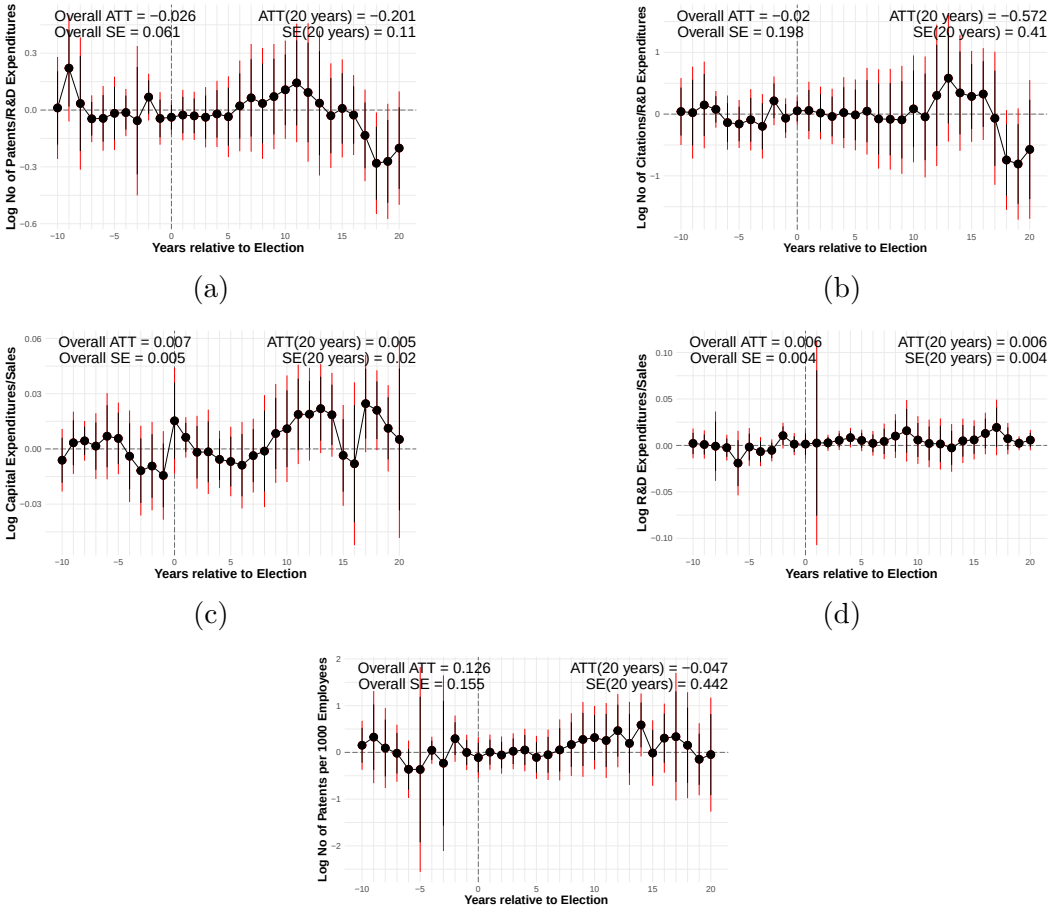
Notes: The figures present dynamic difference-in-difference estimates for the subset of non-manufacturing firms which hold election, under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of *number of patents to R&D expenditures ratio, number of citations to R&D expenditures ratio, capital expenditures to sales ratio, R&D expenditures to sales ratio and number of patents per 1000 employees*. The covariates are size, tangibility, profitability, tobin's q and age of the firm. "Years relative to Election" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after holding the election, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 3.18: Event Study Plots for Subset of large size (above median) firms which hold elections (part 1)



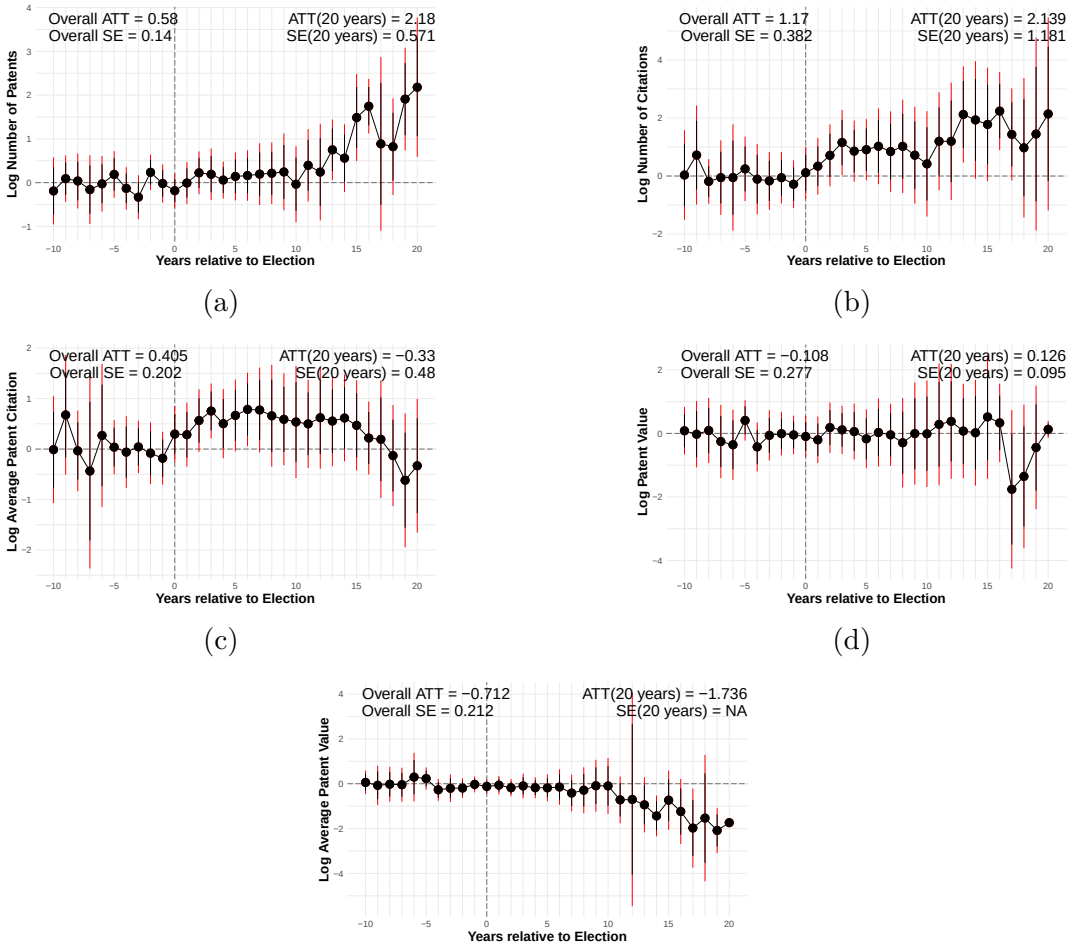
Notes: The figures present dynamic difference-in-difference estimates for the subset of large size (above median) firms which hold election, under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of *number of patents, number of citations, average citations, market value of patents and average value.* The covariates are size, tangibility, profitability, tobin's q and age of the firm. "Years relative to Election" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after holding the election, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 3.19: Event Study Plots for Subset of large size (above median) firms which hold elections (part 2)



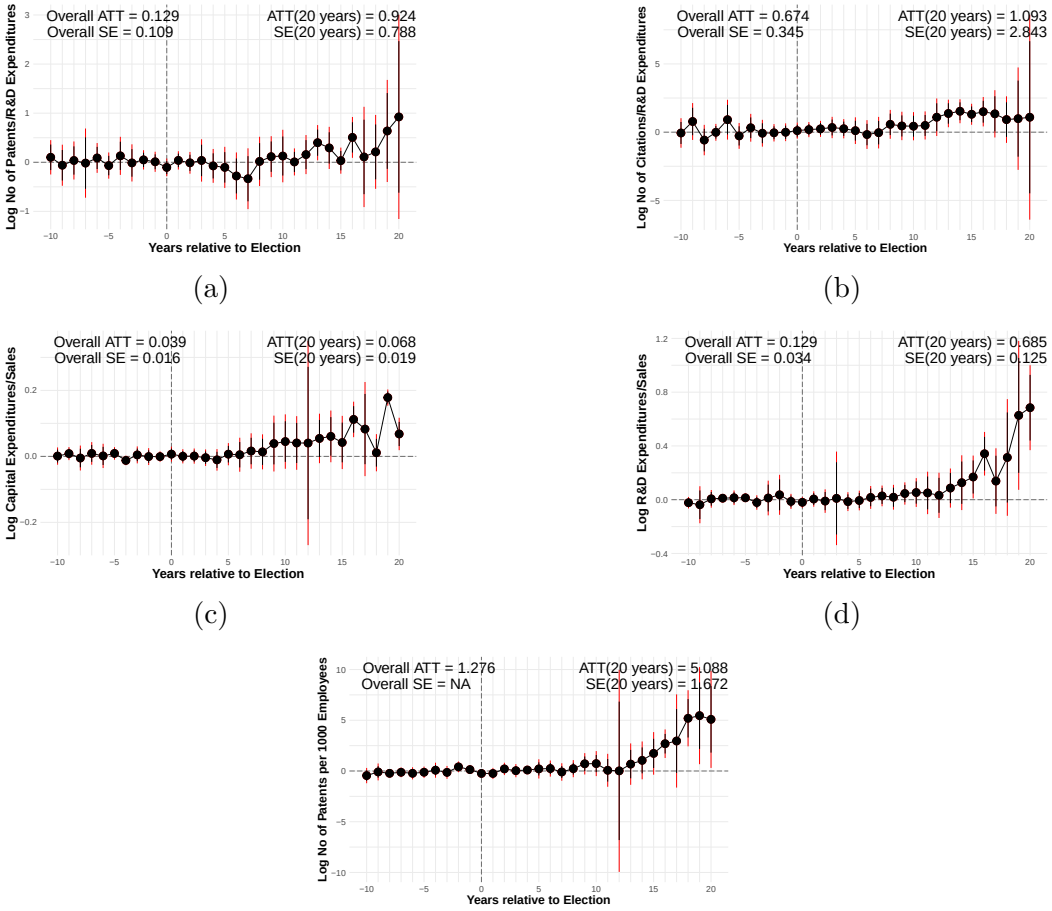
Notes: The figures present dynamic difference-in-difference estimates for the subset of large size (above median) firms which hold election, under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of *number of patents to R&D expenditures ratio, number of citations to R&D expenditures ratio, capital expenditures to sales ratio, R&D expenditures to sales ratio and number of patents per 1000 employees*. The covariates are size, tangibility, profitability, tobin's q and age of the firm. "Years relative to Election" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after holding the election, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 3.20: Event Study Plots for Subset of small size (below median) firms which hold elections (part 1)



