The London School of Economics and Political Science

Essays in Economic Geography: School Vouchers, Student Riots and Maternal Surrogacy

Piero Francesco Montebruno Bondi

A thesis submitted to the Department of Geography and Environment of the London School of Economics for the degree of Doctor of Philosophy, London, September 2016.

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of 65,358 words.

Piero Francesco Montebruno Bondi

Statement of use of third party for editorial help

I can confirm that my thesis was copy edited for conventions of language, spelling and grammar by London School of Economics (LSE) Official Proofreading Service.

Abstract

In this thesis, I investigate spatial aspects of education and family economics. In the first chapter, I explore the effect of voucher school competition on pupil achievement in Chile. Specifically, I create spatial indices to measure spatially determined competition: a choice index which counts the number of schools that are accessible from a given municipality; and a competition index which summarizes the choice index for a given community of students. The chapter tests the hypothesis that schools which spatially compete more are also more efficient. The results show no effect of spatially determined competition on value added. I discuss how the absence or slow response of parents to "poorly performing" schools and a "too low" voucher can be proposed as two of the causes of the poor functioning of the voucher system. In the second chapter, I exploit a police report on occupied schools in the so-called *Chilean Winter*—a huge social outburst of pupil protests, walk-outs, riots and school occupations, which started in early June of 2011—and test the hypothesis that a decrease in attendance has a causal effect on reducing students' performance in standardized tests. My evidence indicates that the performance of pupils affected by missed days from school dropped to nearly 0.18σ, which is sizeable in terms of human capital accumulation. In the last chapter, I produce the first quantitative evaluation of maternal surrogacy. I exploit variation in surrogacy legislation in every US state over time and study surrogacy's causal effect on vital statistics such as marriage, divorce, births and out-of-wedlock births. Using arguably exogenous changes in legislation to identify the causal impact of surrogacy, I show that one additional standard deviation in the surrogacy rate causes an increase of 0.05σ in the number of marriages and of 0.04σ in the number of divorces. It also causes a decrease of -0.02σ in births and of -0.03σ in out-of-wedlock births. The three chapters introduce novel results that advance current knowledge and should be carefully considered by policy makers in these areas.

Acknowledgements

I decided which courses to take in the first year of my Ph.D. during one of my initial meetings with my supervisors, Professor Stephen Gibbons and Dr. Olmo Silva, to whom I am deeply indebted for their guidance, academic generosity and sense of humour. They indicated that I should take a course on causality instead of an advanced R-coding workshop. LSE, they added, has been a causality school "since its origins". Not in vain its motto is "Rerum cognoscere causas", which means "to know the causes of things". In fact, LSE is perhaps the oldest institution in continuous operation to put causality at the center of its debate. Causality has been studied extensively in the past: The Four Causes of Aristotle; the Chinese ideograms, which provide a great territory for visual experimentation on the subject; and the Bhagavad Gita, which distinguishes "causes of things", to cite but a few examples. Hume is modern history's starting point for causality. Recently, Kerouac et al. borrowed from the elaborate Buddhist Chain of Causation. Perhaps it would be of interest to explore whether a society has existed that did not consider the notion of causation, or did not include a word in their language to express "a cause" or "to cause". This early advice from my supervisors proved to be of great help and breadth, and I began to investigate causality and causal econometrics. When a paper is first published, it gets "priority of discovery" over those next-in-line. Therefore, social science is inevitably a dialogue with the past. This is why I mention in the acknowledgements that my inspiration was incited by Jerzy Neyman (1923) in one of the papers of his doctoral thesis for the University of Warsaw which was written in Polish with an extensive comment in French while he was at the Agricultural Institute of Bydgoszcz.

In these "dialogues with old friends" Donald Rubin (1974)'s paper in the *Journal of Educational Psychology* has a prominent position. I have also been impressed by the breath-taking books by Judea Pearl (1988, 2009). Causality quickly brought me to the first and foremost activity of a post-graduate research student: identification strategies crafts. That is, creating useful and credible identification strategies or effective ways of employing observational data to resemble a randomized trial. As a matter of fact, simple economic genius in crafting observational data into credible identification strategies is at the core of economic research, transforming causality from a lost paradise into a heaven found. As an intellectual voyager through both causality and identification strategy, I became addicted to "Mostly Harmless Econometrics" (MHE) by Angrist and Pischke (2009), which may just be the most popular book on campus these days. I was also drawn to a little red and black book with the informative title "Counterfactuals and Causal Inference"(CCI) by Morgan and Winship (2010). MHE's focus is closer to economics, with a great deal of content devoted to building intuition. CCI also takes a statistical approach and includes graphs and visual tools. I realized that many MHE arguments could be translated into new, innovative graphs which could become a pedagogical device. Some examples are presented here. I wish to express my heartfelt gratitude to Conicyt and his Chile Scholarship programme for their outstanding help in financing my studies. I also deeply thank the Chilean Ministry of Education's Agency for Quality in Education (AQE, 1999-2013) for kindly allowing me access to the SIMCE dataset, which I use in the school voucher and student riots chapters. I also thank the Social Observatory of the Ministry of Social Development for allowing me access to the Casen survey (Casen, 2000-2011). These papers have been presented in multiple seminars and have benefited from the comments of Jeff Grogger (Chicago Harris School of Public Policy), Laurent Gobillon (Paris School of Economics), Michela Tincani (University College London) and Corinne Low (Wharton School). I also acknowledge the inspiring teaching of Professor Lena Edlund on the economics of sexuality at Columbia Economics Department. Finally, I want to thank my fellow students in the Department of Geography and Environment for their wonderful insights (especially Nat

Picarelli, Pascal Jaupart, Susanne Frick, Davide Luca, David Nguyen, Paula Prenzel, Pablo Navarrete, Eduardo Ibarra-Olivo, Tanner Regan, Jamila Nigmatulina) and, also, Julia Bird, Kenzo Asahi and Cristián Crespo for excellent conversations on these topics. As stated by Professor Henry Overman during my upgrade meeting, "Writing a good paper is about balancing or treading a tight rope between telling a good story and presenting scientific facts clearly". It is now your turn to decide if I have been successful. Last but not least, any errors are mine.

Contents

Figures

Tables

Introduction and Critical Discussion

In this thesis, I investigate spatial aspects of education and family economics. Firstly, I study the performance of the Chilean school voucher system from a spatial perspective. This is of key public policy interest because the voucher scheme is believed to outperform more centrally planned interventions. The relevant literature is full of Chilean voucher scheme assessments, but until my study there has not been a spatial evaluation of voucher performance. Elucidating whether spatial incentives, i.e. spatial economic stimuli, can increase school value-added is of central importance in the theory of education economics. Assessing whether schools that serve a wider—more distant—audience of students are also the ones adding more value to them, can permit us to address the question of whether spatially determined competition plays a role in Chilean education. Secondly, I am also measuring the decrement in human capital accumulation caused by lost school days in the context of student riots. Since the so-called "Penguin Revolution", an increasingly strained atmosphere between students and the government has resulted in long and ongoing student walk-outs and riots. This thesis tries to measure exactly how much school absence reduces performance among students using the exogenous variation embedded in the student riots. In the Chilean debate, this pure assessment is not part of the discussion. Both students and policy makers will gain from knowing exactly the amount of diminished education caused by missed school days. Finally, I study a novel area of economic assessment: surrogacy and surrogacy legislation in the US.

From a policy perspective, this is the first applied approach with a new dataset on surrogacies in each US state. Surrogacy is a new practice, broadly debated but poorly understood. I use this unique dataset to answer questions regarding whether the marriages, divorces and out-of-wedlock births are affected by the practice of surrogacy. Surrogacy is still an open debate and policy makers will certainly benefit from the outcomes of my study.

More precisely, I explore the spatial relationships between schools and school performance through school choice and competition in the long-standing voucher educational system in Chile. I also study the spatial liaison between student performance on standardized tests and missed school days in the context of student protesting, during the *Chilean Winter*, against what they perceive to be an unfair educational system. Finally, I investigate the spatial characteristics of the relationship between surrogacy in the various US states and its effects on vital statistics such as those related to marriage, divorce and births. The chapters attempt to answer the following questions: "Is spatially determined competition increasing the productivity of schools in the Chilean voucher system?"; "Are the missing school days due to the student riots in the *Chilean Winter* decreasing the performance of those students measured as the standardized SIMCE test?"; and "Is the increase in surrogacy in the US causing more marriages, more divorces and fewer births?". Causal empirical work is needed to attempt to answer these questions. I next introduce the frameworks I will use in my thesis.

A causal dissection of the three chapters

Each of the problems I investigate is subject to endogeneity concerns, in that my causal variable/treatment is partly affected by factors, some unobserved, that also affect the outcome variable. For example, when studying school competition, ability is in the error term and is correlated with the causal variable/treatment, i.e. school

competition is endogenous. Therefore, a central focus of the thesis is on providing estimates of the effects of the causal variable/treatment on the outcome that can be interpreted as causal. By causal, I mean a connection between the causal variable/treatment and the outcome as a functional relationship that describes what a given outcome would be if the causal variable/treatment took a certain value, i.e. if I could change the causal variable/treatment in a perfectly controlled environment, or change the causal variable/treatment randomly so that those with different levels of causal variable/treatment would be otherwise comparable. The framework I am using to discuss causality here follows that of Angrist and Pischke (2009) and Morgan and Winship (2010). As a matter of fact, each of my essays has an explained outcome, a causal variable/treatment, an instrument and specific confounders (or variables that fog the causal link):

chapter	explained outcome	causal variable (treatment)	instrument	confounders
Chilean voucher	SIMCE test scores	spatially determined competition	slope/ruggedness	ability residential location travel time
Chilean Winter	SIMCE test scores	missed school days	school occupation	ability "compensatory" study
Surrogacy in the US	vital statistics (marriage, divorce, births)	surrogacy rate	surrogacy legislation	moral attitudes ("cons/libs")

In the Chilean voucher chapter (Chapter 1), I assess the effect of spatially determined competition—the causal variable which is measured by choice and competition indices—on standardized test scores—the explained outcome—for the city of Santiago. The main confounder is ability. "High-ability" students score higher test scores. At the same time, a pupil with more ability or better family background might travel a longer distance to school. Therefore, more spatially determined competitive schools will have better students on average, causing a positive bias on my coefficients. Class size is another confounder, because more popular schools where students travel longer distances will be crowded and so will have larger class sizes.

The effect of class size is deeply controversial but the estimates may be biased as a result. Residential location is yet another confounder, richer families choosing more competitive schools will most probably also choose closer residential location. So there might be a negative bias of school competition if students living close to the very productive schools lowering its competition indices. Finally, travel time is a related confounder with analogous interpretation and bias with richer pupils with more educated families travelling shorter distances with little travel time. To bypass these threatens to identification, I use slope and ruggedness as instruments to address the endogeneity issue introduced by the indices. Slope and ruggedness pick up building and travel costs, which translate into fewer schools for voucher schools but into school agglomeration for private schools because of competing reputation of being "on the top of the hill". Therefore, using this defensible external variation in the causal variable permits disclosure of the exact impact of spatially determined competition on test scores. Note that in my regressions I control for elevation and distance to border of the city of Santiago so that any link between these variables, family income and school reputation is controlled for. A balancing test shows that the instruments are conditionally uncorrelated to the unobservable. After addressing the endogeneity issues, I measure the causal effect of spatially determined competition on productivity and value added on pupil education. The chapter tests the hypothesis that schools which tend to spatially compete tend to provide more value added, or whether school performance is responsive to market forces.

In the *Chilean Winter* chapter (Chapter 2), I assess the effect of missed school days—i.e. the causal variable—on standardized test scores—i.e. the explained outcome. The confounder is again ability, which could create selection bias if more/less able pupils attend school regularly/miss more school days. In addition, "compensatory" student activities can emerge as confounders if, during school absences, students or students' parents offset school deprivation with out-of-school counterbalancing activities (e.g. extra reading or math exercises). A police report lists the schools which were occupied, providing an external source of variation in decreased

attendance so that I can test if missing school days decreases student performance as measured by the SIMCE, the Chilean standardized test. In short, this chapter aims to measure the causal effect of decreased attendance on test score performance, exploiting the arguably exogenous variation of riots in causing missed school days. The evidence shows that the performance of pupils affected by missed days from school due to the riots dropped by nearly 5%—0.18σ, which is a large number in terms of its cost in human capital. At the centre of the chapter, then, is the effect of diminishing school input on pupil achievement.

Finally, in the surrogacy chapter (Chapter 3) I assess the effect of increased surrogacy in a US state—the causal variable—on vital statistics such as marriage, divorce and births—i.e., the explained outcome. The confounders here are moral attitudes. "Liberal" states, besides having a higher number of surrogacies, also have lower marriage rates—and this situation generates selection bias. Variation in surrogacy legislation (the instrument) is the source of external variation that I use to isolate the causal effect of the surrogacy rate on vital statistics. States which pass more favourable legislation typically show more surrogacies than states where, for instance, surrogacies are criminalized. A balancing test shows that variation in legislation within states over time is arguably random after controlling for key covariates and year and state fixed effects. Exploiting this variation, I found that more surrogacy implies higher marriage rates, higher divorce levels and fewer births.

Final remarks

In summary, I attempt to consolidate a distinct contribution to education and family economics from an economic geography perspective. This thesis seeks to uncover new facts and relationships amongst school vouchers, student riots—as an instrument for missed school days—and maternal surrogacy, through the exercise of solid causal analysis. In each chapter, I present a critical assessment of the

relevant literature, the methodology of the research, the findings and a discussion with the main results and their implications, including external and internal validity and extensive robustness checks. As a conclusion, these essays are applied causal contributions as they answer specific "what-if" questions, and their judgement must be based on the creativity and relevance of the questions and the rigour and accuracy of the answers. It is now the time to prove their value in the community of social scientists and beyond.

Chapter 1

School Competition and Pupil Attainment in the Chilean Voucher System

1.1 Introduction

The market for Chilean education has been widely explored mainly because it is the oldest and most radical voucher system in the world. As schools are paid a subsidy for each pupil that enrols, the schools compete to attract and serve pupils. The voucher system thereby should promote competition among fiscally subsidized schools, with most of the competition exerted through the spatial engagement of pupils and families getting around and choosing schools: pupils and parents vote with their feet. This model predicts that schools that compete in a spatial manner to serve pupils will be more efficient as measured by standardized test results. A key characteristic of the system is its spatial nature and the main scope of this chapter is to systematically address that "spatiality". Previous well-known studies of the Chilean voucher system such as Mizala and Romaguera (2000), Hsieh and Urquiola (2002), Gallego (2006) and Chumacero et al. (2011) have not elucidated this aspect of the market. In fact, they have not used space in defining how competition should

be measured. They have, therefore, missed out on one of the key elements of the Chilean voucher system. If you want to understand whether a school is competitive or not, you must understand how geographically far its market stretches.

This paper, for the first time, uses space to infer competition from choice. My focus is on the Santiago municipalities. Pupils are not constrained to choose a school in the same municipality in which they are living; on the contrary, they can choose a school anywhere. The only major constraint is the potential travel time. Space thereby plays a central role in pupils' choice and school competition. I first infer from geographical patterns how pupils from a given municipality choose their schools. This inference constitutes my measure of how far they are willing to choose to travel. The measure of competition is based on these pupils' revealed preferences averaged across each school. A school will be more competitive if the pupils that are attending the school had more choices. By these means, schools competing harder for pupils in a spatial setting are also the schools that perform better on standardized tests. It is therefore necessary to test how space affects the productivity of the model if the central incentives of the pupils are spatial in nature.

By using choice and competition measures I am also introducing endogeneity to the model. For example, pupils with a better family background will travel from further away, which means that the school that they attend can have a biased measure of competition and a biased effect on test scores making it problematic to uncover the real causal effect. I have two solutions for this endogeneity problem. The first and foremost is an instrumental variable strategy. My instruments are slope and ruggedness, how fast the altitude is increasing and the terrain undulation, both of which are unlikely to be correlated with the unobservables and individual characteristics. These instruments are good predictors of competition because they are negatively correlated with voucher competition: the extra building cost is internalized in the market as fewer built schools, which drives down the indices. After controlling for distance to Santiago's urban footprint and elevation, which are both

proxies for "high class" (families with higher income live in elevated areas and close to the urban border), a balancing test shows the conditional orthogonality of my instruments to the unobservables, thus making the case for a smooth identification strategy. The second strategy comes from the fact that I observe pupils at two stages: in primary school and then again in secondary school. These two stages of observations allow me to use a value-added model to partial out individual fixed characteristics of individuals that affect at the same time the school choice and the educational attainment.

Discussion of the effects of school competition on school performance should envisage an input oriented approach to production of education. First, teacher quality as a confounder will upwardly bias test score outcomes because more competition will also signify more teacher quality and finally better test score outcomes. Schools that spatially compete to a greater extent attract better teachers, and the literature shows that school teacher quality systematically increases education outcomes, e.g. Hanushek (2003) finds a 0.11σ increase. At the same time, schools that spatially compete to a greater extent should attract more students, which will probably imply larger classes, which the literature has shown has disputable effects on student performance. Recall for example, the famous debate between Alan Krueger (2002) and Eric Hanushek (2002) in Mishel and Rothstein (Eds., 2002). In some of the recent literature this issue has been settled by researchers, such as Bingley, Jensen and Walker (2007) who state that reducing class size during compulsory schooling by 5% would increase mean length of education by about 8 days, which translates to approximately a 0.2% increase in lifetime earnings. Another input is accessibility/transport travel time which is also related to schools' spatial determined competition. Asahi (2014) finds that schools that have a large decrease of more than 4.7 km of distance to the nearest subway station—ending at walking distance from the subway network—have test scores 0.15σ lower.

Overall, my OLS results are positive but likely to be biased and my IV results

are marginally negative or insignificant when I measure the effect of voucher competition on voucher pupil test results. I analyse the heterogeneity of my main model and, in Appendix A, I present a model on changes. In Appendix B, I also "disaggregate" competition and pupils, with interesting additional results. In sum, although competition may still work for some specific schools or in some specific sectors, my results show that competition works poorly in general. In short, competition is not a tide that lifts all boats.

Within the Chilean literature, this chapter continues the spatial analysis started by Chumacero et al. (2011) who find that students are willing to "walk" 0.2σ of distance—a quarter of a km—to avoid paying one standard deviation more of the price of the nearest school—50 dollars. Furthermore, parents are willing to pay 1.5σ of the price—75 dollars—more to attend the nearest school if it improved its test in one standard deviation —30 points in SIMCE. Chumacero et al. (2011) acknowledge that they do not measure any causal link between spatially determined competition and test scores, proving that my research is the first to do so. Improving upon this study, my paper aims to measure the relevance of distance from the municipality of residence to school as a predictor of school choice. It also adds a new measurement of voucher efficiency to the well-known studies of Mizala and Romaguera (2000), Hsieh and Urquiola (2002) and Gallego (2006). My results compare best to those of Hsieh and Urquiola (2002). Using panel data for about 150 municipalities, the researchers found no evidence that choice improved average educational outcomes as measured by test scores, repetition rates, and years of schooling. As I said before, what distinguishes my data and results from these previous studies is their spatial nature. Only Chumacero et al. (2011) is spatial in nature, but they do not measure the effect of spatial choice on test scores, which is exactly what I assess in this chapter.

At the same time, this chapter contributes to the broad debate on school competition, e.g., to the work of Gibbons et al. (2006) and Gibbons and Silva (2008),

and aims to advance the extensive literature in the US on spatially determined choice and competition, such as the work of Hoxby (2002) and Cullen et al. (2000). While Hoxby (2002) based her analysis in a Tiebout world where competition comes from mobility, Gibbons et al. (2006) and Cullen et al. (2000) measured choice and competition conditional on where the agent is. Hoxby (2002) finds positive effects of competition, but some critical comments on this paper were subsequently published: Ladd (2003) indicated that she is deeply skeptical about the benefits of an educational market—primary and secondary education are compulsory, parental choices of school are heavily influenced by the composition of students, and absence of pricing mechanism, evidence on "cream-skimming" and impact of choice on students' achievement in choice of school; and Rothstein (2007) said Hoxby's key results do not seem robust to small, reasonable alterations to the sample or to the instrumental variable used. Most importantly, Gibbons et al. (2006) and Cullen et al. (2000) find no general effect of competition, only some effect on heterogeneous and selected groups. My methods, and therefore my results, are more similar to those of the latter papers. In fact, I find a non-negative correlation in primary and secondary schools between voucher competition and voucher test results in an OLS setting, and I find no significant or marginally significant results, in the IV setting which compares closely to Gibbons et al. (2006) results for London.

Lastly it is necessary to include an explanation of why this research paper question centres only on the educational outcomes—test scores—dismissing the costs. This is easily explicated by the fact that for Chile and the period of study no full data for productivity is available. For public voucher schools, the voucher amount is readily available but the often soft municipality budget is not. While for private voucher schools, the voucher is also available, but co-payment and profit are not. Voucher value for the period was around £40 per month for primary schools and £50 per month for secondary schools.[1] Co-payment for a later period, i.e. 2013, was

[1] A complete explanation of how voucher values are calculated can be obtained from the Chilean Ministry of Education Vouchers (2010) "Vouchers".

on average £17 for half of the private voucher schools that charge co-payment—for the other half it is just £0. Profit data are also not available, for a previous period and for US Hanushek (1994) reported profits on average equal to 60% of costs for all private schools.[2]

1.2 The school system

In the Chilean voucher system, there are no districts. As a result, every pupil is free to choose any school in any place he/she prefers. There are three main types of schools: public voucher schools, private voucher schools and private schools. Both types of voucher schools are fiscally subsided. Private schools receive funds directly from parents. Some voucher schools also receive a co-payment from the parents. In terms of the curriculum taught in Chile, the "what" to study is centrally defined but the "how" is flexible and decentralized. The quality of teachers and even their average age are different. Better teachers, for example, are in private schools and older teachers are in public voucher schools. The length of a school day is regulated for voucher schools by the full school day reform passed in 1997. As a result of this legislation, the time spent at school has increased by 30%—115% of the OECD average—since the onset of the reform. Length of school day is not regulated in private schools. Managerial practices vary from municipality administration for public voucher schools, where the mayor is the legal representative, to private administration for private voucher and private schools. Additionally, public voucher schools receive poorer and harder-to-teach students. All of these factors generate a great heterogeneity in the system with a continuum of quality/performance in which private schools are the top performers, followed by private voucher schools, and finally public voucher schools. Primary school starts at 6 years old and finishes eight years later at 14 years old. Then secondary school lasts four years until pupils are 18 years old. Some schools are mixed, enrolling primary and secondary pupils, but

[2]Again the on-going educational reform includes non-profit, non co-payment and non selection and full public accountability of state resources in the system, so full productivity measurement will be readily available.

most of the time (nearly 85%) each pupil chooses a different school for secondary school after primary school. Often, the decision to change schools happens before the complete cycle of primary schools ends; for instance, the so-called emblematic lycées (a group of well-performing public secondary schools) start in 7^{th} grade, a full two years before primary school ends. In reality, there is no rule for the transfer from primary school/education to secondary school/education. At the time of the sample (2002/2008), compulsory education was primary and secondary education; preschool education became compulsory only in 2013. Finally, it is important to note that the standardized tests (SIMCE) are compulsory: all pupils from every school must take them on any of the test dates.

The Chilean education system, as indicated, permits families to freely choose schools that are either near or far away, as selection is rarely linked to proximity of households and schools, so the whole notion of a catchment area is highly implausible in the actual system.[3] Notwithstanding this, residential location may still be a confounder, because families sort in space, i.e. a more educated family will choose a residential location closer to better and/or more competitive and efficient schools, not to belong to a particular catchment area but because travel cost are reduced. This would represent a downward bias of the estimates, in other words, the effect of spatially determined competition would be understated by OLS. At the same time, house prices will be higher in the proximity of more efficient/more competitive schools for the same reason, namely that better educated families will choose residential locations with better schools. Moreover, different concepts of distance to school that are picked up by municipality FE, the control distance to GEO border and travel time distance (public and private transport travel times), all contribute to endogeneity.

I focus on Santiago, the capital of Chile. Less than 7% of pupils travel in or out of Santiago, suggesting that I can assume that its market is an independent one,

[3]But this is going to change with the on-going reform.

and therefore I can try to measure the spatial nature of this educational market and exploit choice and competition indices. After picking one of the 2-period data with a time gap of six years in the SIMCE data (explained below), my sample contains 79,463 pupils in primary school in 2002 and 52,164 pupils in secondary school in 2008. The large attrition of the sample is due mainly to course repetition; for instance, in 2002, 8.6% of overall pupils repeated the year. If I presume the same rate for each of the six following years, a simple calculation explains the attrition almost in full. As Table 1.1 shows, the sample contains 1,762 primary schools (644 public voucher, 793 private voucher and 325 private) and 839 secondary schools (139 public voucher, 484 private voucher and 216 private). This can also be seen spatially in Figure 1.1, which shows graphically that private voucher schools are more numerous and that private schools are located in the so-called "High Neighbourhood", a popular designation for an elevated and extensive north-eastern residential area in Santiago.[4] Table 1.2, which shows the number of pupils per school, confirms that mixed schools (both primary and secondary schools) and public voucher schools are the biggest. While Table 1.1 shows the number of schools, Table 1.2 shows their size. Multiplying both numbers at the beginning and at the end of the sample reveals that during this period there were more students in private voucher schools than in public voucher schools. Although in 2002, public voucher schools enrolled 54% of pupils while private voucher enrolled just 38%, in 2008 this was reversed, and private schools enrolled 49% while public schools only 44%. Private schools maintained a steady 7%. Additionally, though it is not shown in a table, 15.92% of pupils stay in the same school for primary and secondary education. As Table 1.3 indicates, transferring from a primary public voucher school to a secondary private voucher school and vice versa is quite common, with 17.5% and 10.7%, respectively. Staying in public voucher schools for primary and secondary education accounts for 14.4% of pupils. But the most popular choice (39.7%) is to stay in private voucher schools for primary and secondary education. At the same time, private schools are

[4]In recent years there has been an increase in the development of new subcentres outside the "High Neighbourhood" where more affluent people live, but at the time of the sample, rich people "strictly" lived in the area .

an independent but tiny market.

The standardized and centrally and externally marked test, called the SIMCE, was first administered in 1991, testing the last grade of primary school. Since then, the SIMCE schedule has undergone many changes: from one SIMCE per year in the beginning, it has become an intense set of tests spanning five examinations in different cohorts every year. The Chilean Ministry of Education's Agency for Quality in Education (AQE, 1999-2013), has released data for every SIMCE test since 1999. The data is only made available after a careful analysis of a written research proposal sponsored by a tertiary academic institution, and the use of the data is held under a very strict confidentiality agreement. Single schools or pupils cannot be, for any reason, identified. In principle, pupils are trackable along time. The complete dataset contains fifteen 1-period cross-section data, ten 2-period panel data (with a time gap of two, four, and six years) and one 3-period panel data. As I have indicated, I use one 2-period data with a time gap of six years from 4^{th} grade in primary school in 2002 to 10^{th}grade in secondary school in 2008 for any particular reason.

1.3 Methods

In my main equation I explain the outcome, the SIMCE test (a proxy for the productivity of each school), by a spatially determined measure of competition; individual and school characteristics; individual, school type, level of education and municipality fixed effects. The following is the equation I aim to estimate:

$$\mathbf{Y_{ismt}} = \boldsymbol{\alpha_i} + \boldsymbol{\beta_1 comp_{st}} + \mathbf{X'_{it}}\boldsymbol{\Lambda} + \mathbf{Z'_{st}}\boldsymbol{\Theta} + \boldsymbol{\sigma_p} + \boldsymbol{\varphi_t} + \boldsymbol{\delta_m} + \boldsymbol{\varepsilon_{ismt}} \qquad (1.3.1)$$

Where the subscript **i** stands for individual; **s** for school; **m** for municipality; and

$\mathbf{t}$ for level of education (primary or secondary). $\mathbf{Y_{ismt}}$ represents the standardized test score of pupil $\mathbf{i}$ going to school $\mathbf{s}$ in municipality $\mathbf{m}$ at level of education $\mathbf{t}$. $\mathbf{comp_{st}}$ is competition experienced by school $\mathbf{s}$ at level of education $\mathbf{t}$. $\mathbf{X_{it}}$ are individual characteristics, which are possibly time varying. These include prior achievements (SIMCE or grades), school choice, age, gender, family background, books at home, and disadvantaged groups (family receiving subsidy, Chile Solidario, vulnerability index, integrated, etc.). $\mathbf{Z_{st}}$ are school characteristics, including class size. This category also includes the geographic concentration indices, the competition indices and observable teacher evaluations. $\alpha_{\mathbf{i}}$ is an individual fixed effect which causes problems, being potentially correlated with the error term;[5] $\sigma_{\mathbf{p}}$ are school types: public voucher, private voucher and private; $\varphi_{\mathbf{t}}$ are primary or secondary dummies; and $\delta_{\mathbf{m}}$ are municipality dummies. These control for choice because data is not available at postcode level and some municipalities may have more choice as revealed by the preferences of the pupils. In some specifications, I use choice indices instead of the municipality dummies. Finally, $\varepsilon_{\mathbf{ismt}}$ is an error term picking up possible serial correlation and heteroskedasticity, and unobserved characteristics (other teachers' quality besides observable teachers' evaluation, motivation of neighbours, other local or spatially close neighbouring resources that improve the educational outcome of the pupils).[6] A data generating process with pupils sampled from many schools is likely to be correlated; therefore, it should be clustered at the school level.

