Title
Essays in Political Economy and Institutions

Author
Gonzalez, Felipe

Publication Date
2017

Essays in Political Economy and Institutions

by

Felipe Gonzalez

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy
in
Economics
in the
Graduate Division
of the
University of California, Berkeley

Committee in charge:

Professor Edward Miguel, Chair
Professor Frederico S. Finan
Professor Solomon M. Hsiang

Spring 2017

Essays in Political Economy and Institutions

Abstract
Essays in Political Economy and Institutions
by
Felipe Gonzalez
Doctor of Philosophy in Economics
University of California, Berkeley
Professor Edward Miguel, Chair

The first chapter of my dissertation studies the 2011 student movement in Chile, the largest protest mobilization in the country's history, in which hundreds of thousands of students skipped school to protest with the goal of reforming the educational system. Using administrative data on millions of students' daily school attendance decisions on protest and non-protest days, a large network composed by the lifetime history of classmates, and differential network exposure to the first national protest, I employ an instrumental variables approach to test how networks affect protest behavior. The main finding is that individual participation follows a threshold model of collective behavior: students were influenced by their networks to skip school on protest days only when more than 40 percent of the members of their networks also skipped school. Additional findings show that protest participation imposed significant educational costs on students and helped to shift votes towards non-traditional opposition parties. Taken together, results indicate that networks amplify the effect of protests in non-linear ways with potentially significant consequences for institutional change.

The second chapter, co-authored with José Ignacio Cuesta and Cristián Larroulet, investigates the workings of educational institutions in Chile. In education, data on school quality is often gathered through standardized testing. However, the use of these tests has been controversial because of behavioral responses that could distort performance measures. We study the Chilean educational market and document that low-performing students are underrepresented in test days, generating distortions in school quality information. These distorted quality signals affect parents' school choice and induce misallocation of public programs. These results indicate that undesirable responses to test-based accountability systems may impose significant costs on educational markets.

The third chapter, co-authored with Mounu Prem, studies firms during Chile's transition to democracy. Political transitions are associated with significant economic changes, but little is known about how firms fare across regimes. We study Chile's democratization and show that firms in the dictator's network make critical investments in physical capital during the political transition. These investments are made possible by government banks during the dictatorship and allow firms to improve their market position in the new regime. Our results show how market distortions can be transferred across political regimes.

Acknowledgements

I would like to thank my advisors Ted Miguel, Sol Hsiang, and Fred Finan for their constant support throughout the last years. Ted has been incredibly supportive of my research and gave me confidence to keep moving forward, Sol pushed me to think creatively about problems, and Fred inspired me to work hard. All of them provided invaluable feedback to improve my research. I would also like to thank Suresh Naidu for many conversations and for hosting me at Columbia for a couple of weeks. During my time at Berkeley I also learned a lot from talking with Ernesto Dal Bó, David Card, Jeremy Magruder, Andres Rodriguez-Clare, Gerard Roland, Jesse Rothstein, Chris Walters, and Noam Yuchtman. The third chapter of this dissertation benefited from generous funding from the Center for Effective Global Action, the Stanford Center for International Development, and the Economic History Association. I have also benefited from the funding provided by the Chilean government through the Conicyt scholarship.

I am thankful to work with amazing co-authors that I am happy to call my friends. Working with Mounu Prem has been a joy, he made research as fun as it can be, and I look forward to a life of writing papers and playing music. José Ignacio Cuesta has taught me more than he will ever recognize. Special thanks to Guillermo Marshall, who has been my friend since I ran my first regression and a role model for me during my Ph.D. I have also been fortunate to have spent these years at Berkeley talking sense and non-sense, enjoying inside and outside Evans, and discussing research and life with Ceren Baysan, Tamma Carleton, Tomás Monarrez, David Schoenholzer, and everyone at the Global Policy Lab.

Finally, I would like to thank my family. My parents worked hard to give me opportunities, have loved me unconditionally, have supported me in all of my decisions, and taught me about the intrinsic value of education. I will be forever grateful to you. Danae, the love of my life, thanks for absolutely everything that you do, eres la luz de mi vida. Gracias por todo.

Contents

1 Collective Action in Networks **1**

1.1 Introduction 1

1.2 Theoretical framework 4

1.3 The Chilean student movement 7

1.4 Data and descriptive statistics 8

1.5 Collective action in social networks 10

1.6 Multinetworks 15

1.7 Consequences of protests 18

1.8 Conclusion 22

2 Distorted Quality Signals **38**

2.1 Introduction 38

2.2 Institutional context 41

2.3 Data and attendance on test days 43

2.4 Distortions in quality signals 45

2.5 Implications for school choice 51

2.6 Misallocation of public programs 61

2.7 Conclusion 63

3 Losing Your Dictator **75**

3.1 Introduction 75

3.2 Chile's transition to democracy 79

3.3 Theoretical framework 83

3.4 Empirical framework 89

3.5 Firms during political transition 91

3.6 The credit market during political transition 96
3.7 Consequences during the democratic period 98
3.8 Conclusion 100

Bibliography **110**

Chapter 1

Collective Action in Networks

1.1 Introduction

Throughout history, organized groups of individuals have challenged the status quo and achieved significant social, economic, and political transformations. From the French Revolution to the Arab Spring, examples of groups aiming to transform societies are abundant. These organized groups are critical for institutional change and therefore, economic development patterns.

Individual participation in collective action has puzzled social scientists due to the presence of common benefits and private costs. This "collective action problem" has given rise to a large amount of theoretical literature emphasizing that the actions of others are crucial in order to understand individual participation.[1] However, an empirical investigation of how individual participation responds to the participation of others is still lacking. Given the enormous data requirements, the lack of evidence is not entirely surprising.

This paper studies the 2011 student movement in Chile, the largest protest mobilization in the country's history, in which hundreds of thousands of students skipped school to protest with the goal of reforming the educational system. I employ an instrumental variables approach to test for the role of networks in protest behavior using administrative data on millions of students' daily school attendance decisions on protest and non-protest days, a large network composed by the lifetime history of classmates, and differential network exposure to the first national protest.

The main finding is that participation in the student movement followed a pattern consistent with a threshold model of collective behavior. Students were influenced by their networks to skip school in protest days only when more than 40 percent of the members of their networks also skipped school, creating a bimodal distribution in par-

[1] See Olson (1965), Granovetter (1978), Tilly (1978), Kuran (1989), Lohmann (1993), and Marwell and Oliver (1993), among many others.

ticipation across groups. Skipping school imposed significant costs on students but it also shifted vote shares towards non-traditional opposition parties in the 2012 local elections. Taken together, results indicate that networks amplify the effect of protests in non-linear ways with potentially significant consequences for institutional change.

To organize the empirical analysis, I begin by presenting a simple framework that focuses on the individual decision to participate in a social movement. The participation decision has an individual cost and depends on aggregate participation and the participation of others in an individual's network. More participation in the network could decrease or increase individual participation in a linear or non-linear manner. The framework emphasizes a potential differential influence within a network and the possibility of multiple networks affecting the participation decision.

Two key features of the Chilean student movement allow me to empirically study participation in collective action. First, the central government assembles an exceptionally rich dataset of daily school attendance. Thus, I can measure participation in the movement for more than 800,000 high-school students using school absenteeism on days of protest. Second, students interact primarily with classmates (Araos et al., 2014), and information about their *lifetime* history of classmates is available. The latter data allows me to construct a country-wide network with more than 600 billion potential interactions and more than 60 million links between students who have shared a classroom.

The empirical analysis is divided in five parts. The first part uses an instrumental variables approach to estimate how network participation affects individual participation in the context of Manski (1993) "linear-in-means" model. The second part deviates from the linear model and tests for Granovetter (1978) "threshold model of collective behavior" using the non-parametric estimation proposed by Newey et al. (1999). The third part tests for the possibility of weak and strong ties embedded into the threshold model. The fourth part tests for additional influence from the network of neighbors. The fifth part estimates the effect of participation in the movement on students' academic performance and also on electoral outcomes in the 2012 local elections.

A crucial element in the analysis is the use of an instrument that solves the simultaneity of decisions, the possibility of unobservable variables causing a spurious positive correlation between students and their networks, and potential measurement error. The instrument is the exposure of networks to the first protest in May 12, organized by college students. Exposure is measured as absenteeism in May 12 in the set of students that (1) are part of the network of networks, and (2) are attending a different school in 2011. The instrument is similar to the one proposed by Bramoullé et al. (2009) and De Giorgi et al. (2010), although it is a refinement in the sense that it uses variation across protest days and focuses on students in different schools.

Network exposure to the first protest in May 12 is highly predictive of participation in June 16 – the first large national protest – even after controlling for a large set of observable variables for students, networks, schools, and city fixed effects. This is the preferred specification, but all results are qualitatively robust to the inclusion of fixed effects at the neighborhood, school, or school-grade level. Placebo checks using non-protest days confirm the importance of the May 12 protest.

Using the linear-in-means model, the estimates suggest that a 10 percent increase in network participation increases individual participation by 7 percent. This result is robust, statistically precise, and smaller in magnitude than a naive estimation that does not address endogeneity. However, a non-linear estimation reveals that a threshold model is a better representation of individual participation decisions. If the share of students in the network that participates in the movement is lower than 40 percent, the individual decision to participate is not affected by the network. After this threshold, individual participation increases rapidly with network participation. This result suggests that a "critical mass" of individuals is needed to facilitate participation.[2]

The critical mass of 40 percent should be interpreted as an average threshold. Students in larger schools, smaller networks, and smaller cities have lower thresholds. In addition, augmenting the estimation to allow for differential non-linear effects within networks, the estimates suggest that students are more influenced by others that are similar to them, a result that I interpret as evidence of weak and strong ties. Taken together, the results reject a linear-in-means model and suggest a critical mass of others that are similar is needed to foster participation. As a consequence, participation across network groups in the country follows a bimodal distribution with low and high levels of participation.

An exploration of additional sources of influence from students in schools close by confirms the previous results in two different ways. First, results are similar using a different identification assumption. These "multinetwork" results rely on network exposure to the first protest but across (instead of within) networks: participation among students in social networks is instrumented with exposure in *spatial* networks. To the best of my knowledge this "cross-network" identification is novel and potentially useful in other settings. Second, potential non-random measurement error in the network of students could introduce bias (Laumann et al., 1983; Kossinets, 2006; Chandrasekhar and Lewis, 2011). Reassuringly, results are robust to incorporate "neighbor" students into the analysis.

The last part of the analysis studies the consequences of protests. In particular, the focus is on the private costs of participation and the effects of the movement on electoral outcomes. A differences-in-differences analysis among students aged 6–10 and high-school students in the period 2002–2015 reveals that grade repetition increased by 60 percent, from a base of 6 percent, among high-school students in 2011. Using within school variation in 2011, I estimate that participation in the June 16 national protest decreased GPA by 0.1 standard deviations and increased grade repetition by 33 percent. Remarkably, these private costs of participation resemble the "critical mass" patterns previously discussed.

In addition to the private costs that participation had for students, I provide sugges-

[2] This tipping behavior is predicted by models of social interactions (e.g. Brock and Durlauf 2001). However, empirical evidence is limited. A notable exception is Card et al. (2008), who use Census tract data to provide evidence of tipping in the context of Schelling (1971) dynamic model of segregation.

tive evidence that the student movement was able to shift votes towards non-traditional opposition parties, which were relatively more aligned with the movement's demands. A cross-sectional regression using county-level electoral data suggests that a one standard deviation increase in the intensity of the movement in local schools increased vote shares for non-traditional parties by 5–10 percentage points. Interestingly, although arguably speculative, the effect on vote shares seems to be non-linear, a result consistent with the previously described "critical mass" patterns of participation.

This paper contributes to the empirical understanding of participation in collective action. Only a few number of articles have studied *protest* participation.[3] Using an annual panel dataset of geographic cells in Africa, Manacorda and Tesei (2016) show that mobile phone coverage facilitated protests when countries experienced economic downturns. Enikolopov et al. (2016) show that the penetration of an online social network in Russia increased the probability of a protest and the number of protesters across cities. Finally, Acemoglu et al. (2016a) show that citizens' discontentment on Twitter predicts daily protest participation in Tahir Square during Egypt's Arab Spring. To the best of my knowledge this is the first paper to estimate how individual specific networks affect individual participation in collective action.

This paper also speaks to a related literature estimating the consequences of protests.[4] Madestam et al. (2013) uses rainfall shocks as exogenous variation affecting the number of protesters from the Tea Party Movement across U.S. counties to show how the movement affected electoral outcomes and the policies being implemented. Aidt and Franck (2015) show that the Swing riots in early 19th century Britain – credible signals of the threat of a revolution – facilitated democratic reforms. This paper contributes to this literature by providing novel evidence on the individual costs associated with participation and suggestive evidence on the effect of the student movement on electoral outcomes.

The next section presents a theoretical framework for the individual decision to participate in collective action. Section 1.3 provides details about the 2011 student movement in Chile. Section 1.4 presents the data and describes participants. Section 1.5 tests for different models of participation in collective action. Section 1.6 tests for additional complementarities in space. Section 1.7 estimates the costs of participation in the student movement and estimates its effect on the 2012 local elections. Section 1.8 concludes.

1.2 Theoretical framework

This section presents a simple framework for the individual decision to participate in a social movement. The objective is to lay out testable features of participation as a

[3]There are studies of participation in other types of collective action. For example, McAdam (1986) shows that friends' participation in the 1964 Freedom Summer project predicts individual participation, and Yanagizawa-Drott (2014) shows that radios facilitated participation in the Rwandan genocide.

[4]There is, of course, a large theoretical literature studying social unrest and political transformation. See, for example, Acemoglu and Robinson (2009).

function of the participation of others. In this framework, each individual interacts with a specific network and perfectly observe their participation in the movement.[5]

Network participation could affect individual participation for multiple reasons, including pressure to conform, strategic complementarities, and information updating. Although some results in sections 1.5 and 1.6 arguably help to distinguish between these mechanisms, I remain agnostic about which one is relatively more important.

1.2.1 Environment

There are I individuals in a society. Let a_i be an indicator variable that takes the value of one if individual i participates in a movement that is revolting against the status quo and zero otherwise. Each individual i interacts with a group of individuals, i's network. Individual i perfectly observes participation in her network, composed by n_i individuals and denoted by the vector $\vec{a}_{j(i)} = (a_1, \ldots, a_{n_i})$. She also observes aggregate participation, denoted by the scalar $a_{-i} \equiv \frac{\sum_{k \neq i} a_k}{I}$. To simplify notation, let $n_i \equiv n$ and $\vec{a}_{j(i)} \equiv \vec{a}_j$.

The utility of individual i from participating in the movement depends on her own action a_i, aggregate participation a_{-i}, and the participation of others in her network $\vec{a}_j$:

$$u_i(a_i|\vec{a}_j, a_{-i}) \quad = \quad p(a_{-i})B + \left(f(\vec{a}_j) + b_i - c_i\right) a_i \tag{1.1}$$

where $p(a_{-i})$ is the probability of achieving change, B is the benefit of changing the status quo, $f(\vec{a}_j)$ is an unknown function of the participation of others in the network, b_i is an individual benefit derived from participation, and c_i is an individual cost of participation.[6] Note that if $f(\vec{a}_j) = 0$, then only individuals with $b_i > c_i$ participate.

Individual i decides to participate in the movement if the utility from participating $u_i(a_i = 1|\vec{a}_j, a_{-i})$ is higher than the utility from not participating $u_i(a_i = 0|\vec{a}_j, a_{-i})$:

$$f(\vec{a}_j) + b_i - c_i \quad > \quad 0 \tag{1.2}$$

where it is easy to see that if $f(\vec{a}_j) > 0$ (or $f(\vec{a}_j) < 0$) some individuals with $b_i < c_i$ ($b_i > c_i$) will be pushed towards participation (non-participation) because of "network effects." Section 1.7 calculates empirically this share of individuals.

1.2.2 Testable features of participation

This section describes four testable features of participation. First, is individual participation increasing or decreasing in the participation of others in the network? This

[5]For a thorough theoretical analysis, including equilibrium conditions, see Bramoullé et al. (2014) and Blume et al. (2015).

[6]Other predetermined observable variables of networks may also affect utility. These variables are omitted for simplicity but incorporated in the empirical analysis.

question can be answered empirically using a simple "linear-in-means" model. This model was used by Manski (1993) to discuss econometric challenges and assumes that:

$$f(\vec{a}_j) = \beta \cdot \frac{1}{n}(a_1 + \cdots + a_n) = \beta \cdot \overline{a_j} \tag{1.3}$$

This functional form is a modeling choice and may or may not be the best representation of the data. Nevertheless, if $\beta > 0$ then individual participation is increasing in the participation of others in the network and the reverse is true if $\beta < 0$.

A slightly more general representation of the previous model allows for differential influence within i's network. These "weak and strong ties" are emphasized by Granovetter (1973) and can be represented as:

$$f(\vec{a}_j) = \beta \cdot (\omega_1 a_1 + \cdots + \omega_n a_n) \tag{1.4}$$

This function allows individuals to take a *weighted* average of the participation of others. The weights ω_k, with $k = 1, \ldots, n$ and $\sum_k \omega_k = 1$, represent the differential influence that others have in i's participation decision.[7]

A model with differential influence can take many forms, depending on the modeling choice for the weights ω_k. Consider two examples. First, we could try to non-parametrically estimate these weights. After estimation, we could characterize influential individuals. Second, we could parameterize these weights using observable variables. The former approach is being explored by an ongoing research agenda (e.g. Manresa 2016). The latter approach has been, to the best of my knowledge, relatively unexplored empirically. Section 1.5 explores a "homophily model of influence" in which the difference in observable variables between the individual and others in her network determines the strength of the influence. In the absence of this type of influence, differences in observables should not determine the strength of influence.

The third testable feature of the model is a potential non-linearity in f. In seminal studies of collective action, Tilly (1978) and Granovetter (1978) argue that the individual decision to participate in a movement may be influenced by the participation of others only when there is a "critical mass" participating. This is:

$$f(\vec{a}_j) = g\left(\overline{a_j}\right) \cdot \mathbb{1}\left(\overline{a_j} > \alpha_i\right) \tag{1.5}$$

where $\overline{a_j}$ is the share of i's network that participates, and $\mathbb{1}(\overline{a_j} > \alpha_i)$ is an indicator function that takes the value of one if the share participating is larger than $\alpha_i \in [0, 1]$. The threshold α_i may be individual specific and could lead to multiple equilibria in participation.

The final testable feature of participation is that individuals may interact in K differ-

[7] In a study of participation in the 1964 Freedom Summer project, McAdam (1986, p. 88) shows that a strong tie to a participant is a better predictor of individual participation than a weak tie.

ent networks or, equivalently, there are K different types of links:

$$f(\vec{a_j}) = f_1(\vec{a_{j_1}}) + \cdots + f_K(\vec{a_{j_K}}) \tag{1.6}$$

where $f_k(\vec{a}_{j_k})$ represents the influence of others in network j_K. Two examples are social and spatial networks: individuals may be affected by the participation of friends that are not their neighbors or vice versa. If only social networks (say $k = 1$) affect individual participation, then $f(\vec{a_j}) = f_1(\vec{a_{j_1}})$ and $f_k(\vec{a}_{j_k}) = 0 \; \forall k \neq 1$.

1.3 The Chilean student movement

From the Tunisian demonstrations sparking the Arab Spring to Occupy Wall Street triggering a movement against inequality, 2011 will be remembered as the year of the protester. The global wave of citizens demanding a "new democracy" also took place in Chile, where students revolted to reform the educational system installed by the Pinochet dictatorship, nowadays one of the most expensive and segregated in the world (Hsieh and Urquiola, 2006; OECD, 2013). Organized groups of students triggered the largest demonstrations in the country's history, which were recognized worldwide as one of the most important social movements of the year.

The student movement began its protest activities in May of 2011, two months within the academic year and 14 months after a right-wing government took office for the first time in 50 years.[8] The initial demonstrations were triggered by delays in the assignment of students' scholarships and bus passes. The first student-led national protest took place on May 12 and thousands of high-school and university students participated.[9]

The first national protest was organized by the Confederation of Chilean Students, a national student organization, with the objective of exerting pressure on the annual presidential speech of May 21, day in which the government outlines next year policies. Students wrote a document outlining policies to decrease segregation in the educational system and increase government spending on education. After the presidential speech, the confederation of students sent a letter to the ministry of education expressing their discontent with the presidential announcements (Confech, 2011). Students called for another national protest day in June 1, the last rally before the movement expanded in an unprecedented way.

After the national protest of June 1, and a failure to reach an agreement with the ministry of education in meetings held in May 30 and June 8, students intensified their protest activities. The movement was gradually supported by deans – including those in prestigious public universities – teachers, prominent worker unions, and public figures.

[8]Chronicles written by leaders of the student movement are Figueroa (2012), Vallejo (2012), and Jackson (2013). A brief history of movement among high-school students can be found in Simonsen (2012).

[9]For additional context, Figure A.1 plots the daily number of protests in Chile in the period 1979-2013 and Figure A.2 plots economic indicators around the beginning of the student movement of 2011.

Over the weeks that followed, students occupied schools and universities and protest activities spread across the country. In an attempt to prevent occupations, the Ministry of Education asked students "to stop protesting" and the president stated that "countries do not progress by occupying schools." The government's approval was low and continued to plummet after the rise of the movement (Figure A.3). Students called for another national protest day on June 16, at the time the largest mobilization in the country's history. The government responded in June 25 with an offer and students rejected it and called for yet another national protest day on June 30.

Education was the main topic of conversation in the months of July and August. The leaders of the movement were regularly invited to television and radio shows, and diverse protest activities filled the country. Students ran and danced in front of the government palace, kissed in public spaces to gain citizens' support, and exerted pressure on the government using different types of non-violent resistance. Rainy days and school holidays did not stop students' activities, with large protests taking place under thousands of umbrellas. The president replaced the ministry of education in July 18 and the government responded to students' demands with offers on July 5, August 8, and August 17. Students rejected these offers and demonstrations continued after the school break of July, with the largest national protests taking place in August 24 and 25. These two days marked the peak of the student movement and protest activities declined in the following months.

Various reasons explain the decay of the student movement, including the beginning of formal negotiations, the focus of popular media on violent protesters, and students' concerns about grade retention.[10] After months of protests, what were the consequences? Contemporary surveys show that 80 percent of citizens supported the movement (Adimark, 2011) and education became a national priority (Figure A.4). Candidates in the subsequent local elections in 2012 and congress and presidential elections in 2013 were constantly questioned about their ideological position regarding education. Some of the older leaders of the movement founded political parties and four of them won seats at the congress.

1.4 Data and descriptive statistics

The first part of this section describes the data used in the empirical analysis. All seven administrative datasets were provided by the Ministry of Education in Chile. Six datasets contain information about students and one dataset describes schools. The second part of this section describes the participation of students in the student movement.

[10]Jackson (2013, p. 22) states: "the constant emphasis on violence affected the strength of the movement". The government threatened students with grade retention promoting the "Let's save the academic year" plan. In addition, public figures died in an airplane crash in September 2 – shifting public interest away from the movement – the movement's leaders had to face annual elections to renew their leaderships, and summer holidays caused the movement to retreat until the next academic year.

1.4.1 Administrative datasets and cities

The first dataset of students presents information about daily attendance to school in 2011. The second dataset contains students' enrollment information (school, grade, classroom). There are approximately 3,000,000 students every year, and 975,000 high-school students enrolled in 2,700 high-schools in 2011. The third dataset contains information on students' annual academic performance (GPA, school attendance, repetition). These three datasets are available for all students enrolled in the educational system.

Three additional datasets contain more information about high-school students in 2011. The fourth dataset corresponds to students' performance in standardized tests, taken some years by students in specific grades. Approximately 40 percent of high-school students in 2011 took the test before that year. The fifth dataset corresponds to household surveys, conducted in parallel to standardized tests. These surveys allow me to measure household income and students' internet connection at home, data available for 57 percent and 36 percent of high-school students in 2011. The sixth dataset contains self-reported home addresses and it is available for 35 percent of high-school students in 2011.

The seventh and last dataset contains information about schools. Approximately 40 percent of students were enrolled in public schools in 2011, 60 percent in private schools, and 96 percent attended urban schools. School addresses are available and I use these to construct geographic clusters that I refer to as "cities." These cities are isolated components in the network of schools, where schools are linked if these are closer than 5 kilometers from each other (see Figure A.5 for a map of cities). Table 1.1 presents descriptive statistics for students, schools, and cities.

1.4.2 Participation in the student movement

To measure students' participation in the movement I use school absenteeism among high-school students during protest days in 2011. Several patterns in the data suggest this is indeed a useful way to measure participation. First, there are significant spikes in absenteeism during protest days. The upper panel in Figure 1.1 plots absenteeism from the beginning to the end of the 2011 school year. The first two national protest days (May 12 and June 1) are easy to observe in this figure. The sharp increase in school absenteeism between June 1 and June 16 maps corresponds to the real-time escalation of protest activities. Second, some schools were temporarily taken over by students and these closures are observed in the data with the correct dates. As examples, the lower panels in Figure 1.1 present three time series of school-level absenteeism.

Given that the government collects information about students to measure performance and allocate public programs, I am able to describe participation in the movement in an unusually rich way. Let student i's participation be defined as $\max\{0, A_{iT} - A_{i\tau}\}$, where A_{iT} and $A_{i\tau}$ represent student i's absenteeism before and after June 1.[11] This

[11]The purpose of the maximum function is to truncate participation to be positive, although it is in-

measure of participation accounts for heterogeneity in absenteeism at the student level.

The average absenteeism of high-school students before and after June 1 was 13 percent and 40 percent respectively, which means that absenteeism increased by more than 200 percent. The average participation of a student is 0.30 (s.d. 0.32) and 4 percent of students did not go back to school in 2011. The participation of an average school is 0.20 (s.d. 0.22). Figures 1.2-A and 1.2-B plot the distribution of this participation measure for students and schools respectively.

Participation in the student movement was similar across students from different income groups and different performance measures. Figure 1.2-C plots the correlation between students' participation and annual household income. Students from families earning US$10,000 annually skipped 30 days more than usual, while students from families earning more than US$30,000 missed 23 days more than usual. Figure 1.2-D shows that standardized test scores have little predictive power for participation. Both of these correlations are robust to the inclusion of student-level controls and city fixed effects.

Lower quality schools with more students connected to the internet participate more in the movement. Figure 1.2-E plots the correlation between participation and schools' average test scores, a quality measure. A one standard deviation increase in quality signals is associated with a decrease of 4 percentage points in participation. Figure 1.2-F shows that, after controlling for household income, a one standard deviation increase in internet connection is associated with an increase of 8 percentage points in participation.

1.5 Collective action in social networks

This section tests for potential complementarities in school absenteeism decisions between students and their networks during protest days. The focus is on June 16, at the time the largest mobilization in the country's history and the first massive national protest day in a two and a half months period of intense protest activities (details in section 1.3). After describing the main regression equation of interest, I define social networks, discuss the main identification concerns, describe the identification strategy, and present results. In short, the identification strategy uses the exposure of students' networks to the *initial* protest in *other* schools, an exposure that arises due to predetermined switching of students across schools.

1.5.1 Estimating equation and social networks

Consider the following regression relating a student's decision to skip school in the first largest protest day as a function of school absenteeism in her social network:

$$A_{isc} = f(A_{j(i)}) + g(b_i, c_i) + g(b_{j(i)}, c_{j(i)}) + \delta x_s + \zeta_c + \epsilon_{isc} \qquad (1.7)$$

nocuous in the sense that few students decrease their absenteeism after June 1 in 2011.

where A_{isc} is an indicator that takes the value of one if student i in school s located in city c decides to skip school in June 16. In addition, $f(A_{j(i)})$ is a function of a vector containing the absenteeism decisions of i's social network $j(i)$, and $g(b_i, c_i)$ and $g(b_{j(i)}, c_{j(i)})$ are functions of observable variables that account for the benefits and costs that may affect students' decision to participate. Finally, x_s is a vector of control variables at the school level, ζ_c is a city fixed effect, and ϵ_{isc} is an error term clustered at the city level.

The analysis begins using a linear-in-means function f. The vector of control variables b_i and c_i include average school attendance in 2010, GPA in 2010, an indicator for grade retention in 2010, an indicator for female, an indicator for students who switched school in 2010, and age. Averages of the same variables are included in $b_{j(i)}$ and $c_{j(i)}$, although results are robust to use more flexible functions. In addition, student controls also include school absenteeism during the May 12 and June 1 protest days. School-level controls include an indicator for public schools, reported quality signals (i.e. test score averages), the percentage of students who have repeated a grade in the past, and average household income.

Because students interact mostly with other students in their classroom, I define student i's social network $j(i)$ as the lifetime history of classmates. This definition of social networks gives rise to a large network of students linked within *and* across schools. Links across schools arise from switching of students across schools before 2011. Overall, this network contains more than 600 billion potential interactions between students across the country, and more than 60 million existing links in 2011. The average student has 80 other students in her social network, 60 percent attending the same school and 40 percent attending a different school in 2011.[12]

1.5.2 Identification strategy

There are two concerns with a naive estimation of equation (1.7). First, the reflection problem emphasized by Manski (1993): students affect their social networks and social networks affect students. Second, there may be unobservable variables causing students and their networks to make similar decisions. Both concerns imply that an OLS estimation will overestimate the effect of social networks on student's decisions. To solve these issues, I exploit three sources of variation in an instrumental variables approach.[13]

The first source of identifying variation is the exposure of social networks to protests in their social networks. The second source is a restriction to students attending a differ-

[12]The term "social network" is coined for expositional purposes as other individuals may be part of a student's network. The possibility of measurement error in social networks is addressed in the following sections. For computational reasons results use social networks defined as classmates in the previous four years. Results are robust to using more or less previous years.

[13]An additional source of bias is known as "exclusion bias" and causes OLS estimates to be biased *downwards* (Guryan et al., 2009; Angrist, 2014; Stevenson, 2015; Caeyers and Fafchamps, 2016). To deal with this bias I follow Caeyers and Fafchamps (2016) and include the student's value of the instrument (absenteeism in May 12) as an additional control variable.

ent school in 2011. The third source is the first national protest day in May 12, organized outside of the network of high-school students (see section 1.3). All in all, this strategy is similar to the one proposed by Bramoullé et al. (2009) and De Giorgi et al. (2010) with two important differences: the use of variation across days and, to minimize the threat of unobservables, a restriction on the set of second degree connections.

To gain intuition let student i's network be denoted by the set n_i. The exposure of students in n_i is measured by how much *their* networks N_i participated in May 12, with $i \notin N_i$. Students in the set N_i may however still have similar unobservable variables than i. To deal with this concern, I restrict attention to a subset of students. Given the predetermined switching across schools, many students in N_i are attending a different school in 2011. Let M_i be the set of students that attend a different school than i in 2011, with $M_i \subset N_i$ and $n_i \cap M_i = \emptyset$. The identification assumption is thus that school absenteeism in May 12 among students in the set M_i only affects student i absenteeism in June 16 through the absenteeism of n_i.

The first stage using the previously described instrument is strong (see Table A.1), with coefficients having the expected positive sign – higher initial exposure fosters participation – and corresponding F-stats always far from a weak instrument problem (Stock and Yogo, 2005). Reassuringly, the value of the instrument in non-protest days before May 12 does not predict networks' absenteeism in June 16 (see Figure A.6), suggesting that unobservable variables that affect absenteeism on non-protest days are unlikely to be a concern.

1.5.3 Linear estimates

Table 1.2 present estimation results of a linear-in-means model. Panel A presents OLS estimates of different specifications of equation (1.7). Although the focus is on the effect of network, school absenteeism during the initial protests of May 12 and June 1 are also interesting because these could potentially be measures of habit formation in protest activity.[14] As the mean of the dependent variable and the main variable of interest are similar (0.49 and 0.50), point estimates can be interpreted directly as an elasticity.

Column 1 in Table 1.2-A presents estimates without control variables, a regression that explains almost two-thirds of the variation in June 16 absenteeism.[15] The estimated coefficient implies that a one standard deviation increase in network absenteeism (0.31) is associated with an increase of 38 percentage points (0.31×1.23=0.38) in the probability of skipping school. In terms of elasticities, a 10 percent increase in network absenteeism is associated with a 12 percent increase in student absenteeism. In addition, skipping school in the first two protest days increases the probability of skipping school in June 16 by 14 percentage points (0.06+0.08=0.14). Columns 2-5 progressively control for

[14] The number of observations is presented in the bottom of Table 1.2. Differences in observations are due to missing values, which are more common in small schools located in rural areas.

[15] In contrast, a regression on student, network, and school characteristics explains less than one-third of the June 16 variation in absenteeism, suggesting network effects and habit formation are important.

student, network, and school characteristics, and city fixed effects. As a result, the coefficient of networks remains stable. Although regressions control for a large set of observable variables at multiple levels, reflection and potential unobservable variables could cause a comovement of decisions between students and their networks.

A leading concern with estimates in Table 1.2-A are neighborhood unobservable variables causing a spurious positive correlation between students and their networks. To explore this possibility, I geo-coded 50,000 home addresses of students in Santiago, capital of Chile, and construct neighborhood fixed effects using latitude and longitude coordinates, creating areas of approximately 10×10 blocks (see Figure A.7 for a map). Column 7 includes these 714 indicators and estimated coefficients are unchanged, providing some evidence that neighborhood level variables are unlikely to be a concern. However, there may be *additional* complementarities in neighborhoods, i.e. absenteeism may be influenced by neighbors that were not classmates. I explore this possibility in section 1.6.

Table 1.2-B presents instrumental variables results. Table A.1 presents first-stages and reduced forms. As expected, estimated coefficients are positive and smaller in magnitude than their OLS counterparts. Columns 1-6 show the coefficient is robust with an elasticity of 0.6–0.8. Importantly, F-statistics in first-stages are always strong. Column 5 in Table 1.2-B is the preferred specification. This result is robust to excluding schools closed by students in June 16: the 2SLS coefficient for networks is 0.53 (s.e. 0.14) with a first-stage F-stat of 30.2 (see Table A.2). When including school fixed effects the coefficient for networks decreases to 0.07 (s.e. 0.03, first-stage F-stat of 77.7).

The remainder of this section explores two deviations from the linear-in-means model. First, potential non-linear effects of the participation of others. Second, differential influence within students' networks.

1.5.4 Critical mass

To test for non-linear networks effects, I use the nonparametric approach proposed by Newey et al. (1999). In this control function estimation, the coefficients of interest are associated to indicators for different values of absenteeism in social networks. The benchmark estimation uses eleven indicators: the first takes the value of one if absenteeism in social networks is between 0 and 10 percent, the second for 10-20 percent absenteeism, and so on until 100 percent absenteeism. Using equation (1.7):

$$f(A_{j(i)}) = \beta_1 \cdot \mathbb{1}[A_{j(i)} \in [0.1, 0.2)] + \cdots + \beta_9 \cdot \mathbb{1}[A_{j(i)} \in [0.1, 1)] + \beta_{10} \cdot \mathbb{1}[A_{j(i)} = 1] \tag{1.8}$$

where $(\beta_1, \ldots, \beta_{10})$ are the parameters of interest and the omitted category is absenteeism lower than 10 percent in social networks.

The upper-left panel of Figure 1.3 presents these ten estimated coefficients $(\widehat{\beta}_1, \ldots, \widehat{\beta}_{10})$ with their corresponding 95 percent confidence interval using the Newey et al. (1999) approach. The figure also plots the analog OLS estimates for comparison. The exact specification corresponds to column 5 in Table 1.2, which includes student, network, and school controls, and also city fixed effects. Similar to the linear estimates, the control function approach delivers estimates that are lower in magnitude than the OLS counterparts, as expected.

Consistent with the threshold model of collective behavior proposed by Granovetter (1978), the estimated coefficients show "critical mass" patterns. Student's absenteeism is not affected by low values of absenteeism in social networks. In contrast, large values of social network absenteeism have strong effects on students' decisions to skip school. The upper-right panel of Figure 1.3 plots the sequential difference between estimated coefficients to understand the marginal contribution of additional absenteeism in the social network, suggesting that approximately 30-40 percent is the critical mass needed for networks to have an influence on students.

To estimate the average threshold in which networks begin to influence decisions, I repeat the previous estimation strategy but using 51 indicators for social network absenteeism, from 0 to 100 percent absenteeism in steps of 2 percentage points. The lower-left panel of Figure 1.3 presents the coefficients for these indicators. A vertical line marks the point in which the coefficient becomes statistically different from zero: 40 percent of absenteeism in the social network. The lower-right panel plots the distribution of absenteeism in social networks, a bimodal distribution consistent with a critical mass pattern of influence.

These critical mass results hold when omitting schools that were closed by students, and are qualitatively similar when including school or school-grade fixed effects (see panels A, B, and C in Figure 1.4). In addition, the critical mass of 40 percent should be interpreted as the average threshold. Panels D, E, and F in Figure 1.4 shows that students in larger schools, smaller networks, and smaller cities have lower thresholds.[16]

In sum, the estimated coefficients in Figures 1.3 and 1.4 provide evidence of "critical mass" type of complementarities in protest behavior: students' absenteeism in protest days increases with network absenteeism only when more than 40 percent of other students in their networks skipped school.

1.5.5 Differential influence within networks

The empirical regularity of individuals forming links with others with similar characteristics is known as homophily (Jackson, 2010, chapter 6). Empirical work testing for differential *influence* following homophily patterns within networks is, however, more limited. Conditional on a network structure, i.e. links are already formed, does the strength of influence follows homophily patterns? This section tests this hypothesis focusing on three variables: gender, internet connection, and household income.

Table 1.3 presents results for the linear-in-means model. Columns 1 and 2 test for gen-

[16] Figure A.8 shows additional patterns of heterogeneity. In addition, Figure A.9 shows that results are identical in schools with low and high levels of irregular spending of government transfers (CIPER, 2012).

der homophily patterns of influence by estimating a linear-in-means version of equation (1.7), restricting attention to males or females, and splitting the network in males and females. For estimation I use the control function approach proposed by Newey et al. (1999). Under the null hypothesis of equal influence we should observe similar coefficients for the male and the female parts of the network. Results, however, indicate strong homophily patterns: same gender influence is more than ten times stronger than cross gender influence.

Columns 3 and 4 restrict attention to students with and without internet connection respectively, split the network in two, students with and without internet connection, and follow the same estimation strategy than before.[17] Estimated coefficients indicate that the influence of students with internet on other students with internet is three times larger. The influence of students without internet on students without internet is two times larger.

Similar patterns of influence arise when restricting attention to the position of students in the income distribution. Columns 5–7 show that students from poor households are more influenced by students from poor households, and students from rich households are more influenced by students from rich households (and not by students from poor households).[18]

Figure 1.5 present the nonparametric analogue of previous results. Patterns of critical mass are still clear when allowing for differential influence, and the null hypothesis of equal influence is easily rejected. Overall, results in Table 1.3 and Figure 1.5 provide evidence that supports the hypothesis of homophily patterns of influence within networks.

1.6 Multinetworks

This section tests for additional complementarities in other networks. Given the saliency of geographic location, I incorporate *spatial* networks into the analysis. The motivation is that individuals may be influenced by neighbors that are not in the previously defined social network. The findings in this section confirm previous results in two ways. First, these results rely on a different identification assumption. Second, these results are less likely to be subject to measurement error.[19]

[17]This is a partial test for the hypothesis of stronger coordination with internet because students (1) may have internet access at the school, and (2) may coordinate with *other* networks over the internet. Manacorda and Tesei (2016) and Enikolopov et al. (2016) provide city-level evidence of stronger network coordination with more access to cell phones and social media.

[18]Rich households are defined as those with reported annual income higher than US$16,000, poor households with reported annual income lower than US$5,000, and the rest is defined as the middle class.

[19]Laumann et al. (1983) show how "missing" links create bias in network statistics, Kossinets (2006) discusses different sources of measurement error, and Chandrasekhar and Lewis (2011) discusses the implications for regression analysis. Importantly, working with administrative data for the universe of

The empirical challenge is to simultaneously test for complementarities in both social *and* spatial networks. Because home location is not available for all students and schools' exact geographic location is, I now use schools as the unit of observation. First, I discuss the main estimating equation. Second, I propose an instrumental variables strategy that is based on "cross-network" exposure to the initial protest. Third, I present results.

1.6.1 Estimating equation

Consider an extended version of equation (1.7) that includes potential complementarities in social *and* spatial networks:

$$A_{sc} \quad = \quad f(A_{n(s)}) + f(A_{m(s)}) + \gamma x_s + \theta_c + \varepsilon_{sc} \tag{1.9}$$

where A_{sc} represents students' average absenteeism in school s in city c in June 16, $A_{n(s)}$ and $A_{m(s)}$ represent students' absenteeism in spatial $n(s)$ and social $m(s)$ networks in June 16, and x_s is a vector of control variables chosen using the method proposed by Belloni et al. (2013).[20] The vector θ_c controls for city fixed effects and the error term ε_{sc} is allowed to be spatially correlated within cities. The estimation of f corresponds to the statistical tests of interest, with $f' > 0$ providing evidence of complementarities in multinetworks. Given the fewer number of observations I use a linear-in-means function f for most of the analysis, but I also discuss and present estimates of flexible estimation of this function.

1.6.2 Identification strategy

The main concern with estimation of equation (1.9) is the potential existence of unobservable variables that affect both the absenteeism of schools and their networks. This source of bias will cause a spurious positive correlation between schools' and networks' absenteeism. Crucially for the identification strategy I propose, the unobservables causing comovement can differ across networks. For example, the use of police force to decrease absenteeism in a geographic area will cause a comovement in spatial networks but not necessarily in social networks.

Before presenting the key equations of the econometric strategy, let me define two sets of schools. Let $m(s)$ represent the set of schools in the social network of s, $n(s)$ the set of schools in the spatial network of s, and $\ell(s) \equiv m(s) \cup n(s)$. The first set of interest corresponds to schools in the social network of a spatial network, $m(n(s)) \equiv m(n)$ under the previous notation, with $m(n) \cap \ell(s) = \varnothing$. The second set of interest corresponds to schools in the spatial network of a social network, $n(m(s)) \equiv n(m)$ with $n(m) \cap \ell(s) = \varnothing$.

students mitigates a significant number of concerns.

[20] Control variables include: school absenteeism in May 12, school absenteeism in June 1, school absenteeism in 2010, and an indicator for public schools.

Schools in the disjoint sets $n(s)$ and $m(n)$ are linked in the social network, while schools in the disjoint sets $m(s)$ and $n(m)$ are linked in space. With these two sets of schools in mind, consider the following first-stage regressions:

$$A_{m(s)} = \tau_1 Z_{n(m)} + \tau_2 Z_{m(n)} + \phi_1 A_{n(s)} + \gamma_1 x_s + \theta_c + \eta_{sc} \quad (1.10)$$
$$A_{n(s)} = \tau_3 Z_{n(m)} + \tau_4 Z_{m(n)} + \phi_2 A_{m(s)} + \gamma_2 x_s + \theta_c + \eta_{sc} \quad (1.11)$$

where $A_{m(s)}$ and $A_{n(s)}$ represent school absenteeism in the social and spatial networks of s in June 16, Z_{m_n} represents school absenteeism in the spatial network of m_s in May 12, and Z_{n_m} students' average absenteeism in the social network of n_s in May 12. Both Z_{m_n} and Z_{n_m} are the instruments I propose to isolate exogenous variation in network protest activity.

The relevance condition behind the proposed instruments follows a simple logic: if coefficients τ_2 and τ_3 are positive and statistically different from zero, then exposure to the initial protest increases protest participation in June 16. To gain intuition, Figure 1.6 presents this identification strategy graphically. In the upper panel the interest is on the effect of B on A through a link in the social network. The initial participation of B's spatial links (marked in red) affects B's participation, which in turn could affect A's participation. The same logic applies in the lower panel, where the focus is now on the effect of A's spatial links: the initial participation of B's social links (marked in red) affects $B's$ participation, which again could affect A's participation. Then, the exclusion restriction is that B's links affect A only through B.

Equations (1.10) and (1.11) also provide some evidence for this approach. Coefficients τ_1 and τ_4 should not be statistically different from zero. This should be the case because, after controlling for $A_{n(s)}$, we should not observe that social networks have an additional effect because all their influence is through $A_{n(s)}$, implying that $\tau_1 \approx 0$. The same argument applies in equation (1.11): after controlling for $A_{m(s)}$, there should be no effect of spatial networks, which implies that $\tau_4 \approx 0$.

1.6.3 Results

The multinetwork is composed by 2,070 high-schools (i.e. nodes) and two types of links (i.e. edges): spatial and social. Links are defined in the following way. Hypothetical schools A and B are linked in the social network if students transferred between these schools in previous years. In addition, schools A and B are linked in space if these are geographically close enough.[21] Note that two schools can be theoretically linked in two networks and networks are imperfectly overlapped. Figure 1.7 presents a visualization of social and spatial networks. The average school has 3.4 spatial links, 1.7 social links, and a total of 4.6 links. The number of schools without any type of link is 177, with 709 schools having zero spatial links and 453 having zero social links.

[21] I use transfers of students in 2010 to define social links and one kilometer as the threshold for spatial links – roughly 10 blocks – although results are robust to different definitions.

Table 1.4 presents estimation results of three versions of equations (1.10) and (1.11). The first specification includes only the instruments $Z_{n(m)}$ and $Z_{m(n)}$, the second specification adds controls, and the third adds city fixed effects. The column pairs 1-4, 2-5, and 3-6 correspond to the first-stages in the three specifications. The bottom of the table presents the Angrist-Pischke F-statistic to test the statistical strength of the relevant instrument, and the Cragg-Donald F-stat for the combined strength of both first stages. In addition, columns 7-9 present the corresponding reduced form regressions. Overall, estimated coefficients and statistical tests suggest that the instruments are valid. First-stage coefficients are positive as hypothesized and statistically different from zero (i.e. $\tau_2 \gg 0$, $\tau_3 \gg 0$), I can reject the presence of weak instruments, there is evidence to support the approach (i.e. $\tau_1 \approx 0$, $\tau_4 \approx 0$), and the reduced form coefficients are always statistically significant and positive.

Table 1.5 presents two-stage least squares estimates (columns 4-6) and OLS coefficients for comparison (columns 1-3) for the three specifications previously discussed. Overall, estimated coefficients suggest the existence of complementarities in both spatial and social networks, and I cannot statistically reject that coefficients in both networks are different (p-value of 0.63). These linear complementarities are robust to the exclusion of schools without students' attendance in June 16: spatial and social network coefficients are 0.27 (*s.e.* 0.08) and 0.16 (*s.e.* 0.06) respectively. Estimated coefficients in my preferred specification (column 6) imply that a one standard deviation increase in absenteeism in the spatial (social) network causes an increase in school absenteeism of 9 (7) percentage points, an increase of 23 (18) percent. The corresponding elasticities with respect to absenteeism in spatial and social network are 0.3 and 0.2.

Given the number of observations, estimated coefficients are less precise than those in section 1.5 but still statistically different from zero. Although there is less statistical power to estimate potential non-linearities in both networks, for completeness Figure A.10 presents estimated coefficients similar to the ones presented in Figure 1.3. Not surprisingly, estimated coefficients have wide confidence intervals so I cannot reject the existence of non-linearities in spatial networks.

1.7 Consequences of protests

This section estimates the costs of participating in the student movement and its effects on electoral outcomes, i.e. estimates of private costs and common benefits. A cohort analysis in a differences-in-differences framework reveals that grade repetition increased by 60 percent among high-school students in 2011. Within schools in 2011 more participation is associated with lower academic performance and a 33 percent increase in the probability of repeating the grade. An analysis of the 2012 local elections suggests that the movement shifted votes towards non-traditional opposition parties, relatively more aligned with the movement's demands. The section ends with a counterfactual calculation for the contribution of networks to protest activities emphasizing the importance of allowing for non-linearities.

1.7.1 The cost of participation

Cohort analysis. Analysis of administrative data for the period 2002–2015 shows that participation in the movement lead to an increase in grade retention, an outcome causally associated with higher dropout rates, lower educational attainment, and more criminal activities (Manacorda, 2012; Díaz et al., 2016).

To estimate the change in grade retention among high-school students in 2011, consider the following differences-in-differences regression:

$$y_{hst} \quad = \quad \beta_t \times (G_{hs} \times T_t) + \zeta_{hs} + \lambda_t + \varepsilon_{hst} \tag{1.12}$$

where y_{hst} is retention of students in grade h of school s in year t, with h either students in 1st-4th grade (non-participants) or students in 9-12th grade (participants). The indicator G_{sh} is equal to one for grades 9-12th and zero otherwise, T_t is a vector of indicator variables for years $t = 2002, \ldots, 2015$ (with 2010 as the omitted category), ζ_{hs} and λ_t are school-grade and year fixed effects, and ε_{hst} is an error term correlated within schools. An increase in grade retention among high-school students in 2011 translates into $\beta_{2011} > \beta_t$, with $t \neq 2011$.

Figures 1.8-A and 1.8-B present estimated coefficients $\widehat{\beta}_t$ using OLS. Figure 1.8-A uses absenteeism as dependent variable and Figure 1.8-B uses grade retention. Absenteeism among high-school students increased by 4.5 percentage points in annual official statistics, a 60 percent increase from a base of 0.08 absenteeism in 2010.[22] Retention among high-school students increased by 3.5 percentage points in 2011, a 60 percent increase from a base of 0.06 in 2010, an unprecedented increase in the period 2002–2015.

Individual analysis. To estimate the individual costs of participation consider a version of equation (1.7) that includes school fixed effects and uses academic performance at the end of 2011 as dependent variable. The coefficients of interest are again flexible estimates of social network absenteeism in June 16. Figures 1.8-C and 1.8-D present estimates using GPA and grade retention as dependent variables. Estimated coefficients imply that full absenteeism in social networks in June 16 is associated with (1) a decrease of 0.16 standard deviations in academic performance, and (2) an increase in grade retention of 38 percent (from a base retention of 0.06 in 2010).

Now consider a similar regression but using individual participation as independent variable. Estimated coefficients suggest that individual school absenteeism in June 16 leads to (1) a decrease of 0.1 standard deviations in GPA (coefficient of -0.07, p-value of 0.00), and (2) an increase in grade retention of 33 percent (coefficient of 0.02, p-value of 0.00). Results using annual participation, defined in section 1.4.2, imply that a one standard deviation increase in participation decreases GPA by 0.15 standard deviations and increases grade retention by 31 percent.

[22] This increase in absenteeism needs to be interpreted with caution as both the denominator and the numerator are changing. The central government decreased the total number of official days of schooling in 2011, mechanically decreasing an otherwise larger increase in absenteeism.

1.7.2 Political effects of the student movement

The first election after the rise of the student movement of 2011 was held in October of 2012.[23] At these local elections citizens elected mayors in all of 345 counties in the country. Traditional parties, organized in left and right wing coalitions, competed against each other and against candidates from "non-traditional" parties. Although with new leaders and lower participation rates, the student movement was still active and many anticipated it would have an effect on electoral outcomes. The movement showed its discontent with traditional politics and publicly supported non-traditional parties.[24] Despite its contemporary relevance, there is no systematic evidence of the impact the student movement had at these elections.

To estimate the effect of the student movement in the 2012 local elections, consider the regression:

$$V_{c,2012} = \alpha + \beta \cdot \text{Student Movement}_{c,2011} + \gamma V_{c,2008} + \varepsilon_c \qquad (1.13)$$

where $V_{c,2012}$ and $V_{c,2008}$ are electoral outcomes in the 2012 and 2008 local elections in county c, Student Movement$_{c,2011}$ is the county-level average participation of high-school students in the movement (see section 1.4.2), and ε_c is an robust error term. The dependent variable is either the vote share for non-traditional candidates or the percentage of voters in the county population.[25]

The main concern with an OLS estimation of β is the potential existence of omitted variables correlated with the student movement and electoral outcomes. Three exercises suggest this is unlikely to be a major concern. First, regressions control for electoral outcomes in previous elections, which captures cross-sectional variation in political preferences. Second, placebo checks using school absenteeism in previous years support results. Third, I use the method proposed by Altonji et al. (2005) to construct bounds for estimates and qualitative conclusions remain.

Table 1.6 presents regression estimates. Column 2 indicates that a one standard deviation increase in the intensity of the student movement is associated with an increase of 5.3 percentage points in the vote share for non-traditional candidates, an increase of 16 percent on a base of 33 percent in 2008. This increase in vote shares is mostly explained from a decrease in vote shares for right-wing candidates, the coalition of the incumbent president. Column 5 suggest that the same increase in the movement intensity is associated with a decrease of 1.5 percentage points in voters. More speculatively, columns

[23]There was an informal plebiscite organized by citizens in October of 2011. Figure A.4 shows that participation was higher and people agreed more with students' demands in counties with higher participation in the movement.

[24]One popular election involved the non-traditional (independent) candidate Josefa Errázuriz – explicitly supported by the student movement – competing against the traditional (right-wing) candidate Cristián Labbé, mayor of *Providencia* county between 1996 and 2012. Errázuriz won that election.

[25]Electoral outcomes are based on official data reported by the Electoral Service of Chile. Population data come from censuses. Figure A.11 plots the student movement variable for all counties.

3 and 6 provide some suggestive evidence of non-linear effects that are consistent with the previous "critical mass" patterns (see Figure A.12). As placebo checks, I create fake local movements using the differential increase in county-level school absenteeism between 2009 and 2010, i.e. before the rise of student movement. Reassuringly, this "fake movement" does not have an effect on electoral outcomes in the 2012 elections.

Local elections are a natural setting to use the Altonji et al. (2005) method to study a potential bias due to unobservable variables because past electoral outcomes are powerful predictors of outcomes at the county level. Oster (2016) emphasizes that changes in the R-squared from an uncontrolled to a controlled regression can be used to obtain an adjusted coefficient that accounts for unobservables. This "coefficient stability approach" confirms previous results and suggests the effect of the movement is in the range $[0.053, 0.107]$ and the effect on voters is in the range $[-0.015, -0.011]$.[26]

1.7.3 Aggregate network effects

What is the aggregate contribution of networks to the observed daily protest activity? To answer this question, I estimate the difference in students' choice probabilities of skipping school with and without network effects. More precisely, I estimate the following counterfactual difference:

$$\Pr\left(A_i = 1 | A_{j(i)} = a_{j(i)}\right) \quad - \quad \Pr\left(A_i = 1 | A_{j(i)} = 0\right) \tag{1.14}$$

where $A_{j(i)}$ again represents absenteeism in students' social networks. To estimate the choice probabilities, I use a control function approach and a probit model. The exact specification includes student, network, school controls, and city fixed effects. In addition, I compute the difference in equation (1.14) in two different ways: assuming linear network effects and allowing for non-linearities. The difference is informative about the importance of non-linearities.[27]

The model with linear network effects predicts that 55 percent of students will skip school in a protest day. This percentage is equivalent to 440,000 students (0.55×800,000) or 360,000 additional absent students (440,000−80,000) when compared to a regular school day. Now, in a counterfactual scenario without network effects, the model predicts than only 19 percent of students would skip school. These estimates imply that 78 percent of students that skipped school on a protest day would have attended school if there were no network effects ($1 - 80,000/360,000 = 0.78$). In contrast, the model with non-linearities predicts that 61 percent of students will skip school in a protest day. This number decreases to 45 percent when shutting down networks, implying that 31 percent

[26]Bounds use $\widehat{\beta} = \beta_c - (\beta_{nc} - \beta_c)\frac{R_{max} - R_c}{R_c - R_{nc}}$, where β_c and β_{nc} are coefficients from a regression with and without controls with corresponding R-squared of R_c and R_{nc}, and R_{max} is an unknown parameter in the interval $[R_c, 1]$. I use the conservative assumption of $R_{max} = 1$. See Oster (2016) for details.

[27]This calculation is similar to the one made in Yanagizawa-Drott (2014). Figure A.13 presents aggregate network effects for different numbers of initial protesters using estimated coefficients in section 1.5 and 1.7.2.

of students skipping school on a protest day would have attended school in the absence of network effects ($1 - 280,000/408,000 = 0.31$).

Networks amplify protest activity significantly. However, the difference between linear and non-linear network effects is remarkable. Networks explain more than two-thirds of aggregate participation when effects are "constrained" to be linear (78 percent), and less than one-third when effects are non-linear (31 percent). A simple explanation for this pattern is found in Figure 1.4-B, where I plot the distribution of absenteeism in social networks. In the linear case, all students are influenced by their networks, regardless of network absenteeism. In the non-linear case, only students exposed to higher than 40 percent network absenteeism experience changes in their choice probabilities. The aggregate effect of networks captures this differential exposure.

1.8 Conclusion

The individual decision to participate in a social movement is a crucial component behind the rise of groups demanding institutional change. Studying the Chilean student movement of 2011, this paper shows that students were influenced by their networks to participate in the movement only when a "critical mass" of 40 percent of their networks also participated. Overall, results support the popular idea of a tipping point in behavior (Gladwell, 2000) and the importance of strong ties to promote political activism (McAdam, 1986).

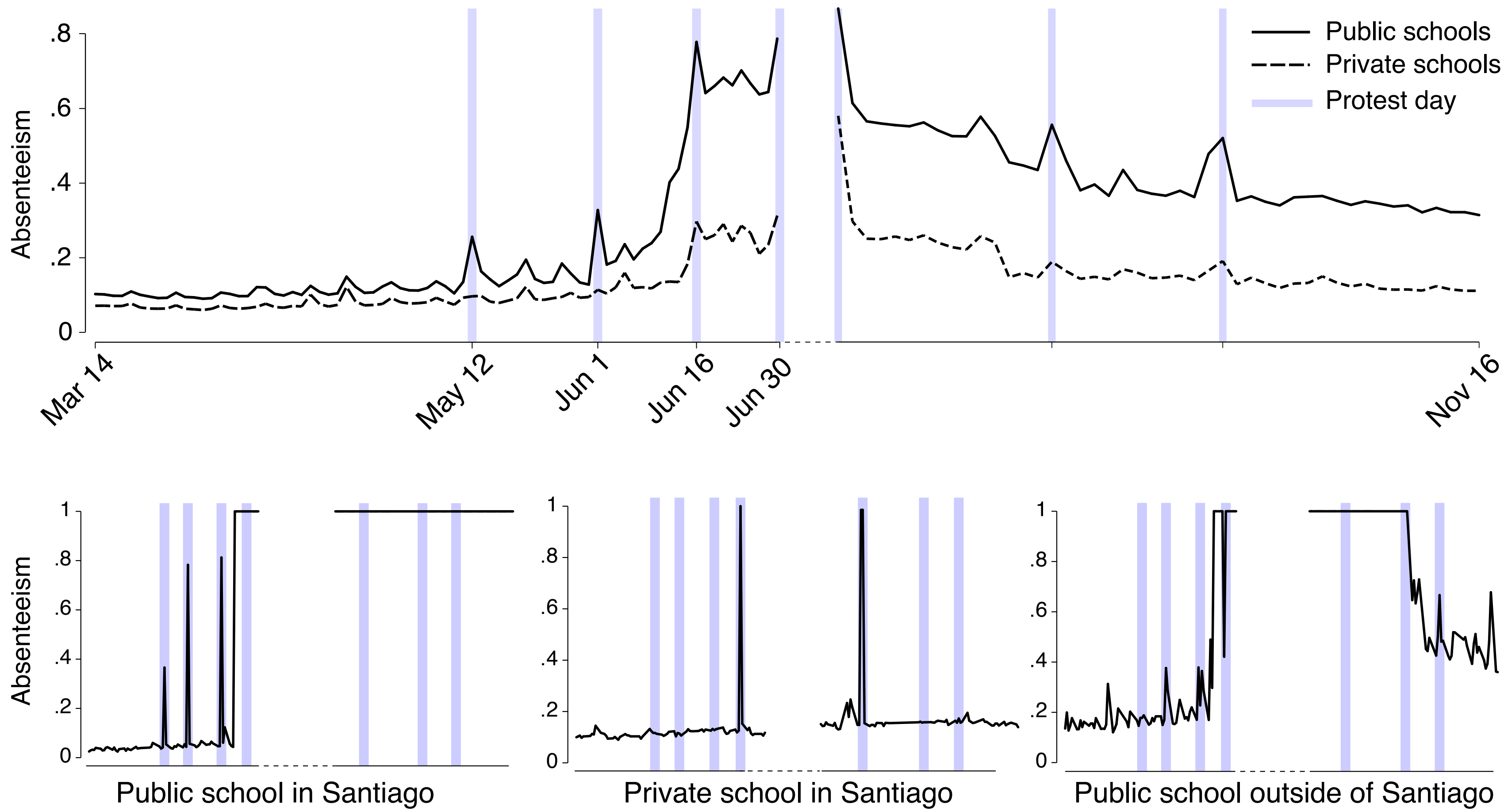
The findings in this paper have at least two implications. First, results are relevant for the modeling of collective action in networks. Theoretical work has emphasized that protest participation may be modeled as a game of strategic complements or strategic substitutes. The "critical mass" type of influence found in this paper suggests that complementarities are relevant for at least some levels of participation. Results also point towards the possibility of protest participation as strategic substitute – i.e. individuals free riding on the participation of others – but only for large values of participation in network groups.