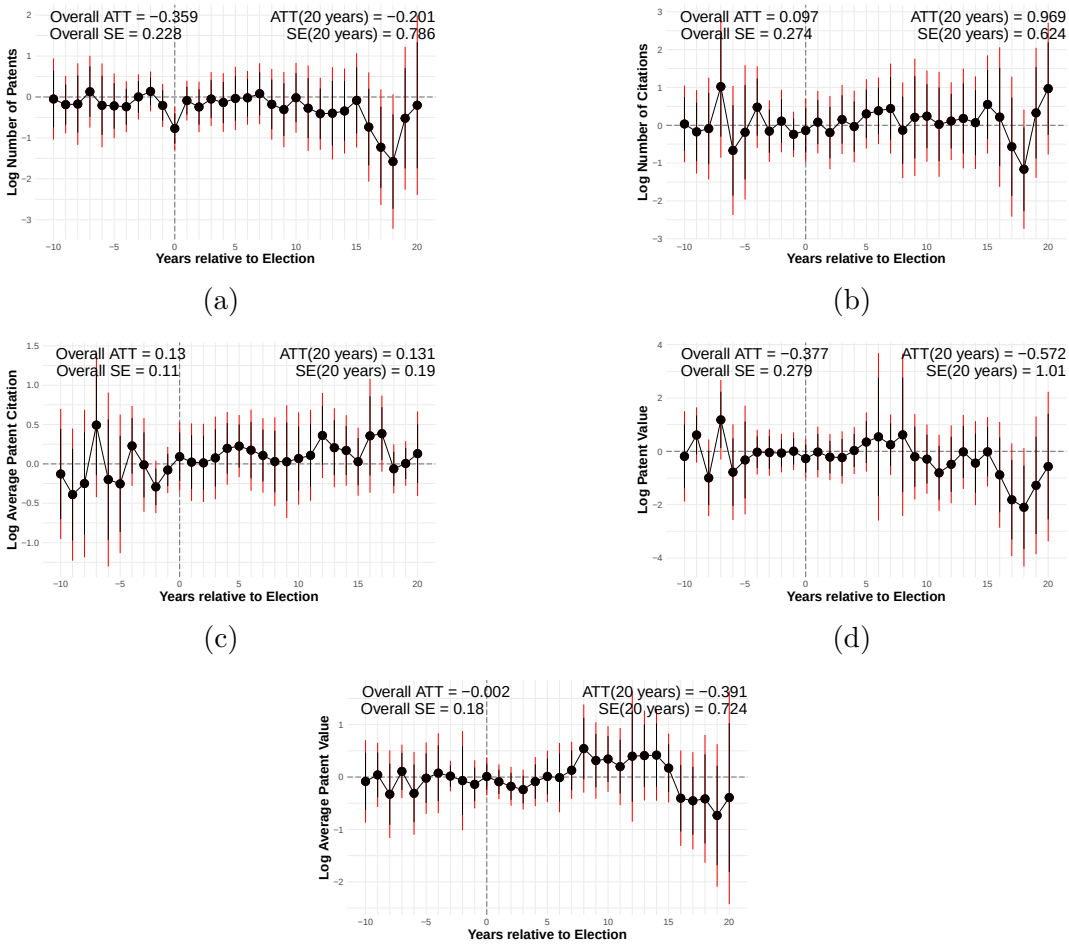
Notes: The figures present dynamic difference-in-difference estimates for the subset of small size (below median) firms which hold election, under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of *number of patents, number of citations, average citations, market value of patents and average value.* The covariates are size, tangibility, profitability, tobin's q and age of the firm. "Years relative to Election" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after holding the election, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 3.21: Event Study Plots for Subset of small size (below median) firms which hold elections (part 2)



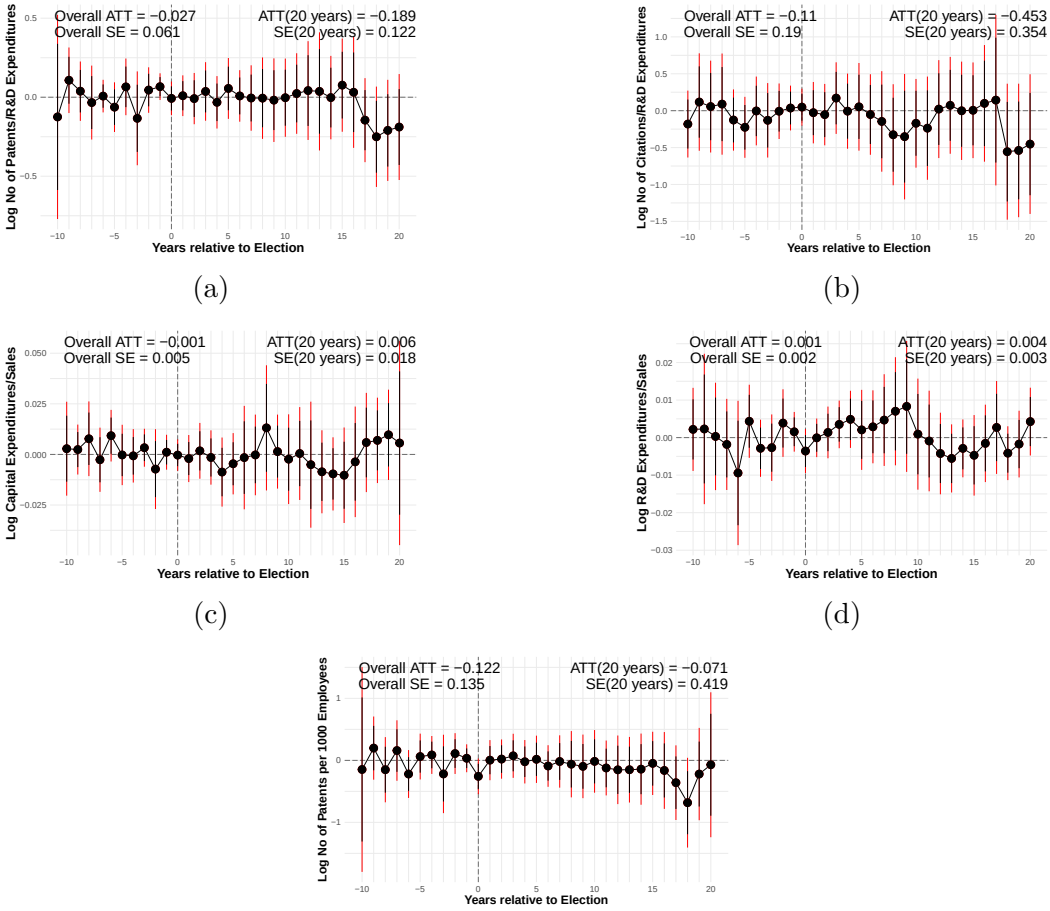
Notes: The figures present dynamic difference-in-difference estimates for the subset of small size (below median) firms which hold election, under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of *number of patents to R&D expenditures ratio, number of citations to R&D expenditures ratio, capital expenditures to sales ratio, R&D expenditures to sales ratio and number of patents per 1000 employees*. The covariates are size, tangibility, profitability, tobin's q and age of the firm. "Years relative to Election" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after holding the election, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 3.22: Event Study Plots for Subset of old age (above median) firms which hold elections (part 1)



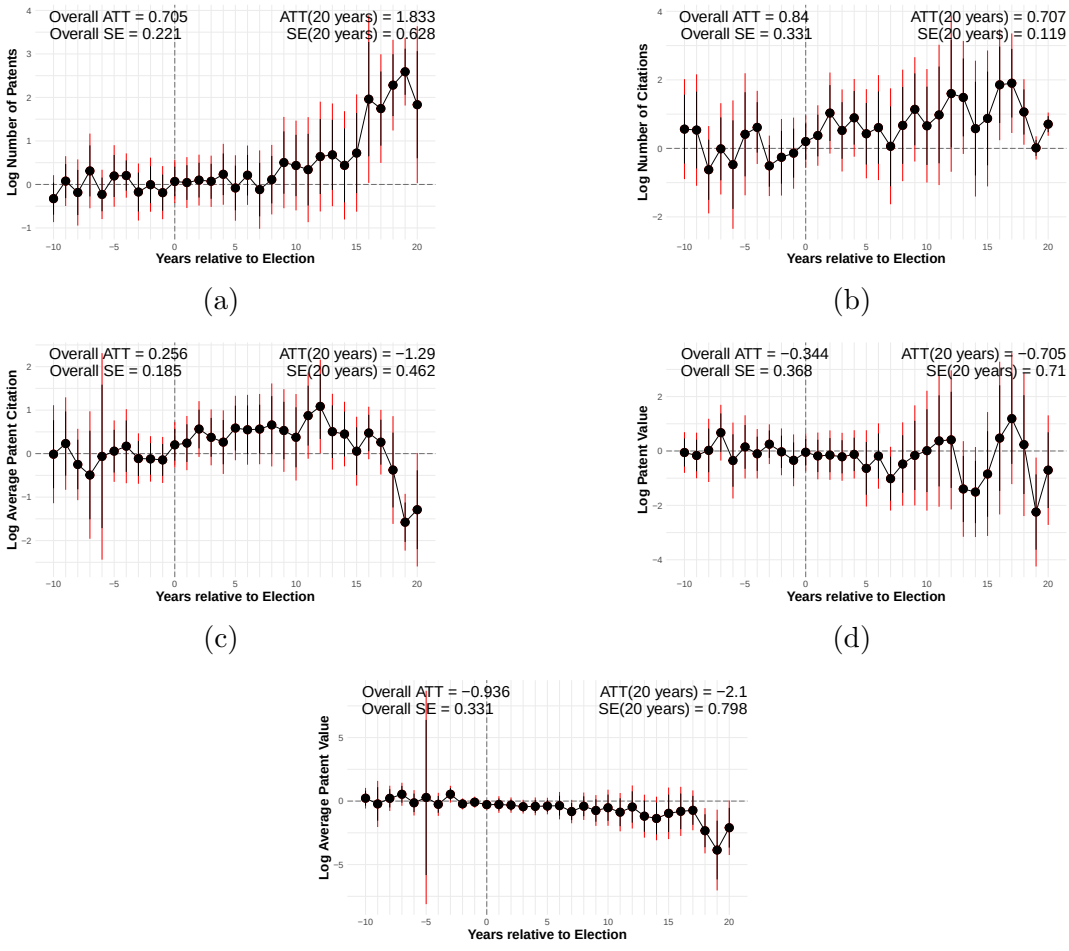
Notes: The figures present dynamic difference-in-difference estimates for the subset of old age (above median) firms which hold election, under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of *number of patents, number of citations, average citations, market value of patents and average value.* The covariates are size, tangibility, profitability, tobin's q and age of the firm. "Years relative to Election" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after holding the election, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 3.23: Event Study Plots for Subset of old age (above median) firms which hold elections (part 2)