In spatial analysis, the researcher can create spatial weights and define spatial lags at the center of spatial analysis. Therefore, inspired by similar indices defined by Gibbons et al. (2006), I define a **choice index** and a **competition index** as follows:

$$\textbf{choice index}_{\mathbf{m}} = \sum_{\mathbf{m=1}}^{\mathbf{M}} \textbf{schools}_{\mathbf{m}} \tag{1.3.2}$$

[5] A correlation between the regressors and the individual fixed effect, unobserved effect or unobserved heterogeneity $\alpha_{\mathbf{i}}$ requires controlling for the fixed effect.

[6] The correlation between the regressors and the idiosyncratic error: $\varepsilon_{\mathbf{ismt}}$ requires instrumental variables.

$$\textbf{competition index}_{\mathbf{s}} = \frac{\mathbf{1}}{\mathbf{I_s}} \sum_{\mathbf{i=1}}^{\mathbf{I_s}} \textbf{choice}_{\mathbf{i}} \tag{1.3.3}$$

The **choice index** for a given municipality **m** is defined as how many schools you can get to from that municipality, i.e. the number of schools which pupils in the given municipality can choose from—in a revealed preference manner—regarding where the schools are located (**M** is equal to the schools inside the given municipality **m**, plus the schools outside the given municipality that the pupils living in that municipality actually choose).[7] Similarly, the **competition index** for a given school **s** is defined as how many alternative choices the students attending that school actually have, i.e. the weighted average of the number of schools each pupil from the given school **s** chooses, i.e. an average of the choice indices of each pupil's municipality for the given school **s** (i.e. for each pupil in a municipality his/her choice index is the same choice index of the municipality), where **i** is the choice index for pupil **i** , $\mathbf{I_s}$ is the number of students in school **s** and the sum is over students within schools.[8] For both indices, I use all pupils of each school, not just the ones that sat for the SIMCE in a particular year. This seems reasonable as distortions could arise doing the contrary because pupils sitting for the SIMCE may not form a representative sample of the pupils of each school. Moreover, schools compete to attract pupils from all grades, not just those tested by the SIMCE, as these grades change yearly.

Figure 1.2 shows the overall competition index for primary and secondary schools. It can be seen that municipalities such as Puente Alto (south-east), La Florida (south-east) and Maipú (south-west) have the schools with the highest competition indices. Gibbons et al. (2006) counted the number of schools that pupils can choose from within the 75th percentile of the home-to-school travel distance, which limited the number of "feasible" schools (i.e. those that can be attended) which in their

[7] For instance, pupils from a given municipality attend 15 schools only situated in the same municipality, so the choice index is 15. Or perhaps students from a given municipality attend a total of 40 schools in six municipalities eventually which may or may not include the given municipality; then the choice index is 40.

[8] For instance, suppose there are three students from the municipality with choice index 40 and two from the municipality of choice index 15. Then the competition index is 30.

case is based, as said, on common travel patterns. In a sense, schools that are very far and attended by few pupils are dropped from the "choice set" (and therefore from the competition index), avoiding very "artificial" large numbers. Since I work at the municipality level, it does not make much sense to use home-to-school distance; instead I use "frequency". For example, for each municipality, the schools are in the "choice set" if at least 5% of the pupils from that municipality chose to attend them. If a school is chosen by only one pupil out of 100 (1%), then I call this an outlier and drop it from the "choice set". Taking this into account, I experimented with different tail dropping techniques, sequentially discarding the lowest 5%, 10%, 20%, 25%, 30% or the lowest 5th, 10th, 20th, 25th and 30th percentiles of the schools according to the revealed preference of the pupils' school choices for each given municipality. Ultimately, I preferred the 25th (lowest) percentile because the "choice set" represents a spatially robust market showing the main and stable schools available in each municipality in repetitive samples. I picked the 25th lowest percentile over the percentage because it gives more flexibility to the distribution by freeing it from being a fixed quantity regardless of the number of chosen schools. Notice that the least chosen schools are systematically located in places of extremely difficult access or are part of hospitals (to benefit children with chronic illnesses who must remain hospitalized) so it seems correct to exclude those schools. There is also a high correlation between the untrimmed and the trimmed version of the ranks. Testing the null hypothesis that both the untrimmed and trimmed version of the ranks are independent, I obtain a rejection of the null meaning that both ranks are closely rank correlated. For example, comparing overall secondary competition with the full sample and the 25th trimming I calculate a Spearman's rho of 0.9230 and a p-value of virtually zero.

As a clarifying example, in a district-shaped educational market, the pupils come only from the same municipality, and accordingly the choice index counts only schools inside that educational district. In fact, it is equal to the total number of schools in the municipality. That, of course, is not the case for the Chilean free-to-

choose market, so both indices are meaningful. At the same time, Figure 1.3 shows that schools have a distribution of competition indices upwardly skewed. Additionally, Figure 1.4 compares the public voucher and private voucher school competition indices for overall schools in 2002 and 2008, showing they are remarkably stable, while private school competition indices are stable with a slight downward trend in the middle of the distribution.

Unfortunately, with the inclusion of competition measures I am also introducing endogeneity to the model. But as I have already said, I have two potent strategies for this endogeneity issue. Firstly, I adopt an instrumental variable strategy; slope (how fast the altitude is increasing) and ruggedness (the terrain undulation) are my instruments, both unlikely to be correlated with the unobservables and individual characteristics. To set the instrumental variable model, I also control by distance to the urban footprint (which I call the GEO border) and elevation. Both controls are proxies for "high class" families (i.e. families with higher income living in elevated areas and close to the GEO border). Secondly, because I observe pupils at two stages, in primary and then again in secondary school, I can implement a value-added model to partial out fixed characteristics of individuals that at the same time affect the school choice and the educational attainment.

1.4 Main results

Descriptive statistics

Table 1.4 provides a statistical summary of the competition indices and test scores. The competition indices remain stable across primary and secondary schools. Specifically, the primary mean for the overall competition index and for the public voucher competition index is the same as the secondary mean at 64 and 16 respectively. The primary and secondary means for the private voucher competition index vary slightly, at 31 and 35 schools respectively. The private competition in-

dex has a slightly higher mean in primary, with 17 versus 12 schools. These results are similar to the results that Figure 1.4 displays graphically. Test scores have a normal distribution around the 50th percentile as expected for a standardized test. Table 1.5 presents the correlation among the competition indices of primary and secondary schools. From this table, it follows that there is a fairly strong correlation in competition between both types of voucher schools and a somewhat weaker negative correlation in competition between the voucher schools and the private schools.

On the other hand, Table 1.6 shows the list of exogenous covariates. These include parents' education, books at home, ethnic background, disabled pupils, sex and age. Parents' education is 11 years in primary and 12 years in secondary; books are 54 units in primary and 62 units in secondary; the male-to-female sex ratio is approximately 1:1; 0.2% of disabled pupils in primary and secondary -when they are called *integrated*-; and pupils are 9.5 years old in primary and 16 years old in secondary.

OLS regressions are contaminated by endogeneity bias because of the standard argument that there is no reasonable proxy for ability, which implies that ability remains in the error term, causing correlation from the competition indices and the error term. For this reason the key identification issue in this paper is to find an appropriate instrument, i.e. one highly correlated with the competition indices and uncorrelated with the error term. After careful theoretical and empirical considerations, I propose the following candidates as instruments:

1. **slope** is the slope of the area centred around each school in Greater Santiago according to the processing by the Slope tool of ArcMap of a DEM provided by the SRTM 90m Digital Elevation Database of the CGIAR Consortium for Spatial Information. Slope is the first derivative of elevation, and a sensible calibration is the use of a Z factor of 0.00002619 according to the data in meters and the Latitude close to 70 as suggested by this ESRI blog (ESRI, 2014), (Figure 1.5b).

2. **ruggedness** is the ruggedness of the terrain calculated as the second derivative of the elevation using the just described DEM, (Figure 1.5c).

In the spirit of Pearl (2009) and Morgan and Winship (2010), the theoretical arguments for selecting slope and ruggedness as valid instruments are explained as follows: slope and ruggedness are negatively correlated with the voucher competition, because the extra cost of building is internalized in the market as fewer schools are built, which drives down the indices. A different outcome is found for private competition, where building a school in an inaccessible place gives the school a reputation linked to the survival of the best fit. Also notice that slope and ruggedness are not necessarily correlated; neither are they correlated with elevation. This is important because I need to use slope or ruggedness as instruments while conditioning on elevation to block other sources of endogeneity. In fact, as shown below, there is some correlation between slope and ruggedness which makes both potentially good instruments, but less between elevation and slope, and very little between elevation and ruggedness to serve as a good conditioning strategy.

Furthermore, these arguments predict something which my empirical findings confirm: the first-stage correlation between the instruments and competition indices, with slope and ruggedness being negatively correlated with overall, public and private voucher competition indices, while being positively correlated with the private competition index. Notice that this correlation not only has the right sign, but is also sufficiently strong, as the first stage F tests show in all the IV tables below.

It could be argued that the instruments are also correlated with the error term (and by this mean correlated with the outcome). A thorough study of the causality model suggests a straightforward way of blocking the other sources of endogeneity by conditioning on dwellers' income, elevation or distance to GEO border to estimate the causal effect of competition on test scores using slope and ruggedness as valid instruments. This is a typical reverse causality relation. Instead of the market structure causing the academic outcome, it is the academic outcome causing the

market structure. In other words, instead of more competition causing more efficient schools, more efficient schools causes a monopolistic structure. Pupils just choose this school, then the competition index is at the minimum. This is the endogeneity of the school market. The key controls needed to identify the causal effect of the competition indices on test scores are as follows:

1. **distance to GEO border** is the distance from each school to the GEO border or border of the urban footprint of Greater Santiago.

2. **elevation** is the elevation of each school in the Greater Santiago area according to a DEM provided by the SRTM 90m Digital Elevation Database of the CGIAR Consortium for Spatial Information, (Figure 1.5a).

These controls are good controls because they are proxies of "high class" and families with higher income live in elevated areas and close to the GEO border.

Figure 1.5 shows the orographic details of slope, ruggedness and elevation. Elevation (Figure 1.5a) is defined by waves of east-west platforms in successively lower altitudes; slope (Figure 1.5b) is marked by higher activity east of the city and by a lower, but still intense, activity in the west; and ruggedness (Figure 1.5c) is characterized by intense activity all over the surface of Greater Santiago, with high activity not just in the east of the city but also in the west. Table 1.7 shows the summary statistics for controls and instruments. The table provides evidence that the mean distance to the GEO border increases from 3.1 km in primary to 3.9 km in secondary school, while the elevation of the school only increases from 577 m to 582 m. Considering the instruments, the mean slope is 2 degrees but with a minimum just over 0 degrees and a maximum of 19-35 degrees; finally, ruggedness is close in mean to 0.5 degrees with a range between almost 0 to 15-12 degrees. Table 1.8 presents the correlation between controls and instruments. Broadly speaking, the correlations are modest; it is reassuring that information is not being lost in the process. Specifically, the correlation within controls is just over 0, within instruments 0.6-0.4 and between controls and instruments -0.1 (distance to GEO border) and 0.2-0.5

(elevation).

Finally, while the causation of slope and ruggedness (i.e. the instruments) to the competition indices is mediated by the building costs as explained above, the causation of elevation and distance to the GEO border (i.e. the control variables) to the test scores is mediated not by construction costs but mainly by environmental amenities (e.g. property rights over clean air and a pollution-free environment) which cause high dwellers' income and more educated family background to be localized in the high areas of the city and/or in the shortest distances to the GEO border. The fact that instruments and controls have different underlying processes is the theoretical explanation of why they are weakly correlated and why the conditioned instruments are excluded and do not cause the test scores other than via competition.

Balancing tests

I have already shown the power of slope and ruggedness by demonstrating their correlation with the competition index. It is now time to defend exogeneity, i.e. to show that these instruments are not correlated with the unobservables. To do so I perform balancing tests using related variables not included in my controls, aiming to test the exogeneity of the instruments. From the surveys conducted of the pupils' cohort I take several related variables aggregated at the school-level that presumably are part of the unobservables which are not outcomes and are pre-determined. I regress them on slope and on ruggedness with controls clustered at the school level. The list of tested variables starts with the average and the standard deviation of the SIMCE for the exact previous cohort, reasons for choosing the school (closeness, good teachers, low cost), whether the household has a personal computer and internet, the total income of the household, the co-payment to the school, other school-related expenses and preschool background. Each tested variable is aggregated at the school-level. These balancing tests permit me to provide a plausible check that the instruments are exogenous with respect to the similar and correlated variables which are subsumed in the error term. These tested variables remain in

the background and are a credible sample of what constitutes the unobservables, to which the instruments are shown to be strictly orthogonal. The results are presented in Table 1.9 and neatly support the claim of exogeneity. None of the independent 19 regressions shows a significant coefficient for the respective instrument. In other words, the correlation between the instruments and the unobservable is likely to be non-existent or weak at most.

Results

The first and foremost result is that the coefficient for "aggregated" school competition (i.e. overall competition index) is non-negative in Table 1.10 (0.060 without controls and 0.053 with controls in primary and 0.100 without controls and 0.108 with controls in secondary) but negative in Table 1.11 (-0.223 without controls and -0.231 with controls in primary and -2.168 without controls and -2.390 with controls in secondary) using slope as the preferred instrument. The reversion of the coefficient indicates that there was a positive selection bias, i.e. pupils from more competitive schools also have more ability and for this reason perform better in test scores. This is a plausible story, because pupils naturally gifted for study look for more competitive schools. But they are not better off: they perform better only because of this extra ability, not by any value-added from the school. This effect is very interesting because it is spatially similar to the literature finding that controlling by peer effects, socio-economic background and non spatial selection bias either dissolves the difference between private voucher and public voucher schools entirely or reduces it to existing only for selected groups. See, for instance, Hoxby (2002), Howell et al. (2001). As presented in Table 1.3 the voucher, i.e. public and private voucher schools, and the private are two separate markets. My results in Table 1.10 confirm this fact, showing different coefficients for both type of schools (for instance, 0.065 v/s 1.281** in secondary schools with controls). As my main interest is to scrutinize the causal effect of spatially determined competition in the voucher system, the results I present subsequently are for voucher schools. In short, I am testing the voucher system. Table 1.10 shows that voucher competition also

causes non-negative reactions in voucher schools in primary schools and in voucher schools in secondary schools. But this is fully permeated by endogeneity. Tables 1.12 and 1.13, for example, show different findings. These two tables display the results of the corrected instrumental variable model, the former table with a choice index to measure and control for choice variation and the latter with a municipality fixed effect to achieve the same. The real effect is mainly marginally significant for primary and for secondary voucher pupils when using the choice index and slope as instrument and not significant in all other specifications. The instrumental variable approach proves all its effectiveness and discloses the underlying causal effect.

1.5 Heterogeneity

In an effort to identify the groups that are in a better position based on the spatially determined competition, I repeat the main regression equation: voucher schools with OLS and IV, by family background and school characteristics. From Table 1.14 it seems that pupils with families having more income, parents with more years spent in school, a personal computer and internet access at home are better off, i.e. the spatially determined competition in the voucher system increases their performance and school achievements. Books at home seems not to factor into the results. From Table 1.15, it seems also that smaller schools and schools with higher achievements in standardized tests and with a median dispersion that is not too heterogeneous or homogeneous in pupil performance[9] are more incentive-driven.

There are two sources of heterogeneity: Competition heterogeneity, or so-called "cause" heterogeneity or test results/pupils heterogeneity, and "reaction" heterogeneity. In reaction heterogeneity, for example, while private schools better serve pupils with an average primary SIMCE percentile of 75th in the $4^{\mathbf{th}}$ grade, private

[9]This result could be of policy interest because a system that increases or reduces the dispersion in the composition of schools could imply large improvements or detriments in school educational productivity. According to study results, it is best to have an inclusive system (to assure equality) but allowing for some schools to select heavily to create centers of excellence, i.e. emblematic lycées should be allowed to fully select their students.

voucher schools serve pupils with an average SIMCE percentile of 50th and public voucher schools serve pupils with an average SIMCE percentile of only 40th. These scores indicate that teaching a pupil in an average public voucher school is more demanding than teaching a pupil in an average private school. The model accounts for these differential costs in teaching by including choice index, municipality and school type fixed effects.

1.6 Discussion

Table 1.10 shows an OLS regression in which more competition implies non-negative test scores for both primary and secondary schools. In Table 1.12, after using IV regressions, the non-negative test score effects reverse into negative.[10] For secondary pupils, the effect is equal to 0 in all the specifications, which are negative but not significant. These are the main results of the paper. How do these results compare with the effect of other determinants of school productivity? One additional standard deviation in teacher quality causally implies an increase of 11% of a standard deviation in pupil performance as many independent assessments have determined; see, for example, Hanushek (2003). So in the worst case, when I use only slope as an instrument, an increase in one standard deviation in teacher quality cancels out with approximately one extra standard deviation in spatially determined competition as it has been understood and implemented in the Chilean educational voucher system in the last few decades for primary school. Again, in every other specification for primary and in all specifications for secondary pupils, the causal effect of spatially determined competition is close to 0, so an increase in teacher quality is not exhausted by an increase or by a decrease in spatially determined competition.

At the same time, the space aggregation process implied by the choice and the

[10] Using slope as instrument suggests a relatively small congestion effect at least for primary school competition. A one standard deviation (21 schools) of the voucher competition index represents a decrease of -4.24 (=-0.202X21) percentiles in test scores, which corresponds to a decrease in -15% (=-4.24/27.83) of a standard deviation (27.83 percentiles) of voucher test results/pupils in primary school for 2002 SIMCE with controls.

competition indices mainly reflects, as has been argued up till now, spatially determined competition. But one could counter that other concepts could also be picked up by this grouping method. For example, it is not easy to rule out the idea that peer effects also drive the indices in the study. So why is spatial competition behind my aggregators? Only the fact that I am aggregating at the school-level, i.e. I use the municipality of the spatial residence of the school, drives the spatial competition interpretation. Aggregation at the student-level, i.e. using the municipality of spatial residence of the student, would tip the scales in favour of a peer effect interpretation. One way to overcome this issue is to stress the significance of controlling for pupil characteristics. In fact, I control for all main student-level variables which are not outcomes in order to avoid bad controls. This approach strengthens the likelihood of my indices picking up spatially determined competition rather than peer effects.

Recently, it has been argued by government officials in Chile that admission selection should disappear completely. This chapter may indeed evaluate this public policy using the evidence gathered and suggest some ideal policy on student mix ratio. First, suppose that there are two types of students: G and B (good and bad). Then also suppose that a G and G interaction increases each test score performance and a B and B interaction decreases each test score performance. Let me finally suppose than a G and B interaction could have either a positive or negative value. Then a certain admission selection policy should be enforced if C-(A-B) is maximized for some mix ratio (if C-(A-B)$<$0, then take -[C-(A-B)]). In other words you calculate C-(A-B) for all mix ratios and take the maximum. Then the argmax is the mix ratio that you should use. Corner solutions are likely not to be desirable solutions. So this chapter suggests that the actual proposed admission selection policy of null percentage for schools in Chile is highly inappropriate.

1.7 Conclusion

This analysis has been one of the first of hopefully many more thorough spatial analyses on the Chilean voucher system. The main spatial variables were the competition indices as determined by the weighted average of schools a typical pupil can choose from based on his or her preferences; the indices were built using a revealed preference strategy. These indices were used in OLS, IV and Changes (see Appendix A) regressions. The spatial analysis gives evidence that helps to answer the fundamental questions on the functioning of the Santiago voucher system. As the data show, for more than a decade a spatial process has been repeating itself over and over: first, a school performs badly; their pupils get a poor quality education; this education translates into poor standardized test results; families that are aware of these results reduce their choice of this school; attendance decreases; the school receives fewer vouchers; the school has fewer resources; the quality of the school's education further declines; finally, the school closes causing a void of space and less competition amongst the surrounding schools. The turnover can be dramatic in size and place. As just described, the voucher system has worked in Chile, producing big changes in the market structure of schools. But as the results here make apparent, these deep structural changes in the market have not brought performance gains in education.

These results suggest that the voucher system is not working because more spatially determined competition does not imply more value-added to the education production functions of the different school types. Only the schools clustered in high-high regions, such as private schools, are more productive; their pupils perform better, and the change in value-added performance is higher, but such schools seem to be in a separate market, with no interaction with the voucher schools, and after controlling for pupils' heterogeneity, not even a particularly successful one. In short, *the voucher system needs to be amended.* The advanced econometric techniques employed in this chapter have yielded evidence of no spatial effects, which means that

increased competition among schools does not increase the value-added.

But is it that the voucher system is not working? The fact that schools at the bottom consistently close should be good overall. But the private voucher systems are not getting better. As Schumpeter (1934) posited, there is competition *for* (market share) and *in* the market (efficiency/productivity). In Chile, the competition *for* the market works due to the voucher system. It is the competition *in* the market which is failing. Why? Perhaps the standards for new schools are low and represent not a better influx of production is replacing the lowest value-added schools. Or think of a model where parents take time to learn whether the school is "poorly performing". In this case bad schools could survive before being expelled from the system. Also, consider another model in which there is no link between the *for* and *in* aspects of competition, in which the production function of the worst performing schools is not linked with the overall system performance. This seems odd, but could still be explained in the case of Chile by the big segmentation of the market; bad schools are squeezed out of the market but their students move to other similar bad schools and never climb the market to well-performing private schools. In other words, poor students are captive in a vicious cycle. Often, for this group of students, the schools that remain are not better than the schools that have exited the market.

One key mechanism that could be causing the poor functioning of the voucher system is the voucher's value of £40 per month for primary schools and £50 per month for secondary schools. The costs of an excellent education are believed to be three times these numbers.[11] The fact that the economic incentive, i.e. spatially determined competition, is not aligned with the economic outcomes, i.e. test scores, does not preclude the voucher system from working consistently if higher demand side vouchers were guaranteed. It is well documented in the health sector that if providers can vary price and quality, they try to increase their market shares by

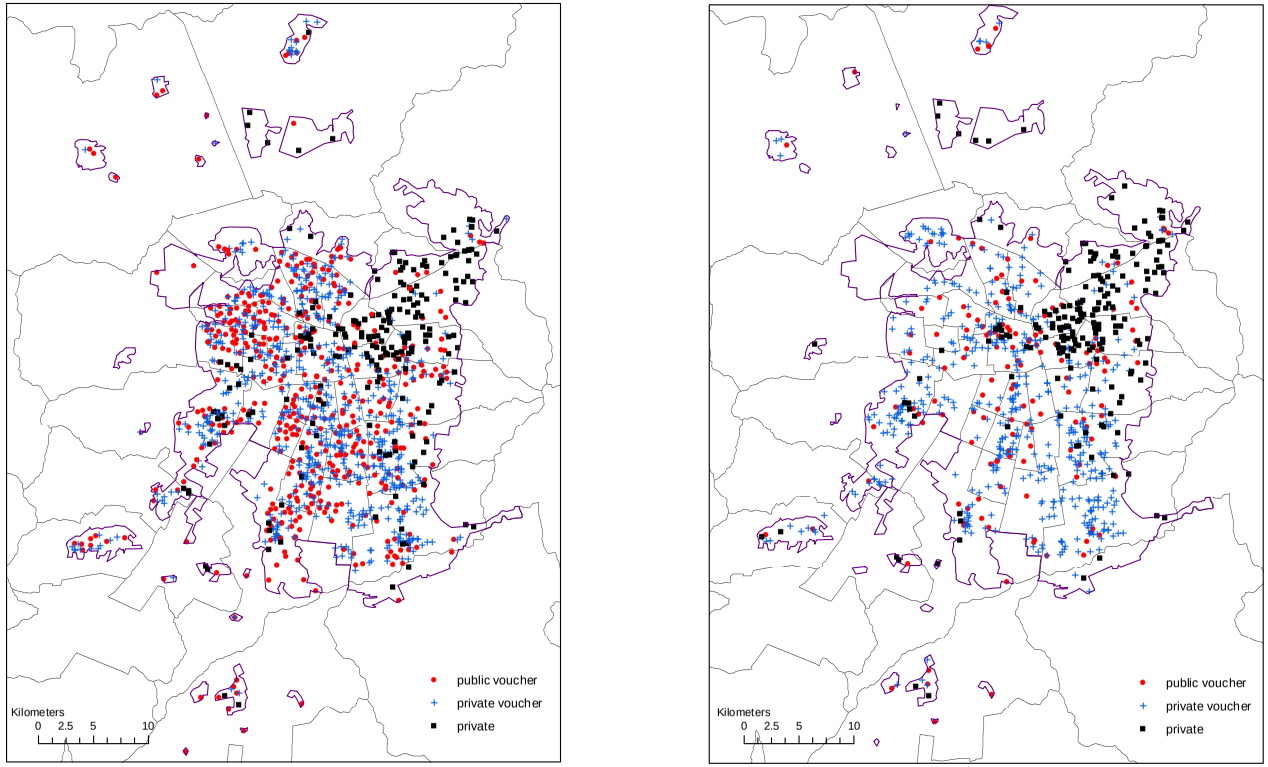
[11] From a personal conversation with a renowned education expert at PUC during my time as a postgraduate pupil in Economics there.

lowering the price, not by improving the quality of the service; this trend has been evidenced in the literature on the NHS internal market reforms of the 1980s and 1990s. See, for example, Propper et al. (2008). Conversely, when the price is fixed, as is the case with the state schools in England, the NHS Payment by Results and Medicare in the US, competition is believed to translate into quality improvements; see Gaynor (2004) for an insightful theoretical explanation. However, in the Chilean educational system, the voucher is extremely low. Some improvements have come from the so-called SEP reform, a law passed to increase the number of the vouchers given to prioritized (very poor and/or disabled) pupils. In fact, for the year 2005—in the middle of my sample—the amount of annual expenditure per student was $1,930, while in the UK it was three times more $6,888 (USD PPP, OECD (2008)).[12] Then, raising the voucher could be one possible way of solving the paradox of more competition not leading to better results because, at least theoretically, such a voucher system seems a reasonable policy.

As mentioned previously, this analysis is one of the first spatial studies of this long-term object of interest. The hypotheses tested are of key importance for policy makers. New authorities in Chile are pushing toward largely uncertain changes in voucher systems. These findings (in particular the negative effect of competition when taking into account the endogeneity problem) can offer guidance as to which changes are most appropriate.

[12]New reforms will increase voucher in 20% in $1^{\mathbf{st}}$-$6^{\mathbf{th}}$ and $9^{\mathbf{th}}$-$12^{\mathbf{th}}$ grades and 80% in $7^{\mathbf{th}}$-$8^{\mathbf{th}}$ grades.