Second, complementarities in protest behavior imply that individuals with larger networks are more influential. This corollary is potentially extremely important for both the organization of a social movement and its disruption. For example, imagine a group of individuals organizing a social movement to bring down a dictatorship, as the Otpor! movement in Serbia in the 1990s. The findings presented in this paper suggest that the marginal return of enrolling one additional citizen in the movement is higher for individuals with larger networks. In addition, an organization may exploit the "critical mass" patterns by exerting effort to go beyond the threshold. In the same way, a state could decrease participation in a social movement by preventing central individuals to participate or by exerting effort to avoid reaching a "critical mass."

Two additional remarks are necessary to interpret results more broadly. Firstly, students may be subject to more or less influence from their networks than the non-student population. More than a concern – after all many important movements have been started by students – the setting may restrict the external validity of results to interpret social movements originating in non-student populations. In the second place, the lack of a precise identification of the mechanisms behind the results may also hinder their external validity. The lack of emphasis on beliefs about the actions of others and the missing dynamics in social networks also prevent us from a full understanding of the decision to participation in a social movement. Nevertheless, this paper is still a clear step forward in the study of social movements.

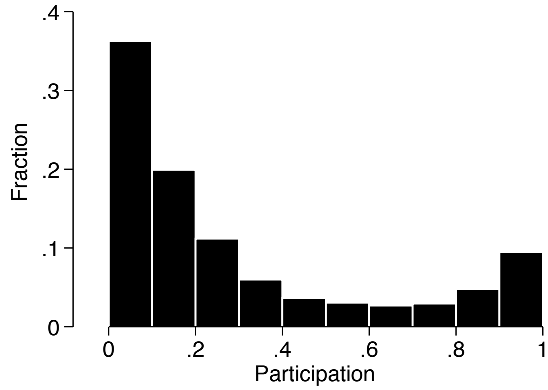
Future empirical studies of social movements may explore how protests create network links between participants and what are the consequences of this, how police violence in protests disrupt (or foster) participation, and how habit formation contributes to the escalation of a mobilization. For now, we have evidence that networks amplify the effect of protests in non-linear ways with potentially significant consequences for institutional change.

Figure 1.1: Absenteeism of high-school students in 2011

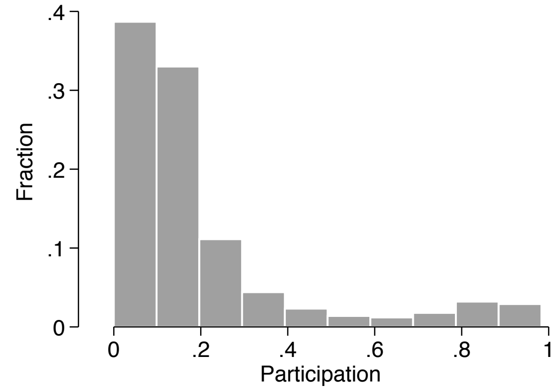


Notes: Own construction using administrative data. High-school students are students enrolled in grades 9-12. The y-axis is average school absenteeism among high-school students (in percentages) and the x-axis represents days in 2011. Vertical lines denote national protest days during week days. The gap in the center of the figures corresponds to winter break. More details about data in section 1.4.

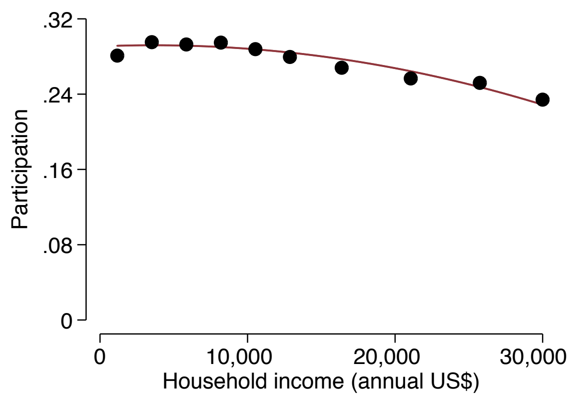
Figure 1.2: Participation in the student movement



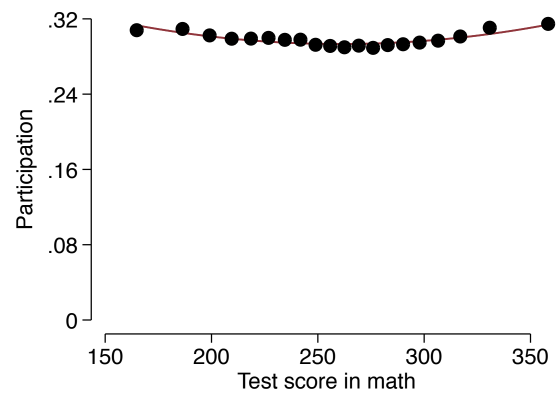
(a) Students (N=836,988)



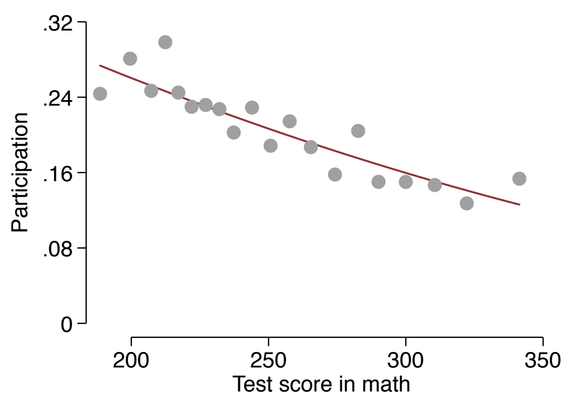
(b) Schools (N=2,590)



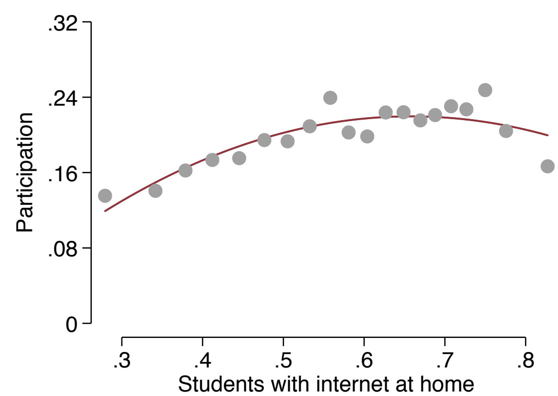
(c) Students' household income (N=481,998)



(d) Students' test scores (N=326,820)



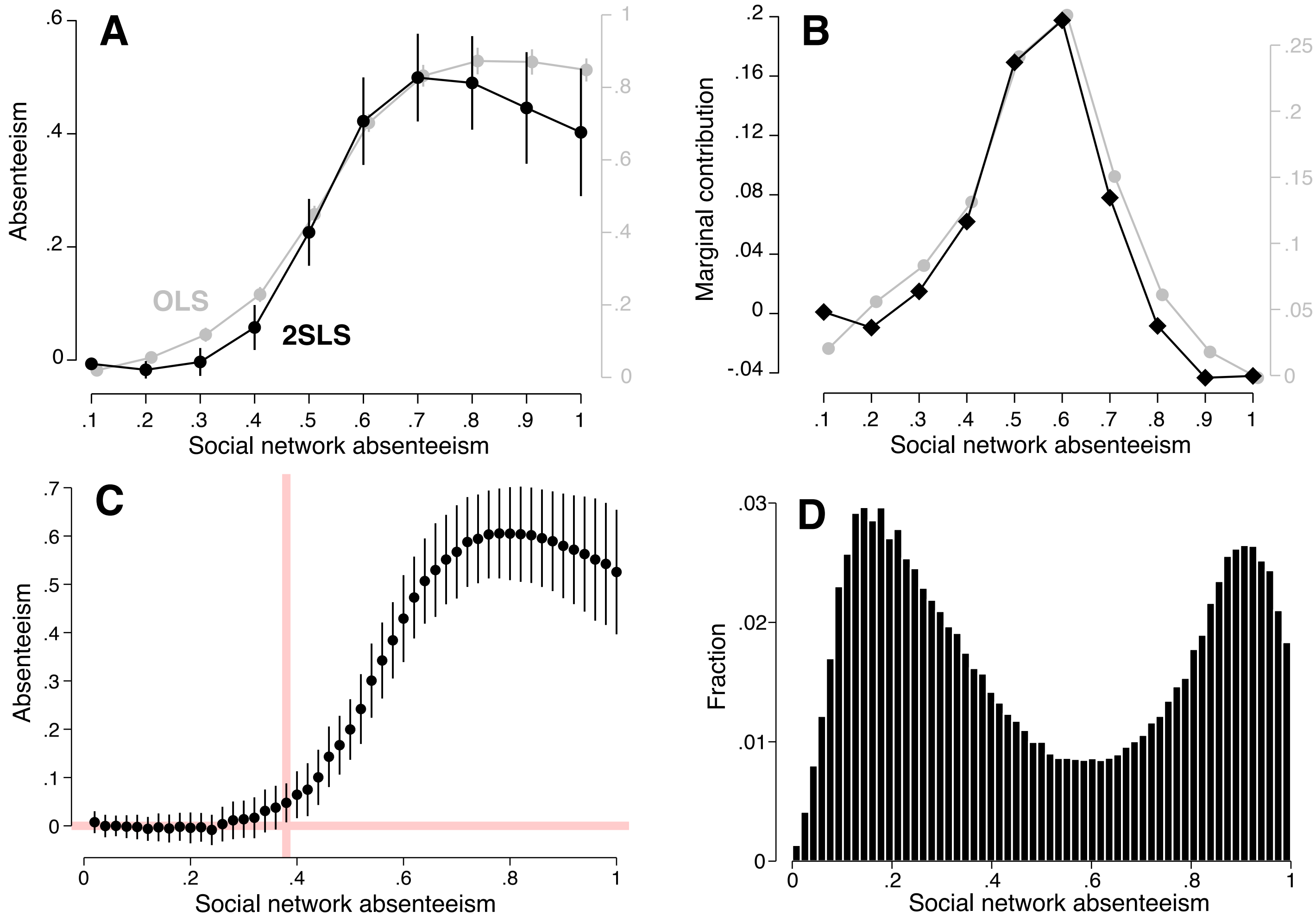
(e) Schools' test scores (N=2,428)



(f) Schools' internet (N=2,589)

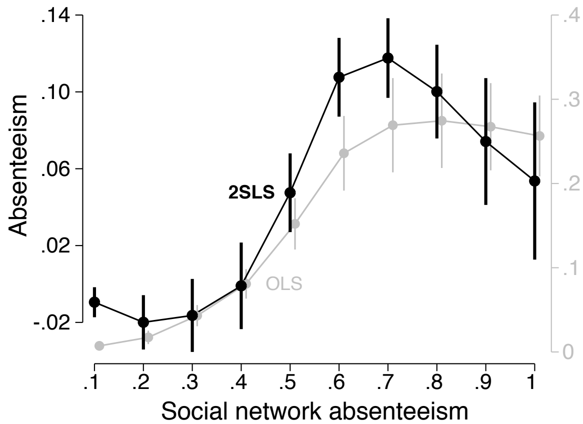
NOTES: *Participation* is defined as additional school absenteeism after the beginning of the student movement. Red lines represent quadratic fits. School variables are averages of high-school students. Tests scores are measured with standardized tests. Household income and internet at home are constructed from household surveys. More details about data in section 1.4.

Figure 1.3: Critical mass – main results

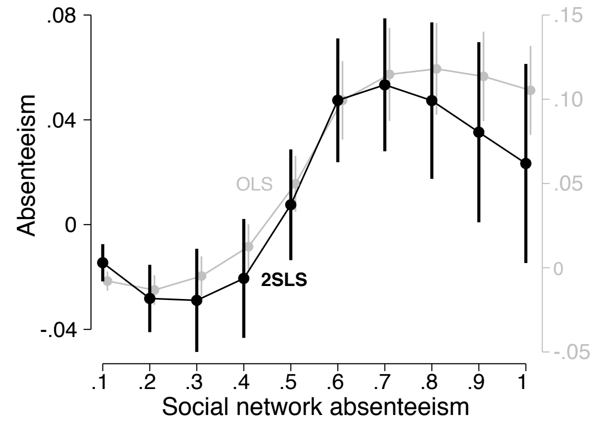


Notes: Panel A plots estimates (OLS and 2SLS) from a regression of individual school absenteeism on 10 indicators of network absenteeism in June 16, controlling for school, network, and school characteristics, and city fixed effects. Vertical lines denote 95 percent confidence intervals (s.e. clustered at the city level). Panel B plots the difference in the estimated coefficients in Panel A. Panel C plots the same coefficients that in Panel A, but using 51 (instead of 10) indicators of network absenteeism. Panel D plots the distribution of social network absenteeism. More details in section 1.5.4.

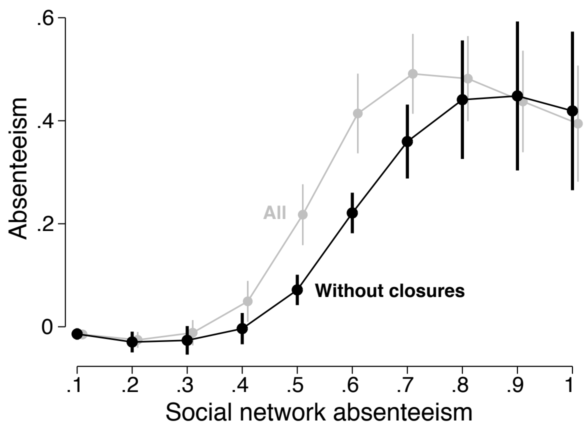
Figure 1.4: Critical mass – additional results



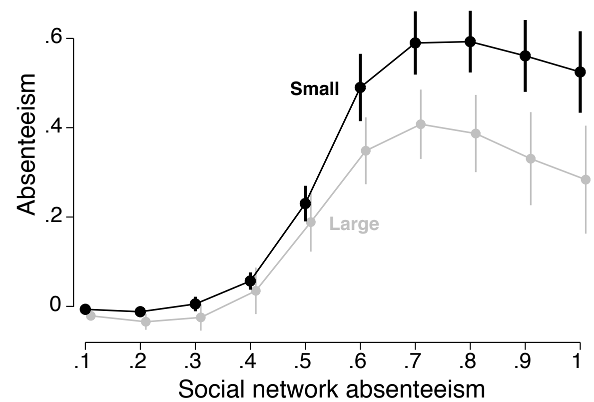
(a) Including school F.E.



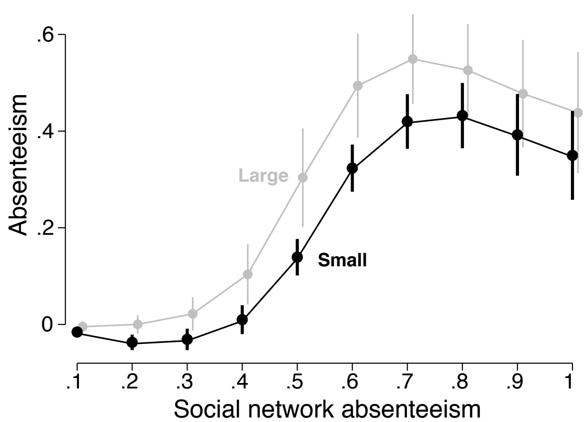
(b) Including school-grade F.E.



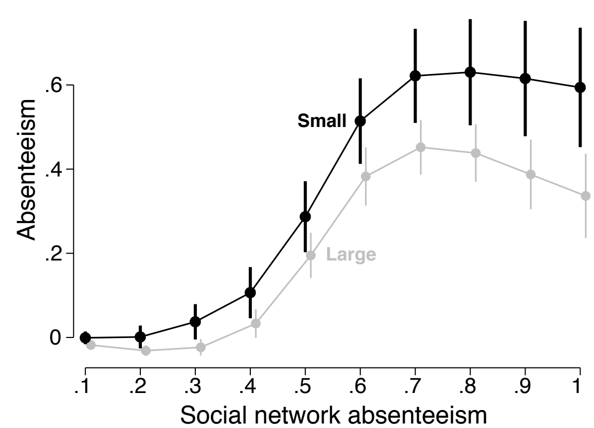
(c) School closures



(d) Network size



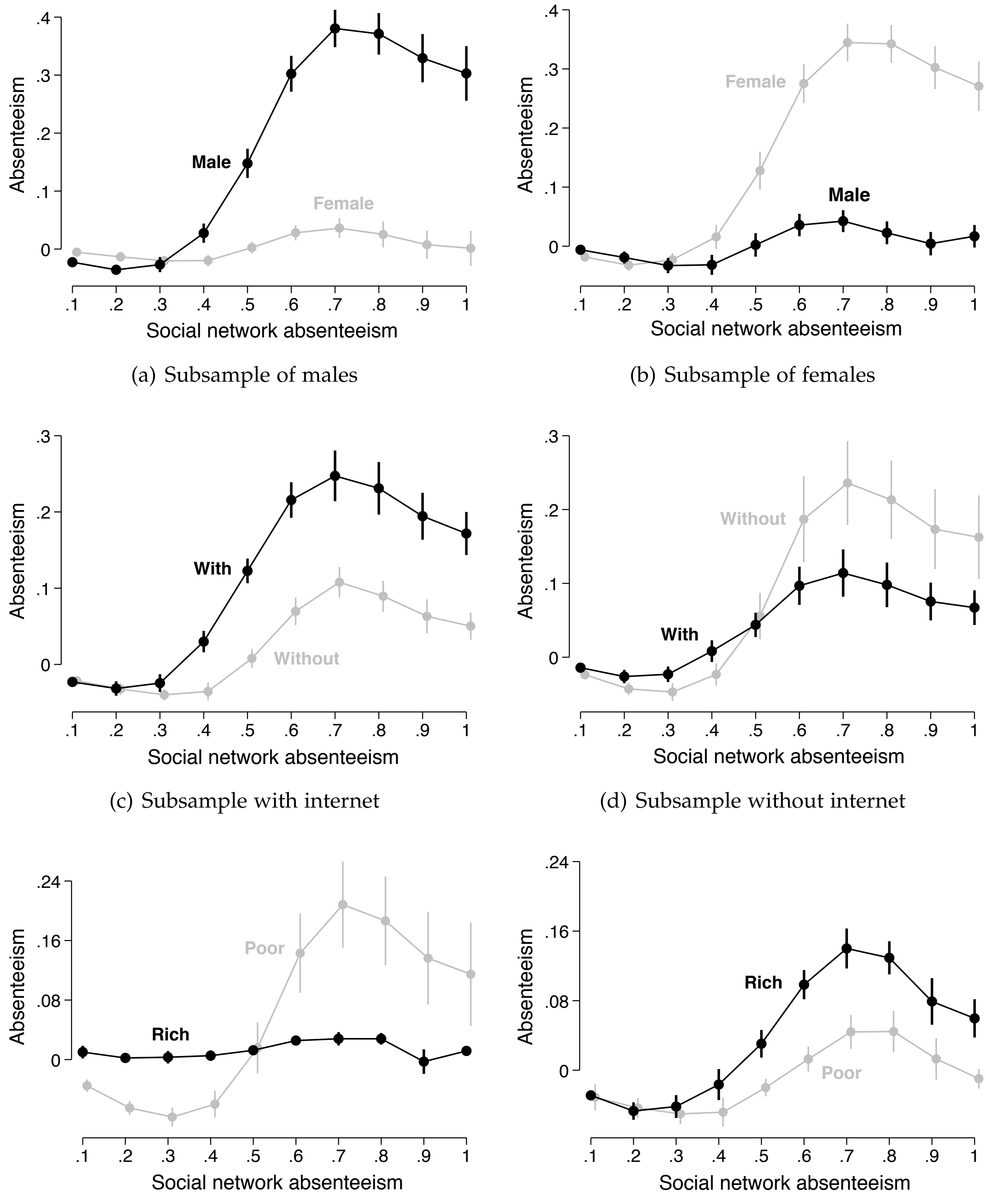
(e) School size



(f) City size

Notes: Panels A and B plot OLS and 2SLS estimates from a regression of individual school absenteeism on 10 indicators of network absenteeism, controlling for school, network, and school characteristics. Panels C-F present 2SLS estimates in sub-samples. More details in section 1.5.4.

Figure 1.5: Differential influence within networks



(a) Subsample of males

(b) Subsample of females

(c) Subsample with internet

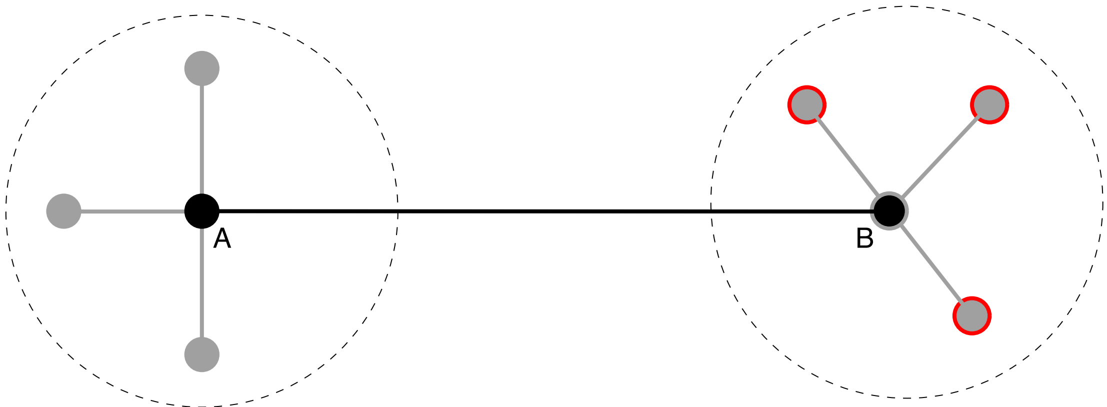
(d) Subsample without internet

(e) Subsample from poor households

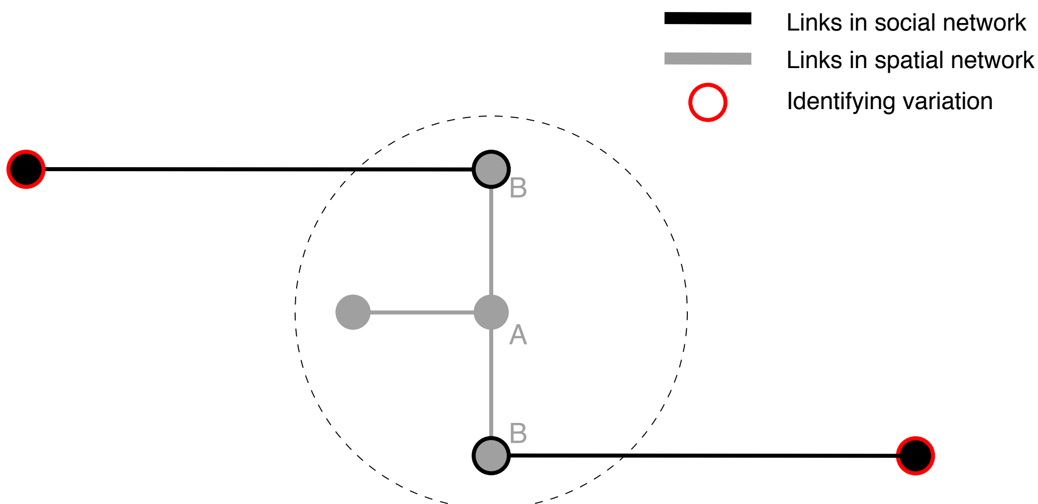
(f) Subsample from rich households

Notes: All panels plot 2SLS estimates from a regression of individual school absenteeism on 10 indicators of network absenteeism, controlling for school, network, and school characteristics, and city fixed effects. Regressions are in sub-samples and split the network in groups. More details in section 1.5.5.

Figure 1.6: Multinetwork instruments



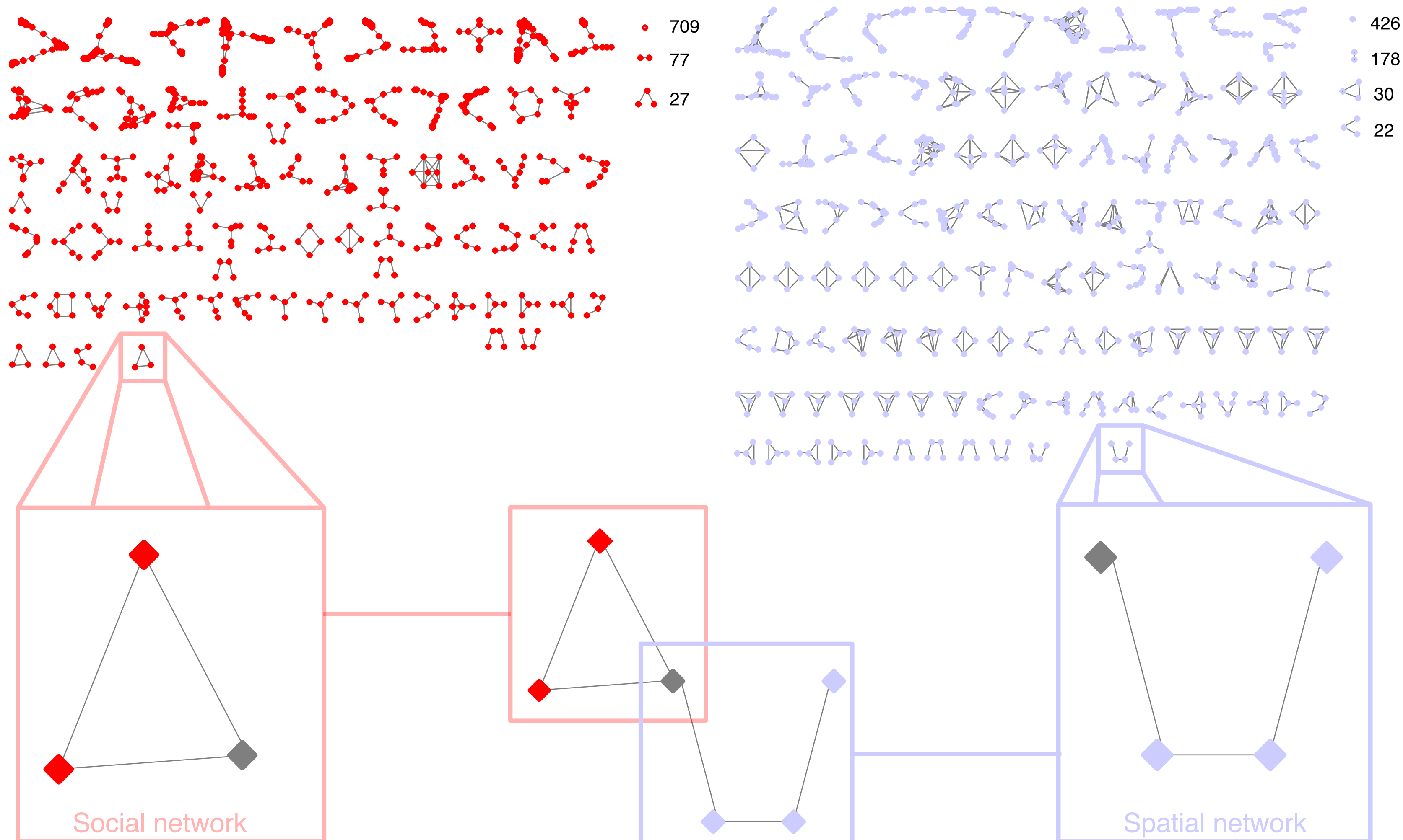
(a) Spatial network of B for social network of A



(b) Social network of B for spatial network of A

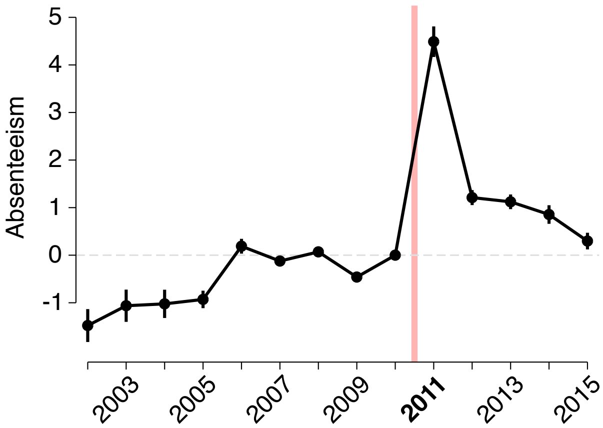
Notes: This figure provides intuition for the multinetwork instruments to identify the causal effect of node B on node A in social networks (Panel A) and spatial networks (Panel B). Each circle represents a node (e.g. school) and each line represents a link in the social network (dark lines) or spatial network (gray lines). The dash circles define the area of spatial networks. For example, two nodes are linked if these are closer than 1 kilometer. The identifying variation is marked in red. Identification in this strategy relies on "cross network" exposure. The additional dimension of time is missing from this figure. More details in section 1.6.2.

Figure 1.7: Multinetworks in the data

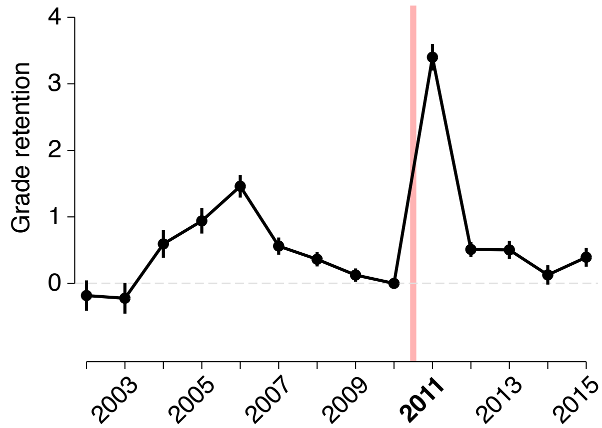


Notes: This figure presents a visualization of high-schools in social networks (left-hand side), and spatial networks (right-hand side). Each node represents a high-school and lines represent social or spatial links. The legend with numbers in the top right of each network represents the number of isolated schools in those types of sub-networks, e.g. there are 709 schools without social links and 426 without spatial links.

Figure 1.8: The cost of participation

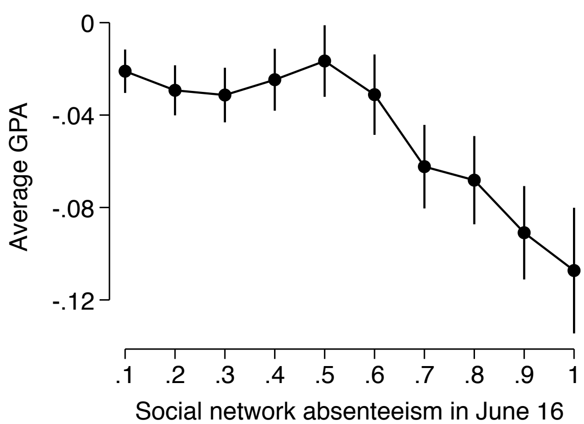


(a) Absenteeism 2002–2015

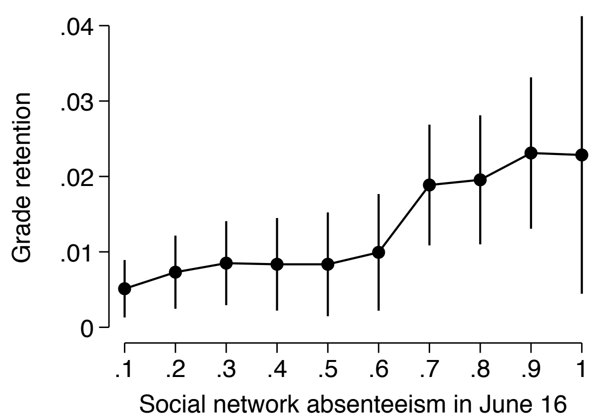


(b) Grade retention 2002–2015

Notes: Panels A and B plot differences-in-differences estimates of absenteeism/retention rates between high-school students (participants in the movement) and students age 6–10 (non-participants) in the period 2002-2015. Vertical lines denote 95 percent confidence intervals and standard errors are clustered at the school level. The omitted category is 2010. In both figures the y-axis is measured in percentage points. More details in section 1.7.1.



(c) GPA (N =755,874)



(d) Retention ($N = 780,306$)

Notes: Panels C and D plot OLS estimates from a regression of academic performance on social network absenteeism in June 16, controlling for student controls, network controls, and school fixed effects. Vertical lines denote 95 percent confidence intervals and standard errors are clustered at the school level. The y-axis in Panel C is GPA. The standard deviation of GPA is 0.67. The y-axis in Panel D is in percentages. More details in section 1.7.1.

Table 1.1: Descriptive statistics

	Mean	St. Dev.	Observations
Students			
School absenteeism:			
May 12, 2011	0.15	0.36	760,801
June 1, 2011	0.19	0.39	
June 16, 2011	0.49	0.50	
Average in 2010	0.07	0.07	
Repeated grade in 2010	0.06	0.23	
GPA in 2010	5.40	0.59	
Switched school after 2010	0.24	0.42	
Female	0.51	0.50	
Age	15.8	1.3	
Household income (annual US$)	7,891	7,892	481,998
Internet connection at home	0.55	0.50	304,448
Schools			
Average test score in standardize test	250	40	2,224
Share of students who repeated grade	0.06	0.05	
Average household income (annual US$)	8,877	5,227	
Public	0.30	0.46	
Students	342	325	
Cities			
High-schools in the city	7.7	44.3	290
High-school students in the city	2,623	16,134	

NOTES: Own construction based on administrative data provided by the Ministry of Education. All variables are measured in 2011 unless otherwise stated. The number of observations is the same as in the first row of each panel unless otherwise stated. More details in section 1.4.

Table 1.2: Linear estimates

Dependent variable is absenteeism in June 16, 2011

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A – OLS estimates						
Network absenteeism in June 16	1.23***	1.22***	1.27***	1.21***	1.24***	1.47***
	(0.02)	(0.02)	(0.02)	(0.04)	(0.05)	(0.03)
Absenteeism in May 12	0.06***	0.05***	0.05***	0.03***	0.04***	0.04***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)
Absenteeism in June 1	0.08***	0.08***	0.07***	0.05***	0.06***	0.06***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Panel B – 2SLS estimates						
Network absenteeism in June 16	0.80***	0.77***	0.69***	0.81***	0.69***	0.63***
	(0.05)	(0.05)	(0.08)	(0.07)	(0.10)	(0.21)
Absenteeism in May 12	0.12***	0.10***	0.09***	0.05***	0.05***	0.06***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Absenteeism in June 1	0.16***	0.15***	0.14***	0.08***	0.08***	0.12***
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)
Student controls		x	x	x	x	x
Network controls			x	x	x	x
School controls				x	x	x
City F.E.					x	
Neighborhood F.E.						x
F-stat 1*st* stage	53.3	50.5	30.6	36.0	24.1	14.0
R-squared (Panel A)	0.626	0.629	0.638	0.645	0.652	0.583
Observations	779,327	779,251	771,121	760,801	760,801	49,273

NOTES: *Student controls* include academic performance, average school attendance in previous years and socioeconomic characteristics. *Network controls* include average student controls at the network level. *School controls* include indicators for publicly managed schools, average academic performance of students, and average socioeconomic characteristics of students. *Neighborhoods* are geographic areas where students live. More details in section 1.5. See Figure A.5 for a map of cities. In column 6, each neighborhood is of size 10×10 blocks. Neighborhood data is only available for some students. See Figure A.7 for a map of neighborhoods. Standard errors clustered at the city level are reported in parentheses. Significance level: *** $p < 0.01$.

Table 1.3: Linear estimates – differential influence within networks

Dependent variable is absenteeism in June 16, 2011

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Gender		Internet connection		Household income		
Social network in June 16	Male	Female	Yes	No	Poor	Middle	Rich
Males	0.44***	0.03					
	(0.05)	(0.02)					
Females	0.03	0.42***					
	(0.03)	(0.04)					
Students with internet			0.27***	0.11***			
			(0.04)	(0.02)			
Students without internet			0.08***	0.22***			
			(0.02)	(0.02)			
Students from poor households					0.15***	0.06***	-0.01
					(0.02)	(0.01)	(0.01)
Students from middle income households					0.06***	0.14***	0.11***
					(0.02)	(0.03)	(0.02)
Students from rich households					0.01***	0.01***	0.09***
					(0.00)	(0.00)	(0.01)
Mean of dep. variable	0.48	0.49	0.42	0.49	0.51	0.45	0.37
Absenteeism previous protests	x	x	x	x	x	x	x
Student, network, and school controls	x	x	x	x	x	x	x
City F.E.	x	x	x	x	x	x	x
Observations	375,737	385,057	229,066	187,034	234,615	257,555	78,760

NOTES: All columns present 2SLS estimates. Regressions are estimated in sub-samples of students and the network is split in groups with similar observable variables. More details in section 1.5.5. Significance level: *** $p < 0.01$, ** $p < 0.05$.

Table 1.4: First-stages and reduced forms with a multinetwork identification strategy

	Absenteeism in spatial network June 16			Absenteeism in social network June 16			School absenteeism June 16		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Absenteeism in spatial network of social network in May 12	0.15	0.16**	0.03	0.87***	0.73***	0.68***	0.71***	0.19***	0.15***
	(0.10)	(0.08)	(0.07)	(0.14)	(0.19)	(0.23)	(0.05)	(0.07)	(0.05)
Absenteeism in social network of spatial network in May 12	0.82***	0.78***	0.60***	-0.02	-0.04	-0.18	0.31***	0.23***	0.12**
	(0.07)	(0.09)	(0.05)	(0.061)	(0.05)	(0.11)	(0.07)	(0.06)	(0.05)
Controls		x	x		x	x		x	x
City F.E.			x			x			x
Angrist-Pischke F-test	137.0	70.4	148.0	37.7	16.4	9.2	–	–	–
Cragg-Donald (F-stat)	98.2	67.9	48.2	98.2	67.9	48.2	–	–	–
R-squared	0.19	0.20	0.40	0.11	0.17	0.32	0.09	0.43	0.56

Notes: Total number of schools is 2,070 and regressions are weighted by the number of students. Standard errors clustered at the city level are reported in parentheses. More details in section 1.6. Significance level: *** $p < 0.01$, ** $p < 0.05$.

Table 1.5: Collective action in multinetworks

Dependent variable is school absenteeism in June 16, 2011

	OLS			2SLS		
	(1)	(2)	(3)	(4)	(5)	(6)
Spatial network absenteeism in June 16	0.13**	0.18***	0.01	0.39***	0.31***	0.27***
	(0.06)	(0.03)	(0.06)	(0.09)	(0.06)	(0.07)
Social network absenteeism in June 16	0.33***	0.10***	0.04	0.75***	0.19*	0.20*
	(0.05)	(0.02)	(0.03)	(0.16)	(0.10)	(0.11)
School absenteeism in June 1		0.17**	0.16***		0.13*	0.13**
		(0.07)	(0.06)		(0.07)	(0.06)
School absenteeism in May 12		0.08	0.09*		0.05	0.05
		(0.06)	(0.05)		(0.06)	(0.05)
Controls		x	x		x	x
City F.E.			x			x
R-squared	0.106	0.446	0.553	–	–	–

Notes: Total number of schools is 2,070 and regressions are weighted by the number of students. Statistical tests and first-stages are presented in Table 1.4. More details in section 1.6. Standard errors clustered at the city level are reported in parentheses (320 clusters). Significance level: *** $p < 0.01$, ** $p < 0.05$.

Table 1.6: The political effects of the student movement

Dependent variables are electoral outcomes in the 2012 local elections

	Vote share non-traditional parties			Voters in population		
	(1)	(2)	(3)	(4)	(5)	(6)
Main estimates						
Student movement	0.025	0.053**	0.026	-0.031***	-0.015***	-0.039***
	(0.029)	(0.022)	(0.034)	(0.008)	(0.008)	(0.006)
Student movement squared			0.016			0.015***
			(0.014)			(0.003)
Placebo						
Δ school absenteeism 2010-2009	-0.013	-0.002	-0.003	-0.022	-0.002	0.001
	(0.026)	(0.021)	(0.019)	(0.016)	(0.005)	(0.004)
Δ school absenteeism 2010-2009 squared			-0.001			0.003
			(0.005)			(0.002)
Dep. variable in 2008 election		x	x		x	x
Student movement (*p*-value)			0.02			0.00
Placebo (*p*-value)			0.98			0.27
Mean dep. variable	0.35	0.35	0.35	0.49	0.49	0.49
R-squared	0.006	0.347	0.351	0.057	0.832	0.849
Counties	345	345	345	345	345	345

NOTES: Regressions are weighted by the total number of voters in 2008. *Student movement* has been standardized to facilitate the interpretation of coefficients. Non-traditional parties correspond to parties that are different from the coalition "Concertación" and the right wing coalition, i.e. independent candidates. The coefficients for *Placebo* estimates come from separate regressions. The "Student movement (p-value)" and "Placebo (p-value)" in the bottom of the table correspond to p-values for the test that linear and quadratic terms are different from zero. Robust standard errors are reported in parentheses. Significance level: *** $p < 0.01$, ** $p < 0.05$.

Chapter 2

Distorted Quality Signals

2.1 Introduction

Information plays a key role in consumer choices. In education, information on school quality is often measured via standardized tests. The use of these tests to assess school performance has became common in recent decades (Figlio and Loeb, 2011). However, accountability systems that resort to these tests have been controversial among academics and educators. Critics argue that high-stakes testing might generate undesirable behavioral responses that introduce distortions in the performance metric itself and thus the accomplishment of its goals (Neal, 2013). This argument stems from Holmstrom and Milgrom (1991), who underscore the role of hidden actions in producing changes in the observed outcome (i.e. test scores) without necessarily improving the real outcome of interest (i.e. learning). Despite increasing evidence of undesirable behavioral responses, quantification of these potential distortions and their consequences is lacking.

How large are these distortions in school quality signals? What are the market consequences of these distortions? We study one of the most developed accountability systems in the world –Chile's market-oriented educational system (Figlio and Loeb, 2011)– and show that behavioral responses are in place and distort key performance metrics. Distortions are large and have significant consequences on school choice and the allocation of public programs. The government relies on standardized testing to generate school-specific quality metrics. These measures are used not only for quality assessment and performance evaluation, but also as a disclosure system in school choice. These features make Chile an ideal setting to quantify the consequences of behavioral responses to accountability systems.

The analysis proceeds in four steps. First, we show that low-performing students are more likely to be absent on test days relative to other students. Using national administrative data on Chilean school children, we compare daily attendance of test takers (fourth graders) and non-takers (third graders) within schools on test and non-test days. High performing test-takers increase their attendance on test days by 0.18

standard deviations more than low performing test-takers. This result suggests that a behavioral response to standardized testing is at work. However, the degree of student non-representativeness varies considerably across schools.

Second, we use a multiple imputation method to predict the test scores of absent students and thus the associated distortions in school quality signals.[1] We find average distortions in the system to be sizable: 0.1 standard deviations of school test scores. Distortions vary widely across schools, but are persistent within schools over time. To better understand these distortions, we construct a panel dataset for all Chilean schools during the period 2005–2013. Public, low quality, and for-profit schools display larger average distortions. Notably, we find some evidence of larger distortions in more competitive markets. In particular, schools facing more quality-elastic consumers display larger distortions in quality signals.[2] In contrast, we find no evidence for potential perverse incentives by teacher performance programs in place.

Third, we estimate a school choice model to quantify the implications of these distortions. We find that providing undistorted school quality information would likely induce three percent of students to switch schools. To estimate the model, we use geocoded addresses of 100,000 students and 1,500 schools, and estimate a discrete choice model in which households trade-off school quality and distance. For identification, we exploit quasi-experimental variation in government programs, climate-induced variation in test scores, and fixed characteristics of competitors. Given the magnitude of distortions and the spatial distribution of schools, the trade-off between distance to school and quality explains the student switching rate among schools. The model suggests that households that would change their choices are willing to pay 117 U.S. dollars annually for undistorted quality information, with non-poor households willing to pay more than poor households due to differences in preferences.

Fourth, we show that two large public programs are significantly misallocated because of distortions. In the first program, the government assigns bonuses to teachers in schools with sufficiently high average test scores. We reallocate bonuses based on removing distortions, and find that 13 percent of resources are misallocated each year, equivalent to $20 million in the last twenty years. In the second program, the government used test scores to classify schools in three quality categories and delivered this information to parents with the objective of assisting school choice. Using the classification algorithm, we show that four percent of schools were incorrectly classified and these errors persuaded two percent of the incoming student cohort to choose a different school.

This paper makes three main contributions. First, we document a novel channel

[1] Multiple imputation methods are routinely used in the Survey of Consumer Finances conducted by the Federal Reserve in the U.S., and in the Household Financial Survey conducted by the Central Bank of Chile, among many others (Kennickell, 1998; Alfaro and Fuenzalida, 2009).

[2] Recent studies have suggested a link between competitive environments and cheating behavior (e.g. Shleifer 2004; Schwieren and Weichselbaumer 2010; Gilpatric 2011; Cartwright and Menezes 2014). The association we find between distortions and quality elasticity is consistent with the framework proposed by Dorfman and Steiner (1954), who study firms' joint decision of price and quality.

through which school performance measures can get distorted: attendance on test days.[3] This type of behavioral response has *not* been found in the U.S., where the most common response has been the selective assignment of students to special education programs (Jacob, 2005; Rockoff and Turner, 2010; Figlio and Loeb, 2011). Second, we propose and implement a statistical method to quantify the magnitude of the distortions in quality signals that arise from non-representative attendance. Third, and most importantly, we estimate the effect of distortions on school choice and the allocation of public programs and thus quantify the market consequences of these behavioral responses. While we implement our analysis in the Chilean educational market, the implications of it go beyond both Chile and schooling. Multiple markets in which quality is imperfectly observed have quality disclosure systems, many of which may create incentives for undesirable behavioral responses (Dranove and Jin, 2010). Moreover, whenever quality signals generated by the disclosure system feed into consumer and government choices, implications similar to those discussed in this paper might arise. Examples of such settings are when quality information is provided to patients for health provider choice or when hygiene information is provided to consumers for restaurant choice (Dranove et al., 2003; Jin and Leslie, 2003).

This study relates to at least three branches of literature. First, is the literature that documents behavioral responses to high-stakes testing. These responses include diversion of resources, cheating, or manipulation of conditions under which the test is taken.[4] Behavioral responses to incentives placed by standardized testing are not, however, the only source of distortions. Mean reversion and random variation in the conditions under which the test is applied can also create distortions.[5] We provide evidence that non-representative test day attendance is an additional behavioral response to accountability systems and compute the implied distortions in school quality signals.

This paper also contributes to the school choice literature. Several authors have shown that fees, distance between home and school, and school quality are the most relevant attributes for school choice.[6] In addition, another set of studies investigates how information affects school choice, yielding mixed results.[7] Our paper emphasizes the importance of *accurate* information in a context in which consumers are actively

[3] In a concurrent paper, Quezada-Hofflinger and Von Hippel (2017) provide complementary evidence for this channel in the case of Chile.

[4] See Figlio and Getzler (2002), Jacob and Levitt (2003), Jacob (2005), Figlio and Winicki (2005), Reback and Cullen (2006), Neal and Schanzenbach (2010), Apperson et al. (2016), Dee et al. (2016), Diamond and Persson (2016), Deming et al. (2016) and Quezada-Hofflinger and Von Hippel (2017), among others.

[5] See Kane and Staiger (2002), Chay et al. (2005), Graff Zivin et al. (2015), and Ebenstein et al. (2016), among others.

[6] See Gallego and Hernando (2009), Neilson (2013) and Feigenberg (2015) for Chile; Bayer et al. (2007), Hastings et al. (2009) and Walters (2014) for the U.S.; and Bau (2015) and Carneiro et al. (2016) for Pakistan, among others.

[7] See Hastings and Weinstein (2008), Jensen (2010), Cooper et al. (2013), Mizala and Urquiola (2013), Andrabi et al. (2017).

choosing.[8]

Finally, our work is related to the literature in industrial organization studying disclosure and advertising (see Dranove and Jin 2010 and Bagwell 2007 respectively for reviews). As mentioned above, work that analyzes the effects of quality disclosure in educational markets is somewhat limited and has yielded mixed results. Our paper relates to the case in which advertising is *informative*. Moreover, following the distinction proposed by Nelson (1970), the fact that schooling is an *experience good* implies that quality is hardly verifiable ex-ante, further implying that information acquired from advertising might be particularly important. This paper adds to this literature by focusing on educational markets, where there is limited work from an advertising perspective, and by measuring the implications of *deceptive advertising*.

The remainder of the paper is structured as follows. Section 2.2 describes school markets and public programs in Chile. Section 2.3 describes the data and shows that low-performing students are underrepresented on test days. Section 2.4 constructs measures of distortions in quality signals and provides an empirical discussion of their determinants. Section 2.5 estimates a school choice model and studies the welfare implications of distorted quality signals. Section 2.6 shows that two large public programs are misallocated because of non-random attendance on test day. The final section concludes.

2.2 Institutional context

2.2.1 School markets

Our analysis focuses on the Chilean primary school market. After a market-oriented reform was implemented in 1980, education has been provided by a mixture of public, private voucher and non-voucher schools. Students can apply and attend any school in the system, although funding varies across school types. Public schools are fully funded by the government. Private voucher schools are privately managed, although eligible for receiving public funding through vouchers. They are allowed to charge fees to parents in the form of copayments, although vouchers are phased out on the basis of those. Private non-voucher schools are not eligible for public funding.

Over the last three decades, the private sector has steadily increased its market share. In 2013, public schools had 38 percent of all students, while private voucher and non-voucher schools enrolled 54 and 8 percent of students respectively (Ministry of Education, 2013).

[8]Our approach to measure the welfare implications of distorted quality signals distinguishes between choice and experience utility (Bernheim and Rangel, 2009). Recent work on the role of information frictions for insurance choice has adopted this insight (Handel and Kolstad, 2015; Spinnewijn, 2016). We adopt it to study information frictions in school choice.

2.2.2 Public programs

Throughout the paper, we will refer to different public programs that are part of the Chilean educational system. For convenience, we briefly describe them in the remainder of this section, providing detail about the relevant institutional features.

Students in the Chilean educational system are eligible for vouchers. Public funding is provided on a per student basis and is linked to student attendance. However, the amount covered by vouchers depends on the characteristics of both students and schools. The baseline voucher program has been in place since the 1980's reforms. During the period we study, the amount of this voucher has varied across schools according to whether they offer full school shifts (*Jornada Escolar Completa*, JEC). Figure B.1 displays the evolution of the amount covered by vouchers during the years included in our dataset. As it can be noted, the amount paid to schools offering JEC is larger than what other schools receive.

In 2008, the Preferential Educational Voucher (*Subvención Escolar Preferencial*, SEP) was enacted as a complementary voucher targeted towards poor households. Eligibility for this program is determined mostly by household income and poverty status. In particular, households in the lowest third of the income distribution or that participate in the main social program offered by the government (*Chile Solidario*) are eligible for SEP vouchers. All public schools are eligible for SEP vouchers, while private voucher school must subscribe in order to become eligible. Subscribing to the SEP program involves additional commitments by schools including limits to fees they might charge and designing resource management plans. SEP vouchers vary according to two school characteristics, namely the share of their students eligible for the SEP voucher and changes in the school's academic performance. Figure B.1 displays the evolution of SEP vouchers through time since their inception.

The National System of Quality Measurement (*Sistema de Medición de la Calidad de la Educación*, SIMCE) has existed since 1988 and gives national standardized tests on different subjects. Tests are implemented every year at the national level for a subset of grades – see Figure B.2 for the timeline of test implementation. Test scores from SIMCE are comparable across schools and years. Tests are implemented by third party personnel. Moreover, average test scores are publicly disclosed and strongly disseminated at the aggregate school level, but are never made available to the public at the student level. Finally, test scores are never disclosed individually to teachers or students.

The National Performance Evaluation System (*Sistema Nacional de Evaluación de Desempeño*, SNED) is a school performance evaluation system that takes the form of a tournament and provides awards to improved schools. SNED operates as follows: (i) groups of *homogeneous* schools are constructed, within which the contest is implemented; (ii) every two years, an index is computed at the school level, which considers academic performance and improvement and socioeconomic integration among other outcomes; (iii) schools are ranked within their groups according to the value of such index; and

(iv) schools covering 25-35 percent[9] of the total enrollment of each group get a monetary prize equivalent to around 40 percent of a teacher's monthly wage for each teacher in the school. Importantly, across dimensions of the index, SIMCE test scores account for as much as 70 percent of the weight of the components used for its calculation (Contreras and Rau, 2012).

The Educational Traffic Lights program (*Semáforo Educacional*, ETL) was announced in April, 2010 and consisted of sending information to all households about local schools. That information included both test scores and a classification of schools as red, yellow or green according to their test scores, with clear cutoffs determining this outcome. An evaluation of this policy by Allende (2012) that uses the discontinuities in such classification for identification, finds that it effectively impacted school enrollment: households in the margin responded by enrolling more in yellow than red schools and more in green than yellow schools.

2.3 Data and attendance on test days

We use four administrative datasets provided by the Ministry of Education. First, is the record of schools operating between 2005 and 2013, in which we observe school type (public, private-voucher, private non-voucher), enrollment, fees, participation in government programs, and school addresses, which we use to construct markets. Second, we use student records between 2005 and 2013 (approximately 3.5 million per year), in which we observe enrollment (school, grade, classroom) and annual average GPA. Third, we use daily school attendance in 2013 to study heterogeneity in attendance on test days across the distribution of potential SIMCE performance. We argue that such heterogeneity is the source of distortions in quality signals. Finally, we use students' performance at SIMCE test as a measure of observed school quality. We focus on $4th$ graders because they are tested every year in the period 2005–2013 and because all schools offering $4th$ grade also offer $1st$ grade, the most relevant margin for school choice.

The focus on test scores as quality signals is appropriate given their contextual relevance. There is an extensive literature studying test scores and value added as quality measures for accountability systems (Meghir and Rivkin, 2011; Figlio and Loeb, 2011). In Chile, however, media outlets and government authorities use *test scores* as quality signals (McEwan et al., 2008) and survey evidence suggests that parents consider test scores important (Centro de Investigación y Desarrollo de la Educación, 2010). Accordingly, evidence shows that test scores affect school choice (Gallego and Hernando, 2009; Chumacero et al., 2011; Gómez et al., 2012). In addition, the government uses these test scores to guide the allocation of public programs. Figure B.3 shows how test scores are publicly disseminated through media outlets, used for advertising by schools, and used as policy tools by the government.[10]

[9]The coverage of the prize was increased to 35 percent of the enrollment of the group since 2006.

[10]The only measures of value added available for Chile are those computed by Neilson (2013). These

2.3.1 Descriptive statistics

Using the previously described administrative records, we construct two datasets: (1) a panel of schools, and (2) a panel of students. Although the former includes all schools operating in the period 2005–2013, the latter is only available for public and voucher schools in 2013, which represent 93 percent of enrollment that year.

The school level dataset contains annual information on schools offering 4*th* grade in urban areas. The entry and exit of schools makes this panel unbalanced. There are 5,386 different schools and, on average, 4,640 schools operating in a given year. Table 2.1-A presents summary statistics for these schools: 39 percent are public, 52 percent are voucher schools, and 9 percent are private. The average school has approximately 50 students in 4*th* grade. More than half of schools charge no fees, and the average monthly fee is approximately $48.[11] The average test score is 255 and the standard deviation is 27.7.

Table 2.1-B presents descriptive statistics for the student level dataset. Students' academic performance is measured by their GPA, which ranges from 1 to 7, with a threshold of 4 to pass a class. The mean of this variable is 5.9. The last two variables are attendance rates on test and non-test days. The former is simply the average of two indicator variables that take the value of one if a student went to school on test days; recall that there are two test days, so this variable has the value of 0, 0.5, or 1 at the student level. The latter is the average attendance in the five non-test days previous to test ones.

2.3.2 Attendance on test days

Schools average test scores (i.e., quality signals) are distorted if attendance on test days is non-random. We now show how attendance of students changes on test days. Because the government attempts to increase attendance on test days, and schools have incentives to prevent low-performing students taking the test, it is not a priori clear if attendance should increase or decrease on test day. Our interest is not focused on the average change in attendance, but rather in the *heterogeneity* behind this average change, both within and across schools.

In order to estimate the average change in students' attendance on test days, we compare the daily attendance rate of 4*th* graders ($\overline{A}_{4t}$, who take the test) to the daily attendance of 3*rd* graders ($\overline{A}_{3t}$, who do not take the test) around test days in 2013 (October 8*th* and 9*th*):

$$\Delta\overline{A} \;=\; \left(\overline{A}_{4T} - \overline{A}_{3T}\right) \;-\; \left(\overline{A}_{4\tau} - \overline{A}_{3\tau}\right) \tag{2.1}$$

---

value added measures are based on administrative data that is unfortunately unavailable to us. Figure B.4 displays the relationship between that measure of value added and test scores, which is positive and strong.

[11] All monetary units in the paper have been properly deflated and are measured in U.S. dollars using the early 2012 exchange rate.

where $t = T$ represents the two test days, and $t = \tau$ represent other days around test days. We calculate $\Delta\overline{A}$ in four subsamples of students: high-performing, above the 90*th* and 75*th* percentile of the GPA distribution; and low-performing, below the 10*th* and 25*th* percentile of the GPA distribution. In addition, to study the heterogeneity behind $\Delta\overline{A}$, we calculate the following school-specific changes in attendance on test day:

$$\Delta\overline{A}_j \;=\; \left(\overline{A}_{j4T} - \overline{A}_{j3T}\right) \;-\; \left(\overline{A}_{j4\tau} - \overline{A}_{j3\tau}\right) \tag{2.2}$$

where $\overline{A}_{jkt}$ is the average attendance rate of k*th* graders in school j and day t. The next section shows how a larger variance in $\overline{A}_j$ translates into more distorted quality signals.

Figure 2.1 displays $\Delta\overline{A}$ and $\Delta\overline{A}_j$. In panel (a), we plot the differential attendance rate around test days. On average, attendance increases by 2 percentage points on test days. Moreover, the increase is larger among high-performing students than among low-performing students. However, despite the intuition that schools might ask low-ability students to stay home on test days, we do not observe on average a decrease in attendance of students with low GPA. These averages, however, mask significant heterogeneity. In panel (b), we plot the distribution of $\Delta\overline{A}_j$. The vertical line denotes the average increase of 2 percentage points.[12]

These attendance patterns on test days suggest that some behavioral response to the test is in place, in which pools of test-takers are not necessarily representative of their schools. Importantly, the fact that this behavior is heterogeneous and non-random across schools causes observed quality signals in the educational market to be distorted.

2.4 Distortions in quality signals

2.4.1 Estimating undistorted quality signals

Quality signals are *undistorted* if all or a random sample of students take the test. However, the patterns described in section 2.3.2 suggest that absenteeism on test days is not random. The empirical challenge to recover undistorted quality signals consists in estimating test scores for absent students.[13] If we can recover missing test scores, we can estimate undistorted quality signals that would be equivalent to the signals in a world with full or random attendance on test day. Our strategy to estimate missing test scores

[12]In order to assess the robustness of this result, we implemented the same calculations for two alternative class days: one the date of a soccer game between Chile and England and a class day exactly two weeks after test days. Figure B.3 displays the results for these placebo tests. In both cases, there is no differential attendance pattern across 4*th* and 3*rd* grades. Moreover, the distribution of $\Delta\overline{A}_j$ for those cases is symmetric and centered around zero. This provides further support for considering the evidence presented here as a source of non-random distortions in quality signals.

[13]Although daily attendance is not available for all years, it is possible to identify absenteeism on test days at the student level using the administrative records of annual academic performance and test scores: students with academic performance data but without test scores were absent on test days.

consists in using the multiple imputation methods developed by Rubin (1987). Using this strategy, we construct a panel dataset of distortions in quality signals for 2005–2013.

Estimating missing test scores

Let us begin with the estimation of missing test scores. Let q_{ijt} be the test score of student i in school j and year t, and x_{ijt} be a vector of variables that predict test scores at the student level and we observe for *all* students. Then, we estimate the following equation in the sample of test takers for each school in our dataset:

$$q_{ijt} \quad = \quad f(x_{ijt}; \gamma_j) + \lambda_t + \eta_{ijt} \tag{2.3}$$

where γ_j is a school specific vector of parameters, λ_t are year fixed effects, and η_{ijt} is a random error term with mean zero. Importantly, the vector x_{ijt} needs to contain strong predictors of test scores and be available for *all* students. We choose GPA and the following indicator variables: students who were in *4th* grade the previous year and students who studied at a different school the previous year. Note that equation (2.3) allows for the gradient of test scores to covariates in x_{ijt} to vary across schools. There are 7,500 schools in our dataset with, on average, 270 test takers between 2005 and 2013. This means that our imputation method relies on 7,500 regression equations that use on average 270 observations and that we estimate using OLS.