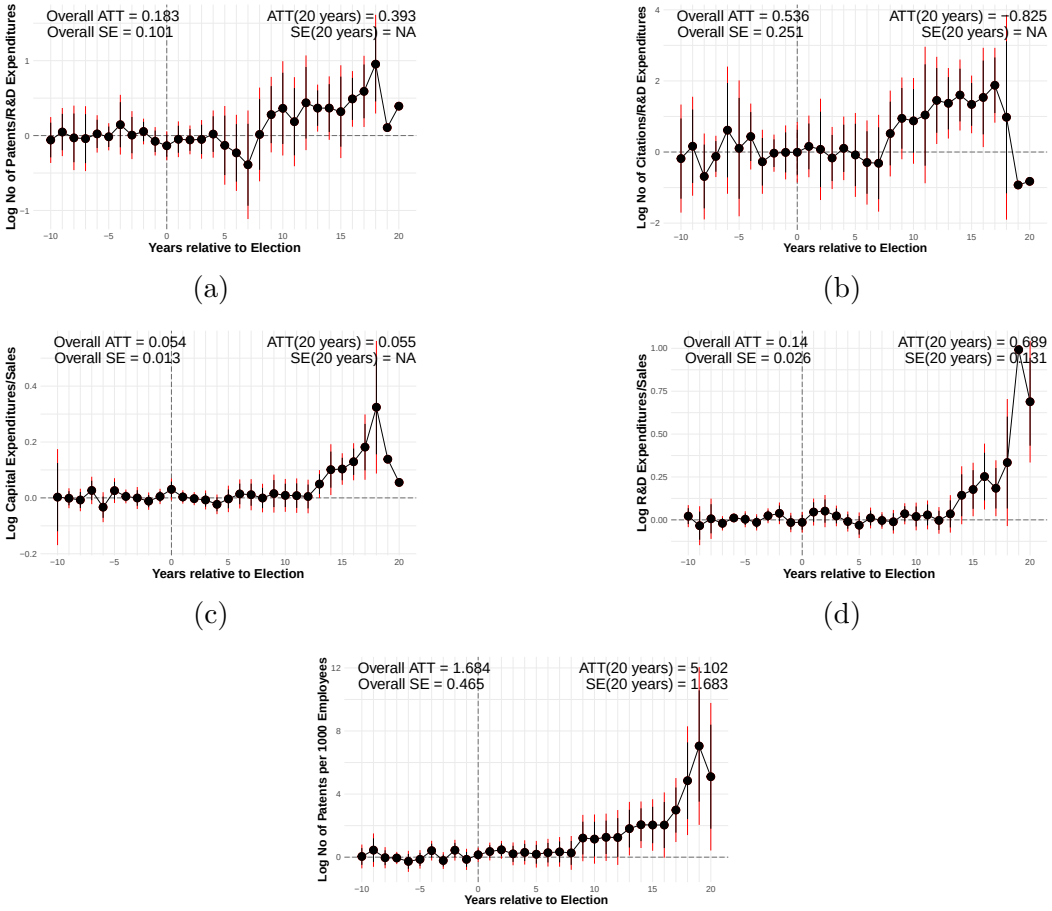
Notes: The figures present dynamic difference-in-difference estimates for the subset of old age (above median) firms which hold election, under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of *number of patents to R&D expenditures ratio, number of citations to R&D expenditures ratio, capital expenditures to sales ratio, R&D expenditures to sales ratio and number of patents per 1000 employees*. The covariates are size, tangibility, profitability, tobin's q and age of the firm. "Years relative to Election" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after holding the election, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 3.24: Event Study Plots for Subset of young age (below median) firms which hold elections (part 1)



Notes: The figures present dynamic difference-in-difference estimates for the subset of young age (below median) firms which hold election, under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of *number of patents, number of citations, average citations, market value of patents and average value.* The covariates are size, tangibility, profitability, tobin's q and age of the firm. "Years relative to Election" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after holding the election, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure 3.25: Event Study Plots for Subset of young age (below median) firms which hold elections (part 2)



Notes: The figures present dynamic difference-in-difference estimates for the subset of young age (below median) firms which hold election, under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of *number of patents to R&D expenditures ratio, number of citations to R&D expenditures ratio, capital expenditures to sales ratio, R&D expenditures to sales ratio and number of patents per 1000 employees*. The covariates are size, tangibility, profitability, tobin's q and age of the firm. "Years relative to Election" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after holding the election, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

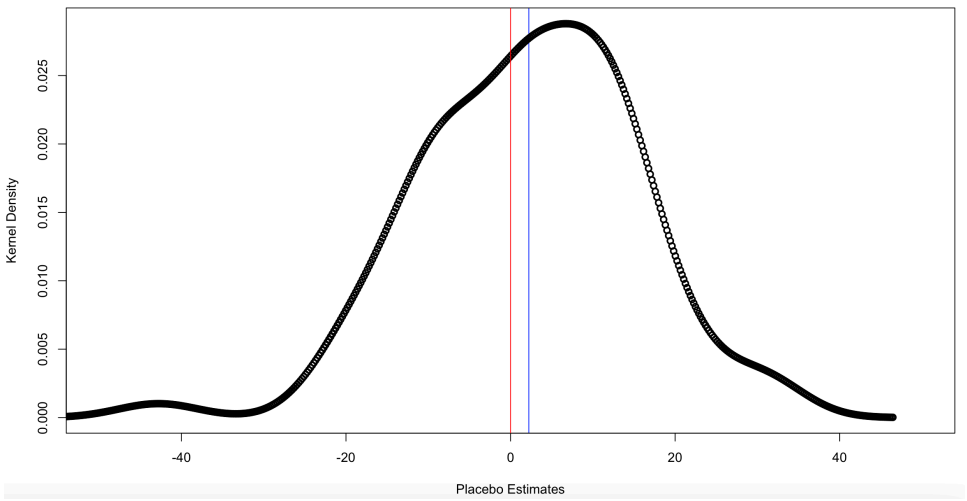
Appendices

Appendix A

Chapter 1 Appendix

A.1 Placebo Test

Figure A.1: Distribution of Placebo Treatment Effects (Effect after 4-6 years) on Vantage Score



Notes: The figure plots the density distribution of the estimated aggregated dynamic treatment effects after 4-6 years from loan receipt, obtained from a randomization placebo test. We randomly select a placebo treatment group (from g=2 to g=5) without considering the true time of the loan receipt, and estimate the CS dynamic treatment effects. This estimation process is repeated 80 times, and the graph shows the distribution of placebo treatment effects on vantage score. The red vertical line is drawn at the placebo estimate value of zero and the blue vertical line corresponds to the value that indicates the mean of all aggregated dynamic placebo treatment effects(=2.2).

A.2 Sensitivity Analysis

A.2.1 Sensitivity of the 2-4 year post-treatment loan effects

Figure A.2: Sensitivity Analysis for Vantage Score (e=2)

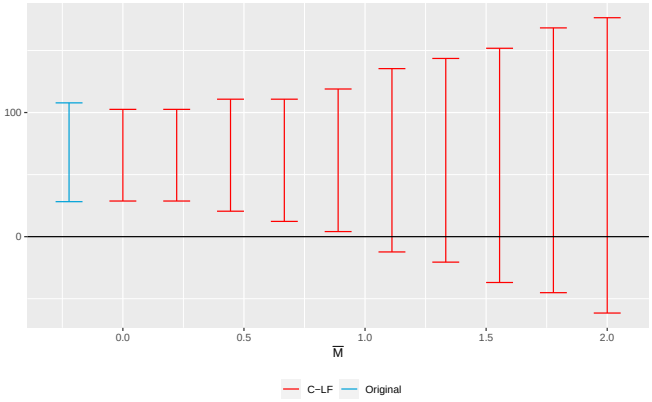
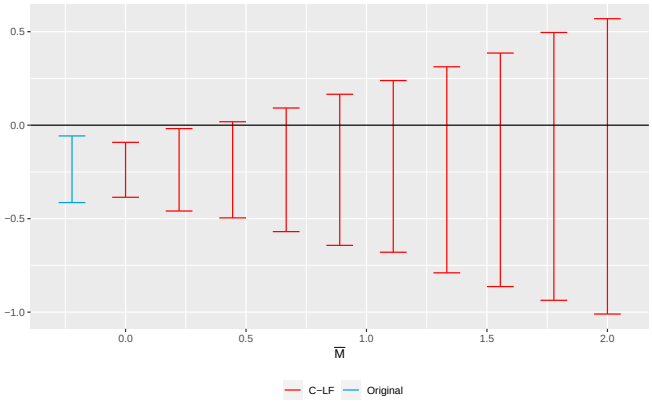


Figure A.3: Sensitivity Analysis for Utilization ratio(18) (e=2)



Notes: Figures A.2 and A.3, offer a formal sensitivity analysis that relates the magnitude of violations of parallel trends to the robustness of treatment estimates in post-treatment periods (Rambachan and Roth [78]). In Figure A.2, we report the sensitivity of the 2-4 year post-treatment loan effects on Vantage score to the different degree of violation of the *PTA*. The graph reports robust confidence sets under varying restrictions on the set of possible violations of parallel trends (different values of $\bar{M}$). When $\bar{M} = 0$, it implies that the parallel trend assumption holds exactly in post-treatment periods regardless of what happened in pre-treatment periods. The value of $\bar{M} = 1$ allows for violations of *PTA* in the post-treatment periods whose magnitude is as large as the largest violation of parallel trends on pre-treatment periods, and so on. In Figure A.3, we report the sensitivity of the 2-4 year post-treatment loan effects on Utilization Ratio to the different degree of violation of the *PTA*.

A.2.2 Sensitivity of the 4-6 year post-treatment loan effects

Figure A.4: Sensitivity Analysis for Vantage Score (e=4)

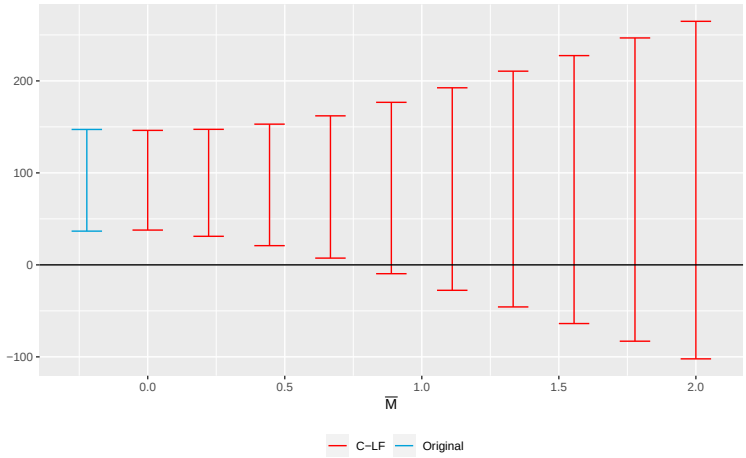
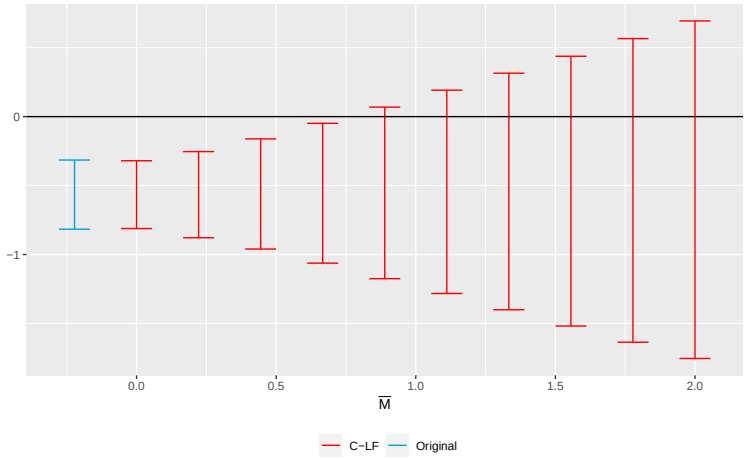


Figure A.5: Sensitivity Analysis for Utilization ratio(18) (e=4)



Notes: Figures A.4 and A.5, offer a formal sensitivity analysis that relates the magnitude of violations of parallel trends to the robustness of treatment estimates in post-treatment periods (Rambachan and Roth [78]). In Figure A.4, we report the sensitivity of the 4-6 year post-treatment loan effects on Vantage score to the different degree of violation of the *PTA*. The graph reports robust confidence sets under varying restrictions on the set of possible violations of parallel trends (different values of $\bar{M}$). When $\bar{M} = 0$, it implies that the parallel trend assumption holds exactly in post-treatment periods regardless of what happened in pre-treatment periods. The value of $\bar{M} = 1$ allows for violations of *PTA* in the post-treatment periods whose magnitude is as large as the largest violation of parallel trends on pre-treatment periods, and so on. In Figure A.5, we report the sensitivity of the 4-6 year post-treatment loan effects on Utilization Ratio to the different degree of violation of the *PTA*.