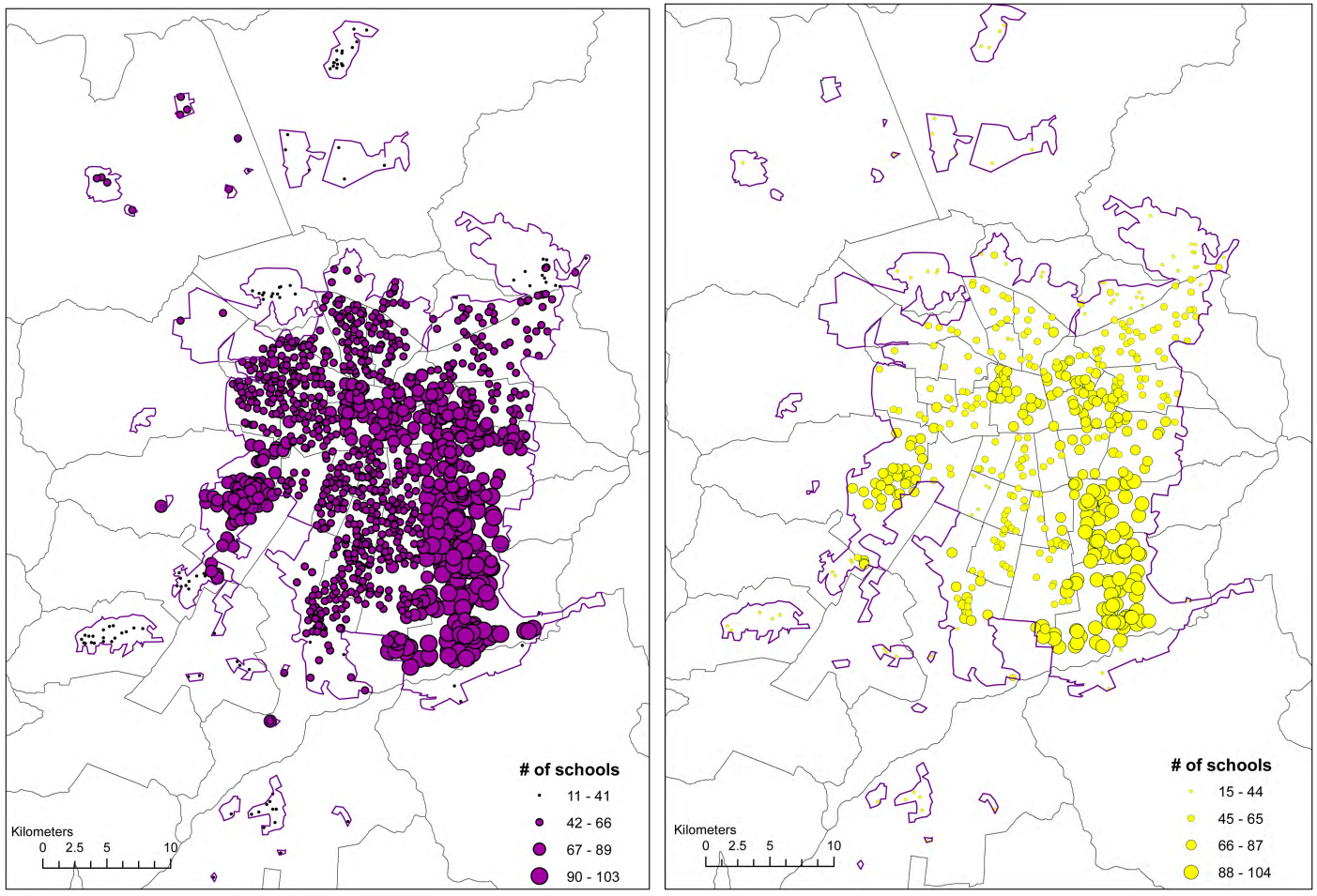
Figure 1.1: Schools by school type



(a) Primary schools by school type

(b) Secondary schools by school type

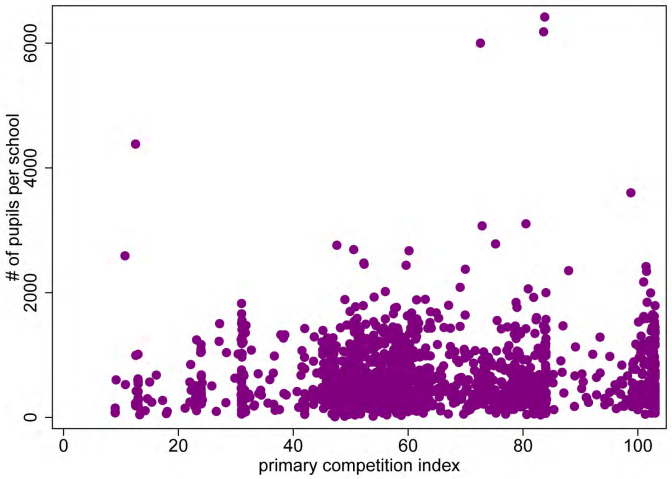
Figure 1.2: Overall competition indices



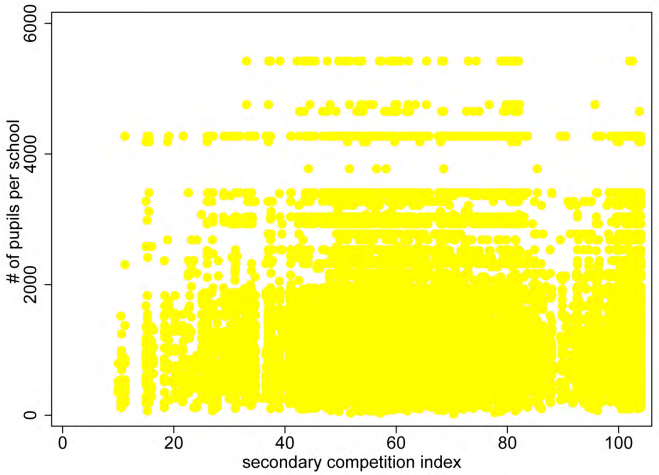
(a) Primary school overall competition index

(b) Secondary school overall competition index

Figure 1.3: Overall competition indices v/s number of pupils per school

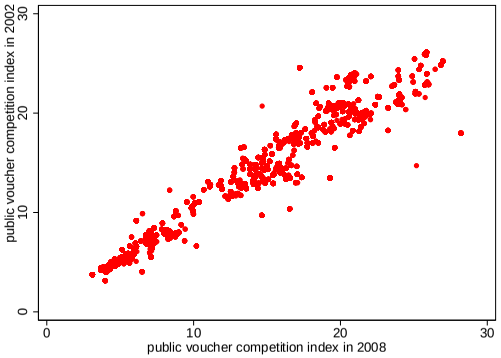


(a) Primary comp index v/s # pupils/sch

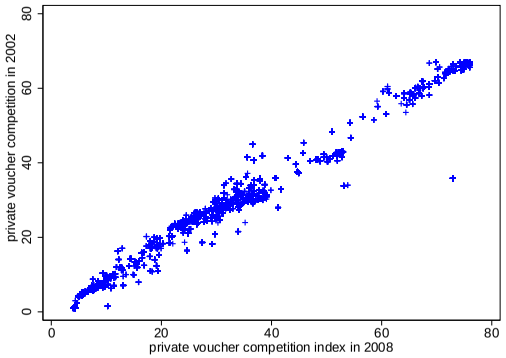


(b) Secondary comp index v/s # pupils/sch

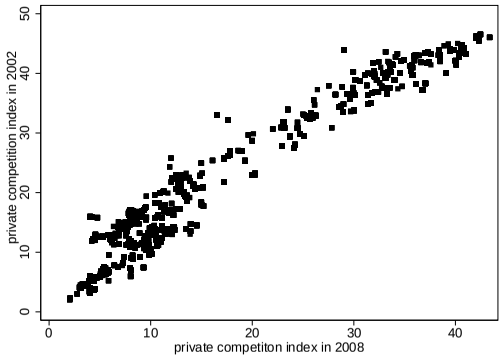
Figure 1.4: Stability of competition indices by school type, 2002 v/s 2008



(a) Public voucher competition index

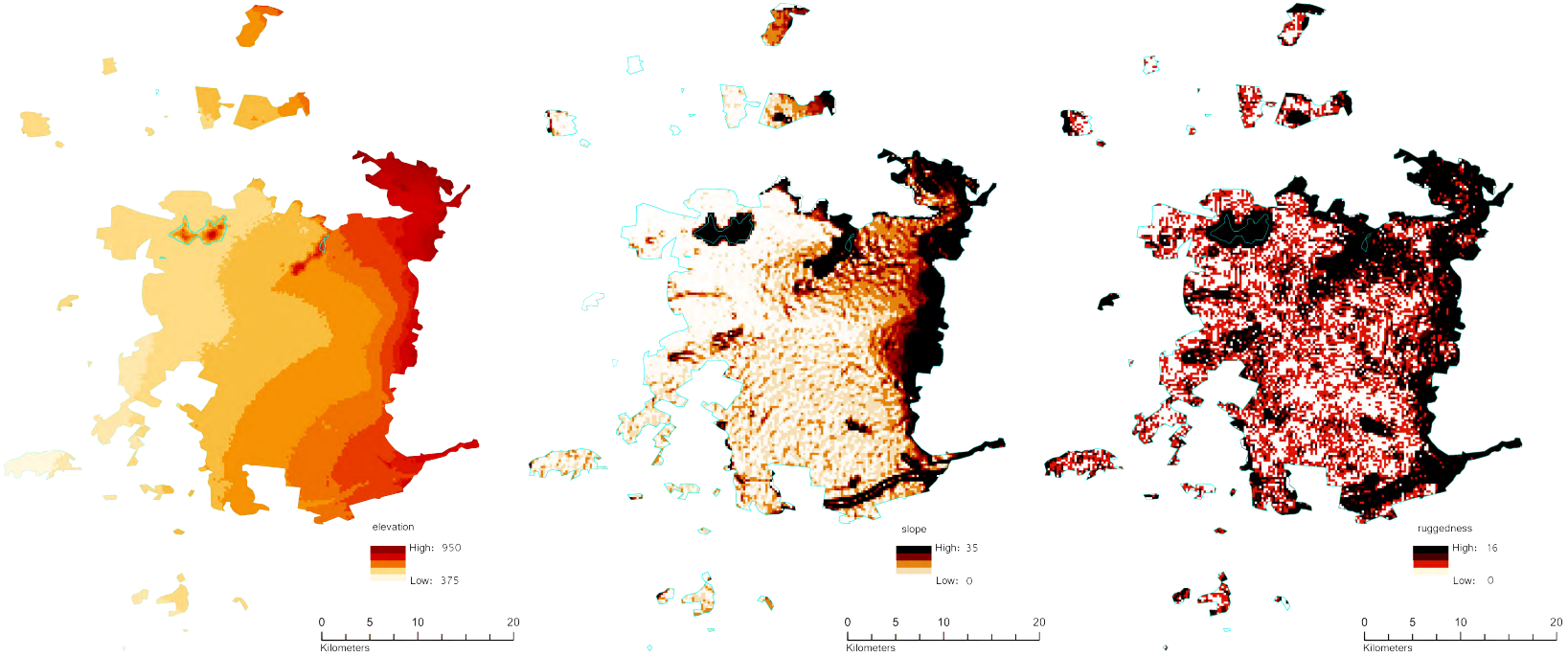


(b) Private voucher competition indices



(c) Private competition indices

Figure 1.5: Control and instruments



(a) Elevation (b) Slope (c) Ruggedness

Table 1.1: Number of schools by school type

	primary	secondary
public voucher	644	139
private voucher	793	484
private	325	216
Total	1,762	839

Table 1.2: Mean number of pupils by school type

	primary	secondary	primary & secondary
public voucher	428	626	932
private voucher	362	367	819
private	101	246	592

Table 1.3: Switching from primary to secondary school by school type, (percentage)

from\to	pub vou	pri vou	pri
pub vou	14.4	17.5	0.1
pri vou	10.7	39.7	0.8
pri	0.8	3.3	12.2

Table 1.4: Summary statistics of competition indices and test scores

variable	mean	std. dev.	min.	max.	n
competition indices					
primary					
overall competition index	64.832	21.654	9	103	79,463
public voucher competition index	16.32	5.954	2.988	28.201	79,463
private voucher competition index	31.342	16.603	0.968	67	79,463
private competition index	17.17	10.062	2	46.526	79,463
secondary					
overall competition index	64.033	21.35	10	104	52,164
public voucher competition index	16.517	6.157	3.11	30	52,164
private voucher competition index	35.002	18.935	4	76	52,164
private competition index	12.514	9.239	2	43.363	52,164
test scores					
SIMCE primary test scores	50.118	28.831	1	100	79,463
SIMCE secondary test scores	50.214	28.895	1	100	52,164

Table 1.5: Cross-correlation of school competition indices

variables		(1)	(2)	(3)	(4)
primary					
overall competition index	(1)	1.000			
public voucher competition index	(2)	0.648	1.000		
private voucher competition index	(3)	0.891	0.658	1.000	
private competition index	(4)	0.295	-0.283	-0.122	1.000
secondary					
overall competition index	(1)	1.000			
public voucher competition index	(2)	0.649	1.000		
private voucher competition index	(3)	0.922	0.641	1.000	
private competition index	(4)	-0.012	-0.481	-0.345	1.000

Table 1.6: Summary statistics of the full set of exogenous covariates

variable	mean	std. dev.	min.	max.	n
primary					
father's education	11.73	3.938	0	28	79,463
mother's education	11.461	3.828	0	28	79,463
books	54.125	67.505	0	250	79,463
$>$ 100 books	0.162	0.368	0	1	79,463
disabled	0.002	0.04	0	1	79,463
male	0.499	0.5	0	1	79,463
age in years	9.517	0.522	6.045	14.773	79,463
secondary					
father's education	12.515	3.708	1	23	52,164
mother's education	12.28	3.555	1	23	52,164
books	62.815	62.389	0	180	52,164
$>$ 100 books	0.198	0.399	0	1	52,164
Indigenous father	0.047	0.212	0	1	52,164
Indigenous mother	0.028	0.165	0	1	52,164
integrated	0.002	0.042	0	1	52,164
male	0.48	0.5	0	1	52,164
age in years	16.041	0.403	10.303	20.167	52,164

Table 1.7: Summary statistics of controls and instruments

variable	mean	std. dev.	min.	max.	n
primary					
distance to GEO border	3,177.77	2,252.561	9.746	9,493.505	79,463
elevation	577.659	85.132	376	952	79,463
slope	2.256	2.453	0.18	35.315	79,463
ruggedness	0.474	1.159	0.004	15.526	79,463
secondary					
distance to GEO border	3,934.351	2,527.029	20.744	9,404.745	52,164
elevation	582.093	81.988	375	942	52,164
slope	2.367	2.073	0.127	19.473	52,164
ruggedness	0.493	1.093	0.005	12.848	52,164

[1] Distance to GEO border in meters, elevation in meters above sea level, slope and ruggedness in arc degrees of inclination to the horizontal over meters (°/m) and sq meters (°/m^2).

Table 1.8: Cross-correlation of primary school controls and instruments

variables		(1)	(2)	(3)	(4)
primary					
dist to GEO border	(1)	1.000			
elevation	(2)	0.057	1.000		
slope	(3)	-0.157	0.49	1.000	
ruggedness	(4)	-0.151	0.229	0.652	1.000
secondary					
dist to GEO border	(1)	1.000			
elevation	(2)	-0.094	1.000		
slope	(3)	-0.143	0.542	1.000	
ruggedness	(4)	-0.124	0.248	0.491	1.000

Table 1.9: Balancing test

	primary		secondary	
Selected test variables	slope	ruggedness	slope	ruggedness
at a school-level	with controls	with controls	with controls	with controls
SIMCE average,	-0.004	-0.004	0.005	-0.000
(previous cohort)	(0.004)	(0.003)	(0.004)	(0.003)
	[0.409]	[0.206]	[0.221]	[0.952]
SIMCE std. dev.,	0.026	0.004	-0.009	-0.008
(previous cohort)	(0.019)	(0.012)	(0.014)	(0.010)
	[0.175]	[0.746]	[0.550]	[-0.76]
choice,	0.566	0.299	0.394	-0.005
closeness	(0.396)	(0.212)	(0.308)	(0.237)
	[0.153]	[0.160]	[0.202]	[0.983]
choice,	-0.350	-0.136	-0.109	-0.033
good teachers	(0.350)	(0.209)	(0.231)	(0.140)
	[0.318]	[0.517]	[0.638]	[0.813]
choice,	0.181	0.084	-0.244	-0.335
low cost	(0.229)	(0.145)	(0.350)	(0.217)
	[0.430]	[0.561]	[0.487]	[0.124]
pc at home	-0.353	-0.226	-0.023	-0.226
	(0.299)	(0.173)	(0.345)	(0.203)
	[0.237]	[0.191]	[0.946]	[0.265]
internet at home	-0.416	-0.276	0.220	-0.089
	(0.414)	(0.213)	(0.264)	(0.186)
	[0.316]	[0.194]	[0.404]	[0.630]
total income	-0.000	-0.000	0.000	-0.000
of households	(0.000)	(0.000)	(0.000)	(0.000)
	[0.262]	[0.361]	[0.320]	[0.678]
co-payment	-0.004	-0.002		
	(0.003)	(0.002)		
	[0.169]	[0.253]		
other school	-0.007	-0.002		
related expenditures	(0.008)	(0.005)		
	[0.421]	[0.682]		
preschool	-0.004	-0.017		
	(0.342)	(0.202)		
	[0.991]	[0.931]		

This table shows the balancing test between the instruments for the competition measure—slope and ruggedness—and selected test variables thought to belong to the unobservable for primary and secondary schools and which are not used as controls in any preferred specification: previous cohorts' SIMCE (level and std dev), main reason for school choice (closeness, good teachers and low cost), IT at home, total income of HHs, co-payment, other school related expenditures and preschool. The tests make several separate regressions of each selected variable averaged at the school level on each instruments and controls (distance, elevation and municipality fixed effect), clustered at the school level. Not all variables are available for secondary schools. *** p<0.1 Std errs in round parenthesis, p-values in square brackets.**

Table 1.10: Effect of competition on test results with/out controls. OLS

	primary		**secondary**	
	overall pupil test results			
overall competition index	0.060	0.053	0.100	0.108
	(0.070)	(0.063)	(0.111)	(0.099)
	voucher and private pupil test results			
voucher competition index	0.070	0.051	0.058	0.065
	(0.101)	(0.092)	(0.132)	(0.117)
private competition index	0.278	0.305	1.435**	1.281**
	(0.306)	(0.286)	(0.621)	(0.543)
	voucher pupil test results			
voucher competition index	0.073	0.054	0.014	0.006
	(0.101)	(0.092)	(0.130)	(0.115)
controls				
distance to GEO border	Yes	Yes	Yes	Yes
elevation	Yes	Yes	Yes	Yes
list of covariates	No	Yes	No	Yes
municipality FE	Yes	Yes	Yes	Yes

Source: Author's calculation from research data provided by the Chilean Ministry of Education's Agency for Quality in Education. This table shows the effect of "aggregated" overall competition on "aggregated" educational test scores for pupils who sat for the SIMCE Language and Math Tests (unified in one ranking) in the $4^{\mathbf{th}}$ grade of primary school in 2002 and for the SIMCE Language and Math Tests (unified in one ranking) six years later in the $10^{\mathbf{th}}$ grade of secondary school in 2008 for all schools located in Greater Santiago. It also shows the effect of "disaggregated" in voucher and private competition to voucher and private pupil test results. Finally, it shows the effect of voucher competition on voucher pupil test results. The table shows the OLS regressions in levels for 2002 and 2008 regarding the effect of "aggregated" overall competition indices (overall is unified public voucher, private voucher and private, not disaggregated), "disaggregated" voucher and private competition and voucher competition on overall, voucher and private pupil test scores without and with controls. Controls include municipality fixed effect, distance to GEO border, elevation and the full set of exogenous covariates described in Table 1.6. Overall, voucher and private competition indices are the number of schools an average pupil from a given overall, voucher and private school can choose. They are a by-municipality weighted average of the choice indices that measures the revealed preference number of each of the three school types that pupils from a given municipality are effectively choosing. *** p<0.1, ** p<0.05, *** p<0.01**

Table 1.11: Effect of "aggregated" competition on "aggregated" test results with/out controls. Instrumental variables

	primary		**secondary**	
instrument:	overall pupil test results			
slope				
overall competition index	-0.223	-0.231	-2.168	-2.390
tests				
First stage coeff.	-0.379***	-0.366***	-0.397***	-0.388***
First stage F t.(F-s r.)	16.29	15.03	13.64	12.26
Kleibergen-Paap rk LM s.(Ui)	22.96	20.95	16.66	15.60
controls				
distance to GEO border	Yes	Yes	Yes	Yes
elevation	Yes	Yes	Yes	Yes
list of covariates	No	Yes	No	Yes
municipality FE	Yes	Yes	Yes	Yes

[1] **Source:** Author's calculation from research data provided by the Chilean Ministry of Education's Agency for Quality in Education. The data is available only to researchers after submitting a written proposal.

[2] Instrument slope. First stage with F test. Choice index measuring choice variation.

[3] s.=statistics, t.=test, F-s r.=First-stage relevance, Ui=Under identification test.

Table 1.12: Effects of "voucher" competition on "voucher" test results with/out controls. IV. Choice Index

	primary		**secondary**	
instrument:	voucher test results			
slope				
voucher competition index	-0.223*	-0.202*	-0.686**	-0.558
tests				
First stage coeff.	-1.466***	-1.423***	-1.479***	-1.333***
First stage F t.(F-s r.)	20.42	18.92	23.80	22.12
Kleibergen-Paap rk LM s.(Ui)	54.62	51.48	19.52	20.05
ruggedness				
voucher competition index	-0.323	-0.286	-1.259	-1.002
tests				
First stage coeff.	-1.35***	-1.321***	-1.225**	-1.144**
First stage F t.(F-s r.)	14.57	13.90	11.47	11.21
Kleibergen-Paap rk LM s.(Ui)	20.10	19.22	11.56	12.19
controls				
distance to GEO border	Yes	Yes	Yes	Yes
elevation	Yes	Yes	Yes	Yes
list of covariates	No	Yes	No	Yes
choice index	Yes	Yes	Yes	Yes

[1] **Source:** Author's calculation from research data provided by the Chilean Ministry of Education's Agency for Quality in Education. The data is available only to researchers after submitting a written proposal.
[2] Instrument slope, ruggedness and both. First stage with F test. Choice index measuring choice variation.
[3] s.=statistics, t.=test, F-s r.=First-stage relevance, Ui=Under identification test.

Table 1.13: Effects of "voucher" competition on "voucher" test results with/out controls. IV. Municipality fixed effects

	primary		**secondary**	
instrument:	voucher test results			
slope				
voucher competition index	-0.140	-0.183	-1.670	-1.541
tests				
First stage coeff.	-0.193**	-0.191**	-0.474***	-0.464***
First stage F t.(F-s r.)	10.88	12.68	13.44	13.44
Kleibergen-Paap rk LM s.(Ui)	11.53	11.36	13.01	13.35
ruggedness				
voucher competition index	1.773	1.728	-8.91	-6.88
tests				
First stage coeff.	-0.338***	-0.334***	-1.003***	-1.000***
First stage F t.(F-s r.)	9.64	9.61	34.64	36.34
Kleibergen-Paap rk LM s.(Ui)	10.45	10.15	28.87	30.75
controls				
distance to GEO border	Yes	Yes	Yes	Yes
elevation	Yes	Yes	Yes	Yes
list of covariates	No	Yes	No	Yes
municipality FE	Yes	Yes	Yes	Yes

[1] **Source:** Author's calculation from research data provided by the Chilean Ministry of Education's Agency for Quality in Education. The data is available only to researchers after submitting a written proposal.
[2] Instrument slope, ruggedness and both. First stage with F test. Municipality fixed effect measuring choice variation.
[3] s.=statistics, t.=test, F-s r.=First-stage relevance, Ui=Under identification test.

Table 1.14: Heterogeneity (family background)

variable	primary		secondary	
	OLS	IV	OLS	IV
Baseline	0.054	-0.183	0.006	-1.541
income pcm				
Above 95% (≈ £2,000-£5,000)	0.203	0.138	0.058*	-0.321
Above median (≈ £200- £350)	0.089	-0.132	0.039**	-0.98
Below median	0.001	-0.179*	0.023	-1.05
Below 5% (≈ £50- £65)	-0.025	-0.184*	0.001	-1.45**
average years of school				
Above 95% (17 years)	0.110	-0.140	0.03*	-0.856
Above median (12-12.5 years)	0.069	-0.201	0.027	-0.844
Below median	0.023	-0.230	0.018	-1.09
Below 5% (6-7.5 years)	-0.064	-0.349**	-0.019	-1.038*
books at home				
Above 95% (180-250)	0.194	-0.464	0.022	-0.71
Above median (35-40)	0.099	-0.506	0.018	-0.893
Below median	0.041	-0.749	0.015	-1.41
Below 5% (2-3)	-0.161	-0.825	-0.007	-1.91
pc at home				
Yes	0.049	-0.121	0.007	-0.919
No	0.015	-0.219	0.01	-1.1
internet at home				
Yes	0.092	-0.155	0.036	-0.84
No	0.051	-0.193	0.021	-1.139

[1] Effect of competition on test scores by family background: Income, average years of school, books at home, pc at home, internet at home. Baseline has OLS and IV (with slope and ruggedness as instruments) results for voucher school test scores on voucher school competition.

Table 1.15: Heterogeneity (school characteristics)

variable	**primary**		**secondary**	
	OLS	IV	OLS	IV
Baseline	0.054	-0.183	0.006	-1.541
mean number of pupils				
Above 95% (2,172-3,037)	-3.869***	-1.93	0.013	-2.336
Above median (829-1,085)	-0.087	-0.389***	0.032*	-1.868
Below median	0.132	-0.137	0.051*	-0.163
Below 5% (248-389)	-0.170	-1.97***	0.081**	-0.215
SIMCE average				
Above 95% (80.84-85.04)	0.207***	0.139***	0.060***	0.998***
Above median (48.95-47.25)	0.032	-0.045	0.041***	-0.071
Below median	-0.011	-0.06	0.03**	-0.047
Below 5% (24.33-19.59)	-0.198***	-0.087	0.006	-0.075
SIMCE standard deviation				
Above 95% (27.83-28.3)	-0.671*	-0.29	0.094	-1.262
Above median (23.5-20.52)	-0.02	-0.224	0.014	-0.153
Below median	-0.02	-0.248	0.054***	-2.3
Below 5% (15.77-12.23)	0.358	-0.074	-0.136	-9.679

[1] Effect of competition on test scores by school characteristics: Size of school (number of pupils), average SIMCE, standard deviation of SIMCE. Baseline has OLS and IV (with slope and ruggedness as instruments) results for voucher school test scores on voucher school competition.

Appendix

A. Changes model: A robustness check

Note that regression in changes is also a suitable model for dealing with endogenous regressors. I construct the competition indices in changes as the difference of each secondary competition index and the respective primary competition index; see Table A.1. Notice that while the private competition index has decreased by five percentiles, the private voucher index has risen by four percentile points and the public voucher competition has not presented any major variation. Notice that the distance to the GEO border, elevation, the school type FE, and the municipality FE drop in the 'in changes' model. The list of covariates for the model in changes is given in Table A.2. Households on average have more parental education,[13] books and more than a hundred books, personal computers and internet access. At the same time, schools are bigger.

Table A.3 shows the results from a model in changes that has the same negative but not significant effect on the value-added to all and to voucher schools. Note the similarities with the previous IV results. Table A.4 shows the heterogeneity regression where the voucher system is more incentive-driven for more affluent families, those with higher parent average years at school, personal computers and internet access at home, smaller schools, schools with a higher SIMCE average and near median dispersion in SIMCE test scores.

[13] Of course, the years of each parent education can not defer by more than positive six years so the min and max are inaccuracies in giving the information. Anyway they are by far the exceptions in the whole dataset.

Table A.1 Summary statistics of competition indices in changes

Variables in changes	Mean	Std. Dev.	Min.	Max.
overall competition index	-0.719	5.32	-32.738	38.59
voucher competition index	4.288	4.241	-23.237	42.928
public voucher competition index	0.298	1.761	-7.348	11.842
private voucher competition index	3.991	3.888	-20.04	37.114
private competition index	-5.008	3.493	-16.916	2.148

Table A.2 Summary statistics list of covariates in changes

Variables in changes	Mean	Std. Dev.	Min.	Max.
pupils per school	236.612	1000.27	-6140	5187
father education	0.159	2.926	-18	21
mother education	0.195	2.824	-16	18
books at home	0.393	69.164	-250	178
households with $>$ 100 books	-0.056	0.484	-1	1
households with pc	0.318	0.547	-1	1
households with internet	0.264	0.533	-1	1

Table A.3 Overall and voucher in changes

Secondary-Primary

	SIMCE in changes	
	all pupils	
overall competition index in changes	-0.123	-0.147
	(.088)	(.091)
	SIMCE in changes	
	voucher pupils	
voucher competition index in changes	-0.067	-0.072
	(.056)	(.065)

[1] **Source:** Author's calculation from research data provided by the Chilean Ministry of Education's Agency for Quality in Education. The data is available only to researchers after submitting a written proposal.

[2] The table shows the effect of changes in competition between primary, in 2002, and secondary school, in 2008, on value-added in educational test scores without and with covariates for pupils who sat for the SIMCE Language and Math Tests (unified in one ranking) in the $4^{\mathbf{th}}$ grade of primary school in 2002 and (but not or) for the SIMCE Language and Math Tests (unified in one ranking) six years later in the $10^{\mathbf{th}}$ grade of secondary school in 2008 for all and voucher schools located in Greater Santiago.