Unsurprisingly, GPA is the strongest predictor of test scores at the student level, as displayed by Figure B.6. Moreover, given the quadratic empirical relationship between test scores and GPA, we include this variable as a quadratic polynomial. Two statistical exercises provide support for this specification. First, the R-squared of the 7,500 linear regressions we estimate are high (approximately 0.51) and are always higher in the polynomial model (see Figure B.7). Second, we implement a cross-validation exercise in which we assume test takers are the universe of students and we delete the test scores of ten percent of students with low GPA, essentially mimicking real world patterns. Reassuringly, in this exercise the quadratic polynomial specification has a lower mean squared error than the linear model.[14]

We use equation (2.3) to predict test scores for absent students in the period 2005–2013. In order to account for the uncertainty related to the estimation of missing test scores, we estimate these test scores multiple times by drawing from the asymptotic variance of the estimated parameters $\hat{\gamma}_j$, an approach similar to that in Mas and Moretti

[14]A concern related to the proposed model of test scores is that of selective attendance. To test for selection, we re-estimated equation (2.3) using a Heckman selection correction and found evidence supporting our model. The excluded variable when calculating the Heckman corrected distortions is an indicator for students living outside of the school's county, which effectively predicts attendance on test days. These Heckman corrected distortions are remarkably similar to the uncorrected ones –but noisier, as expected– and both are highly correlated, as displayed by Figure B.10. Finally, our cross validation exercise shows remarkably similar results for both models in terms of mean squared error. Given this evidence, we utilize distortions estimated without this selection correction for the remainder of the paper.

(2009).[15] More precisely, for each absent student in our dataset, we generate one hundred estimated test scores based on equation (2.3).

Estimating distortions

After estimating test scores of absent students, we estimate "undistorted" quality signals using a simple simulation estimator. Let $\tilde{q}_{jt}^{(n)}$ be the average test score of school j in year t calculated using draw $n = 1, \ldots, 100$. Then, our estimate for an undistorted quality signal is:

$$\tilde{q}_{jt} = \frac{1}{100} \sum_{n=1}^{100} \tilde{q}_{jt}^{(n)}$$

The uncertainty of our estimates corresponds to the variance of the imputations $\tilde{q}_{jt}^{(n)}$. We order $\tilde{q}_{jt}^{(n)}$ from lowest to highest within a school and take the percentiles 2.5 and 97.5 to generate a 95 percent confidence interval for our estimate $\tilde{q}_{jt}$.

We define distortions in quality signals as $\psi_{jt} \equiv q_{jt} - \tilde{q}_{jt}$, where q_{jt} is the observed (distorted) quality signal of school j in year t. Thus, a school with a positive distortion is one that signaled a higher quality than its true quality through its test score. Each distortion in our dataset has an associated distribution and a corresponding confidence interval.

2.4.2 Descriptive statistics of distortions

Table B.1 presents descriptive statistics for distortions for different tested subjects. The average distortion has a value of 2.7 test score points. The distribution of distortions is remarkably similar across subjects as displayed by Figure B.8. Moreover, the correlation of distortions is high, as documented by Figure B.9. In what follows, we use the average of distortions across mathematics and language in $4th$ grade, as those tests were taken during all years in our sample.

The average distortion is equivalent to 0.1 standard deviations (σ) of test scores at the school level. To assess their relative relevance, we compare the size of distortions to the impact of educational policies. Bellei (2009) evaluates a program that substantially extended school days in public and voucher schools in Chile and finds an impact of 0.06σ on test scores. Contreras and Rau (2012) find that the impact of SNED on test scores was between 0.14σ and 0.25σ. More broadly, Kremer and Holla (2009) and Glewwe and Muralidharan (2015) review multiple educational interventions in developing countries and a significant share display impacts smaller that 0.20σ. Similarly, a survey of field experiments in developed countries by Fryer (2016) finds that average treatment effects from

[15]Alternatively, we could use a bootstrap procedure. We have done this as a robustness check and results are similar.

school-based interventions are between 0.05σ and 0.07σ. Then, distortions in quality signals are of a relevant economic magnitude.

Figure 2.2-A presents estimated distortions for all schools in our data set. The y-axis represents distortions (in test score points), while the x-axis orders schools from lowest to highest distortion. In addition, distortions in green (gray) are (not) statistically different from zero. Approximately 31 percent of distortions are *statistically* larger than zero, and 80 percent of schools have a positive distortion. Figure 2.2-B presents the distribution of distortions. That (i) the average distortion is different from zero, and (ii) the distribution is not normal, make it clear that distortions in quality signals are not random variation in test scores. Moreover, we should note that *relative*, not *absolute*, distortions are relevant in terms of their potential implications. In Figure B.11, we present an empirical analysis of the rank correlation between undistorted and distorted quality signals at the market-year level (see section 2.5.1 for details on market definition). Approximately 60 percent of rank correlations are different from one, which suggests distortions in quality signals cause changes in the rankings of schools. Moreover, Figure B.12 shows that there were ranking changes in almost all large markets and in a sizable share of small markets.

Finally, we relate the estimated distortions with the motivating evidence presented in section 2.3.2. We would expect schools with higher differential changes in attendance in test days for high performing students (i.e. the difference between $\Delta\overline{A}_j^{high}$ and $\Delta\overline{A}_j^{low}$) to display larger distortions in quality signals. In this line, we start by calculating the difference in $\Delta\overline{A}_j$ between students above the 75*th* percentile and below the 25*th* percentile of the school's GPA distribution. Then, we study the relationship between this measure and our estimated distortions. This relationship is displayed in Figure 2.3. Schools with the largest increases in relative attendance of high with respect to low ones on test days are also on average those with the highest estimated distortions, which provides evidence for our methodology for estimating distortions in quality signals.

2.4.3 Understanding distortions

What explains the variation in distortions in quality signals? We now present a discussion of the empirical determinants of distortions. For this, we employ the panel dataset of distortions at the school level between 2005 and 2013.

Schools' characteristics

A significant share of the variation in distortions is explained by school time-invariant characteristics. If we regress distortions on school indicators, we can explain 36 percent of the variance. If we restrict attention to schools with distortions statistically positive, we can explain 60 percent of the variance. These percentages are large, especially considering that the maximum variation that can be explained is probably lower than one due to measurement error in the dependent variable. Which fixed characteristics of schools predict distortions? Consider the following regression:

$$\psi_{jmt} = X'_{jt}\theta + \nu_{mt} + \varepsilon_{jmt}$$

where X_{jt} is a vector school attributes in year t and ν_{mt} is a market-year fixed effect. Markets are defined as isolated groups of schools, i.e., with no schools closer than 3 miles as discussed in section 2.5.1. In order to account for the uncertainty in ψ_{jmt}, we present estimates weighted by (the inverse of) the 95 percent distortion confidence interval, thus accounting for the uncertainty associated to each distortion.

Results are presented in Table 2.2-A and show that distortions are larger in small public schools, for-profit schools, schools serving relative low-income households, and schools with historic low attendance rates. These correlations are larger in schools with distortions that are statistically different from zero. Additionally, Table 2.2-B presents the autocorrelation of distortions, which is always positive and statistically different from zero. This positive autocorrelation serves as additional evidence that distortions are non-random but rather associated to school characteristics.

Additionally, we study whether variation in distortions can be explained by within-school-variation in observable characteristics including school fees, socioeconomic composition, undistorted quality, and measures of attendance and class size. In particular, we estimate:

$$\psi_{jt} = \beta X_{jt} + \eta_j + \nu_t + \varepsilon_{jt} \tag{2.4}$$

where X_{jt} is the covariate of interest, and η_j and ν_t are school and time fixed effects. Figure B.13 show basically no relationship between any of these variables and estimated distortions.[16]

Competitive environment

An alternative explanation is that market environment creates incentives for schools to introduce distortions and signal higher quality (Shleifer, 2004). The market-oriented nature of the system suggests that schools facing more competition might choose to increase their quality signals using distortions. Dorfman and Steiner (1954) provide a useful framework to study firm behavior in contexts in which price and quality are jointly determined. The authors show that firms offer higher quality when facing more quality elastic consumers.[17] This section tests for this "quality elasticity" and related hypotheses.

[16] The only clear relationship is that between the number of students missing on test days and the magnitude of the distortion, which is positive as expected: missing students are a necessary condition for this distortion.

[17] Dorfman and Steiner (1954) analyze the behavior of a monopolist and argue that quality is optimally set following the condition:

$$q = \frac{p}{c_q}\frac{\eta^q}{\eta^p}$$

We exploit within school variation in variables related to the competitive environment. We proceed by estimating regressions following equation (2.4). The variables we consider include the number of schools in the market, average quality, fees and distortion of rivals, and the position of a school in the distribution of fees and quality in the market. We also employ the estimates from our school choice model in section 2.5 to calculate quality demand elasticities.

Figure B.14 displays results graphically. Although changes in the number of schools in the market and changes in average attributes of competitors are uncorrelated with distortions, demand quality elasticity is strongly correlated with distortions. The latter result is consistent with Dorfman and Steiner (1954): schools facing higher quality elasticity optimally choose to signal higher quality. This result is reinforced by the fact that schools in higher percentiles of the market-level quality distribution also seem to introduce higher distortions.

School and teacher incentives

Finally we exploit quasi-experimental variation from two government programs to understand the role of perverse incentives. The first is the SNED teacher incentives program explained in section 2.2.2. This program effectively increases teachers' wages if students in the school obtain high test scores, and it provides variation in incentives depending on the probability that a school will earn the prize (Contreras and Rau, 2012). Teachers might react to the likelihood of obtaining these prizes by affecting student attendance patterns on test days in order to increase the school's average test score. If anything, we would hypothesize that schools closer to the prize threshold would display larger distortions. However, our results show that distortions are not higher in schools that are more likely to win the prize, providing some evidence against the hypothesis that teachers manipulate attendance to increase test scores. See Figure B.15 for more details and results.

The second government program we exploit is the ETL information program, also described in section 2.2.2. This program classified schools according to test scores: red (bad), yellow (regular), and green (good). This information was disseminated to households in order to aid school choice. The cutoffs of these categories provide quasi-experimental variation in the incentives to manipulate test scores.[18] We find some evidence that low quality schools have higher distortions around the cutoff between red

where q is quality, p is price, c_q is the cost of quality, and η^q and η^p are the quality and price demand elasticities, respectively. In our interpretation, however, we use their result to approximate the case of imperfect competition with multiple firms and the analysis of a particular firm facing residual demand which is one way of modeling school behavior in this market setting (Neilson, 2013). In our setting, we argue that observed quality q can be increased by either increasing true quality or introducing higher distortions.

[18]The timing for this exercise is relevant. The SIMCE test was taken shortly after ETL report cards had been distributed to households, in the same academic year. Therefore, any distortions schools could have introduced in October, 2010 would affect test scores before any households reactions in terms of school choice.

and yellow schools, but no differential distortions in the cutoff between yellow and green schools. Moreover, these differences mostly disappear once school characteristics are controlled for. Overall, these patterns do not provide evidence for this mechanism being the main driver of distortions. See Figure B.16 for further details and results.

Discussion

The empirical patterns presented in this section improve our understanding of distortions in several dimensions. First, distortions are a non-random phenomenon. Second, fixed school characteristics are strongly correlated with distortions. Third, distortions are not explained by within-school variation in observable school characteristics. Fourth, we provide suggestive evidence that the market environment is correlated with distortions through the quality demand elasticity that schools face. Fifth, we provide evidence against perverse incentives induced by public programs driving variation in distortions. And finally, because individual test scores are never disclosed, we can rule out consequences of test scores at the individual level as a driver.

Unfortunately we are not able to rule out other potential drivers of distortions. For instance, it is possible that the existence of different school *unobserved types*, where some choose to distort test scores and some choose not to do so. These types might be related to, for example, school principals, who we do not observe. Alternatively, schools' heterogeneous responses to idiosyncratic events might generate non-random changes in attendance on test days. For example, schools might react to government attempts to increase attendance. While our setting does not allow us to study these alternative explanations, that does not prevent us from estimating the market consequences of distortions, which is what we do in the next sections.

2.5 Implications for school choice

In this section, we address the implications of distortions in quality signals on school choice. We estimate a school choice model to evaluate those impacts. Using estimated preferences, we implement the counterfactual exercise of providing households with information on undistorted quality signals. This allows us to compare observed with counterfactual outcomes, as well as to compute the welfare loss caused by distortions.

2.5.1 School choice model

We estimate a model of school choice in the lines of Bayer et al. (2007). When constructing the model, we impose certain assumptions, some of which are related to the Chilean institutional framework. First, we assume that households are informed regarding both available schools and their *observed* characteristics. Second, we assume that schools do not select students based on attributes and do not face capacity constraints, i.e. households can enroll their children in any school in their choice set. As discussed by Gallego and Hernando (2009) and Neilson (2013), this assumption is likely to hold in the Chilean school system. Finally, we assume the household's location choice is independent of the school choice problem. This assumption is supported by the fact that there are no institutional constraints on the choice set of schools based on residential location.

Let households be indexed by i and schools by j. Household utility depends on school fees, quality, and distance to school, denoted respectively p_j, q_j and d_{ij}. They also derive utility from other school characteristics W_j. For notational simplicity, we denote $X_j = [p_j, q_j, W_j]$, which includes K attributes. Preferences are heterogeneous depending on household type, indexed by r. In our model, only observed heterogeneity in preferences is considered, as explained below. Moreover, we allow for households to derive utility from schools' characteristics that are unobserved to the econometrician, ξ_j. Finally, each household has an idiosyncratic preference shock, ε_{ij}, which we assume is distributed T1EV.

Under these assumptions, the indirect utility of household i of type r from enrolling their children in school j is:

$$u_{ij}^r = \sum_k x_{k,j}\beta_k^r + \xi_j^r + \beta_d^r d_{ij} + \varepsilon_{ij} \tag{2.5}$$

where the first two terms measure utility from characteristics that depend only on the school and are therefore constant across households of type r for a given school j, while the third term measures disutility from distance between household i and school j for households of type r, which varies across households. We can therefore rewrite equation (2.5) as follows:

$$u_{ij}^r = \delta_j^r + \beta_d^r d_{ij} + \varepsilon_{ij} \tag{2.6}$$

such that the parameters of the model are contained in the vector β^r, but can be alternatively represented by the vector δ^r and by β_d^r. Note that δ_j^r is the component of utility derived from choosing school j that is constant across households, the mean value of school j for households of type r.

The probability of household i choosing school j can be derived analytically using households indirect utility.[19] The choice probability of school j by household i of type r predicted by the model is a function of school and household characteristics:

$$P_{ij}^r(\mathbf{ffi^r}, \mathbf{d^r}, \beta_d^r) = \frac{\exp(\delta_j^r + \beta_d^r d_{ij})}{\sum_{l \in \mathcal{J}_i} \exp(\delta_l^r + \beta_d^r d_{il})} \tag{2.7}$$

where $\mathcal{J}_i$ is the set of schools in the market where household i is located. We use this result in the next subsections for both estimating the model and for computing the counterfactual exercise of interest.

[19] In the context of school choice, there is no obvious outside option. Therefore, we follow Neilson (2013) and instead normalize $\delta_1 = 0$ within each market.

Estimation

We estimate the parameters of the model using a two-step procedure. First, we estimate standard conditional logit models for each group r in each market and year in the data, to recover schools' mean values. Second, we exploit the assumed linear functional form of households' indirect utility function in order to estimate the relationship between schools' mean values and and school attributes and recover preference parameters .

The first stage of the estimation procedure consists of estimating equation (2.7), which can be done by maximum likelihood. In order to allow for heterogeneity in preferences, this procedure is implemented within each of multiple cells defined on the basis of R socioeconomic levels, T time periods, and M markets. The former is determined by the eligibility of a student for the SEP program, which is determined fundamentally by participation in social programs aimed at supporting poor households. Therefore, we estimate $R \times T \times M$ conditional logit models in the first stage, which yields the same number of estimates for δ^r and β^r_d.

The second stage exploits the assumed linear functional form of the utility function in order to estimate the following linear regression:

$$\delta^r_{jmt} = \delta^r_{0,mt} + \sum_k x_{k,jmt}\beta^r_k + \epsilon^r_{jmt} \tag{2.8}$$

where $\delta^r_{0,mt}$ is a constant term specific to each market, year, and household type; β^r_k measures the effect of x_k on school mean value for households of type r and maps to the preference parameters of our model; and ϵ^r_{jmt} is a mean-zero error term. Note that $\delta^r_{0,mt} + \epsilon^r_{jmt}$ maps to the unobserved school characteristic ξ^r_{jmt} in our model.

A concern with this type of regression is the potential endogeneity of school characteristics, particularly of prices and quality. Therefore, we estimate this regression using an instrumental variables approach, using various instruments. First, for each school, we include the fixed non-price and non-quality characteristics of other schools in the market, in line with instruments suggested in Berry et al. (1995). In particular, we compute the share of religious schools, schools with gender constraints, and public schools in the market for each school in the sample, using them as instruments. Second, we follow Neilson (2013) and use average teacher hourly wages, which arguably operates as a cost shifter for schools, such that it might affect their choices of fees. Third, we use the amounts of funding provided by different voucher program components, which display within market variation due to school characteristics that are fixed in the short run. In particular, we include the baseline voucher and two additional components related to a school being part of the SEP program and to a school having a concentration of SEP students above a threshold. Moreover, we utilize county temperature data on test days as an instrument for quality. While the data provides support for a relationship between temperature and test scores, it would be hard to argue that temperature on test days could otherwise be correlated with unobserved school attributes. This instrument is motivated by a literature that studies the relationship between climate and academic achievement, as discussed in Graff Zivin et al. (2015).[20] Finally, we use an indicator variable for whether a school was awarded a SNED prize in its most recent version. This instrument is motivated by Contreras and Rau (2012) who show how these prizes impact quality in subsequent years.[21]

We estimate the model using data for 2011 through 2014, the only years in which student home address data is available. In addition, we only utilize data for students in 1*st* grade in order to focus on the margin in which most school choices are made. In terms of covariates to be included in the vector X_j, we include school fees, quality as measured by the school's average SIMCE test score, whether the school has a religious orientation, whether the school has any gender constraints, whether a school is public, and whether a school is part of the SEP program.[22] Finally, we are able to compute the distance between households and schools using geo-referenced data on their addresses.[23]

Market definition and estimating dataset

Determining which suppliers belong to the consumers' choice set in context of spatial competition is not straightforward. In contrast to other school systems, in Chile there are not any institutional constraints that limit the extent to which students can travel. Therefore, we need to define markets.

We adopt an approach based on the spatial distance between schools, similar to that in Neilson (2013). Distance has been shown to be a relevant determinant of school choice in the literature (Gallego and Hernando, 2009; Neilson, 2013). In our data, students' average distance to chosen schools is 1.3 miles and the 90*th* percentile of such distribution is 3 miles. Therefore, it makes sense to argue that schools located far enough from each other might belong to different educational markets. We define an educational market as a cluster of schools in a closed polygon with no other school closer than 3 miles from its boundaries. Operationally, a market is uniquely identified from the adjacency matrix of schools, where links are defined as two schools being closer than 3 miles from each other. In implementing this procedure, and therefore in estimation as well, we only con-

---

[20] We construct this variable using data from the Berkeley Earth dataset, which provides population-weighted estimates of daily temperature at the county level. In implementing this regression, we include both temperature and temperature squared in order to account for non-linear effects of temperature on academic achievement as documented in Graff Zivin et al. (2015).

[21] In practice, we utilize the residual of a regression of the SNED award indicator on quality in the year of the award in order to further control for quality differences between SNED awardees and non-awardees which might be driven by other factors that could be persistent in time.

[22] We use data on monthly copayments faced by households as a measure of school fees. Moreover, we use data on students' eligibility for SEP in order to adjust school fees accordingly; eligible students do not pay any school fees in schools that operate under the SEP regime.

[23] We compute the Euclidean distance between every household and school in each market. We then proceed to clean these results by (i) removing mass points, which arise from imperfect geo-reference; and (ii) removing students located further than 55 kilometers from the median household location in the market.

sider urban schools. Specifically, we only include markets with at least 20 schools and for which we have data for at least 300 students. Table B.2 displays summary statistics for markets.[24]

A description of the resulting sample is displayed in Table 2.3. The number of household types is $R = 2$, the number of markets included is $M = 25$, and the number of periods covered is $T = 4$. Therefore, the estimating dataset is comprised of 200 cells. The estimating dataset includes 1,556 schools and 97,471 students. On average, 33 percent of the students attending schools in markets in our sample are included, and 92 percent of the schools operating in each market. Moreover, an average of 49 percent of students included in the sample across markets are eligible for the SEP program.[25]

Results

Given that the most relevant dimension of household heterogeneity is socioeconomic status, we present all the results for poor and non-poor households separately. Figure B.17 displays the resulting coefficients in each market for distance between households and schools for both poor and non-poor households. In all these cases, the coefficient is negative, which reflects a decreasing utility for choosing a school further away from home. Poor households are on average 14 percent more distance-sensitive than non-poor households.

Table 2.4 presents results for different specifications of instrumental variables linear regressions of the estimates of δ_{jmt} on different sets of school characteristics and fixed effects. Columns 1 through 3 display results for all households in the sample, columns 4 through 6 display results for poor households, and columns 6 through 9 for non-poor households. Overall, results point in the expected direction: household utility decreases with school fees and increases with their reported quality. Both adding market-year fixed effects and other school attributes to the regression increase the magnitude of point estimates with respect to the baseline case.[26] Overall, the model provides a good fit for observed enrollment shares, as displayed by Figure B.18. The correlation between observed and predicted enrollment shares is of 0.88.

[24] As a robustness exercise, we estimated the model using counties as markets. For estimation, we included counties for which a large share of students resided in the market (at least 90 percent) and where we had available data for more than 300 students. Results were quantitatively similar.

[25] We tested for differences in observables across students included and excluded in the sample within each market. While some of the differences across groups are statistically significant, they are not economically significant and do not show a clear pattern. Results are available upon request.

[26] Table B.3 and Table B.4 display results from the first stage of the IV estimation for school fees and quality respectively. The bottom rows in Table 2.4 show the respective F-tests for the subsets of instrumental variables utilized for school fees and quality respectively. Moreover, we further assess the strength of the instruments by reporting the Cragg and Donald (1993) eigenvalue statistic for each specification. Stock and Yogo (2005) provide critical values for rejection of this test. In our setting, the critical value for rejection is 29.32, always below the reported values for the Cragg-Donald statistic. Finally, Table B.5 displays the results from estimating the second stage of the model by OLS. As expected, the OLS estimates are in general smaller in magnitude than the IV ones.

There are interesting patterns of heterogeneity across poor and non-poor households. For example, our preferred specifications in columns 6 and 9 imply that poor households are 88 percent more price-sensitive than non-poor households. Inversely, poor households are estimated to be 37 percent less quality-sensitive than non-poor households. These results imply in turn that non-poor households' willingness to pay for quality is three times higher than that of poor households. This heterogeneity suggests that quality disclosure policies will have heterogeneous effects across these demographic groups. These patterns of heterogeneity coincide with previous findings within the school choice literature (e.g. Gallego and Hernando 2009, Hastings et al. 2009, and Neilson 2013).[27]

2.5.2 Counterfactual analysis

In our setting, schools quality signals are distorted and therefore households are choosing schools on the basis of a misperceived vector of attributes. A key aspect of the situation, however, is that while perceived school quality might be different than true quality, the value that households ultimately obtain from a school is the true quality of their school choice. This is related to the distinction stated by Bernheim and Rangel (2009), by which some elements of the choice environment may be relevant for constructing positive descriptions of choice behavior, but not for welfare analysis. Throughout this section, we emphasize this aspect and account for it when measuring implications of distorted quality signals.

In order to compute the effects of distorted quality signals on choices and welfare, we define two scenarios: *baseline* and *counterfactual*. The former corresponds to an environment in which households actually choose schools. The latter corresponds to a counterfactual world in which households are provided with undistorted information about school quality. This exercise rules-out changes in other variables (e.g. school fees and school investments) as well as the existence of capacity constraints. While those might be relevant margins of supply side behavior in this market, we argue that the impacts of the policy we evaluate in this counterfactual exercise would induce remarkably small equilibrium responses by schools on these margins.

Throughout this section, we utilize our estimates for δ^r and β^r_d, and the observed vector of school characteristics X_j to compute choice probabilities and consumer welfare for the baseline scenario. For the counterfactual scenario, calculations additionally use estimates of β^r_k from the second stage of the school choice model, and a counterfactual vector of school characteristics $\tilde{X}_{ij} = [p_j, \tilde{q}_j, W_j]$, where $\tilde{q}_j$ stands for the undistorted quality of school j.[28]

[27]As a robustness check on the results, we study the correlation in estimates of unobserved school characteristics ξ^r_{jmt} across poor and non-poor households. While there is variation in results across both groups, there is a positive correlation of 0.57 between the estimates of ξ^r_{jmt}. This is, while services provided by schools might be differently valued across consumer types, those values are strongly correlated across them.

[28]More precisely, we utilize the results for the second stage from our preferred specifications: columns

Choices

We begin the analysis by examining school choice probabilities by households across both scenarios. We do so by adjusting the choice probabilities predicted by equation (2.7) of our school choice model and using parameter estimates and data on school attributes for both scenarios. Following equation (2.7), choice probabilities are therefore computed as $P^r_{ijmt}(d^r, \hat{\delta}^r, \hat{\beta}^r_d)$ and $P^r_{ijmt}(d^r, \tilde{\delta}^r, \hat{\beta}^r_d)$, where $\tilde{\delta}^r_{jmt} = \sum_k \tilde{x}_{k,jmt}\hat{\beta}^r_k + \hat{\xi}^r_{jmt}$ is the mean utility of school j in market m in period t, computed using preferences estimates and data on counterfactual school quality.

Figure B.19 displays the computed changes in choice probabilities between both scenarios. It is interesting to note that, despite the fact that the magnitude of estimated distortions is moderate, there is significant heterogeneity. This pattern holds when restricting the analysis to the set of schools actually chosen by parents as displayed by Figures B.19-C and B.19-D. This shows that changes in the quality disclosure system would induce changes in households' choices. However, given that households have a limited number of schools in their choice sets, these changes in choice probabilities might only induce actual changes for a small fraction of households. Those marginal changes in the observed vector of school quality might not be strong enough as to induce households to actually change their school choices. Note that non-poor households display more variation in the computed changes, which is driven by their higher quality sensitivity. This stands in contrast with potential gains from the policy, as the average distortion in poor household choice sets are 0.33σ higher than those in non-poor household choice sets. Despite that difference, a simple simulation based on the proposed model and our estimates shows that 3.3 percent of poor households and 3 percent of non-poor households would be induced to change their school choice when provided undistorted quality information. The higher willingness to pay for quality of non-poor households explains these similar switching rates despite the large gap in distortions faced by both groups. We denote this subpopulation as *switchers*.[29]

We compute the predicted attributes of schools chosen by households under both scenarios. Table 2.5 displays results from these calculations for both poor and non-poor households. We report both the average across all households as well as the average within the switcher subpopulation. Columns 1 and 3 in Table 2.5 display results for the average across all households. It is easy to note that changes in predicted distance to chosen schools and fees are small. This is expected since non-switching households are unaffected by the information policy we evaluate. The average changes in attributes of chosen schools by poor and non-poor households are not larger than 0.03σ for any of the attributes considered.

When considering the subpopulation of switchers, however, changes in chosen schools'

6 and 9 of Table 2.4.

[29]We calculate switching rates by simulating choices of consumers in our sample in both the baseline and counterfactual scenarios. Reported results correspond to average switching rates for poor and non-poor households over 200 simulations across all households in the sample.

attributes between both scenarios are substantive. Columns 2 and 4 in Table 2.5 display results for these household groups. First, note that in the baseline scenario, switchers were receiving substantially less quality than the average household, which suggests that switchers mainly had chosen schools that had highly distorted quality signals. Conditional on switching, we observe that households are willing to travel longer distances to chosen schools, to pay higher fees and, importantly, that they choose schools with remarkably higher true quality. In particular, our results show that poor (non-poor) switchers would choose schools with 0.71σ (0.74σ) higher true quality in the counterfactual than the baseline scenario. This would be coupled by an increase in fees paid to chosen schools of 0.2σ (0.49σ) for poor (non-poor) switchers and, similarly, an increase in distance travelled to chosen schools of 0.04σ (0.05σ). These results imply that the subpopulation of switchers would change their choices in a substantial way. Switchers move towards higher-quality schools, for which they are willing to both travel more and pay higher fees.

Welfare

We now calculate the welfare changes of providing undistorted quality signals. In the baseline scenario households choose schools using the observed measure of school quality, which, as discussed, is distorted. However, consumers' effective utility is determined by undistorted school quality. Thus, our baseline scenario is a case in which *choice* utility and *experience* utility differ (Bernheim and Rangel, 2009). This is not the case in the counterfactual scenario in which households choice and experience utility coincide.

Let u_{ij} be the utility of household i from school j under distorted school quality, choice utility. Similarly, let $\tilde{u}_{ij}$ be the utility of household i from school j under undistorted school quality, experience utility. In our setting, these two utilities are related. Given that the only difference between choice and experience utility is the misperception of quality under the former, we know that $\tilde{u}_{ij} = u_{ij} + \tau_j$, where τ_j measures the wedge between choice and experienced utility from school j. Under the utility function assumed in section 2.5.1, we know that $\tau_j = \beta_q(\tilde{q}_j - q_j)$.[30]

The choices household i would make in each scenario would be:

$$\begin{aligned} j_i^* &= \arg\max_j \{u_{ij}\}_{j \in \mathcal{J}_i} \\ \tilde{j}_i^* &= \arg\max_j \{\tilde{u}_{ij}\}_{j \in \mathcal{J}_i} \end{aligned}$$

which might or might not differ. Importantly, if the choice is the same in both scenarios then there is no welfare loss from distorted quality signals for household i, as experience utility is the same in both cases. This makes it clear that welfare losses will be driven by

[30]These linear relationships between observed and true quality and between choice and experience utility are similar to those analyzed in Train (2015). From this expression for τ_j, it becomes clear that at baseline all schools with positive distortions have $\tau_j < 0$, such that experience utility from those schools is lower than choice utility from them.

households that were changing their behavior due to distorted quality signals.

The change in household welfare from providing undistorted information would therefore be the difference in experience utility between the counterfactual and baseline scenarios, $\tilde{u}_{i\tilde{j}^*} - \tilde{u}_{ij^*}$. Using the fact that $\tilde{u}_{ij^*} = u_{ij^*} + \tau_{j^*}$, we can compute the expected monthly welfare change as:

$$E[\Delta W_i] = \frac{1}{\beta_p}\left[\log\sum_j \exp(\tilde{v}_{ij}) - \log\sum_j \exp(v_{ij}) - \sum_j P_{ij}\tau_j\right] \tag{2.9}$$

where we define $\tilde{v}_{ij} \equiv \tilde{\delta}_j + \beta_d d_{ij}$ and $v_{ij} \equiv \delta_j + \beta_d d_{ij}$ for notational simplicity. The first and second terms measure consumer surplus under undistorted and distorted school quality information respectively, and the results follow from the inclusive value formula in Small and Rosen (1981) given the assumed utility function. The third term measures the expected difference between choice and experience utility at baseline. Dividing by β_p simply transforms the welfare loss to monetary units. Equation (2.9) calculates the average welfare gain over the whole household sample. We can then compute average welfare gains for switchers or aggregate these gains across different dimensions. These welfare gains can alternatively be interpreted as the average willingness to pay of households for undistorted quality information.

Results from welfare calculations are displayed by Table 2.6 and show that expected welfare would increase in the counterfactual scenario for all households. This is as expected: non-switchers will obtain the same welfare in both scenarios, while switchers will be strictly better off. For poor households, the average yearly welfare gain we estimate is $1.7. The average yearly welfare gain for non-poor households is $5.3. Scaling up these results for the educational system, welfare gains would add up to $7 million annually.

These results suggest small gains from providing undistorted information across all households. However, the relevant population for an intervention like the one proposed in our counterfactual is not the average household; but rather the subpopulation of switchers. Welfare gains are larger for this subpopulation. The average yearly welfare gain for switchers is $53 among poor households and of $174 among non-poor households. Gains for switchers are thus sizable: poor (non-poor) switchers would experiment welfare gains of 11 (36) percent of the average school fee in our sample.

Heterogeneity in welfare gains

The fact that non-poor households benefit more than poor households from the information policy is evident, and the magnitude of the differences is large. There are two potential explanations for this. First, the former are more quality-sensitive, and less price and distance-sensitive than the latter. Therefore, they will be more willing to take advantage of relative changes in perceived quality of schools in the market. Second, the spatial distribution of households and schools in the market differs systematically across poor and non-poor households, giving them potentially differential opportunities to improve their choices in the counterfactual.

We can use our model and estimates to explore how heterogeneity in preferences and market opportunities determine the observed gap in welfare gains from disclosure of true quality. Results from these additional counterfactual calculations are displayed in Table 2.6. We start by studying how differences in preferences determine lower welfare gains for poor households. First, we let poor households be as quality-sensitive as non-poor ones. The share of switchers among poor households would increase by 0.8 percentage point to 4.1 percent, and the average yearly welfare gains for switchers would increase to $101.[31]

Second, we let poor households have the same preferences as non-poor households on all school attributes. The share of switchers increases by 0.6 percentage point to 3.8 percent. Average yearly gains for poor switchers in this counterfactual would climb to $181, more than three times those in the first counterfactual and higher than those for non-poor switchers.[32] These results imply that differences in preferences are enough to explain the gap across groups in welfare gains from the proposed information policy. Moreover, they highlight the key role that households' quality-elasticity plays in determining the impacts of information policies for school choice.

Finally, we explore the role that the spatial distribution of schools and households play in explaining the gap in welfare gains across groups. We measure welfare gains from the evaluated policy for poor households if they were located in the same place as non-poor households. Our results show that average welfare gains in that setting would be essentially the same that we found in our baseline results above. The share of switchers in this case would be lower than in the first counterfactual, at 2.4 percent, while yearly welfare gains for poor switchers would be only slightly larger than in such counterfactual, $65. This result implies that, in our setting, differences in market opportunities faced by poor and non-poor households play a minor role in explaining the gap in welfare gains from undistorted quality information.

Discussion

We have estimated a school choice model and studied a counterfactual exercise by which information on undistorted quality signals is provided to households. Results point in three directions. First, distortions in quality signals have effects on choices, as choice

[31]Recall that in conditional logit models, coefficients are normalized by the standard deviation of the idiosyncratic preference shock, σ_ε^r, which may vary across household types. Thus, in practice, this counterfactual is not exactly letting the poor have the quality preference of the non-poor, but rather the estimated normalized preference coefficient of such group. This is equivalent to making poor households almost twice as price sensitive as estimated.

[32]The fact that welfare gains for the poor when endowed with preferences of non-poor households are larger than those when endowed with such preference only over school quality comes partly from the fact that we estimate non-poor households to be less price-sensitive. This implies that the willingness to pay for a given increase in quality is higher than under poor preferences as can be noted in equation (2.9).

probabilities would change in the counterfactual scenario. Second, households would react to the change in the quality disclosure system mostly by increasing the probability of choosing higher quality schools. There would thus be a shift of students towards relatively high quality schools available in the market. Third, our welfare calculations point towards small average gains across households but sizable gains for switchers. In both cases, gains are larger for non-poor households, which is driven by them being more quality-sensitive and less price-sensitive. Complementary policies that could increase poor households quality-sensitivity might increase welfare gains from this policy.

Throughout this section, we have assumed that households are not informed about distortions in quality signals. If they were informed, they would optimally incorporate that information and adjust their choices according to true school quality. Because calculating distortions is a complex task, we argue that parents are unlikely to incorporate them in their decisions. In an intermediate scenario, if households had partial knowledge about distortions, then welfare gains for switchers would certainly be lower and our estimates would be an upper bound.

The magnitude of welfare gains for switchers already suggests that it might be socially beneficial to invest in quality disclosure systems that reduce distortions in educational markets. Note, however, that our counterfactual policy does not evaluate the welfare effects of the disclosure system in place, but rather the welfare effects of distorted quality signals given the current school quality disclosure system. Moreover, note that these welfare calculations do not consider the social costs of potential hidden actions that might be driving distortions. In that sense, our results provide a lower bound for welfare gains from correcting distortions in this market.

2.6 Misallocation of public programs

There is a second set of implications of quality signal distortions. Multiple public programs are allocated using rules that follow directly from test scores. Thus, distortions in test scores will induce misallocation of public programs. This section studies the extent of such misallocation for two relevant public programs: teacher bonuses and additional information for school choice. While these measures are policy relevant, we note that distortions might induce other unobserved behavioral responses by school personnel that might also be socially costly. We cannot, however, measure these hidden actions with available data.

2.6.1 Monetary incentives for teachers

As explained in section 2.2.2, teachers are awarded bonuses by the SNED program depending on their school's average test score. In 2012, the total amount of public resources transferred to schools in the form of teacher bonuses reached 15 million U.S. dollars. The sharp discontinuity to assigning resources is based on the following index for each school:

$$I_{jgt}(q_{j\tau}, q_{j\tau-1}, X_{j\tau}) \;=\; \omega_1 q_{j\tau} + \omega_2(q_{j\tau} - q_{j\tau-1}) + \boldsymbol{\omega}_3' \mathbf{X}_{j\tau} \tag{2.10}$$

where I_{jgt} is the index of school j, in group g, and year t; $q_{j\tau}$ is the average test score in year τ; $\mathbf{X}_{j\tau}$ is a vector of attributes; and $(\omega_1, \omega_2, \boldsymbol{\omega}_3)$ are weights chosen by the government, with $\omega_k \in (0,1)$ and $\sum_k \omega_k = 1$. Note that: (i) $t > \tau$, otherwise the index cannot be computed as the inputs to calculate it are not observed, (ii) all input variables are mapped to the $[0,1]$ interval before computing the index, and (iii) groups g are defined by the government using schools attributes.

We say there is misallocation of public resources if teacher bonuses were given to schools that would not have receive bonuses in a counterfactual scenario without any distortions in quality signals. In particular, using our estimates for undistorted quality signals $(\tilde{q}_{j\tau}, \tilde{q}_{j\tau-1})$, we calculate schools undistorted indices using equation (2.10), $\tilde{I}_{jgt} = I_{jgt}(\tilde{q}_{j\tau}, \tilde{q}_{j\tau-1}, \mathbf{X}_{j\tau})$, and reallocate bonuses based on these undistorted measures.

Figures 2.4-A and 2.4-B present the actual and the counterfactual assignment of bonuses. To the left of the threshold (vertical line) are the schools that did not get bonuses, and to the right are the schools that did. The percentage of public resources that were misallocated is the total amount of money that was incorrectly given to some schools over the total amount of resources that schools received. We estimate that 13 percent of teacher bonuses were misallocated.

Although intuitive, our method to calculate misallocation of public resources still needs to account for the uncertainty associated with the estimation of undistorted quality signals. For this, recall that each school-year distortion has an associated distribution. We proceed by taking 1,000 independent draws of distortions from their school-year distribution – a normal distribution with a school-year specific mean and standard deviation – and calculate the percentage of misallocated public resources 1,000 times. Bounds for our misallocation estimates can be constructed using the estimated distribution of misallocation.

Our estimates indicate that 13 percent of teacher bonuses were delivered to the incorrect schools, which is equivalent to $2 million every two years or approximately $20 million since this public program started in 1996. This estimate is significantly different from zero and precise: we can rule out misallocation of public resources being less than 11 percent.

2.6.2 Information for school choice

A quality disclosure program was implemented in 2010 ("Educational Traffic Lights"), aimed at providing simpler information about school quality (more details in section 2.2). Schools were classified, based on the average test scores of 4*th* and 8*th* graders, into three mutually exclusive categories. Maps with school categories were disseminated across counties with the explicit objective of affecting parents information set.

Let $c_j = \{r, y, g\}$ be the category of school j (red, yellow, green). Schools were assigned to categories using the following formula:

$$c_j(q_{jt}) \quad = \quad r \cdot 1[q_{jt} < \underline{s}] + y \cdot 1[\underline{s} < q_{jt} < \bar{s}] + g \cdot 1[q_{jt} > \bar{s}] \tag{2.11}$$

where q_{jt} is the average test score of school j in year $t = 2009$, and $(\underline{s}, \bar{s})$ were thresholds decided by the government. These thresholds corresponded to one standard deviation lower ($\underline{s}$) and higher ($\bar{s}$) than the average test score of all schools.

Equation (2.11) makes it clear that the provided information is directly linked to distorted quality signals. Because the formula used to categorize schools is known, we can replace distorted quality signals by undistorted ones, assign undistorted categories $\tilde{c}_j = c_j(\tilde{q}_{jt})$, and calculate the percentage of schools that were incorrectly categorized. In order to account for the uncertainty in our undistorted quality signals, we follow the same strategy as in the previous section.

Figures 2.4-C and 2.4-D present our results. Our estimates indicate that approximately 4 percent of schools were assigned to an incorrect category. Moreover, we can rule out that fewer than 3 percent of schools were misassigned. Using the causal effects reported in Allende (2012) we calculate that, as a consequence of this misallocation of categories, approximately 5,000 students (two percent of the *1st* grade cohort) attended schools in misallocated categories. The welfare implications for the compliers are, however, not straightforward to calculate as some children attended higher-quality and some attended lower-quality schools.

2.7 Conclusion

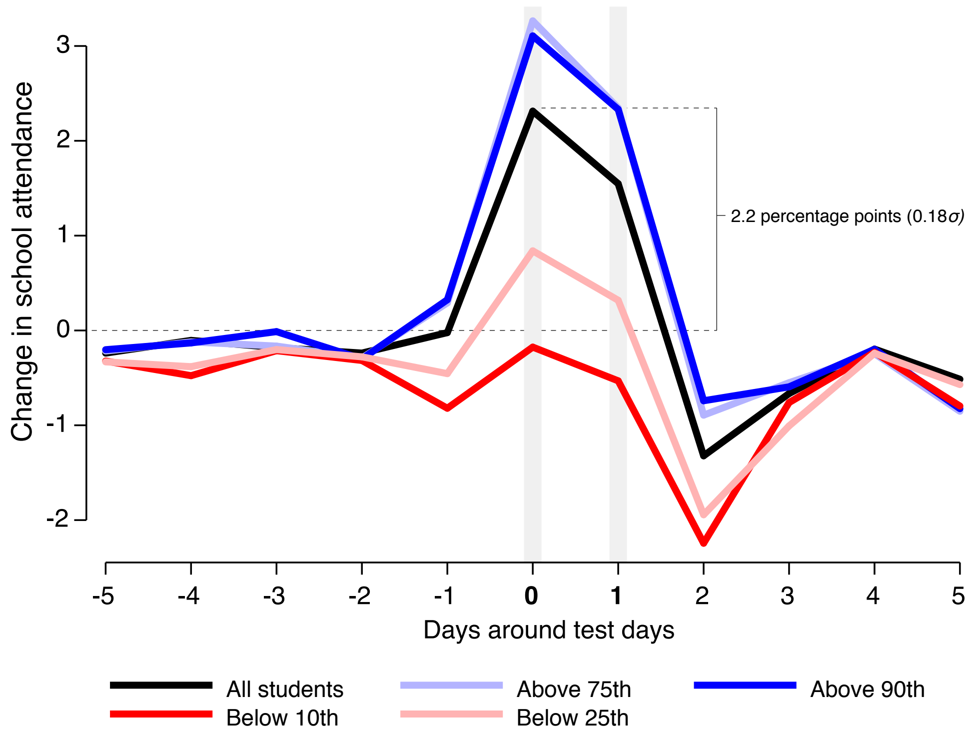
We have shown that significant distortions in quality signals are in place in the Chilean educational market, which is dependent on high-stakes testing. In particular, we have quantified how non-random attendance on test day causes school quality signals to be distorted. Our results are consistent with the so-called Campbell's Law: the higher the stakes are for an indicator of a social phenomenon, the more liable it is to be distorted (Campbell, 1979). Distortions, however, are not *per se* a reason of concern. To claim distortions have costs, we need to study the impacts they have on decisions. The Chilean market-oriented educational system is particularly interesting to study the consequences of distortions because: test scores are not just used for the two objectives of quality assessment and performance evaluation emphasized by Neal (2013), but rather for three, as these also feed school choice. We show that distortions have negative impacts on school choice and also induce misallocation of public programs. These findings allow us to conclude that distortions can impose significant costs in educational markets.

Our study is, to the best of our knowledge, the first to quantify the *market* consequences from distortions in quality signals. Further research is required to quantify other distortions and to address other margins of educational markets. Quantifying the costs associated to other hidden actions is also necessary to fully characterize the costs of accountability systems. We highlight that the institutional environment might determine the magnitude and impacts of distortions.[33] Market-oriented educational systems such as the one we have studied –where test scores play a key role as quality signals in disclosure policies– might be particularly prone to exacerbating the consequences of distortions.

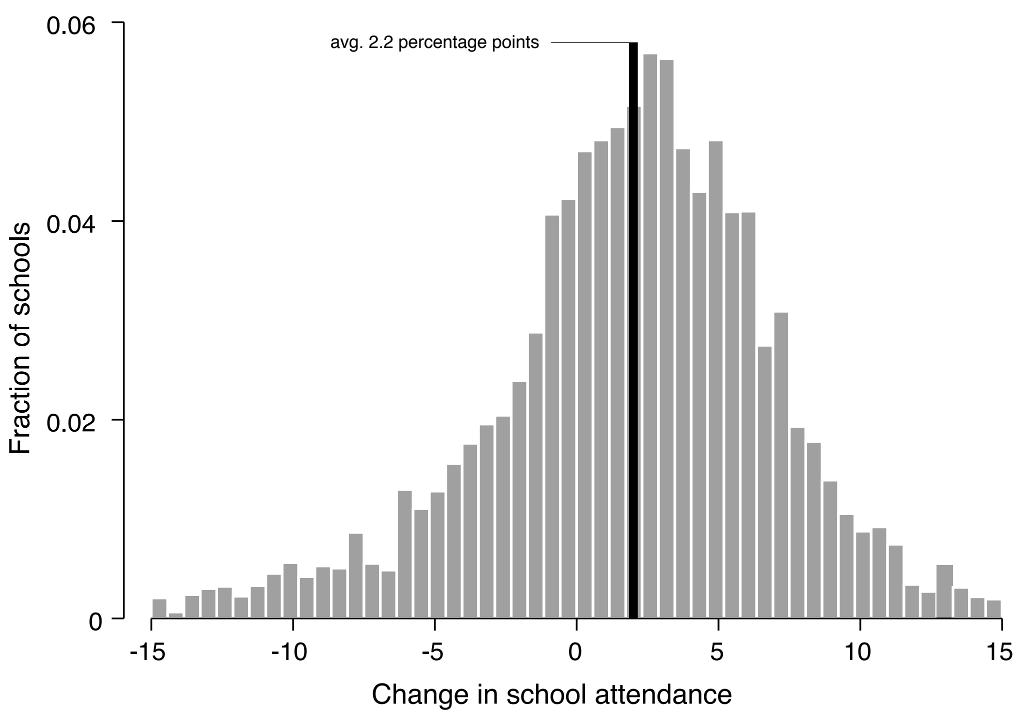
Our results have several policy implications. Previous work has emphasized the importance of providing information to parents, while our work emphasizes the importance of providing *undistorted* information. A simple solution within the current system is to calculate quality signals as a *corrected* average of test scores instead of an *arithmetic* average. One way to accomplish this correction is using the imputation method we have proposed. This seems to be a better solution than requiring a minimum attendance rate (e.g., 95 percent in No Child Left Behind) in contexts where test scores can affect school choice. In addition, we emphasize that the magnitude of elasticities determines the extent to which households can benefit from information policies. In school markets, we argue that complementary policies that increase quality-sensitivity of the poor might enable them to benefit more from accurate information. Finally, our results on misallocation of public programs provide an argument against sharp assignment rules for public programs based on variables prone to distortions. Multiple programs in different countries and sectors are assigned through such kind of rules and might be subject to the similar misallocation as described in this paper.

[33] A relevant institutional dimension is the level of corruption. Interestingly, Chilean counties with higher levels of corruption have larger distortions in quality signals (see Tables B.6 and B.7). This suggests that our findings might be exacerbated in settings with different levels of corruption.

Figure 2.1: School attendance around test days

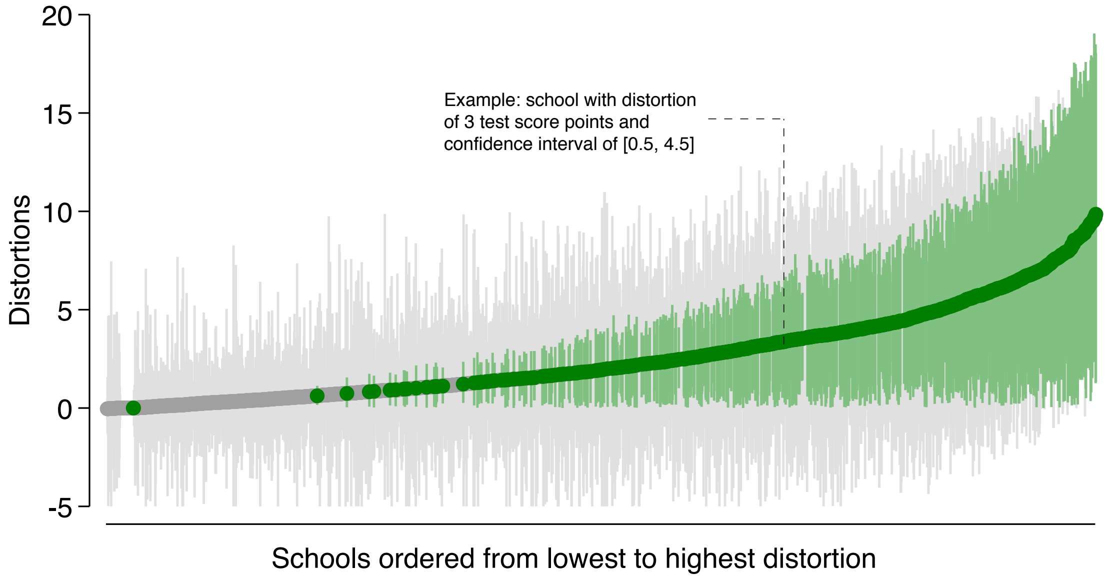


(a) Difference in average attendance rate (y-axis, in percentage points) between $4th$ graders (test takers) and $3rd$ graders (non-takers) around the two test days in 2013 (x-axis). Students are grouped by their position in the school GPA distribution.

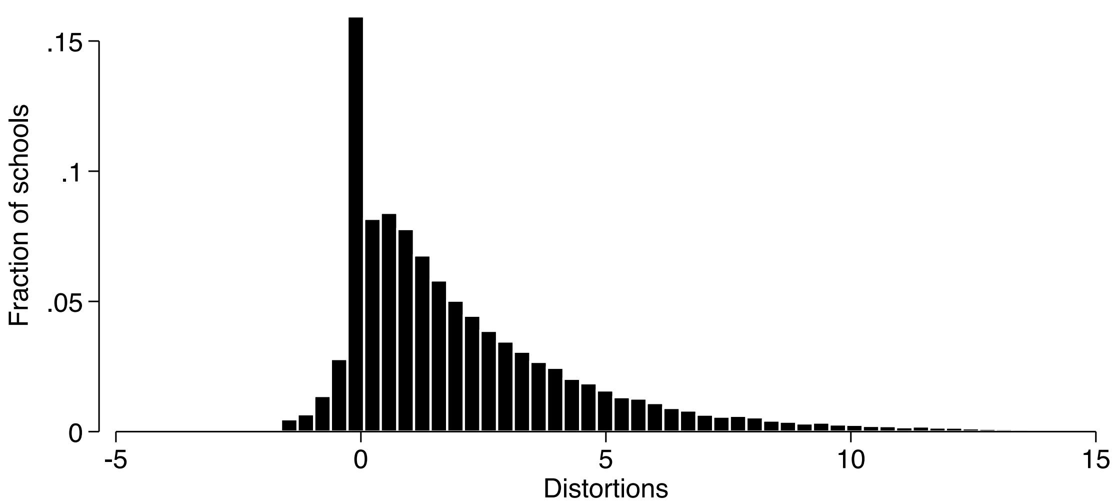


(b) Distribution of changes in school attendance in test days in 2013 (in percentage points).

Figure 2.2: Distortions in quality signals

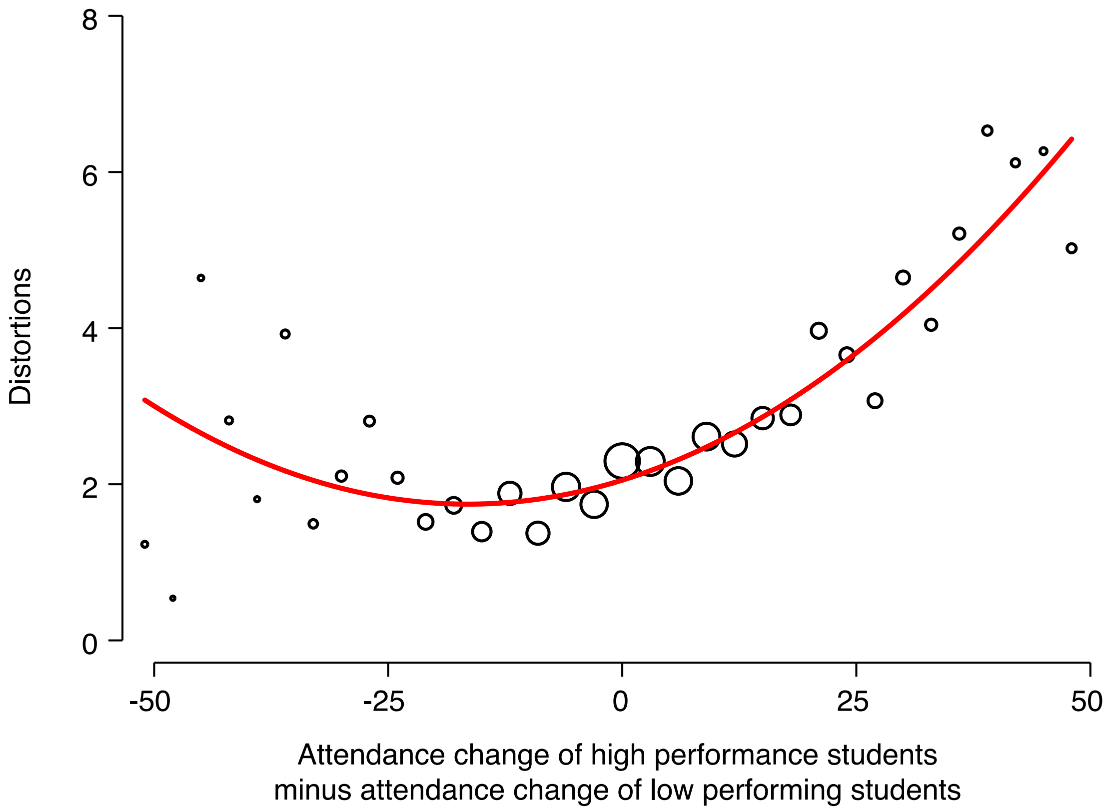


(a) Distortion in quality signals (y-axis, in test score points) are defined as (minus) the difference between school's observed test score and school's counterfactual test score. Schools are ordered from lower to higher distortions in the x-axis. Vertical lines represent the 95 percent confidence interval. Green (gray) lines represents distortions that are (not) statistically different from zero. The figure includes a random sample of distortions for 3,000 school-years.



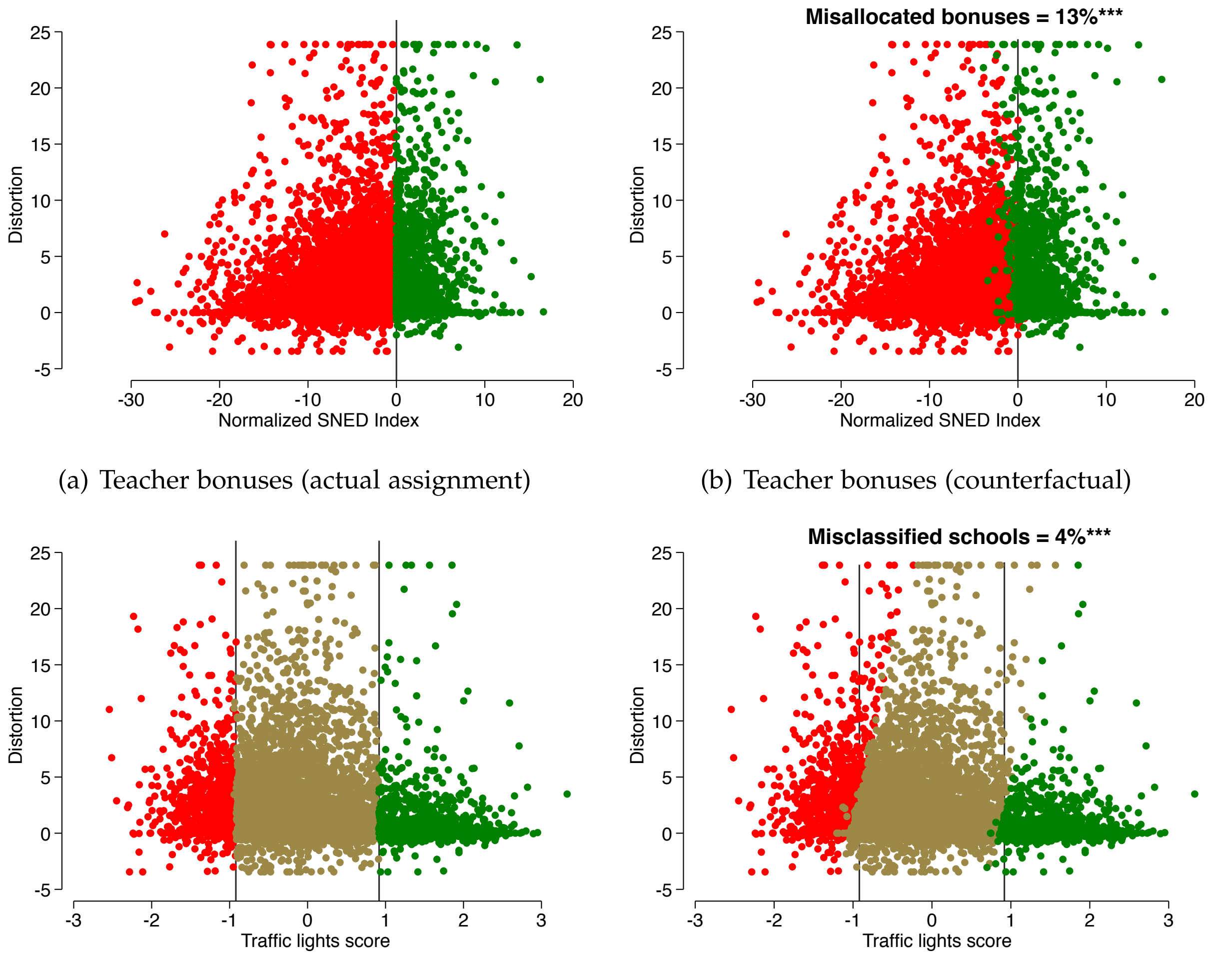
(b) Distribution of distortions in quality signals. Each observation is a school in a specific year between 2005 and 2013.

Figure 2.3: Distortions and attendance in test days



Notes: This figure displays the differential test-day attendance of students above the 75*th* percentile and below the 25*th* percentile of the GPA distribution (*x*-axis, in percentage points) and distortions in quality signals (*y*-axis, in test score points). We include all schools in 2013. The coefficients (robust standard errors) of a linear regression of distortions on a linear and quadratic term of differential changes in test-day attendance are 4.38 (0.36) and 3.96 (1.19) respectively.

Figure 2.4: Misallocation of public programs



(a) Teacher bonuses (actual assignment)

(b) Teacher bonuses (counterfactual)

(c) Information (actual provision)

(d) Information (counterfactual)

Notes: In panels (a) and (b) we plot school distortions (y-axis), school scores to assign teacher bonuses (x-axis), and the threshold of the assignment (red schools did not get bonuses, green schools did get bonuses) using the actual and counterfactual quality signals. In panels (c) and (d) we plot school distortions (y-axis), school scores (x-axis), and their actual and counterfactual categories (red, yellow, and green).