A.3 WWBIC Loan Application details

A.3.1 Existing Firms

WWBIC Loan Application Facts and Required Documents List (RDL)

EXISTING BUSINESS…

1) The first step in the loan application process is to complete a loan application: Please apply online at: https://wwbic.com/fund-your-business/loans/ Please pay the non-refundable $75 loan application fee online or if you prefer to pay via phone with credit or debit card please call 414-263-5450. Business Hours: Monday through Friday 8:30 AM-5:00 PM.
2) Forward Loan Officer (LO) a copy of the receipt for the application fee via email so that he can start pulling your credit report.
3) To process a loan request LO will need a complete Loan Packet which includes the following:

Personal Financial Items:

- ☐ (PFS) Personal Financial Statement (*unless you applied online*)*
- ☐ 3 most recent years of personal tax returns with W-2
- ☐ Most current paystub or other proof of personal incomes, if any
- ☐ Resume
- ☐ Most recent 3 months of personal bank statements
- ☐ (CIF) Client Information Form (unless you applied online)**
- ☐ Driver Licenses

Married?

- ☐ Both spouses complete the PFS together
- ☐ Paystub or other proof of income from spouse, if any

Business has more than one owner?

For each additional owner that has 20% ownership or more (or would like to guarantee the loan) we need a full set of personal financial items above

Business Financial Items

- ☐ Executive Summary of Business Plan*
- ☐ 3 years of projections (must include the first year's 12-month cash flow plus 2 additional years)*
- ☐ List of advisors including lawyer, accountant, banker, marketing, insurance agent with contact (phone and/or email); if do not have any please think about who you will use if you have a chance.
- ☐ Copy of signed lease if applicable
- ☐ 3 most recent years of business tax returns (if not on the schedule C of form 1040; we need either form 1065, 1120s, or 1120)
- ☐ Year to date profit and loss statement
- ☐ Most current balance sheet
- ☐ Business debt schedule
- ☐ Most recent 3 months of business bank statements
- ☐ (BIF) Business Information Form**
- ☐ Collateral and Inventory Form**
- ☐ Articles of Incorporation
- ☐ Operating Agreement or Bylaws
- ☐ EIN
- ☐ (EIFs) Employee Information Form A & Employee Information Form B**
- ☐ Proof of owner injection

Purchasing an existing Business?

- ☐ Seller's current year's financial statements
- ☐ Business Tax returns for the past 3 years.
- ☐ Purchase Agreement

Please note:

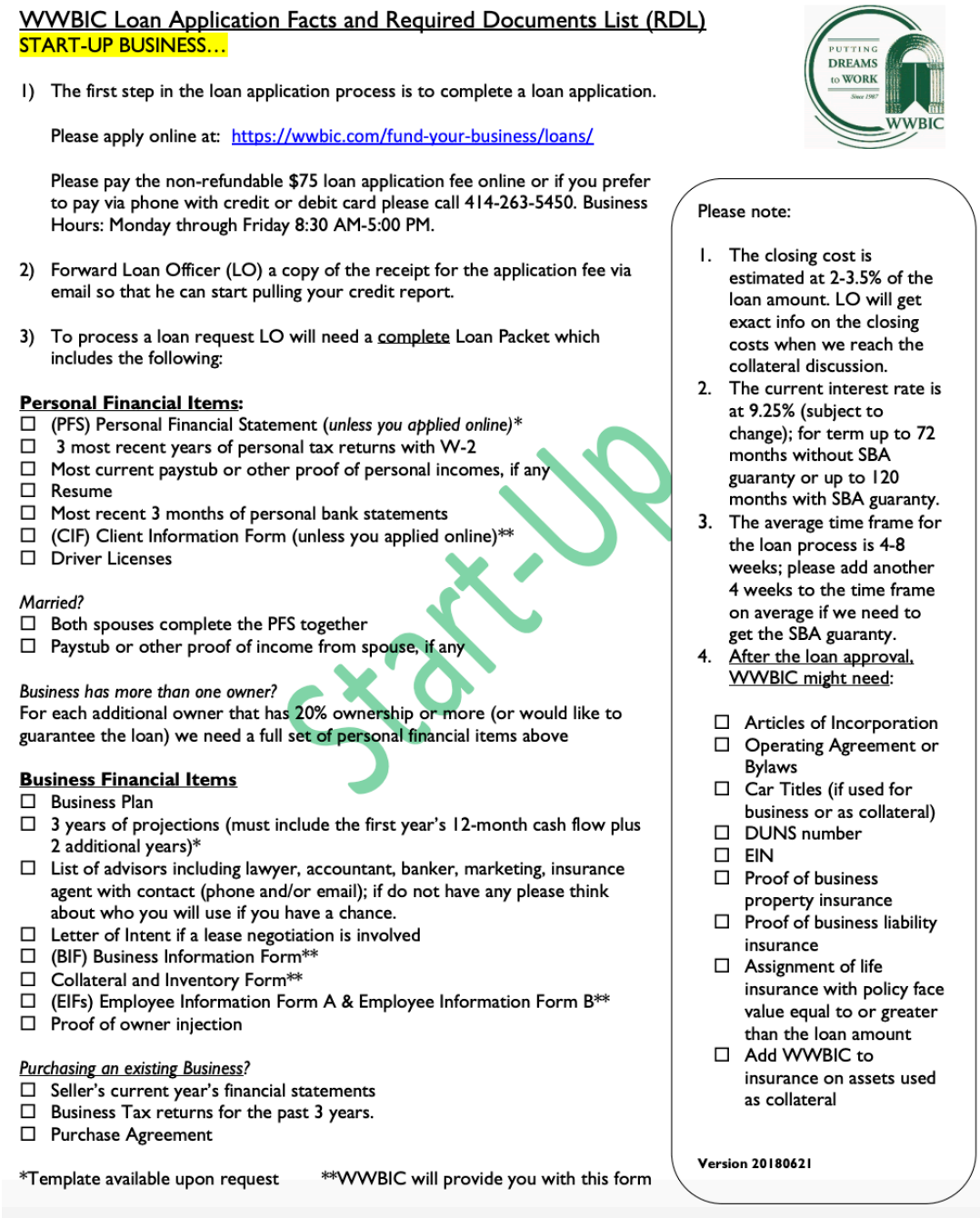
1. The closing cost is estimated at 2-3.5% of the loan amount. LO will get exact info on the closing costs when we reach the collateral discussion.
2. The current interest rate is at 9.25% (subject to change); for term up to 72 months without SBA guaranty or up to 120 months with SBA guaranty.
3. The average time frame for the loan process is 4-8 weeks; please add another 4 weeks to the time frame on average if we need to get the SBA guaranty.
4. After the loan approval, WWBIC might need:

   - ☐ Car Titles (if used for business or as collateral)
   - ☐ DUNS number
   - ☐ Proof of business property insurance
   - ☐ Proof of business liability insurance
   - ☐ Assignment of life insurance with policy face value equal to or greater than the loan amount
   - ☐ Add WWBIC to insurance on assets used as collateral

*Template available upon request **WWBIC will provide you with this form

Version 20180621

A.3.2 Start-up Firms



WWBIC Loan Application Facts and Required Documents List (RDL)
START-UP BUSINESS…

1) The first step in the loan application process is to complete a loan application.

   Please apply online at: https://wwbic.com/fund-your-business/loans/

   Please pay the non-refundable $75 loan application fee online or if you prefer to pay via phone with credit or debit card please call 414-263-5450. Business Hours: Monday through Friday 8:30 AM-5:00 PM.

2) Forward Loan Officer (LO) a copy of the receipt for the application fee via email so that he can start pulling your credit report.

3) To process a loan request LO will need a complete Loan Packet which includes the following:

Personal Financial Items:
- ☐ (PFS) Personal Financial Statement (*unless you applied online*)*
- ☐ 3 most recent years of personal tax returns with W-2
- ☐ Most current paystub or other proof of personal incomes, if any
- ☐ Resume
- ☐ Most recent 3 months of personal bank statements
- ☐ (CIF) Client Information Form (unless you applied online)**
- ☐ Driver Licenses

Married?
- ☐ Both spouses complete the PFS together
- ☐ Paystub or other proof of income from spouse, if any

Business has more than one owner?
For each additional owner that has 20% ownership or more (or would like to guarantee the loan) we need a full set of personal financial items above

Business Financial Items
- ☐ Business Plan
- ☐ 3 years of projections (must include the first year's 12-month cash flow plus 2 additional years)*
- ☐ List of advisors including lawyer, accountant, banker, marketing, insurance agent with contact (phone and/or email); if do not have any please think about who you will use if you have a chance.
- ☐ Letter of Intent if a lease negotiation is involved
- ☐ (BIF) Business Information Form**
- ☐ Collateral and Inventory Form**
- ☐ (EIFs) Employee Information Form A & Employee Information Form B**
- ☐ Proof of owner injection

Purchasing an existing Business?
- ☐ Seller's current year's financial statements
- ☐ Business Tax returns for the past 3 years.
- ☐ Purchase Agreement

*Template available upon request **WWBIC will provide you with this form

Please note:

1. The closing cost is estimated at 2-3.5% of the loan amount. LO will get exact info on the closing costs when we reach the collateral discussion.
2. The current interest rate is at 9.25% (subject to change); for term up to 72 months without SBA guaranty or up to 120 months with SBA guaranty.
3. The average time frame for the loan process is 4-8 weeks; please add another 4 weeks to the time frame on average if we need to get the SBA guaranty.
4. After the loan approval, WWBIC might need:

   - ☐ Articles of Incorporation
   - ☐ Operating Agreement or Bylaws
   - ☐ Car Titles (if used for business or as collateral)
   - ☐ DUNS number
   - ☐ EIN
   - ☐ Proof of business property insurance
   - ☐ Proof of business liability insurance
   - ☐ Assignment of life insurance with policy face value equal to or greater than the loan amount
   - ☐ Add WWBIC to insurance on assets used as collateral

Version 20180621

Start-Up

A.4 Important components of Credit Score

How to Build and Establish Credit:

"Your FICO credit score, the one most commonly used by lenders, is made up of five factors: (Source)

- **Payment history (35% of your score):** Your payment history is the most important factor in your credit score. Lenders want to know their borrowers will be reliable in paying back their debts, and a long history of on-time payments lets them know that.
- **Credit utilization (30% of your score):** This measures the balances on your revolving debt accounts (credit cards or lines of credit) compared with your borrowing limits. Creditors like to see this number at 30% or lower, but the lower, the better for your scores.
- **Length of credit history (15% of your score):** The length of time you've been managing credit accounts also plays a role in your scores. This factor includes the age of your newest and oldest accounts and the average age for all of them. A lengthy credit history is good for your credit health.
- **Credit mix (10% of your score):** Creditors like to see a mix of revolving (credit cards) and installment accounts (car loans, student loans, mortgages). A blended portfolio of credit accounts may help boost your score.
- **New credit (10% of your score):** This includes the number of recently opened accounts and hard inquiries generated each time you apply for credit. Many recent hard inquiries and new accounts can affect your credit approval. "

Appendix B

Chapter 2 Appendix

B.1 Data and Matching Details Appendix

B.1.1 NLRB Election Case Numbers

An NLRB Case ID Number is represented by the ID variable in the election data. This case number is assigned after the first election petition is filed. A single case number, on the other hand, may incorporate multiple different vote counts. For example, there could be (1) several tallies of the same election or (2) multiple elections with the same case number.[1] Furthermore, separate lections for multiple bargaining units may be filed under the same case number (e.g., if a union initially filed a petition for one bargaining unit but the NLRB later split bargaining unit). As a result, it is critical to select the vote count that corresponds to the outcome of the certification election.