[3] *** p<0.1, ** p<0.05, *** p<0.01**

Table A.4 Heterogeneity in changes

variable	changes
Baseline	-0.072
family background	
income pcm	
Above 95% (≈ £2,000)	2.822
Above median (≈ £200)	0.113
Below median	-0.05
Below 5% (≈ £50)	-0.09
average years of school	
Above 95% (17 years)	0.955**
Above median (12 years)	0.132
Below median	-0.017
Below 5% (6 years)	-0.125
books at home	
Above 95% (180)	0.687**
Above median (35)	-0.025
Below median	0.08
Below 5% (2)	-0.143
pc at home	
Yes	0.122
No	-0.339
internet at home	
Yes	0.222
No	-0.318
school characteristics	
mean number of pupils	
Above 95% (2,172)	-0.248
Above median (829)	-0.052
Below median	0.061
Below 5% (248)	0.173
SIMCE average	
Above 95% (80.84)	0.238
Above median (48.955)	0.177
Below median	-0.048
Below 5% (24.33)	-0.274**
SIMCE standard deviation	
Above 95% (27.83)	0.629
Above median (23.5)	0.192
Below median	-0.42
Below 5% (15.77)	0.145

[1] Effect of competition on test scores by family background and school characteristics in changes with controls. Baseline is regression in changes for voucher school test scores on voucher school competition.

[2] *** p<0.1, ** p<0.05, *** p<0.01**

B. Disaggregating

Tables B.1 and B.2 show the "disaggregated" twice competition, firstly in voucher and private competition and secondly in public and private vouchers. Doing so permits the different causes of a better performance to be disentangled to identify, for example, whether the competition among voucher or private schools or, perhaps, public voucher or private voucher schools leads to better or worse performance in overall pupils. Table B.1 shows similar results for primary and secondary school with voucher school competition—after including controls and municipality FE—does not have any effect on overall pupil test results while private competition has a strong positive effect on them. In the other form of splitting overall competition, i.e. splitting voucher competition into public voucher and private voucher leaving private competition outside the model, the main findings are that public voucher competition causes pupils to perform worse while private voucher competition causes overall pupils to perform better overall. Therefore, in summary, vouchers and public voucher competition cause a negative effect (i.e. lead to a decreased performance) in pupils overall, while private vouchers and private competition have a positive effect (i.e. lead to an increased performance). But all these is permeated with bias so it is needed to use an IV procedure. Table B.2 indeed shows that in contrast there are no positive effect on either of the four types of competition on overall pupil tests, which confirm the main paper results.

Tables B.3 and B.4 show both the cause and effect "disaggregated", so it is possible to ask whether public voucher competition affects private voucher pupils, or any of the overall 48 possibilities (in primary/secondary, voucher-private/public-private voucher) possibilities. As shown, the outcome is that voucher competition exerts a negative effect on public voucher pupils, a positive effect on private voucher pupils and a significant negative effect on private school pupils in primary school. Moreover, public voucher competition provokes a clear negative effect in private school pupils while a mixed effect on public voucher and private voucher school pupils.

Private voucher competition causes a mainly positive effect on the three types of test results and pupils. The most remarkable results are the positive effect of private competition on public voucher, private voucher and the same private test scores and the mixed effects of voucher competition on the "disaggregated" test results and pupils. It is also noticeable that public and private voucher competitions have a rather neutral effect on all types, but a negative effect on private voucher schools. This result may be attributed to the private schools' location, namely the "High Neighbourhood".

Table B.3 shows that voucher competition causes a negative effect on private pupils on primary school and a positive effect on voucher pupils and a negative effect on private school pupils on secondary. While private competition creates primarily a positive effect on the three types of pupils at both level of school education. At the same time, public voucher competition causes various negative effects on the "disaggregated" pupils in both levels, while private voucher competition exerts a positive effect on voucher school pupils and private school pupils mainly in secondary school. This is for endogenously biased estimates, so there is need to at least check the regularity of these results with the instrumental variable regressions. Table B.4 shows that for primary schools, there is a relative confirmation of the OLS findings with some exceptions: There is a full positive effect on the "disaggregated" test results and pupils. At the same time, public voucher competition exerts positive effects on the "disaggregated" test results and pupils, while private voucher competition causes negative effects on almost every type of pupil. Conversely, it is in secondary school were the IV results show more novelties: Voucher competition provokes primarily negative effects, being positive only on private voucher school pupils, while private competition loses almost all its strength but for the effect on public voucher school pupils. Public voucher competition exerts positive effects but on its own public voucher school pupils, while private vouchers positively affect public voucher school pupils, negatively affect private school pupils and provides mixed effects to its own private voucher school pupils.

Table B.1 Effect of "disaggregated" competition on "aggregated" test results with/out controls

"disaggregated" competition	overall pupil test results		"disaggregated" competition	overall pupil test results	
primary					
voucher competition index	-0.043*	-0.037	public voucher competition index	-0.433***	-1.556***
	(0.023)	(0.081)		(0.112)	(0.373)
private competition index	0.695***	0.699***	private voucher competition index	-0.003	0.496***
	(0.048)	(0.207)		(0.039)	(0.155)
secondary					
voucher competition index	0.061***	0.022	public voucher competition index	-1.003***	-2.654***
	(0.020)	(0.048)		(0.115)	(0.073)
private competition index	1.130***	0.904***	private voucher competition index	0.071**	0.631***
	(0.047)	(0.102)		(0.028)	(0.074)
controls					
distance to GEO border	No	Yes		No	Yes
elevation	No	Yes		No	Yes
list of covariates	No	Yes		No	Yes
municipality FE	No	Yes		No	Yes

[1] **Source:** Author's calculation from research data provided by the Chilean Ministry of Education's Agency for Quality in Education.
[2] The table shows the effect of "disaggregated" competition on "aggregated" educational test score for pupils who sat for the SIMCE Language and Math Tests (unified in one ranking) in the $4^{\mathbf{th}}$ grade of primary school in 2002 and for the SIMCE Language and Math Tests (unified in one ranking) six years later in the $10^{\mathbf{th}}$ grade of secondary school in 2008 for all schools located in Greater Santiago. The table shows the OLS regressions in levels for 2002 and 2008 regarding the effect of "disaggregated" by voucher (public voucher and private voucher) and private school competition indices on overall test scores (not disaggregated by school type) with/out controls (distance to GEO border, elevation, list of covariates and municipality FE). A competition index is the number of schools an average pupil from a given school can choose. It is a by-municipality weighted average of the choice indices that measures the revealed preference number of schools that pupils from a given municipality are effectively choosing. Clustered at the school level.
[3] *** p<0.1, ** p<0.05, *** p<0.01**

Table B.2 Effects of "disaggregated" competition on "aggregated" test results with/out controls

instruments: slope/ruggedness	overall pupil test results			overall pupil test results	
"disaggregated" competition			"disaggregated" competition		
primary					
instrumented			**instrumented**		
voucher competition index	-0.202*	-0.183*	public voucher competition index	-4.835	-2.238**
private competition index	0.342	2.216	private voucher competition index	1.873	-2.282
tests					
First stage coeff.	-2.038***/1.637***	-1.977/1.568***		27.11/20.36	4.32/7.07
First stage F t.(F-s r.)	14.33	1.61		1.40	1.63
secondary					
instrumented			**instrumented**		
voucher competition index	-1.354*	-1.161*	public voucher competition index	-7.053**	-10.308***
private competition index	0.977*	2.259***	private voucher competition index	1.488	0.966
tests					
First stage coeff.	-1.331***/-0.087	-1.223***/-0.133		34.52/33.53	3.14/6.08
First stage F t.(F-s r.)	9.93	7.22		3.00	7.23
controls					
list of covariates	No	Yes		No	Yes
municipality FE	No	Yes		No	Yes
distance to GEO border	Yes	Yes		Yes	Yes
elevation	No	Yes		No	Yes

[1] **Source:** Author's calculation from research data provided by the Chilean Ministry of Education's Agency for Quality in Education. The data is available only to researchers after submitting a written proposal.
[2] Instrument slope and ruggedness. First stage with F test.
[3] s.=statistics, t.=test, F-s r.=First-stage relevance, Ui=Under identification test.

Table B.3 Effects of "disaggregated" competition on "disaggregated" test results with/out controls

"disaggregated" competition	pupils "disaggregated" by school type					
	pub vou	pri vou	pri	pub vou	pri vou	pri
primary						
voucher competition index	-0.006	0.015	-0.112***	-0.063	0.094	-0.267**
	(0.032)	(0.032)	(0.041)	(0.145)	(0.119)	(0.109)
private competition index	0.397***	0.1693**	0.362***	-0.794**	0.719**	0.001
	(0.075)	(0.078)	(0.074)	(0.392)	(0.315)	(0.184)
public voucher competition index	0.122	0.199	-1.132***	-0.573	-0.700	-1.034*
	(0.126)	(0.153)	(0.354)	(0.595)	(0.463)	(0.594)
private voucher competition index	-0.043	-0.021	0.140	0.128	0.348*	0.026
	(0.051)	(0.051)	(0.143)	(0.269)	(0.182)	(0.267)
secondary						
voucher competition index	0.070***	0.101***	-0.102***	0.031	0.068	-0.203***
	(0.021)	(0.022)	(0.031)	(0.055)	(0.059)	(0.072)
private competition index	0.731***	0.692***	0.267***	1.132***	1.208***	-0.081
	(0.100)	(0.075)	(0.076)	(0.292)	(0.187)	(0.127)
public voucher competition index	-0.065	0.107	-0.996***	-1.622***	-2.542***	-2.778***
	(0.109)	(0.119)	(0.316)	(0.236)	(0.238)	(0.217)
private voucher competition index	0.048*	0.054*	0.094	0.399***	0.702***	0.788***
	(0.028)	(0.032)	(0.103)	(0.101)	(0.099)	(0.210)
controls						
distance to GEO border	No	No	No	Yes	Yes	Yes
elevation	No	No	No	Yes	Yes	Yes
list of covariates	No	No	No	Yes	Yes	Yes
municipality FE	No	No	No	Yes	Yes	Yes

[1] **Source:** Author's calculation from research data provided by the Chilean Ministry of Education's Agency for Quality in Education. The data is available only to researchers after submitting a written proposal.
[2] The table shows the effect of "disaggregated" competition on "disaggregated" educational test scores without covariates for pupils who sat for the SIMCE Language and Math Tests (unified in one ranking) in the $4^{\mathbf{th}}$ grade of primary school in 2002 and for the SIMCE Language and Math Tests (unified in one ranking) six years later in the $10^{\mathbf{th}}$ grade of secondary school in 2008 for all schools located in Greater Santiago. The table shows the OLS regressions in levels for 2002 and 2008 regarding the effect of "disaggregated" competition indices (public voucher, private voucher and private competition indices) on "disaggregated" by voucher (public voucher and private voucher) and private school competition indices on "disaggregated" test scores by school type (public voucher, private voucher and private) with/out controls (distance to GEO border, elevation, list of covariates and municipality FE). Clustered at the school level.
[3] *** p<0.1, ** p<0.05, *** p<0.01**

Table B.4 Effects of "disaggregated" competition on "disaggregated" test results with/out controls

instruments: slope/ruggedness						
"disaggregated" competition	pupils "disaggregated" by school type					
	pub vou	pri vou	pri	pub vou	pri vou	pri
primary						
voucher competition index	-0.052	-0.055	0.719	1.233	1.337	0.944
private competition index	0.076	0.066	4.302	4.020	3.838	-0.753
tests						
First stage coeff.	5.08/6.45	10.30/24.56	7.64/1.33	1.42/0.46	2.75/4.68	3.04/1.50
First stage F t.(F-s r.)	3.81	5.30	6.14	1.12	7.98	6.51
public voucher competition index	0.320	0.317	12.902	-5.650	8.519	5.715
private voucher competition index	-0.192	-0.166	-9.078	1.684	-1.041	-0.501
tests						
First stage coeff.	6.81/4.63	10.00/10.83	8.62/6.89	0.74/1.64	1.05/3.36	2.85/2.98
First stage F t.(F-s r.)	2.51	1.72	4.34	6.08	1.37	5.99
secondary						
voucher competition index	-0.137	-0.271	-0.813	-126.603	1.908	-1.116
private competition index	3.745	-0.595	-2.363	42.485	-0.208	-1.537
tests						
First stage coeff.	5.94/0.90	10.51/14.95	7.31/0.45	0.10/2.97	5.09/5.96	1.79/2.92
First stage F t.(F-s r.)	6.17	3.56	7.08	5.01	5.18	7.46
public voucher competition index	-11.024	108.711	2.163	-54.177	10.152	5.399
private voucher competition index	1.178	-22.580	-1.401	17.128	1.236	-2.661
tests						
First stage coeff.	4.02/5.82	8.23/10.40	8.29/6.68	0.31/0.06	0.32/6.61	2.44/1.59
First stage F t.(F-s r.)	6.43	5.00	3.93	4.12	7.04	3.06
controls						
distance to GEO border	Yes	Yes	Yes	Yes	Yes	Yes
elevation	Yes	Yes	Yes	Yes	Yes	Yes
list of covariates	No	No	No	Yes	Yes	Yes
municipality FE	No	No	No	Yes	Yes	Yes

[1] **Source:** Author's calculation from research data provided by the Chilean Ministry of Education's Agency for Quality in Education. The data is available only to researchers after submitting a written proposal.
[2] Instrument slope, ruggedness and both. First stage with F test.
[3] s.=statistics, t.=test, F-s r.=First-stage relevance, Ui=Under identification test.

Chapter 2

Effect of diminishing School Attendance on Test Scores: Evidence from the Chilean School Occupations

2.1 Introduction

This chapter is the result of my attempt to measure the effect of a dramatic decrease in school attendance related to the so-called "Chilean Winter", a huge social outburst that shook the country with an explosion of pupil protests which started in early June of 2011 and continued with sequential but diminishing bursts until the end of 2012—and even had smaller replicas until fairly recently. The question to be answered relates to whether or not there is a causal relationship between lost school days in the context of protests and school occupations and a decrease in the standardized test performance for those pupils whose schools were occupied. I do not want to deter anybody from participating in politics. Indeed, this paper says nothing about the righteousness or evil of school occupation. This paper is about school attendance and school occupation is used only as an external exogenous variation

to disentangle the effect of a worsening in school attendance on pupil achievement. I just want to put forwards some facts and causal effects of school closure due to school occupation in terms of test score achievements and human capital accumulation.

It is difficult to imagine a timeless production function for education. New technology has changed the paradigm, but traditional face-to-face teacher/pupil interaction is still considered essential to learning. A recent report by the UK Department of Education, *Improving Attendance at School*, illustrates the fact that despite a diminishing trend in cancelled school days in England, there were still 57 million days of school missed in 2009/2010. Second, there is a clear link—but not necessarily causal—between poor attendance at school and lower academic achievement. Third, of pupils who miss more than 50 per cent of school, only three per cent manage to achieve five or more GCSEs at grades A* to C, including mathematics and English. In contrast, 73 per cent of pupils who have over 95 per cent attendance achieve five or more GCSEs at grades A* to C. Fourth, when considering attendance, it is worth noting what a one percentage point improvement means in terms of days missed. An average-sized secondary school that manages to improve its attendance by one percentage point represents an additional 1,300 pupil-days spent in school in a year. That is a significant amount of education.

Some of this evidence can be directly extrapolated to the case in Chile. School absence is detrimental to pupil performance. Children who are persistently absent perform worse at school and have worse job prospects, which has the potential to increase poverty and crime rates. However, this study does not focus on chronic absence, but is concerned with the hypothesis that poor test scores can be directly linked to an acute—but not necessarily short—episode of days missed, namely the episode suffered within the 205 occupied schools during the so-called *Chilean Winter*. Estimates indicate that during this sole episode, almost 8 million pupil-days were missed from these 205 occupied schools. To put this into perspective, while in

England there are 8.2 million pupils attending 24,372 schools, Chile has roughly half of both, with 3.6 million pupils attending 12,063 schools. As noted above, in England there were a total of 57 million pupil-days of school missed in 2009/2010. To restate, in Chile 205 out of 12,063 schools represented almost 8 million riot-related pupil-days missed. It is therefore a very plausible hypothesis that this period of absence could have directly and negatively impacted the process of education in Chile. Therefore, this study seeks to establish a causal relationship between lost school days and standardized test results.

Thanks to the national register of attendance, the number of cancelled days experienced by the identified schools can be measured. At a school-level, cancelled days refer specifically to days missed because the school was occupied. It does not include pupil sick days or absences for any other cause. On the other hand, when examining the data at pupil-level, it is easy to detect different types of cancelled days. For instance, school occupation cancelled days manifest themselves as the sharp end of variation in individual attendance and create a distinct difference between individual and average attendance for each occupied school. In Chile, there is not a fixed number of school days, but there is a regulatory framework issued for each administrative division which outlines general rules. Schools are then free to choose a calendar under these general rules. The regulatory framework fixes two weeks of winter holidays, the starting and ending days of classes, and some national holidays to be followed. In other words, there is no absolute baseline for attended school days to make comparisons, and this study must rely on a relative method to identify cancelled days. Unfortunately, this method fails to tell if some controls should better be included as additional occupied schools. For this reason my estimates may underestimate the real figures. Cancelled days can be identified in the national register because schools report few open or school working days during occupations. Therefore, cancelled days are calculated directly as the difference in open or working days in the national registry of attendance between occupied and non-occupied schools. The study then holds the following unifying definition. Pupil-

missed school days, pupil forced-to-cancel school days, and pupil-lost days, have the same meaning in this study. At the same time, "missed"/"cancelled"/"lost" school days are used interchangeably, meaning non-attended school days because of school occupations.

It is well known in the literature that there is a positive correlation between school attendance and school performance, i.e. pupils who attend more classes are better off: Goodman—using Massachusetts data for American students and an instrumental variable exploiting the fact that moderate snowfall induces student absences while extreme snowfall causes school closures—shows that exogenous snow days' disruption and the absence provoked by bad weather reduces math achievement by $0.05\boldsymbol{\sigma}$, Goodman (2014). At the same time, longer school days and school years are also associated with better pupil achievement: Hincapié shows for Colombia—where municipalities were given more flexibility to choose the length of the school day for their schools following the rescindment in 1994 of the full school day reform; some schools offer a full school day (7 hours), meanwhile others have half school days (or two separate 4 or 5 hour shifts)—that the cohorts exposed to full school days have test scores that are about $0.1\boldsymbol{\sigma}$ higher than cohorts that attended half school days, Hincapié (2016). While Pischke uses variation introduced by two short school years in Germany—shorter by a third—finding more repetition and less continuing education, but not finding any adverse effect on long-term outcomes, e.g. earning or employment, Pischke (2007). How does this compare with my results? After using different econometric techniques I have very robust estimates which for "hardliner"—students who on average occupied their school for 48.08 days: more than 2 months and a week—imply a decrease in $0.18\boldsymbol{\sigma}$ on the standardized test scores. So my estimates are somehow higher most probably because the treatment is more intense and perdurable .

In this chapter, using a unique list of occupied schools filtered by the Chilean police, I implement a Difference-in-Difference (DiD), a Difference-in-Difference-in-

Difference (DiDiD) and synthetic control estimates to assess the following research question: how did lost school days in the context of protests and school occupations affect overall pupil performance on standardized tests during the *Chilean Winter* (June 2011-November 2012)? I classify schools as occupied and non-occupied according to a leaked police report. I further distinguish between occupied-S (occupied schools that gave up the occupation and sat for the first standardized test, the SIMCE, after the onset of the revolts; because they allowed the standardized test to take place I also called them moderates) and occupied-NS (occupied schools that did not give in so they did not take the first standardized test, the SIMCE, after the onset of the revolts; because they did not surrender, I call them hard-liners). The evidence provided by the DiD indicates that the average test results of the hard-liner occupied group who suffered an average of 48.08 days of school cancellations due to the riots, i.e. 2 calendar months and 1 week of missed days, decreased by 5.4 percentile; this result is strongly statistically significant at the 5% level and sizeable in magnitude. The performance of these pupils dropped, as commented on before, almost 0.18 standard deviation, which is large in terms of its cost in human capital. This means that there is also sufficient evidence not to reject the hypothesis that protests caused a decrease in the performance of the pupils in occupied schools. The DiDiD estimates also give similar results. For example, for the occupied-NS $10^{\mathbf{th}}$-$4^{\mathbf{th}}$ the DiDiD estimate shows a decrease of 5.78 percentile or 0.19 of a standard deviation. Additionally, the synthetic control method gives a reduction of 4.6 points in the standardized tests which is a reduction of 0.1 standard deviation in test results.

Throughout this paper, I use the definition of causality presented in Lewis (1973), Pearl (2009) and Woodward (2003). For the Difference-in-Difference estimate, I closely follow the work of Angrist and Pischke (2009), Card (1990) and Card and Krueger (1994). For the synthetic control estimate, I employ the work of Abadie et al. (2011), Diamond et al. (2010) and Abadie et al. (2003). Furthermore, there is also companion literature on the effects of compulsory attendance in a quasi-experimental setting, for example, Angrist and Krueger (1991) and Card and Lemieux (2001).

2.1.1 Background and theory

During the *Chilean Winter* and within a matter of days riots reached the national level with hundreds of thousands of pupils occupying schools, marching on the streets and demanding better education. The protests lasted for different periods of time in different schools, but as a general rule most schools were no longer occupied by the end of 2012 (see Appendix IV).

The SIMCE, the outcome this chapter examines, is the Chilean standardized test which takes place regularly each year but for different cohorts. Chilean primary education is divided into eight grades, identified as $1^{\mathbf{st}}$-$2^{\mathbf{nd}}$-$3^{\mathbf{rd}}$-$4^{\mathbf{th}}$-$5^{\mathbf{th}}$-$6^{\mathbf{th}}$-$7^{\mathbf{th}}$-$8^{\mathbf{th}}$. Secondary schools have four grades which will be identified here as $9^{\mathbf{th}}$-$10^{\mathbf{th}}$-$11^{\mathbf{th}}$-$12^{\mathbf{th}}$. This amounts to twelve years of education. The SIMCE test sometimes covers languages, math and science, as in $4^{\mathbf{th}}$ and $8^{\mathbf{th}}$ grades, and sometimes covers just languages and math, as in $10^{\mathbf{th}}$ grade; these are the three grades analysed.

Pupils want better education, and in the event just described they fought for public education and against for-profit schools. I will not assert if they are right or wrong. But their voice stands from the fact that the students have been the protagonists of the disputed Chilean voucher experiment. Academia, including my personal work on chapter 1 of this thesis, has found that the voucher system does not seem to respond to spatial incentives and educational outcomes are unimpressive. In other words, while competition among schools, matching families to schools, the closing of bad schools, mainly public schools, and the entrance to the market of new but seemingly not more productive schools, have all occurred, they have been unsuccessful in improving the performance of the whole educational system. In fact, the vast majority of poor pupils have received an extremely bad education which has condemned them to further poverty. What pupils want is an opportunity society. They want to boost their talents. They are asking only for a chance at a better life

through education. But some of them are doing so by hampering their own personal education with a dramatic excision of their own number of school days: a break on their academic life through a radical blow to their school attendance.

I also acknowledge the existence of "compensatory" activities by students (or students' parents) offsetting a reduced investment in human capital with greater out-of-school educational effort. In other words, parents or the students themselves, may be able to exploit the school absence to reduce the lack of investment due to lost school days with special educational aids and commitments. In fact, better educated parents are expected to compensate more than less educated ones. This activity will bias downwardly the estimates, causing the effect of school absence to be less steep than it really is.

There may be two different channels to link the lost days due to school occupation with the diminishing scores of the outcome exam. The main and straightforward channel is through pupils' school absence. The longer the absence, the greater the impairment in human capital accumulation. This can be tested by the dose effect model discussed below. There is also a type of school deferral, which implies that the activities in which the pupils are involved are completely different from the activities, for instance, during a holiday recess. One way is to add depth at the recess in academic activities when a school occupation is in place: one week of school occupation, or just the threat of occupation, could influence more than the customary absences from school. This effect can be captured just by the treated dummy variable of belonging or not to the police report list of occupied schools, irrespective of the days of school missing.

This model could not exhaust itself on the demonstration of protests and school occupation as instruments on missed school days to explain bad performance. In fact, the direct relation of absenteeism caused by school occupation is also likely to take place. Absenteeism *due to* school occupation can explain the variation in pupil

achievement. I argue that the character of the absenteeism is not different from other instigators. In a typical holiday period, the composition of the spare time could not be radically different from the composition of spare time during a school occupation absence. A school occupation is a radical step, potentially altering many strata of life, such as sociability, friends and classmates, significant others and parents. All these relations could become tense to the maximum if the school has been occupied. But the core common causal effect is the same: lost school days or reduced school attendance.

In Section 2, the model is outlined and in Section 3 the treatment and the outcome are defined. Section 4 offers pupil-level data; an analyses for school-level data; a robustness check; an evaluation of the identifying assumptions; and an evaluation of missed days as an extension of causality. Finally, Section 5 provides a conclusion. Five appendices include the complete police report list of 205 occupied schools in the central region of Santiago (I), the synthetic control R code (II), a robustness check for the synthetic control employed here (III), an outburst review of the Chilean Winter in the media at a national level (IV) and the geography of the occupied schools (V).

2.2 Empirical methods

2.2.1 The model

All my regressions are restrictions of the following general model.

$$\mathbf{Y_{ist}} = \boldsymbol{\alpha_i} + \boldsymbol{\beta}\mathbf{missed_{ist}} + \mathbf{X'_{it}}\boldsymbol{\Lambda} + \mathbf{Z'_{st}}\boldsymbol{\Theta} + \boldsymbol{\varphi_s} + \boldsymbol{\delta_t} + \boldsymbol{\varepsilon_{ist}} \tag{2.2.1}$$

The indices are as follows: $\mathbf{i}$ is individual; $\mathbf{s}$ is school; $\mathbf{t}$ is level of education/cohort, primary or secondary. The dependent variable is $\mathbf{Y_{ist}}$, a standardized test score of pupil i going to school s at level of education t. The main independent variable is a continuous variable $\mathbf{missed_{ist}}$ which represents missed school days or the causal variable of interest (which will be proxied by a dummy if referring to an occupied school during the *Chilean Winter*, $\mathbf{occupy_{st}}$). The control variables are $\mathbf{X_{it}}$, individual characteristics, and $\mathbf{Z_{st}}$, school characteristics. Individual characteristics include prior achievements (SIMCE or grades), school choice, age, gender, family background, books at home, or disadvantaged group member (family receiving subsidy, Chile Solidario, vulnerability index, integrated, etc.). School characteristics include same-grade average standardized test score, size and school type dummies PubV, PriV and Pri (the three sectors of Chilean educational system: public voucher, private voucher and private. See Section 1.2. The school system for further details). Fixed effects include $\boldsymbol{\alpha_i}$ (an individual fixed effect which causes problems when correlated with the error term),[1] $\boldsymbol{\varphi_s}$ (a school fixed effect), and $\boldsymbol{\delta_t}$ (a level of education/cohort fixed effect). Finally, the error term $\boldsymbol{\varepsilon_{ist}}$ represents possible serial correlation, heteroskedasticity, or unobserved characteristics (other teacher qualities outside of observable teacher evaluations, motivation of neighbours, or other local resources that improve pupil educational outcomes).[2] A data generating process with pupils sampled from many schools is likely to be correlated. Therefore, standard errors should be clustered at the school level.

This study also considers the following relationship, which demonstrates the identification strategy for uncovering the causal relation of interest.

$$\mathbf{Y_{ist}} = \boldsymbol{\alpha_i} + \boldsymbol{\beta}\mathbf{occupy_{st}} + \mathbf{X'_{it}}\boldsymbol{\Lambda} + \mathbf{Z'_{st}}\boldsymbol{\Theta} + \boldsymbol{\varphi_s} + \boldsymbol{\delta_t} + \boldsymbol{\varepsilon_{ist}} \tag{2.2.2}$$

[1] The correlation between the regressors and the individual fixed effect, unobserved effect or unobserved heterogeneity ($\boldsymbol{\alpha_i} = \boldsymbol{\alpha} + \boldsymbol{\gamma A_i}$) requires controlling for the fixed effect, DiD and synthetic control method.

[2] As an assumption, there is no correlation between the causal regressor and the idiosyncratic error $\boldsymbol{\varepsilon_{ist}}$. Therefore, I can proceed with DiD and synthetic control analysis.

The reduced form equation shows how school occupation in the *Chilean Winter* can be a an instrument for missed school days (see discussion of a fully IV model at the end of the chapter as an extension). The equation is used to disentangle the model's unobservables.