Table 2.1: Descriptive statistics

	Observations	Mean	St. dev.	p10	p50	p90
A – Schools (2005-13)						
Test score (SIMCE)	38,416	254.8	27.7	219.5	254.0	292.5
Students in $4th$ grade	38,616	50.4	35.5	17.0	40.0	91.0
Students absent in test days	38,616	3.7	4.5	0.0	3.0	8.0
Class size	38,609	30.4	8.0	19.4	31.0	40.3
Average annual attendance	38,616	93.3	3.1	89.6	93.6	96.7
Students in $1st$–$8th$ grades	38,616	415.5	283.8	143.0	335.0	748.00
Public	38,616	0.39	0.49	0.0	0.0	1.0
Voucher	38,616	0.52	0.50	0.0	1.0	1.0
Private	38,616	0.09	0.28	0.0	0.0	0.0
Religious	37,401	0.44	0.50	0.0	0.0	1.0
Monthly fee (U.S. dollars)	38,341	48.46	92.3	0.0	0.0	182.1
Distortion in test score	60,813	2.7	4.2	0.0	1.1	7.7
B – Students (2013)						
Test score (SIMCE)	140,982	263	46	200	267	321
GPA	159,356	5.9	0.6	5.1	5.9	6.5
Attendance in test-day	137,604	0.95	0.20	1.0	1.0	1.0
Attendance in non-test days	137,127	0.92	0.17	0.8	1.0	1.0

Notes: Own construction based on administrative data provided by the Ministry of Education. We restrict the data to schools with zero distortion or with sufficient data to calculate it. Distortions are measured in test score points and we estimated them using the methodology described in section 2.4.1. See Figure B.2 for a timeline of standardized tests. See section 2.4 for details. There are 8,254 schools in the period 2005–2013.

Table 2.2: Understanding distortions

Dependent variable: distortions in quality signals (in test score points)

	All		Distortions> 0	
A – School attributes	(1)	(2)	(3)	(4)
Public	1.55***	1.31***	0.70**	0.57*
	(0.13)	(0.14)	(0.31)	(0.33)
Religious	0.03	-0.11	0.17	0.03
	(0.08)	(0.08)	(0.17)	(0.20)
For-profit	0.28**	0.36***	0.76**	0.91***
	(0.11)	(0.11)	(0.31)	(0.32)
Log parents income	-0.78***	-0.68***	-0.82***	-0.85***
	(0.04)	(0.04)	(0.11)	(0.12)
Average annual attendance	-0.17***	-0.20***	-0.32***	-0.30**
	(0.04)	(0.05)	(0.09)	(0.12)
Students in $4th$ grade	-0.12	-0.11	-2.30***	-2.22***
	(0.16)	(0.15)	(0.25)	(0.26)
Enrollment in grades $1st$-$8th$	-0.51***	-0.51***	-0.11	0.15
	(0.16)	(0.16)	(0.25)	(0.27)
Constant	1.77***	1.95***	6.25***	6.34***
	(0.12)	(0.13)	(0.31)	(0.33)
B – Autocorrelation				
Lagged distortion	0.41***	0.38***	0.39***	0.37***
	(0.01)	(0.02)	(0.03)	(0.03)
Constant	1.97***	2.06***	6.25***	6.30***
	(0.04)	(0.05)	(0.13)	(0.14)
Mean of dep. variable	2.18	2.18	5.11	5.11
Market-year F.E.	No	Yes	No	Yes
Variance explained by schools F.E.	0.36	0.36	0.60	0.60
Schools	3,417	3,417	2,339	2,339
Observations	29,588	29,579	5,929	5,927

Notes: Estimation includes all urban schools. All non-indicator variables have been normalized (except for lagged distortion). All regressions are weighted by the inverse of the uncertainty associated to the calculation of distortions, where uncertainty is the size of the confidence interval. Columns 3-4 restrict the data to school-year observations with distortions statistically different from zero. Robust standard errors in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Table 2.3: Summary statistics for estimation of school choice model

Variable		Mean	St. dev.	p10	p50	p90
Students	In sample	1,009	844	324	665	2,446
	Coverage rate	0.33	0.12	0.17	0.31	0.48
Schools	In sample	63	62	19	45	134
	Coverage rate	0.92	0.13	0.72	0.97	1.00
Poor students	In sample	479	391	166	323	1,184
	Sample share	0.49	0.11	0.35	0.50	0.60

Notes: This table displays market-level summary statistics for the sample we use to estimate the school choice model. This sample includes 25 markets in the period 2011–2014. For the number of students and schools per market, we provide summary statistics in levels and coverage rate of the complete market. For the number of poor students, we provide summary statistics of levels and their share over the sample market size.

Table 2.4: IV results from the second stage of school choice model

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	All			Poor students			Non-poor students		
Fee	-0.003***	-0.004***	-0.006***	-0.006***	-0.007***	-0.010***	-0.002***	-0.003***	-0.005***
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)
Quality	0.012***	0.019***	0.019***	0.004**	0.011***	0.015***	0.021***	0.028***	0.023***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Religious			-0.054**			-0.086***			-0.019
			(0.024)			(0.030)			(0.029)
Gender constraint			0.148***			0.121**			0.161***
			(0.047)			(0.059)			(0.055)
Public			0.089***			0.229***			-0.061
			(0.031)			(0.040)			(0.039)
SEP school			-0.325***			-0.533***			-0.587***
			(0.065)			(0.094)			(0.066)
Market-year F.E.	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Observations	10,774	10,774	10,774	5,335	5,335	5,335	5,439	5,439	5,439
First stage tests									
F-test fee	1566.15	2031.73	395.15	484.55	582.07	73.91	1285.81	1593.35	329.15
F-test quality	70.62	17.69	15.03	33.81	9.68	8.17	36.50	8.03	6.86
Cragg-Donald EV	283.97	232.38	203.91	146.05	127.12	101.57	139.36	106.47	98.52

Notes: Instrumental variable estimates. We use two sets of instruments: (i) the amount awarded by school vouchers, mean fixed characteristics of rivals in the market (i.e. BLP instruments) and rivals market wages are used as instruments for schools fees; and (ii) a linear and quadratic term on county-specific temperature and the residual of a regression of being awarded a SNED prize in the previous period on lagged school quality are use as instruments for school quality. F-tests are computed separately for each first stage for the respectively excluded instruments. All regressions are weighted by school enrollment. Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 2.5: Means of predicted school attributes of households choices

		(1)	(2)	(3)	(4)
		Poor students		Non-poor students	
Attribute	Scenario	Average	Switchers	Average	Switchers
Distance	Baseline	2.36	2.00	2.58	2.2
(in kilometers)	Counterfactual	2.36	2.07	2.59	2.31
	Change	0.00	0.07	0.00	0.11
Fee	Baseline	17.08	6.58	71.43	39.02
(in U.S. dollars)	Counterfactual	17.52	22.52	72.89	81.95
	Change	0.43	15.95	1.47	42.93
Quality	Baseline	254.77	242.72	267.13	252.62
(in test score points)	Counterfactual	255.25	260.15	267.78	271.53
	Change	0.48	17.43	0.65	18.91

Notes: Columns 1 and 3 (2 and 4) display the average attributes of chosen schools for poor and non-poor households (poor and non-poor switchers). Results for distance are measured in kilometers, results for school fees are measured in US dollars and results for quality are measured in SIMCE test scores, net of distortions.

Table 2.6: Yearly welfare gains of a policy that provides undistorted quality signals

	(1)	(2)	(3)	(4)	(6)	(6)
	Poor students			Non-poor students		
Comparison	Average $E[\Delta W_i]$	Switchers $E[\Delta W_i]$	Switch rate	Average $E[\Delta W_i]$	Switchers $E[\Delta W_i]$	Switch rate
Counterfactual scenario	\$1.73	\$53.2	3.25%	\$5.29	\$173.94	3.04%
Poor households with non-poor quality preferences	\$4.13	\$100.94	4.10%	-	-	-
Poor households with non-poor preferences	\$6.91	\$181.37	3.81%	-	-	-
Poor households with non-poor market opportunities	\$1.55	\$ 64.95	2.39%	-	-	-

Notes: Welfare results are measured in U.S. dollars per year. Columns 1 and 3 display average welfare gains for poor and non-poor households. Columns 2 and 4 display average welfare gains for poor and non-poor switchers. Columns 3 and 6 display the share of switchers for poor and non-poor households respectively.

Chapter 3

Losing Your Dictator

3.1 Introduction

Political transitions are ubiquitous and are associated with significant changes in the economy, but little is known about the behavior of key economic actors during these times.[1] This knowledge would be crucial in understanding the sustainability of democracy, a result which, as the Arab Spring has reminded us recently, is far from guaranteed. By studying firms during political transition, we can also improve our understanding about the relationship between political power and the distribution of economic power.

Scholars have long argued that firms with links to a non-democratic regime benefit from a number of distortions, such as corruption in procurement, preferential lending, and preferential access to information.[2] The anticipation that these distortions will disappear if there is a regime change could lead firms to "prepare" for the new state of the world. If firms successfully prepare, this anticipation might affect market structures even well after the regime change. Firms would be able to transfer distortions across political regimes, limiting the benefits of a democratization and the market changes it creates. However, observing how firms adjust their inputs to better position themselves under the new regime is typically difficult.

In this paper, we exploit an unexpected democratization announcement and the subsequent peaceful transition from dictatorship to democracy in Chile to study how firms in the dictator's network prepare for a future democratic period. Our findings indicate

[1] There have been, on average, 3.7 (2.0) transitions to (out of) democracy per year in the last 25 years (see Appendix Figure C.1). A large literature in economics and political science study the effects of political regimes on economic growth. See Acemoglu et al. (2015) for a discussion of the literature and the more recent estimates and Papaioannou and Siourounis (2008) for an empirical assessment of democratization theories.

[2] The empirical literature in economics can be traced to Fisman (2001), who shows that firms close to Indonesia's dictator Suharto experienced substantial decreases in their value when he suffered heart attacks, a source of a potential regime change.

that firms close to the old regime increased their productive capacity during the transition to democracy, and these investments improved their market position in the new regime. We provide suggestive evidence that these differential adjustments in capacity were made possible by government banks during the dictatorship.

Chile's transition to democracy provides a unique opportunity to measure and study the interactions between a dictatorship and firms. Vast amounts of (previously unexploited) information exists about firms that operated during and after Augusto Pinochet's dictatorship (1973–1990). The existence of records with information about individuals working for Pinochet assures that interactions between the state and firms are measurable. The timing of this political transition gives us an opportunity to measure firm responses after a democratization announcement but before the new democratically elected government takes office. We exploit all these features when analyzing capital, labor, productivity, and profits around Chile's political transformation to show how firms shaped during the Pinochet regime moved towards the new democratic era.

After fifteen years in power, Augusto Pinochet called for elections in 1988, where he would run as the unique candidate to transform his autocratic regime into a democratic one for the next eight years. Contrary to everyone's expectations, Pinochet not only lost the election, but also acknowledged his defeat. Pinochet's defeat at this election, known the "1988 plebiscite," marked the beginning of Chile's transition to democracy. The plebiscite's outcome changed the post-1990 years from a Pinochet regime to a democracy.

To motivate our analysis of firms, we first collect daily stock prices around the 1988 plebiscite to investigate how financial investors reacted. We use the board of directors to construct Pinochet's network of firms before the plebiscite, and we document that firms with first *or* second degree connections to Pinochet suffered a substantial decrease in abnormal returns in the days that followed. Although decreases in stock prices of connected firms after negative political events have been documented before, there are two surprising facts about our findings. First, firms with *second* degree connections also suffered a substantial decrease in their stock prices. This finding suggests the existence of more complex political networks than in previous research. Second, the value of connected firms *increases* twelve weeks after the plebiscite. Moreover, the value of firms with first degree connections increases by more than the value of firms with second degree connections. This suggests that (1) connected firms react to political events in ways not anticipated by financial investors, and (2) firms with different degrees of connectedness react differently. We discuss alternative interpretations that are hard to reconcile with the data.

To understand how firms reacted to the defeat of Pinochet at the plebiscite—effectively a democratization announcement—we build a simple model of firms making decisions during a political transition. As we observed a decrease in stock prices of firms with first and second degree connections, we incorporate three types of firms into the model: unconnected, connected (first degree), and indirectly connected (second degree). Because the credit market has been shown to be a source of advantage for connected firms, an insight particularly relevant in the Chilean case, we model links between firms and the government as a parameter that determines the cost of capital. We interpret the cost of capital as a sufficient statistic for credit market relations, and we study the credit market empirically in the last part of our analysis.

In the model, there are different political periods. During a first period of dictatorship, firms compete *à la* Cournot, using their political connections to determine their productive capacities. After an unexpected democratization announcement, firms learn that they will lose their connection—after a period of transition where connections are still in place—and that the new democratic period triggers an increase in firm entry, an empirical fact in our context consistent with Acemoglu (2008). Consequently, competition turns into a Stackelberg game, where incumbents exploit their connections to improve their market position in the new democratic era. The primary insight we take from the model, therefore, is that increasing productive capacity becomes a dominant strategy for incumbent firms with political connections after a democratization is announced.

Empirically, we study changes in productive capacity using a differences-in-differences econometric strategy with three types of firms and three political periods, i.e., dictatorship (five years), transition (one-and-a-half years), and democracy (five years). We construct a panel dataset of publicly listed firms observed between 1985 and 1994 using two sources of information: (1) quarterly data from balance sheets, and (2) data from annual statements. Our main findings exploit within firm variation over time and indicate that firms with first degree political connections increased their productive capacity by 0.4 standard deviations after the democratization announcement, with no changes in either productivity or labor. These results account for any heterogeneous effects the political transition might have had across industries (and other observable variables) and are robust to a wide range of specification and robustness checks. Moreover, a comparison of the magnitude of results between firms with first and second degree connections provides additional support for the insights we obtained from the model. This increase in productive capacity provides one explanation for the short- and long-run movements in stock prices after the plebiscite. We discuss several alternative interpretations that are not supported by the data.

We provide evidence that government banks are key to understand the investment patterns we document. This mechanism is consistent with corruption cases that involve loans from the largest public bank to politically connected firms (e.g., Leon-Dermota 2003). To study the credit market, we make use of data we have digitized from annual statements. Following Khwaja and Mian (2005), we exploit time variation in a panel dataset of debt between firms and two types of banks: government-owned and others. The former set of banks is composed of banks in which the Pinochet regime had more decision power. Consistent with the mechanism in our model, our findings indicate that connected firms obtained substantially more loans from government-owned banks after the democratization announcement but before the new democratically elected government took office. This result confirms that political connections create distortions in the credit market and suggests one mechanism through which firms can transmit their economic power across political regimes.

In our final analysis, we explore the consequences of increasing capacity after the democratization announcement. Our theoretical framework suggests that firms with links to the old regime increased their capacity to improve their economic outcomes under the new political regime. Whether these connections indeed provided firms with a long-term advantage is an empirical question. We analyze two outcomes: survival and profitability. We show that connected firms were 15 percentage points more likely to remain operating in the market *after* our period of analysis. We also determined that firms that increased their capacity after the democratization announcement experienced an increase in profits of approximately 0.2 standard deviations, a return larger than business as usual.

The main contribution of this paper is to show how economic power and distortions can be transferred from dictatorship to democracy, an important and understudied issue in a world with four democratizations every year. Because elections in authoritarian regimes have become common in recent years (Lindberg, 2009), our results are of general interest for countries where a democratic transition is likely to occur. In particular, our findings suggest that regulation of the credit market might be necessary to minimize the transfer of distortions from dictatorship to democracy.

Our paper is related to literature in several fields. First, and most importantly, our work contributes to the political economy literature studying (1) the effects of political transitions (e.g., Acemoglu 2008, Acemoglu et al. 2011, 2015), (2) the legacies of non-democratic regimes (e.g., O'Donnell and Schmitter 1986, Martínez Bravo 2014), and (3) rent extraction (e.g., Brough and Kimenyi 1986, Burgess et al. 2015, Mocetti 2016). Second, our analysis of capacity responses across firms is related to a literature in industrial organization studying strategic investments to achieve entry deterrence (e.g., Dixit 1980).[3] Third, by showing that connected firms are not only relatively unproductive, but also more likely to survive Chile's transition to democracy, our results highlight a new dimension of inefficiency arising from smooth political transitions (see Roland 2002 for a survey).

Methodologically, our work uses tools from different literature. We consider a network approach with multiple degrees in our analysis of firms during political transition, something motivated by movements in stock prices and the interlocking directors literature in corporate finance (e.g., Haunschild and Beckman 1998). In addition, by using stock prices to motivate both the network analysis and the study of firm reactions to a democratization announcement, our work relates to the literature estimating the value of political connections (e.g., Fisman 2001, Dube et al. 2011, Acemoglu et al. 2016a,b). Lastly, we take insights from the literature linking political connections and firm-level variables to motivate the study of mechanisms (e.g., Khwaja and Mian 2005, Do et al. 2015).

The next section presents the main features of the context under study, including the historical background and our methodology to construct links between Pinochet and

[3]Consistent with this literature, we find that firm entry is lower during the democratic period in industries with a higher share of connected firms before the democratization took place. However, this result needs to be interpreted with caution, because we observe a limited number of industries in our dataset.

firms. In section 3.3, we present comparative statics of a simple model of firms making decisions around a political transition. In section 3.4, we discuss our main dataset and empirical strategy. In section 3.5, we present our main set of results with a large set of robustness checks and a discussion of identification. In section 3.6, we analyze the credit market as a mechanism behind our main result. In section 3.7, we study firm outcomes in the democratic period. Finally, in section 3.8, we offer some conclusions.

3.2 Chile's transition to democracy

In this section, we provide historical details about Augusto Pinochet's rise to power, the political foundations of his economic policy, and Chile's transition to democracy. In addition, we discuss our methodology to measure Pinochet's network of firms, and estimate how the political transition affected stock prices in the short- and long-run.

3.2.1 The Pinochet dictatorship

Augusto Pinochet's dictatorship began in Chile after a coup d'état against democratically elected socialist Salvador Allende on September 11*th*, 1973.[4] After the coup, Pinochet was part of a military *junta* that ruled the country from September of 1973 until June of 1974. During these months, the *junta* needed to fill positions in the government and decide which policies to implement. There was, however, a tremendous amount of disorganization (Cavallo et al., 2011). These months were crucial for the right wing to negotiate the economic policies that would be be implemented in the following years. In a context of macroeconomic imbalances, economists trained at the University of Chicago (known as Chicago Boys) offered a plan to control inflation, cut government spending, and privatize the economy. After consolidating his power at the military *junta*, Pinochet agreed to follow their recommendations. Pinochet ruled the country until March of 1990.

Following Silva (1996), we can divide Pinochet's dictatorship into three periods: installation and repression (1973–1975), implementation of radical economic policies (1976–1982), and implementation of pragmatic policies (1983–1989). In the first period, Pinochet's primary objective was to restore internal and external balance. To achieve this goal, domestic commodity and financial markets were liberalized, government spending was drastically reduced, and government assets were sold. In the second period, tariffs were uniformly reduced to 10 percent, and the exchange rate was fixed to be the main instrument in stabilizing the economy. The fixed exchange rate was controversial and has been linked to the 1982 economic crisis that hit the Chilean economy, in which gross domestic product decreased 14 percent. In the years before the crisis, inflation decreased, and the growth of real GDP averaged 7 percent (see Appendix Figures C.2 and C.3 for macroeconomic indicators during this period). In the third and final period,

[4]This event took place in the context of the Cold War. For the Western Block, Allende's government was an ally of the Soviet Union, and thus a threat for its coalition.

the fixed exchange rate system was abandoned, reforms were partially reversed, and the economic policies implemented were less radical than in the previous period.

3.2.2 Pinochet's network of firms

Although many researchers have discussed the macroeconomic reforms implemented during Augusto Pinochet's dictatorship (e.g., Ramos 1980, Corbo 1985, Edwards 1986), very few have researched the practices of firms during this period (Ossandón and Tironi, 2013). As our analysis exploits variation in the network of firms operating during Pinochet's dictatorship, we now provide historical details about firms and our methodology to measure Pinochet's network.

Besides the macroeconomic reforms implemented, for which we account explicitly in our empirical analysis, the most controversial aspect of the dictatorship's economic policy was the privatization process. The controversy relies on the fact that Pinochet's allies bought profitable firms at prices lower than market prices, effectively transferring wealth from the State to a group of politically connected individuals (Mönckeberg, 2015). During these processes, individuals in Pinochet's social network began working on the boards of directors at several different firms. Pinochet's former son-in-law—Julio Ponce Lerou, who, according to Forbes, is now one of the richest people in the world, with a wealth of $3.3 billion—is perhaps the most famous case. Ponce Lerou, currently the principal shareholder of the Chemical and Mining Society of Chile, the world's largest producer of potassium nitrate, iodine, and lithium, entered the board of directors in the process of a privatization.[5]

Despite the commonly acknowledged existence of a network of firms close to Pinochet during his regime, our paper is the first comprehensive attempt to measure it. Empirically, we conducted three steps to construct Pinochet's network of firms. First, we gathered data on the universe working directors from financial statements—name and unique national ID, approximately 10 per firm. Second, we performed a Google search of those working in 1987 and classified them as being politically connected if (1) he/she worked for Pinochet's dictatorship before 1988, or (2) was a member of Pinochet's close family.[6] Using this procedure, 10 percent of director positions were classified as politically connected. Third, we classified a firm as politically connected if at least one member of its board of directors had a political connection.[7]

[5]Importantly for our empirical analysis, not all firms privatized during this period are classified as part of Pinochet's network, and not all firms in Pinochet's network were privatized during his regime. This means we can control explicitly for the privatization process in our analysis.

[6]Other papers have classified individuals as politically connected in a similar way (e.g., Fisman 2001, Bertrand et al. 2007, Acemoglu et al. 2016a,b).

[7]An example of a politically connected firm is presented in Table C.1. Investigation of the employment history of directors before 1988 is possible due to the vast amount of information available online about individuals closely related to Pinochet's dictatorship. We examined this information using Google as an oracle for standardized queries. In particular, we performed searches in *incognito* mode to avoid personalized searches and facilitate replication. More precisely, we look in the first page of results using

In addition to *direct* political connections, we say a firm has an *indirect* political connection if none of its directors was connected, but at least one also worked for a connected firm in 1987. We decided to consider indirectly connected firms as a separate category in order to avoid "contaminating" our control group with somewhat connected firms. Contamination is more likely if indirect connections matter. In this paper, indirect connections are defined using interlocking directors (i.e., same director in different firms), which has been shown to affect firm outcomes through an information mechanism (e.g., Khwaja et al. 2011, Patnam 2013, and Fracassi 2014). Overall, out of the 118 firms in our dataset, a total of 43 (36%) had a *direct* political connection, 33 (28%) had an *indirect* political connection, and 42 (36%) were unconnected.[8] This network of firms was affected by Chile's transition to democracy, a process that began after the 1988 plebiscite.

3.2.3 The 1988 plebiscite

In 1988, and as stated in the 1980 Constitution, Pinochet called for elections in which he would run as the only candidate—a Yes/No option known as the "1988 plebiscite" that took place on October 5*th*. Pinochet's goal was to internationally validate his regime and become president of Chile for the period 1988–1996. He did not accomplish his goal, however, because he lost. In an election in which more than 90 percent of the voting-age population registered to vote, 56 percent rejected Pinochet's continuation. Then, in December of 1989, a presidential election with candidates from all parties took place, an election in which Pinochet could not run. As expected, the opposition won, and the new democratically elected president Patricio Aylwin took office in March of 1990. Importantly, between the plebiscite and the arrival of the new government, seventeen months transpired in which firms could adjust their decisions for the new economic environment. We call this period "lame duck," as it was known that Pinochet would leave office, and it is a crucial part of our analysis.

Pinochet's defeat at the plebiscite was unexpected for several reasons. First, there was no legal institution in charge of regulating the election. Second, previous surveys did not state a clear prediction (Cauce, 1988). Third, most people believed that Pinochet

two different queries: (1) director's full name, and (2) director's full name + Pinochet. Empirically, there is a large set of reports documenting the name and specific job that all individuals with power performed during the dictatorship (e.g., "Los 100 rostros de la dictadura" and "Memoria Viva"). In addition, heads of government departments and some army officers are extremely well known and, consequently, have an employment history that is easy to track. Overall, we found that 78 directors were politically connected to Pinochet in 1987. These directors had different jobs in the dictatorship: 22 were army officers working directly for Pinochet, 9 were close (economic or legal) advisors, 24 worked as head of government departments (ministers), 19 worked as politicians (e.g. local politicians), 2 were Pinochet's sons-in-law, and 2 collaborated with money and press.

[8] This strategy of distinguishing between direct and indirect political connections is novel and did *not* drive our main results. It did, however, increase the precision of our estimates. Nevertheless, in what follows, we present results using only direct political connections for completeness and to facilitate comparison with the literature.

was not going to acknowledge a negative result.[9] And fourth, on election day, most preliminary results showed that Pinochet was winning, and the opposition's victory was only recognized on October 6 at around 2 a.m. (Méndez et al., 1988; Engel and Venetoulias, 1992).

In addition to this historical evidence, we provide empirical evidence for the unexpectedness of the plebiscite's outcome by analyzing stock market returns of firms with and without political connections to Pinochet's dictatorship. To conduct this analysis, we hand-collected data on daily stock market prices from contemporary newspaper *El Mercurio*, publicly available at Chile's National Library.

3.2.4 Short- and long-run effects in stock prices

Following the political connections literature, we restrict attention to firms that were traded four months before the plebiscite to analyze abnormal returns—i.e., the difference between returns and expected returns. This reduced our data to 80 firms. Figure 3.1-A shows a significant drop in (abnormal) stock returns for firms politically connected to Pinochet. This drop corresponds to a decrease of three standard deviations and is similar for firms with direct and indirect connections. Figure 3.1-B shows that this effect persisted for twelve days. This drop was unique to the plebiscite, since there were no significant heterogeneous returns between connected and unconnected firms around other major political events (see Appendix Figure C.4). We interpret this result as (1) evidence that the plebiscite's outcome was unexpected, and (2) validation of our connection measure.[10]

The stock market differences we observe between firms with and without political connections are in line with most findings in the literature. How long-lasting are these effects? Theoretically, stock prices should remain low if the event under study is unexpected and the present value of future cash flows is permanently lower. Nevertheless, actions in the aftermath of the event could have easily reversed the initial drop in stock prices. To analyze how permanent the effect was in our case, Figure 3.1-C plots the weekly price relative to the price one week before the plebiscite. The initial drop lasted only around twelve weeks, which suggests that once firms lose their dictator, there is a change in the observed behavior by investors such that stock prices revert to their pre-event levels.[11] This empirical fact also serves as motivation to understand how firms respond after learning a political transition will occur.

---

[9] According to declassified documents posted by the U.S. National Security Archive, Pinochet stated, "I'm not leaving power, no matter what." Different political forces (including the navy) pushed him to finally accept the result (Huneeus, 2006).

[10] See the appendix for a network graph of firms in 1987 (Figure C.5), and regression results (Table C.2). Abnormal returns of stock i in day t are defined as: $AR_{it} \equiv R_{it} - (\hat{\alpha}_i + \hat{\beta}_i R_{mt})$, where R_{mt} is the return on the market for day t.

[11] These patterns could also be consistent with an overreaction of investors. We argue this is unlikely to be the case because (1) there are no significant differences in stock prices during other important political events, and (2) the observed heterogeneity is hard to reconcile with an overreaction mechanism.

A final remark is important. A political transition dramatically changes the economic environment (e.g., Acemoglu and Robinson 2009). Empirically, we observe a sharp increase in the number of firms operating after the new democratic government takes office (Figure 3.1-D). Many potential explanations could explain this empirical fact. As we observe a single political transition, we do not attempt to provide a full explanation. It is, however, interesting to note that industrial policies implemented during the late period of Pinochet's dictatorship and the beginning of the democratic period were fairly similar.[12] This similarity in policies implies that the increase in firm entry was less likely to be driven by endogenous government actions (e.g., Acemoglu 2008). Accordingly, we take this increase in firm entry as an (exogenous) effect of democracy.

3.3 Theoretical framework

How do firms react to an announced political transition, and how does that vary with a firm's degree of connection to the incumbent regime? This section presents a theoretical framework to answer these questions. There are two key assumptions in our model. First, firms close to the non-democratic regime enjoy differential access to finance, which disappears after a democratization. Second, there is an increase in firm entry during the democratic period.[13] The main insight we obtain is that increasing productive capacity becomes a dominant strategy for connected firms after a democratization announcement.

3.3.1 Environment

Let there be N_t incumbent firms and three different periods $t = 1, 2, 3$. In the first period, a dictator is in power and N_1 firms operate in the market. In the second period, all firms learn that a new democratic government will take office in the third period. Following our setting, we assume this is an unanticipated democratization announcement. As it is public knowledge that the dictator will leave office, we call this period "lame duck." Potential entrants also learn about the political transition and update their entry decisions accordingly. We call the third period "democracy," where a newly democratically elected government rules the country and new firms enter the market.

---

[12] Some features of the transition explain this: (1) army officers could not be removed from their positions, which explains the credible threat of political unrest emphasized by Ellman and Wantchekon (2000); (2) there were former army officers designated in the Congress, and they could not be removed; and (3) the National Security Council, in which Pinochet participated, was capable of blocking the President's actions. The 1989 statement of an unconnected firm in our data puts it succinctly: "Considering the political changes in the country, and given we do *not* expect significant economic changes, operations should continue as usual" (our italics).

[13] Khwaja and Mian (2005) show that banks controlled by the government tend to favor politically connected firms when allocating loans, while Claessens et al. (2008) show that firms connected to politicians have a higher access to bank finance. The change in firm entry is motivated by both theoretical research (e.g., Acemoglu 2008) and empirical patterns observed in our data.

In periods 1 and 2, firms can have different degrees of connections to the dictator. No firm has connections to the democratic government in period 3. Let connections of firm i be represented by $\gamma_i \in [0, r]$, where $\gamma_i = 0$ represents no connection, and $\gamma_i > 0$ some connection between a firm and the dictatorship. In each period, firms compete à la Cournot by choosing their input subject to a given private demand $Q_t = a - bP_t$. Let the production technology be $q_t^i = K_t^i$, where K_t^i is the stock of capital of firm i in period t. The marginal cost of producing one extra unit is zero if production is below a firm's capacity, and infinite otherwise.

The cost of capital for firms is $R_i \equiv r - \gamma_i$. We interpret this lower cost of capital as the combination of two factors: (1) connected firms have more access to credit, and (2) connected firms have relatively better information about investment opportunities. We will discuss how we can place bounds on these mechanisms exploiting the network analysis.[14]

3.3.2 Timing

In period 1, there are N_1 firms competing in quantities. Private demand for the homogeneous product is fixed. Firm i chooses K_1^i to maximize the discounted present value of profits, and expect the dictator to be ruling indefinitely. Then, a firm's problem is:

$$\max_{K_1^i} \ \Pi_1^i \ = \ \frac{1}{1-\delta}\left[b^{-1}\left(a - \sum_j^{N_1} K_1^j\right) K_1^i - R_i K_1^i\right] \tag{3.1}$$

where $\delta \in (0, 1)$ represents the discount factor. Let the term in square brackets be denoted by $\Omega(K_1^i, N_1|\gamma_i)$ to facilitate exposition. For simplicity, let $\gamma_i \in \{0, \bar{\gamma}, r\}$, with $\bar{\gamma} \in (0, r)$. Then, there are three types of firms: directly connected (C: $\gamma_i = r$), indirectly connected (I: $\gamma_i = \bar{\gamma}$), and unconnected (U: $\gamma_i = 0$). Thus, $i \in \{C, I, U\}$.[15]

Period 1: Nash equilibrium in dictatorship

During dictatorship, firms compete *à la* Cournot by setting their productive capacities. Before the political transition is announced, firms operate in equilibrium, i.e., productive capacities have reached a steady state. To find this (Nash) equilibrium, we obtain firms best response function. Then, we use these best response functions to solve for

[14] Non-democratic regimes might also shift demand towards connected firms (e.g., Cingano and Pinotti 2013). Although this is an unobservable dimension in our setting, it implies capacity should *decrease* among connected firms after a democratization announcement, attenuating our estimates.

[15] Note that firms do not expect a political transition to take place in the foreseeable future. Then, assuming that the free entry condition is binding, we can interpret this period as a steady state.

productive capacities. Equilibrium capacities are:

$$K_1^C = \frac{a + b(2r - \bar{\gamma})}{4}$$

$$K_1^I = \frac{a - b(2r - 3\bar{\gamma})}{4}$$

$$K_1^U = \frac{a - b(2r + \bar{\gamma})}{4}$$

Total quantity offered in the market is defined as:

$$Q_1^* = \sum_{i \in \{C,I,U\}}^{N_1} K_1^i$$

Note that there are N_1 incumbent firms in dictatorship, and each one is either connected (C), indirectly connected (I), or unconnected (U). The equilibrium price is determined by the aggregate demand function, and profits are computed as in equation (3.3).

Period 2: democratization announcement

In period 2, firms learn that a democratization will take place with certainty. This means connected firms will lose their political connections. Specifically, we assume:

Assumption 1. *There is an exogenous democratization announcement.*

Assumption 2. *It is common knowledge that firm entry is exogenously higher in democracy.*

Firms can adjust their productive capacity in the second period, when connections are still in place. This could be a firm's optimal response in order to compete with new entrants. The new equilibrium is similar to the equilibrium of a Stackelberg game. In our case, incumbent firms are first movers and entrants are followers. Then, firms internalize future competition and solve the following problem:

$$\max_{K_2^i} \Pi_2^i = \Omega(K_2^i, N_2|\gamma_i) + \frac{\delta}{(1-\delta)}\Omega(K_2^i, N_3|\gamma_i = 0) \tag{3.2}$$

where $N_2 = N_1$ and N_3 is the number of firms operating in the democratic period.

In period 3, all connections disappear (i.e., $\gamma_i = 0 \ \forall i$), and production and entry decisions are decided as a function of the actions taken by incumbent firms in period 2. Former incumbent firms face the same objective function.

After an exogenous democratization announcement, incumbent firms expect new firms to enter the market. The number of entrants is exogenous. However, before new entry occurs during the democratic period, incumbent firms can adjust their quantities.

We now solve for the Stackelberg game played by incumbent firms and entrants. In particular, we obtain productive capacities and prices for period 2. In order to do this, we first need to define the objective function of entrants (E):

$$\max_{K_2^E} \ \Pi \ = \ \frac{1}{1-\delta}\left[b^{-1}(a - \sum_{i\in\{C,I,U\}}^{N_3} K_2^i)K_2^i - rK_2^E\right] - F \tag{3.3}$$

where F represents the fix entry cost of entering the market. Then, equilibrium capacities for the period after the democratization announcement are defined by:

$$\begin{aligned}
K_2^C &= \frac{1}{4}\left(a + b\left(r + \frac{2(1+\psi)(r-\bar{\gamma}) + 2\psi r}{2+\psi}\right)\right) \\
K_2^I &= \frac{1}{4}\left(a + b\left(r - \frac{6(1+\psi)(r-\bar{\gamma}) - 2\psi r}{2+\psi}\right)\right) \\
K_2^U &= \frac{1}{4}\left(a + b\left(-3r + \frac{2(1+\psi)(r-\bar{\gamma}) - \psi r}{2+\psi}\right)\right) \\
K_2^E &= \frac{1}{8}\left(a + b\left(-3r + \frac{2(1+\psi)(r-\bar{\gamma}) - \psi r}{2+\psi}\right)\right)
\end{aligned}$$

where $\psi \equiv \delta(1-\delta)^{-1}$. Then, total quantity offered in the market during the lame duck and democratic periods are defined respectively by:

$$\begin{aligned}
Q_2^* &= \sum_{i\in\{C,I,U\}}^{N_2} K_2^i \\
Q_3^* &= \sum_{i\in\{C,I,U,E\}}^{N_3} K_3^i
\end{aligned}$$

where note that there are $N_2 = N_1$ incumbent firms in the second period, and N_3 incumbent firms in the third period. Finally, the equilibrium price is determined by the aggregate demand function, and profits are computed as in equation (3.3).

3.3.3 Comparative statics

Let K_t^i —the solution of the game— denote the capital stock of firm i during period t. Then:

Proposition 3.3.1. *Under assumptions 1 and 2 capital adjustment is a dominant strategy. Firms* *with different degrees of political connections adjust differently:*

$$K_2^C > K_1^C \quad ; \quad K_2^I \lesseqqgtr K_1^I \quad ; \quad K_2^U < K_1^U$$

Exists $\gamma^* \in (0,1)$ s.t. if $\bar{\gamma} > \gamma^*$ then $K_2^I > K_1^I$, if $\bar{\gamma} < \gamma^*$ then $K_2^I < K_1^I$, and if $\bar{\gamma} = \gamma^*$ then $K_2^I = K_1^I$.

Proof: *Using the equilibrium capacities from the previous section, we can compare how capacity changes between periods for firms with different degrees of political connections. We begin by calculating the difference for firms with first degree connections:*

$$K_2^C - K_1^C \quad = \quad \frac{\psi b}{4(2+\psi)}(3r - \bar{\gamma})$$

Note that if $\delta = 0$, we have that $\psi = 0$, then $K_2^C = K_1^C$. Therefore, if $\delta > 0$, and $\bar{\gamma} \leq r$, we have that $(3r - \bar{\gamma}) > 0$. This means that $K_2^C > K_1^C$. In the case of firms with second degree connections, this inequality becomes ambiguous. To see this more clearly, let us subtract the equilibrium capacities for these firms in the two periods of interest:

$$K_2^I - K_1^I \quad = \quad \frac{\psi b}{4(2+\psi)}(9\bar{\gamma} - 7r)$$

Note that if $\delta = 0$, we have that $\psi = 0$, then $K_2^I = K_1^I$. Therefore, if $\delta > 0$, the difference of interest will be positive if and only if $(9\bar{\gamma} - 7r) > 0$. This means that if $\bar{\gamma} > \frac{7}{9}r$ we have that $K_2^I > K_1^I$, if $\bar{\gamma} = \frac{7}{9}r$ we have that $K_1^I = K_2^I$, and if $\bar{\gamma} < \frac{7}{9}r$ we have that $K_2^I < K_1^I$. Note that, in terms of the paper's notation, $\gamma^* = \frac{7}{9}r$. Finally, unconnected firms decrease their productive capacity. To see this, let us again subtract the equilibrium capacities in the two periods of interest:

$$K_2^U - K_1^U \quad = \quad -\frac{\psi b}{4(2+\psi)}(r + \bar{\gamma})$$

Note that if $\delta = 0$, we have that $\psi = 0$, then $K_2^U = K_1^U$. Therefore, if $\delta > 0$, we have that $(r + \bar{\gamma}) > 0$. This means that $K_2^U < K_1^U$. $\square$

Connected firms increase their capital stock in period 2 because of (i) the increase in firm entry in period 3 and (ii) the lower cost of capital they face. Unconnected firms adjust their capital stock downwards to keep prices high when new firms enter the market. A corollary of Proposition 3.3.1, which we take to the data, is:

$$\underbrace{K_2^C - K_1^C}_{>0} \quad > \quad \underbrace{K_2^I - K_1^I}_{\gtreqless 0} \quad > \quad \underbrace{K_2^U - K_1^U}_{<0} \tag{3.4}$$

In addition, to give us insights about mechanisms behind the lower cost of capital, the network analysis is useful. The lower cost of capital can be decomposed as $R - r = \bar{\gamma} + (\gamma - \bar{\gamma})$. Recall that firms with direct and indirect connections share directors, which implies that information flows freely among them. This means that their differential investment reaction places a bound to the role of information. Specifically, the higher the difference in investment between directly and indirectly connected firms, the lower the role of information.

The following proposition summarizes the predictions for profits:

Proposition 3.3.2. *Under assumptions 1 and 2, there exist $\bar{\gamma}$ such that profits are higher for connected firms during the lame duck period:*

$$\Pi_2^C(\bar{\gamma}) \;>\; \Pi_1^C(\bar{\gamma}) \quad ; \quad \Pi_2^I(\bar{\gamma}) \;>\; \Pi_1^I(\bar{\gamma}).$$

Unconnected firms obtain decreasing profits $\forall\gamma$, $\Pi_3^U < \Pi_2^U < \Pi_1^U$.

Proof: *Profits for firms with different types of political connections can be easily calculated from equilibrium capacities and the equilibrium price we computed in each period. Let us start by calculating the change in profits for firms with first degree connections. To do this, we need to take the difference between Π_2^C and Π_1^C. Note that if $\delta = 0$, $\Pi_2^C = \Pi_1^C$. Now let us assume $\delta > 0$ and take the derivative of the difference in profits with respect to ψ:*

$$\frac{\mathrm{d}(\Pi_2^C - \Pi_1^C)}{\mathrm{d}\psi} \;=\; \frac{1}{16b}\left(\frac{4b(a+br)}{(2+\psi)^2}(r-\bar{\gamma}) + \frac{8b^2\psi}{(2+\psi)^3}((1-\psi)r^2 + (1+\psi)(r-\bar{\gamma})^2)\right)$$

where the last term in the big parenthesis is always positive because $\psi < 1$. Then, given that ψ is increasing in δ, we can use the chain rule to conclude that $\Pi_2^C - \Pi_1^C$ increases with δ. This means that as δ increases, Π_2^C becomes larger than Π_1^C. Intuitively, the more connected firms value the future the more they are going to invest during the lame duck period in order to deter entry in period 3, this will lead to an increase in their profits. To facilitate the proof for indirectly connected firms, let us now move to the analysis of unconnected firms. Note that if $\delta = 0$ we have that $K_2^U = K_1^U$ and $P_2 = P_1$. Therefore, $\Pi_2^U = \Pi_1^U$. Assume that $\delta > 0$. Then, it is easy to see that $P_2 < P_1$ and $K_2^U < K_1^U$. Therefore, it follows that $\Pi_2^U < \Pi_1^U$. Unconnected firms do not have access to preferential credit which leads them to reduce their capital, reducing their profits. Using previous results for connected and unconnected firms, we can conclude that for any given $\delta > 0$, if $\gamma \to 0$, the difference in profits converges to the one of unconnected firms, meaning that $\Pi_2^I < \Pi_1^I$. If $\gamma \to r$, on the other side, the difference in profits converges to the one of connected firms, which implies that $\Pi_2^I > \Pi_1^I$. Therefore, for any given δ, there must be a $\bar{\gamma}$, such that for $\gamma > \bar{\gamma}$ the difference in profits is positive and for $\gamma < \bar{\gamma}$ is negative. Finally, since $P_2 > P_3$, due to the entry of new firms, we conclude that $\Pi_2^C < \Pi_3^C$, $\Pi_2^I < \Pi_3^I$, $\Pi_2^U < \Pi_3^U$. $\square$

Two forces drive profits: prices and capacity. Prices decrease during periods 2 and 3. Proposition 3.3.2 shows that for some $\bar{\gamma}$, profits increase during period 2. In particular, the higher $\bar{\gamma}$, the lower the profits for connected firms during period 2, because a higher $\bar{\gamma}$ implies indirectly connected firms increase their capacity by more, which lowers prices. A corollary of this proposition is:

$$\Pi_2^C - \Pi_1^C \;>\; \Pi_2^I - \Pi_1^I \;>\; \Pi_2^U - \Pi_1^U. \tag{3.5}$$

The difference in profits between the lame duck and democratic period, on the other hand, depends on the number of firms that enter the market in period 3. The number of entrants could be driven by, for example, lower entry costs.[16]

3.4 Empirical framework

3.4.1 Data and descriptive statistics

We constructed a panel dataset of firms listed in the Chilean stock market between 1985 and 1994. Our main analysis uses a balanced panel of 118 firms.[17] We collected information from two different sources: (1) quarterly balance sheets from the Chilean stock market's regulatory agency—to measure assets, physical capital, and profits—and (2) annual reports, which are required by law and audited by an international firm (e.g., Ernst & Young).[18] From annual reports, we hand-collected firms outstanding borrowing from banks, bond and equity issuance, number of workers, year of foundation, and information about whether or not the firm exports.

In the first part of our analysis, we use firm investment in physical capital (i.e., change in capacity), profits, logarithm of number of workers, productivity, and misallocation wedges as dependent variables. Firm investment in physical capital is defined as the logarithmic change in fixed capital (land, machinery, and buildings) between two quarters. This definition of investment is similar to the one used in Banerjee and Duflo (2014). Profits are defined as earnings before interests, taxes, and depreciation, although results are robust to different definitions. Revenue productivity was estimated using the Olley and Pakes (1996) procedure, but results are similar when we use a simple Solow residual.

In addition, two misallocation wedges were constructed using the Hsieh and Klenow (2009) methodology. The first one, a capital wedge, measures a distortion in the marginal product of capital relative to labor. This wedge would be higher for firms without access to capital, and lower for firms with access to cheap credit. The second one, an output wedge, is a distortion that changes both the marginal product of labor and capital by the same proportion. It would be higher for firms that face government restrictions and lower for those firms that benefit for output subsidies. All variables were winsorized at

[16]Some auxiliary predictions can be derived from the model. First, increases in capacity during the lame duck period are associated with more profits during the democratic period. Second, capacity increases are a function of how capital intensive the industry is. We test for these predictions in section 3.7.

[17]To avoid confounding factors from the recovery period after the 1982 economic crisis, we start our analysis in 1985, when GDP growth reached its pre-crisis level (Figure C.2). As we focus on physical capital, we exclude all firms operating in the financial services industry.

[18]Chile's regulatory agency is called *Superintendencia de Valores y Seguros*. The US equivalent is the Securities and Exchange Commission. All variables are measured in 1998 Chilean pesos and were transformed using the consumer price index constructed by the Central Bank of Chile.

2.5 percent of the empirical distribution to handle outliers.[19]

Figure 3.2 shows time series of our main variables for the three types of firms. For investment and profits (Panels A and B), there is no clear pre-trend between groups, something we statistically show later on. The time series for firms with direct and indirect connections, however, diverge after the plebiscite. Productivity (Panel C) and the logarithm of labor (Panel D) also show similar patterns before the democratization announcement. Throughout the period, unconnected firms are the most productive, and connected firms are the least productive. These differences in productivity are not driven by selection into industries.

To improve our understanding of firms, we also constructed firm characteristics: logarithm of total assets (firm size), year of foundation, indicator for exporting firms, indicator for firms privatized during Pinochet's dictatorship, and existing business groups in 1987. A total of 40 firms were privatized during the dictatorship, and 32 firms were part of nine different business groups.[20]

Table 3.1 presents summary statistics before the plebiscite. As expected, political connections were not randomly distributed. Using univariate regressions, we observe that connected firms were larger and older, as well as more likely to have been exporters, privatized by Pinochet, and part of a business group. Also, connected firms were less productive and accrued more debt from banks. Differences between firms with first and second degree connections were considerably smaller. We used a variety of approaches to show our results are not driven by any of these differences in observable variables.

In Table 3.2, we study differences in observable variables by estimating OLS regressions for the dictatorship and lame duck periods separately, with and without firm controls. During Pinochet's dictatorship, connected firms benefited from cheaper access to credit and higher subsidies, as measured by capital and output wedges. Connected firms also had more access to credit, both from government and private banks, and were less productive. All differences tend to be larger for firms with first degree connections. Interestingly, during the lame duck period, we observe larger differences in capital distortions and access to government credit between connected and unconnected firms.

Finally, we classified all firms into two-digit industries following the international standard industrial classification of all economic activities (United Nations, 2008). The industries in our dataset are: accommodation and food service activities; agriculture, forestry, and fishing; arts, entertainment and recreation; construction; education; electricity, gas, steam, and air conditioning supply; human health and social work activities; information and communication; manufacturing; mining and quarrying; real estate activities; transportation and storage; and wholesale and retail trade, repair of motor

[19] Misallocation wedges were constructed using equations (17) and (18) in Hsieh and Klenow (2009).

[20] To identify privatized firms, we used data from a commission in charge of investigating the privatization process (CEME, 2004). Business groups were identified using Superintendencia de Valores y Seguros (1988).

vehicles, and motorcycles.[21]

3.4.2 Empirical strategy

Our econometric strategy exploits within firm variation and the 1988 plebiscite as an exogenous democratization announcement, resulting in a standard differences-in-differences with three time periods and three types of firms. However, as firms were not randomly assigned to each group, we perform a variety of robustness exercises, including matching estimators and placebo checks, to show that observed responses correspond to differences in political connections.

The main regression equation we estimate is:

$$Y_{ijkt} \quad = \quad \boldsymbol{\beta}_T \times P_i + \boldsymbol{\gamma}_T \times \widetilde{P_i} + \boldsymbol{\psi}_k \times Post_t + \xi_i + \lambda_t + \varepsilon_{ijkt} \tag{3.6}$$

where Y_{ijkt} is an outcome for firm i—which is part of business group j and operates in industry k—during period t. The time subscript T groups years into political periods (i.e., $T = \{lame\ duck,\ democracy\}$), with the dictatorship period as the omitted category. The vectors of parameters $\boldsymbol{\beta}^k_T = (\beta_{lame} \quad \beta_{dem})'$ and $\boldsymbol{\gamma}_T = (\gamma_{lame} \quad \gamma_{dem})'$ contain the coefficients of interest. The indicators P_i and $\widetilde{P}$ are equal to one if firm i had a first or second degree political connection in 1987, mutually exclusive categories. The vector $\boldsymbol{\psi}_k$ is composed by industry fixed effects, $Post_t$ is an indicator for the period after the plebiscite, and ξ_i and λ_t are firm and time fixed effects. Finally, ε_{ijkt} is an error term clustered at the business group level.[22]

In addition, we present two variations of equation (3.6). The first classifies firms with second degree connections as unconnected firms, to explicitly show the importance of indirect connections. The second includes second degree connections but omits $\boldsymbol{\psi}_k \times Post_t$. When compared to our main regression, this last specification addresses concerns about industry shocks driving our results.[23]

3.5 Firms during political transition

3.5.1 Main results

In Table 3.3-A, we present our main results. In column 1, we show how changes in productive capacity changed following the plebiscite during both the lame duck period

[21]See Appendix Table C.3 for the distribution of connections by industry.

[22]Any firm that is not part of a business group is assumed to be a business group on its own. Overall, we have a total of 104 clusters in our dataset.

[23]For example, one might worry that firms in the energy sector anticipate increases in demand after the plebiscite and decide to increase their productive capacity accordingly. Including industry fixed effects after the plebiscite addresses this type of concern.

and after the new government takes office. Firms with first degree connections increase their capacity right after the democratization announcement, although the increase in the lame duck period is only marginally significant (p-value of 0.11). In terms of its magnitude, note that the standard deviation of changes in capacity was 0.05 before the plebiscite among connected firms. This means that changes in capacity increase by 0.22 standard deviations when only first degree connections are considered. Another way to interpret the economic magnitude of these coefficients is by transforming quarterly coefficients to more aggregated changes. For example, a coefficient of 0.012 implies a 7.5 percent increase in investment during the lame duck period using the formula for semilogarithmic equations in Kennedy (1981).

Column 2 allows for the indirectly connected firms to respond differently than unconnected firms. As expected, coefficients have a positive sign and are smaller in magnitude in the lame duck period. In addition, the coefficients for firms with direct political connections increase by 50 percent, and the coefficients on indirectly connected firms are significant or marginally significant (γ_{lame} has a p-value of 0.13). We interpret these results as evidence that indirect connections matter.[24]

Column 3 includes industry fixed effects after the plebiscite. Results are similar to those in column 2. Nevertheless, because column 3 compares firms within the same industry, this is our preferred specification. Firms with direct political connections increase their capacity by 11.4 percent (0.36 standard deviations) during the lame duck period, which is statistically significant at conventional levels. Moreover, firms with second degree connections increase their capacity by 8.7 percent (0.22 standard deviations), which is smaller, as expected, and only marginally significant (p-value of 0.14).[25]

It is important to highlight one characteristic of our results that is not captured by the theoretical framework. Even though there are multiple periods in the model, the setup is essentially static. Although this has the benefit of simplicity, the downside is that the model does not generate predictions for changes in capacity during the democratic period. Nevertheless, after reading annual reports, we realized that (1) most investments in physical capital take place across multiple years, and (2) investments are usually followed by complementary investments.[26] Then we adopt this interpretation of investment and expand the predictions in Proposition 3.3.1 to argue that capacity among connected firms increases during the democratic period, as well.

---

[24]As robustness check for the clustering of standard errors at the business group level, we constructed clusters of firms using a community detection algorithm (Newman, 2004). In Appendix Table C.4, we show standard errors are extremely similar.

[25]In Appendix Table C.5, we present a similar set of regressions where we (1) pool directly and indirectly connected firms into one type of connection, and (2) exclude indirectly connected firms from our analysis. Our main results remain the same. A different approach to measure political connections is to use the response of abnormal returns the day after the plebiscite. We performed this exercise, and results are qualitatively similar (Appendix Table C.6). When restricting attention to firms with observed abnormal returns, however, we are left with only half of observations.

[26]For example, *Chilgener* decided to invest in the *Alfalfal* project in 1988, and investment took place across multiple years.

The following three columns in Table 3.3-A replace the dependent variable by profits, standardized using the mean and standard deviation of all firms in the dictatorship period. Results in columns 5 and 6 indicate that profits are higher for connected firms, especially during the lame duck period, the effect being smaller for indirectly connected firms. Point estimates are economically significant: directly (indirectly) connected firms increase profits by 0.3 (0.1) standard deviations during the lame duck period.[27]

Summarizing, we find evidence supporting the predictions in Proposition 3.3.1 and 3.3.2. In what follows, we test for the robustness of these results in order to be confident in interpreting them as a causal effect of political connections.

3.5.2 Specification checks

Our theoretical framework abstracts from two variables in the production function: productivity and labor. These omissions are a threat to our findings if a connected firm's productivity or labor is affected by the democratization announcement, and this causes changes in capacity decisions. This could be the case if, for example, workers attach a premium to work for connected firms, the premium disappears with a democratization, and firms substitute labor for capital as a consequence.

In Table 3.3-B, we test formally for differential changes in productivity and labor using our annual dataset. The first three columns show no differential changes in productivity. The exception is the coefficient on firms with second degree connections during the lame duck period, but the coefficient is smaller than 0.15 standard deviations. The remaining point estimates are all smaller than 0.05 standard deviations. In columns 4 to 6, we present results using the logarithm of workers as dependent variable, in which case point estimates are approximately interpreted as elasticities. There are no statistically robust patterns in labor adjustments among connected firms following the plebiscite. In sum, productivity and labor are unlikely to be omitted variables driving our results.[28]

In Table 3.3-B, we employ a balanced panel dataset. When focusing on firms that are always operating throughout the period under study, we are effectively excluding the extensive margin of adjustment in capacity: closing the firm. In Appendix Table C.9, we present results using an unbalanced panel of 145 firms; results remain qualitatively similar, and point estimates are slightly attenuated but still statistically significant at conventional levels.[29] Consistent with this result, we find no differential attrition by

---

[27]To test for the prediction of decreasing profits for unconnected firms, we estimated a version of equation (3.6), where we replace the quarterly fixed effects λ_t by $\boldsymbol{\lambda}_T = (\lambda_{lame}\ \lambda_{dem})'$ and find that, consistent with predictions, $\lambda_{lame} = -0.27$ and $\lambda_{dem} = -0.24$ ($p-$values of 0.001 and 0.003, respectively).

[28]Results are similar when we use a Solow residual to estimate productivity (Appendix Table C.7). Results are also robust to use different parts of the empirical distribution to winsorize the dependent variables (Appendix Table C.8).

[29]The inclusion criteria in this case is that a firm must be observed operating at least six quarters before and after the plebiscite to be included. We do this to be sure we can estimate firm fixed effects and exploit within firm variation.

connected status using the unbalanced panel (Appendix Table C.10).

A final specification check is presented in column 1 of Table 3.4, where we collapse the data to three periods (dictatorship, lame duck, and democracy) to deal with potentially inconsistent standard errors due to serially correlated outcomes (Bertrand et al., 2004). Estimated coefficients are similar to previous results. Therefore, by using the entire panel, we are not introducing bias in standard errors.

3.5.3 Identification

As shown in Table 3.1, firms not only differ in their links to Pinochet's dictatorship, but also on other observable variables. We now present and discuss a large set of empirical exercises that suggest our findings are driven by political connections and not other variables. Results in this and the following sections give us confidence that our findings support the causal chain in our argument.

In columns 2-5 of Table 3.4, we add control variables to study the influence of observables in our estimates. Column 2 controls for an interaction between an indicator for large firms, defined as those above the median of the firm size distribution before the plebiscite, and an indicator for the period after the plebiscite. We follow the same strategy in columns 3, 4, and 5, but use indicators for firms privatized during the dictatorship, firms participating in a business group, and exporting firms. Moreover, in column 6, we include all previously mentioned control variables. In all these cases, results are unchanged.

We perform three additional exercises making use of differences in observable variables across firms. First, we control in a flexible way for the probability of being connected based on observables.[30] Results are presented in column 7 and are similar. In the second exercise, we restrict attention to the subset of firms with overlap in the propensity score distribution using the procedure in Crump et al. (2009).[31] Results are presented in column 8, and the coefficient of interest is 25 percent larger for directly connected firms. In the third exercise, we use a synthetic control approach proposed by Abadie and Gardeazabal (2003) and Abadie et al. (2010). Results can be found in Appendix Table C.13 and are again similar.

In addition, in Appendix Table C.14, we present two placebo treatments: the first examines the time before and after the third quarter of 1986, restricting attention to the period 1985–1988 (Panel A); and the second, the time before and after the 1994 presidential election, restricting attention to the period 1990–1997 (Panel B). In both cases, point estimates are smaller and not statistically significant. The former placebo exercise also serves as a test for the parallel trend assumption in our differences-in-differences

[30]We estimate two probit models to predict (direct and indirect) political connections using observable variables before the plebiscite (Appendix Table C.11). We then interact this propensity score with a linear trend and include it as control variables.

[31]We present a version of Table 3.1 for the balanced sample in Appendix Table C.12. Most firm characteristics are balanced except for the logarithm of workers and total assets.

strategy. In addition, our main results are: (1) robust to measuring political connections in 1986 (Table C.15), (2) robust to measuring political connections as share of directors that is connected (Table C.16), (3) larger among firms managed by connected presidents of the board or connected through a former high-level politician (Tables C.17 and C.18), and (4) not driven by firms substituting connections from Pinochet's dictatorship to the new democratic regime (Table C.19).

3.5.4 Alternative explanations

Although previous results suggest firms in the dictator's network make critical investments during the transition, alternative explanations are still a concern. A first concern relates to how connections were formed. It is possible that Pinochet's regime positioned individuals in firms that were expected to do well in the future. A second concern relates to the effect of (political) uncertainty on investment (e.g., Bloom 2009, Julio and Yook 2012).

Two pieces of evidence suggest that the strategic positioning of individuals is unlikely to play a major role in our context. First, Pinochet's regime should have been able to identify (1) firms that will do well in the future, and (2) how the future will be. Although possible, the variables determining how a firm will fare need not to be related to the observable variables we consider in the previous analysis. In addition, connections were formed before 1988, a time of considerable uncertainty. Therefore, identifying which firms will do well in the future was difficult in practice. Second, stock prices of connected firms *decreased* following the plebiscite. If firms in Pinochet's network were expected to do well in a democratic world, we should have observed an *increase* following the plebiscite.

Political uncertainty could theoretically explain the investment patterns we have documented. Empirically, this is not the case. Because the effect of political uncertainty on firm investment is captured by time fixed effects, the main threat to our interpretation relates to the differential impact of uncertainty on firms in the dictator's network. If politically connected firms are delaying investments until political uncertainty is resolved, we should observe an increase in their liquid assets before the plebiscite. The difference in liquid assets between connected and unconnected firms is, however, not statistically different from zero before 1988. If anything, politically connected firms have *less* liquid assets than unconnected firms before the plebiscite.

In summary, our main set of findings is robust to different specifications, control variables, and estimation techniques. In addition, results are mostly consistent with our causal argument and less consistent with alternative explanations. Remarkably, the magnitude of capacity responses follow the order suggested by our theoretical framework. In the next section, we provide evidence for the main mechanism driving our results, and we explore the consequences of our findings for connected firms during the democratic period.

3.6 The credit market during political transition

Why were politically connected firms able to increase their productive capacity, but not firms which lacked political connections? In the model, this difference in reactions is driven by the lower cost of capital faced by connected firms. In this section, we show that politically connected firms indeed had a differential relationship with banks during the political transition. We begin by presenting narrative evidence of favoritism from government banks to connected firms during the political transition. Then, we move to an empirical test that analyzes loans between government banks and firms over time.

3.6.1 Government banks

Three government banks are active in our data: the Bank of the State, the Central Bank, and the Production Development Corporation. The Bank of the State granted 83 percent of loans from government banks between 1988 and 1990. Executives at these banks were directly appointed by Pinochet and were in charge of the review and approval of loan petitions (Law No. 2079, enacted in 1978).