In case where we have several observations of the same case number, there is information showing why there are many observations for a same case number in the public NLRB data (the "Public Data"). As a result, we select the final tally of the last election for each case number for a certain bargaining unit. This assures that in situations where there are

[1]There could be multiple tallies for the same election due to challenged votes. The first tally would not include challenged votes, while the final tally would include challenged votes that were determined to be valid. Additionally, there could be multiple elections for the same case number if an NLRB director orders a second election due to objections to the first election.

multiple counts of the same election or multiple ordered elections for the same bargaining unit, we pick the vote tally that determines the unions' certification. Within each case number, we then take the results from the election at the largest bargaining unit in cases where there are distinct bargaining units for a single case.

B.1.2 Accessing Compustat Data

Wharton Research Data Services (WRDS) is online platform that provides access to analytics and historical financial and accounting data for corporations and banks, historical economic data and more. WRDS incorporates multiple datasets. We use WRDS[2] to access the Compustat data. After creating the account and entering the login information, from the get data tab, we choose the "Compustat- Capital IQ" option, then we choose "North America- Daily" and primarily we are interested in " Fundamentals- Annual", we wil go with this option. This way, we can have access to the Compustat database of interest. We can choose our desired data range and "search the entire database" option to recover all the available data for firms. In the next step, we should select the variables that we need information on. Finally, we should tell the system the format of the output (we choose Stata file), then we submit the query to receive the data in our desired format via email. For more details, you can refer to this link: https://libguides.stanford.edu/library/wrds

B.1.3 Standardizing and Cleaning the Name of the Firms

The data on union elections do not have unique identifiers to link establishments to Compustat firms. Instead, we relied on another matching criteria to link each election to its corresponding firm in Compustat.

[2] https://wrds-www.wharton.upenn.edu/

Name and Address String Cleaning: We start by standardizing and cleaning the name strings. Our cleaning procedure builds on the `stnd_compname` Stata name standardization programs [93].

We clean the firm names as follows:

- Remove most symbols, non-numeric or letter characters, and non-standard ASCII characters.
- Remove the portion of company names in parentheses. The union election data often contain supplemental information in the parentheses portion of the name (e.g., (`wage employees only`)).
- Remove the portion of company names following `DOING BUSINESS AS (DBA)` or `A DIVISION OF`
- Combine consecutive singleton letters and symbols separated by spaces (e.g., `A T & T` $\Rightarrow$ `AT&T` and `D R HORTON` $\Rightarrow$ `DR HORTON`).
- Remove company entity types (e.g., `CORP`, `INC`, etc.), articles, and standard common compnay names (e.g., `MANUFACTURERS` $\Rightarrow$ `MANUFACTURING`).

B.1.4 Matching Criteria

In order to establish accurate matches between the union and Compustat datasets, we took a different approach by creating a new matching criteria instead of relying on potentially flawed fuzzy matching algorithms. Our aim was to ensure reliable links between the datasets while avoiding incorrect matches. To begin, we standardized the names of the firms in both datasets as explained in section B.1.3. This process helped eliminate inconsistencies and

variations that could hinder the matching process. With standardized names, we proceeded to devise a new matching criteria based on the number of tokens/words present in the company names.

When a company name consisted of a single token, we used that token as the basis for establishing the link between the two datasets. For names with two tokens, we utilized both tokens as the matching key. In cases where the number of tokens exceeded two, we selected the first three tokens of the company names as the matching criteria. To ensure the effectiveness of this approach, we also manually checked the firm names. This step allowed us to verify that our chosen matching criteria adequately served our purpose. By carefully examining the names, we could make informed decisions about which tokens to include in the matching process.

Additionally, we encountered situations where the second or third token of a company name included the symbol "&". To handle these cases, we expanded the matching criteria to include the first four tokens of each firm name. This adjustment ensured accurate matches by considering a broader portion of the name when the "&" symbol was present. The combination of these restrictions and considerations resulted in a new matching identifier that we utilized to link the union and Compustat datasets. While this approach did yield a relatively smaller sample, it significantly reduced the risk of incorrect matches that can arise from using fuzzy matching algorithms solely based on string similarities.

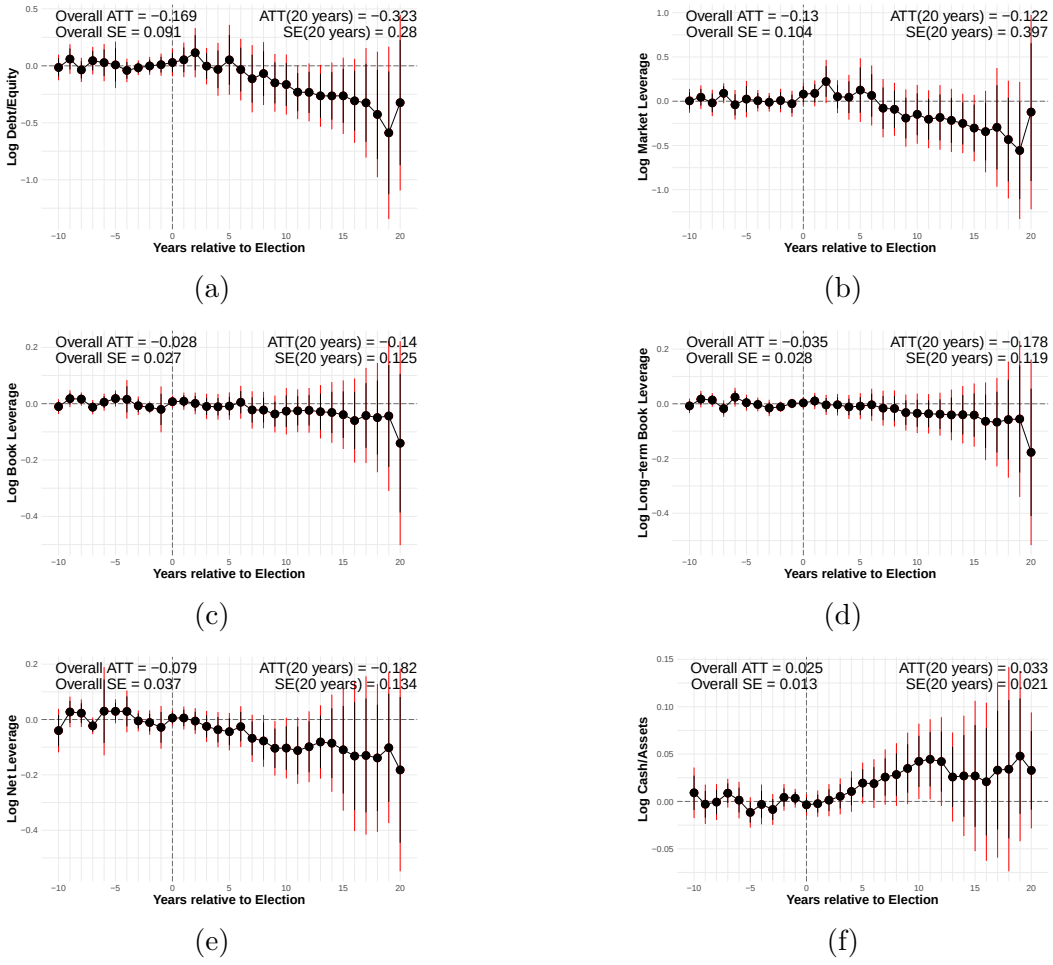
The decision to move away from fuzzy matching algorithms and focus on token-based matching criteria has several advantages. By considering the composition and structure of company names, we gain valuable insights beyond simple string similarity. This approach leverages the meaningful information contained within the tokens themselves, improving the accuracy of the matches. Furthermore, the manual inspection of firm names adds an additional layer of verification. It allows us to account for exceptions that may not be captured by the es-

tablished matching criteria. Human oversight ensures that the matching process is tailored to the specific context of the datasets, enhancing the reliability of the results. While it is true that this new matching approach yields a smaller sample size compared to using fuzzy matching algorithms, the focus here is on the quality and accuracy of the matches rather than the quantity of data. By prioritizing accurate matches, we can have greater confidence in the integrity and reliability of the resulting dataset.

In conclusion, our adoption of a new matching criteria based on token counts, manual verification, and adjustments for exceptional cases provides a robust approach to link the union and Compustat datasets accurately. By incorporating specific rules and addressing variations within the company names, we mitigate the risk of incorrect matches associated with fuzzy matching algorithms. Although the sample size may be reduced, the trade-off is a dataset that is more reliable, ensuring the integrity of subsequent analysis and research endeavors.

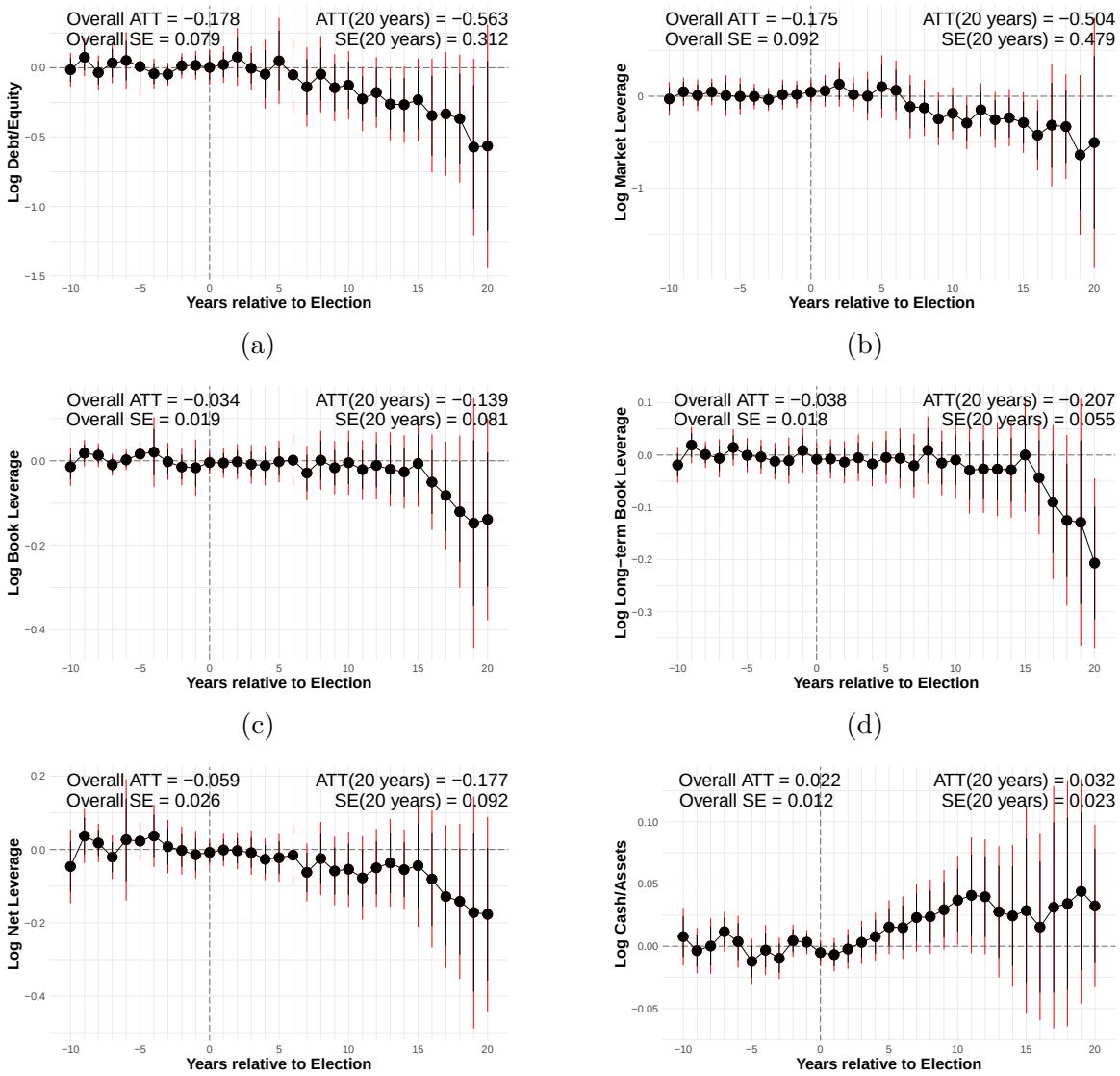
B.2 Supplementary Event Study Results

Figure B.1: Event Study Plots for Subset of Firms Holding Elections for unions- Unconditional Parallel Trends Assumption