Finally, in order to clarify the main DiD model is as follows:[3]

$$\mathbf{Y_{ist}} = \boldsymbol{\gamma_s} + \boldsymbol{\lambda_t} + \boldsymbol{\delta}\mathbf{occupy_{st}} + \boldsymbol{\varepsilon_{ist}} \tag{2.2.3}$$

where $\mathbf{i}$ is a student, $\mathbf{s}$ denotes school and $\mathbf{t}$ level of education/cohort and $\mathbf{E}(\boldsymbol{\varepsilon_{ist}}|\mathbf{s},\mathbf{t}) = \mathbf{0}$. $\boldsymbol{\gamma_s}$ is a occupied/non-occupied time-invariant school group effect and $\boldsymbol{\lambda_t}$ a level of education/cohort that is common across schools. $\mathbf{occupy_{st}}$ is a dummy for occupied schools and level of education/cohort and $\boldsymbol{\delta}$ is a constant equal to $\mathbf{E}[\mathbf{Y_{1ist}} - \mathbf{Y_{0ist}}|\mathbf{s},\mathbf{t}]$. The additive structure permits for each non-occupied school:

$$\begin{aligned}\mathbf{E}[\mathbf{Y_{ist}}|\mathbf{s} = \mathbf{non}\quad\mathbf{occupied}\quad\mathbf{schools},\mathbf{t} = \mathbf{10^{th}}]-\\ \mathbf{E}[\mathbf{Y_{ist}}|\mathbf{s} = \mathbf{non}\quad\mathbf{occupied}\quad\mathbf{schools},\mathbf{t} = \mathbf{4^{th}}] = \boldsymbol{\lambda}_{\mathbf{10^{th}}} - \boldsymbol{\lambda}_{\mathbf{4^{th}}}\end{aligned} \tag{2.2.4}$$

and for each occupied school:

$$\begin{aligned}\mathbf{E}[\mathbf{Y_{ist}}|\mathbf{s} = \mathbf{occupied}\quad\mathbf{schools},\mathbf{t} = \mathbf{10^{th}}]-\\ \mathbf{E}[\mathbf{Y_{ist}}|\mathbf{s} = \mathbf{occupied}\quad\mathbf{schools},\mathbf{t} = \mathbf{4^{th}}] = \boldsymbol{\lambda}_{\mathbf{10^{th}}} - \boldsymbol{\lambda}_{\mathbf{4^{th}}} + \boldsymbol{\delta}\end{aligned} \tag{2.2.5}$$

[3] -1 refers to previous cohort or level of education

Then the population DiD is the causal effect $\boldsymbol{\delta}$ of $\mathbf{occupy_{st}}$:

$$\begin{aligned}
&\{\mathbf{E}[\mathbf{Y_{ist}}|\mathbf{s = occupied \quad schools, t = 10^{th}}]- \\
&\mathbf{E}[\mathbf{Y_{ist}}|\mathbf{s = occupied \quad schools, t = 4^{th}}]\}- \\
&\{\mathbf{E}[\mathbf{Y_{ist}}|\mathbf{s = non \quad occupied \quad schools, t = 10^{th}}]- \\
&\mathbf{E}[\mathbf{Y_{ist}}|\mathbf{s = non \quad occupied \quad schools, t = 4^{th}}]\} \\
&= \boldsymbol{\delta} \quad \square
\end{aligned} \tag{2.2.6}$$

2.3 Data

2.3.1 Treatment: Occupied schools

This study strictly follows the filtered Carabineros/LA TERCERA list of occupied schools in Santiago (see Appendix I). The first standardized test -called the SIMCE- was scheduled on October 19$^{\text{th}}$ 2011, four months after the onset of the protests. In relation to this test, the list can be further divided into 108 occupied schools that gave up occupation and held the test on that day, which this study calls Occupied-S/moderate schools (Occupied-S meaning occupied according to the police report + SIMCE). There are 97 Occupied-NS/hard-liner schools (Occupied-NS meaning occupied according to the police report + No SIMCE). The control group and all three *treatment* groups (O, Occupied-S and Occupied-NS) have no mechanically similar observable characteristics. Nevertheless, a comparison of the control Occupied-S and Occupied-NS groups in Table 2.1 indicates that the groups are indeed comparable. Only slight differences exist between the control group and both occupied groups in the percentage of school type, but not between the groups themselves. There are some specific differences in treatment (post-treatment SIMCE). There is also a particularly higher income in the occupied-NS group which shows that families with above average income have their children in the hard-liner schools (which in Chile is called an aspirational family), but this unevenness can be eas-

ily overcome using methods and assumptions such as DiD, matching and synthetic control. Occupations continued after this first SIMCE test and into 2012, the next year. However, on the next test date of November 6thth, 2012, almost all of the 205 schools sat for the test. Figure 2.1 shows a map with the treated groups: Occupied-S and Occupied-NS.

Note that while the decision to enter into the treated group was sometimes made through a democratic process and sometimes through the decision of a few, as soon as the school was occupied, no more academic work could be performed there. Lectures and classes all stopped immediately and could not resume until the occupation had surrendered. School days were in fact cancelled for every pupil in the occupied school. This signifies that, at least for those forced to comply with the school occupation, the treatment can be viewed as if randomly assigned in the sense that, in principle, missed days from school are assumed to be uncorrelated with the other determinants of standardized test performance changes over this period. Being able to identify this margin could be a significant extension of this paper.

2.3.2 The SIMCE

As mentioned above, the SIMCE is this chapter's outcome. The SIMCE is the Chilean standardized test, which takes place regularly each year for different cohorts. Figure 2.2 shows the relevant SIMCE exams. It includes two pupil-level cohorts in the periods of 2006-2012 and 2007-2011. Occupied-S schools sat for both SIMCEs (October 19th, 2011 and November 6thth, 2012). Occupied-NS schools did not sit for the October SIMCE test, because they did not surrender their school occupation. Figure 2.2 shows nine 2-period cohorts and one 3-period cohort.

Pupil-level cohorts have two or three observations during the cohort time lapse. One of the two cohorts were for 4th-year students in 2007, 8th-year students in 2011 and 10th-year students in 2013. The other cohorts included pupils that were in 4th

level in 2006 and in level 10^{th} in 2012. Both cohorts can be tracked at a pupil level through both the SIMCE results and several pupil, parent and teacher questionnaires. Pupil-level cohorts provide an insight into the micro-economic foundations of the data which model the individual decision of the agents for the naïve estimator, DiD, matching, and regression with pupil-level controls and dose-response analysis. The other margin at which the data can be handled is on the school level. This approach allows the creation of new cohorts along the diagonal, or following a given school at different years across pupil-level cohorts. This is more suitable for synthetic control, naïve estimator, DiD, matching and regressions with school-level controls and dose-response analysis.

2.4 Results

2.4.1 Pupil-level analysis

Difference-in-Difference

This section presents the results of the DiD estimates. Table 2.2 shows the DiD estimate of the average test results per school before and after occupation for Occupied-S schools on October 19^{th}, 2011, where just Occupied-S schools sat for the test. The results (for the first row) show that the pre-treatment test results (the 2007 SIMCE for 4^{th} grade of primary school in languages, math, and science) were lower, higher and higher respectively in the 108 Occupied-S schools than the results in the control group (CG). The difference is strongly significant at the 5% level. Later, during the protests and school occupations that started in early June of 2011 and lasted until the date of the tests (October 19th 2011), an average of at least 11.58 days (or 2 calendar weeks) of normal school activities were lost in the schools that appeared on the police report list of occupied schools. These losses took place at the end of June, the climax of the occupations, and the occupations were suspended in order to sit for the post-treatment 2011 languages, math and science SIMCE tests. The post-treatment test results for Occupied-S schools are

significantly lower than those of the control group. The third row shows that the difference in test results from pre- to post-treatment years is negative, showing a decrease in outcome for the treated units. That is, for Occupied-S, the three test results decrease with a strongly significant drop in scores. Under the parallel trend assumption, this means that there is sufficient evidence to reject the *null* hypothesis that lost school days during school occupations do not affect pupil performance on standardized tests during the *Chilean Winter*. In particular, the average of 11.58 days of school occupation leads to a decrease in 1.56, 2.23 and 2.62 percentile points in the test result outcomes for the 108 Occupied-S schools.

Table 2.3 shows the same 108 Occupied-S schools after 9.56 lost school days. This is less time lost than in the previous sub-period, because the Occupied-S schools implemented extra time activities to catch up with the control group. Again, the first-row results show that the pre-treatment test results for the 2006 SIMCE are lower and lower in the 108 Occupied-S schools than those of the control group (CG). These results represent the 2006 SIMCE for 4^{th} grade of primary school, which took place six (not four) years before the post-treatment test in languages and math (science was excluded from the 10^{th} SIMCE test). For the second row, post-treatment test results for Occupied-S schools are respectively lower than the control group. The third row shows that the difference in test results from pre- to post-treatment years is slightly positive in the control group as well. For Occupied-S schools, the two test results drop by 1.45 (language) and 1.10 (math) relative to the control group and before and after the treatment. Under the parallel trend assumption, this means that there is sufficient evidence to reject the *null* hypothesis that lost school days during the school occupations did not affect pupil performance on standardized tests during the *Chilean Winter*. As a matter of fact, lost school days during the school occupations decreased pupil performance for Occupied-S schools during the period from June 2011 to November 6^{th}, 2012.

Finally, Table 2.4 overviews the 97 Occupied-NS schools that did not take the

2011 SIMCE but did take the November 6^{th}, 2012 SIMCE. The results (for the first row) show that the pre-treatment test results (the 2006 SIMCE for 4^{th} grade of primary school in languages and math) are lower and higher in the 97 Occupied-NS schools than those results in the control group (CG). The difference is strongly significant at the 5% level. During the period from June 2011 to the date of the second test after the onset of the revolts on November 6^{th}, 2012, at least an average of 48.08 days (or 2 calendar months) of normal school activities were lost in the schools that appeared on the police report list of occupied schools at the end of June, the climax of the occupations. These absences took place in schools that did not take the 2011 SIMCE. For the second row, the post-treatment test results for Occupied-NS schools remain strongly, significantly lower than those of the control group. The third row shows that the difference in test results from post- to pre-treatment years is negative, indicating that treated units decrease in outcome. That is, for Occupied-NS schools, the two test results plunge with a strongly significant drop in scores by 2.14 (languages) and 5.40 (math) percentile points. Under the parallel trend assumption, this means that there is sufficient evidence to reject the *null* hypothesis that lost school days during the school occupations did not affect the pupil performance on standardized tests. As a result, this study's main results hold for both school types. There is strong evidence that school occupations during the *Chilean Winter* significantly decreased the performance of the pupils of these two types of treated schools on their standardized tests. As a comparison, Hanushek (2003) suggests that a standard deviation increase of 1 in overall teacher quality is associated with a 0.11 standard deviation increase in student performance. This compares to an 0.18 decrease in standard deviation to which the 5.40 lower percentile points are equal. Therefore, an increase in 1.5 standard deviations in teacher quality is cancelled out by being on the occupied school list. Therefore, the effect outlined in this chapter is a particularly strong one.

Naïve estimators

The naïve estimators are a by-product of the DiD, or the mean of the treated minus the mean of the untreated *after* the treatment. Thus, it is expressed by the second row of each DiD estimator. Table 2.5 summarizes the results.

Regression with controls

To replicate all of these test results, the alternative regression approach can be considered, or a regression of the difference in the school occupation dummy. The results are exactly the same as those received using the DiD estimator. The advantage of this specification is that controls can be added to the model. Table 2.5 shows the results of adding nine controls, which are books, HH income, being indigenous, mother and father's education, pc at home, internet at home, preschool and repeating a year.

DiDiD

As seen above, there are losses in value-added for the cohorts immediately after the treatment, or "during" cohorts. If the pre- and post-cohort DiD estimates are added to these estimates, it can be seen that there are no gains or losses in value-added for pre-treatment cohorts for Occupied-S 8^{th} 2009-4^{th} 2005 and Occupied-S and Occupied-NS 10^{th} 2008-4^{th} 2002. The novelty is that these losses persist for the cohorts two years later, Occupied-S 8^{th} 2013-4^{th} 2009 and Occupied-S and Occupied-NS 10^{th} 2013-4^{th} 2007. This indicates that the effect of school occupation is persistent, as seen in Table 2.6. These are post cohorts because there is thought to be sufficient time for the effect to fade away. This study seeks to answer the question of precisely whether there is a long term effect. In Figure 2.2, Occupied-S and Occupied-NS are in orange (6-year gap with tenth to fourth 2-period cohorts) and Occupied-S are in yellow (4-year gap with eighth to fourth 2-period cohorts) with the last group sharing the green three-stage cohort. It is important to notice

that it is possible for only pupils to be followed from primary to secondary school, not schools. Applying a DiDiD approach, which subtracts the during-DiD estimate from the pre-DiD estimate, the data can be controlled for unobserved divergent or convergent trends in the series (data that do not have parallel trends). The following is the DiDiD procedure:

$$\text{DiDiD} = \text{DiD}_{\text{during}} - \text{DiD}_{\text{pre}} \tag{2.4.1}$$

The DiDiD strategy can also be used to analyse the only three-wave cohort which is precisely amid the onset of more active protests, $4^{\text{th}}07$-$8^{\text{th}}11$-$10^{\text{th}}13$, see Figure 2.7. Unfortunately, this cohort contains no Occupied-NS schools which did not sit for the October 2011 test. They continue to display a strong decay in performance as a result of the protests after controlling for unobserved divergent or convergent trends in the data series.

2.4.2 School-level analysis, i.e. same-grade-level analysis

Difference-in-Difference

This section presents the results of the DiD estimates for school-level analysis. These are not value-added components. Instead, they represent simple comparisons between one-year performance figures for a grade in a definite school and another cohort performance for the same grade and school in another year. These are not any value-added estimates because different cohorts are examined, and pupil averages or schools are analysed, not pupils themselves. Table 2.8 shows the DiD estimate of average test results per school before and after school occupations for Occupied-S schools on October 19^{th}, 2011, where just Occupied-S schools sat for the test. The shortest diagonal cohort of the same grade is used, 8^{th} in 2011.

The results (for the first row) show that the pre-treatment test results (the 2009 SIMCE for 8^{th}-year primary school in languages, math, and science) are, significantly, 5% lower in the 108 Occupied-S schools than in the control group (CG). Afterwards and during the protests and school occupations, which started in early June of 2011 and lasted until the date of the tests (October 19^{th}, 2011), at least an average of 11.58 days (2 calendar weeks) of normal school activities were lost in the schools that appeared on the police report list of occupied schools at the end of June, the climax of the occupations. The occupations of these schools were ended in order to allow students to sit for the post-treatment 2011 languages, math and science SIMCE tests. In the second row, the post-treatment test results for Occupied-S schools are significantly lower than those of the control group, and with similar magnitudes to those above. The third row shows that the difference in test results from pre- to post-treatment years is not significant at a 5% level. Under the parallel trend assumption, this means that there is not sufficient evidence to reject the *null* hypothesis that lost school days during the school occupations do not affect pupil performance on standardized tests—at least for these schools and time periods. In particular, the average of 11.58 days of school occupation leads to an insignificant decrease of 1.31, an increase of 0.67, and again a decrease of 1.78 in percentile points in the test result outcomes for the 108 Occupied-S schools.

Table 2.9 shows the same 108 Occupied-S schools, but now after 9.56 lost school days (approximately 2 calendar weeks). That is less time lost than before, because the schools implemented extra time activities to catch up with the control group for the shortest diagonal cohort, 10^{th} grade, which is 2010.

Again, the results for the first row show that the pre-treatment test results (the 2010 SIMCE for 10^{th} grade of secondary school) were lower in the 108 schools (Occupied-S) than those of the control group (CG). However, this time the results come from the period two years before the post-treatment test in languages and math (science was excluded from the 10^{th} SIMCE test). The difference is strongly

significant, at a 5% level. During the period from June 2011 to the date of the second test after the onset of the revolts on November 6th, 2012, at least an average of 9.56 days (approximately 2 calendar weeks) of normal school activities were lost in the schools that appeared on the police report list of occupied schools at the end of June, the climax of the occupations. These schools also sat the previous year's SIMCE test. For the second row, post-treatment test results for the Occupied-S schools are significantly lower than the control group with similar magnitudes as those above. The third row shows that the difference in test results from pre- to post-treatment years was not significant at a 5% level. Under the parallel trend assumption, this means that there is not sufficient evidence to reject the *null* hypothesis that lost school days during the school occupations do not affect pupil performance on standardized tests—at least for these schools and this time period. In particular, the average of 9.56 days of school occupation led to a insignificant increase of 1.75 and 0.53 in percentile points in the test result outcomes for the 108 Occupied-S schools.

Finally, Table 2.10 overviews the 97 Occupied-NS schools that did not take the 2011 SIMCE but did take the November 6th, 2012 SIMCE for the shortest diagonal cohort of 10th grade, which is 2010. The DiD estimate shows that before and after, the average test results of the Occupied-NS schools are 3.24 and 4.89 percentile lower relative to the control group, indicating a peak in decrement .

The results for the first row show that the pre-treatment test results (the 2010 SIMCE for grade 10th of secondary school, but two years before the post-treatment test in languages and math) are lower in the 97 Occupied-NS schools. The difference is strongly significant at the 5% level. During the period from June 2011 to the date of the second test after the onset of the revolts on November 6th, 2012, at least an average of 48.08 days (or over 2 calendar months) of normal school activities were lost in the schools that appeared on the police report list of occupied schools. Again, this loss took place at the end of June at the climax of the occupations, and these

schools did not take the 2011 SIMCE. For the second row, the post-treatment test results for the Occupied-NS schools are more significantly lower than those of the control group. The third row shows that the difference in test results from pre- to post-treatment years is slightly positive in the control group. For the Occupied-NS, the two test results plunge with a strongly significant drop in scores. Under the parallel trend assumption, this means that there is sufficient evidence to reject the *null* hypothesis that lost school days during the school occupations do not affect pupil performance on standardized tests. As a result, this study's main result holds, at least for the Occupied-NS school type. There is strong evidence that school occupations significantly decrease the performance of these schools' pupils on their standardized tests using a school level grade cohort approach. This effect is both significant and sizeable, representing an almost 0.05 decrease in standardized test scores. This result is very close to the pupil-level approach and a particularly strong one.

Naïve estimators

For these school-level estimates, the naïve estimator is the second row of each DiD estimator. Table 2.11 summarizes the results.

It is important to note that different cohorts bear different biases. In this case, they are mainly upward-biased and move in a direction opposite to that of the pupil-level averages, which are mainly downward-biased. This is because pupil-level cohorts do not have a one-to-one correspondence to school-level cohorts because the latter is the average of the former. For instance, suppose there are just two schools, one with only one pupil in a percentile of 40th and the other, a larger school, with nineteen pupils all in the same lower percentile of 20th. The school-level average is a percentile of 30th while the pupil-level average has a percentage of only 21th. This variation happens because bias exists in different directions with the same DiD estimates.

Regression with controls

This section describes alternative approaches using reduced form models similar to what was employed in the pupil-level approach. This section focuses on Occupied-NS schools and math tests, which had the greatest DiD estimates of -4.89. The first alternative approach is shown in Table 2.11, a regression of the difference in the math test percentile before and after the treatment (2010 and 2012 SIMCEs) and the binary variable for occupied-NS schools. This strategy yields an estimate very close to the DiD estimates after applying 10 controls. The controls include the previous 2006 exam, books, HH income, being indigenous, mother and father's education, pc at home, internet at home, preschool and repeating a year.

2.4.3 Robustness check

Parallel trends

To implement DiD estimates, it is first necessary to give credibility to the parallel trends assumption. This assumption is untestable. Moreover, the pupil cohorts are observed just once before the treatment. Thus, indirect testing is not possible. Nevertheless, a similar test can be implemented for school cohorts by simply comparing the trends for different school-level cohorts who sat for the SIMCE in 10^{th} and 8^{th} grades.

A full sample for 10^{th} grade pupils shows that the truth of the parallel trends assumption is highly probable, see Figure 2.3. From 2001 to 2010, five points with their line segments between them follow a strict parallel trends pattern. Then there is a divergence in 2011, but again the trend returns. As mentioned above, I am not proving parallel trends, because this assumption is untestable. Parallel trends represent a "what if" question and will remain unanswered as a consequence of the

fundamental problem of causal inference. In summary, from 2001 to 2013, there is an absolute parallelism which is broken only in the treated units just before and after the treatment period. These findings reaffirm that—at least for 10^{th} grade pupils—DiD is the most effective option.

A full sample of 8^{th} grade pupils shows that the parallel trends assumption is somehow less likely in this case, see Figure 2.4. From 2000 to 2013, there is merely an approximate parallelism in the different cohorts who sat for the 8^{th} grade test. These findings suggest that for 8^{th} grade pupils, the DiD results should be treated with caution. At least one robust non-compliance method should be attempted with the parallel trends assumption, such as matching or the synthetic control method.

Although parallel trend is an untestable assumption, credibility is added to its validity in this study by performing a falsification exercise. This exercise first repeats the DiD estimates. However, they are repeated for the SIMCE tests held in a full previous pupil cohort, which are SIMCEs from the 4^{th} grade level in 2002 and 10^{th} grade in 2008. In short, the third line of the DiD tables is reported in Table 2.12. The results are self-affirming, because the DiD estimates are strictly insignificant with an almost negligible absolute value. This cannot prove the parallel trend assumption. However, it both justifies confidence in such a trend and provides a benchmark against which the principal results can be tested.

Synthetic control

For 8^{th} grade pupils, the fact that the parallel trends assumption is broken suggests that another method, such as synthetic control, should be enforced. The synthetic control method was introduced by Diamond et al. (2010) and implemented by Abadie et al. (2011). It is of special interest Abadie et al. (2003)'s R implementation paper on the Basque economy under terrorism. This method could be implemented, closely following that example after some R coding (see Appendix II for full code) and

using twenty-one predictors. The twenty-one predictors include SIMCE 10^{th}, repeat, study time, income, time to home, indigent, poor, unemployment, indigenous, pc, internet, President scholarship, indigenous scholarship,[4] tuition scholarship, other scholarship, public insurance a, public insurance b, public insurance c, public insurance d,[5] income with subsidies, and percentage public schools. The summary statistics of the final variables are in Table 2.13.

The simple average of the SIMCE math and language results are used as the variable to be predicted with the synthetic model. Also, the method is slightly modified through building a panel of 7 years (2000, 2004, 2007, 2009, 2011, 2012, 2013) and 52 municipalities plus the hypothetical territory of occupied schools. A control is then synthesized from the 52 municipalities.

The result shows a strict matching of occupied schools and controls before the treatment year, a sharper decrease in 2011, and a return to the first trend after the treatment year. The estimated treatment effect (or the difference between the occupied schools and the synthetic control in the treatment year) is -4.6. This treatment effect is significant at any significance level and implies a 0.1 reduction in the standard deviation in the test results, being approximately two thirds of the DiD estimates for 8^{th} grade (0.18 v/s 0.1 reduction on a standard deviation in test results). All of this implies that missed school days have a cost, and that the estimation of this cost is consistent between the several methods presented here.

[4]The President scholarship is a scholarship given to high performance poor students and the indigenous scholarship is a scholarship given to indigenous people.

[5]The Chilean public health insurance which includes 90% of the population is divided into 4 categories: a, b, c and d according to income. a are homeless people while d working people with high earnings. There are also many private insurances that cover the other 10% of the population. Mostly affluent workers.

Matching: Midway between naïve and DiD estimators

From Table 2.11, it can be seen that the naïve or 2012 same-year estimate of the effect of the treatment for Occupied-NS schools is -20.07 percentile points on the 10^{th} grade 2012 SIMCE. This is due to lower ability in Occupied-NS pupils, which is eliminated in the DiD estimates. To improve upon the naïve or same year 2012 estimate, it is necessary to control for covariates in the same spirit that was used for the reduced form models. An estimate including previous test scores, in particular 4^{th} grade SIMCE in 2006 for the same cohort, plus gender and income of the HHs, gives -6.65 percentile points on the 10^{th} grade 2012 SIMCE. Improving upon these results requires matching. Using the *Matching* R package,[6] the sample is first balanced and then the matching treated group coefficient is estimated at -5.75 percentile points on the 10^{th} grade 2012 SIMCE. The values for the treated group coefficient for Occupied-NS schools are summarized in Table 2.14. It can be seen that the coefficient converges from below from the raw naïve (or same-year comparison of means between the same year SIMCE and the Occupied-NS treated group) of -20.07, to the DiD estimate of -4.89. The coefficient uses the regression with covariates with special interest on the previous year's SIMCE scores and the matching models. It is important to stress the key effect of the previous year scores on the estimates. In fact, the DiD estimate is mainly this effect in its own right.

Finally, Table 2.15 shows the identification strategy of each method used above. Naïve estimation requires independence between the treatment and any function of the potential outcomes, or random experimentation. In these settings, this is difficult to satisfy. The next method is adjusted regression, which implies selection on observables, unconfoundedness or conditional independence. Conditional independence is independence between the treatment and any function of the potential outcome, conditional on some selected covariates. Matching implies the same, but it is done in a local neighbourhood of matched pairs using the propensity score to

[6] Jasjeet S. Sekhon. 2011. *"Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching Package for R."* Journal of Statistical Software, 42(7): 1-52.

reduce the search to a single variable or the Mahalanobis distance. Lastly, the DiD method implies selection on unobservables: γ_s and λ_t are wiped out by the double difference remaining the causal effect δ of $\text{occupy}_{\text{ist}}$ (see the previous subsection The model for details).

Apart from the selection on observables assumption, the matching estimators need another assumption, namely the common support Pr(D=1|X)<1. This is illustrated in Figure 2.6 for the propensity score approach.

2.4.4 Evaluating the identifying assumptions

Can the treatment effect be estimated for this study's sample? It is generally accepted that internal validity fails when there are differences between the treated components and controls (other than the treatment itself). These differences affect the outcome, and this cannot be controlled for. This is tested in general with a favourable result, as seen in Table 2.1. But any unbalance remaining can be dealt with by the identification strategy. In particualr, difference-in-differences has two additional threats to validity. The first threat is credibility of parallel trends, which has been discussed above. The second threat is compositional difference, which is briefly discussed in this subsection. Repeated cross-sections are only valid when the composition of the target population does not change between the two periods. This condition is tested in this study by looking at the distribution of the control and treated groups, which suggests that their distributions are the same before the treatment. Table 2.16 (an elaboration of Table 2.1) shows the compositional difference of the control and occupied groups for control variables. Control and occupied groups are homogeneous with each other before and after the treatment. The comparison between each type before and after evidences the natural growth of the growing variables such as income, pc and internet. Overall, the compositional difference is satisfied.

The SUTVA assumption is crucial. The treatment mechanism of assignment can be individual, executed by third parties inside the model or by a researcher. In this case, there is a mixture of self-decision (a majority vote school decision) and third-party decision (a radical minority decision). The reduction in standardized tests in school A is unaltered regardless of whether school B is occupied. This could be challenged if there is friendship or partnership. If a boy's school and girl's school in proximity to each other are both occupied, it is possible that couples decide, for instance, to go to the beach. If just one school is occupied, students from the partner school may decide to study alone. The reduction in standardized tests for any school may or may not be the same whether it is occupied after a majority vote decision or under a third-party decision.

At the same time, external validity should be contrasted. In other words, can this study's estimates be extrapolated to other populations? This is possible, missed school days is quite a common event in schools all over the world, and this study uncovers its causal effect on test scores. Comparable time deprivations in the context of school occupations, snow storms, natural disasters, most probably have a common path to diminishing test scores. A more general claim to include disruptions in other working environments, in the context of strikes, that involve a country health system and how does this affect the productivity in hospitals or perhaps pilots and airplanes' security needs further studies.

2.4.5 Extension: Missed school days

Figure 2.7 graphically illustrates the lost school days for the pre-treatment when there were no missed school days and during the *Chilean Winter*. The number of missed school days can be identified until the October 19$^{\text{th}}$, 2011 SIMCE and until November 6$^{\text{th}}$ 2012. The graph also shows the intensity of the treatment, which I call the *dose* of the treatment.

Table 2.17 summarizes the key information on the amount of time attended by the control group and lost by the relevant 3 treated groups. From June 2011 to the 8^{th} grade SIMCE Exam on October 19^{th}, 2011, 11.58 days were lost in Occupied-S (occupied schools that sat the 8^{th} grade 2011 October SIMCE test). This represents over two calendar weeks of lost school days. In addition, 40.58 days were lost in Occupied-NS (occupied schools that did not sit the 8^{th} grade 2011 October SIMCE test), or approximately two calendar months of lost school days. From June 2011 to the 10^{th} grade SIMCE exam on November 6th, 2012, the former treatment plus the cancelled days from October 19^{th}, 2011 to November 6^{th}, 2012, 9.56 (the previous 11.58 days minus 2.02 "catch-up" days) days were lost in Occupied-S (occupied schools that sat for the 8^{th} grade 2011 October SIMCE test). This amounts to approximately two weeks of lost school days. During this period, Occupied-S schools actually attended more time relative to the control group. This is probably an attempt to catch up with missing class material. Even so, the overall attendance rate was negative. Finally, 48.08 (the previous 40.58 days plus an additional 7.50 lost days) days were lost in Occupied-NS (occupied schools that did not sit for the 8^{th} grade 2011 October SIMCE test). This was approximately two calendar months and one calendar week of lost school days.