The President of the Bank of the State during the transition was Alvaro Bardón, former President of the Central Bank (1977–1981), Undersecretary of Finance (1982), and member of the Chicago Boys. Bardón was appointed president one month after the plebiscite (November 7, 1988) and remained in this position until March 1990. This appointment has been the focus of controversy due to the bank's financial operations during the transition. The controversy lies on the privatization of *El Mercurio* and *La Tercera* (the two largest newspapers), bankrupted by the time of the transition. These newspapers were bailed out after the 1982 financial crisis and, as a consequence, were heavily indebted to the Bank of the State. These debts meant that the opposition party could have owned a significant part of the written media after taking office in 1990. To prevent this scenario, Bardón used debt swaps to transfer the ownership of newspapers to politically connected firms. These financial operations were implemented between November 1989 and March 1990 and, because of significant mispricing, cost the Bank of the State approximately 26 million USD (Leon-Dermota, 2003).[32]

The newspapers case exemplifies how the Pinochet regime used government banks to gain an advantage during the political transition. Leon-Dermota (2003, p.143) puts it succinctly: "The connection between *El Mercurio* and the military regime facilitated access to credit that was used to invest and gain an advantage over competitors." The difference in debt from government and non-government banks during the political transition is our test of differential access to credit among connected firms.

[32]Price Waterhouse was in charge of estimating the value of debts. Bardón and his team were investigated for state fraud in 1991. In a controversial ruling, the Supreme Court decided to exonerate them. Leon-Dermota (2003) argues that this exoneration is an example of Pinochet's power in the new democratic era.

3.6.2 Econometric strategy

As discussed in section 3.4.1, we have hand-collected data on outstanding loans between firms and banks for every year in our period of study (1985–1994). The majority of firms have some positive level of debt with multiple banks. We collapsed the universe of banks in our dataset to two types: government banks and others (i.e., two firm-bank relationships for each firm every year). We exploit within firm-bank variation over time in the following regression:

$$\begin{aligned} Y_{ijkbt} &= \boldsymbol{\rho}_T \times P_i \times G_b + \boldsymbol{\lambda}_T \times \widetilde{P}_i \times G_b \\ &+ \boldsymbol{\pi}_{1T} \times P_i + \boldsymbol{\pi}_{2T} \times \widetilde{P}_i + \boldsymbol{\pi}_{3T} \times G_b \\ &+ \boldsymbol{\psi}_k \times Post_t + \theta_{ib} + \eta_T + \varepsilon_{ijkbt} \end{aligned} \tag{3.7}$$

where Y_{ijkbt} is debt of firm i—which operates in business group j and industry k—with bank b in year t.[33] We use three dependent variables: (1) monetary value of debt in Chilean pesos, (2) an indicator for positive amounts of debt, and (3) the logarithm of debt.

In addition, T indexes a time period (lame duck, democracy); P_i and $\widetilde{P}$ are indicators for firms with first and second degree connections, respectively; G_b is an indicator for debt with government banks; the vector $\boldsymbol{\psi}_k$ is composed by industry fixed effects; $Post_t$ is an indicator for the period after the plebiscite; θ_{ib} is a set of firm-bank fixed effects; η_T are period fixed effects, and ε_{ijkbt} is an error term clustered at the business group level. Coefficients are defined as: $\boldsymbol{\rho}_T = (\rho_{lame} \quad \rho_{dem})'$, $\boldsymbol{\lambda}_T = (\lambda_{lame} \quad \lambda_{dem})'$, and $\boldsymbol{\pi}_{\ell T} = (\pi_{\ell,lame} \quad \pi_{\ell,dem})'$, with $\ell = 1, 2, 3$. If connected firms had a different relationship with government banks during the lame duck period, then $\rho_{lame} > 0$ and $\lambda_{lame} > 0$.

3.6.3 Results

Table 3.5 presents results. Column 1 presents estimates using debt in Chilean pesos as dependent variable. As we hypothesized, directly connected firms have substantially more debt with government banks during the lame duck period, providing an explanation for the differential increase in capacity that we have documented. This effect is large, as can be seen from a firm's average debt. The probability of having a positive amount of debt is also larger for directly connected firms in this period (column 2). Debt over assets also increases significantly after the plebiscite (column 3). Remarkably, coefficients are always positive but smaller in magnitude for indirectly connected firms, although marginally significant at conventional levels (p-values of 0.16, 0.23, and 0.07,

[33] This is similar to the statistical test in Khwaja and Mian (2005). The main difference is that we exploit *within firm-bank* variation *over time* (i.e., before and after the plebiscite), and not only *within firm* variation.

respectively).[34]

Following our sensitivity analysis in section 3.5, in Appendix Table C.21 we perform a series of robustness checks to study the stability of these estimates. Reassuringly, debt is always larger for connected firms during the lame duck period, larger in firms with first degree connections than in firms with second degree connections, and estimated coefficients are of similar magnitude and statistically significant at conventional levels. In addition, we also explore the effect of the plebiscite on other sources of funding, such as stocks and bond issuances (Appendix Table C.22). Results are not significantly different between connected and unconnected firms, which emphasize the importance of the relationship between connected firms and government banks.

Why are government banks lending money to connected firms? Throughout the paper we have emphasized that these additional funds allowed firms to make critical investments during the political transition. Another interpretation is that Pinochet's regime is tunneling money to friends with no investment plans. One way to test for this alternative interpretation is by studying changes in extraordinary dividends after the 1988 plebiscite. If this interpretation is correct, we should observe an increase in extraordinary dividends after 1988. Results, however, suggest that tunneling is unlikely to be a concern in this context (see Table C.23).

3.7 Consequences during the democratic period

3.7.1 Firm survival

Are connected firms more likely to survive in the new democratic regime? To answer this question, we focus on the probability that a firm is operating in the stock market in the years following the plebiscite. If connected firms gained an advantage by increasing their capacity after the democratization announcement, we should expect these firms to remain operating for a longer time. To test for this hypothesis, we estimate the following cross-sectional regression every year between 1994 and 2008:

$$Y_{ijt} \quad = \quad \alpha_t + \beta_t P_i + \gamma_t \widetilde{P}_i + \gamma_t' \mathbf{X}_i + \boldsymbol{\psi}_j + \varepsilon_{ijt} \tag{3.8}$$

where Y_{ijt} is an indicator that equals one if firm i in industry j is operating in year t, P_i and $\widetilde{P}$ are indicators for first and second degree connections, $\mathbf{X}_i$ is a set of control variables (firm size, indicator for privatized firms) for the pre-plebiscite period, and $\boldsymbol{\psi}_j$ is a set of industry fixed effects.

Figure 3.3-A presents OLS estimates of coefficients $(\hat{\beta}_{1994}, \ldots, \hat{\beta}_{2008})$. Although these

[34]In Appendix Table C.20 we show our results are robust to identify banks closely related to the Pinochet regime in three other ways: (1) banks that were privatized or received financial help during the 1982 economic crisis, (2) state-owned banks and large banks that were privatized during Pinochet's regime, and (3) state-owned banks and banks with politically connected directors.

estimates are somewhat imprecise, it is clear that firms with direct connections were more likely to remain operating. In particular, these firms are approximately 10 percentage points more likely to be active by the year 2000, an effect that increases to slightly more than 20 percentage points when the year is moved forward to 2008. This effect is similar among indirectly connected firms until the year 2000, but remains around 10 percentage points towards 2008 (Appendix Figure C.6).

3.7.2 Profits

If firms that increased their capacity obtained an advantage over firms that did not, we should expect the former to have higher profits in the new democratic era. We test for this auxiliary prediction of our model by estimating the correlation between profits and firm-specific capacity responses. Because investments are expected to increase profits (at least on average), we compare the profits response to investments during the lame duck period to the profits response to investments in a placebo period before the plebiscite (3*rd* quarter of 1986).

We proceed in three steps. First, to estimate firm-specific responses, we augment equation (3.6) by interacting time period indicators with firm specific indicators. This allow us to estimate firm specific capacity responses $\beta_{i,lame}$ with $i = 1, \ldots, 118$—in Appendix Figure C.7 we plot these coefficients. Second, we construct profits in year t by adding up quarterly profits. Third, we estimate the following cross-sectional regression each year between 1990 and 1994:

$$\Pi_{it} = \alpha_t + \tau_t \widehat{\beta}_{i,lame} + \eta_{it} \tag{3.9}$$

where Π_{it} represents profits in year t for firm i, α_t is a constant term, and $\widehat{\beta}_{i,lame}$ is our estimate of firm-specific capacity responses. To facilitate the interpretation of coefficients, we have standardized capacity responses and yearly profits. The coefficient of interest is τ_t, and our theoretical framework implies that $\tau_t > 0$. Standard errors for τ_t are calculated using a bootstrap procedure to account for the uncertainty in our estimation of firm-specific capacity responses.

Figure 3.3-B presents OLS estimates of coefficients $(\hat{\tau}_{1989}, \ldots, \hat{\tau}_{2008})$. There are three interesting patterns to note from this figure. First, there is, as expected, a positive and statistically significant relationship between capacity responses and profits. Second, the effect is roughly constant in the five years after the plebiscite. Third, the effect is large when compared to investments in a different period. In particular, a one standard deviation increase in a firm's capacity response to the plebiscite is associated with an increase of 0.2 standard deviations in profits, which is larger than an increase of 0.1 standard deviations to investments in a different period.

3.7.3 Entry barriers

One key assumption in entry deterrence models is that investment in capacity is an effective way to reduce firm entry. To test for this, we follow Lambson and Jensen (1998), and construct a proxy for average sunk costs by industry using average property, plant, and equipment in the period 1985–1987. We then divide industries into those with more and less average sunk costs and use this variable to augment our main specification with a triple difference. This exercise is useful to test if firms in industries with higher entry costs (or more capital intensive) tend to invest more during the lame duck period to deter entry. Our results suggest that the increase in capacity among connected firms is indeed higher in more capital intensive industries during the lame duck period (see Appendix Table C.24).

In a related exercise, we estimate industry-level regressions to test whether a higher presence of connected firms is associated with lower firm entry after the democratization. In particular, we estimate a regression of log number of firms on the share of connected firms. We do this in our sample of publicly listed firms and in a panel of firms constructed from the Chilean annual manufacturing census, effectively performing an out-of-sample test. We find suggestive evidence that those industries with a higher presence of connected firms tend to have less entry in the new democratic period. However, results need to be interpreted with caution, because we have a limited number of industries in our dataset.

3.8 Conclusion

In order to improve our understanding of the effects that democracies have on the economy, we need to study the behavior of key economic actors during political transitions. In this paper, by studying firms, we have taken a first step in this direction. Our empirical analysis focuses on Chile's transition to democracy, one of the most well-known political transitions in the last three decades. This transition offers a unique opportunity to measure not only the degree of connectedness that firms have with the dictatorship, but also key variables such as profits, inputs, and loans. Our results show that firms in Pinochet's network significantly increased their capital stock during the political transition and that these investments improved their market position into the new political regime.

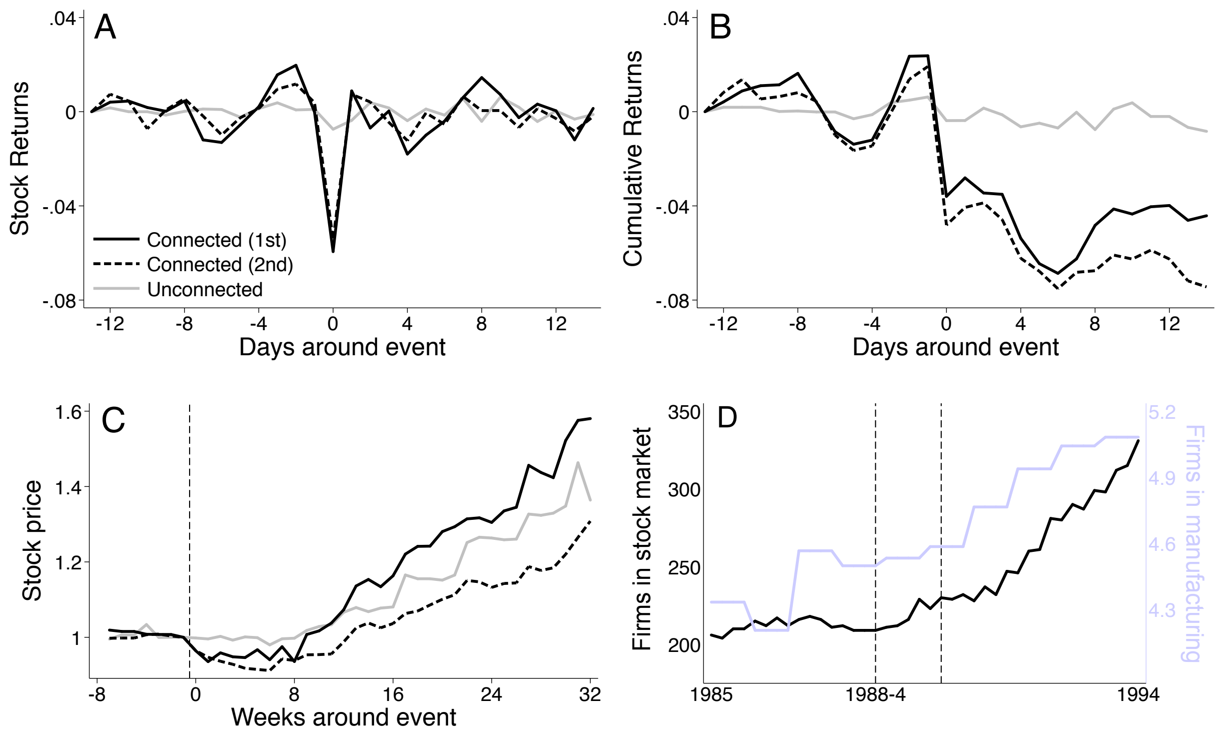
Although increasing productive capacity might seem like a surprising strategy from a firm's perspective, we show that it is a perfectly rational—but not necessarily welfare-maximizing—response to the news of a future change in political regime. When studying the mechanisms at work, we explain this increase in productive capacity through a simple entry-deterrence motive for incumbent firms, combined with favoritism of connected firms by government banks during the dictatorship. Government banks have been suggested as important actors during Chile's transition to democracy and have empirical support as a mechanism behind our findings.

Overall, our results show that firms successfully improve their market position using their political ties, which allows them to transfer market distortions from the old to the new political regime. Although connected firms are relatively unproductive during dictatorship, we cannot claim their investments are necessarily inefficient for the market. In order to compute the complete welfare implications of our findings, we would need to fully characterize demand and supply in different industries, which is beyond the scope of this paper. In that sense, the study of the demand side and the labor market during political transition seems promising for future research. Nevertheless, taken together, our findings suggest significant negative welfare impacts associated to investments before a political transition takes place.

The reader might worry that Chile's transition to democracy differs from other transitions and our findings have little external validity. Although certainly unique, we believe the time structure in Chile's democratization provides a valuable starting point to evaluate the behavior of firms before a regime change. If firms in a dictator's network have more accurate information about the future than other firms—the most likely case in our view—our framework would predict larger effects than those documented in this paper. Conversely, if the new regime is fragile and a reversal probable, increasing investment might not be the best response from a firm's perspective. Careful regulation of the credit market during a democratization seems like a potentially effective policy to increase the benefits of democracies. One way to achieve this regulation is with government audits of investment projects, which have been shown to reduce corruption (Olken, 2007).

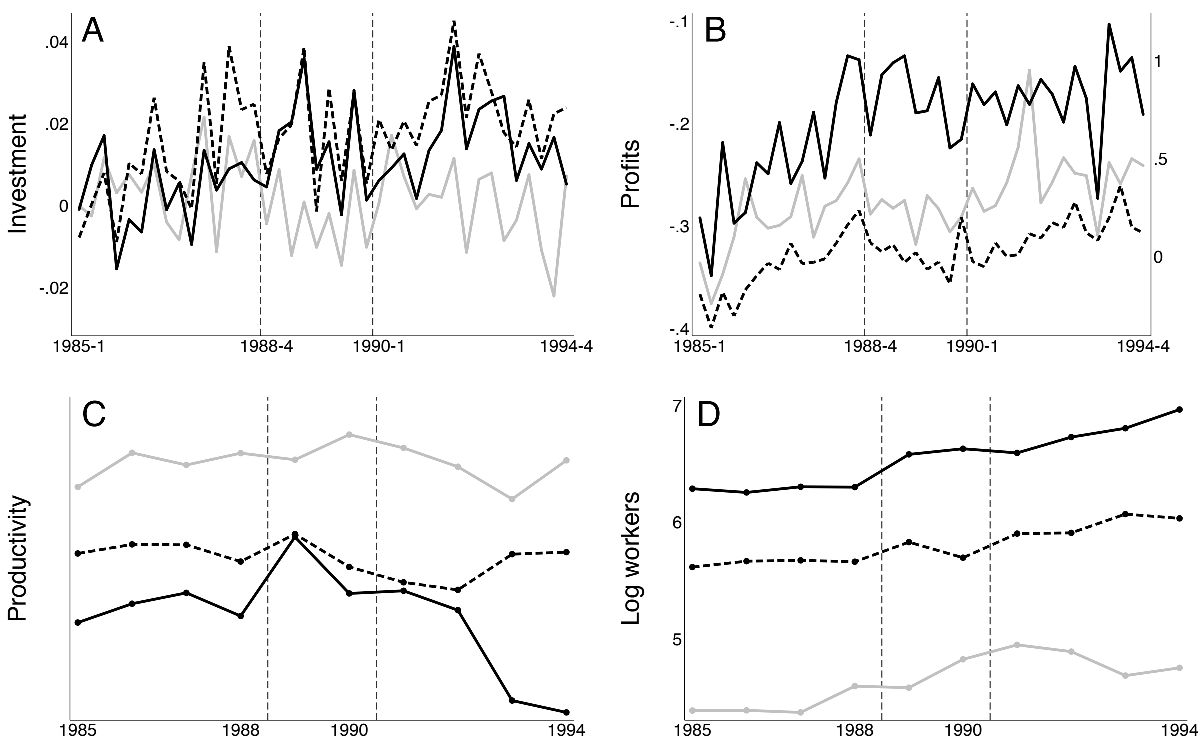
Besides the outcomes we have analyzed, there could be other economic and political areas affected in the democratic period. We believe the political arena is particularly important in not only the Chilean case, but potentially in other settings as well. Economic and political power usually coevolve in complex ways. If economic power translates into political power during the new regime, then the old political regime could still exert its influence. Recent corruption scandals in Chile suggests this is indeed the case, as several firms have been accused of (illegally) financing electoral campaigns. Therefore, we believe that tracking wealth across political regimes is yet another interesting avenue for future research, and it would help us improve our understanding of the coevolution of economic and political power.

Figure 3.1: Stocks and firms after the 1988 plebiscite



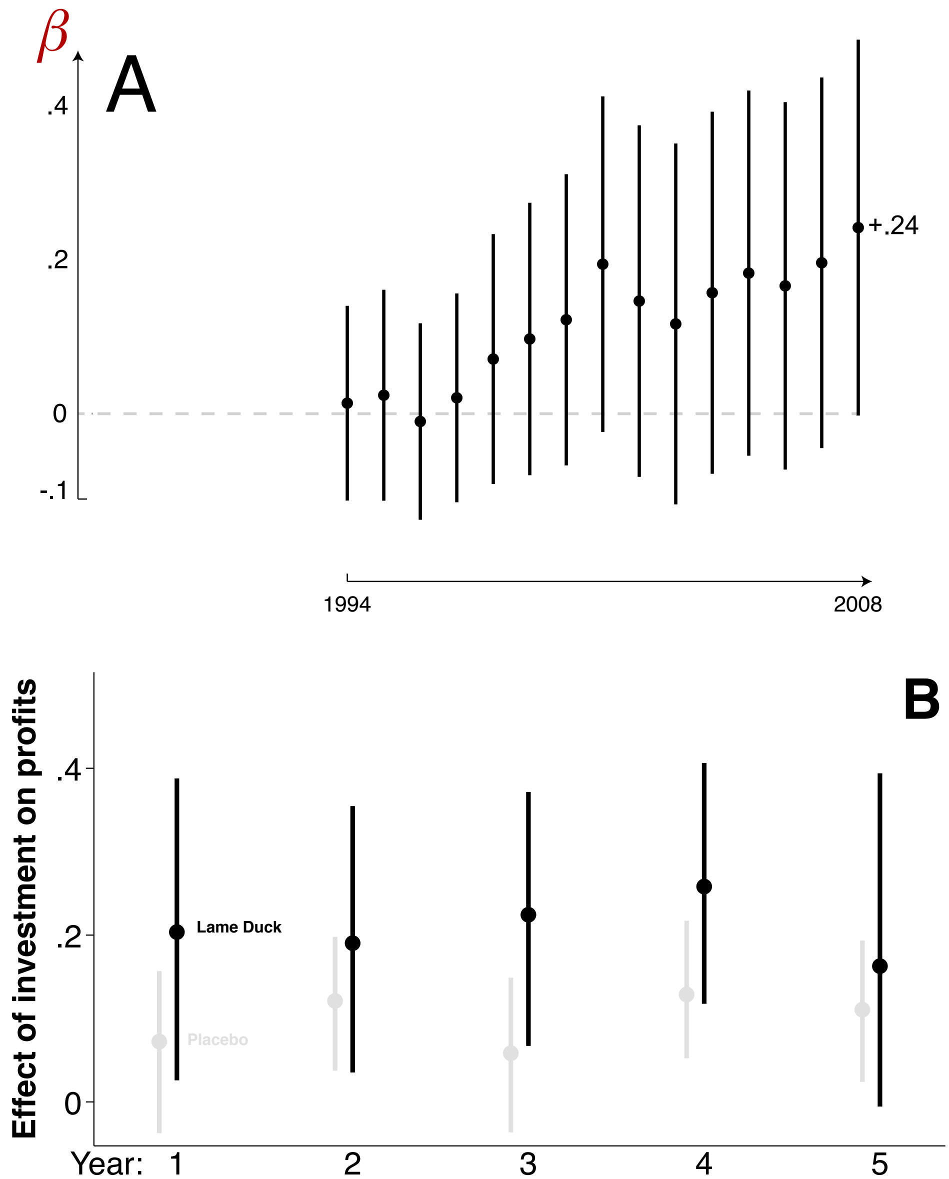
Notes: Panel A presents daily stock market abnormal returns for firms with direct connections (solid black), indirect connections (dash black), and no connections (gray) around the 1988 Plebiscite (October 5th, 1988). Panel B presents cumulative abnormal returns, which are defined as $CR_{it} = \sum_{k=-12}^{t} R_{ik}$. Panel C presents weekly stock prices relative to the price the week before the plebiscite. Panel D presents the number of firms operating in the stock market during the period under study. Vertical dash lines represent the time of the plebiscite and the beginning of the democratic period (March 11th, 1990). The second y-axis (gray) presents the number of firms operating in the manufacturing census (*Encuesta Nacional de la Industria Manufacturera, ENIA).*

Figure 3.2: Time series for firm outcomes



Notes: Average investment in physical capital in Panel A; profits before interests, taxes and depreciation (standardized) in Panel B; productivity using Olley and Pakes (1996) procedure in Panel C (no meaningful absolute value, so numbers in the *y*-axis are omitted); and the logarithm of workers in Panel D. Legend for different groups of firms: direct connections (solid black), indirect connections (dash black), and no connections (gray). The events under study are the plebiscite (1988-4) and the beginning of the democratic period in Chile (1990-1). Panels A and B use quarterly variation. Panels C and D use annual variation.

Figure 3.3: Consequences in democracy



Notes: Panel A presents OLS $\hat{\beta}_t$ coefficients from $Y_{ijt} = \alpha_t + \beta_t P_{i,1987} + \delta_t \widetilde{P}_{i,1987} + \gamma_t \mathbf{X}_{i,1987} + \boldsymbol{\zeta}_j + \varepsilon_{it}$, where Y_{it} is an indicator that takes the value of one if firm i is operating in year t, $P_{i,1987}$ and $\widetilde{P}_{i,1987}$ are indicators for first and second degree connections, $\mathbf{X}_{i,1987}$ is a set of control variables (firm size, indicator for privatized firms), and $\boldsymbol{\zeta}_j$ is a set of industry fixed effects. Panel B presents OLS $\hat{\tau}_t$ coefficients from $\Pi_{it} = \alpha_t + \tau_t \hat{\beta}_{i,lame} + \eta_{it}$, where Π_{it} is (standardized) profits in year t for firm i, α_t is a constant term, and $\hat{\beta}_{i,lame}$ is our estimate of firm-specific capacity responses after the plebiscite. The standard error for τ_t is calculated using a bootstrap procedure. For comparison, we plot the effect of investment in 1986–3 on profits in the following years.

Table 3.1: Firm characteristics

Mean and standard deviation for our main variables in the period 1985–1987

	No link	Direct link to Pinochet (P)	Indirect link to Pinochet ($\widetilde{P}$)	Uni-variate regression		
	(1)	(2)	(3)	(2) – (1)	(3) – (1)	(2) – (3)
A. Quarterly dataset						
Investment	0.00	-0.00	0.01	-0.00	0.00	-0.01
	(0.05)	(0.05)	(0.06)	(0.01)	(0.01)	(0.01)
Profits	-0.32	0.39	-0.11	0.70***	0.20**	0.50**
	(0.24)	(1.48)	(0.52)	(0.17)	(0.08)	(0.20)
Log assets	14.52	17.55	16.82	2.99***	2.27***	0.72*
	(2.10)	(1.87)	(1.37)	(0.43)	(0.41)	(0.37)
B. Annual dataset						
Log workers	4.38	6.27	5.65	1.90***	1.27***	0.63***
	(1.99)	(1.65)	(1.39)	(0.29)	(0.29)	(0.24)
Productivity	-0.47	-1.52	-1.08	-1.05***	-0.61**	-0.45
	(1.69)	(1.92)	(1.54)	(0.29)	(0.27)	(0.27)
Debt with government banks	3.7	17.0	20.6	13.4***	17.0***	3.6
	(17.2)	(48.0)	(45.9)	(4.9)	(4.8)	(6.6)
Debt with other banks	17.8	87.8	63.0	70.0***	45.2***	24.8*
	(46.4)	(117.7)	(92.0)	(12.1)	(10.1)	(15.0)
C. Time invariant						
Age in 1987	39	53	49	14**	10	4
	(27)	(30)	(29)	(6)	(7)	(7)
Exporter	0.26	0.48	0.57	0.27**	0.34***	-0.07*
	(0.43)	(0.50)	(0.50)	(0.10)	(0.11)	(0.11)
Privatized	0.11	0.56	0.33	0.44***	0.21**	0.22*
	(0.31)	(0.50)	(0.47)	(0.09)	(0.10)	(0.11)
Business group	0.02	0.21	0.39	0.19***	0.37***	-0.18*
	(0.15)	(0.41)	(0.49)	(0.07)	(0.09)	(0.11)

Notes: 118 firms in Panel A, 99 in the first two rows of Panel B, 113 in the last two of rows of Panel B. Debt is measured in billions of Chilean pesos. Standard deviation in parentheses in columns 1-3, and standard error in parentheses in the last three columns. Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3.2: Firm differences during dictatorship

Differences between firms during dictatorship (1985–1987) and lame duck period (1989), with and without control variables

	Misallocation wedges				*Debt with banks*					
	Capital		*Output*		*Government*		*Other*		*Productivity*	
Dictatorship	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
P	-0.403***	-0.292**	-0.117	-0.215*	12.56**	8.71	66.83***	34.14**	-1.220**	-0.969**
	(0.139)	(0.116)	(0.089)	(0.112)	(5.33)	(5.60)	(13.43)	(13.26)	(0.448)	(0.418)
$\widetilde{P}$	-0.394***	-0.282**	-0.006	-0.036	17.15***	11.99**	39.43***	16.40	-0.784*	-0.467
	(0.149)	(0.128)	(0.058)	(0.091)	(5.84)	(5.07)	(12.26)	(16.04)	(0.90)	(0.423)
Lame Duck										
P	-0.453*	-0.250	0.038	-0.054	26.43**	16.15	65.41***	49.53**	-0.970***	-0.520*
	(0.233)	(0.205)	(0.039)	(0.104)	(13.05)	(15.44)	(13.43)	(22.10)	(0.347)	(0.271)
$\widetilde{P}$	-0.471**	-0.390*	-0.134	-0.254	24.46*	16.57	26.60*	17.63	-0.789**	-0.425
	(0.231)	(0.207)	(0.199)	(0.231)	(14.17)	(10.81)	(13.98)	(12.99)	(0.372)	(0.292)
Firms	91	88	91	88	110	110	110	110	86	86
S.D. dep. var.	0.821	0.821	0.571	0.571	39.95	39.95	95.08	95.08	1.78	1.78
Firm controls	No	Yes	No	Yes	No	Yes	No	Yes	No	No
Industry F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Capital and output wedges were computed using Hsieh and Klenow (2009) formula. P and $\widetilde{P}$ are indicators for firms with direct and indirect political connections, respectively. Firm controls include firm age and set of indicators for exporters, privatized, and large firms. Robust standard errors are reported in parentheses. Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3.3: Firms During Political Transition

Chile's transition to democracy and firms in Augusto Pinochet's network

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A	*Investment*			*Profits*		
$P \times$ Lame Duck	0.012	0.018**	0.018**	0.308***	0.350***	0.290***
	(0.007)	(0.007)	(0.008)	(0.099)	(0.099)	(0.100)
$P \times$ Democracy	0.011*	0.020**	0.020***	0.197	0.254*	0.194
	(0.006)	(0.006)	(0.007)	(0.128)	(0.135)	(0.124)
$\widetilde{P} \times$ Lame Duck		0.013	0.014		0.096**	0.115*
		(0.008)	(0.009)		(0.044)	(0.069)
$\widetilde{P} \times$ Democracy		0.019**	0.020**		0.129	0.147
		(0.008)	(0.008)		(0.127)	(0.095)
Observations	4,694	4,694	4,694	4,692	4,692	4,692
P : Δ Investment lame duck (%)	7.5	11.4	11.4			
$\widetilde{P}$: Δ Investment lame duck (%)	–	8.1	8.7			
Panel B	*Productivity*			*Log workers*		
$P \times$ Lame Duck	0.046	-0.068	-0.065	0.022	0.020	0.028
	(0.158)	(0.174)	(0.177)	(0.106)	(0.106)	(0.108)
$P \times$ Democracy	0.009	-0.013	0.130	-0.040	-0.016	-0.057
	(0.314)	(0.345)	(0.376)	(0.110)	(0.116)	(0.104)
$\widetilde{P} \times$ Lame Duck		-0.224*	-0.230*		-0.006	-0.004
		(0.119)	(0.118)		(0.089)	(0.089)
$\widetilde{P} \times$ Democracy		-0.043	-0.017		0.044	0.048
		(0.214)	(0.229)		(0.111)	(0.110)
Observations	792	792	792	792	792	792
Firm & time F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Industry F.E. $\times$ Post	No	No	Yes	No	No	Yes

Notes: Panel A uses a quarterly panel with 118 firms (104 clusters). Panel B uses an unbalanced annual panel with 99 firms (88 clusters). P and $\widetilde{P}$ are indicators for firms with direct and indirect political connections respectively. Productivity was calculated following Olley and Pakes (1996). Sample period is 1985–1994. Robust standard errors are clustered at the business group level and are reported in parentheses. Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3.4: Robustness checks

Dependent variable is firm investment/profits. Column titles refer to the corresponding robustness exercise

		Control variables						
	Pre/Post	*Large firms*	*Privatized*	*Bus. Group*	*Exporter*	*All*	*Pscore*	*Matching*
Investment	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$P \times$ Lame Duck	0.019**	0.019**	0.017**	0.018**	0.018**	0.018*	0.017*	0.019**
	(0.008)	(0.009)	(0.008)	(0.009)	(0.008)	(0.009)	(0.009)	(0.009)
$P \times$ Democracy	0.019***	0.021***	0.019***	0.020***	0.020***	0.020**	0.016*	0.021***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)	(0.008)
$\widetilde{P} \times$ Lame Duck	0.015*	0.015	0.014	0.014	0.014	0.013	0.013	0.015
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.010)	(0.009)	(0.010)
$\widetilde{P} \times$ Democracy	0.019**	0.021**	0.020**	0.020**	0.020**	0.021**	0.017**	0.016**
	(0.008)	(0.009)	(0.008)	(0.008)	(0.008)	(0.009)	(0.008)	(0.007)
Control $\times$ Post		-0.003	0.002	0.000	-0.000		0.000	
		(0.007)	(0.006)	(0.005)	(0.005)		(0.000)	
Profits								
$P \times$ Lame Duck	0.297***	0.218**	0.228**	0.341***	0.282***	0.218**	0.252**	0.136
	(0.101)	(0.100)	(0.113)	(0.105)	(0.101)	(0.108)	(0.100)	(0.120)
$P \times$ Democracy	0.233*	0.122	0.132	0.245**	0.186	0.121	0.055	0.061
	(0.128)	(0.118)	(0.118)	(0.125)	(0.127)	(0.112)	(0.101)	(0.112)
$\widetilde{P} \times$ Lame Duck	0.106	0.062	0.095	0.194*	0.107	0.109*	0.082	0.070
	(0.064)	(0.078)	(0.080)	(0.110)	(0.080)	(0.106)	(0.083)	(0.072)
$\widetilde{P} \times$ Democracy	0.168*	0.094	0.128	0.226**	0.137	0.140	0.023	0.185
	(0.100)	(0.089)	(0.090)	(0.107)	(0.098)	(0.091)	(0.105)	(0.103)
Control $\times$ Post		0.254***	0.200	-0.194	0.176*		0.016**	
		(0.077)	(0.126)	(0.151)	(0.096)		(0.007)	
Observations	354	4,692	4,692	4,692	4,692	4,692	4,692	3,100
Firm & time F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry F.E. $\times$ Post	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Column 1 collapses the 40-periods panel to a 3-periods panel (dictatorship, lame duck, democracy). Column 2 controls for an indicator that equals one for firms above the median of the firm size distribution before 1988. Column 3 controls for an indicator that equals one for firms that were privatized during Pinochet's dictatorship. Column 4 controls for an indicator that equals one for firms part of a business group at the end of 1987. Column 5 controls for an indicator that equals one for exporter firms. Column 6 controls for all previous variables. Controls in columns 2-6 are always interacted with an indicator for the period after the 1988 plebiscite. Column 7 controls for the propensity score of political connections, interacted with a linear trend. Column 8 uses a matching procedure that drops firms in the tails of the propensity score distribution. The optimal bounds were computed following Crump et al. (2009) and leaves us with 78 firms. Sample period is 1985–1994. Robust standard errors are clustered at the business group level and are reported in parentheses. Columns 1 to 7 have 104 clusters and column 8 has 75 clusters. Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3.5: The Credit Market During Political Transition

Dependent variable (Debt) is total debt with banks, measured from annual statements

	Debt	$\mathbf{1}[\text{Debt} > 0]$	Debt over assets
	(1)	(2)	(3)
$P \times$ Lame Duck $\times$ Government bank	29.77***	0.20**	0.09**
	(10.21)	(0.08)	(0.03)
$\widetilde{P} \times$ Lame Duck $\times$ Government bank	28.40	0.09	0.07*
	(20.10)	(0.07)	(0.04)
$P \times$ Democracy $\times$ Government bank	4.45	0.22**	0.07*
	(14.96)	(0.09)	(0.04)
$\widetilde{P} \times$ Democracy $\times$ Government bank	17.16	0.14	0.08
	(19.47)	(0.09)	(0.05)
$P \times$ Lame Duck	-18.98**	-0.20**	-0.08**
	(9.54)	(0.07)	(0.04)
$P \times$ Democracy	-7.41	-0.20**	-0.05
	(15.94)	(0.08)	(0.04)
$\widetilde{P} \times$ Lame Duck	-21.36	-0.03	-0.07*
	(15.68)	(0.08)	(0.04)
$\widetilde{P} \times$ Democracy	-14.82	-0.10	-0.07
	(15.50)	(0.07)	(0.05)
Lame Duck $\times$ Government Bank	2.44	-0.06	-0.02
	(5.69)	(0.04)	(0.03)
Democracy $\times$ Government Bank	1.94	-0.08**	-0.02
	(6.29)	(0.04)	(0.03)
Firm-bank F.E.	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes
Industry F.E. $\times$ Post	Yes	Yes	Yes
Mean of dep. variable	29.17	0.38	0.06
Firms	113	113	113
Observations	2,075	2,075	2,075

Notes: Estimation using the annual panel dataset of firms in the period 1985–1994. P and $\widetilde{P}$ are indicators for firms with direct and indirect political connections, respectively. Robust standard errors are clustered at the business group level and are reported in parentheses. The number of clusters is 99. Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Bibliography

Abadie, Alberto; Diamond, Alexis and Hainmueller, Jens: ≪Synthetic control methods for comparative case studies: estimating the effect of California's tobacco control program≫. *Journal of the American Statistical Association*, 2010, **105(490)**.

Abadie, Alberto and Gardeazabal, Javier: ≪The economic costs of conflict: a case study of the Basque Country≫. *American Economic Review*, 2003, **93(1)**, pp. 113–132.

Acemoglu, Daron: ≪Oligarchic versus democratic societies≫. *Journal of the European Economic Association*, 2008, **6(1)**, pp. 1–44.

Acemoglu, Daron; Hassan, Tarek and Tahoun, Ahmed: ≪The power of the street: evidence from Egypt's Arab Spring≫. *Working Paper*, 2016a.

Acemoglu, Daron; Johnson, Simon; Kermani, Amir; Kwak, James and Mitton, Todd: ≪The value of connections in turbulent times: evidence from the United States≫. *Journal of Financial Economics*, 2016b.

Acemoglu, Daron; Naidu, Suresh; Restrepo, Pascual and Robinson, James: ≪Democracy does cause growth≫, 2015.

Acemoglu, Daron and Robinson, James: *Economic Origins of Dictatorship and Democracy*. Cambridge University Press, 2009.

Acemoglu, Daron; Ticchi, Davide and Vindigni, Andrea: ≪Emergence and persistence of inefficient states≫. *Journal of the European Economic Association*, 2011, **9(2)**, pp. 177–208.

Adimark: ≪Encuesta: evaluación gestión del gobierno≫. *Adimark GfK*, 2011.

Aidt, Toke S. and Franck, Raphaël: ≪Democratization under the threat of revolution: evidence from the great reform act of 1832≫. *Econometrica*, 2015, **83(2)**, pp. 505–547.

Alfaro, Rodrigo and Fuenzalida, Marcelo: ≪Imputación múltiple en encuestas microeconómicas≫. *Cuadernos de Economía*, 2009, **46**, pp. 273–288.

Allende, Claudia: ≪The impact of information on academic achievement and school choice: evidence from Chilean "Traffic lights"≫, 2012.

Altonji, Joseph G.; Elder, Todd E. and Taber, Christopher R.: ≪Selection on observed and unobserved variables: assessing the effectiveness of catholic schools≫. *Journal of Political Economy*, 2005, **113(1)**, pp. 151–184.

Andrabi, Tahir; Das, Vishnu and Khwaja, Asim: ≪Report cards: the Impact of Providing School and Child Test Scores on Education Markets≫. *American Economic Review*, 2017.

Angrist, Joshua D.: ≪The perils of peer effects≫. *Labour Economics*, 2014, **(30)**, pp. 98–108.

Apperson, Jarod; Bueno, Carycruz and Sass, Tim R.: ≪Do the cheated ever prosper? The long-run effects of test-score manipulation by teachers on student outcomes≫, 2016.

Araos, Consuelo; Cea, Macarena; Fernández, Matías and Valenzuela, Eduardo: ≪The role of school context on marijuana use in Chile: a classroom-level analysis≫. *Deviant Behavior*, 2014, **35**, pp. 412–432.

Bagwell, Kyle: ≪The Economic Analysis of Advertising≫. In: M. Armstrong and R. Porter (Eds.), *Handbook of Industrial Organization*, volume 3, pp. 1701–1844. Elsevier. ISSN 1573-448X, 2007. doi: http://dx.doi.org/10.1016/S1573-448X(06)03028-7.
`http://www.sciencedirect.com/science/article/pii/S1573448X06030287`

Banerjee, Abhijit V. and Duflo, Esther: ≪Do firms want to borrow more? Testing credit constraints using a directed lending program≫. *Review of Economic Studies*, 2014, **(81)**, pp. 572–607.

Bau, Natalie: ≪School Competition and Product Differentiation≫, 2015.

Bayer, Patrick; Ferreira, Fernando and McMillan, Robert: ≪A Unified Framework for Measuring Preferences for Schools and Neighborhoods≫. *Journal of Political Economy*, 2007, **115(4)**, pp. 588–638. ISSN 00223808.
`http://www.jstor.org/stable/10.1086/522381`

Bellei, Cristián: ≪Does lengthening the school day increase students' academic achievement? Results from a natural experiment in Chile≫. *Economics of Education Review*, 2009, **28(5)**, pp. 629–640.

Belloni, Alexandre; Chernozhukov, Victor and Hansen, Christian: ≪Inference on treatment effects after selection among high-dimensional controls≫. *Review of Economic Studies*, 2013, **81(2)**, pp. 608–650.

Bernheim, B. Douglas and Rangel, Antonio: ≪Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics*≫. *The Quarterly Journal of Economics*, 2009, **124(1)**, p. 51.

Berry, Steven; Levinsohn, James and Pakes, Ariel: ≪Automobile Prices in Market Equilibrium≫. *Econometrica*, 1995, **63(4)**, pp. 841–890.

Bertrand, Marianne; Duflo, Esther and Mullainathan, Sendhil: «How much should we trust differences-in-differences estimates?» *Quarterly Journal of Economics*, 2004, **119(1)**.

Bertrand, Marianne; Kramarz, Francis; Schoar, Antoinette and Thesmar, David: «Politicians, firms, and the political business cycle: evidence from France», 2007.

Bloom, Nicholas: «The impact of uncertainty shocks». *Econometrica*, 2009, **77(3)**, pp. 623–685.

Blume, Lawrence E.; Brock, William A.; Durlauf, Steven N. and Jayaraman, Rajshri: «Linear social interactions models». *Journal of Political Economy*, 2015, **123(2)**, pp. 444–496.

Bramoullé, Yann; Djebbari, Habiba and Fortin, Bernard: «Identification of peer effects through social networks». *Journal of Econometrics*, 2009, **(150)**, pp. 51–55.

Bramoullé, Yann; Kranton, Rachel and D'Amours, Martin: «Strategic interaction and networks». *American Economic Review*, 2014, **104(3)**, pp. 898–930.

Brock, William A. and Durlauf, Steven N.: «Discrete choice with social interactions». *Review of Economic Studies*, 2001, **68**, pp. 235–260.

Brough, Wayne T. and Kimenyi, Mwangi S.: «On the inefficient extraction of rents by dictators». *Public Choice*, 1986, **(48)**, pp. 37–48.

Burgess, Robin; Jedwab, Remi; Miguel, Edward; Morjaria, Ameet and i Miquel, Gerard Padró: «The value of democracy: Evidence from road building in Kenya». *American Economic Review*, 2015, **105(6)**, pp. 1817–1851.

Caeyers, Bet and Fafchamps, Marcel: «Exclusion bias in the estimation of peer effects». *NBER working paper 22565*, 2016.

Cameron, A Colin; Gelbach, Jonah B and Miller, Douglas L: «Bootstrap-based improvements for inference with clustered errors». *The Review of Economics and Statistics*, 2008, **90(3)**, pp. 414–427.

Campbell, Donald: «Assessing the impact of planned social change». *Evaluation and Program Planning*, 1979, **2**, pp. 67 – 90.

Card, David; Mas, Alexandre and Rothstein, Jesse: «Tipping and the dynamics of segregation». *Quarterly Journal of Economics*, 2008, **123(1)**, pp. 177–218.

Carneiro, Pedro; Das, Jishnu and Reis, Hugo: «The Value of Private Schools: Evidence from Pakistan», 2016.

Cartwright, Edward and Menezes, Matheus L.C.: «Cheating to win: dishonesty and the intensity of competition». *Economics Letters*, 2014, **(122)**, pp. 55–58.

Cauce: «Mito, Temores y Encuestas». *Cauce Magazine*, 1988.

Cavallo, Ascanio; Salazar, Manuel and Sepúlveda, Oscar: *La Historia Oculta del Régimen Militar: Memoria de una Época 1973–1988*. Uqbar editores, 2011.

CEME: «Informe de la comisión investigadora encargada de analizar presuntas irregularidades en las privatizaciones de empresas del estado ocurridas con anterioridad al año 1990». *Centro de Estudios Miguel Enriquez*, 2004.

Centro de Investigación y Desarrollo de la Educación: «VIII Encuesta a Actores del Sistema Educativo 2010», 2010.

Chandrasekhar, Arun and Lewis, Randall: «Econometrics of sampled networks». *Working Paper*, 2011.

Chay, Kenneth Y.; McEwan, Patrick J. and Urquiola, Miguel: «The central role of noise in evaluating interventions that use test scores to rank schools». *American Economic Review*, 2005, **95(4)**, pp. 1237–1258.

Chumacero, Rómulo A.; Gómez, Daniel and Paredes, Ricardo: «I would walk 500 miles (if it paid): Vouchers and school choice in Chile». *Economics of Education Review*, 2011, **30**, pp. 1103–1114.

Cingano, Federico and Pinotti, Paolo: «Politicians at work: the private returns and social costs of political connections». *Journal of the European Economic Association*, 2013, **11(2)**, pp. 433–465.

CIPER: «Ley SEP: Contraloría revela masiva pérdida de recursos destinados a educación de los más pobres». *Centro de Investigación Periodística*, 2012.

Claessens, Stijn; Feijen, Erik and Laeven, Luc: «Political connections and preferential access to finance: the role of campaign contributions». *Journal of Financial Economics*, 2008, **88**, pp. 554–580.

Confech: «Confederation of Chilean students to Joaquín Lavín». Letter, 2011.

Contreras, Dante and Rau, Tomás: «Tournament Incentives for Teachers: evidence from a Scaled-Up Intervention in Chile». *Economic Development and Cultural Change*, 2012, **61(1)**, pp. 219–246.

Cooper, Ryan; Gallego, Francisco and Neilson, Christopher: «Types of Information and School Choice: an Experimental Study in the Chilean Voucher System», 2013.

Corbo, Vittorio: «Reforms and macroeconomic adjustments in Chile during 1974-84». *World Development*, 1985, **13(8)**, pp. 893–916.

Cragg, John G. and Donald, Stephen G.: «Testing Identifiability and Specification in Instrumental Variable Models». *Econometric Theory*, 1993, **9**, pp. 222–240. ISSN 1469-4360.

Crump, Richard K.; Hotz, V. Joseph; Imbens, Guido W. and Mitnik, Oscar A.: ≪Dealing with limited overlap in estimation of average treatment effects≫. *Biometrika*, 2009, **96(1)**, pp. 187–199.

De Giorgi, Giaocomo; Pellizzari, Michele and Redaelli, Silvia: ≪Identification of social interactions through partially overlapping peer groups≫. *American Economic Journal: Applied Economics*, 2010, **(2)**, pp. 241–275.

Dee, Thomas S.; Dobbie, Will; Jacob, Brian A. and Rockoff, Jonah: ≪The causes and consequences of test score manipulation: Evidence from the New York regents examinations≫, 2016.

Deming, David J.; Cohodes, Sarah; Jennings, Jennifer and Jencks, Christopher: ≪School Accountability, Postsecondary Attainment and Earnings≫. *Review of Economics and Statistics*, 2016.

Diamond, Rebecca and Persson, Petra: ≪The long-term consequences of teacher discretion in grading of high-stakes tests≫, 2016.

Díaz, Juan; Grau, Nicolás; Reyes, Tatiana and Rivera, Jorge: ≪2016≫. *Working Paper*, 2016.

Dixit, Avinash: ≪The role of investment in entry deterrence≫. *Economic Journal*, 1980, **90(357)**, pp. 95–106.

Do, Quoc-Anh; Lee, Yen-Teik and Nguyen, Bang Dang: ≪Political connections and firm value: evidence from the regression discontinuity design of close gubernatorial elections≫, 2015.

Dorfman, Robert and Steiner, Peter O.: ≪Optimal Advertising and Optimal Quality≫. *The American Economic Review*, 1954, **44(5)**, pp. 826–836.

Dranove, David and Jin, Ginger Zhe: ≪Quality Disclosure and Certification: theory and Practice≫. *Journal of Economic Literature*, 2010, **48(4)**, pp. 935–963.

Dranove, David; Kessler, Daniel; McClellan, Mark and Satterthwaite, Mark: ≪Is More Information Better? The Effects of Report Cards on Health Care Providers≫. *Journal of Political Economy*, 2003, **111(3)**, pp. 555–588.

Dube, Arindrajit; Kaplan, Ethan and Naidu, Suresh: ≪Coups, corporations, and classified information≫. *Quarterly Journal of Economics*, 2011, **126(3)**, pp. 1375–1409.

Ebenstein, Avraham; Lavy, Victor and Roth, Sefi: ≪The long run consequences of high-stakes examinations: Evidence from transitory variation in pollution≫. *American Economic Journal: Applied Economics*, 2016.

Edwards, Sebastian: ≪Monetarism in Chile, 1973–1983: some economic puzzles≫. *Economic Development and Cultural Change*, 1986, **34(3)**, pp. 535–559.

Ellman, Matthew and Wantchekon, Leonard: ≪Electoral competition under the threat of political unrest≫. *Quarterly Journal of Economics*, 2000, **115(2)**, pp. 499–531.

Engel, Eduardo and Venetoulias, Achilles: ≪The Chilean plebiscite: projections without historic data≫. *Journal of the American Statistical Association*, 1992, **87(420)**, pp. 933–941.

Enikolopov, Ruben; Makarin, Alexey and Petrova, Maria: ≪Social media and protest participation: evidence from Russia≫. *Working Paper*, 2016.

Feigenberg, Ben: ≪Priced Out: Aggregate Income Shocks and School Pricing in the Chilean Voucher Market≫, 2015.

Figlio, David and Loeb, Susanna: ≪School Accountability≫. In: Stephen Machin Eric A. Hanushek and Ludger Woessmann (Eds.), *Handbook of the Economics of Education*, volume 3, pp. 383–421, 2011.

Figlio, David and Winicki, Joshua: ≪Food for thought: the effects of school accountability plans on school nutrition≫. *Journal of Public Economics*, 2005, **89(2-3)**, pp. 381–394.

Figlio, David N. and Getzler, Lawrence S.: ≪Accountability , Ability and Disability: Gaming the System≫. *Working Paper 9307*, National Bureau of Economic Research, 2002.

Figueroa, Francisco: *Llegamos para quedarnos: crónicas de la revuelta estudiantil*. LOM ediciones, 2012.

Fisman, Raymond: ≪Estimating the value of political connections≫. *American Economic Review*, 2001, **91(4)**, pp. 1095–1102.

Fracassi, Cesare: ≪Corporate finance policies and social networks≫, 2014.

Fryer, Roland: ≪The Production of Human Capital in Developed Countries: Evidence from 196 Randomized Field Experiments≫. volume 1 of *Handbook of Field Experiments*. Elsevier, 2016.

Gallego, Francisco and Hernando, Andrés: ≪School Choice in Chile: Looking at the Demand Side≫. *Pontificia Universidad Catolica de Chile Documento de Trabajo*, 2009, **(356)**.

Gilpatric, Scott M.: ≪Cheating in contests≫. *Economic Inquiry*, 2011, **49(4)**, pp. 1042–1053.

Gladwell, Malcolm: *The Tipping Point: How Little Things Can Make a Big Difference*. Little, Brown and Company, 2000.

Glewwe, Paul and Muralidharan, Karthik: ≪Improving School Education Outcomes in Developing Countries: Evidence, Knowledge Gaps, and Policy Implications≫, 2015.

Gómez, Daniel; Chumacero, Rómulo and Paredes, Ricardo: ≪School choice and information≫. *Estudios de Economía*, 2012, **39(2)**, pp. 143–157.

Graff Zivin, Joshua; Hsiang, Solomon and Neidell, Matthew: ≪Temperature and human capital in the short- and long-run≫, 2015.

Granovetter, Mark: ≪Threshold models of collective behavior≫. *American Journal of Sociology*, 1978, **83(6)**, pp. 1420–1443.

Granovetter, Mark S.: ≪The strength of weak ties≫. *American Journal of Sociology*, 1973, **78(6)**, pp. 1360–1380.

Guryan, Jonathan; Kroft, Kory and Notowidigdo, Matthew J.: ≪Peer effects in the workplace: evidence from random groupings in professional golf tournaments≫. *American Economic Journal: Applied Economics*, 2009, **1(4)**, pp. 34–68.

Handel, Benjamin and Kolstad, Jonathan: ≪Health Insurance for"Humans": Information Frictions, Plan Choice, and Consumer Welfare≫. *American Economic Review*, 2015, **105(8)**, pp. 2449–2500.

Hastings, Justine; Kane, Thomas J. and Staiger, Douglas O.: ≪Heterogeneous Preferences and the Efficacy of Public School Choice≫, 2009.

Hastings, Justine S. and Weinstein, Jeffrey M.: ≪Information, School Choice, and Academic Achievement: Evidence from Two Experiments≫. *Quarterly Journal of Economics*, 2008, **123(4)**, pp. 1373–1414.

Haunschild, Pamela R. and Beckman, Christine M.: ≪When do interlocks matter?: Alternate sources of information and interlock influence≫. *Administrative Science Quarterly*, 1998.

Holmstrom, Bengt and Milgrom, Paul: ≪Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design≫. *Journal of Law, Economics, & Organization*, 1991, **7**, pp. 24–52.

Hsieh, Chang-Tai and Klenow, Peter J: ≪Misallocation and manufacturing TFP in China and India≫. *Quarterly Journal of Economics*, 2009, **124(4)**, pp. 1403–1448.

Hsieh, Chang-Tau and Urquiola, Miguel: ≪The effects of generalized school choice on achievement and stratification: evidence from Chile's voucher program≫. *Journal of Public Economics*, 2006, **90**, pp. 1477–1503.

Huneeus, Carlos: *The Pinochet Regime*. Lynne Rienner Publishers, 2006.

Jackson, Giorgio: *El país que soñamos*. Debate, 2013.

Jackson, Matthew O.: *Social and Economic Networks*. Princeton University Press, 2010.

Jackson, Matthew O and Rogers, Brian W: ≪The economics of small worlds≫. *Journal of the European Economic Association*, 2005, **3(2-3)**, pp. 617–627.

Jacob, Brian A.: ≪Accountability, incentives and behavior: the impact of high-stakes testing in the Chicago Public Schools≫. *Journal of Public Economics*, 2005, **89**, pp. 761 – 796. ISSN 0047-2727.

Jacob, Brian A. and Levitt, Steven D.: ≪Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating≫. *Quarterly Journal of Economics*, 2003, **118(3)**, pp. 843–877.

Jensen, Robert: ≪The (perceived) returns to education and the demand for schooling≫. *Quarterly Journal of Economics*, 2010, **125(2)**, pp. 515–548.

Jin, Ginger Zhe and Leslie, Phillip: ≪The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards≫. *The Quarterly Journal of Economics*, 2003, **118(2)**, pp. 409–451.

Julio, Brandon and Yook, Youngsuk: ≪Political uncertainty and corporate investment cycles≫. *Journal of Finance*, 2012, **LXVII(1)**, pp. 45–83.

Kane, Thomas J. and Staiger, Douglas O.: ≪The Promise and Pitfalls of Using Imprecise School Accountability Measures≫. *The Journal of Economic Perspectives*, 2002, **16(4)**, pp. 91–114.

Kennedy, Peter E.: ≪Estimation with correctly interpreted dummy variables in semilogarithmic equations≫. *American Economic Review*, 1981, **71(4)**, p. 801.

Kennickell, Arthur B.: ≪Multiple imputation in the Survey of Consumer Finances≫. *Woking Paper*, 1998.

Khwaja, Asim Ijaz and Mian, Atif: ≪Do lenders favor politically connected firms? Rent provision in an emerging financial market≫. *Quarterly Journal of Economics*, 2005, **120(4)**, pp. 1371–1411.

Khwaja, Asim Ijaz; Mian, Atif and Qamar, Abid: ≪Bank credit and business networks≫, 2011.

Kossinets, Gueorgi: ≪Effects of missing data in social networks≫. *Social Networks*, 2006, **28**, pp. 247–268.

Kremer, Michael and Holla, Alaka: ≪Improving Education in the Developing World: What Have We Learned from Randomized Evaluations?≫ *Annual Review of Economics*, 2009, **1(1)**, pp. 513–542.

Kuran, Timur: ≪Sparks and prairie fires: a theory of unanticipated political revolution≫. *Public Choice*, 1989, **61**, pp. 41–74.

Lambson, Val Eugene and Jensen, Farrell E: ≪Sunk costs and firm value variability: theory and evidence≫. *American Economic Review*, 1998, **88(1)**, pp. 307–313.

Laumann, Edward O.; Marsden, Peter V. and Prensky, David: ≪The boundary specification problem in network analysis≫. In: Ronald S. Burt and Michael J. Minor (Eds.), *Applied Network Analysis: A Methodological Introduction*, Wiley-Interscience, 1983.

Leon-Dermota, Ken: *...And Well Tied Down: Chile's Press Under Democracy*. Praeger, 2003.

Lindberg, Staffan I.: *Democratization by Elections: A New Mode of Transition*. Johns Hopkins University Press, 2009.

Lohmann, Susanne: ≪A signaling model of informative and manipulative political action≫. *American Political Science Review*, 1993, **87(2)**, pp. 319–333.

Madestam, Andreas; Shoag, Daniel; Veuger, Stan and Yanagizawa-Drott, David: ≪Do political protest matter? Evidence from the Tea Party Movement≫. *Quarterly Journal of Economics*, 2013, **128(4)**, pp. 1633–1685.

Manacorda, Marco: ≪The cost of grade retention≫. *Review of Economic and Statistics*, 2012, **94(2)**, pp. 596–606.

Manacorda, Marco and Tesei, Andrea: ≪Liberation technology: mobile phones and political mobilization in Africa≫. *Working Paper*, 2016.

Manresa, Elena: ≪Estimating the structure of social interactions using panel data≫. *Working Paper*, 2016.

Manski, Charles: ≪Identification of endogenous social effects: the reflection problem≫. *Review of Economic Studies*, 1993, **(60)**, pp. 531–542.

Martínez Bravo, Monica: ≪The role of local officials in new democracies: evidence from Indonesia≫. *American Economic Review*, 2014, **104(4)**, pp. 1244–87.

Marwell, Gerald and Oliver, Pamela: *The critical mass in collective action: a micro-social theory*. Cambridge University Press, 1993.

Mas, Alexandre and Moretti, Enrico: ≪Peers at Work≫. *American Economic Review*, 2009, **99(1)**, pp. 112–145.

McAdam, Doug: ≪Recruitment to high-risk activism: the case of Freedom Summer≫. *American Journal of Sociology*, 1986, **92(1)**, pp. 64–90.

McEwan, Patrick J.; Urquiola, Miguel and Vegas, Emiliana: ≪School Choice, Stratification and Information on School Performance: Lessons from Chile≫. *Economia*, 2008.

Meghir, Costas and Rivkin, Steven: ≪Econometric Methods for Research in Education≫. volume 3 of *Handbook of the Economics of Education*, pp. 1 – 87. Elsevier, 2011.

Méndez, Roberto; Godoy, Oscar; Barros, Enrique and Fontaine, Arturo: ≪¿Por qué ganó el no?≫ *Centro de Estudios Públicos*, 1988.

Ministry of Education: ≪Estadísticas de la Educación 2013≫, 2013.

Mizala, Alejandra and Urquiola, Miguel: ≪School markets: The impact of information approximating schools' effectiveness≫. *Journal of Development Economics*, 2013, **103(0)**, pp. 313–335.

Mocetti, Sauro: ≪Dynasties in professions and the role of rents and regulation: Evidence from Italian pharmacies≫. *Journal of Public Economics*, 2016, **(133)**, pp. 1–10.

Mönckeberg, María Olivia: *El Saqueo de los Grupos Económicos al Estado Chileno*. De Bolsillo, 2015.

Neal, Derek: ≪The Consequences of Using one Assessment System to Pursue two Objectives≫. *Journal of Economic Education*, 2013, **44(4)**, pp. 339–352.

Neal, Derek and Schanzenbach, Diane Whitmore: ≪Left behind by design: proficiency counts and test-based accountability≫. *Review of Economics and Statistics*, 2010, **92(2)**, pp. 263–283.

Neilson, Christopher: ≪Targeted Vouchers, Competition Among Schools, and the Academic Achievement of Poor Students≫, 2013.