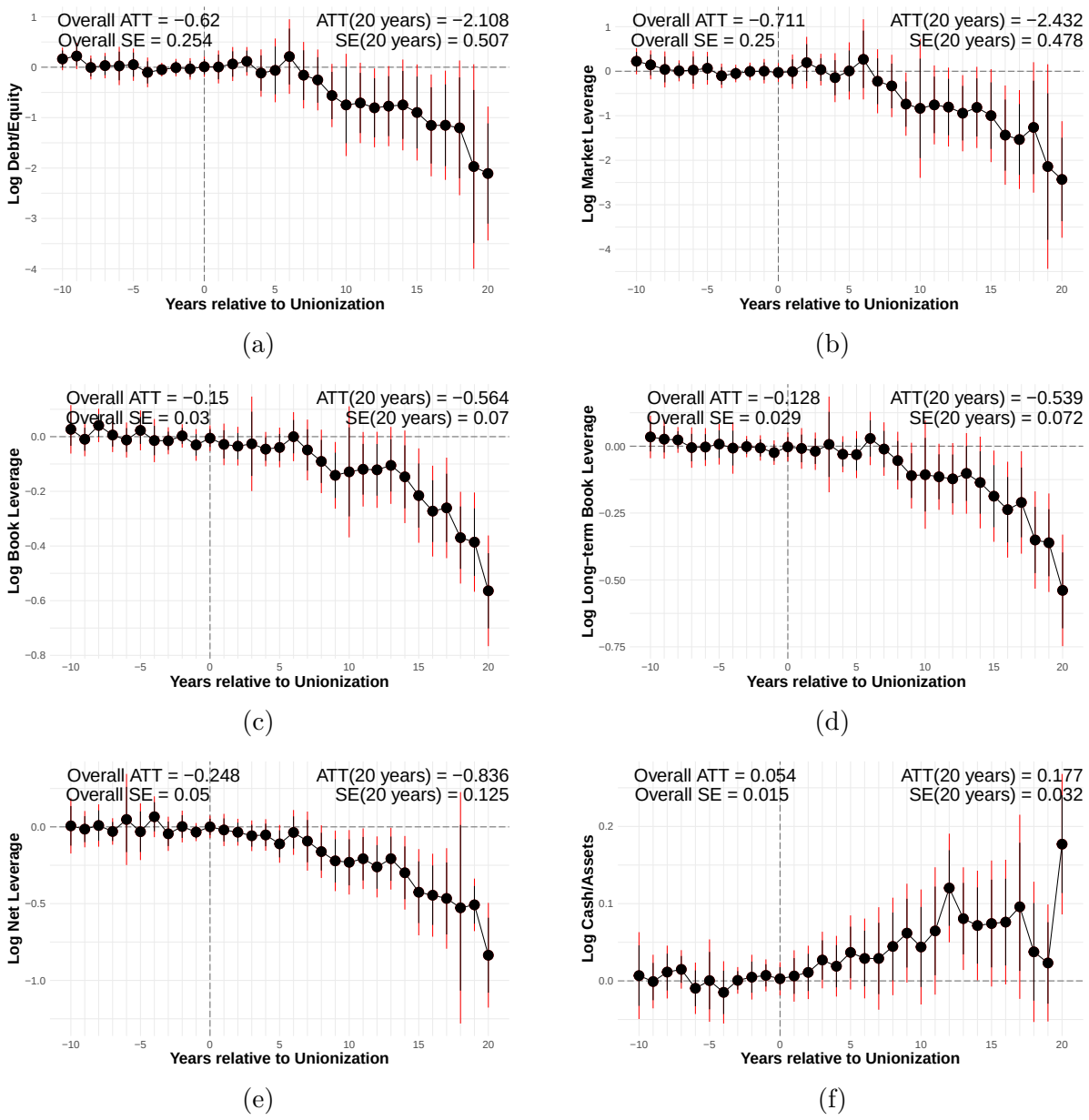
Notes: The figures present dynamic difference-in-difference estimates for the subset of firms holding elections for unionization, under unconditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of D/E, market leverage, book leverage, long-term book leverage, net leverage, and cash/asset ratio. "Years relative to Election" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after holding the election, and so on. For each e, the *ATT* is calculated by averaging dynamic treatment effects across all groups. The overall *ATT* is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure B.2: Event Study Plots for Subset of Firms Holding Elections for unions- Conditional Parallel Trends Assumption



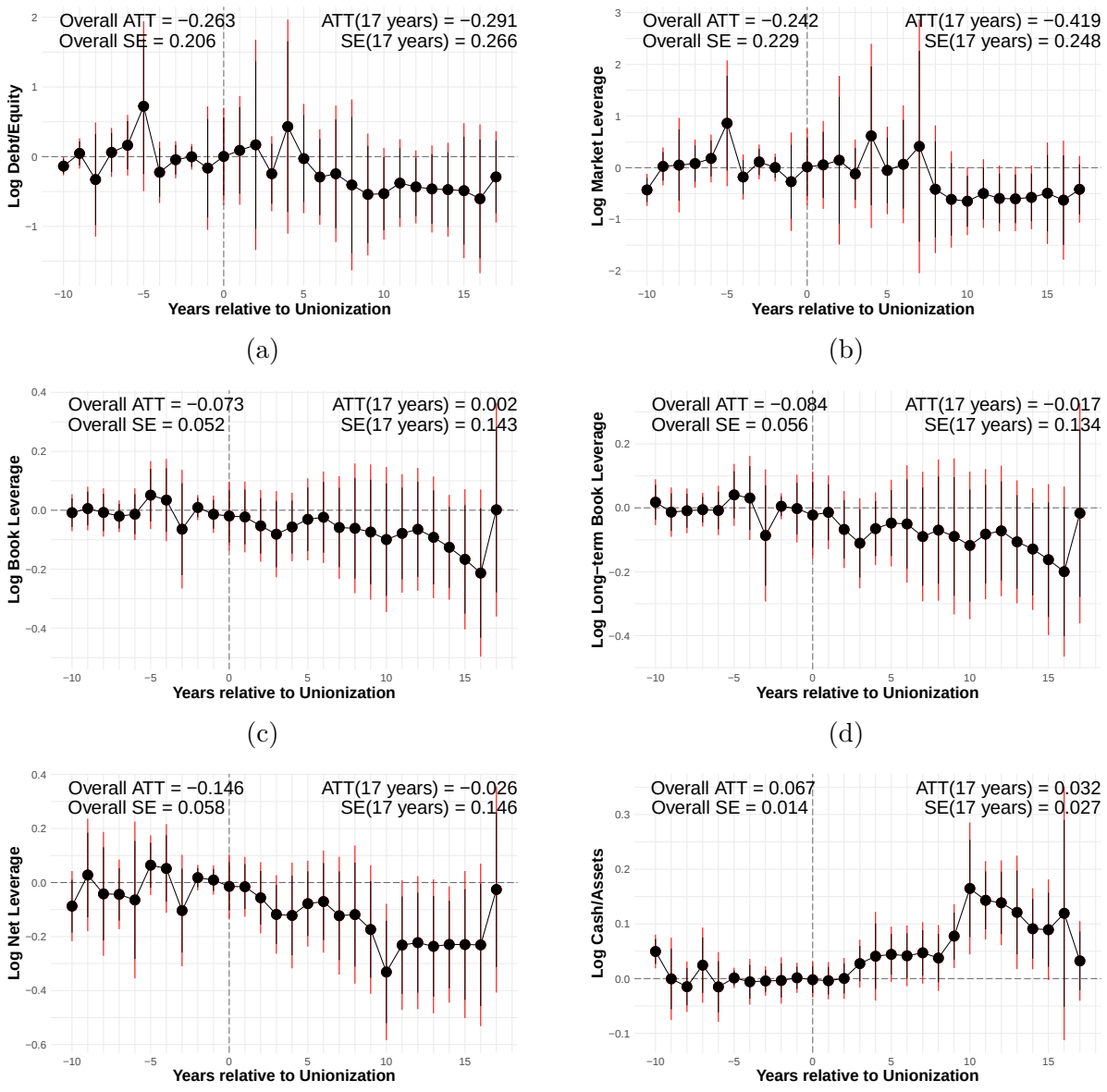
Notes: The figures present dynamic difference-in-difference estimates for the subset of firms holding elections for unionization, under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of D/E, market leverage, book leverage, long-term book leverage, net leverage, and cash/asset ratio. The covariates are size, tangibility, profitability, and tobin's q. "Years relative to Election" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after unionization, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure B.3: Event Study Plots for Subset of Unionized Firms with High Support



Notes: The figures present dynamic difference-in-difference estimates for the subset of unionized firms with "high" support, under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of D/E, market leverage, book leverage, long-term book leverage, net leverage, and cash/asset ratio. The covariates are size, tangibility, profitability, and tobin's q. "Years relative to Unionization" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after unionization, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Figure B.4: Event Study Plots for Subset of Unionized Firms with Low Support