Pupil-level reduced form models

In this section, I describe alternative approaches using reduced form models. This study focuses on Occupied-NS schools and math tests, which have the greatest DiD estimates of -5.40. The first alternative approach is **reg diff ns**, i.e., a regression of the difference (**diff**) in the math test percentile before and after the treatment (the 2006 and 2012 SIMCEs) and the binary variable (**ns**) for Occupied-NS schools. The results are statistically significant and similar to the stand-alone model.

The second alternative approach used is a dose-response treatment. This study uses an IV approach. First, a number of missed days variable is built by setting the

attendance at the 90th percentile and taking the difference between this number, 278, and the real attended days of each school.

$$\text{days missed} = (\text{attend}^{90^{\text{th}}\text{pctile}} - \text{attend}_{\text{s}}) \tag{2.4.2}$$

Table 2.18 shows the percentiles of attended and missed school days respectively. The instrument is the occupied school dummy. The sign of the coefficients is the appropriate: more missed days implies fewer test results. In fact, one standard deviation in missed days 14.59 multiplied by -0.711, the coefficient with controls and municipal fixed effect, results in a decrease of -10.37 which is 0.36 of a standard deviation of the test results (10.37/28.86). These calculated effects are similar to the calculated estimate using the DiD estimate (-5.40). Table 2.19 summarizes the above results. And the following is the IV model constructed in this study:

Reduced Form

$$\text{SIMCE} = \delta \text{ occupied} + \text{controls} \tag{2.4.3}$$

First Stage

$$\text{days missed} = \gamma \text{ occupied} + \text{controls} \tag{2.4.4}$$

Second Stage

$$\text{SIMCE} = \beta \widehat{\text{days missed}} + \text{controls} \tag{2.4.5}$$

School-level reduced form models

One standard deviation in missed days 14.59 multiplied by -0.826 (the coefficients with controls and the municipal fixed effect) results in a decrease of -12.05 which is 0.42 of a standard deviation of the test results (12.05/28.86). These calculated effects are similar to the estimate calculated using the DiD estimate (-5.40). Table 2.20 summarizes the above results.

Exclusion restriction

When using IV, Z (the instrument or occupied schools) must only affect D (the treatment or missed days) and not Y (the outcome or the SIMCE). There are two caveats. One caveat is that the treated and control groups are not equilibrated, meaning that there is an effect of Z on the known covariates. For instance, income is not equilibrated, meaning the occupied dummy causes variation in income, which on its own causes variation in the SIMCE. This is solved by conditioning on income. The second caveat is the undetermined effect, which means that the occupied dummy has an effect on the SIMCE through other covariates not included in the model. Such a covariate, for example, could be an ideological channel.

2.5 Conclusion

This chapter has addressed the cost of lost school days using an external variation engaging in social disruptive activities. The average effect is a decrease of 5.40 percentile points in standardized tests, a 0.18σ in performance.

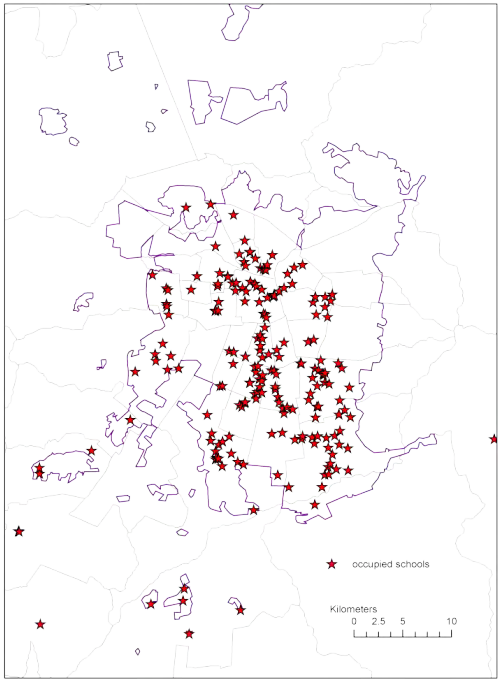
As stated above, compared with other effects such as the increase in value-added by better teacher quality Hanushek (2003)—0.11σ—and comparable literature values on snow days disruption Goodman (2014)—0.05σ, length of school day in Colombia Hincapié (2016)—0.1σ—and length of school year in Germany Pischke (2007)—more repetitions and less continuing education but the same earnings and employment, the decision to miss (and force others to miss) academic work and go on strike is similarly deleterious to losing one and half standard deviations in teacher quality during the same period. To put it in an explicit way, if the Instituto Nacional—whose pupils were leading the protests and walkouts—was suddenly completely deprived of its top teachers, education would suffer from a similar drop

in value added as it did in the years 2011 and 2012 when it was on strike with its students involved in walk-outs and protests. Therefore, the production function of education in Chile was seriously impeded due to school occupation.

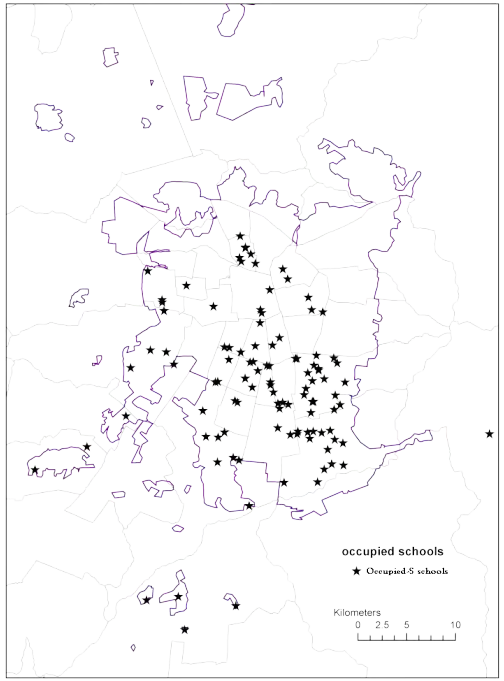
This research is part of the causality literature implementing various techniques to attempt to establish a causal relationship among the relevant variables in use. Comparing the different methods included in this research, the pupil-level Occupied-NS largest DiD estimate drop in 5.40 percentiles points can be compared to the school-level estimate of 4.89. At the same time, the DiDiD estimates also show a relevant decrease in school performance for pupils participating in these protests from Occupied-NS 10^{th}-4^{th} grades, at 5.78 percentile points or 0.19 standard points. The synthetic control method shows a decrease in 4.6 percentile points (or 0.1 of a standard deviation) in the results of the occupied schools. All DiD, DiDiD and synthetic control estimates are significantly similar to each other, which confirms the robustness of the estimators. Moreover, they are all strongly statistically significant. Additionally, this study verified the robustness of these estimates and discussed the validity of the parallel trend assumption and the DiD and DiDiD estimates.

In summary, this research focused on, and hopefully identified, the cost of missing school days in the context of students immersed in the 2011/2012 Chilean student riots called the *Chilean Winter*. Missed school days can be tracked as causal effects which decrease student standardized test performance. This study does not claim that student protest is a negative activity or that students should not involve themselves in these activities, but merely seeks to disclose the relevant average cost of lost school days. This research could assist both interested scholars and students in recognizing and balancing the real costs of lost school days due to specific protest and riot activities.

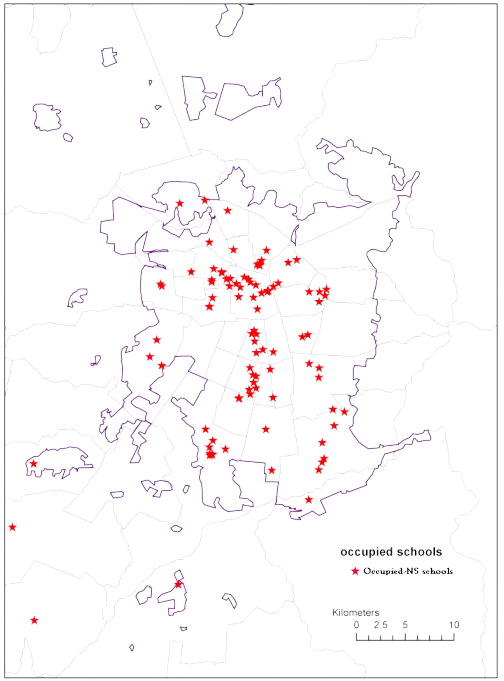
Figure 2.1: Occupied schools



(a) Occupied schools.



(b) Occupied-S schools.



(c) Occupied-NS schools.

Figure 2.2: Outcome: The SIMCE (The Chilean standardized test)

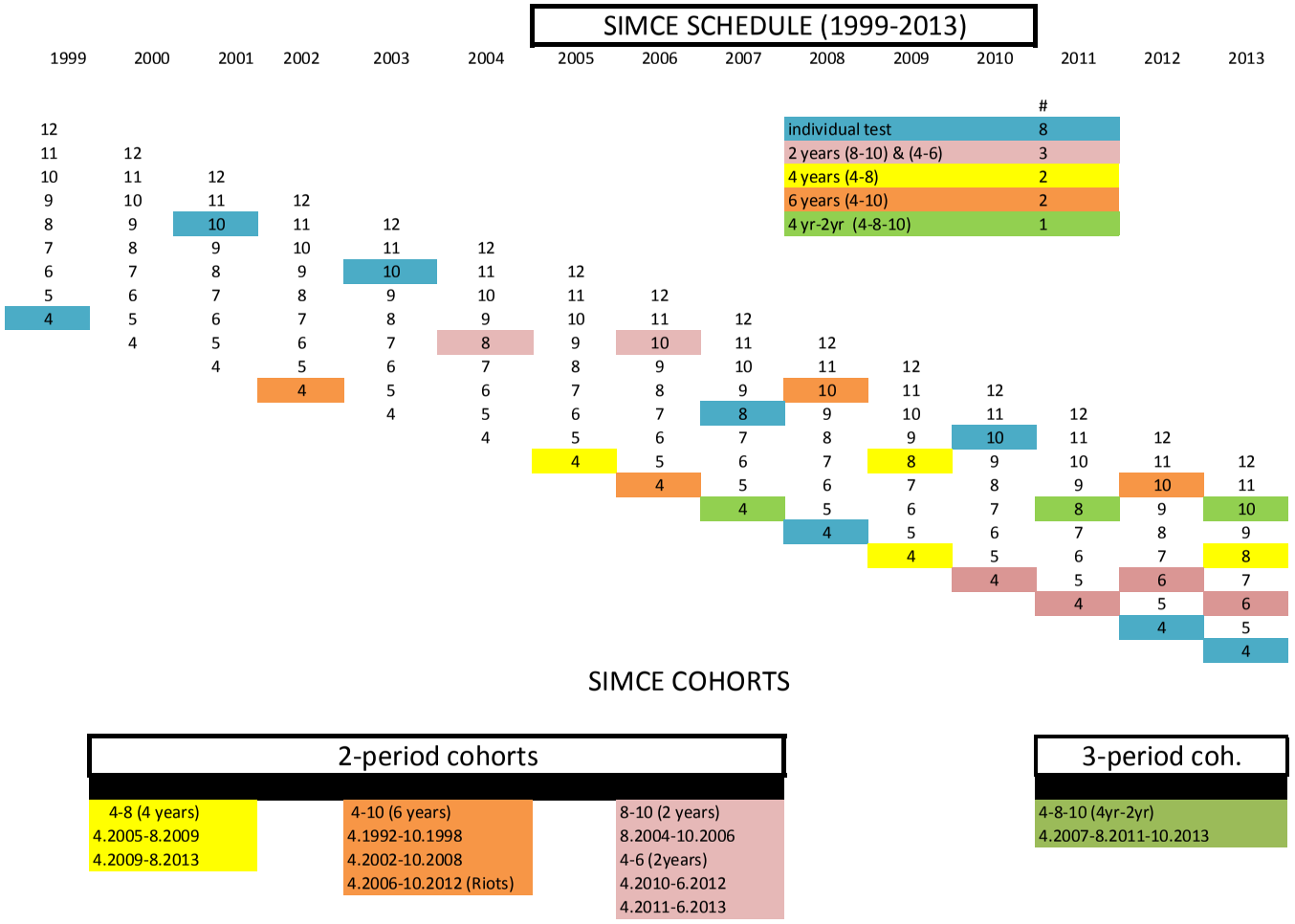


Figure 2.3: Parallel Trends: 10^{th} grade

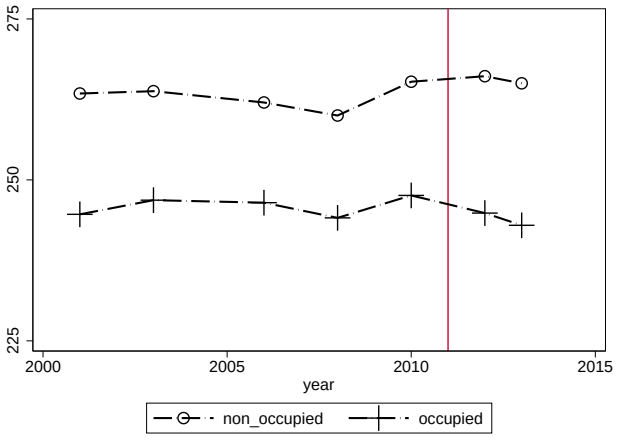


Figure 2.4: Non parallel trends: 8^{th} grade

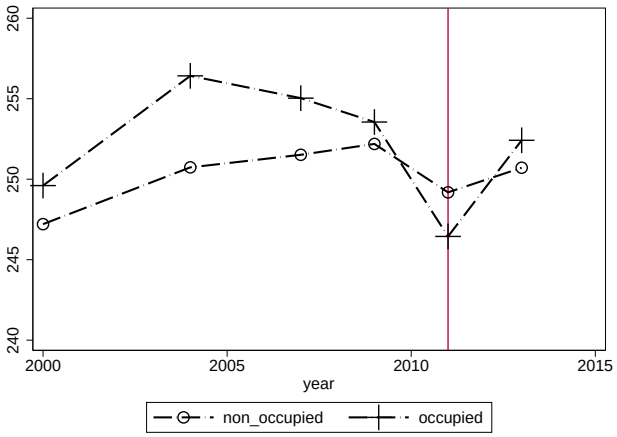


Figure 2.5: Synthetic control: 8^{th} grade

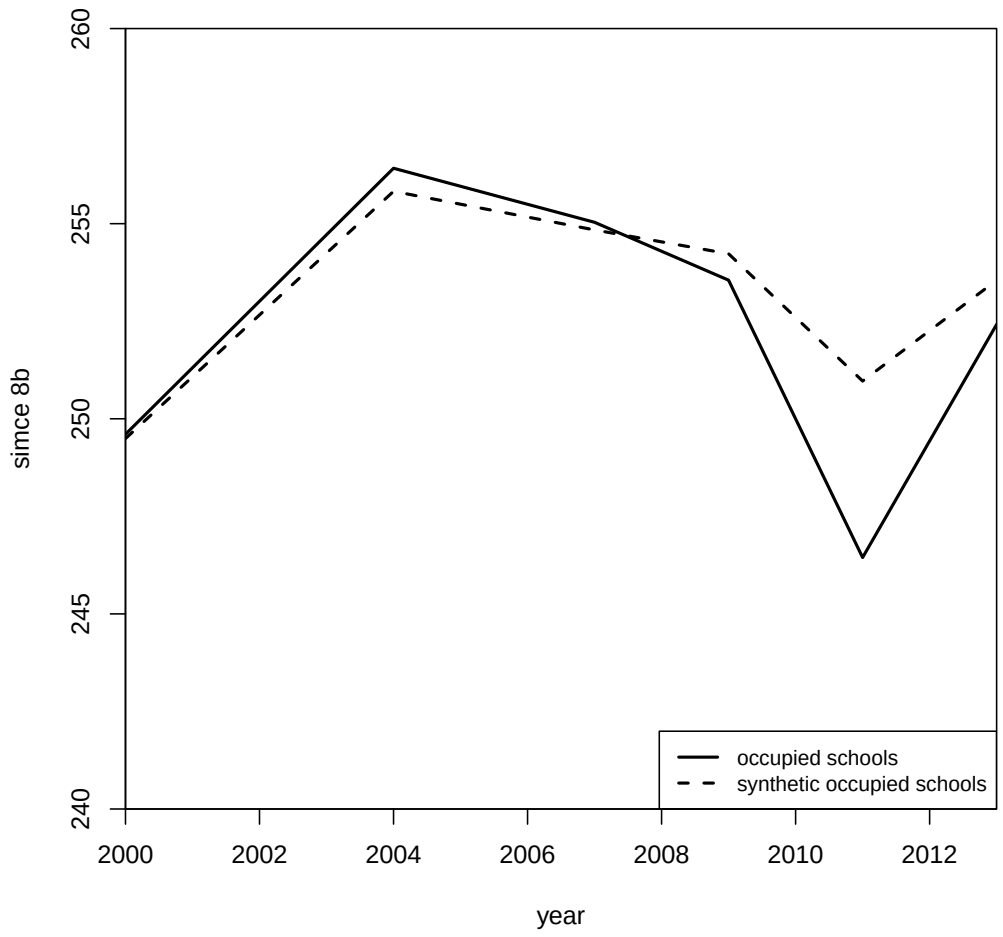


Figure 2.6: Common support

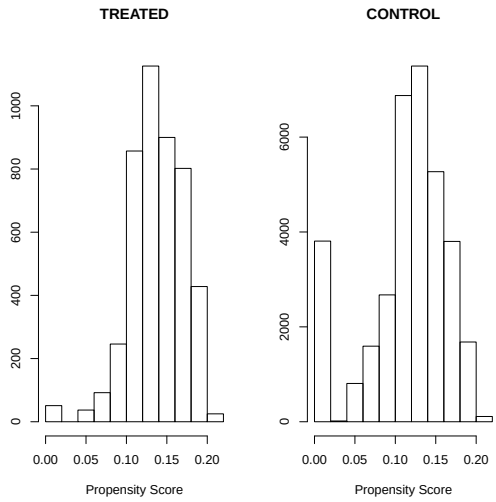


Figure 2.7: Monthly lost school days, (Mar 2011/Nov 2012)

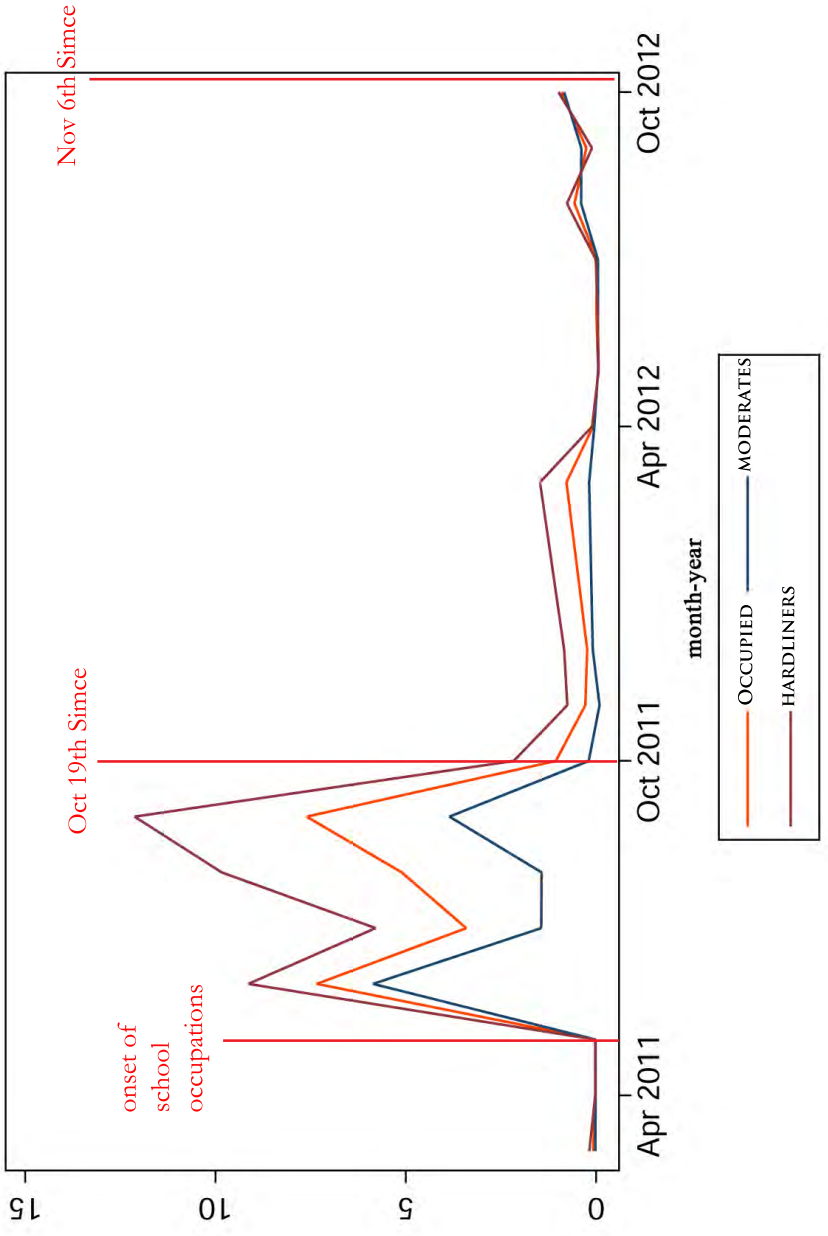


Table 2.1: Comparison statistics control group, occupied-S and occupied-NS

Variable	Control Group	Occupied-S	Occupied-NS
books	42.85	49.94	64.79
income	500.15	566.81	837.93
mother indigenous	0.08	0.07	0.08
father education	12.07	12.34	12.77
mother education	11.78	12.11	12.35
pc	0.73	0.77	0.79
internet	0.57	0.62	0.67
years of preschool	1.30	1.28	1.26
repeat year	0.48	0.45	0.44
percentage public voucher schools	25.88	44.97	61.58
percentage private voucher schools	67.41	53.91	38.42
percentage private schools	6.71	1.12	0
SIMCE perc. pre-treatment Lang. 4^{th}	50.51	48.41	49.08
SIMCE perc. pre-treatment Math 4^{th}	50.43	49.01	50.99
SIMCE perc. Language 10^{th}	50.67	47.05	47.09
SIMCE perc. Math 10^{th}	50.74	47.97	45.90
SIMCE perc. difference 10^{th}-4^{th} Lang.	0.15	-1.36	-1.99
SIMCE perc. difference 10^{th}-4^{th} Math	0.31	-1.04	-5.09

[1] Author's calculation from research data provided by the Chilean Ministry of Education's Agency for Quality in Education. The data is available only to researchers after submitting a written proposal.

[2] The table shows the summary statistics of the control group, occupied-S and occupied-NS. It lists the type of school, pre- and post-SIMCE, difference in percentiles and control variables.

Table 2.2: Pupil-level 8th11-4th07 Occupied-S value-added DiD estimates, Oct. 19th 2011

Pupil-level Occupied-S schools

11.58 lost days (over 2 calendar weeks)

	Language test			Math test			Science test		
	control and occupied schools			control and occupied schools			control and occupied schools		
	CG	Occupied-S	Difference Occupied-S/CG	CG	Occupied-S	Difference Occupied-S/CG	CG	Occupied-S	Difference Occupied-S/CG
SIMCE 4th grade '07,	50.50	50.16	-0.34	50.44	52.17	1.73	50.43	52.33	1.90
	(0.07)	(0.38)	(0.39)	(0.07)	(0.38)	(0.39)	(0.07)	(0.38)	(0.38)
SIMCE 8th grade '11,	50.55	48.65	-1.90	50.51	50.00	-0.51	50.52	49.80	-0.72
	(0.07)	(0.39)	(0.39)	(0.07)	(0.38)	(0.38)	(0.07)	(0.39)	(0.40)
Change in mean,	0.05	-1.51	-1.56***	0.07	-2.16	-2.23***	0.08	-2.54	-2.62***
	(0.05)	(0.32)	(0.32)	(0.05)	(0.30)	(0.30)	(0.05)	(0.31)	(0.32)

[1] Author's calculation from research data provided by the Chilean Ministry of Education's Agency for Quality in Education. The data is available only to researchers after submitting a written proposal.

[2] The table shows the Difference-in-Differences (DiD) estimates of average test results before and after the school occupations and protests of 2011 which amount to at least 11.58 (over 2 calendar weeks) lost school days from normal activities since the on-set of the pupil protests in the so-called *Chilean Winter* in the early days of June of 2011. The results are shown by schools and are calculated at a pupil level. The first two rows show all pupil observations and the third the balanced sample of pupils who sat for both the 2007 and 2011 SIMCEs. The table shows the average test result difference between the control group (2,197 schools) and the occupied schools that sat for the 8th grade 2011 SIMCE on October 19th of that year (108 schools of the 205 police report list). The other 97 schools did not sit for the test so they are not considered. The test results are shown disaggregated by subject (language, math and science).

Table 2.3: Pupil-level 10^{th}12-4^{th}06 Occupied-S value-added DiD estimates, Nov. 6^{th} 2012

	Pupil-level Occupied-S schools					
	9.56 lost days (approximately 2 calendar weeks)					
	Language test			Math test		
	control and occupied schools			control and occupied schools		
	CG	Occupied-S	Difference Occupied-S/CG	CG	Occupied-S	Difference Occupied-S/CG
SIMCE 4^{th} grade '06,	50.51	48.41	-2.10	50.51	49.01	-1.50
	(0.07)	(0.45)	(0.45)	(0.07)	(0.40)	(0.41)
SIMCE 10^{th} grade '12,	50.61	47.05	-3.55	50.57	47.97	-2.60
	(0.07)	(0.40)	(0.41)	(0.07)	(0.40)	(0.40)
Change in mean,	0.09	-1.36	-1.45***	0.06	-1.04	-1.10***
	(0.07)	(0.34)	(0.35)	(0.06)	(0.32)	(0.33)

[1] Author's calculation from research data provided by the Chilean Ministry of Education's Agency for Quality in Education. The data is available only to researchers after submitting a written proposal.

[2] The table shows the Difference-in-Differences (DiD) estimates of average test results before and after the school occupations and protests of 2011/2012, which amount to at least 9.56 (approximately 2 calendar weeks) lost school days from normal activities since the on-set of the pupil protests. The table shows the average test result difference between the control group (2,197 schools) and the occupied schools that sat for the 8^{th} grade 2011 SIMCE on October 19th of that year and for the current November 6^{th} 2012 SIMCE test (108 Occupied-S schools of the 205 police report list). The other 97 schools did sit for the current test and are considered in the next table. The test results are shown disaggregated by subject (language and math).

Table 2.4: Pupil-level 10^{th}12-4^{th}06 Occupied-NS value-added DiD estimates, Nov. 6^{th} 2012

Pupil-level Occupied-NS schools

48.08 lost days (over 2 calendar months and 1 week)

	Language test			Math test		
	control and occupied schools			control and occupied schools		
	CG	Occupied-NS	Difference Occupied-NS/CG	CG	Occupied-NS	Difference Occupied-NS/CG
SIMCE 4^{th} grade '06,	50.51	49.08	-1.43	50.43	50.99	0.56
	(0.08)	(0.33)	(0.34)	(0.08)	(0.33)	(0.34)
SIMCE 10^{th} grade '12,	50.67	47.09	-3.58	50.74	45.90	-4.84
	(0.08)	(0.34)	(0.35)	(0.08)	(0.33)	(0.34)
Change in mean,	0.15	-1.99	-2.14***	0.31	-5.09	-5.40***
	(0.07)	(0.27)	(0.28)	(0.06)	(0.26)	(0.27)

[1] Author's calculation from research data provided by the Chilean Ministry of Education's Agency for Quality in Education. The data is available only to researchers after submitting a written proposal.

[2] The table shows the Difference-in-Differences (DiD) estimates of average test results before and after the school occupations and protests of 2011/2012, which amount to at least 48.08 (over 2 whole calendar months and 1 week) lost school days from normal activities since the on-set of the pupil protests. The table shows the average test result difference between the control group (2,197 schools) and the occupied schools who sat only for the 10^{th} grade 2012 SIMCE on November 6^{th} of 2012, but not for the 2011 SIMCE (97 schools of the 205 police report list). The test results are shown disaggregated by subject (language and math).