Nelson, Phillip: ≪Information and Consumer Behavior≫. *Journal of Political Economy*, 1970, **78(2)**, pp. 311–329.

Newey, Whitney; Powell, James and Vela, Francis: ≪Nonparametric estimation of triangular simultaneous equations models≫. *Econometrica*, 1999, **67(3)**, pp. 565–603.

Newman, Mark EJ: ≪Fast algorithm for detecting community structure in networks≫. *Physical review E*, 2004, **69(6)**, p. 066133.

O'Donnell, Guillermo and Schmitter, Phillipe C.: *Transitions from Authoritarian Rule: Tentative Conclusions about Uncertain Democracies*. The Johns Hopkins University Press, 1986.

OECD: ≪Education at a glance 2013: OECD indicators≫. *OECD publishing*, 2013.

Olken, Benjamin A.: ≪Monitoring corruption: Evidence from a field experiment in Indonesia≫. *Journal of Political Economy*, 2007, **115(2)**, pp. 200–249.

Olley, G. Steven and Pakes, Ariel: ≪The dynamics of productivity in the telecommunications equipment industry≫. *Econometrica*, 1996, **64**, pp. 1263–1297.

Olson, Mancur: *The logic of collective action. Public goods and the theory of groups*. Cambridge, MA: Harvard University Press, 1965.

Ossandón, José and Tironi, Eugenio (Eds.): *Adaptación: La Empresa Chilena Después de Friedman*. Ediciones Universidad Diego Portales, 2013.

Oster, Emily: ≪Unobservable selection and coefficient stability: theory and evidence≫. *Journal of Business Economics and Statistics*, 2016.

Papaioannou, Elias and Siourounis, Gregorios: ≪Economic and social factors driving the third wave of democratization≫. *Journal of Comparative Economics*, 2008, **36**, pp. 365–387.

Patnam, Manasa: ≪Corporate networks and peer effects in firm policies≫, 2013.

Quezada-Hofflinger, Alvaro and Von Hippel, Paul T.: ≪The response to high stakes testing in Chile, 2005-2013: Legitimate and illegitimate ways to raise test scores≫, 2017.

Ramos, Joseph R.: ≪The economics of hyperstagflation: stabilization policy in post 1973 Chile≫. *Journal of Development Economics*, 1980, **7**, pp. 467–488.

Reback, Randall and Cullen, Julie Berry: ≪Tinkering toward accolades: school gaming under a performance accountability system≫. *Working Papers 0601*, Barnard College, Department of Economics, 2006.

Rockoff, Jonah and Turner, Lesley J: ≪Short-run impacts of accountability on school quality≫. *American Economic Journal: Economic Policy*, 2010, **2(4)**, pp. 119–147.

Roland, Gérard: ≪The political economy of transition≫. *Journal of economic Perspectives*, 2002, pp. 29–50.

Rubin, Donald B.: *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley & Sons, 1987.

Schelling, Thomas C.: ≪Dynamic models of segregation≫. *Journal of Mathematical Sociology*, 1971, **1**, pp. 143–186.

Schwieren, Christiane and Weichselbaumer, Doris: ≪Does competition enhance performance or cheating? A laboratory experiment≫. *Journal of Economic Psychology*, 2010, **(31)**, pp. 241–253.

Shleifer, Andrei: ≪Does Competition Destroy Ethical Behavior?≫ *American Economic Review*, 2004, **94(2)**, pp. 414–418.

Silva, Eduardo: ≪From dictatorship to democracy: the business-state nexus in Chile's economic transformation, 1975–1994≫. *Comparative Politics*, 1996, **28(3)**, pp. 299–320.

Simonsen, Elizabeth: *Mala educación: historia de la revolución escolar*. Debate, 2012.

Small, Kenneth and Rosen, Harvey: ≪Applied Welfare Economics with Discrete Choice Models≫. *Econometrica*, 1981, **49(1)**, pp. 105–30.

SPINNEWIJN, JOHANNES: «Heterogeneity, Demand for Insurance, and Adverse Selection». *American Economic Journal: Economic Policy*, 2016.

STEVENSON, MEGAN: «Test of random assignment to peers in the face of mechanical negative correlation: an evaluation of four techniques». *Working Paper*, 2015.

STOCK, J. and YOGO, M.: «Testing for weak instruments in linear IV regression». In: *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, chapter 5, pp. 80–108, 2005.

SUPERINTENDENCIA DE VALORES Y SEGUROS: «Circular Número 766», 1988.

TILLY, CHARLES: *From mobilization to revolution*. McGraw-Hill, 1978.

TRAIN, KENNETH: «Welfare Calculations in Discrete Choice Models when Anticipated and Experienced Attributes Differ: A Guide with Examples». *Journal of Choice Modelling*, 2015, **16**, pp. 15–22.

UNITED NATIONS: «International standard industrial classification of all economic activities». *Statistical Papers Series M No.4/Rev.4*, Department of Economic and Social Affairs, 2008.

VALLEJO, CAMILA: *Podemos cambiar el mundo*. Ocean Sur, 2012.

WALTERS, CHRISTOPHER: «The Demand for Effective Charter Schools», 2014.

YANAGIZAWA-DROTT, DAVID: «Propaganda and conflict: evidence from the Rwandan genocide». *Quarterly Journal of Economics*, 2014, **129(4)**, pp. 1–48.

Table A.1: Linear estimates – reduced form and first stage

Dependent variable is network absenteeism (Panel A) and student absenteeism (Panel B) in June 16, 2011

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A – First stage							
Instrument	1.00***	0.96***	0.68***	0.79***	0.69***	0.62***	0.47***
	(0.14)	(0.14)	(0.12)	(0.13)	(0.14)	(0.16)	(0.05)
Panel B – Reduced form							
Instrument	0.80***	0.74***	0.47***	0.64***	0.48***	0.39*	0.03***
	(0.15)	(0.15)	(0.14)	(0.15)	(0.16)	(0.22)	(0.01)
Student controls		x	x	x	x	x	x
Network controls			x	x	x	x	x
School controls				x	x	x	
City F.E.					x		
Neighborhood F.E.						x	
School F.E.							x
F-stat 1*st* stage	53.3	50.5	30.6	36.0	24.1	14.0	77.7
Observations	779,327	779,251	771,121	760,801	760,801	49,273	771,121

Notes: *Student controls* include academic performance, average school attendance in previous years and socioeconomic characteristics. *Network controls* include average student controls at the network level. *School controls* include indicators for publicly managed schools, average academic performance of students, and average socioeconomic characteristics of students. *Neighborhoods* are geographic areas where students live. More details in section 1.5. See Figure A.5 for a map of cities. In column 6, each neighborhood is of size 10×10 blocks. Neighborhood data is only available for some students. See Figure A.7 for a map of neighborhoods. Standard errors clustered at the city level are reported in parentheses. Significance level: *** $p < 0.01$.

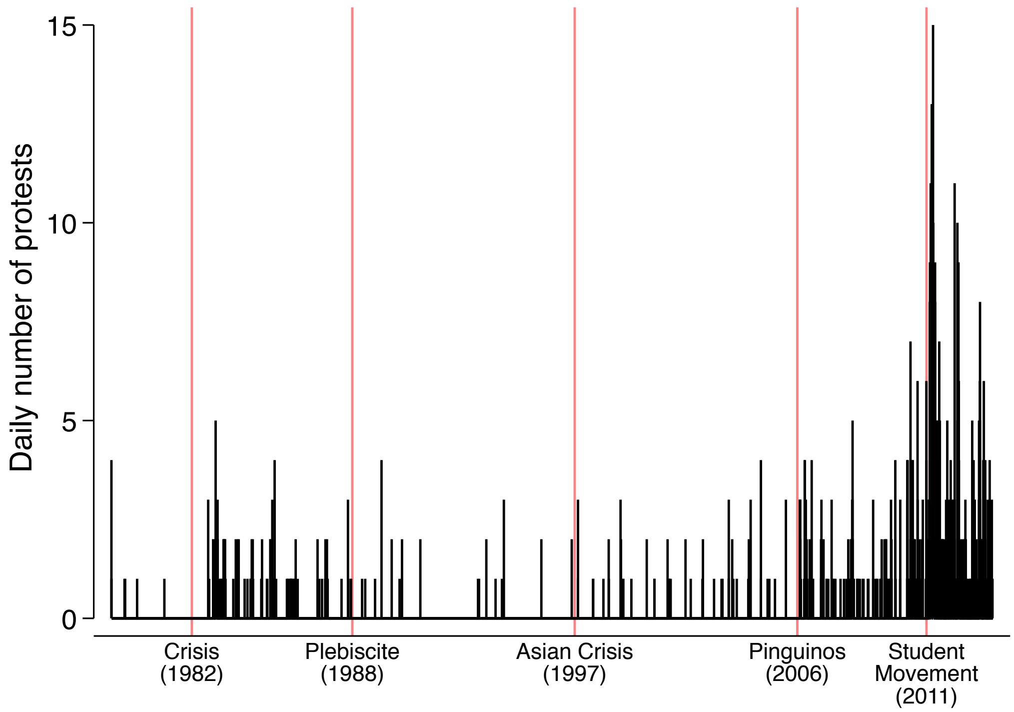
Table A.2: Linear estimates – robustness to school closures

Dependent variable is student absenteeism in June 16, 2011

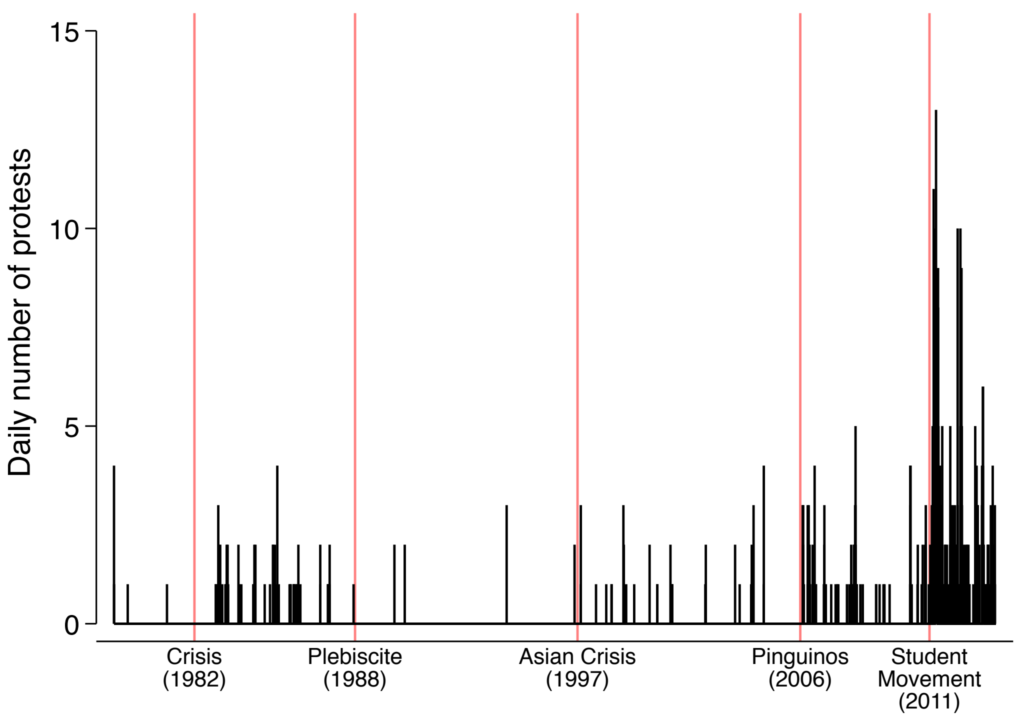
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Network absenteeism in June 16	0.66***	0.57***	0.56***	0.63***	0.53***	0.68***	0.08**
	(0.10)	(0.10)	(0.13)	(0.11)	(0.14)	(0.18)	(0.04)
Absenteeism in May 12	0.14***	0.12***	0.10***	0.09***	0.09***	0.07***	0.09***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
Absenteeism in June 1	0.18***	0.16***	0.14***	0.12***	0.12***	0.13***	0.12***
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.00)
Student controls		x	x	x	x	x	x
Network controls			x	x	x	x	x
School controls				x	x	x	
City F.E.					x		
Neighborhood F.E.						x	
School F.E.							x
F-stat 1*st* stage	23.7	23.7	23.8	24.7	30.2	28.5	183.3
Observations	505,643	505,610	500,834	492,903	492,902	34,145	500,798

Notes: *Student controls* include academic performance, average school attendance in previous years and socioeconomic characteristics. *Network controls* include average student controls at the network level. *School controls* include indicators for publicly managed schools, average academic performance of students, and average socioeconomic characteristics of students. Estimation excludes schools with absenteeism larger than 0.99. *Neighborhoods* are geographic areas where students live. More details in section 1.5. See Figure A.5 for a map of cities. In column 6, each neighborhood is of size 10×10 blocks. Neighborhood data is only available for some students. See Figure A.7 for a map of neighborhoods. Standard errors clustered at the city level are reported in parentheses. Significance level: *** $p < 0.01$.

Figure A.1: Protests in Chile 1979–2013



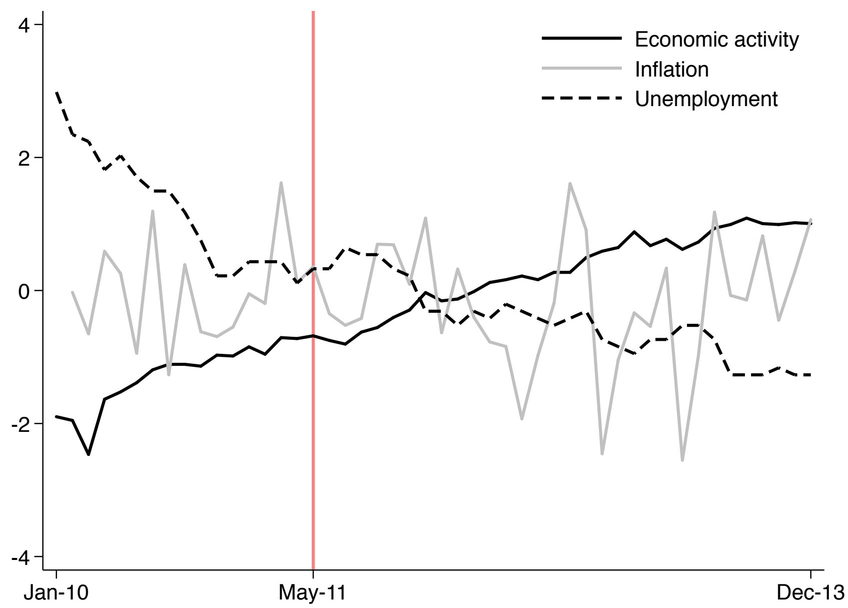
(a) Any type of protest event



(b) Protest events related to education

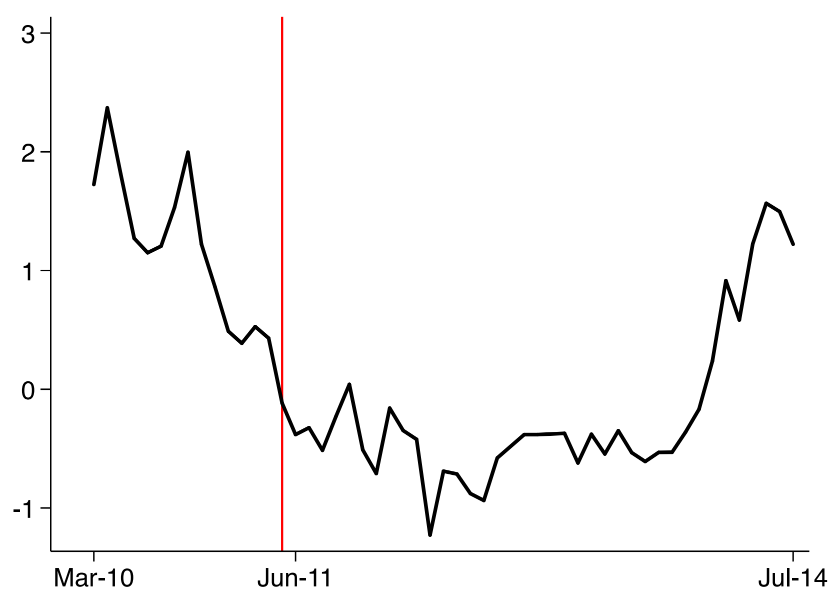
Notes: Data from the Global Dataset of Events, Language, and Tone.

Figure A.2: Economic indicators



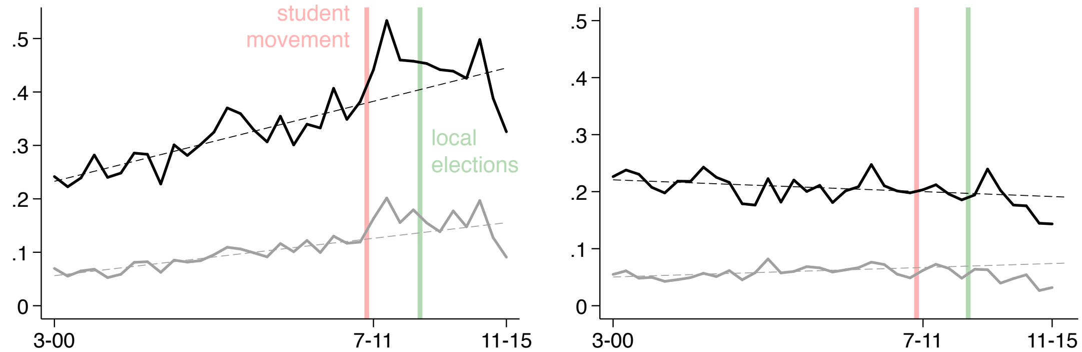
Notes: Data from the Central Bank of Chile. All variables have been normalized. The vertical red line denotes the beginning of the student movement.

Figure A.3: Citizens' evaluation of incumbent politicians

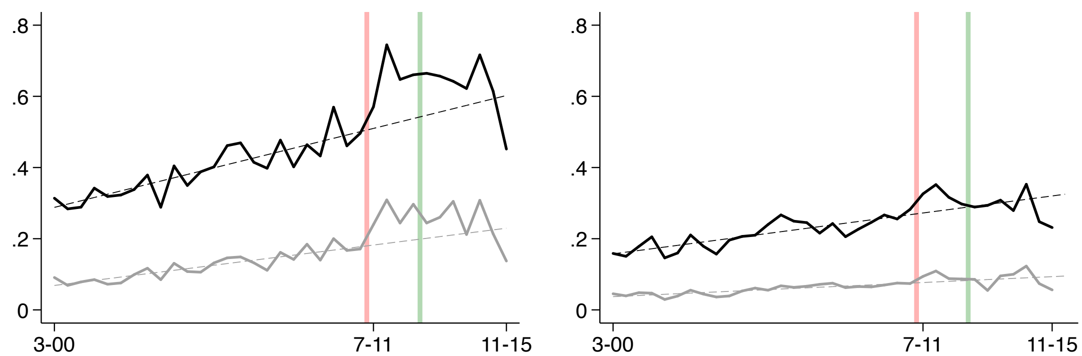


Notes: Normalized index for the approval of incumbent politicians. Data from the Centro de Estudios Públicos and Adimark.

Figure A.4: Survey evidence for the impact of the student movement

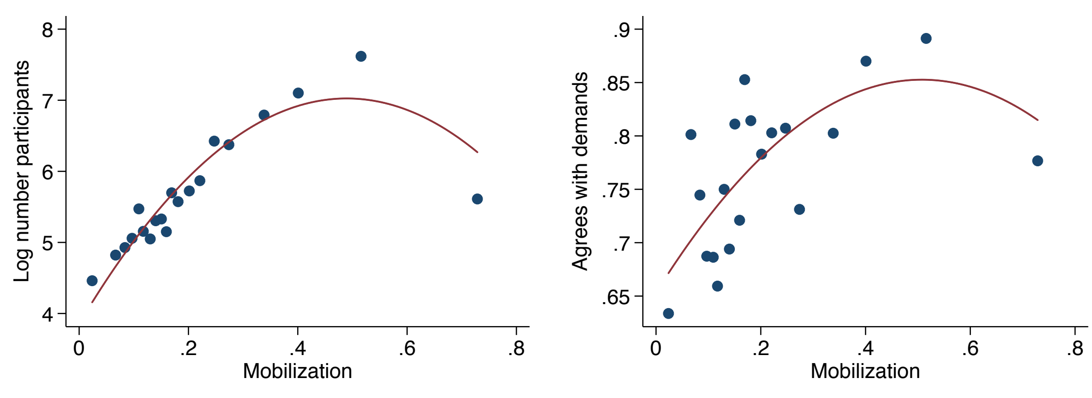


(a) Education should be government's priority (b) Placebo



(c) Individuals 18–44 years old (median) (d) Older than 44 years old (median)

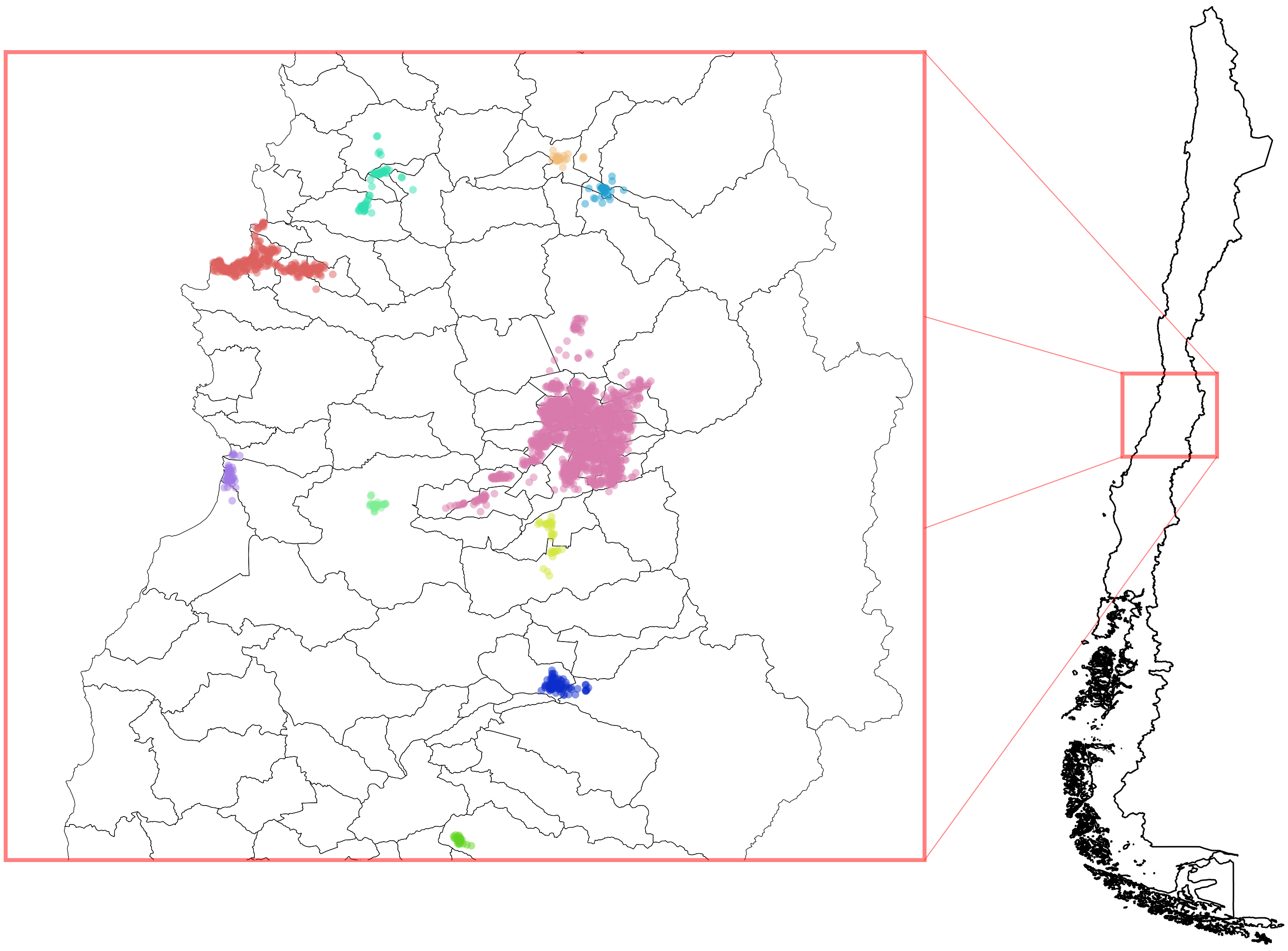
Notes: Panels A-D plot the percentage of people that answer the question "What should be the government's priority?" with "Education" ("Drugs" in Panel B).



(e) Participation in plebiscite (f) Agreement with students' demands

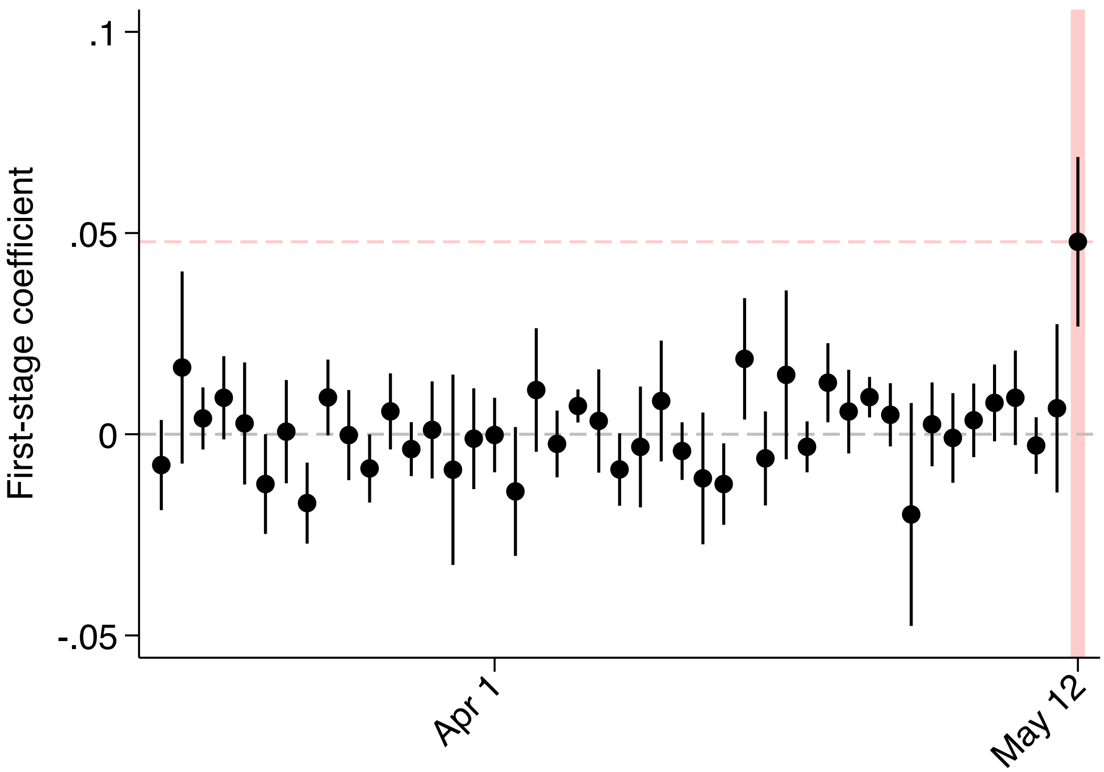
Notes: Panels E and F plot the participation in the "National plebiscite for education" in October of 2011 at the county level (E) and the percentage of people that agrees with the students' demands among those who participated (F).

Figure A.5: Cities



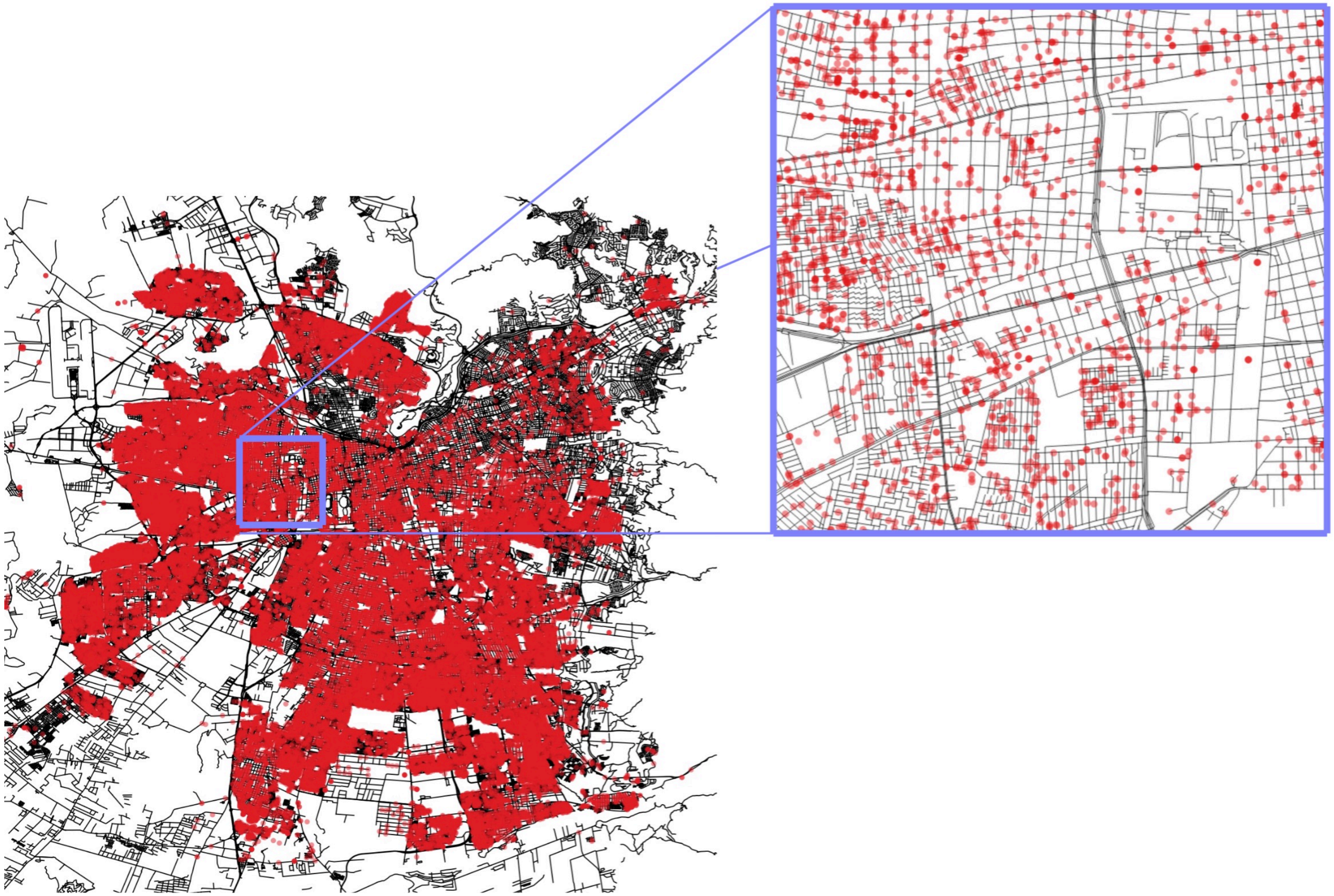
Notes: This map plots the ten largest cities in the most populated area of the country. Cities are defined as closed geographic polygons without any other close by school.

Figure A.6: Placebos for first-stage



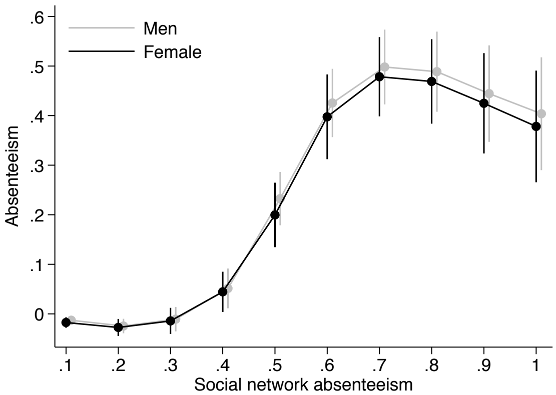
Notes: This figure plots OLS estimates from a single cross-sectional regression. The dependent variable is June 16 school absenteeism in students' social networks. The figure presents standardized coefficients for absenteeism in May 12 among out-of-school students in the social network of social networks. Regression includes student absenteeism in May 12 and June 1, student controls, network controls, school controls, and city fixed effects. Vertical lines denote 95 percent confidence intervals with standard errors clustered at the city level. The coefficient highlighted in red (May 12) corresponds to the first-stage. All other coefficients are placebos for the first-stage. As expected, only 5 percent of coefficients are different from zero.

Figure A.7: Location of students in Santiago

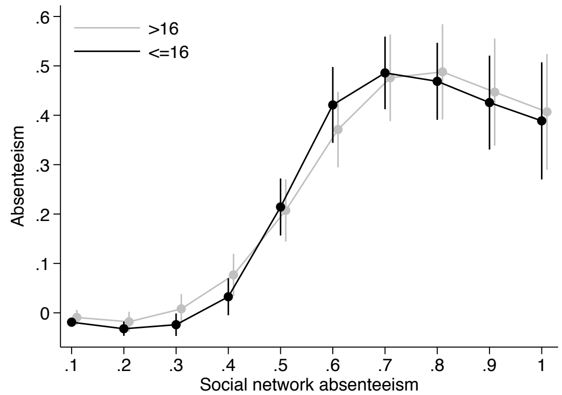


Notes: Home addresses for approximately 50,000 students in Santiago in 2011. The road network is plotted in black for geographic reference.

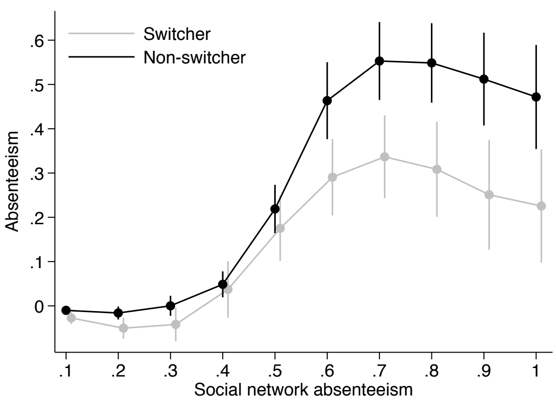
Figure A.8: Threshold model heterogeneity



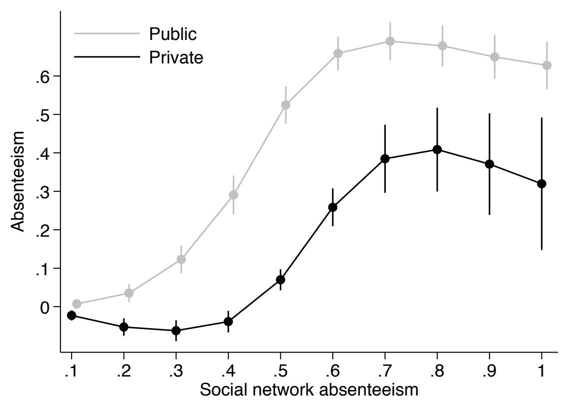
(a) Gender sub-samples



(b) Age sub-samples



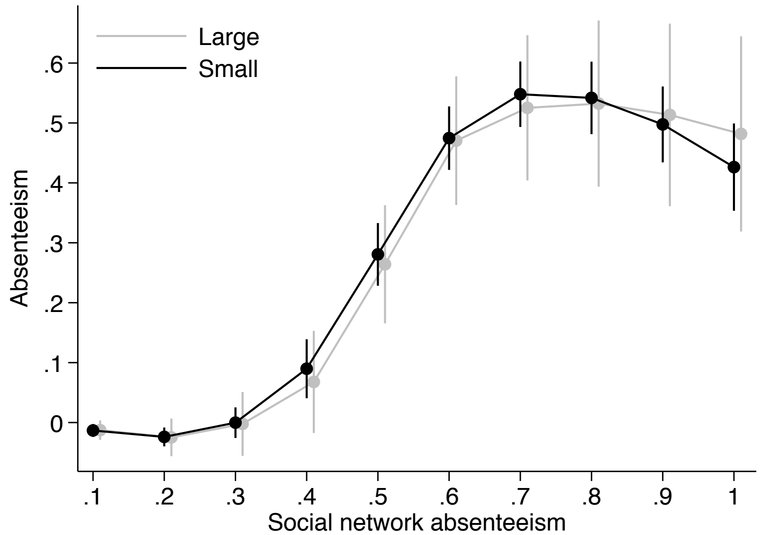
(c) Switcher sub-samples



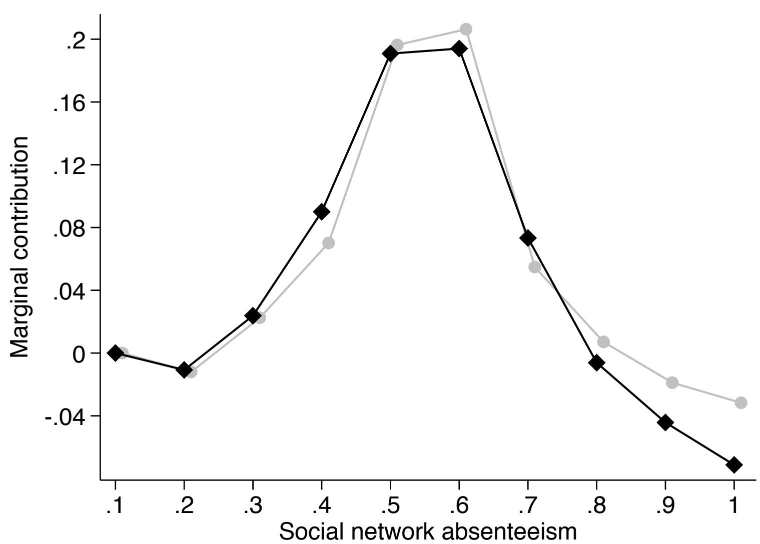
(d) Sub-samples of public and private schools

Notes: All panels plot 2SLS estimates from a regression of individual school absenteeism on 10 indicators of network absenteeism, controlling for school, network, and school characteristics (see equations 1.7 and 1.8) in sub-samples. More details in section 1.5.4.

Figure A.9: Heterogeneity by irregular spending



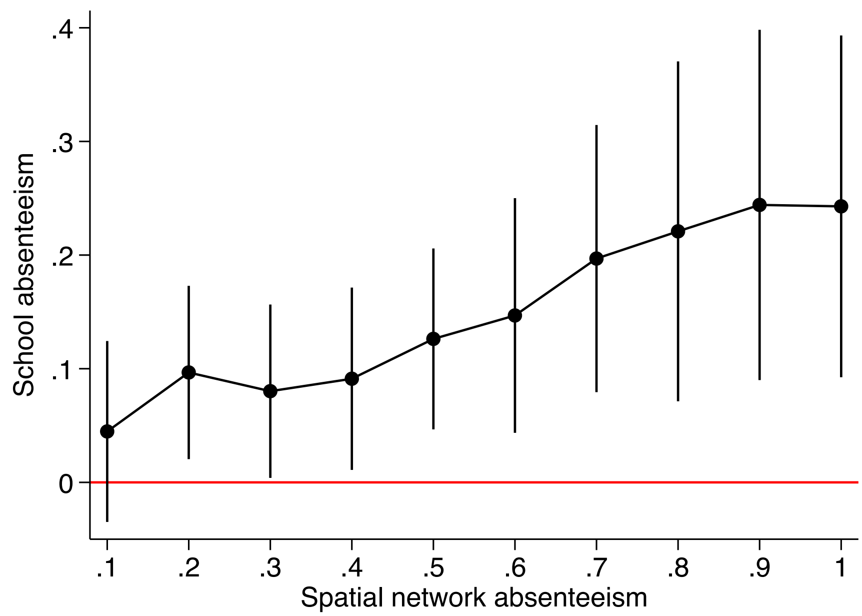
(a) Estimated parameters



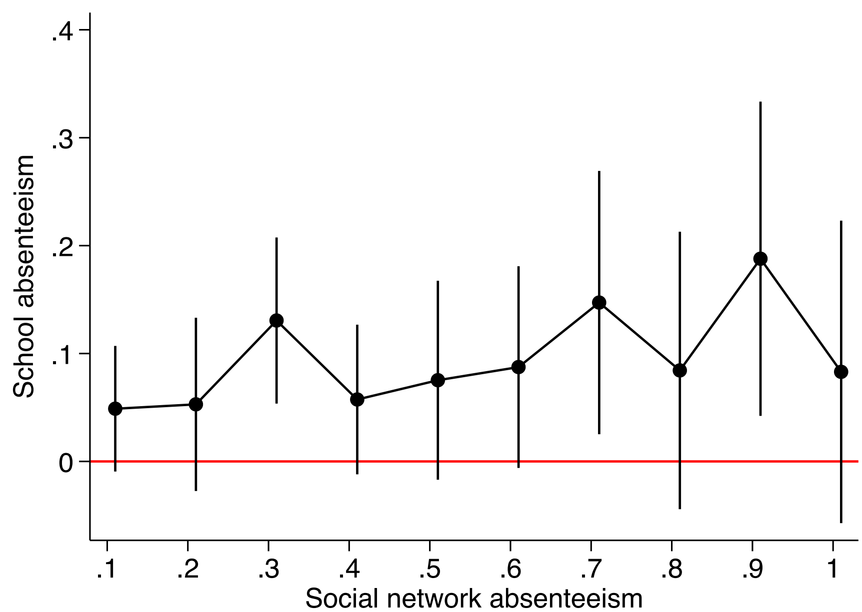
(b) Marginal contribution of additional network absenteeism

Notes: Panel A plots 2SLS estimates from a regression of individual school absenteeism on 10 indicators of network absenteeism in June 16, controlling for school, network, and school characteristics, and city fixed effects (see equations 1.7 and 1.8) in sub-samples. Vertical lines denote 95 percent confidence intervals (s.e. clustered at the city level). Panel B plots the difference in the estimated coefficients in Panel A. *Large* and *small* refer to the sub-samples of counties with large or small percentage of government spending classified as "irregular."

Figure A.10: Multinetworks non-linear estimates



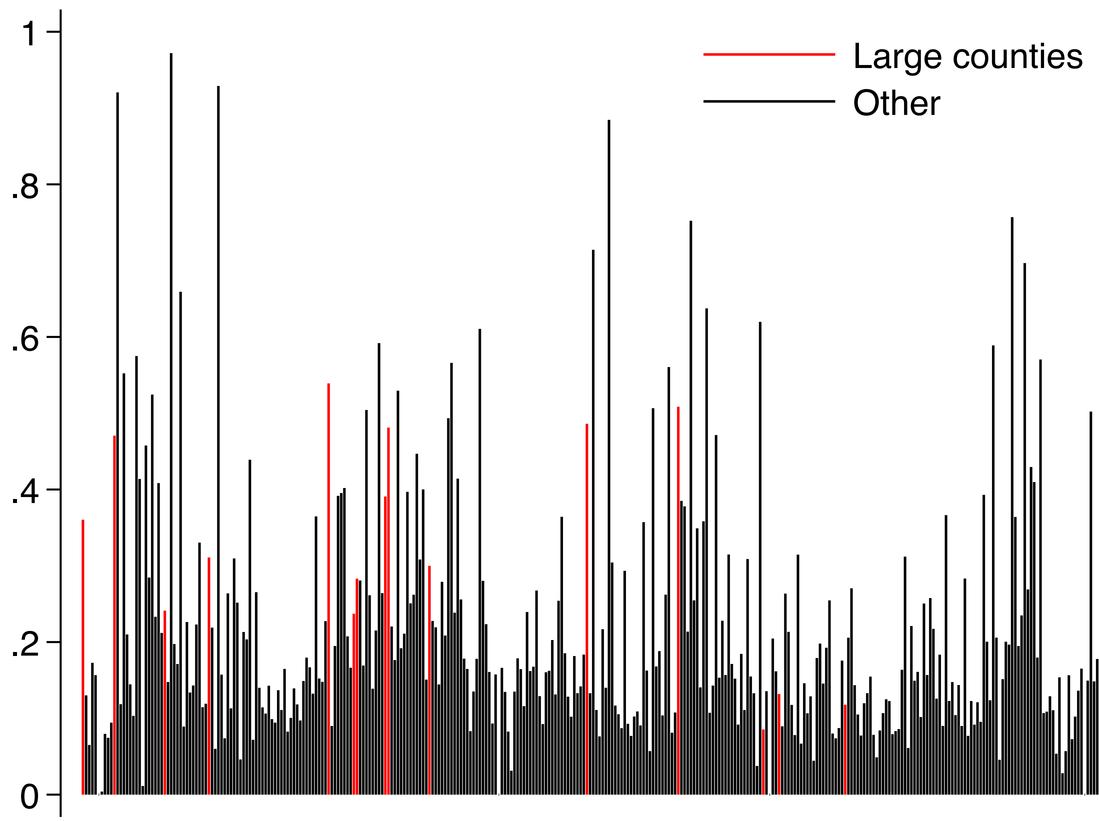
(a) Coefficients for spatial network absenteeism



(b) Coefficients for social network absenteeism

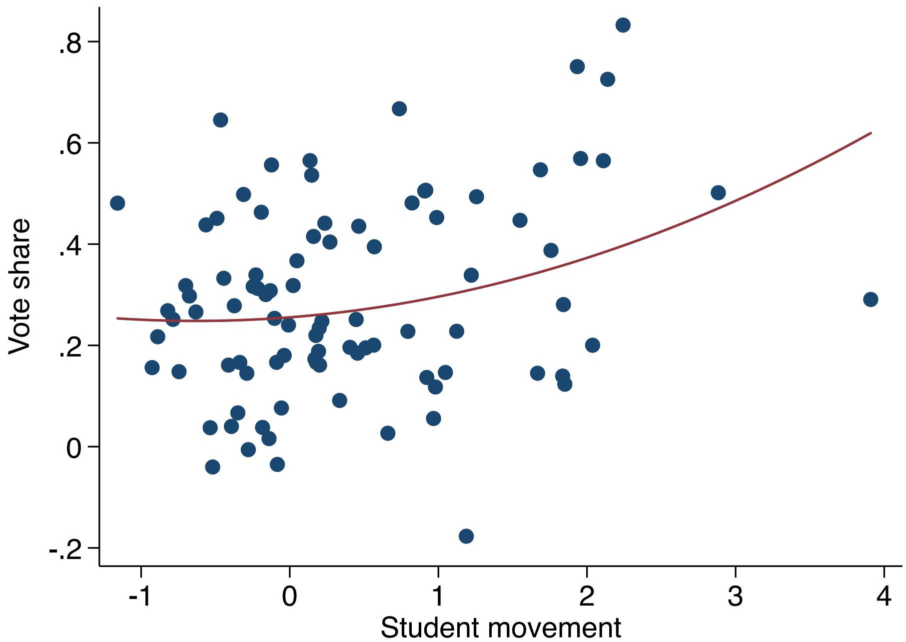
Notes: This figure plots 2SLS estimates from a regression of school absenteeism on 10 indicators of social and spatial network absenteeism in June 16, controlling for school absenteeism before June 16, school characteristics, and city fixed effects. Vertical lines denote 95 percent confidence intervals with standard errors clustered at the city level.

Figure A.11: Participation by county

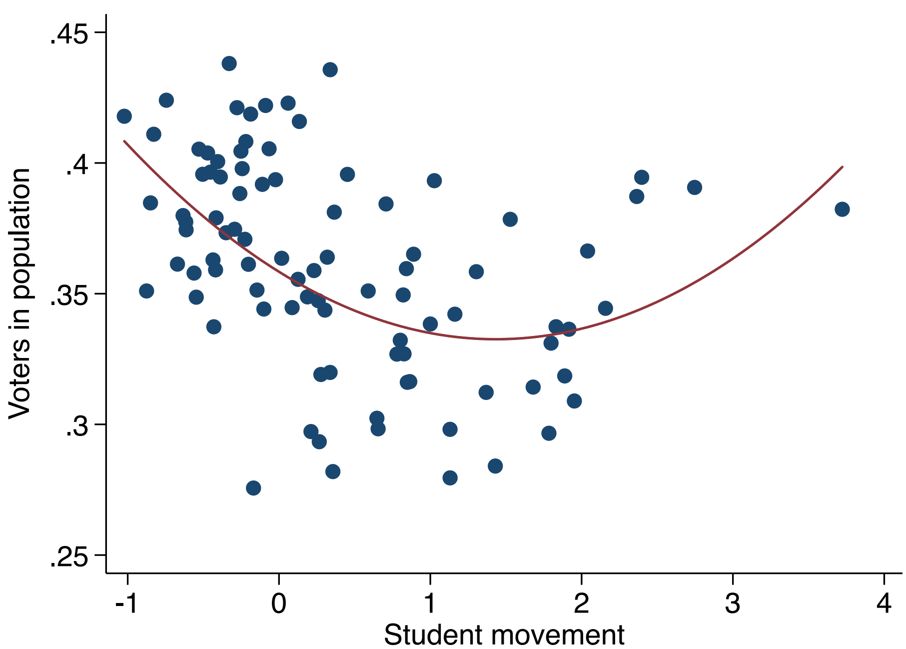


Notes: Own construction based on administrative data. Counties are ordered from north to south in the *x*-axis. There are 324 (out of 346) counties with positive participation in the student movement. Large counties are defined as the fifteen counties with the largest number of students.

Figure A.12: The political effects of the student movement



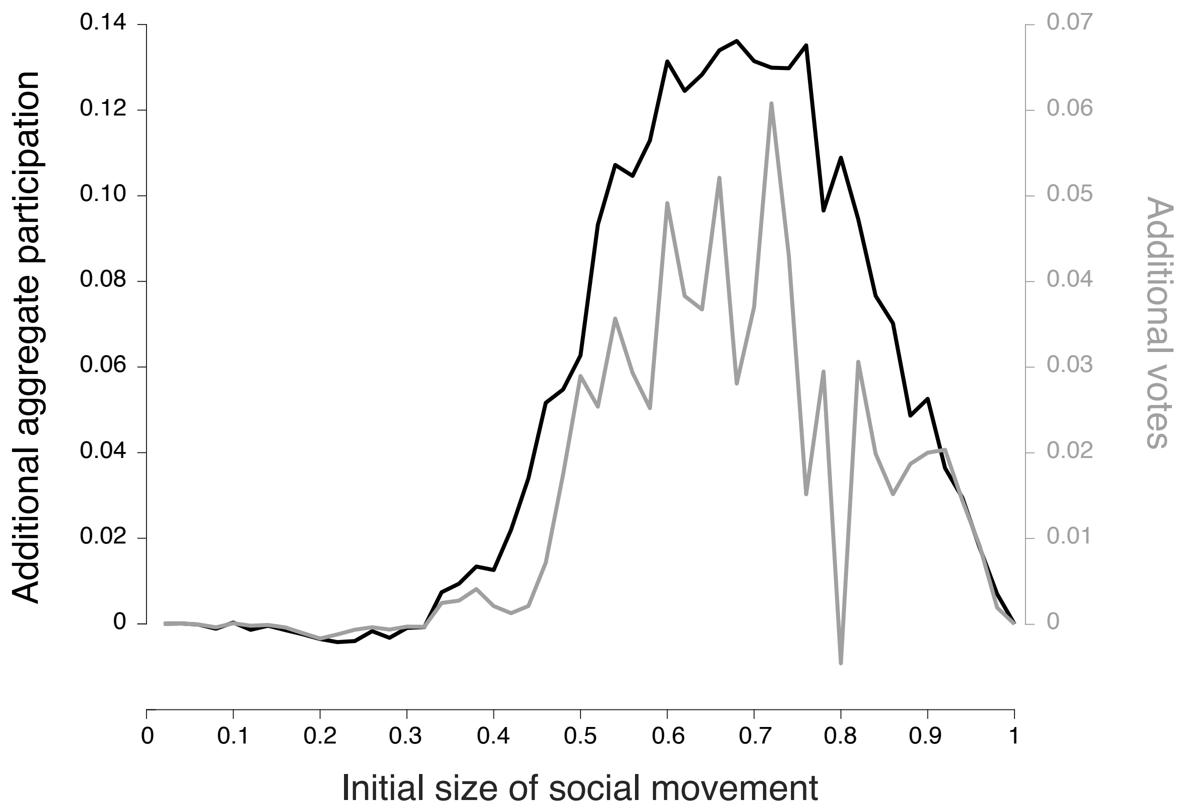
(a) Vote share for non-traditional candidates



(b) Voters in population

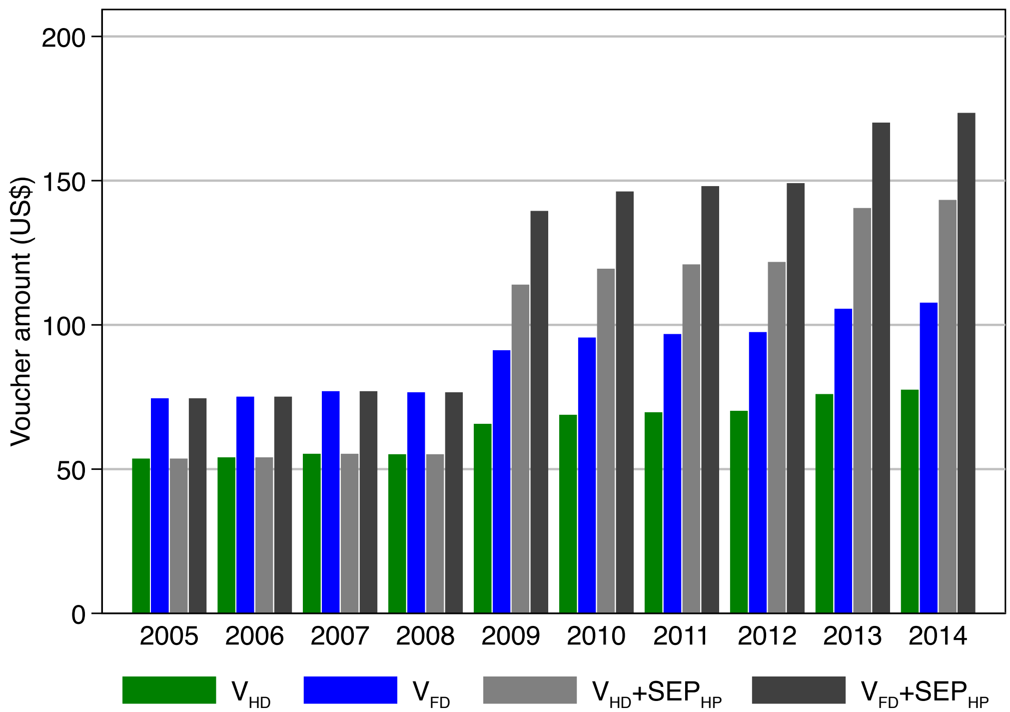
Notes: Quadratic fit of electoral outcomes in the 2012 elections on a measure of the intensity of the student movement in 2011. There are 345 counties in the country.

Figure A.13: Aggregate effects of networks



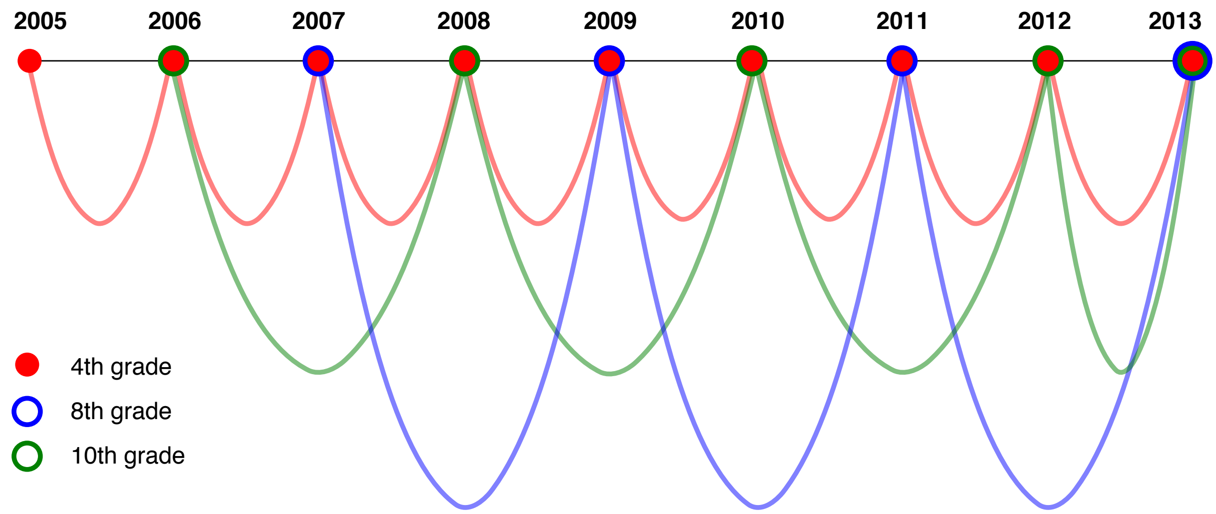
Notes: Simulation results of random initial protesters in the network of students. I proceed in four steps. In the first step, I choose the size of initial protesters, which varies in each simulation. The x-axis measures the percentage of the population of students that is initially protesting ("initial size of social movement"). In the second step, I calculate the choice probabilities of skipping school for all students in the country using the estimated parameters for the effect of networks (non-linear results). In the third step, I calculate how many *additional* students skip school because of network effects (y-axis in the left-hand side). In the fourth step, I take aggregate participation in the movement – measured as the percentage of students skipping school – and use the estimated coefficients to calculate additional votes for non-traditional parties (in percentage points in the right-hand side y-axis).

Figure B.1: Evolution of vouchers



Notes: Amount covered by different types of vouchers in the system. In particular, four types are displayed, covering the interaction of schools offering half and full school shifts (i.e. HD and FD) according to the JEC program, and school subscribed and not subscribed to the SEP program. This figure displays the voucher amount for SEP school with high performance. Note that this figure do not display all voucher types: the voucher amount for low performing SEP schools and the component of SEP vouchers related to the concentration of SEP students in schools are not reported.

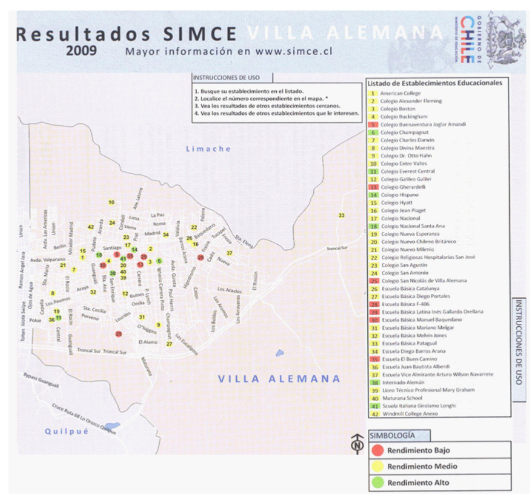
Figure B.2: Timeline of standardized test scores



Notes: Year and grade of students taking the national standardized test (SIMCE) in the period 2005–2013. Math and language tests are always taken by students. Natural and social sciences tests are taken by subsets of students. Additional tests have been applied to 2*nd* and 6*th* grade students since 2012, but we omit them from our analysis because they are relatively new.