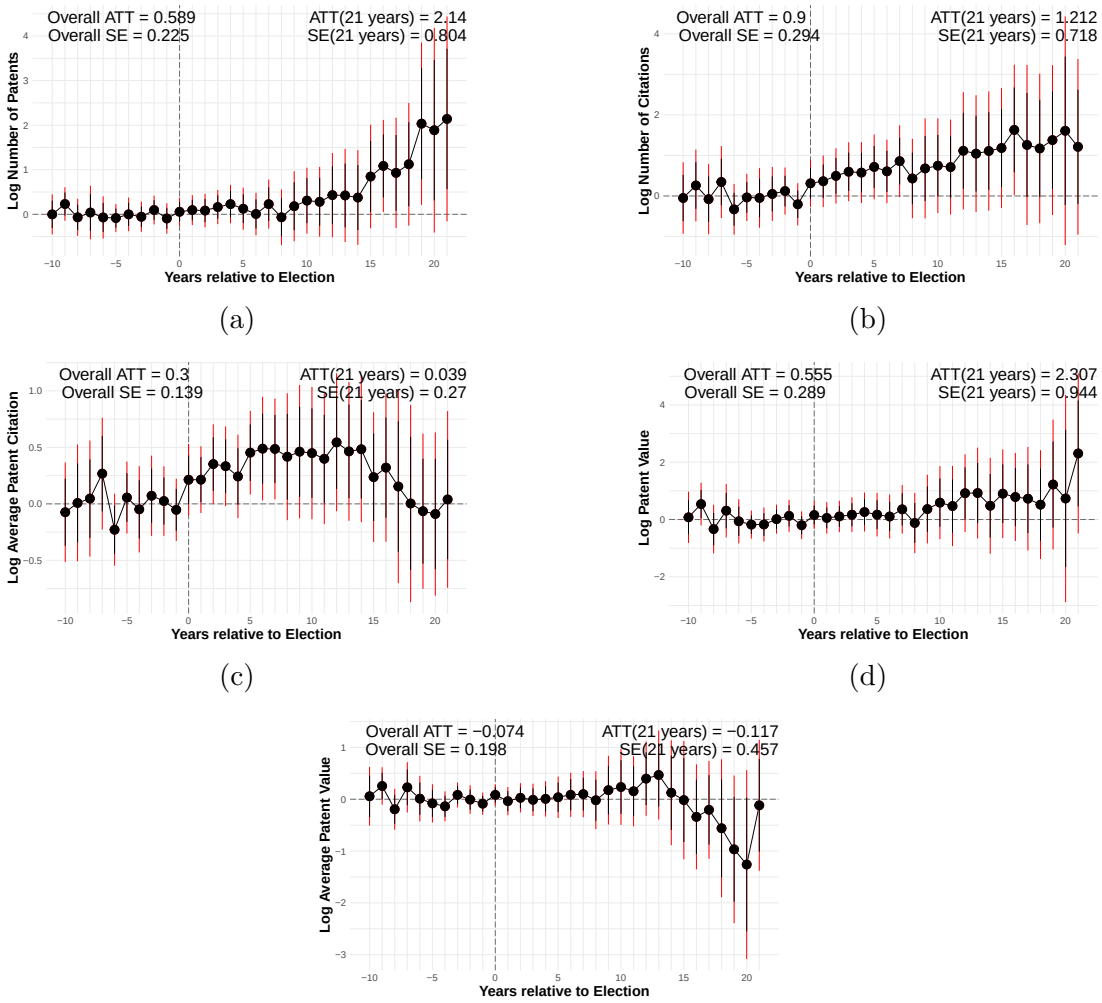
Notes: The figures present dynamic difference-in-difference estimates for the subset of unionized firms with "low" support, under conditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of D/E, market leverage, book leverage, long-term book leverage, net leverage, and cash/asset ratio. The covariates are size, tangibility, profitability, and tobin's q. "Years relative to Unionization" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after unionization, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Appendix C

Chapter 3 Appendix

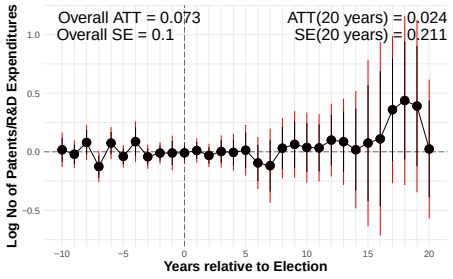
C.1 Supplementary Event Study Results

Figure C.1: Event Study Plots for Innovation Outcome Variables- Subset of Firms Holding Elections for unions- Unconditional Parallel Trends Assumption (part 1)

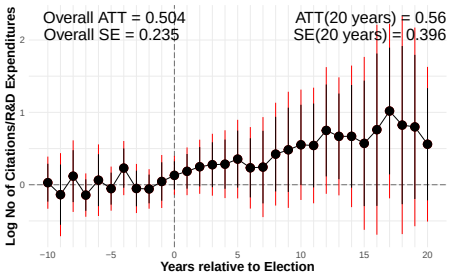


Notes: The figures present dynamic difference-in-difference estimates for the subset of firms holding elections for unionization, under unconditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of *number of patents, number of citations, average citations, market value of patents and average value.* "Years relative to Election" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after holding the election, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

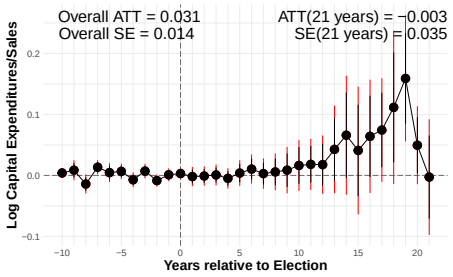
Figure C.2: Event Study Plots for Innovation Outcome Variables- Subset of Firms Holding Elections for unions- Unconditional Parallel Trends Assumption (part 2)



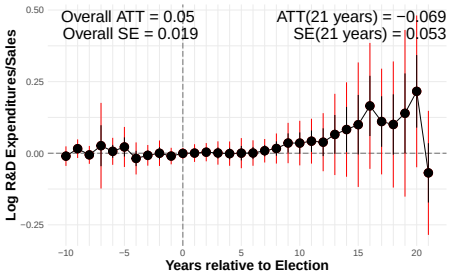
(a)



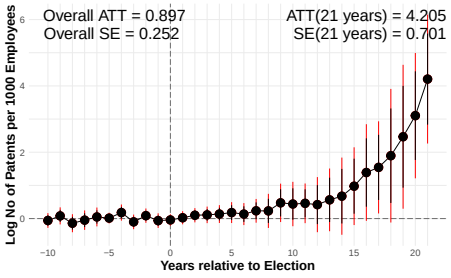
(b)



(c)



(d)



(e)

Notes: The figures present dynamic difference-in-difference estimates for the subset of firms holding elections for unionization, under unconditional parallel trends. The red lines represent simultaneous 95% confidence bands, while the black lines represent pointwise confidence bands. Standard errors are clustered at the firm level. The outcome variables include the logarithm of *number of patents to R&D expenditures ratio, number of citations to R&D expenditures ratio, capital expenditures to sales ratio, R&D expenditures to sales ratio and number of patents per 1000 employees.* "Years relative to Election" indicate the exposure to the treatment (e); for instance, estimates corresponding to $e = 0$ represent the weighted average of group-time average treatment effects *(ATTs)* in the same year of holding the election. Similarly, $e = 2$ represents the weighted average of group-time *ATTs* two years after holding the election, and so on. For each e, the ATT is calculated by averaging dynamic treatment effects across all groups. The overall ATT is obtained by averaging dynamic treatment effects across all event times. The average treatment effects are not statistically significant when the confidence bands include 0.

Bibliography

[1] Abadie, Alberto (2005), "Semiparametric difference-in-differences estimators." *The Review of Economic Studies*, 72, 1–19.

[2] Acharya, Viral V, Ramin P Baghai, and Krishnamurthy V Subramanian (2013), "Labor laws and innovation." *The Journal of Law and Economics*, 56, 997–1037.

[3] Acharya, Viral V, Ramin P Baghai, and Krishnamurthy V Subramanian (2014), "Wrongful discharge laws and innovation." *The Review of Financial Studies*, 27, 301–346.

[4] Addison, John T, Paulino Teixeira, Katalin Evers, and Lutz Bellmann (2017), "Collective bargaining and innovation in germany: a case of cooperative industrial relations?" *Industrial Relations: A Journal of Economy and Society*, 56, 73–121.

[5] Aidt, Toke and Zafiris Tzannatos (2002), *Unions and collective bargaining: Economic effects in a global environment*. Washington, DC: World Bank.

[6] Alimov, Azizjon (2015), "Labor protection laws and bank loan contracting." *The Journal of Law and Economics*, 58, 37–74.

[7] Antonioli, Davide, Massimiliano Mazzanti, and Paolo Pini (2011), "Innovation, industrial relations and employee outcomes: evidence from italy." *Journal of Economic Studies*, 38, 66–90.

[8] Athey, Susan and Guido W Imbens (2018), "Design-based analysis in difference-in-differences settings with staggered adoption." Technical report, National Bureau of Economic Research.

[9] Autor, David H, John J Donohue III, and Stewart J Schwab (2006), "The costs of wrongful-discharge laws." *The review of economics and statistics*, 88, 211–231.

[10] Banerjee, Abhijit, Emily Breza, Esther Duflo, and Cynthia Kinnan (2019), "Can microfinance unlock a poverty trap for some entrepreneurs?" Technical report, National Bureau of Economic Research.

[11] Banerjee, Abhijit, Esther Duflo, Rachel Glennerster, and Cynthia Kinnan (2015), "The miracle of microfinance? evidence from a randomized evaluation." *American economic journal: Applied economics*, 7, 22–53.

[12] Banerjee, Abhijit, Dean Karlan, and Jonathan Zinman (2015), "Six randomized evaluations of microcredit: Introduction and further steps." *American Economic Journal: Applied Economics*, 7, 1–21.

[13] Banerjee, Abhijit V, Emily Breza, Esther Duflo, and Cynthia Kinnan (2017), "Do credit constraints limit entrepreneurship? heterogeneity in the returns to microfinance." *Heterogeneity in the Returns to Microfinance (September 1, 2017). Global Poverty Research Lab Working Paper*.

[14] Banerjee, Abhijit Vinayak (2013), "Microcredit under the microscope: what have we learned in the past two decades, and what do we need to know?" *Annu. Rev. Econ.*, 5, 487–519.

[15] Beath, John, Yannis Katsoulacos, and David Ulph (1989), "Strategic r & d policy." *The Economic Journal*, 99, 74–83.

[16] Berglund, Tomas and Bengt Furåker (2016), "Employment protection regulation, trade unions and tenure of employment: An analysis in 23 european countries." *Industrial Relations Journal*, 47, 492–512.

[17] Bernanke, Ben (2007), "Microfinance in the united states." Technical report, Board of Governors of the Federal Reserve System (US).

[18] Berton, Fabio, Andrea Ricci, and Stefano Dughera (2021), "Are unions detrimental to innovation? theory and evidence."

[19] Besley, Timothy and Robin Burgess (2004), "Can labor regulation hinder economic performance? evidence from india." *The Quarterly journal of economics*, 119, 91–134.

[20] Booth, Alison L (1995), *The economics of the trade union*. Cambridge University Press.

[21] Borusyak, Kirill and Xavier Jaravel (2017), "Revisiting event study designs." *Available at SSRN 2826228*.

[22] Bradley, Daniel, Incheol Kim, and Xuan Tian (2017), "Do unions affect innovation?" *Management Science*, 63, 2251–2271.

[23] Bronars, Stephen G and Donald R Deere (1991), "The threat of unionization, the use of debt, and the preservation of shareholder wealth." *The Quarterly Journal of Economics*, 106, 231–254.

[24] Bronfenbrenner, Kate and Tom Juravich (1998), "It takes more than house calls: Organizing to win with a comprehensive union-building strategy."

[25] Bryson, Alex (2007), "The effect of trade unions on wages." *Reflets et perspectives de la vie économique*, 46, 033–045.

[26] Bryson, Alex, Erling Barth, and Harald Dale-Olsen (2013), "The effects of organizational change on worker well-being and the moderating role of trade unions." *ILR Review*, 66, 989–1011.

[27] Bryson, Alex and Harald Dale-Olsen (2020), "Unions, tripartite competition and innovation." Technical report, IZA Discussion Papers.