Table 2.5: Effects of school occupation on student test scores: Regression estimates

	lang			math			science		
	naïve	reg+ctls	DiD	naïve	reg+ctls	DiD	naïve	reg+ctls	DiD
Occupied-S	-1.90***	-4.55*	-1.56***	-0.51	-2.97	-2.23***	-0.72*	-4.54	-2.62***
8^{th}11-4^{th}07	(0.39)	(2.70)	(0.32)	(0.38)	(3.18)	(0.30)	(0.40)	(2.96)	(0.32)
Occupied-S	-3.55***	-1.49***	-1.45***	-2.60***	-1.81	-1.10***			
10^{th}12-4^{th}06	(0.41)	(0.60)	(0.35)	(0.40)	(1.21)	(0.33)			
Occupied-NS	-3.58***	-4.03**	-2.14***	-4.84***	-4.07***	-5.40***			
10^{th}12-4^{th}06	(0.35)	(1.55)	(0.28)	(0.34)	(1.53)	(0.27)			

[1] Author's calculation from research data provided by the Chilean Ministry of Education's Agency for Quality in Education. The data is available only to researchers after submitting a written proposal.

[2] The table shows the naïve, the regresssion with controls and the Difference-in-Differences (DiD) estimates of average test results before and after the school occupations and protests of 2011/2012. The naïve estimators are a by-product of the DiD being the mean of the treated minus the mean of the untreated *after* the treatment, so it is the second row of each DiD estimator. The 9 controls included are books, HH income, being indigenous, mother and father education, pc at home, internet at home, preschool and repeating a year. The test results are shown disaggregated by subject (language, math and science).

Table 2.6: DiD for value-added: Pre, during and post

	Occupied-S 8^{th}-4^{th}			Occupied-S 10^{th}-4^{th}		Occupied-NS 10^{th}-4^{th}	
	language	math	science	language	math	language	math
POST	-0.94**	-0.10	-0.80*	-1.44***	-1.73***	-1.80***	-4.19***
DURING	-1.56***	-2.23***	-2.62***	-1.45***	-1.10***	-2.14***	-5.40***
PRE	-0.63	-0.09	-0.40	-0.38	0.25	0.50	0.38
DiDiD	-0.93**	-2.14***	-2.22***	-1.07**	-1.35***	-2.64***	-5.78***

[1] The table shows the Difference-in-Difference-in-Difference (DiDiD) estimates for value-added results pre, during and post the school occupations and protests. The left part of the table shows Occupied-S 8-4 (pre: 8^{th} 2009-4^{th} 2005; during: 8^{th} 2011-4^{th} 2007; post: 8^{th} 2013-4^{th} 2009), the middle part Occupied-S 10-4 (pre: 10^{th} 2008-4^{th} 2002; during: 10^{th} 2012-4^{th} 2006; post: 10^{th} 2013-4^{th} 2007), the right part of the table Occupied-NS 10^{th}-4^{th} (pre: 10^{th} 2008-4^{th} 2002; during: 10^{th} 2012-4^{th} 2006; post: 10^{th} 2013-4^{th} 2007). Test subjects are language, math and science (just for 8^{th}-4^{th}).

Table 2.7: DiDiD in the three wave cohort

	Occupied-S 8^{th}-4^{th}	
	language	math
$8^{th}11$-$4^{th}07$	-1.56***	-2.23***
$10^{th}13$-$8^{th}11$	-0.35	-0.08
DiDiD	-1.21***	-2.15***

[1] The table shows the Difference-in-Difference-in-Difference (DiDiD) for value-added results for the only three wave cohort ($4^{th}07$-$8^{th}11$-$10^{th}13$). This gives me the opportunity to control for unobserved divergent or convergent trends in a single cohort or three wave series of data. This series is also important because it pertains to the exact moment of high protests. Data for Occupied-S protester is only available because Occupied-NS did not sit for the SIMCE 8^{th} grade test. Test subjects are language and math because science is not available for upper grades(10^{th}).

Table 2.8: School-level 8th grade '11-'09 Occupied-S non value-added DiD estimates, Oct. 19th 2011

	School-level Occupied-S schools								
	11.58 lost days (over 2 calendar weeks)								
	Language test			Math test			Science test		
	control and occupied schools			control and occupied schools			control and occupied schools		
	CG	Occupied-S	Difference Occupied-S/CG	CG	Occupied-S	Difference Occupied-S/CG	CG	Occupied-S	Difference Occupied-S/CG
SIMCE 8th grade '09,	50.31	46.02	-4.29	50.59	45.37	-5.22	50.33	46.86	-3.48
	(0.74)	(2.62)	(2.73)	(0.74)	(2.64)	(2.74)	(0.74)	(2.68)	(2.78)
SIMCE 8th grade '11,	50.37	44.77	-5.60	50.30	45.75	-4.55	50.23	44.97	-5.25
	(0.74)	(2.58)	(2.68)	(0.74)	(2.60)	(2.71)	(0.74)	(2.62)	(2.72)
Change in mean,	0.07	-1.25	-1.31	-0.28	0.39	0.67	-0.11	-1.89	-1.78
	(0.49)	(1.87)	(1.93)	(0.43)	(1.44)	(1.50)	(0.44)	(1.69)	(1.75)

[1] The table shows the Difference-in-Differences (DiD) estimates of average test results before and after the school occupations and protests of 2011 which amount to at least 11.58 (over 2 calendar weeks) lost school days from normal activities since the on-set of the pupil protests. The table shows the average test result difference between the control group (2,197 schools) and the occupied schools that sat for the 8th grade '09 and '11 SIMCEs. The other 97 Occupied-NS schools did not sit for the test so they are not considered. The test results are shown disaggregated by subject (language, math and science).

Table 2.9: School-level 10^{th} grade '12-'10 Occupied-S non value-added DiD estimates, Nov. 6^{th} 2012

	Language test			Math test		
	control and occupied schools			control and occupied schools		
	CG	Occupied-S	Difference Occupied-S/CG	CG	Occupied-S	Difference Occupied-S/CG
School-level Occupied-S schools						
9.56 lost days (approximately 2 calendar weeks)						
SIMCE 10^{th} grade '10,	52.37	36.41	-15.96	52.11	38.18	-13.93
	(0.98)	(2.50)	(2.69)	(0.98)	(2.56)	(2.75)
SIMCE 10^{th} grade '12,	51.90	37.70	-14.21	52.02	37.55	-14.47
	(0.99)	(2.57)	(2.76)	(0.99)	(2.54)	(2.73)
Change in mean,	-0.47	1.29	1.75	0.10	0.63	0.53
	(0.46)	(1.41)	(1.48)	(0.36)	(1.19)	(1.24)

[1] The table shows the Difference-in-Differences (DiD) estimates of average test results before and after the school occupations and protests of 2011/2012, which amount to at least 9.56 (approximately 2 calendar weeks) lost school days from normal activities since the on-set of the pupil protests. The table shows the average test result difference between the control group (2,197 schools) and the occupied schools that sat for the 8^{th} grade 2011 SIMCE on October 19th of that year and for the current November 6^{th} 2012 SIMCE test (108 Occupied-S schools of the 205 police report list). The other 97 schools did sit for the current test and are considered in the next table. The test results are shown disaggregated by subject (language and math).

Table 2.10: School-level 10^{th} grade '12-'10 Occupied-NS non value-added DiD estimates, Nov. 6^{th} 2012

School-level Occupied-NS schools						
48.08 lost days (over 2 calendar months and 1 week)						
	Language test			Math test		
	control and occupied schools			control and occupied schools		
	CG	Occupied-NS	Difference Occupied-NS/CG	CG	Occupied-NS	Difference Occupied-NS/CG
SIMCE 10^{th} grade '10,	52.14	36.80	-15.34	52.08	36.90	-15.18
	(0.97)	(3.07)	(3.22)	(0.97)	(3.00)	(3.16)
SIMCE 10^{th} grade '12,	52.15	33.57	-18.58	52.37	32.30	-20.07
	(0.97)	(2.99)	(3.15)	(0.98)	(2.74)	(2.91)
Change in mean,	0.01	-3.23	-3.24**	0.29	-4.60	-4.89***
	(0.47)	(1.17)	(1.26)	(0.37)	(0.94)	(1.01)

[1] The table shows the Difference-in-Differences (DiD) estimates of average test results before and after the school occupations and protests of 2011/2012, which amount to at least 48.08 (over 2 whole calendar months and 1 week) lost school days from normal activities since the on-set of the pupil protests. The table shows the average test result difference between the control group (2,197 schools) and the occupied schools who sat only for the 10^{th} grade 2012 SIMCE on November 6^{th} of 2012, but not for the 2011 SIMCE (97 schools of the 205 police report list). The test results are shown disaggregated by subject (language and math).

Table 2.11: School-level effects of school occupation on student test scores: Regression estimates

	lang			math			science		
	naïve	reg+ctls	DiD	naïve	reg+ctls	DiD	naïve	reg+ctls	DiD
Occupied-S	-5.60**	-0.99	-1.31	-4.55*	-2.97**	0.67	-5.25*	-1.64	-1.78
$8^{th}11$-$8^{th}09$	(2.68)	(1.95)	(1.93)	(2.71)	(1.51)	(1.50)	(2.72)	(1.75)	(1.75)
Occupied-S	-14.21***	-0.87	1.75	-14.47***	-1.66	0.53			
$10^{th}12$-$10^{th}10$	(2.76)	(2.60)	(1.91)	(2.73)	(1.21)	(1.24)			
Occupied-NS	-18.58***	-1.69	-3.24**	-20.07***	-6.65***	-4.89***			
$10^{th}12$-$10^{th}10$	(3.15)	(3.46)	(1.26)	(2.91)	(2.52)	(1.01)			

[1] Author's calculation from research data provided by the Chilean Ministry of Education's Agency for Quality in Education. The data is available only to researchers after submitting a written proposal.

[2] The table shows the naïve, the regresssion with controls and the Difference-in-Differences (DiD) estimates of average test results before and after the school occupations and protests of 2011/2012 for the school-level effects of school occupatiion. The naïve estimators are a by-product of the DiD being the mean of the treated minus the mean of the untreated *after* the treatment, so it is the second row of each DiD estimator. The 10 controls included are previous 2006 exam, books, HH income, being indigenous, mother and father education, pc at home, internet at home, preschool and repeating a year. The test results are shown disaggregated by subject (language, math and science).

Table 2.12: Falsification exercise. Previous treatment cohort. Occupied-S and Occupied-NS pupil-level

	NO missed days					
	Language test			Math test		
	CG	Occupied-S	Difference Occupied-S/CG	CG	Occupied-S	Difference Occupied-S/CG
Occupied-S. Change in mean,	-0.45	-0.50	-0.05	-0.51	-0.51	-0.00
	(0.30)	(0.21)	(0.37)	(0.21)	(0.30)	(0.37)
	CG	Occupied-NS	Occupied-NS/CG	CG	Occupied-NS	Occupied-NS/CG
Occupied-NS. Change in mean,	-0.43	-0.41	-0.02	-0.51	-0.51	-0.00
	(0.30)	(0.25)	(0.38)	(0.30)	(0.21)	(0.37)

[1] The table shows the Difference-in-Differences (DiD) estimates as a falsification exercise calculating average test results before and after NO treatment (on average the control group and each subdivisions of the *treated group* had the same number of lost school days). These estimates were calculated for the SIMCE 2002 and 2008 at least three years before the on-set of the protests.

Table 2.13: Summary statistics

Variable	Mean	Std. Dev.	N
SIMCE 8^{th}	250.075	14.767	318
SIMCE 10^{th}	252.68	21.949	370
repeat	0.189	0.051	159
study time	6.209	0.489	53
income	431.643	373.608	212
time to home	17.961	3.121	53
indigent	0.032	0.022	208
poor	0.103	0.053	208
unemployment	0.093	0.041	208
indigenous	0.033	0.021	208
pc at home	0.374	0.277	208
internet at home	0.063	0.067	208
President scholarship	0.002	0.003	208
indigenous scholarship	0.001	0.001	155
tuition scholarship	0.001	0.001	102
other scholarship	0.005	0.005	208
public insurance a	0.245	0.109	208
public insurance b	0.217	0.065	208
public insurance c	0.119	0.041	208
public insurance d	0.088	0.029	208
subsidy income	8.68	6.859	208
percentage public schools	0.437	0.216	213

Table 2.14: Treated group coefficient on several methods

method	coefficient on Occupied-NS occupied schools
naïve estimator (same year scores on treated group)	-20.07
plus covariates (previous scores, sex, income)	-6.65
plus matching estimator (propensity score)	-5.75
plus matching estimator (mahalanobis)	-5.12
DiD estimator	-4.89

Table 2.15: Method identification strategy

method	identification strategies
naïve estimator (same year scores on treated group)	(Y1,Y0)$\perp\!\!\!\perp$ D, D$\rightarrow$Y independence, causal effect
plus covariates (adjusted-regression)	[(Y1,Y0)$\perp\!\!\!\perp$ D]|X selection on observables, unconfoundedness, conditional independence
plus matching estimator (propensity score)	[(Y1,Y0)$\perp\!\!\!\perp$ D]|Pr(D=1|X) locally selection on observables
plus matching estimator (mahalanobis)	[(Y1,Y0)$\perp\!\!\!\perp$ D]|MAHALANOBIS(X) locally selection on observables, unconfoundedness, conditional independence
DiD estimator	$\gamma_s, \lambda_t, \delta, \text{occupy}_{ist}$ selection on unobservables

Table 2.16: Compositional difference: Occupied group before and after treatment

	Before, 4		After, 10	
Variable	Control group	Occupied	Control group	Occupied
books	32.83	36.33	42.85	52.15
income	362.03	325.07	500.15	607.25
mother indigenous	0.07	0.05	0.08	0.07
father education	12.36	12.95	12.07	12.40
mother education	12.20	12.80	11.78	12.15
pc	0.43	0.52	0.73	0.77
internet	0.20	0.23	0.57	0.63

[1] Author's calculation from research data provided by the Chilean Ministry of Education's Agency for Quality in Education. The data is available only to researchers after submitting a written proposal.

[2] The table shows the compositional difference of the control group and the occupied group. It lists the control variables pre- and post-treatment.

Table 2.17: Average attended for control group (CG) and lost school days for treated group (Occupied, Occupied-S and Occupied-NS for the relevant periods)

	June-until '11 SIMCE				
	obs	level	s.e	t-test	p-value
CG	2,589	80.83***	0.2268	353.33	0.000
Occupied	205	-26.09***	0.837	-31.16	0.000
Occupied-S	108	-11.58***	1.290	-8.97	0.000
Occupied-NS	97	-40.58***	1.161	-34.83	0.000
	June-until '12 SIMCE				
	obs	level	s.e	t-test	p-value
CG	2,589	263.64***	0.648	406.64	0.000
Occupied	205	-28.80***	2.410	-11.95	0.000
Occupied-S	108	-9.56***	3.330	-2.87	0.004
Occupied-NS	97	-48.08***	3.411	-14.09	0.000

[1] CG=Control Group. O=Occupation (police report).

[2] Occupied-S=Occ (p.r.) + SIMCE. Occupied-NS=Occ (p.r.) + No SIMCE.

[3] *** p<0.1, ** p<0.05, *** p<0.01**

Table 2.18: Missed Days

	attended days	Percentiles	missed days
	201	1%	77
	242	5%	36
	255	10%	23
	264	25%	14
	270	50%	8
	275	75%	3
attend$^{90^{th}\text{pctile}}$	278	90%	0
	279	95%	-1
	282	99%	-4

Table 2.19: Effect of # missed days on 10^{th} grade test results

instrument:	overall pupil 10^{th} grade test results			
occupied schools				
missedDays(dose)	-0.284***	-0.441***	-0.583***	-0.711***
tests				
First stage coeff.	14.850***	11.028***	16.374***	11.765***
First stage F t.(F-s r.)	219.26	59.75	181.11	50.87
controls				
fatheduc, books	No	Yes	No	Yes
muni FE	No	No	Yes	Yes

[1] *** p<0.1, ** p<0.05, *** p<0.01**

[2] This table shows the results of a dose-response treatment, specifically an IV approach. The missed school days are calculated by fixing attendance at the 90th percentile, 278, and then getting the difference between this 90^{th} percentile and the actual value of attended days. The first stage regress number of missed days on the instrument, occupied schools. The second stage regress the SIMCE on the estimated number of missed school days. I also implement sequentially controls (father education and books) and municipality fixed effects.

Table 2.20: School-level effect of # missed days on 10^{th} grade test results

instrument:	overall pupil 10^{th} grade test results			
occupied schools				
missedDays(dose)	-0.214***	-0.403*	-0.246***	-0.826**
tests				
First stage coeff.	24.264***	16.408***	23.097***	12.504***
First stage F t.(F-s r.)	269.15	24.83	238.28	12.76
controls				
indigenous mother, pc at home, repeat a year	No	Yes	No	Yes
muni FE	No	No	Yes	Yes

[1] *** p<0.1, ** p<0.05, *** p<0.01**

[2] This table shows the results of a dose-response treatment for a school-level model, specifically its IV approach. The missed days are calculated by fixing attendance at the 90th percentile, 278, and differencing with the actual value of missed days. The IV model has as instrument the occupied school dummy. The first stage regress occupied schools on number of missed days. The second stage the estimated number of school days to the SIMCE. I also implement sequentially controls (indigenous mother, pc at home and repeat a year) and municipality fixed effects.

APPENDIX I
THE JUNE 28TH 2011 POLICE REPORT LISTS 205 OCCUPIED SCHOOLS IN SANTIAGO

The original list released in LA TERCERA newspaper on June 28th 2011* included 231 occupied schools that the Chilean police declared were *under control by students* in the Metropolitan Region of Santiago. Small inconsistencies and inaccuracies in the list reduced the number of occupied schools to 205. It was the only "official" list released throughout the protests.

* (http://www.latercera.com/iphone/noticia/educacion/2011/06/657-376037-9-conoce-la-lista-de-los-231-colegios-metropolitanos-que-se-mantienen-en-toma.shtml)

(List by municipalities)

BUIN

COLEGIO DE MAIPO
LICEO 131
LICEO ALTO JAHUEL
LICEO FRANCISCO JAVIER KRÜGGER ALVARADO
LICEO POLIVALENTE LOS GUINDOS
LICEO TECNICO PROFESIONAL DE BUIN

CERRO NAVIA

LICEO POLITECNICO SAN FRANCISCO SOLANO

CONCHALÍ

COLEGIO CRISTOBAL COLON
LICEO AGUSTIN EDWARDS

EL BOSQUE

CENTRO EDUCACIONAL MATIAS COUSIÑO
COLEGIO VILLA SANTA MARIA
LICEO CHRISTA MC AULIFFE
LICEO JUAN GOMEZ MILLAS

EL MONTE

LICEO POLIVALENTE LUIS HUMBERTO ACOSTA GAY

ESTACIÓN CENTRAL

ESCUELA PARTICULAR JOSE ANTONIO LECAROS
LICEO COMERCIAL B-72
LICEO DE ADULTOS ESTACION CENTRAL
LICEO POLIVALENTE A N°71 GUILLERMO FELIU CRUZ

INDEPENDENCIA

LICEO GABRIELA MISTRAL
LICEO IGNACIO CARRERA PINTO
LICEO MIGUEL RAFAEL PRADO
LICEO POLIVALENTE A80 PRESIDENTE JOSE MANUEL BALMACEDA
LICEO ROSA ESTER ALESANDRI RODRIGUEZ
LICEO SAN FRANCISCO DE QUITO
LICEO SANTA TERESITA

ISLA DE MAIPO

CENTRO DE EDUCACION GENERAL BASICA
CENTRO EDUCACIONAL ISLA DE MAIPO

LA CISTERNA

CENTRO EDUCACIONAL LINCOLN COLLEGE
CENTRO POLITECNICO PARTICULAR SAN RAMON
CHILEAN EAGLES COLLEGE N.3
COLEGIO JOSE LUIS LEGRANGE DE LA CISTERNA
COLEGIO SANTA ISABEL DE HUNGRIA
ESCUELA TECNICA SANTA ROSA
LICEO IND DE ELECTROTECNIA RAMON B
LICEO POLITECNICO ABDON CIFUENTES
LICEO POLITEC CIENCIA Y TECNOLOGIA
LICEO POLITECNICO GALVARINO N.2
LICEO POLIVALENTE LA CISTERNA N°1
LICEO POLIVALENTE OLOF PALME
LICEO PORTAL DE LA CISTERNA

LA FLORIDA

CHILEAN EAGLES COLLEGE
COLEGIO PART. ANDARES DE LA FLORIDA
COLEGIO PART. FAMILIA DE NAZARETH
COLEGIO PART. NEW LITTLE COLLEGE
COLEGIO PART. SANTA LUCIA DE LO CAÑAS
COLEGIO PARTICULAR ANTILHUE
COLEGIO PARTICULAR NUEVA ERA SIGLO XXI
COLEGIO POLIVALENTE EDUCADORA ELENA ROJAS
COLEGIO QUINTO CENTENARIO CORDILLERA
COLEGIO SAN CRISTOBAL DE LAS CASAS
COLEGIO SANTA CECILIA DE LA FLORIDA
COLEGIO SANTA MARIA
COLEGIO SHIRAYURI
COMPLEJO EDUC. MUN CARD.A.SAMORE
ESCUELA LAS ARAUCARIAS
ESCUELA MARCELA PAZ
ESCUELA PARTIC PHILIPPE COUSTEAU
ESCUELA RAIMAPU-TIERRA FLORIDA
LICEO ANDRES BELLO
LICEO BENJAMIN VICUNA MACKENNA
LICEO INDIRA GANDHI
LICEO NUEVO AMANECER
LICEO POLIVALENTE LOS ALMENDROS

LA GRANJA

COLEGIO CHRISTIAN GARDEN SCHOOL
COLEGIO NUESTRA SENORA DE GUADALUPE
LICEO LA GRANJA
LICEO POLIVALENTE FRANCISCO FRIAS V.
SAINT CHRISTIAN COLLEGE

LA PINTANA

CENTRO EDUCACIONAL MUN.MARIANO LATORRE
COLEGIO ALTO GABRIELA
COLEGIO SANTO TOMAS
ESCUELA PARTICULAR PDTE.J. J.PRIETO

LO ESPEJO

COLEGIO PARTICULAR KENNEDY
LICEO POLIVALENTE
LICEO TENIENTE FCO. MERY AGUIRRE

LO PRADO

COMPLEJO EDUCACIONAL PEDRO PRADO

MACUL

COMPLEJO EDUC. JOAQUIN EDWARDS BELLO.
ESCUELA VILLA MACUL
LICEO POLIVAL MERC. MARIN DEL SOLAR

MAIPÚ

	LICEO INDUSTRIAL ALBERTO WIDMER
COLEGIO, LICEO COMERCIAL SAN JOSE	LICEO JOSE IGNACIO ZENTENO
ESC. BASICA BOSTON COLLEGE MAIPU	LICEO NACIONAL DE MAIPU
ESCUELA EL LLANO DE MAIPU	LICEO SANTIAGO BUERAS Y AVARIA

ÑUÑOA

	LICEO COMERCIAL GABRIEL GONZALEZ VIDELA
ESCUELA BASICA JOSE TORIBIO MEDINA	LICEO LENKA FRANULIC
ESCUELA JUAN MOYA MORALES	LICEO REPUBLICA DE SIRIA
LICEO AUGUSTO D HALMAR	LICEO TECNICO B N° 58 JOSE MARIA NARBONA

PADRE HURTADO

COLEGIO LOS ROBLES DEL CURATO	LICEO PAUL HARRIS

PEDRO AGUIRRE CERDA

CENTRO EDUC OCHAGAVIA	COLEGIO GRACE SCHOOL	ESCUELA VILLA SUR

PEÑAFLOR

COLEGIO JOSE MANUEL BALMACEDA	LICEO MUNICIPALIZADO PEÑAFLOR

PIRQUE

ESCUELA AGROECOLOGICA DE PIRQUE

PROVIDENCIA

COLEG POLIV PROF GUILL GONZALEZ HEINRICH	LICEO DE NIÑAS N° 7 LUISA SAAVEDRA DE GONZALEZ
LICEO B 42 TAJAMAR	LICEO JOSE VICTORINO LASTARRIA
LICEO CARMELA CARVAJAL DE PRAT	LICEO POLIVALENTE ARTURO ALESSANDRI P.

PUDAHUEL

	LICEO DE ADU ALBERTO GALLEGUILLOS J.
COLEGIO POLIV. SAN LUIS BELTRAN	LICEO MONSEÑOR ENRIQUE ALVEAR
COLEGIO SANTIAGO DE PUDAHUEL	LICEO MUN. CENTRO EDUC PUDAHUEL

PUENTE ALTO

CENTRO EDUC. PRINCIPADO DE ASTURIAS
CENTRO EDUCACIONAL FERNANDO DE ARAGON
CENTRO EDUCACIONAL SAN CARLOS DE ARAGON
COLEGIO EL SEMBRADOR
COLEGIO ENSENADA
COLEGIO MAIPO
COLEGIO NUEVA ERA SIGLO XXI SEDE PUENTE ALTO
COLEGIO OBISPO ALVEAR
COLEGIO PART. ACROPOLIS
COLEGIO PARTICULAR MIRADOR
COLEGIO POLIV. EL ALBORADA
COLEG POLIV PROF ILDEFONSO CALDERON
COLEGIO SANTA MARIA DE LA CORDILLERA
COLEGIO SENDA DEL SABER
ESCUELA CONSOLIDADA
ESCUELA DOMINGO FAUST SARMIENTO
ESCUELA TECNICA LAS NIEVES
LICEO COMERCIAL DE PUENTE ALTO
LICEO IND. MUNICIPALIZADO A N° 116
LICEO MUN.ING.MILITAR JUAN MACKENNA O.
LICEO MUNICIPAL CHILOE
LICEO PUENTE ALTO
LICEO SAN GERONIMO

QUILICURA

COMPLEJO EDUCACIONAL J. MIGUEL CARRERA
LICEO ALCALDE JORGE INDO

QUINTA NORMAL

LICEO EXPERIMENTAL ARTISTICO B-65
LICEO GUILLERMO LABARCA HUBERTSON
LICEO INDUSTRIAL BENJAMIN FRANKLIN
LICEO INDUSTRIAL VICENTE PEREZ ROSALES
LICEO POLIVALENTE JUAN A.RIOS

RECOLETA

LICEO COMERCIAL LUIS CORREA PRIETO
LICEO COMERCIAL NORA VIVIANS MOLINA
LICEO INDUS Y DE MINAS IGNACIO DOMEYKO
LICEO PAULA JARAQUEMADA
LICEO VALENTIN LETELIER

RENCA

LICEO INDUSTRIAL BENJAMIN DAVILA LARRAIN

SAN BERNARDO

CENTRO EDUC. PADRE ALBERTO HURTADO DE SAN BDO
CENTRO EDUCACIONAL BALDOMERO LILLO
COLEGIO ADULTOS INST. BARROS ARANA
COLEGIO NOBEL GABRIELA MISTRAL
COLEGIO PARTICULAR SANTA LUCIA
COLEGIO POLIV. PDTE. JOSE MANUEL BALMACEDA
COLEGIO SEBASTIAN EL CANO SAN BERNARDO
ESCUELA DE PARV. Y ESP. EDIAL
LICEO CLARA SOLOVERA
LICEO COMERCIAL DE SAN BERNARDO
LICEO COMERCIAL GABRIELA MISTRAL
LICEO ELVIRA BRADY MALDONADO-SN.BDO
LICEO IND. MIGUEL AYLWIN GAJARDO
LICEO INDUSTRIAL HARDWARE
LICEO POLIV. LUCILA GODOY ALCAYAGA
LICEO POLIV A-127 FIDEL PINOCHET LE-BRUN

SAN JOAQUÍN

LICEO INDUS DE SAN MIGUEL AGUSTIN ED
CENTRO EDUCACIONAL HORACIO ARAVENA A.
LICEO MUNICIPAL SAN JOAQUIN

SAN JOSÉ DE MAIPO

COLEGIO PART. ANDINO ANTUQUELEN
LICEO POLIVALENTE SAN JOSE DE MAIPO

SAN MIGUEL

CENTRO EDUC. PARTICULAR SAN LUIS
LICEO BETSABE HORMAZABAL DE ALARCON
ESCUELA E INSTITUTO DE MADRID
LICEO COMERCIAL INST. SUP. DE COM. DE CHILE (EX A99)
ESCUELA PARTICULAR Y COLEGIO CHILE
LICEO LUIS GALECIO CORVERA
LICEO ANDRES BELLO
LICEO TECNICO A-100 DE SAN MIGUEL

SAN RAMÓN

CENTRO EDUCACIONAL MIRADOR
LICEO MUNICIPAL PURKUYEN
ESCUELA COLEGIO ALBERTO BLEST GANA
LICEO MUNICIPALIZADO ARAUCANIA
ESCUELA ESPECIAL DE ADULTOS
LICEO SAN FRANCISCO

SANTIAGO

COLEGIO METODISTA DE SANTIAGO
LICEO INDUSTRIAL A-22 DE SANTIAGO
COLEGIO POLIV. MANUEL BAQUEDANO
LICEO INDUSTRIAL ELIODORO GARCIA ZEGERS
COLEGIO SANTA MARIA DE SANTIAGO
LICEO INSTITUTO NACIONAL
ESCUELA BASICA REPUBLICA DE MEXICO
LICEO ISAURA DINATOR DE GUZMAN
INST.SUP.DE COMERCIO EDUARDO FREI M.
LICEO JAVIERA CARRERA
INSTITUTO FEMENINO SUPERIOR DE COMERCIO E
LICEO MANUEL BARROS BORGONO
INTERNADO NACIONAL BARROS ARANA
LICEO MIGUEL DE CERVANTES Y SAAVEDRA
LICEO CONFEDERACION SUIZA
LICEO POLITEC. PDTE. GABRIEL GONZALEZ VIDELA
LICEO DARIO SALAS
LICEO POLIV.LIB. GRAL JOSE DE SAN MARTIN
LICEO DE APLICACION RECTOR JORGE E SCHNEIDER
LICEO TERESA PRAT DE SARRATEA

TALAGANTE

LICEO POLITECNICO DE TALAGANTE
LICEO POLIVALENTE TALAGANTE

Appendix II

Synthetic Control R Code

```
library(foreign)
occupy<-read.dta("occupy13.dta")
head(occupy)
library(Synth)
dataprep.out <- dataprep(
 foo = occupy,
 predictors = c("SIMCE10", "repeat", "studytime", "income",
  "timetohome", "indigent", "poor",
  "unemployment", "indigenous", "pc", "internet",
   "Presidentscholarship",
  "indigenousscholarship",   "tuitionscholarship",
   "otherscholarship", "publicinsurancea",
  "publicinsuranceb", "publicinsurancec",
   "publicinsuranced", "incomewithsubsidies",
  "percentagepublicschools"),
 predictors.op = "mean",
 time.predictors.prior = c(2000,2004,2007,2009),
 dependent = "SIMCE8",
 unit.variable = "regionno",
 unit.names.variable = "regionname",
 time.variable = "year",
 treatment.identifier = 13606,
 controls.identifier = c(13101:13132, 13201:13203, 13301:13303,
  13401:13404, 13501:13505,  13601:13605),
 time.optimize.ssr = c(2000,2004,2007,2009),
 time.plot = c(2000,2004,2007,2009, 2011,2013))
synth.out <- synth(data.prep.obj = dataprep.out, method = "BFGS")
```

```
path.plot(synth.res = synth.out, dataprep.res = dataprep.out,
 Ylab = "SIMCE 8", Xlab = "year",
 Ylim = c(240, 260), Legend = c("occupied schools",
 "synthetic occupied schools"), Legend.position = "bottomright")
pdf("fullsyncon.pdf")
dev.off()
save.image()
savehistory(file="occupy13.txt")
```

Appendix III

Robustness check for synthetic control

Table 2.21 shows the predictor mean for the treated group (205 occupied schools in the police report), the synthetic group (the weighted average of the fifty-two control municipalities after the optimization procedure), and a rough average of the fifty-two control municipalities.