Figure B.3: Test scores as quality signals

(a) Public dissemination of test scores

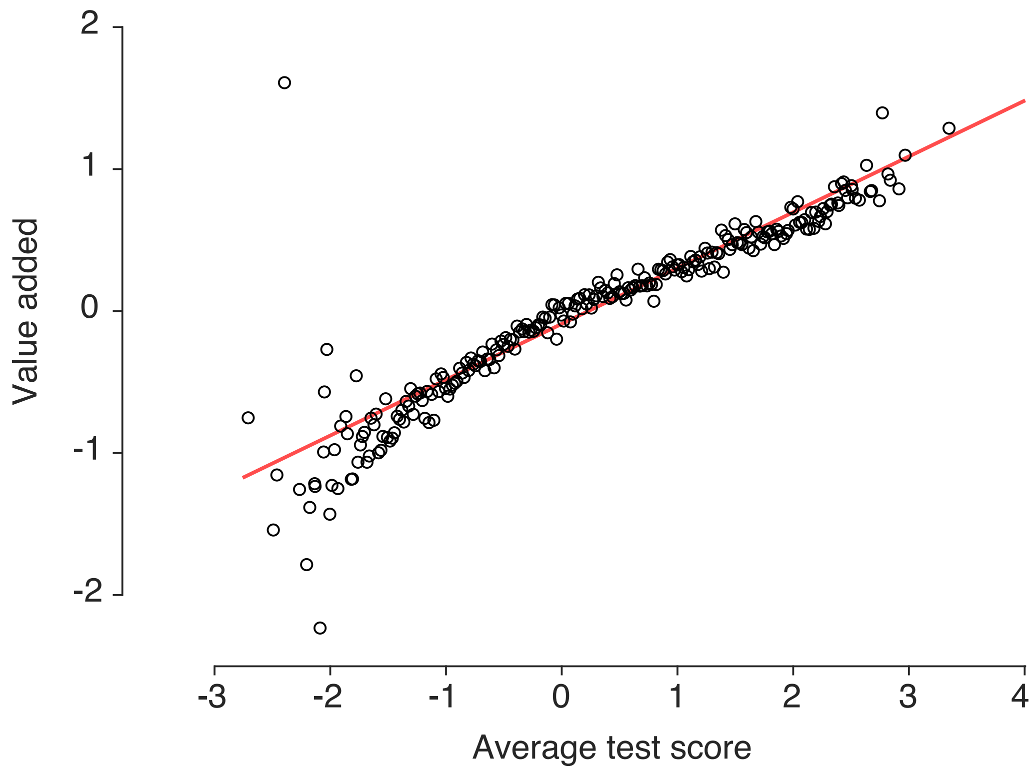


(b) Test scores as an advertising device

(c) Test scores as a policy tool

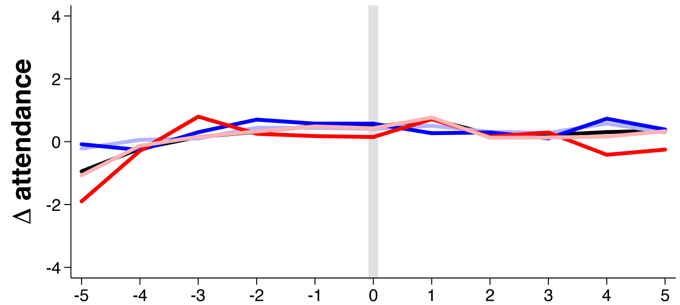
Notes: This figure displays different features of the role test scores in the Chilean educational system. Panel (a) displays the front page of *La Segunda*, a popular newspaper, advertising the disclosure of school level test scores for all schools. Panel (b) displays an advertising banner placed on the front of a school reporting on successful results obtained by the schools in SIMCE as a means of advertising its quality to households. Panel (c) displays an example of of the Educational Traffic Lights policy, which utilizes SIMCE test scores as an input for quality disclosure.

Figure B.4: Test scores as quality signals

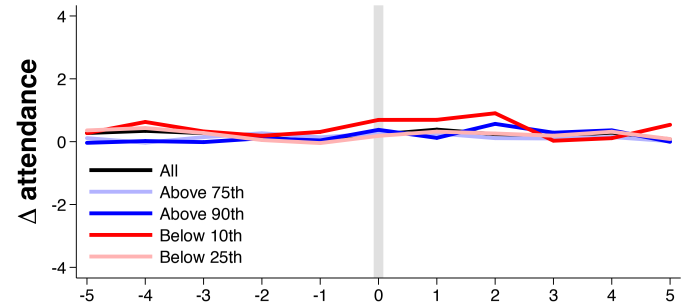


Notes: This figure displays the relationship between test scores and the only available measure of value added in Chile, from Neilson (2013). We thank the author for providing us with this figure.

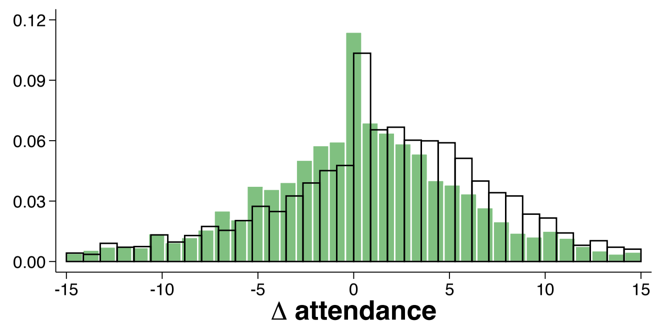
Figure B.5: Comparison of school absenteeism on test day with two other events



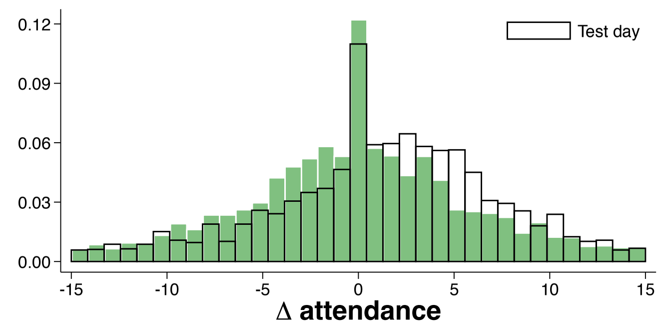
(a) One week after test day



(b) Soccer match day (Chile vs. England)



(c) One week after test day



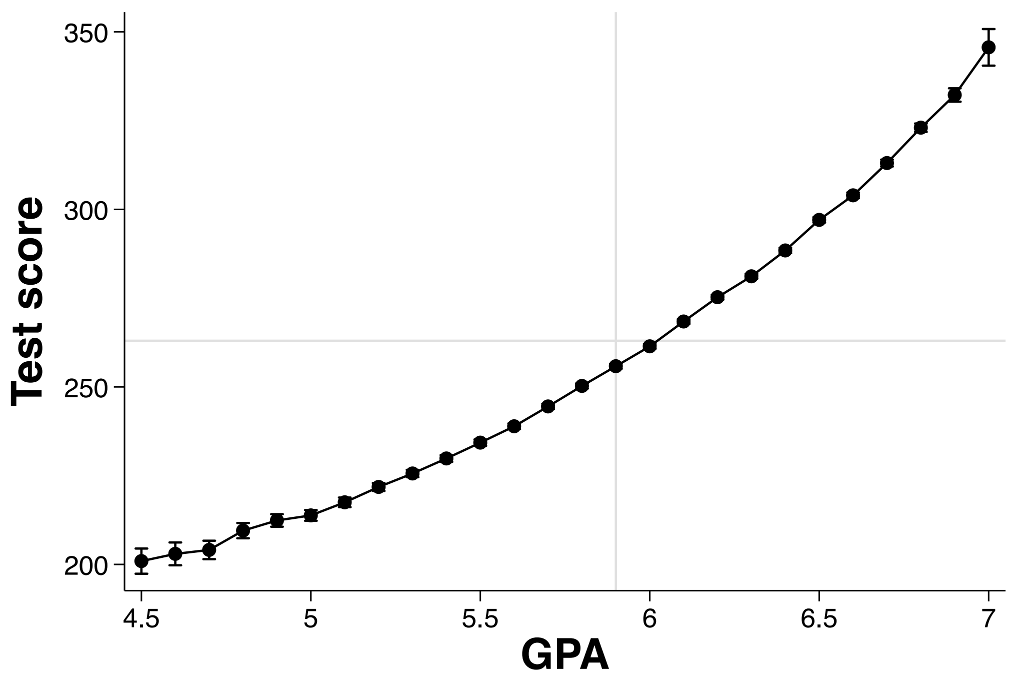
(d) Soccer match day (Chile vs. England)

Notes: Panels (a) and (b) present the difference in absenteeism rates between $4th$ and $3rd$ graders across the GPA distribution around two events: (a) one week after test day, and (b) a soccer match day. The histograms in panels (c) and (d) presents the distribution of the following differences-in-differences estimate at the school level:

$$\Delta\overline{A}_j = (\overline{A}_{j4T} - \overline{A}_{j4t}) - (\overline{A}_{j3T} - \overline{A}_{j3t})$$

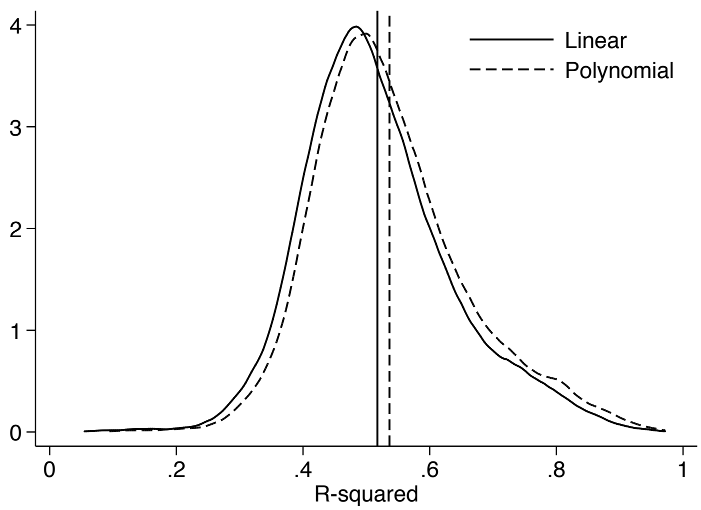
where $\overline{A}_{jkt}$ is the average absenteeism rate of kth graders in school j in day t. Day $t = T$ represents the day of the event analyzed. A Kolmogorov-Smirnov test rejects the equality of distributions in both cases (p-values < 0.01).

Figure B.6: Predictability of test scores

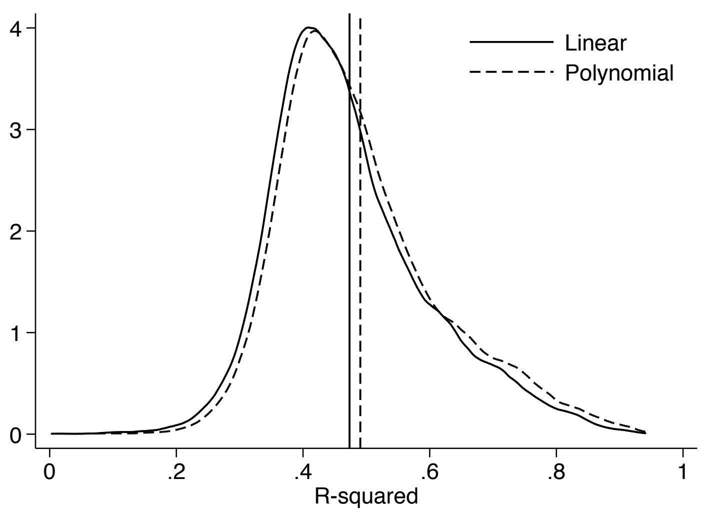


Notes: Coefficient estimates and 95 percent confidence interval of a linear regression of test score on (1) a full set of indicators for a student's GPA, and (2) school fixed effects. Standard errors are clustered at the school level. Gray lines indicate the mean.

Figure B.7: Prediction model



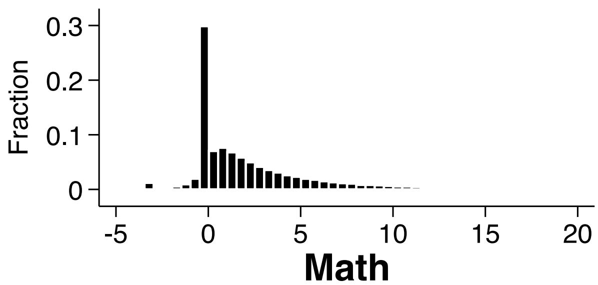
(a) Math



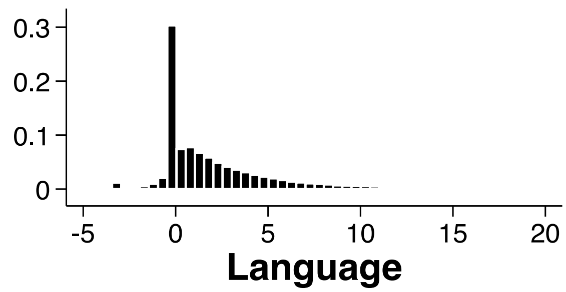
(b) Language

Notes: This figure presents the distribution of R-squared for all regressions of test scores on observable variables (i.e. predictors) among test takers in each school in our data. We include predictors linearly (solid line) or as a polynomial (dash line). Recall that these predictions include GPA, indicators for school switchers and students who are repeating the grade, and year fixed effects. Vertical lines denote the average R-square in the corresponding panel. Panel (a) plots the R-squared for the mathematics test and panel (b) plots the R-squared for the language test. There are a total of 7,493 regressions in each panel.

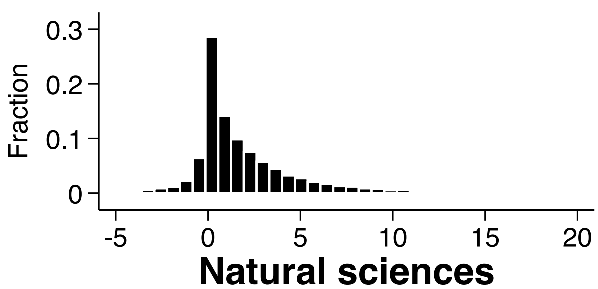
Figure B.8: Distribution of distortions by subject in 4*th* grade



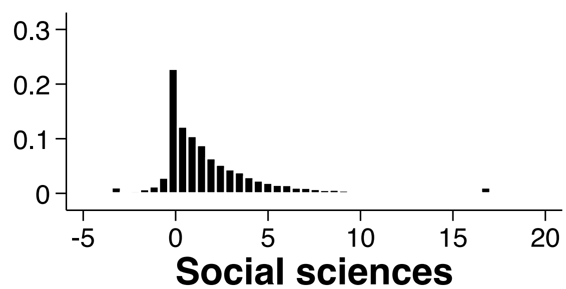
(a) 4*th* grade



(b) 4*th* grade



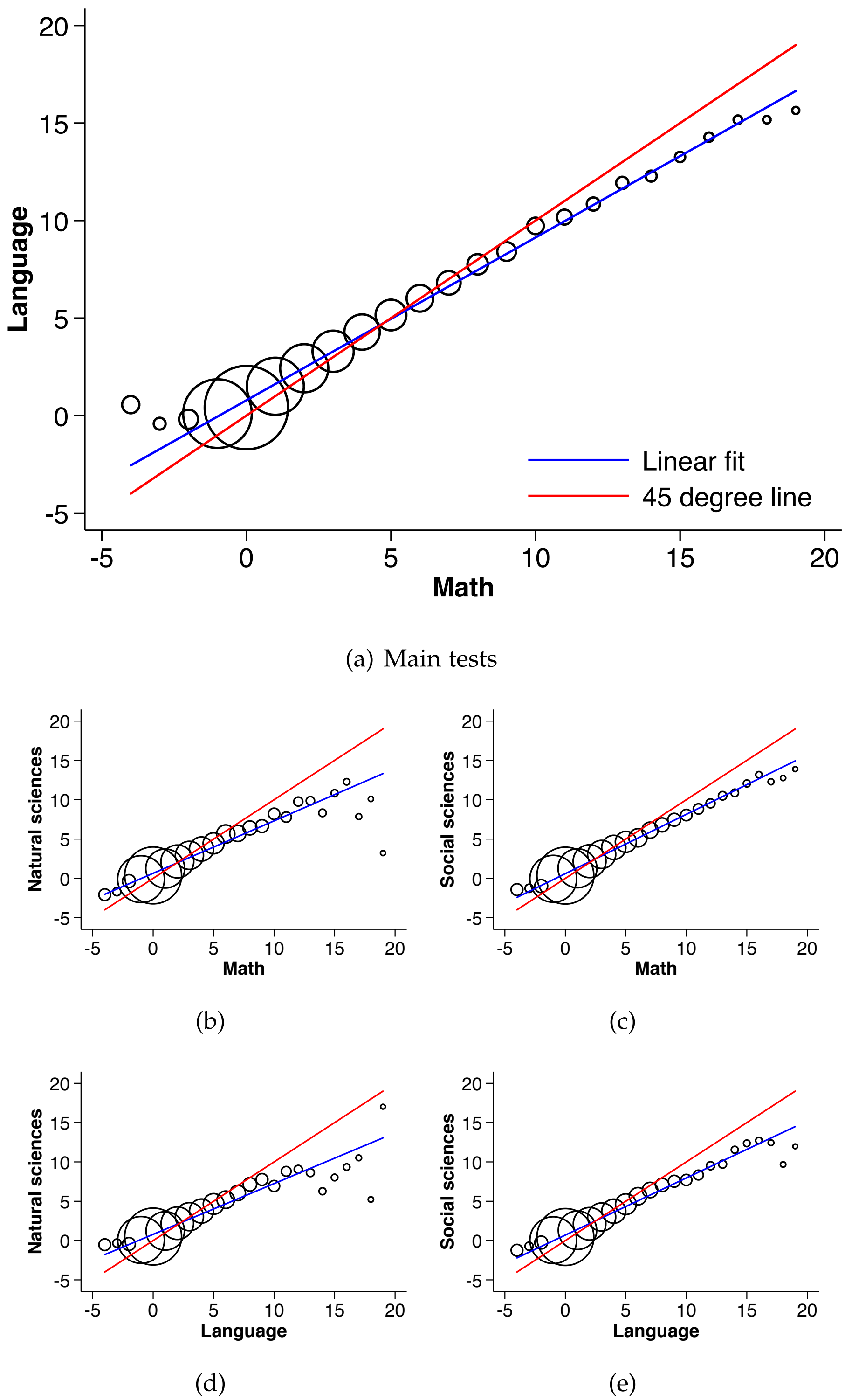
(c) 4*th* grade



(d) 4*th* grade

Notes: We estimate distortions by subject of SIMCE using the methodology described in section 4 of the paper. Distortions in quality signals correspond to the average distortion in mathematics and language. We provide descriptive statistics for distortions by subject in Table B.1.

Figure B.9: Correlation between distortions in different tests



(a) Main tests

(b)

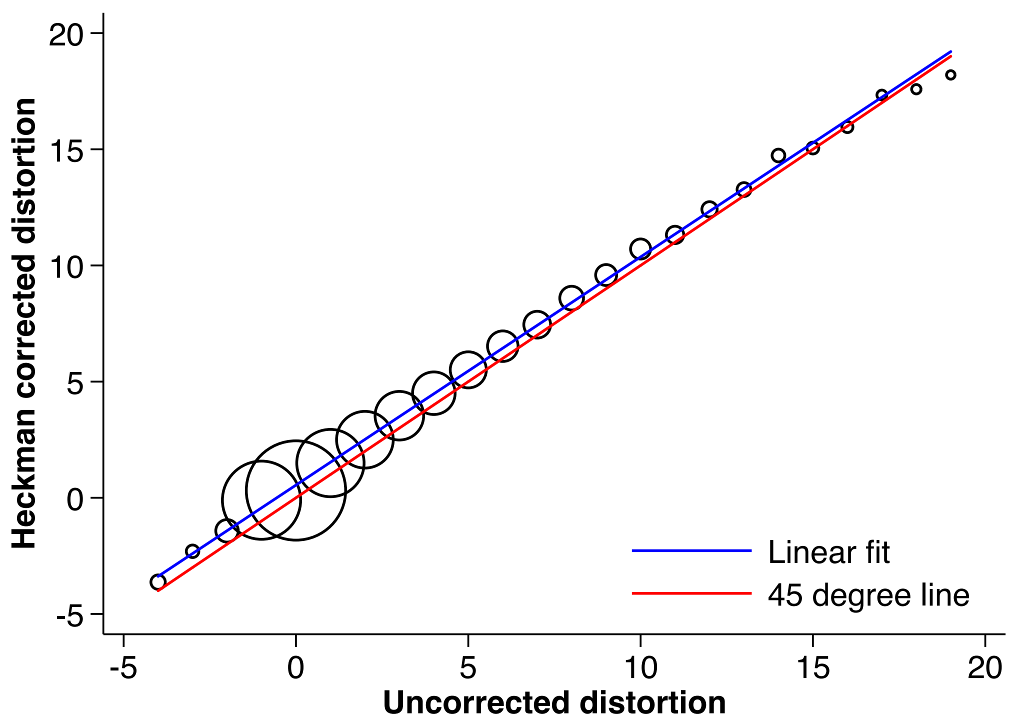
(c)

(d)

(e)

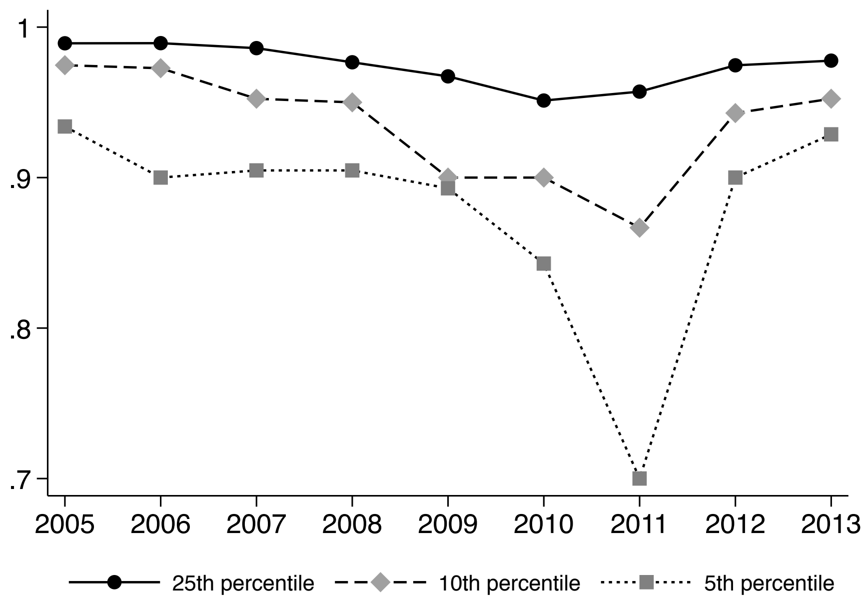
Notes: These figures displays the relationships between estimated distortions in test scores for different subjects of SIMCE.

Figure B.10: Heckman corrected distortions

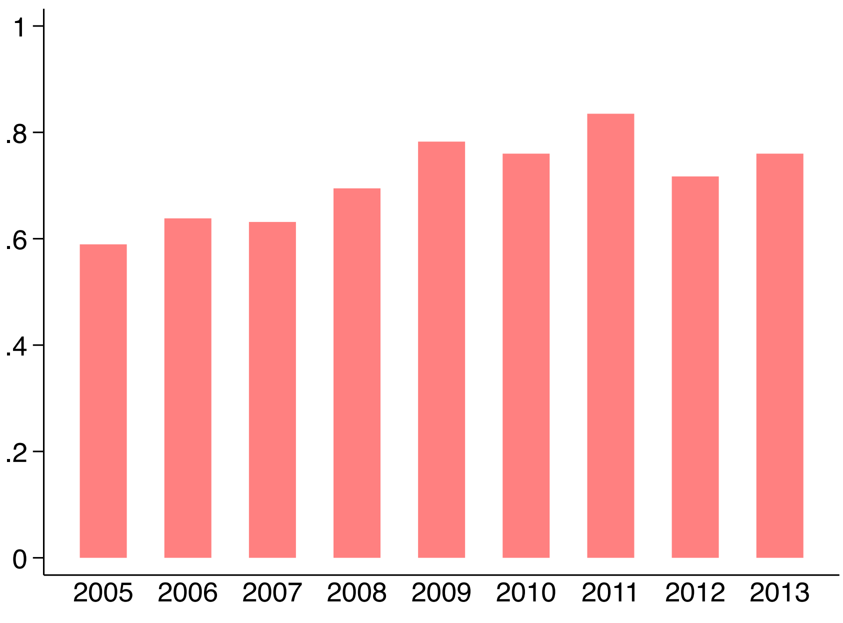


Notes: The excluded variable when calculating the Heckman corrected distortions is an indicator variable that takes the value of one for students that live outside of the municipality of the school.

Figure B.11: Distribution of rank correlations over time



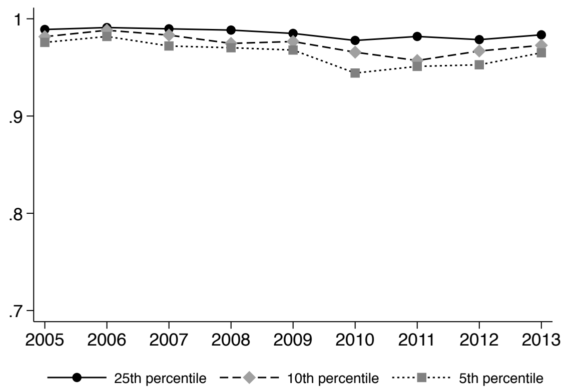
(a) Percentiles in rank correlation distribution $f(\rho_{mt})$



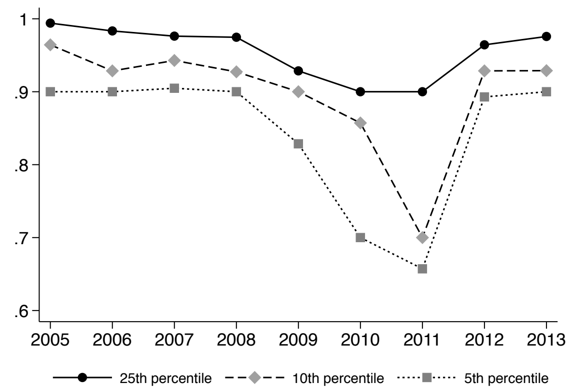
(b) Percentage of markets with changes in ranking (i.e., $\rho_{mt} < 1$)

Notes: Let ρ_{mt} be the rank correlation of distorted and undistorted quality in market m and year t. We observe approximately 210 markets every year.

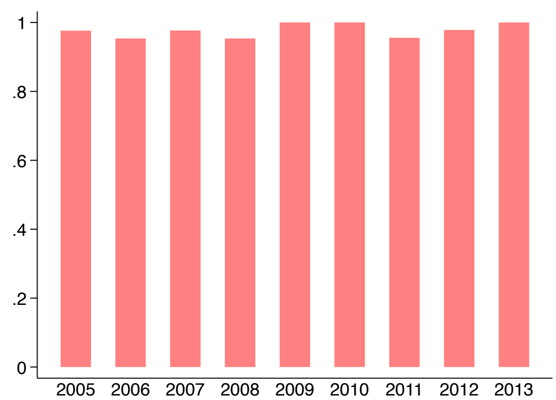
Figure B.12: Distribution of rank correlations by market type



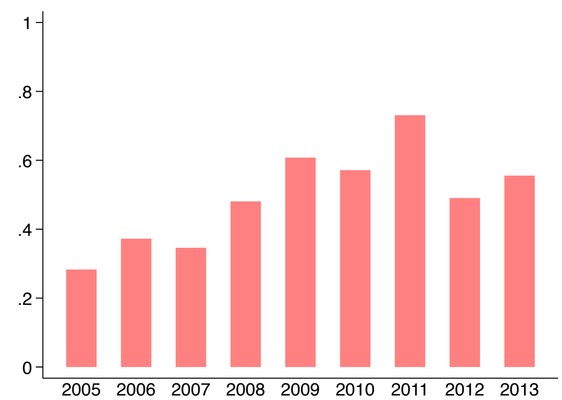
(a) Percentiles in large markets



(b) Percentiles in small markets



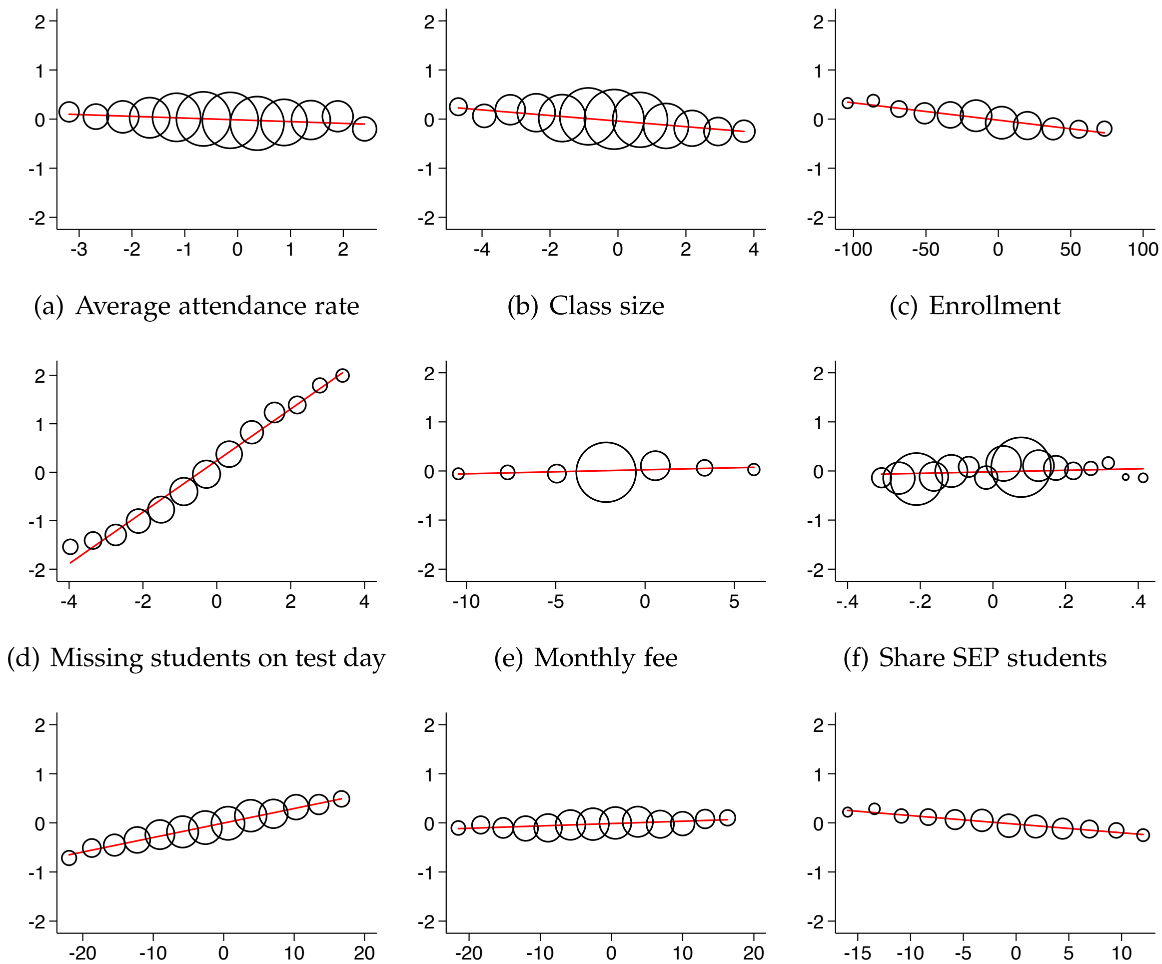
(c) Changes in ranking in large markets



(d) Changes in ranking in small markets

Notes: Let ρ_{mt} be the rank correlation of distorted and undistorted quality in market m and year t. We observe approximately 210 markets every year. "Percentiles in small/large markets" plot the percentiles in the rank correlation distribution $f(\rho_{mt})$ in market m and year t. "Changes in ranking in small/large markets" plot the percentage of markets with changes in ranking, i.e., $\rho_{mt} < 1$. Large (small) markets are defined as market-year observations with more (less) than 10 schools, the median number of schools.

Figure B.13: Distortions and school attributes



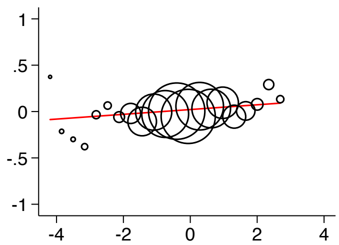
(a) Average attendance rate (b) Class size (c) Enrollment

(d) Missing students on test day (e) Monthly fee (f) Share SEP students

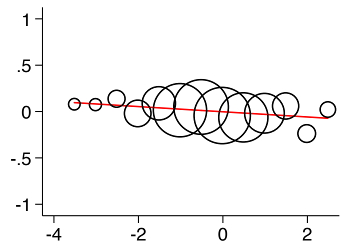
(g) Distorted test score (h) Undistorted test score (i) Students in *4th* grade

Notes: These figures display the relationship between relevant school characteristics and distortions in quality signals. All variables have been residualized with school and year fixed effects. The size of markers indicates the number of students in it. The mean of distortion (*y*-axis) is 2.7 test score points.

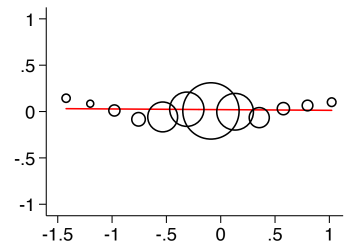
Figure B.14: Distortions and attributes of schools within 3km



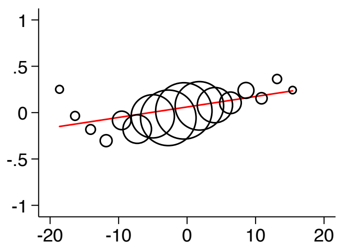
(a) Public schools $-i$



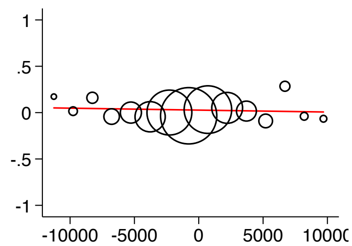
(b) Private-voucher $-i$



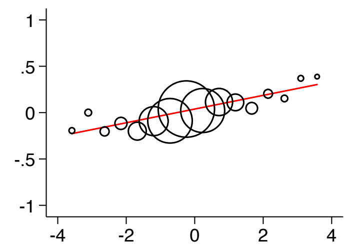
(c) Private $-i$



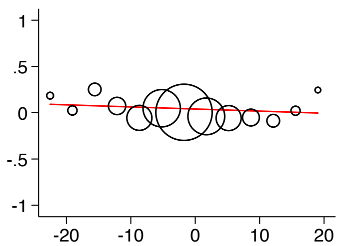
(d) Average test score $-i$



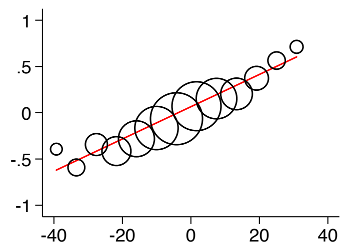
(e) Average fees $-i$



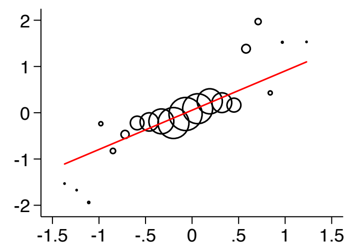
(f) Average distortion $-i$



(g) Perc. fees i



(h) Perc. test score i



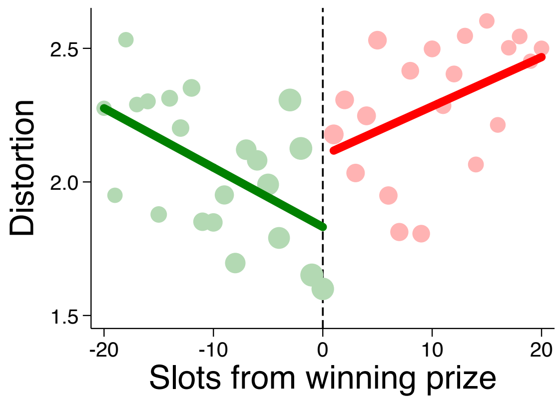
(i) Quality demand elasticity i

Notes: These figures display the relationship between relevant market characteristics and distortions in quality signals. All variables have been residualized with school and year fixed effects. The size of markers indicates the number of students in it. The mean of distortion (y-axis) is 2.7 test score points. Variables in panels (a) through (f) correspond to market aggregates excluding the reference school. Quality demand elasticities in panel (i) are calculated using the sample and estimates from the school choice model in section 5, as:

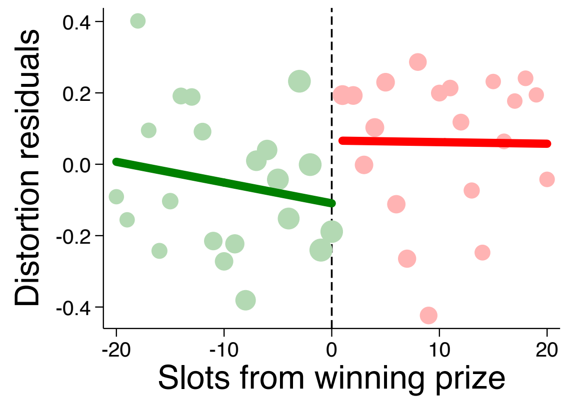
$$\eta^q_{jmt} = \frac{\partial s_{jmt}}{\partial q_{jmt}} \frac{q_{jmt}}{s_{jmt}} = \left(\sum_r \pi^r_{mt} \frac{1}{N^r_{mt}} \sum_{i \in \mathcal{I}^r_{mt}} \frac{\partial P^r_{ijmt}(d^r, \hat{\delta}^r, \hat{\beta}^r_d)}{\partial q_{jmt}} \right) \frac{q_{jmt}}{s_{jmt}}$$

where π^r_{mt} is the share of households of type r in market m and year t, while N^r_{mt} and $\mathcal{I}^r_{mt}$ are the number and the set of such households respectively. The expression in brackets is thus simply a type-share-weighted average of the partial derivative of choice probabilities for school j with respect to quality. In the plot, both variables are residualized by removing school and year fixed effects.

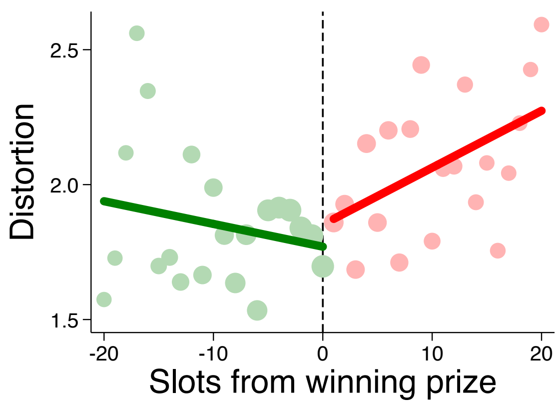
Figure B.15: Monetary incentives for teachers



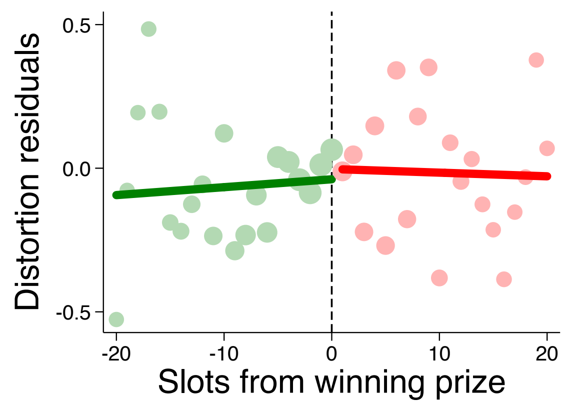
(a) Distortion in year $t+1$



(b) Distortion in year $t+1$



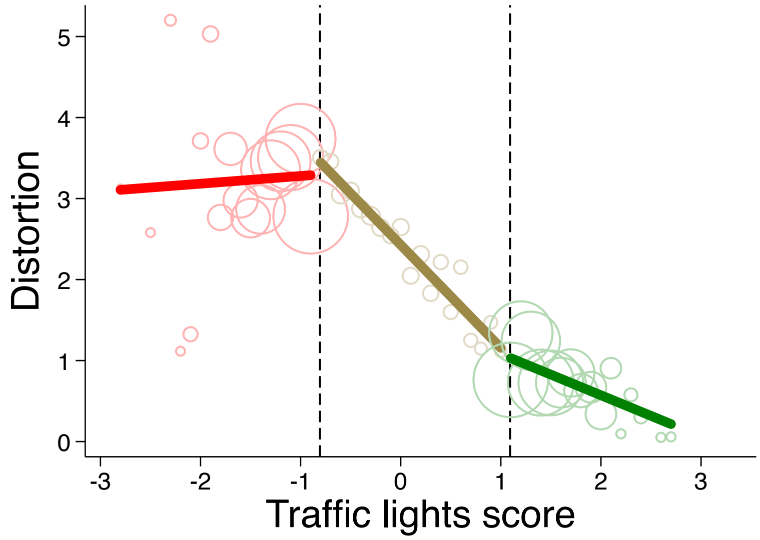
(c) Distortion in year $t+2$



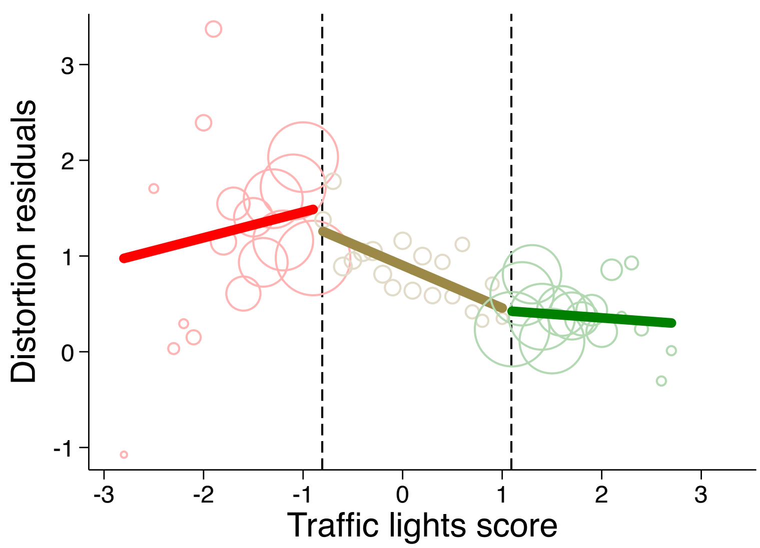
(d) Distortion in year $t+2$

Notes: Regression kink design to test for the effect of monetary teacher incentives on distortions in quality signals (y-axis). The x-axis represents a measure of the probability of winnning the prize (i.e. teacher bonuses). Schools to the left (right) of the thresholds won (did not win) the prize in the previous tournament. We present more details about this public program in section 2 of the paper. Left panels correspond to changes in the slope without controls while right panels control for a set of school fixed effects. The null hypothesis of incentives affecting distortions implies an "inverted V" relationship between "slots from winning prize" and distortions around the kink. We strongly reject the hypothesis that teacher incentives cause distortions in quality signals.

Figure B.16: "Educational Traffic lights" policy



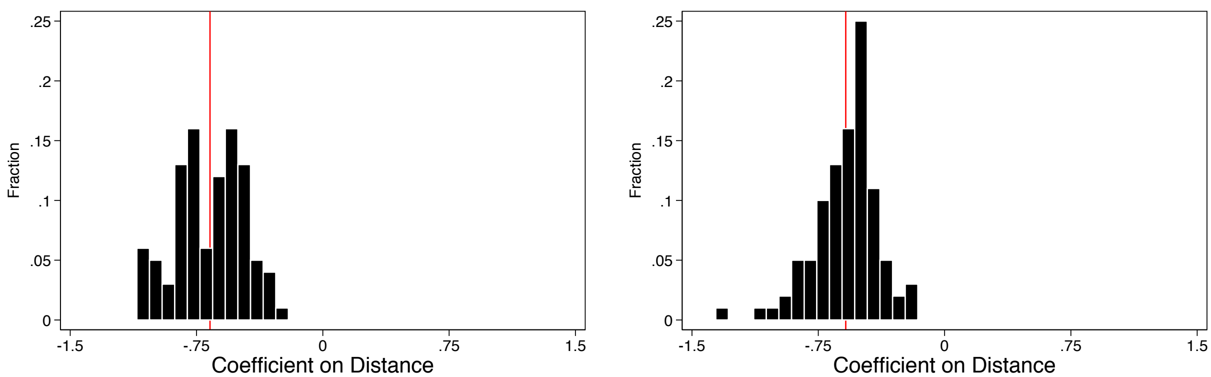
(a) Distortion in year 2010



(b) Distortion in year 2010

Notes: Regression kink design to test for the hypothesis of manipulation of test scores to be classified in a "higher" category. The *x*-axis represents school scores which fully determines their category. We present more details about the policy in section 2 of the paper. The null hypothesis of manipulation implies an "inverted V" relationship between school scores and distortions in quality signals. The upper panel corresponds to the test without controls while the lower panel controls for a basic set of pre-determined school characteristics. We strongly reject the hypothesis of manipulation of test scores for the school to be classified in a higher category.

Figure B.17: Estimated coefficients on distance from the first stage



(a) Poor students, distance

(b) Non-poor students, distance

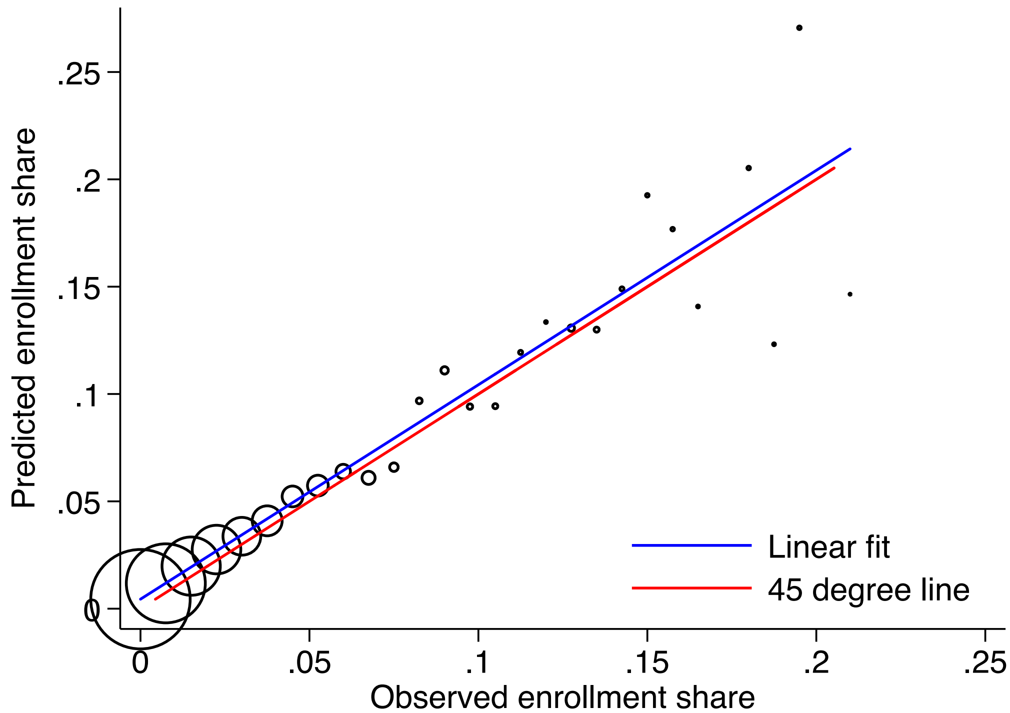
Notes: These figures display resulting estimates for β_d^r from the first stage of the school choice model. Each observation is the estimated coefficient for an estimating cell comprised by a market, year and household type. The red line indicates the average coefficient.

Table B.1: Descriptive statistics for distortions by subject

	Obs.	Mean	St. Dev.	Min	Max	Years
Mathematics	60,741	2.7	4.4	-3.5	23.9	2005–2013
Language	60,760	2.6	4.4	-3.4	23.8	2005–2013
Natural sciences	5,902	2.1	3.9	-7.8	20.6	2008, 2010
Social sciences	10,033	2.1	3.2	-3.5	17.0	2009

Notes: Distortions are measured in test score points and we estimated them using the methodology described in section 4. See Figure B.2 for a timeline of standardized tests.

Figure B.18: Observed and predicted school enrollment

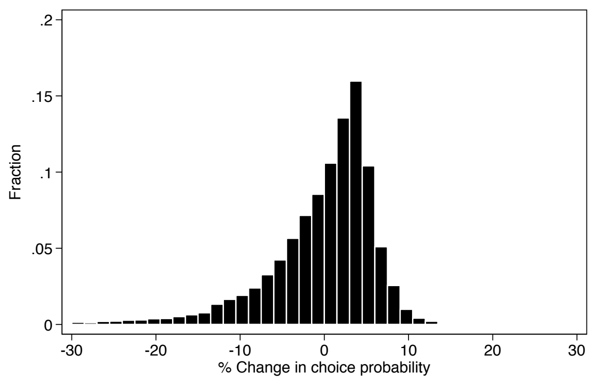


Notes: This figure displays the relationship between observed school enrollment shares and predicted school enrollment shares using model estimates. Predicted enrollment shares are calculated as:

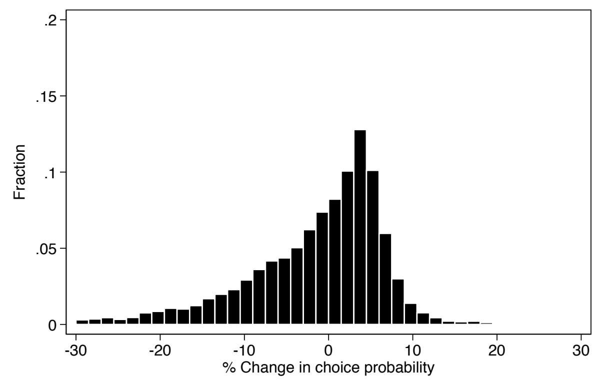
$$s_{jmt}(\hat{\delta}, \hat{\beta}_d) = \sum_r \pi^r_{mt} \frac{1}{N^r_{mt}} \sum_{i \in \mathcal{I}^r_{mt}} P^r_{ijmt}(d^r, \hat{\delta}^r, \hat{\beta}^r_d)$$

where π^r_{mt} is the share of households of type r in market m and year t, while N^r_{mt} and $\mathcal{I}^r_{mt}$ are the number and the set of such households respectively. The expression is thus simply a type-share-weighted average of average choice probabilities for school j.

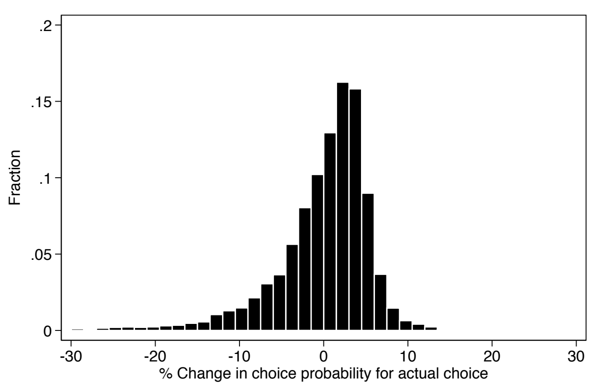
Figure B.19: Changes in choice probabilities



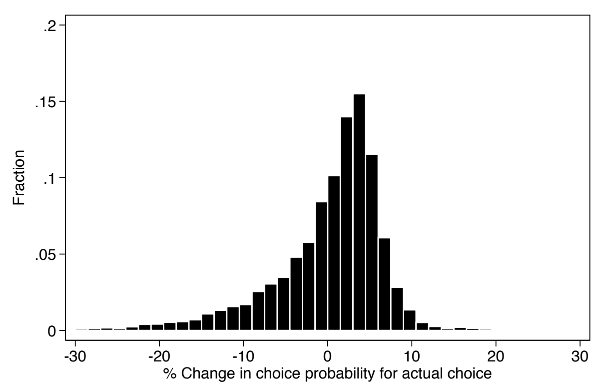
(a) Poor students, all schools



(b) Non-poor students, all schools



(c) Poor students, chosen school



(d) Non-poor students, chosen school

Notes: These figures display change in school choice probabilities between the counterfactual and baseline scenarios we analyze. Each observation is the percentage change in the choice probability of a school by a household in the estimating dataset. Panels (a) and (b) include results for all schools in the dataset, while panels (c) and (d) focus only on schools chosen by household in the baseline scenario.

Table B.2: School markets as connected components

	3km	4km	5km	6km	7km	8km	9km	10km
Markets	451	413	380	348	322	295	273	251
Markets with > 1 schools	262	248	233	219	208	196	191	181
Markets with > 5 schools	106	104	99	93	90	88	86	86
Markets with > 10 schools	63	63	60	55	52	49	48	50
Markets with > 20 schools	36	36	33	31	30	29	28	29

Notes: Let **A** be a $N \times N$ matrix representing the network of N =5,416 urban schools in Chile in the period 2005–2013. In network theory, **A** is referred to as adjacency matrix. This adjacency matrix represents an undirected network, i.e., **A** is a symmetric matrix. The element $A(i, j)$ in this adjacency matrix takes the value of one if school i and j are closer than κ kilometers from each other and zero otherwise. A "component" or "connected component" of **A** is a sub-network in which any two schools are connected to each other through some other school, i.e., we can always find a "path" that connects any two pair of schools in the sub-network. A market is defined as a connected component of **A**. In the paper, we use $\kappa = 5$ (highlighted in gray), but results are robust to different definitions.

Table B.3: IV results from the second stage of school choice model - First stage for school fees

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	All			Poor students			Non-poor students		
Rivals' characteristics									
Religious	-19.876**	79.768*	100.580	-11.500	68.481	11.209	-26.668**	85.263	195.296*
	(7.810)	(46.975)	(70.650)	(10.365)	(57.551)	(79.821)	(11.316)	(71.460)	(108.364)
Gender constraint	-80.777***	463.861***	780.010***	-90.315***	513.487***	617.324**	-62.196**	418.205**	922.088***
	(19.195)	(121.244)	(222.163)	(24.978)	(149.890)	(257.538)	(28.246)	(182.350)	(335.240)
Public	-17.294***	233.888***	-83.801	-9.406	-114.488**	-75.163	-21.657**	614.363***	-68.761
	(6.075)	(45.861)	(61.668)	(7.921)	(48.773)	(70.400)	(9.017)	(74.814)	(94.836)
Teacher wages									
Average hourly wage	0.094***	0.123***	0.135***	0.060	0.078*	0.089**	0.131***	0.170***	0.184***
	(0.031)	(0.032)	(0.030)	(0.042)	(0.043)	(0.041)	(0.045)	(0.047)	(0.045)
Voucher									
Baseline	-1.335***	-1.473***	-1.454***	-0.975***	-1.177***	-1.166***	-1.535***	-1.627***	-1.603***
	(0.031)	(0.030)	(0.029)	(0.059)	(0.055)	(0.054)	(0.037)	(0.036)	(0.035)
SEP school	0.062*	-0.080**	0.007	-0.093***	-0.217***	-0.077**	0.210***	0.056	0.090*
	(0.032)	(0.033)	(0.032)	(0.034)	(0.035)	(0.033)	(0.052)	(0.052)	(0.049)
SEP share	-10.034***	-9.616***	-1.556***	-9.927***	-9.825***	-0.856	-10.113***	-9.368***	-2.342***
	(0.185)	(0.187)	(0.546)	(0.235)	(0.241)	(0.636)	(0.279)	(0.279)	(0.815)
Temperature									
Linear	4.167***	2.684	0.621	3.980**	3.871	0.808	4.284**	0.709	-0.311
	(1.308)	(4.491)	(4.381)	(1.778)	(6.120)	(5.784)	(1.874)	(6.490)	(6.466)
Quadratic	-0.261***	-0.163	-0.095	-0.256***	-0.214	-0.110	-0.260***	-0.088	-0.058
	(0.054)	(0.179)	(0.174)	(0.074)	(0.243)	(0.230)	(0.078)	(0.259)	(0.258)
SNED program									
Prize residual	-5.773***	-4.227***	-5.363***	-3.588**	-2.150	-3.406**	-7.605***	-5.997***	-7.023***
	(1.211)	(1.159)	(1.143)	(1.560)	(1.499)	(1.472)	(1.816)	(1.727)	(1.698)
Controls	No	No	Yes	No	No	Yes	No	No	Yes
Market-Year F.E.	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Observations	10,774	10,774	10,774	5,335	5,335	5,335	5,439	5,439	5,439
R-squared	0.685	0.729	0.739	0.665	0.716	0.730	0.706	0.748	0.758

Notes: All regressions are weighted by school enrollment. Columns 3, 6 and 9 include other school attributes in the corresponding second stage specifications, namely indicators for schools being religious, public, gender constrained or part of the SEP program. Results not reported in this table. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B.4: IV results from the second stage of school choice model - First stage for school quality

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	All			Poor students			Non-poor students		
Rivals' characteristics									
Religious	1.173	12.207	62.650**	2.547	9.472	54.491	-0.046	14.997	71.121*
	(3.687)	(20.849)	(29.730)	(5.269)	(29.957)	(42.769)	(5.157)	(29.337)	(41.727)
Gender constraint	-42.162***	-26.845	375.535***	-45.462***	-27.096	372.044***	-38.185***	-26.725	377.581***
	(8.451)	(46.254)	(77.523)	(12.077)	(65.914)	(110.647)	(11.852)	(65.569)	(109.687)
Public	-7.425***	322.369***	-34.034	-6.927*	321.522***	-46.263	-7.653**	324.807***	-20.461
	(2.729)	(22.419)	(30.724)	(3.881)	(32.046)	(43.895)	(3.834)	(31.777)	(43.497)
Teacher wages									
Average hourly wage	0.060***	0.064***	0.072***	0.059***	0.062***	0.071***	0.061***	0.067***	0.074***
	(0.013)	(0.014)	(0.013)	(0.018)	(0.019)	(0.019)	(0.018)	(0.019)	(0.019)
Voucher									
Baseline	-0.212***	-0.247***	-0.234***	-0.183***	-0.231***	-0.219***	-0.227***	-0.254***	-0.241***
	(0.009)	(0.009)	(0.009)	(0.016)	(0.016)	(0.015)	(0.012)	(0.012)	(0.011)
SEP school	0.715***	0.604***	0.625***	0.720***	0.615***	0.636***	0.709***	0.593***	0.614***
	(0.018)	(0.017)	(0.017)	(0.025)	(0.024)	(0.024)	(0.025)	(0.025)	(0.024)
SEP share	-4.663***	-4.149***	-0.157	-4.694***	-4.200***	-0.177	-4.628***	-4.092***	-0.134
	(0.082)	(0.079)	(0.202)	(0.115)	(0.113)	(0.289)	(0.117)	(0.113)	(0.286)
Temperature									
Linear	4.151***	1.492	0.620	4.134***	1.727	0.620	4.166***	1.196	0.554
	(0.494)	(1.630)	(1.489)	(0.712)	(2.339)	(2.115)	(0.686)	(2.298)	(2.122)
Quadratic	-0.208***	-0.091	-0.064	-0.206***	-0.105	-0.067	-0.209***	-0.076	-0.059
	(0.021)	(0.066)	(0.060)	(0.030)	(0.094)	(0.085)	(0.029)	(0.093)	(0.085)
SNED program									
Prize residual	1.713***	3.029***	2.412***	1.861***	3.143***	2.492***	1.598**	2.935***	2.349***
	(0.486)	(0.447)	(0.431)	(0.687)	(0.636)	(0.614)	(0.689)	(0.634)	(0.612)
Controls	No	No	Yes	No	No	Yes	No	No	Yes
Market-Year F.E.	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Observations	10,774	10,774	10,774	5,335	5,335	5,335	5,439	5,439	5,439
R-squared	0.391	0.474	0.512	0.367	0.452	0.492	0.412	0.493	0.529

Notes: All regressions are weighted by school enrollment. Columns 3, 6 and 9 include other school attributes in the corresponding second stage specifications, namely indicators for schools being religious, public, gender constrained or part of the SEP program. Results not reported in this table. Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table B.5: OLS results from the second stage of school choice model

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	All sample			Poor students			Non-poor students		
Fee	-0.004***	-0.004***	-0.003***	-0.009***	-0.010***	-0.009***	-0.002***	-0.001***	-0.003***
	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)
Quality	0.010***	0.011***	0.011***	0.003***	0.005***	0.006***	0.017***	0.017***	0.016***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Religious			-0.040*			-0.074**			0.000
			(0.023)			(0.029)			(0.028)
Gender constraint			0.233***			0.237***			0.222***
			(0.043)			(0.053)			(0.050)
Public			0.039			0.126***			-0.049
			(0.026)			(0.032)			(0.033)
SEP school			0.012			0.028			-0.247***
			(0.042)			(0.054)			(0.045)
Market-Year F.E.	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Observations	11,041	11,041	11,041	5,461	5,461	5,461	5,580	5,580	5,580
R-squared	0.025	0.370	0.373	0.072	0.488	0.494	0.080	0.561	0.567

Notes: All regressions are weighted by school enrollment. Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table B.6: Schools in corrupt municipalities have larger distortions

Dependent variable is distortions (in test score points)

	Years with transfers	Before audits revealed	After audits revealed
	(1)	(2)	(3)
Irregular payments	0.04***	0.05***	0.02
	(0.01)	(0.02)	(0.02)
Government transfers	0.08***	0.09***	0.07***
	(0.01)	(0.02)	(0.02)
Schools	2,345	2,283	2,239
Municipalities	76	76	76
Observations	11,834	7,588	4,246

Notes: All variables have been normalized. All regressions are weighted by the inverse of the size of the confidence interval of distortions to account for estimation of the dependent variable. Audits in 76 randomly chosen municipalities were implemented by the Comptroller General of Chile to disclose irregular payments from government transfers. The time of disclosure of irregular payments was May of 2012. "Years with transfers" correspond to the period 2008–2013. Column 2 restricts attention to years 2008–2012, and column 3 restricts attention to years 2012–2013. Robust standard errors in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

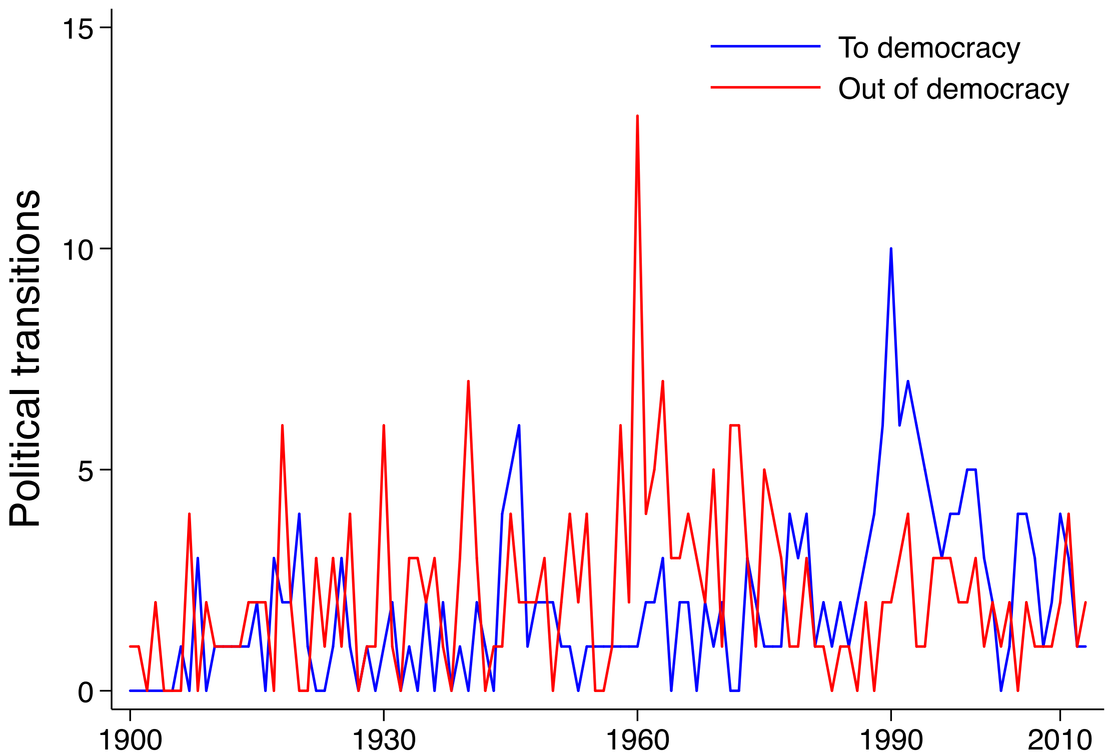
Table B.7: Differences-in-differences of audits

Dependent variable is distortions (in test score points)

	All schools		Schools in audited municipalities	
	(1)	(2)	(3)	(4)
Audit × Post	0.07	0.04		
	(0.05)	(0.03)		
Corrupt × Post			-0.17**	-0.12**
			(0.07)	(0.05)
Post	-0.04*	-0.04**	0.09	0.06*
	(0.02)	(0.02)	(0.06)	(0.03)
Mean of dep. variable	2.9	2.9	3.0	3.0
School-level controls	No	Yes	No	Yes
Municipality F.E.	Yes	Yes	Yes	Yes
Municipalities	344	344	76	76
Schools	7,357	7,357	2,239	2,239
Observations	40,705	37,448	12,865	11,834

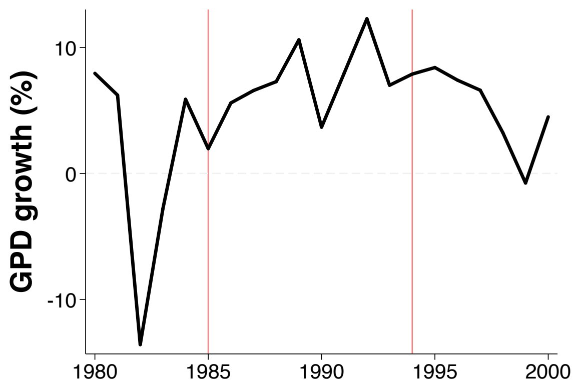
Notes: These regressions restrict attention to the period in which the government transferred monetary resources to be spent under the *Subvención Escolar Preferencial* program (2008–2013). All regressions are weighted by the inverse of the size of the confidence interval of distortions to account for estimation of the dependent variable. Audits in 76 randomly chosen municipalities were implemented by the Comptroller General of Chile to disclose "irregular" expenditures of government transfers. The time of disclosure of irregular payments was May of 2012. The post period are years 2012 and 2013. The "Corrupt" indicator takes the value of one if a municipality has more than 10 percent of the government transfers under "irregular payments." More about irregular payments can be found in CIPER (2012). Standard errors clustered at the municipality level in parentheses. Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Figure C.1: Political transitions in the world (1900–2010)

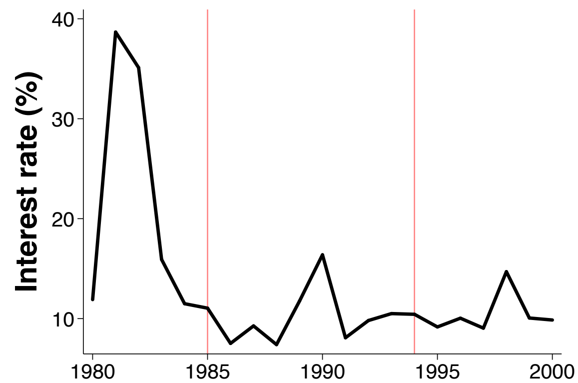


Notes: Data from the Polity IV Project "Political Regime Characteristics and Transitions, 1800–2013." The year of a political transition to democracy is defined as a value of the variable `democ` that is positive in year t and in the set $[-88, -77, -66, 0]$ in year $t-1$. The year of a political transition out of democracy is defined as a value of the variable `democ` that is in the set $[-88, -77, -66, 0]$ in year t and positive in year $t-1$.