[28] Bryson, Alex, John Forth, and Simon Kirby (2005), "High-involvement management practices, trade union representation and workplace performance in britain." *Scottish journal of political economy*, 52, 451–491.

[29] Callaway, Brantly and Pedro HC Sant'Anna (2021), "Difference-in-differences with multiple time periods." *Journal of Econometrics*, 225, 200–230.

[30] Connolly, Robert A, Barry T Hirsch, and Mark Hirschey (1986), "Union rent seeking, intangible capital, and market value of the firm." *The Review of Economics and Statistics*, 567–577.

[31] Craig, Ben R, William E Jackson III, and James B Thomson (2007), "Small firm finance, credit rationing, and the impact of sba-guaranteed lending on local economic growth." *Journal of small business management*, 45, 116–132.

[32] Cull, Robert and Jonathan Morduch (2018), "Microfinance and economic development." In *Handbook of finance and development*, 550–572, Edward Elgar Publishing.

[33] Daniel, Bradley, Kim Incheol, and Tian Xuan (2015), "Do unions affect innovation." *Management Science, Forthcoming*.

[34] De Chaisemartin, Clément and Xavier d'Haultfoeuille (2020), "Two-way fixed effects estimators with heterogeneous treatment effects." *American Economic Review*, 110, 2964–96.

[35] Dinlersoz, Emin, Jeremy Greenwood, and Henry Hyatt (2017), "What businesses attract unions? unionization over the life cycle of us establishments." *ILR Review*, 70, 733–766.

[36] Doucouliagos, Christos and Patrice Laroche (2003), "What do unions do to productivity? a meta-analysis." *Industrial Relations: A Journal of Economy and Society*, 42, 650–691.

[37] Doucouliagos, Hristos and Patrice Laroche (2013), "Unions and innovation: New insights from the cross-country evidence." *Industrial Relations: A Journal of Economy and Society*, 52, 467–491.

[38] Ederer, Florian and Gustavo Manso (2013), "Is pay for performance detrimental to innovation?" *Management Science*, 59, 1496–1513.

[39] Fang, Tony and Ying Ge (2012), "Unions and firm innovation in china: Synergy or strife?" *China Economic Review*, 23, 170–180.

[40] Farber, Henry S and Daniel H Saks (1980), "Why workers want unions: The role of relative wages and job characteristics." *Journal of Political Economy*, 88, 349–369.

[41] Feigenbaum, James, Alexander Hertel-Fernandez, and Vanessa Williamson (2018), "From the bargaining table to the ballot box: Political effects of right to work laws." Technical report, National Bureau of Economic Research.

[42] Frandsen, Brigham R (2017), "Party bias in union representation elections: Testing for manipulation in the regression discontinuity design when the running variable is discrete." In *Regression discontinuity designs*, volume 38, 281–315, Emerald Publishing Limited.

[43] Frank, Murray Z and Vidhan K Goyal (2009), "Capital structure decisions: which factors are reliably important?" *Financial management*, 38, 1–37.

[44] Freeman, Richard B and James L Medoff (1984), "What do unions do." *Indus. & Lab. Rel. Rev.*, 38, 244.

[45] Gardner, John (2022), "Two-stage differences in differences." *arXiv preprint arXiv:2207.05943*.

[46] Goodman-Bacon, Andrew (2018), "Difference-in-differences with variation in treatment timing." Technical report, National Bureau of Economic Research.

[47] Graham, John R and Campbell R Harvey (2001), "The theory and practice of corporate finance: Evidence from the field." *Journal of financial economics*, 60, 187–243.

[48] Grout, Paul A (1984), "Investment and wages in the absence of binding contracts: A nash bargaining approach." *Econometrica: Journal of the Econometric Society*, 449–460.

[49] Hall, Bronwyn H, Adam B Jaffe, and Manuel Trajtenberg (2001), "The nber patent citation data file: Lessons, insights and methodological tools."

[50] Haucap, Justus and Christian Wey (2004), "Unionisation structures and innovation incentives." *The Economic Journal*, 114, C149–C165.

[51] Heckman, James J, Hidehiko Ichimura, and Petra E Todd (1997), "Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme." *The review of economic studies*, 64, 605–654.

[52] Heider, Florian and Alexander Ljungqvist (2015), "As certain as debt and taxes: Estimating the tax sensitivity of leverage from state tax changes." *Journal of Financial Economics*, 118, 684–712.

[53] Hirsch, Barry T (2008), "Sluggish institutions in a dynamic world: Can unions and industrial competition coexist?" *Journal of Economic Perspectives*, 22, 153–176.

[54] Holman, David and Anthony Rafferty (2018), "The convergence and divergence of job discretion between occupations and institutional regimes in europe from 1995 to 2010." *Journal of Management Studies*, 55, 619–647.

[55] Holmes, Thomas J (1998), "The effect of state policies on the location of manufacturing: Evidence from state borders." *Journal of political Economy*, 106, 667–705.

[56] Imai, Kosuke and In Song Kim (2020), "On the use of two-way fixed effects regression models for causal inference with panel data." *Political Analysis*, 1–11.

[57] Islam, Emdad and Jason Zein (2020), "Inventor ceos." *Journal of Financial Economics*, 135, 505–527.

[58] J OSWALD', ANDREW and PETER J TURNBULL (1985), "Pay and employment determination in britain: What are labour 'contracts'really like?" *Oxford Review of Economic Policy*, 1, 80–97.

[59] Klasa, Sandy, William F Maxwell, and Hernán Ortiz-Molina (2009), "The strategic use of corporate cash holdings in collective bargaining with labor unions." *Journal of financial economics*, 92, 421–442.

[60] Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman (2017), "Technological innovation, resource allocation, and growth." *The Quarterly Journal of Economics*, 132, 665–712.

[61] Kuzmina, Olga et al. (2013), "Operating flexibility and capital structure: Evidence from a natural experiment." *Columbia Business School Research Paper*, 13–69.

[62] Lavetti, Kurt, Carol Simon, and William D White (2020), "The impacts of restricting mobility of skilled service workers evidence from physicians." *Journal of Human Resources*, 55, 1025–1067.

[63] Layard, Richard, P Richard G Layard, and Richard Jackman (2005), *Unemployment: macroeconomic performance and the labour market*. Oxford University Press on Demand.

[64] Lee, David S and Alexandre Mas (2012), "Long-run impacts of unions on firms: New evidence from financial markets, 1961–1999." *The Quarterly Journal of Economics*, 127, 333–378.

[65] Lee, YS (2013), "Entrepreneurship, small business, and urban growth." *Unpublished Working Paper, Williams College*.

[66] Lemmon, Michael L, Michael R Roberts, and Jaime F Zender (2008), "Back to the beginning: persistence and the cross-section of corporate capital structure." *The journal of finance*, 63, 1575–1608.

[67] Levitt, Martin Jay, Terry Conrow, and Terry C Toczynski (1993), *Confessions of a union buster*. Crown.

[68] Logan, John (2002), "Consultants, lawyers, and the 'union free'movement in the usa since the 1970s." *Industrial relations journal*, 33, 197–214.

[69] Manso, Gustavo (2011), "Motivating innovation." *The journal of finance*, 66, 1823–1860.

[70] Matsa, David A (2010), "Capital structure as a strategic variable: Evidence from collective bargaining." *The Journal of Finance*, 65, 1197–1232.

[71] Menezes-Filho, Naercio and John Van Reenen (2003), "Unions and innovation: a survey of the theory and empirical evidence." *International handbook of trade unions*, 293–334.

[72] Millán, Ana, José María Millán, Concepción Román, and André van Stel (2013), "How does employment protection legislation influence hiring and firing decisions by the smallest firms?" *Economics Letters*, 121, 444–448.

[73] Mukherjee, Arijit and Enrico Pennings (2011), "Unionization structure, licensing and innovation." *International Journal of Industrial Organization*, 29, 232–241.

[74] Myers, Stewart C (1977), "Determinants of corporate borrowing." *Journal of financial economics*, 5, 147–175.

[75] Öztekin, Özde (2015), "Capital structure decisions around the world: which factors are reliably important?" *Journal of financial and quantitative analysis*, 50, 301–323.

[76] Porter, Michael E (1992), "Capital disadvantage: America's failing capital investment system." *Harvard business review*, 70, 65–82.

[77] Rajan, Raghuram G and Luigi Zingales (1995), "What do we know about capital structure? some evidence from international data." *The journal of Finance*, 50, 1421–1460.

[78] Rambachan, Ashesh and Jonathan Roth (2022), "A more credible approach to parallel trends." Technical report, Working Paper.

[79] Romer, Paul M (1986), "Increasing returns and long-run growth." *Journal of political economy*, 94, 1002–1037.

[80] Sandver, Marcus Hart and Kathryn J Ready (1998), "Trends in and determinants of outcomes in multi-union certification elections." *Journal of Labor Research*, 19, 165–172.

[81] Sant'Anna, Pedro HC and Jun Zhao (2020), "Doubly robust difference-in-differences estimators." *Journal of Econometrics*, 219, 101–122.

[82] Schmalz, Martin C (2018), "Unionization, cash, and leverage." *Available at SSRN 3106798*.

[83] Schmitt, John, Ben Zipperer, et al. (2009), "Dropping the ax: Illegal firings during union election campaigns, 1951-2007." Technical report, Center for Economic and Policy Research (CEPR).

[84] Schnabel, Claus and Joachim Wagner (1994), "Industrial relations and trade union effects on innovation in germany." *Labour*, 8, 489–504.

[85] Simintzi, Elena, Vikrant Vig, and Paolo Volpin (2015), "Labor protection and leverage." *The Review of Financial Studies*, 28, 561–591.

[86] Slichter, Sumner Huber, James J Healy, and Edward Robert Livernash (1960), *The impact of collective bargaining on management*, volume 4. Brookings Inst Press.

[87] Solow, Robert M (1957), "Technical change and the aggregate production function." *The review of Economics and Statistics*, 312–320.

[88] Sun, Liyang and Sarah Abraham (2021), "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects." *Journal of Econometrics*, 225, 175–199.

[89] Sunder, Jayanthi, Shyam V Sunder, and Jingjing Zhang (2017), "Pilot ceos and corporate innovation." *Journal of Financial Economics*, 123, 209–224.

[90] Traxler, Franz (1994), "Collective bargaining: Levels and coverage." *Employment Outlook*, 1994, 167–194.

[91] Tressel, Thierry and Stefano Scarpetta (2004), *Boosting productivity via innovation and adoption of new technologies: any role for labor market institutions?*, volume 3273. World Bank Publications.

[92] Ulph, Alistair M and David T Ulph (1994), "Labour markets and innovation: ex-post bargaining." *European Economic Review*, 38, 195–210.

[93] Wasi, Nada and Aaron Flaaen (2015), "Record linkage using stata: Preprocessing, linking, and reviewing utilities." *The Stata Journal*, 15, 672–697.

[94] Woods, Keegan, Kelvin Jui Keng Tan, and Robert Faff (2019), "Labor unions and corporate financial leverage: The bargaining device versus crowding-out hypotheses." *Journal of Financial Intermediation*, 37, 28–44.

[95] Young, Andrew T, Matthew J Higgins, Donald J Lacombe, and Briana Sell (2014), "The direct and indirect effects of small business administration lending on growth: Evidence from us county-level data." Technical report, National Bureau of Economic Research.

[96] Young, Samuel Goericke et al. (2022), *Essays on Labor Market Institutions*. Ph.D. thesis, Massachusetts Institute of Technology.