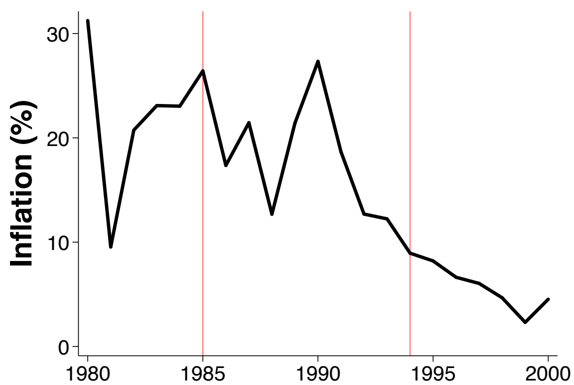
Figure C.2: Macroeconomic indicators (1980–2000)



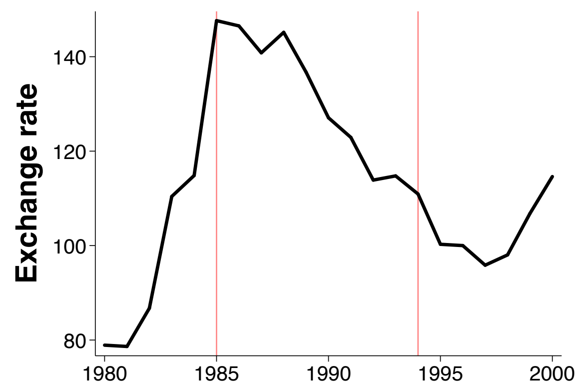
(a)



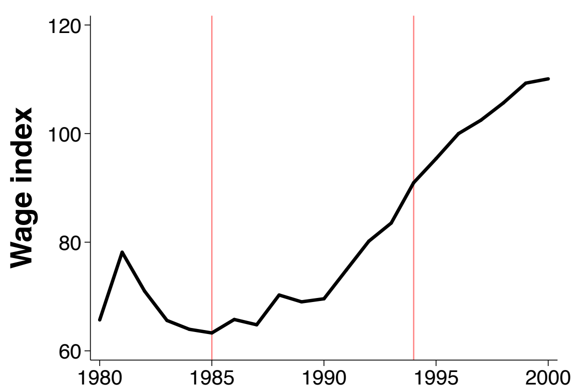
(b)



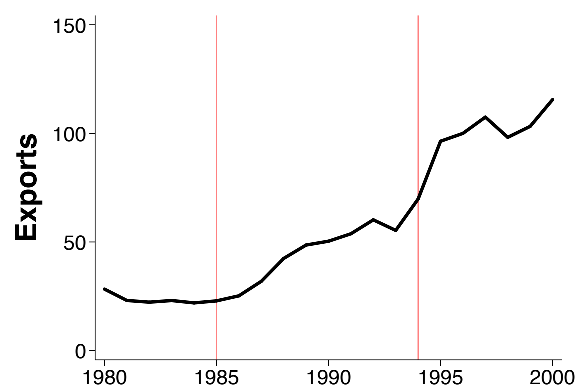
(c)



(d)



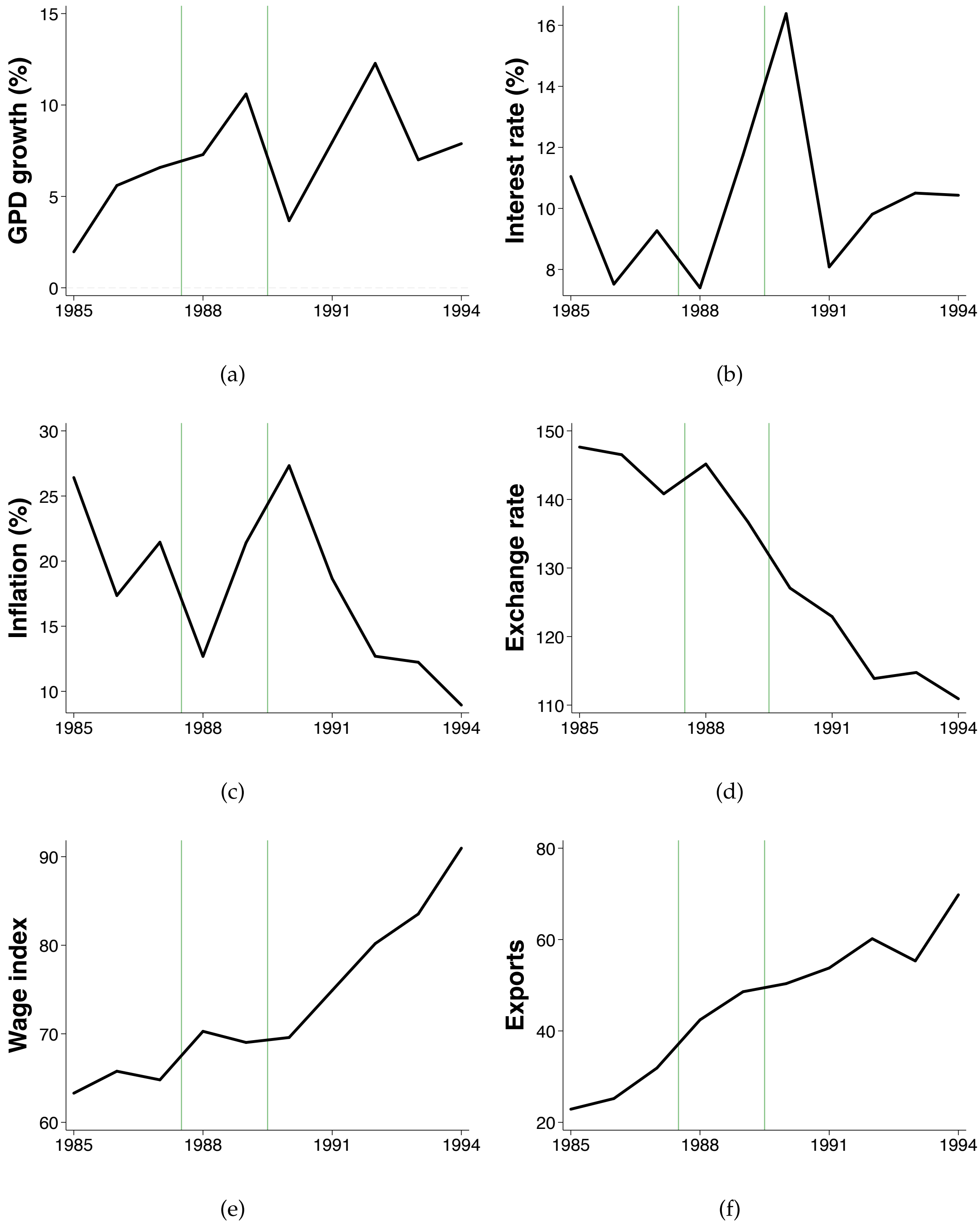
(e)



(f)

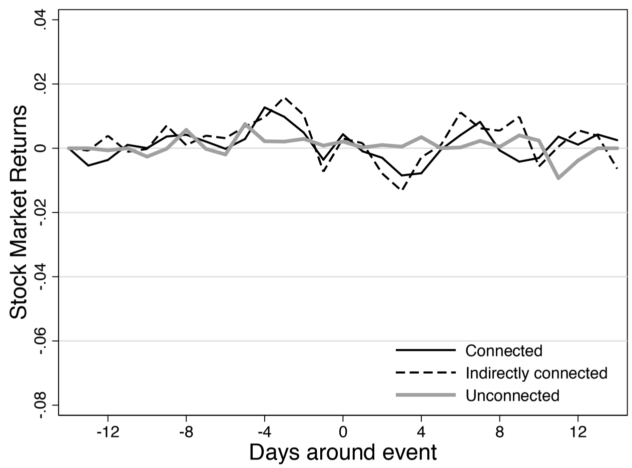
Notes: Red vertical lines represent the first and last year in our analysis.

Figure C.3: Macroeconomic indicators (1985–1994)

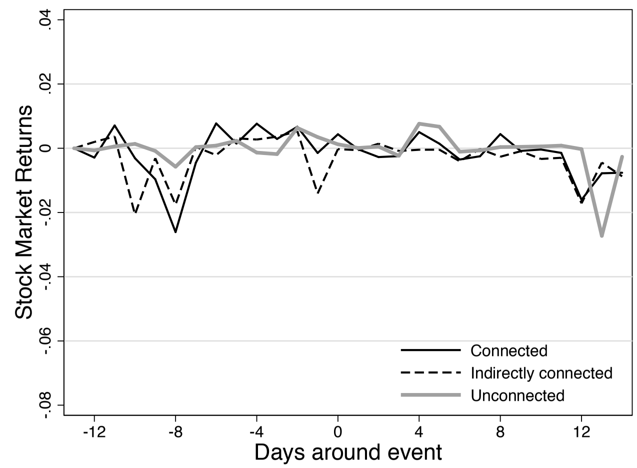


Notes: Green vertical lines represent the 1988 plebiscite and the beginning of democracy in Chile.

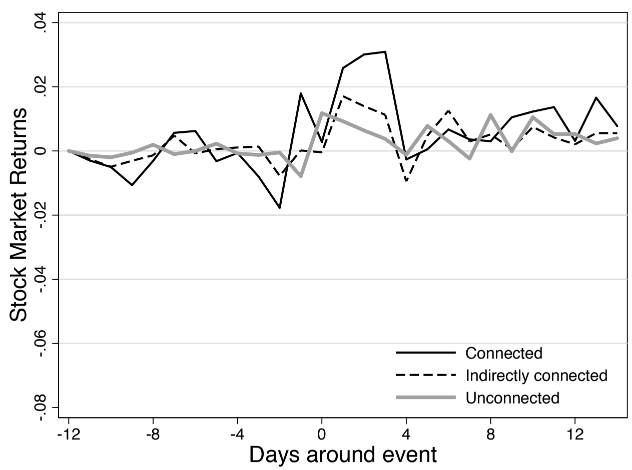
Figure C.4: Stock returns around other political events



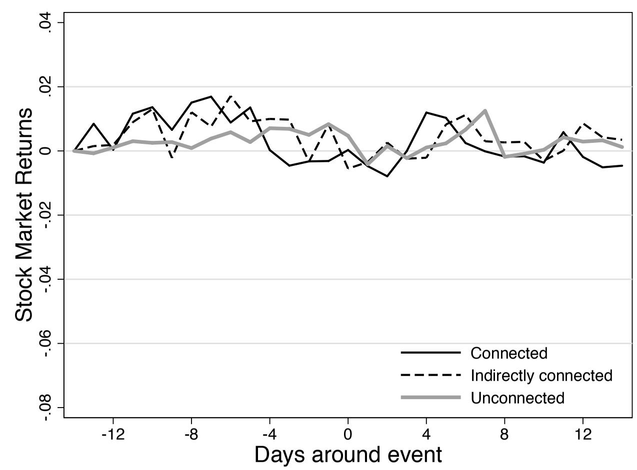
(a) Pinochet's nomination (08/30/1988)



(b) Constitutional reform (07/28/1989)



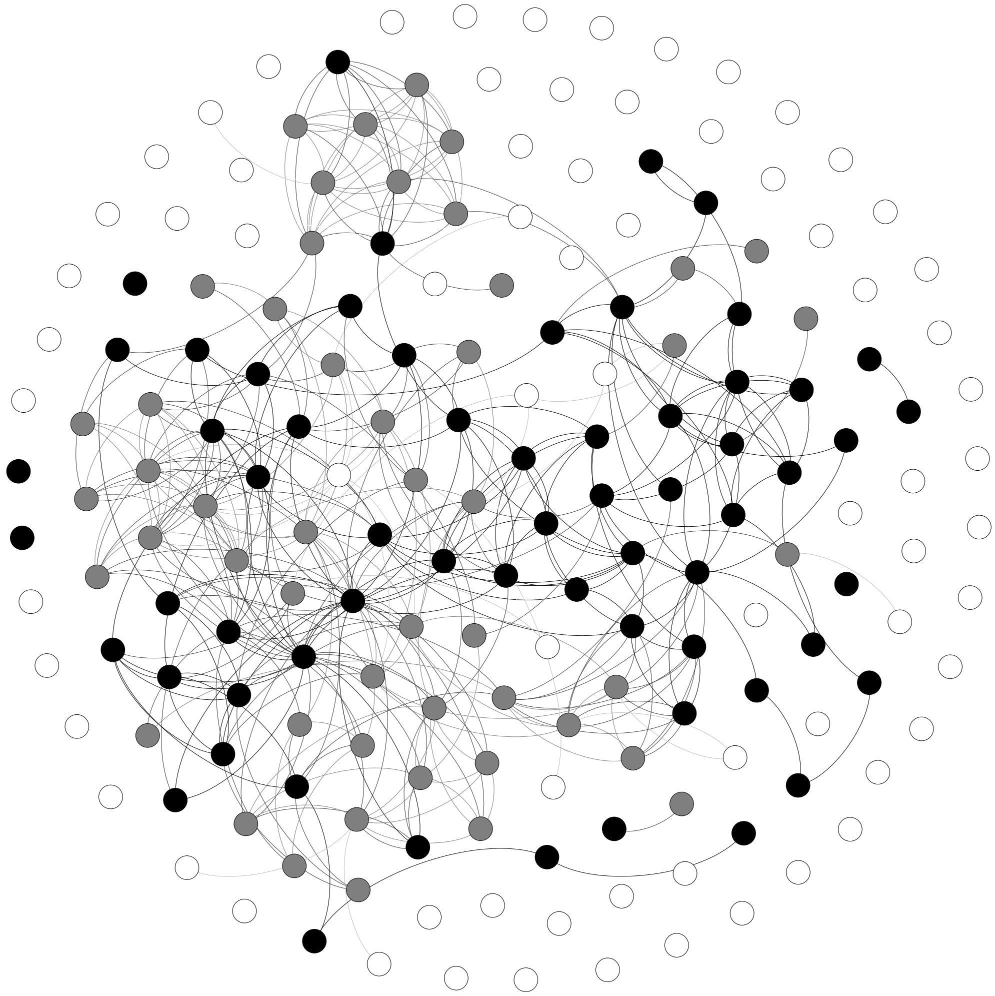
(c) Presidential election (12/13/1989)



(d) Pinochet leaves power (11/03/1990)

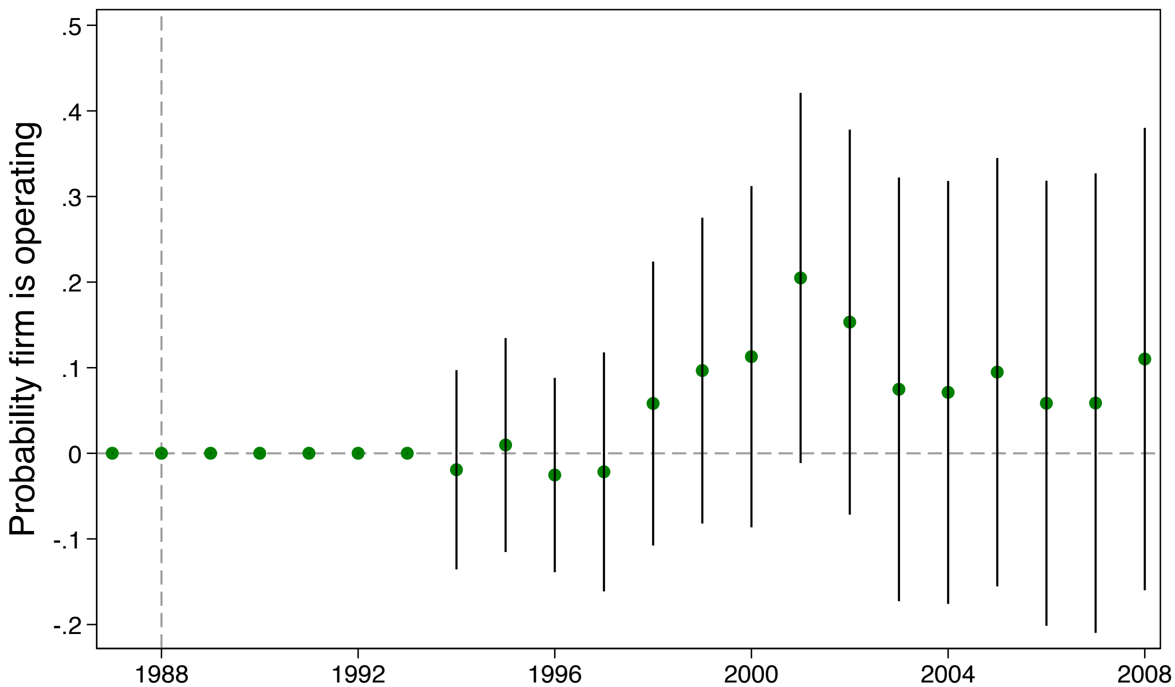
Notes: Stock returns are $R_{it} \equiv \ln S_{it} - \ln S_{it-1}$, where S_{it} is stock price i in day t.

Figure C.5: Network of firms during dictatorship



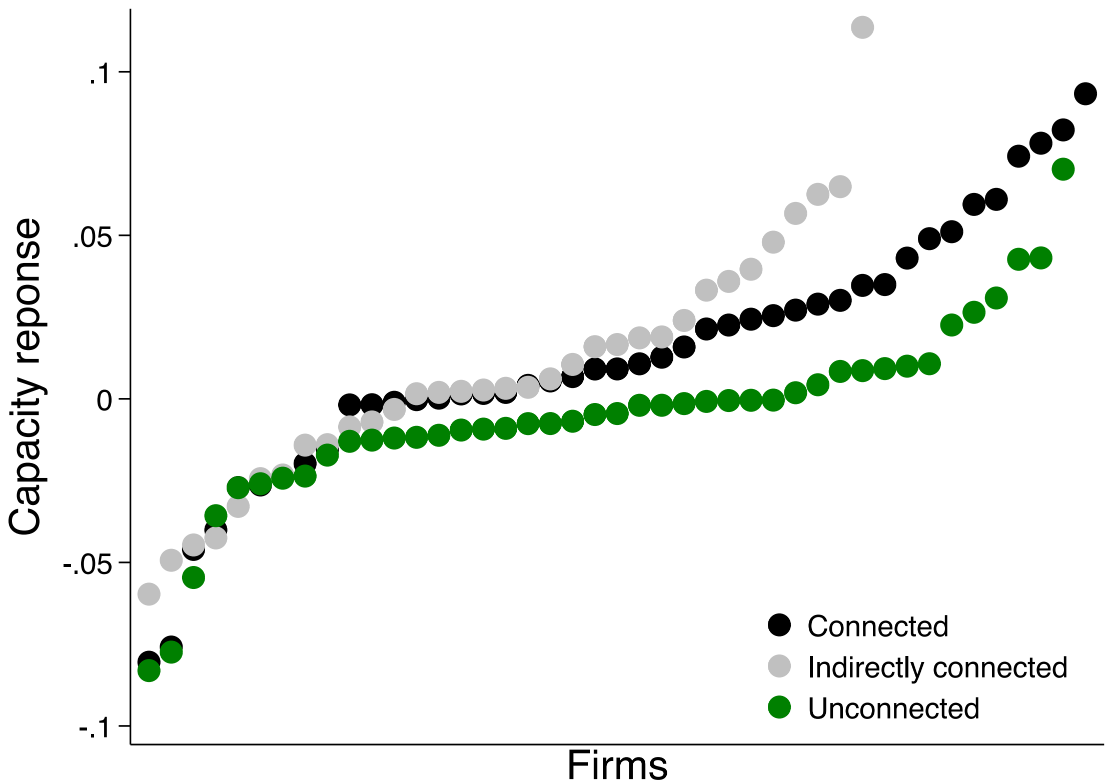
Notes: Undirected graph of firms operating in 1987. A link between firms is defined using board linkages. Firms denoted by are connected to Pinochet's dictatorship, firms denoted by are unconnected but have a link with a connected firm, and firms denoted by are unconnected and have no link to a connected firm. Firms are on average connected with 4.7 other firms by interlocking directors. The average number of links (distance) between any two firms that can be connected is 3.3, while the maximum distance (diameter) between any two firms is 9. The global clustering coefficient is 0.48. Finally, the fraction of firms in the giant network, network of fully connected firms, is 0.44. Similar to Khwaja et al. (2011), our network of interlocking directors shows some features of a "small world," low diameter and high clustering (Jackson and Rogers, 2005).

Figure C.6: Firm survival (indirect connections)



Notes: We present OLS $\hat{\delta}_t$ coefficients from $Y_{ijt} = \alpha_t + \beta_t P_{i,1987} + \delta_t \widetilde{P}_{i,1987} + \boldsymbol{\gamma}_t \mathbf{X}_{i,1987} + \boldsymbol{\psi}_j + \varepsilon_{it}$, where Y_{ijt} is an indicator that takes the value of one if firm i is operating in year t, $P_{i,1987}$ and $\widetilde{P}_{i,1987}$ are indicators for first and second degree connections, $\mathbf{X}_{i,1987}$ is a set of control variables (firm size, indicator for privatized firms), and $\boldsymbol{\psi}_j$ is a set of industry fixed effects.

Figure C.7: Firm capacity responses



Notes: Firm specific capacity responses during the lame duck period. Firms are ordered by the size of capacity responses, from firms decreasing their capacity in the left-hand side to firms increasing their capacity in the right-hand side.

Table C.1: Example of a politically connected firm

Director Name	Job in the Pinochet regime	Years in job
Guillermo Letelier	Army Officer	1980s
Sergio Melnik	Minister of Planning	1987
Julio Ponce Lerou	Pinochet's son in law	1969–1991
Enrique Valenzuela	Minister of Mining	1975–1978
Sergio Valenzuela	Minister of Planning	1985
Nine other directors in Board	Unconnected	–

Notes: Names and connections of individuals working in the Board of Directors of the *Chemical and Mining Society of Chile* in 1987, a Chilean chemical company and supplier of industrial chemicals. Board of Directors data comes from *Superintendencia de Valores y Seguros*.

Table C.2: Daily stock returns around the 1988 plebiscite

Dependent variable is cumulative abnormal stock returns

Days after event:	Same day	0-5 days	0-10 days
P	-0.05***	-0.08***	-0.07***
	(0.02)	(0.02)	(0.03)
$\widetilde{P}$	-0.05**	-0.08***	-0.10***
	(0.01)	(0.02)	(0.03)
Firms	80	80	80

Notes: This table shows a cross section regression of abnormal cumulative stock returns, defined as in Acemoglu et al. (2016b), on political connections. Abnormal returns were computed for the same day of the event, days 0 to 5, and days 0 to 10. P is a dummy for 1st degree (direct) political connections, and $\widetilde{P}$ for 2nd degree (indirect) political connections. Robust standard errors are reported in parentheses. Significance level: *** $p < 0.01$, ** $p < 0.05$.

Table C.3: Political connections by industry

Industry:	Unconnected	Directly link to Pinochet (P_i)	Indirect link to Pinochet ($\widetilde{P}_i$)	Total number of firms
Accommodation and food service activities	0	2	1	3
Agriculture, forestry, and fishing	3	4	8	15
Arts, entertainment, and recreation	0	1	0	1
Construction	2	1	0	3
Education	1	0	0	1
Electricity, gas, steam and air conditioning supply	4	6	2	12
Human health and social work activities	1	0	0	1
Information and communication	0	2	2	4
Manufacturing	13	19	13	45
Mining and quarrying	2	3	2	7
Real estate activities	14	1	3	18
Transportation and storage	1	3	1	5
Wholesale and retail trade	1	1	1	3
Total:	42	43	33	118

Notes: Number of firms by industry and connection type in our balanced panel of firms.

Table C.4: Robustness in standard errors

Dependent variable is investment or profits

	(1)	(2)	(3)
	Business group (benchmark)	Community detection algorithm	Wild cluster [p-value]
Investment			
P × Lame Duck	0.018**	0.018***	0.018***
	(0.008)	(0.005)	[0.000]
P × Democracy	0.020***	0.020***	0.020*
	(0.007)	(0.007)	[0.073]
$\widetilde{P}$ × Lame Duck	0.014	0.014	0.014
	(0.009)	(0.009)	[0.205]
$\widetilde{P}$ × Democracy	0.020**	0.020***	0.020**
	(0.008)	(0.007)	[0.044]
Profits			
P × Lame Duck	0.290***	0.290***	0.290***
	(0.100)	(0.068)	[0.000]
P × Democracy	0.194	0.194*	0.194**
	(0.124)	(0.085)	[0.031]
$\widetilde{P}$ × Lame Duck	0.115*	0.115**	0.115**
	(0.069)	(0.061)	[0.031]
$\widetilde{P}$ × Democracy	0.147	0.147	0.147
	(0.095)	(0.125)	[0.234]
Observations	4,694	4,694	4,694
Firm & time F.E.	Yes	Yes	Yes
Industry F.E. × Post	Yes	Yes	Yes

Notes: Balanced quarterly panel with 118 firms observed in the period 1985–1994. P and $\widetilde{P}$ are indicators for firms with direct and indirect political connections, respectively. Column 1: clusters are real business groups (104 clusters). Column 2: we use a community detection algorithm (Newman, 2004) to detect clusters in our network of firms (42 clusters). Column 3: column 2 with correction for the small number of clusters (Cameron et al., 2008). Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.5: Specification check for political connections

Dependent variable is investment or profits

	(1)	(2)
	Investment	Profits
Pooling direct and indirect connections		
$\hat{P}$ × Lame Duck	0.016**	0.213***
	(0.007)	(0.067)
$\hat{P}$ × Democracy	0.020***	0.173*
	(0.006)	(0.092)
Firms	118	118
Observations	4,694	4,692
Excluding indirect connections		
P × Lame Duck	0.019**	0.267**
	(0.008)	(0.106)
P × Democracy	0.021***	0.171
	(0.006)	(0.127)
Firms	85	85
Observations	3,378	3,376
Firm & time F.E.	Yes	Yes
Industry F.E. × Post	Yes	Yes

Notes: Balanced quarterly panel observed in the period 1985–1994. P and $\widetilde{P}$ are indicators for firms with direct and indirect political connections, respectively, and $\hat{P} = P + \widetilde{P}$. Robust standard errors are clustered at the business group level and are reported in parentheses. A total of 104 (88) clusters in the upper (lower) panel. Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.6: Using observed abnormal returns as political connections

Dependent variable is investment or profits

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Investment				Profits			
Abnormal Return × Lame Duck	-0.062	-0.053			-2.125*	-1.768*		
	(0.047)	(0.054)			(1.120)	(0.931)		
Abnormal Return × Democracy	0.014	0.023			0.355	0.713		
	(0.041)	(0.038)			(1.567)	(1.380)		
Indicator for low returns × Lame Duck			0.017*	0.016			0.361*	0.290*
			(0.009)	(0.010)			(0.199)	(0.170)
Indicator for low returns × Democracy			0.001	0.000			0.033	-0.038
			(0.007)	(0.007)			(0.278)	(0.234)
Firms	58	58	58	58	58	58	58	58
Observations	2,308	2,308	2,308	2,308	2,308	2,308	2,308	2,308
Firm & time F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry F.E. × Post	No	Yes	No	Yes	No	Yes	No	Yes

Notes: We use abnormal returns for the day after the plebiscite. *Indicator for low returns* is an indicator that takes the value of one for abnormal returns below the 33th percentile of the empirical distribution. Average abnormal returns have a mean of -0.06 and a standard deviation of 0.083. Robust standard errors are clustered at the business group level and are reported in parentheses (49 clusters). Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.7: Using Solow residual to measure productivity

	(1)	(2)	(3)
	Productivity		
$P \times$ Lame Duck	0.018	-0.084	-0.079
	(0.150)	(0.166)	(0.169)
$P \times$ Democracy	-0.059	-0.047	0.081
	(0.283)	(0.314)	(0.339)
$\widetilde{P} \times$ Lame Duck		-0.202*	-0.204*
		(0.114)	(0.113)
$\widetilde{P} \times$ Democracy		0.017	0.045
		(0.209)	(0.220)
Observations	792	792	792
Firm & time F.E.	Yes	Yes	Yes
Industry F.E. $\times$ Post	No	No	Yes

Notes: Unbalanced annual panel with 99 firms observed in the period 1985–1994. P and $\widetilde{P}$ are indicators for firms with direct and indirect political connections, respectively. Productivity was calculated estimating a Solow residual using capital stock and total labor. The mean (standard deviation) of productivity before the plebiscite is 0.028 (1.531). Robust standard errors are clustered at the business group level and are reported in parentheses (88 clusters). Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.8: Winsorising in different parts of empirical distribution

Dependent variable is investment or profits

	(1)	(2)	(3)	(4)
	Investment		Profits	
Winsorized at:	1%	5%	1%	5%
P × Lame Duck	0.024	0.014**	0.338**	0.338**
	(0.016)	(0.006)	(0.134)	(0.134)
P × Democracy	0.039***	0.012**	0.253*	0.253*
	(0.012)	(0.005)	(0.135)	(0.135)
$\widetilde{P}$ × Lame Duck	0.018	0.010	0.103	0.103
	(0.015)	(0.006)	(0.070)	(0.070)
$\widetilde{P}$ × Democracy	0.033***	0.014**	0.117	0.117
	(0.012)	(0.006)	(0.086)	(0.086)
Observations	4,694	4,694	4,692	4,692
Firm & time F.E.	Yes	Yes	Yes	Yes
Industry F.E. × Post	Yes	Yes	Yes	Yes

Notes: Balanced quarterly panel of 188 firms observed in the period 1985–1994. P and $\widetilde{P}$ are indicators for firms with direct and indirect political connections, respectively. Dependent variables are winsorized at 1% or 5% of the empirical distribution (bechmark is 2.5%). Robust standard errors are clustered at the business group level and are reported in parentheses. Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.9: Main results using the unbalanced panel of firms

Dependent variable is investment or profits

	(1)	(2)	(3)	(4)	(5)	(6)
	Investment			*Profits*		
P × Lame Duck	0.008	0.012*	0.015*	0.252***	0.336***	0.269***
	(0.007)	(0.007)	(0.007)	(0.086)	(0.094)	(0.098)
P × Democracy	0.006	0.012**	0.014**	0.233***	0.284**	0.211***
	(0.006)	(0.006)	(0.006)	(0.118)	(0.128)	(0.113)
$\widetilde{P}$ × Lame Duck		0.011	0.014*		0.178***	0.209**
		(0.007)	(0.008)		(0.066)	(0.080)
$\widetilde{P}$ × Democracy		0.014*	0.017**		0.100	0.122
		(0.008)	(0.007)		(0.118)	(0.091)
Observations	5,601	5,601	5,601	5,594	5,594	5,594
Firm & time F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Industry F.E. × Post	No	No	Yes	No	No	Yes

Notes: Unbalanced quarterly panel with 145 firms observed in the period 1985–1994. Each firm is observed at least six months before the plebiscite and six months in democracy. P and $\widetilde{P}$ are indicators for firms with direct and indirect political connections, respectively. Robust standard errors are clustered at the business group level and are reported in parentheses (126 clusters). Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.10: Attrition in the unbalanced panel

Dependent variable is an indicator for exit before 1994 – Probit regressions

	(1)	(2)
P	-0.096	-0.034
	(0.070)	(0.082)
$\widetilde{P}$	-0.123*	-0.119
	(0.074)	(0.079)
Indicator for large firms		-0.100
		(0.075)
Indicator for privatized firms		-0.087
		(0.067)
Observations	145	145
Industry F.E.	No	Yes

Notes: Marginal effects from probit regressions are presented. The indicator for large firms was constructed using the firm size distribution before the plebiscite. P and $\widetilde{P}$ are indicators for firms with direct and indirect political connections, respectively. Standard errors in parentheses. Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.11: Predicting political connections

Dependent variable is an indicator for political connections – Probit regressions

	(1) Direct connection	(2) Indirect connection
Log assets	0.186***	0.165***
	(0.044)	(0.045)
Privatized	0.134	0.012
	(0.152)	(0.164)
Business group	0.278	0.637***
	(0.175)	(0.116)
Leverage	0.140**	0.217*
	(0.059)	(0.122)
Exporter	-0.192	-0.006
	(0.168)	(0.157)
Age in 1987	0.005*	0.005*
	(0.003)	(0.002)
Firms	85	75

Notes: Probit regressions using indicators for connected and indirectly connected firms as dependent variables. Marginal effects are presented. In column 1 (2), we omit firms with indirect (direct) connections. We use the average of firm characteristics during the period 1985–1987. Standard errors in parentheses. Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.12: Firm characteristics using matching sample

Mean and standard deviation for the main variables in the period 1985–1987

	No link	Direct link to Pinochet (P)	Indirect link to Pinochet ($\widetilde{P}$)	Uni-variate regression		
	(1)	(2)	(3)	(2) – (1)	(3) – (1)	(2) – (3)
A. Quarterly dataset						
Investment	0.01	0.00	0.01	-0.01	0.00	-0.01
	(0.05)	(0.05)	(0.06)	(0.01)	(0.01)	(0.01)
Profits	-0.29	0.03	-0.19	0.31**	0.10	0.21
	(0.28)	(0.74)	(0.40)	(0.12)	(0.07)	(0.14)
Log assets	15.62	17.09	16.59	1.47***	0.96**	0.51
	(1.26)	(1.58)	(1.27)	(0.38)	(0.37)	(0.41)
B. Annual dataset						
Log workers	5.11	6.17	5.56	0.98***	0.13	0.84***
	(1.34)	(1.27)	(1.30)	(0.25)	(0.27)	(0.27)
Productivity	-0.08	-1.29	-0.98	-1.20***	-0.89***	-0.30
	(1.02)	(1.86)	(1.74)	(0.29)	(0.28)	(0.35)
Debt with government banks	4.88	8.32	13.54	3.43	8.65*	-5.21
	(19.71)	(27.90)	(13.54)	(3.93)	(5.13)	(5.68)
Debt with other banks	22.76	60.62	59.17	37.86***	36.41***	1.45
	(51.76)	(83.15)	(87.34)	(11.30)	(12.90)	(15.11)
C. Time invariant						
Age in 1987	38.65	49.99	47.75	11.54	9.15	2.40
	(28.08)	(28.14)	(23.74)	(7.71)	(7.44)	(7.74)
Exporter	0.37	0.40	0.55	0.08	0.11	-0.03
	(0.48)	(0.49)	(0.50)	(0.13)	(0.14)	(0.15)
Privatized	0.15	0.43	0.24	0.28**	0.07	0.21
	(0.36)	(0.50)	(0.42)	(0.12)	(0.12)	(0.14)
Business group	0.03	0.15	0.14	0.11	0.11	0.01
	(0.18)	(0.36)	(0.35)	(0.08)	(0.09)	(0.10)

Notes: Summary statistics and univariate regressions using a balanced sample based on the matching procedure proposed by Crump et al. (2009). Optimal bounds were computed using a propensity score estimated on the following pre-plebiscite characteristics: total assets, leverage, privatization indicator, exporting indicator, business group indicator, and years since foundation. Debt is measured in billions of Chilean pesos. Standard deviation in parentheses in columns 1-3, and standard errors in parentheses in the last three columns. Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.13: Main results using synthetic controls

	Direct connection		*Indirect connection*	
	Unweighted	Weighted	Unweighted	Weighted
A. Investment				
Lame Duck	0.020	0.017	0.018	0.012
	[0.004, 0.046]	[0.003, 0.045]	[0.005, 0.042]	[-0.004, 0.040]
Democracy	0.000	0.003	0.008	0.010
	[-0.008, 0.040]	[-0.006, 0.041]	[-.002, .040]	[-0.004, 0.057]
B. Profits				
Lame Duck	0.849	0.018	0.019	-0.005
	[0.78, 1.89]	[0.01, 0.56]	[-0.13, 0.58]	[-0.05, 0.46]
Democracy	0.425	-0.002	-0.073	-0.087
	[0.11, 1.63]	[-0.19, 0.45]	[-0.42, 0.79]	[-0.21, 0.74]

Notes: Average difference in quarterly investment for directly (indirectly) connected firm and synthetic controls. Following Abadie and Gardeazabal (2003) and Abadie et al. (2010), we construct synthetic controls for each connected and indirectly connected firms based on a set of firm characteristics in the period 1985–1987. In particular, we use log of total assets, leverage, and indicators for being part of a business group, being an exporter, and have been privatized during dictatorship. We present two average differences between treated and synthetic controls:

- Unweighted: $M = \frac{1}{N \times T} \sum_i^N \sum_t^T Inv_{it} - \hat{Inv}_{it}$, where $\hat{Inv}_{it} = \sum_{j \in Controls} w_j^i Inv_{jt}$ and w_j^i are weights based on the synthetic controls algorithm.

- Weighted: $Mw = \sum_i \alpha_i \times \sum_t \frac{Inv_{it} - \hat{Inv}_{it}}{T}$ where $alpha_i = \frac{1/\sigma_i}{\sum_i 1/\sigma_i}$ and σ_i is the goodness of fit of each synthetic control.

To compute confidence intervals, we conduct the following procedure:

1. Generate a bootstrapped sample, b, from control group.
2. Estimate w^i and compute M_b, Mw_b.
3. Repeat the procedure $B = 2,000$ times.
4. Compute [2.5, 97.5] percentiles of empirical distribution over the bootstrapped sample of M_b and Mw_b. These intervals are presented in brackets below the means.

Table C.14: Placebo treatments

Dependent variable is investment or profits

	(1)	(2)	(3)	(4)
	Investment		Profits	
Placebo in dictatorship				
$P \times$ Post	-0.005	0.001	0.157	0.236
	(0.007)	(0.008)	(0.145)	(0.156)
$\widetilde{P} \times$ Post		0.010		0.151**
		(0.008)		(0.064)
Placebo in democracy				
$P \times$ Post	0.007	0.012	0.089	0.159
	(0.008)	(0.010)	(0.109)	(0.099)
$\widetilde{P} \times$ Post		0.009		0.136
		(0.010)		(0.090)
Firm & time F.E.	Yes	Yes	Yes	Yes
Industry F.E. $\times$ Post	Yes	Yes	Yes	Yes

Notes: Balanced quarterly panel with 118 firms. We create a placebo treatment in the upper panel by splitting the pre-plebiscite period in two. Our placebo starts in 1986 third quarter, and samples cover the period 1985–1987. In the lower panel, we create a placebo treatment by defining a post period in December of 1993 after elections covering the period 1990–1997. P and $\widetilde{P}$ are indicators for firms with direct and indirect political connections, respectively. Robust standard errors are clustered at the business group level and are reported in parentheses (104 clusters). Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.15: Using political connections in 1986

Dependent variable is investment or profits

	(1) Investment	(2) Profits
Connections in 1986		
P × Lame Duck	0.017***	0.245**
	(0.008)	(0.102)
P × Democracy	0.016*	0.171
	(0.008)	(0.123)
$\widetilde{P}$ × Lame Duck	0.011	0.043
	(0.010)	(0.078)
$\widetilde{P}$ × Democracy	0.011	0.028
	(0.008)	(0.084)
Connections in 1986 or 1987		
P × Lame Duck	0.019**	0.243**
	(0.008)	(0.092)
P × Democracy	0.020***	0.207*
	(0.006)	(0.120)
$\widetilde{P}$ × Lame Duck	0.016*	0.073
	(0.008)	(0.069)
$\widetilde{P}$ × Democracy	0.018**	0.119
	(0.007)	(0.102)
Firms	118	118
Observations	4,694	4,692
Firm & time F.E.	Yes	Yes
Industry F.E. × Post	Yes	Yes

Notes: Balanced quarterly panel with 118 firms observed in the period 1985–1994. P and $\widetilde{P}$ are indicators for firms with direct and indirect political connections, respectively. In the upper panel, we use directors in 1986 to define degrees of political connections. In the lower panel, we use both 1986 and 1987 to define them. Robust standard errors are clustered at the business group level and are reported in parentheses (104 clusters). Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.16: Share of board with political connections

Dependent variable is investment or profits

	(1) Investment	(2) Profits
Share P × Lame Duck	0.076**	1.143***
	(0.031)	(0.435)
Share P × Democracy	0.040	1.369**
	(0.033)	(0.616)
Share $\widetilde{P}$ × Lame Duck	0.025	0.130
	(0.017)	(0.144)
Share $\widetilde{P}$ × Democracy	0.020**	0.093
	(0.009)	(0.167)
Observations	4,694	4,692
Firm & time F.E.	Yes	Yes
Industry F.E. × Post	Yes	Yes

Notes: Balanced quarterly panel with 118 firms observed in the period 1985–1994. Share P and Share $\widetilde{P}$ are total number of directors with political connections over board size in 1987. The mean (standard deviation) for these variables are 0.07 (0.11) and 0.14 (0.27) for direct and indirect connections. Robust standard errors are clustered at the business group level and are reported in parentheses (104 clusters). Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.17: Type of political connections I

Dependent variable is investment or profits

	(1)	(2)	(3)
	President	*Director*	*CEO*
$P \times$ Lame Duck	0.025***	0.009	0.017
	(0.009)	(0.009)	(0.022)
$P \times$ Democracy	0.018**	0.008	0.017
	(0.007)	(0.007)	(0.013)
$\widetilde{P} \times$ Lame Duck	0.011	0.008	0.005
	(0.009)	(0.009)	(0.008)
$\widetilde{P} \times$ Democracy	0.014*	0.013*	0.010
	(0.008)	(0.008)	(0.008)
Observations	4,694	4,694	4,694
Firm & time F.E.	Yes	Yes	Yes
Industry F.E. $\times$ Post	Yes	Yes	Yes

Notes: Balanced quarterly panel with 118 firms. P and $\widetilde{P}$ are indicators for firms with direct and indirect political connections, respectively. In each column we define P using different position of the connections: president of board of directors, director, and CEO. Sample period is 1985–1994. Robust standard errors are clustered at the business group level and are reported in parentheses. The number of clusters is 104. Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.18: Type of political connections II

Dependent variable is investment

	(1)	(2)	(3)	(4)	(5)
	Army Officer	*Minister*	*Politician*	*Advisor*	*Family*
$P \times$ Lame Duck	0.009	0.030***	0.018*	-0.000	0.008
	(0.010)	(0.010)	(0.009)	(0.014)	(0.012)
$P \times$ Democracy	-0.006	0.018**	0.011**	0.019**	-0.027
	(0.011)	(0.008)	(0.005)	(0.009)	(0.016)
$\widetilde{P} \times$ Lame Duck	0.006	0.012	0.007	0.004	0.004
	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)
$\widetilde{P} \times$ Democracy	0.008	0.014*	0.011	0.012	0.009
	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)
Observations	4,694	4,694	4,694	4,694	4,694
Firm & time F.E.	Yes	Yes	Yes	Yes	Yes
Industry F.E. $\times$ Post	Yes	Yes	Yes	Yes	Yes

Notes: Balanced quarterly panel with 118 firms observed in the period 1985–1994. P and $\widetilde{P}$ are indicators for firms with direct and indirect political connections, respectively. In each column, we define P using different types of connections: army officers, ministers, politicians, advisors, and close family. Robust standard errors are clustered at the business group level and are reported in parentheses. The number of clusters is 104. Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.19: Substitution of political connections

Dependent variable is investment or profits

	(1)	(2)	(3)	(4)	(5)	(6)
	Investment			Profits		
P × Lame Duck	0.018**	0.019**	0.016*	0.290***	0.282**	0.291**
	(0.008)	(0.008)	(0.008)	(0.100)	(0.110)	(0.119)
P × Democracy	0.020***	0.021***	0.021***	0.194	0.186	0.184
	(0.007)	(0.007)	(0.007)	(0.124)	(0.138)	(0.137)
$\widetilde{P}$ × Lame Duck	0.014	0.014	0.016*	0.115*	0.116*	0.161**
	(0.009)	(0.009)	(0.009)	(0.069)	(0.068)	(0.069)
$\widetilde{P}$ × Democracy	0.020**	0.020**	0.022***	0.147	0.149	0.194**
	(0.008)	(0.008)	(0.008)	(0.095)	(0.095)	(0.096)
$\widehat{P}$ × (Lame Duck + Democracy)		-0.006			0.053	
		(0.010)			(0.237)	
Firms	118	118	111	118	118	111
Observations	4,694	4,694	4,419	4,692	4,692	4,415
Firm & time F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Industry F.E. × Post	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Balanced quarterly panel with 118 firms observed in the period 1985–1994. P and $\widetilde{P}$ are indicators for firms with direct and indirect political connections, respectively. $\widehat{P}$ is an indicator for firms that substituted political connections from the old to the new regime. Columns 2 & 5: we include a control variable for the firms that substituted connections for the period after the democratization announcement. Columns 3 & 6: we exclude firms that substituted connections. Robust standard errors are clustered at the business group level and are reported in parentheses (104 clusters in columns 1-2 and 4-5, and 98 clusters in columns 3 and 6). Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.20: Government banks

Dependent variable is total debt with banks, measured from annual statements

	Bailouts and Privatizations	Large Privatizations	Connected Banks
	(1)	(2)	(3)
P × Lame Duck × Gov. bank	38.33**	40.10***	37.02***
	(17.07)	(15.05)	(12.91)
$\widetilde{P}$ × Lame Duck × Gov. bank	24.45	28.93	33.49*
	(20.02)	(19.52)	(17.96)
P × Democracy × Gov. bank	26.35	16.99	22.07
	(24.11)	(23.89)	(22.01)
$\widetilde{P}$ × Democracy × Gov. bank	17.69	9.64	23.31
	(19.90)	(17.79)	(18.45))
P × Lame Duck	-25.00**	-24.33**	-22.35**
	(10.08)	(9.49)	(9.20)
P × Democracy	-23.85	-18.78	-20.00
	(24.91)	(24.02)	(21.41)
$\widetilde{P}$ × Lame Duck	-25.94	-27.85	-27.99
	(18.92)	(18.39)	(17.24)
$\widetilde{P}$ × Democracy	-24.89	-20.04	-24.04
	(19.89)	(18.81)	(17.41)
Lame Duck × Gov. Bank	1.65	-3.74	-3.66
	(6.67)	(6.58)	(6.92)
Democracy × Gov. Bank	2.14	-2.85	-4.17
	(7.76)	(7.46)	(7.90)
Firm-bank & year F.E.	Yes	Yes	Yes
Industry F.E. × Post	Yes	Yes	Yes
Mean of dep. variable	40.25	39.44	37.59

Notes: Estimation using the annual panel of firms in the period 1985–1994. P and $\widetilde{P}$ are indicators for firms with direct and indirect political connections, respectively. Column 1: we define government banks as state-owned banks and all banks privatized and/or that experienced a bailout during the dictatorship. Column 2: we define government banks as state-owned banks and all large banks privatized during the dictatorship (CEME, 2004). Column 3: we define government banks as state-owned banks and politically connected banks. We have information about the board of directors of five banks, one of them connected to the dictatorship. Robust standard errors are clustered at the business group level and are reported in parentheses (99 clusters). Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.21: Robustness checks for the credit market

Dependent variable is total debt with banks, measured from annual statements

	Pre/Post	*Large firms*	*Privatized*	*Bus. Group*	*Exporter*	*All*	*Pscore*	*Matching*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$P \times$ Lame Duck $\times$ Government bank	26.805**	25.407**	17.467	36.240***	27.178**	22.023*	30.666***	18.067
	(11.381)	(11.626)	(12.394)	(11.000)	(10.561)	(12.174)	(10.924)	(11.367)
$\widetilde{P} \times$ Lame Duck $\times$ Government bank	30.896	24.648	22.730	40.513	24.055	31.408	29.175	15.869
	(19.822)	(18.668)	(17.616)	(25.132)	(18.061)	(20.562)	(20.040)	(14.801)
$P \times$ Democracy $\times$ Government bank	2.315	-0.080	-7.690	10.818	2.040	-3.135	6.658	-3.662
	(16.261)	(15.531)	(15.928)	(16.068)	(15.827)	(15.870)	(14.467)	(17.396)
$\widetilde{P} \times$ Democracy $\times$ Government bank	16.491	13.315	11.502	29.163	12.862	20.082	19.067	11.549
	(18.998)	(17.920)	(17.636)	(23.074)	(17.845)	(18.896)	(19.074)	(16.470)
$P \times$ Lame Duck	-15.337	-12.238	-12.056	-25.090**	-16.765*	-13.921	-17.614*	-16.445
	(10.652)	(11.510)	(11.500)	(10.157)	(9.783)	(11.775)	(10.503)	(11.811)
$P \times$ Democracy	-3.808	-0.454	-0.578	-13.391	-5.524	-2.402	-3.201	5.811
	(16.961)	(16.560)	(16.619)	(16.954)	(16.066)	(16.731)	(14.529)	(17.740)
$\widetilde{P} \times$ Lame Duck	-23.046	-16.317	-18.169	-31.661	-17.671	-23.460	-20.009	-10.736
	(15.920)	(15.407)	(15.048)	(19.118)	(14.605)	(17.123)	(16.568)	(16.215)
$\widetilde{P} \times$ Democracy	-12.736	-9.623	-11.619	-25.008	-11.244	-16.788	-10.991	-11.635
	(15.533)	(15.112)	(15.242)	(17.690)	(14.617)	(15.854)	(17.260)	(17.817)
Lame Duck $\times$ Government Bank	2.336	-0.922	-1.683	3.326	-2.445	-5.089	2.951	3.165
	(5.510)	(4.772)	(5.623)	(5.281)	(6.069)	(5.663)	(5.314)	(7.356)
Democracy $\times$ Government Bank	1.929	-1.361	-2.374	2.872	-3.141	-5.875	3.140	2.494
	(5.921)	(5.213)	(5.984)	(5.861)	(6.527)	(5.890)	(6.699)	(8.000)
Control $\times$ Post $\times$ Government Bank		9.791	28.784	-31.025	15.504		-0.731	
		(11.689)	(19.475)	(19.049)	(13.475)		(3.456)	
Control $\times$ Post		16.008	-16.303	27.231*	-22.989*		-1.583	
		(11.056)	(15.585)	(13.851)	(13.307)		(3.656)	
Firm-bank & time F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry F.E. $\times$ Post	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firms	113	113	113	113	113	113	113	72
Observations	2,075	2,075	2,075	2,075	2,075	2,075	2,075	

Notes: Estimation using the annual panel of firms in the period 1985–1994. P and $\widetilde{P}$ are indicators for firms with direct and indirect political connections, respectively. Robust standard errors are clustered at the business group level and are reported in parentheses (99 clusters in columns 1 and 2, 75 clusters in column 3). Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.22: Other sources of funding

Dependent variable: issuance (total amount issued in shares and bonds) or an indicator for issuance greater than zero

	Shares			Bonds		
	Issuance	$\mathbf{1}$[Issuance $>$ 0]	log(Issuance)	Issuance	$\mathbf{1}$[Issuance $>$ 0]	log(Issuance)
	(1)	(2)	(3)	(4)	(5)	(6)
P_i × Lame Duck	-28.60	0.00	-9.01***	2.83	0.02	1.63*
	(59.34)	(0.04)	(2.38)	(2.21)	(0.04)	(0.94)
P_i × Democracy	198.34	0.08	-2.92**	2.27*	-0.03	–
	(158.07)	(0.05)	(1.14)	(1.15)	(0.03)	–
$\widetilde{P}_i$ × Lame Duck	118.57	0.07	-2.52**	0.19	0.05	0.29**
	(99.92)	(0.05)	(0.97)	(0.62)	(0.04)	(0.13)
$\widetilde{P}_i$ × Democracy	9.83	0.12***	-3.69***	0.62	0.05*	–
	(58.19)	(0.05)	(0.74)	(0.72)	(0.03)	–
Observations	1,107	1,107	100	1,107	1,107	54
Firms	112	112	53	112	112	29
Firm & time F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Industry F.E. × Post	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Estimation using the balanced annual panel with 112 firms in the period 1985–1994. P and $\widetilde{P}$ are indicators for firms with direct and indirect political connections, respectively. Issuances are measured in $MMM Chilean pesos. Robust standard errors are clustered at the business group level and are reported in parentheses. In columns 1, 2, 4, and 5, the number of clusters is 98; in column 3 (6) is 48 (28). Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.23: Extraordinary dividends

Dependent variable is payment of extraordinary dividends

	Extraordinary Dividends		Payment of Extraordinary Dividends	
	(1)	(2)	(3)	(4)
P × Lame Duck	0.014	0.018	-0.090	-0.187
	(0.015)	(0.018)	(0.116)	(0.136)
P × Democracy	-0.041***	-0.030*	-0.291***	-0.369**
	(0.013)	(0.015)	(0.107)	(0.145)
$\widetilde{P}$ × Lame Duck		0.009		-0.208
		(0.015)		(0.133)
$\widetilde{P}$ × Democracy		0.021		-0.158
		(0.017)		(0.164)
Observations	581	581	581	581
Firm & time F.E.	Yes	Yes	Yes	Yes
Industry F.E. × Post	Yes	Yes	Yes	Yes

Notes: Balanced quarterly panel with 72 firms. P and $\widetilde{P}$ are indicators for firms with direct and indirect political connections, respectively. A mandatory dividend of 30% of earnings is part of the Chilean regulation. The board of the company can decide to pay extraordinary dividends above this threshold. Extraordinary dividends are defined as this payment over total assets. Columns (3) and (4) use a dummy for the payment of extraordinary dividends. Sample period is 1985–1994. Robust standard errors are clustered at the business group level and are reported in parentheses. The number of clusters is 65. Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.24: Entry barriers at the industry level

Dependent variable is investment

	(1)	(2)	(3)
P × Lame Duck × High Entry Costs	0.011	0.016	0.024*
	(0.012)	(0.011)	(0.011)
P × Democracy × High Entry Costs	-0.005	0.007	0.015
	(0.011)	(0.011)	(0.011)
$\widetilde{P}$ × Lame Duck × High Entry Costs		0.001	0.010
		(0.028)	(0.030)
$\widetilde{P}$ × Democracy × High Entry Costs		0.020	0.029
		(0.018)	(0.018)
P × Lame Duck	0.002	0.005	-0.003
	(0.008)	(0.005)	(0.009)
P × Democracy	0.015	0.015*	0.007
	(0.008)	(0.008)	(0.009)
$\widetilde{P}$ × Lame Duck		0.013	0.008
		(0.026)	(0.029)
$\widetilde{P}$ × Democracy		0.003	-0.002
		(0.015)	(0.017)
Firms	118	118	118
Observations	4,694	4,694	4,694
Firm & time F.E.	Yes	Yes	Yes
Industry F.E. × Post	No	No	Yes

Notes: Balanced quarterly panel with 118 firms observed in the period 1985–1994. P and $\widetilde{P}$ are indicators for firms with direct and indirect political connections, respectively. We follow Lambson and Jensen (1998) and define average sunk cost in each industry by taking the average of gross book value of property, plant, and equipment using our unbalanced panel during the period 1985 to 1987. High entry costs is defined as those industries with an average above percentile 50. Robust standard errors are clustered at the business group level and are reported in parentheses (104 clusters). Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.