Table 2.22 explores the weights for the twenty-one predictors. It can be seen that pc at home and internet at home are the two predictors with the highest weights, but an additional twelve predictors also have positive weights and only seven predictors have weights of zero.

Finally, Table 2.23 displays the subset of the donor pool, which possesses positive weights. In fact, only six municipalities in the entire fifty-two control municipality donor pool do so, which roughly illustrates why they have been selected. Santiago was the epicentre of the revolts, with twenty occupied schools, and there was also some action in San Ramón, with a total of six occupied schools. Padre Hurtado represents one peripheral municipality of action, having two occupied schools. Figure 2.8a demonstrates the behaviour of the treated group and the average of the fifty-two municipalities, as opposed to a synthetic average. Figure 2.8b offers a graphical illustration of how the treated and synthetic controls are related. The aim of the study is to compute the difference between the treated SIMCE and the synthetic SIMCE. The results should show a gap close to zero up to the treatment date if the optimization process is successful, which translates as a low mean squared prediction error (MSPE). If the treatment affects the outcome, there should be a noticeable bending down after the treatment.

The synthetic model is appealing because it is easily tested with placebo tests. The three most immediate placebo tests (see Figure 2.9) are based on each of the following falsification strategies.

- Placebo-in-outcome (see Figure 2.9a): Based on changing the outcome. In-

stead of the outcome SIMCE 8^{th}, the outcome is not altered by the intervention, such as the poverty percentage. In the short term, families in Chile tend to depend on parents. Thus, whether a pupil is attending school may, in principle, be uncorrelated to poverty.

- Placebo-in-region (see Figure 2.9b): Based on changing the treated region for any of the regions in the donor pool. For instance, instead of using the occupied school territory, Lo Prado can be used. Lo Prado is centrally located but has only one occupied school and, importantly, is not considered in the effective donor pool and has zero weight in the synthetic control.

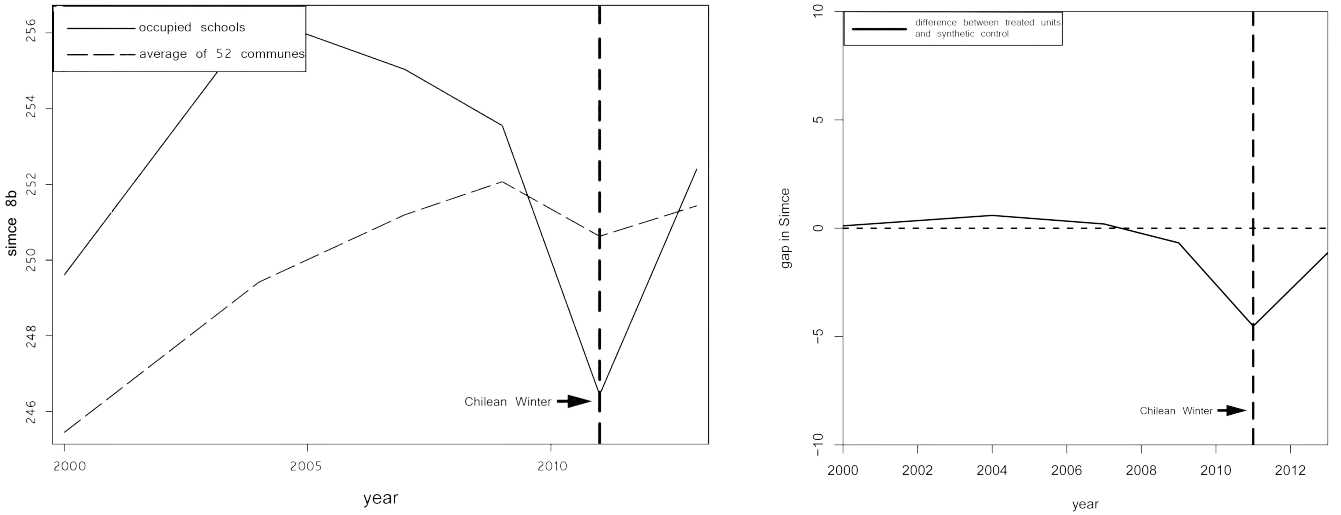
- Placebo-in-time (see Figure 2.9c): Based on changing the time component of the intervention, such as the date. For instance, instead of having the intervention in 2011, this value can be changed to 2009 to look for an absence in impact for this falsified intervention.

Next the permutation test is introduced to this robust scanning of the synthetic control method, as illustrated by Figure 2.10. This is a valid inference test which consists of iteratively running the synthetic method to each control municipality in the "donor pool" in order to gain possession of a distribution of placebo effects. Then, the gap between the occupied schools and the placebo gaps are compared. Relative to the estimated effect for a municipality chosen at random, this study expects a large estimated impact of the synthetic control for the occupied schools affected by the treatment. In other words, control municipalities from the "donor pool" should have random behaviour, with less gap activity before the treatment period and some random behaviour afterwards. The gap activity of the occupied schools should be sufficiently divergent from the random sample. Then, the permutation test graphs the school attainment gap for the occupied schools and for fifty-two (see Figure 2.10a), thirty (see Figure 2.10b), twenty (see Figure 2.10c), or eleven (see Figure 2.10d) municipalities depending on whether the mean square prediction error (MSPE) is at its maximum or less or equal than twenty-five, ten, or two times the MSPE for the occupied schools. Because the intervention is acute and

not continuous, it is important to graphically prove that at the time of the protests (year 2011), the gap is precisely lower for the occupied schools with a bouncing trend afterwards.

Finally, there is an inference test which calculates the ratio between post/pre MSPE for each of the fifty-two municipalities and the occupied schools. This test should show that the occupied schools represent some of the highest ratios, confirming that the permutation test is robust, as confirmed by Figure 2.11.

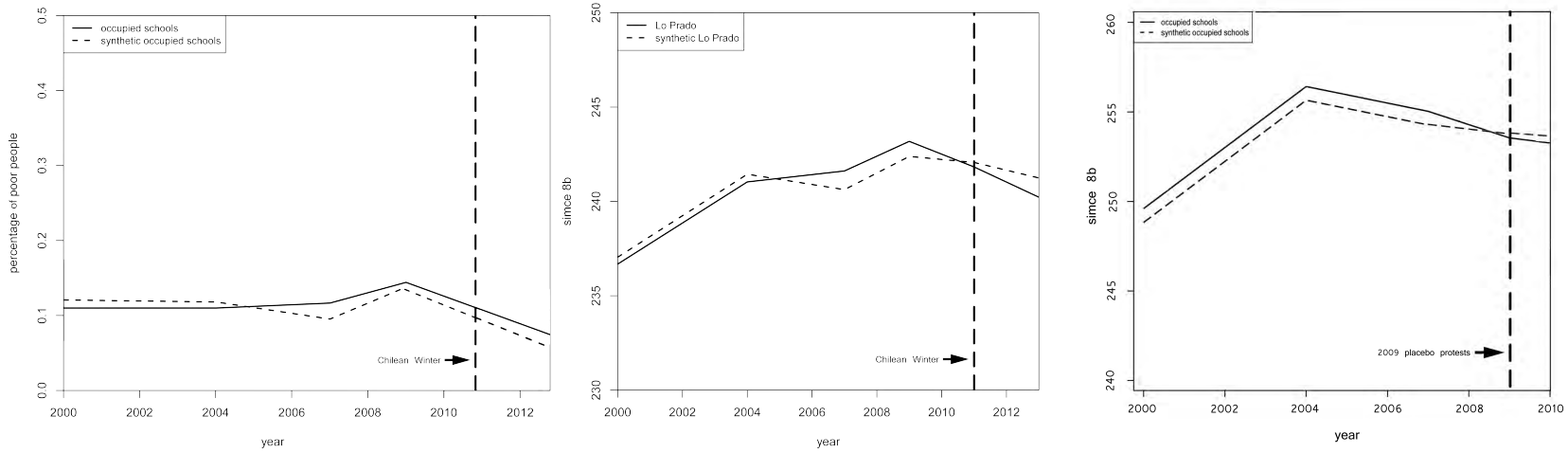
Figure 2.8: Synthetic control



(a) Average of the 52 municipalities

(b) School attainment gap

Figure 2.9: Placebo tests

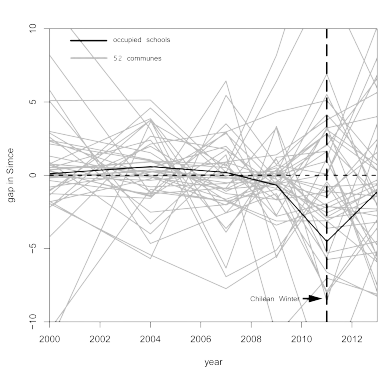


(a) Placebo-in-outcome

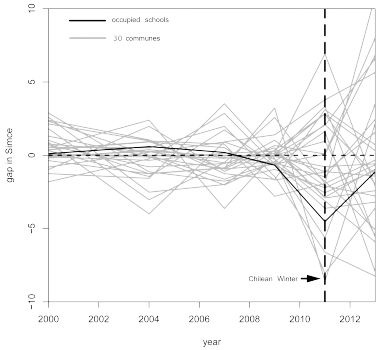
(b) Placebo-in-region

(c) Placebo-in-time

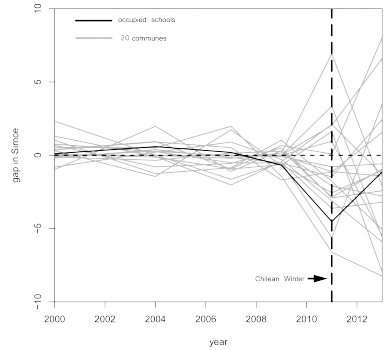
Figure 2.10: Permutation test



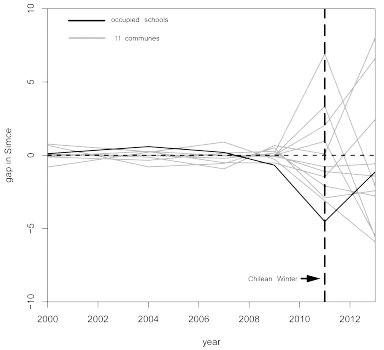
(a) All municipalities in the donor pool



(b) Municipalities MSPE $\leq$ 25 times occupied schools MSPE



(c) Municipalities MSPE $\leq$ 10 times occupied schools MSPE



(d) Municipalities MSPE $\leq$ 2 times occupied schools MSPE

Figure 2.11: Post/pre MSPE

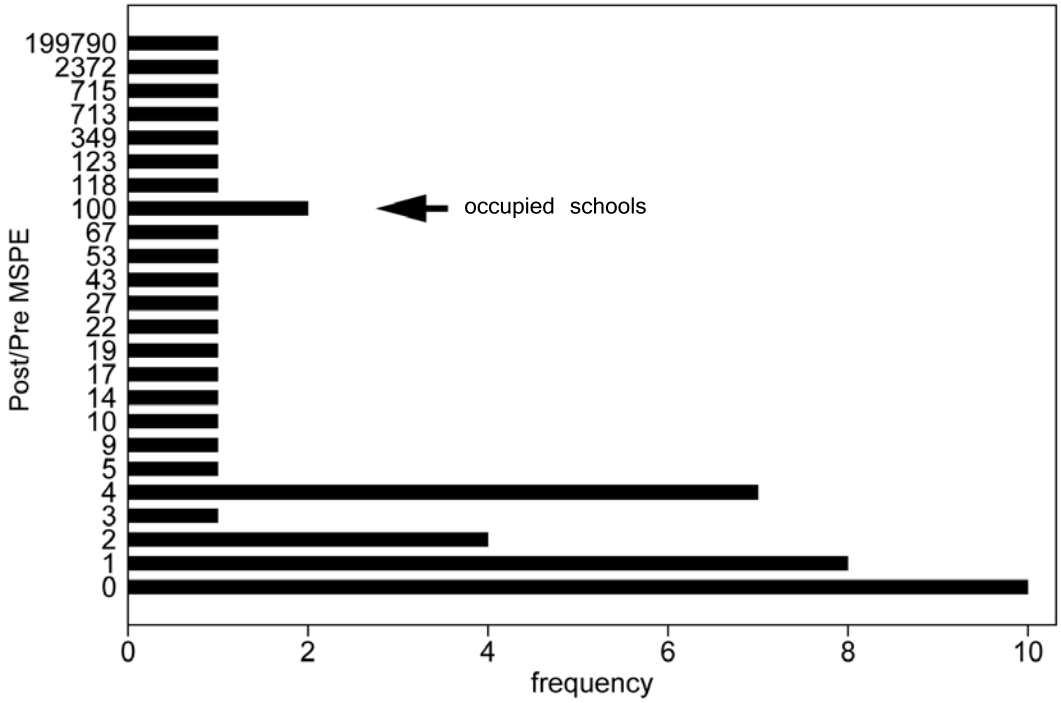


Table 2.21: Treated, synthetic and average predictors

predictors	occupied schools		average of 52 control municipalities
	treated	synthetic	
SIMCE 10^{th}	245.52	251.61	250.19
repeatyear	0.17	0.17	0.19
studytime	6.33	6.60	6.21
income	332.11	357.29	433.56
timetohome	20.96	23.02	17.90
indigent	0.03	0.03	0.03
poor	0.12	0.12	0.10
unemployment	0.29	0.10	0.09
indigenous	0.05	0.04	0.03
pc at home	0.42	0.38	0.38
internet at home	0.01	0.06	0.07
President scholarship	0.01	0.00	0.00
indigenous scholarship	0.00	0.00	0.00
tuition scholarship	0.00	0.00	0.00
other scholarship	0.02	0.00	0.00
public health insurance a	0.25	0.24	0.24
public health insurance b	0.22	0.22	0.22
public health insurance c	0.17	0.12	0.12
public health insurance d	0.08	0.08	0.09
subsidy income	8.94	8.95	8.59
public voucher schools	0.56	0.36	0.43

Table 2.22: Predictors and weights

predictors	weights
SIMCE 10^{th}	0.04
repeatyear	0.02
studytime	0.00
income	0.07
timetohome	0.00
indigent	0.01
poor	0.04
unemployment	0.01
indigenous	0.07
pc at home	0.42
internet at home	0.16
President scholarship	0.02
indigenous scholarship	0.00
tuition scholarship	0.01
other scholarship	0.00
public health insurance a	0.00
public health insurance b	0.05
public health insurance c	0.00
public health insurance d	0.05
subsidy income	0.04
public voucher schools	0.00

Table 2.23: Municipalities with positive weights in synthetic control

w.weights	municipalities	id
0.44	SANTIAGO	1
0.01	EL BOSQUE	5
0.22	SAN RAMÓN	31
0.03	SAN PEDRO	47
0.03	EL MONTE	49
0.26	PADRE HURTADO	51

Appendix IV

The outburst of the Chilean Winter in the media

Table 2.24: The onset and spread of school occupations in June 2011 at a national level in Chile (media reports)

Date				# of occupied schools
June	6^{th}	2011	(1),(2)	3
-"-	7^{th}	-"-	(3)	5
-"-	9^{th}	-"-	(4)	26
-"-	10^{th}	-"-	(5)	40
-"-	13^{th}	-"-	(6)	≈100
-"-	25^{th}	-"-	(7)	≈600

[a] **Source:** Wikipedia from the following media reports:

[1] *75 pupils arrested, 2 injured and millions in losses in protesters eviction from Barros Borgoño Lycée*. BíoBío Radio. June 6^{th} 2011. Retrieved September 3, 2014.

[2] *Eviction is requested for the Lycée Enrique Molina of Concepción*. BíoBío Radio. June 6^{th} 2011. Retrieved September 3, 2014.

[3] *pupils occupy the Amunategui and Aplicación Lycées*. ADN Radio. June 7^{th} 2011. Retrieved September 3, 2014.

[4] *Occupations spread: there are already 26 occupied schools at a national level*. BíoBío Radio. June 9^{th} 2011. Retrieved September 3, 2014.

[5] *Minister Lavín confirms that there are 40 occupied schools at a national level*. La Tercera Newspaper. June 10^{th} 2011. Retrieved September 3, 2014.

[6] *Secondary pupils have occupied approximately a hundred schools across the country*. El Mercurio Newspaper. June 13^{th} 2011. Retrieved September 3, 2014.

[7] *600 occupied schools in Chile*. Argentinian Public TV. June 25^{th} 2011. Retrieved September 3, 2014.

[b] The table shows the exponential growth and spread of school occupations at a national level in Chile for the month of June of 2011 when protests started. The first schools were occupied at the beginning of June. Each new day dozens of new schools were occupied in an explosive pattern. By the end of the month approximately 600 schools were already occupied at a national level.

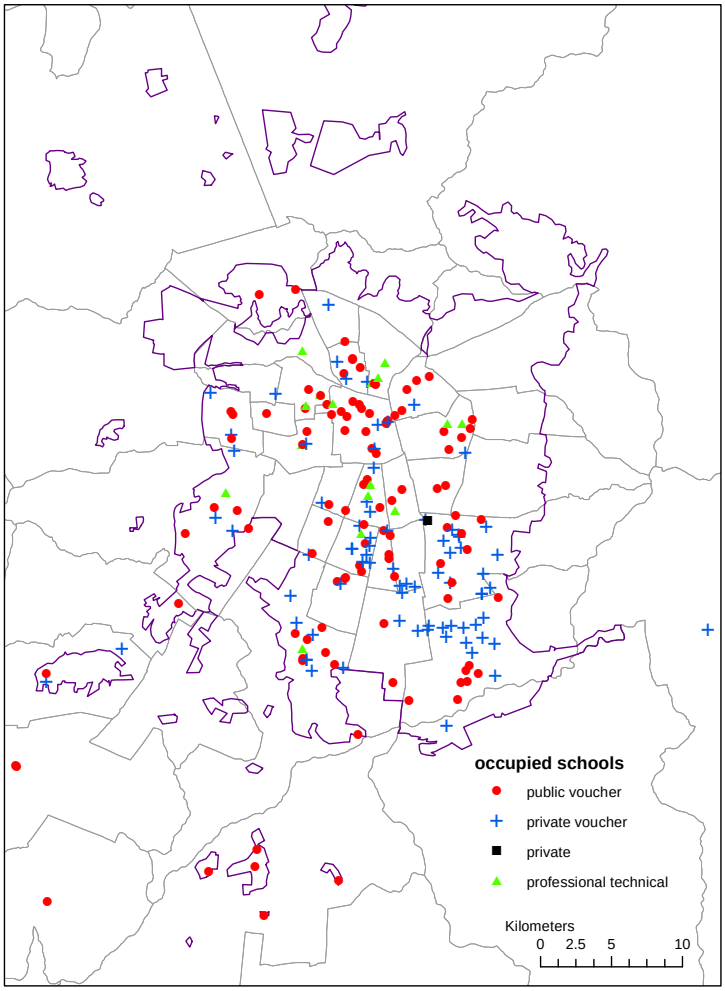
Appendix V

The geography of the *Chilean Winter*

Table 2.25: Overall occupied schools, Occupied-NS and Occupied-S by municipality

municipality	occupied schools	Occupied-S	Occupied-NS
LA FLORIDA	23	19	4
PUENTE ALTO	23	17	6
SANTIAGO	20	3	17
SAN BERNARDO	16	8	8
LA CISTERNA	13	6	7
SAN MIGUEL	8	2	6
INDEPENDENCIA	7	6	1
MAIPÚ	7	5	2
ÑUÑOA	7	2	5
SAN RAMÓN	6	5	1
BUIN	6	4	2
PROVIDENCIA	6	2	4
LA GRANJA	5	3	2
PUDAHUEL	5	3	2
QUINTA NORMAL	5	0	5
RECOLETA	5	0	5
LA PINTANA	4	3	1
EL BOSQUE	4	2	2
ESTACIÓN CENTRAL	4	1	3
PEDRO AGUIRRE CERDA	3	3	0
LO ESPEJO	3	3	0
SAN JOAQUÍN	3	2	1
MACUL	3	1	2
PADRE HURTADO	2	2	0
ISLA DE MAIPO	2	1	1
SAN JOSÉ DE MAIPO	2	1	1
PEÑAFLOR	2	1	1
TALAGANTE	2	1	1
CONCHALÍ	2	1	1
QUILICURA	2	0	2
LO PRADO	1	1	0
EL MONTE	1	0	1
CERRO NAVIA	1	0	1
RENCA	1	0	1
PIRQUE	1	0	1

Figure 2.12: Occupied schools by school type



This figure shows 205 overall occupied schools by school type: 110 public voucher, 78 private voucher, 16 vocational and 1 private schools. From this figure and Figure 2.1 on the main part of the paper, one can infer that the Occupied-NS schools populate the central municipalities around Central Santiago being in their great majority public voucher schools, while the Occupied-S schools are located to the periphery around Puente Alto being in their great majority private voucher schools. This is a real metaphor for the characters these pupils plays in the Chilean pupil movement.

Chapter 3

Surrogacy in the United States: Exploring the Effects of Legislation and Documenting the Consequences on Marriage, Births, Out-of-wedlock Births and Divorce

3.1 Introduction

Surrogacy[1] is an arrangement in which a woman carries and delivers a child for another couple. Although this practice is now medically feasible, it remains rare. California, the most surrogacy friendly state in the US, had an average of only 211 surrogacies per year for the period of 2001-2013. At the country level and according to available national figures, there have been 14,076 surrogacies from 1997 to 2013,

[1]In this chapter I use both the absolute number of surrogacies and the surrogacy rate. The surrogacy rate is the relative number of surrogacy out of 100,000 births.

with an average of 828 surrogacies per year and a total increase of the practice by 50% during the period. Nevertheless, this increase has not been steady. The trend exhibits a peak (e.g. in the year 2000, 1,210 surrogacy cycles were started) and a trough (e.g. in the year 2002—just two years later—only 548 surrogacy cycles were started), but it has stabilized at around 900 surrogacies per year during recent periods.

A mechanism that could explain this pattern is the introduction of either restrictive or permissive state legislation, or favourable or unfavourable court decisions regarding surrogacy. Some legislation may make surrogacy contracts valid, whereas others may forbid them entirely. To whom the birth certificate is extended is of crucial importance (Figure 3.1).[2] Permissive legislation or favourable court decisions tend to guarantee that the commissioning or intended parents are listed in this certificate. Restrictive legislations or unfavourable court decisions give the surrogate mother all parental rights over the baby.

Legislation is incipient, and the legal reforms have been slow, having peaked immediately after the famous Baby M case in New Jersey (Markens, 2007). Baby M was the pseudonym of the baby whose custody case in 1987 became the first American court ruling on the validity of surrogacy. The New Jersey court ruled that the surrogacy contract was invalid according to public policy and recognized the surrogate mother as the child's legal mother. Finally, the commissioning father, who was also the biological father, was awarded custody, with the surrogate mother having visitation rights.

To understand why legislation is incipient it is important to stress that according to Edlund and Korn (2002), marriage is a contract that gives parental rights to an

[2]See section **8a** indicating mother's current legal name, section **10a**, where the father's name is listed, and section **15**, which asks whether the mother is married at birth, conception or any time between. There is also a specification of whether paternity acknowledgement has been signed in the hospital.

uncertain father. The birth of a child is a public event, but the sexual intercourse in which the cell is fertilized is a private event with a potentially uncertain outcome. There is generally an automatic certainty associated with who a baby's mother is, which is never the case for the father: *Mater semper certa est, pater numquam.* In most Western countries, the institution of marriage, enshrined in law, resolves this issue by granting parental rights to the husband. Surrogacy is a mechanism that produces the same outcome without a marriage. As a result of surrogacy, single persons can acquire parental rights without marriage, and couples can deliver parental rights to the father, also without marriage. Nevertheless, most favourable state legislations demand marriage as a prerequisite to grant parental rights to the commissioning parents.

This paper is the first to study surrogacy and has two main aims. First, it documents the effect of legislation on surrogacy, indicating that restrictive legislation reduces surrogacy rates and that permissive legislation increases surrogacy rates. For this purpose, categories of legislation are created based on their impact on surrogacy contracts. This study uses pure restrictive/permissive legislation categories and extended categories with additional characteristics pertaining to surrogacy contracts, such as whether the surrogacy contracts are compensated (commercial surrogacy) or uncompensated (altruistic surrogacy), and whether the surrogate mother has a period of three to five days to change her mind and decide to keep the baby. Second, the causal effect of surrogacy on vital statistics (marriage, births, out-of-wedlock births and divorce rate) is assessed. This study's data are taken from the Clinic Table Data from the US Centers for Disease Control and Prevention for the years from 2001 to 2013. These are the most recent data available and they are the only data available on a state level basis in the US. The study's Surrogacy Data Set (Table 3.1 and Figure 3.2) is the first attempt to gather significant data on this issue. Data on surrogacy are elusive, most likely because it is a controversial issue.

In order to understand the effects of surrogacy on vital statistics, this paper

produces a theoretical model with the following assumptions: infertile couples by definition are infertile, which means that without help they cannot have children or, at least, their fertility is reduced; surrogacy increases fertility for these infertile couples because another woman carries a baby engendered by the infertile couple. Notwithstanding this fact, this may not cause an increase in overall fertility because infertile couples seeking surrogacy have a low fertility rate, and they produce fewer children on average than normal couples. As a result, this chapter aims to display whether there is a crowding out of fertile couples by infertile couples. If a marriage containing an infertile woman and a fertile man replaces a marriage containing a fertile woman with the same fertile man, then fertility increases for this couple but decreases for all couples. Another aim of this chapter is to identify if surrogacy increases marriages because infertile women gain hope of fertility through surrogacy and because of that enter the marriage market. At the same time, surrogate mothers do not reduce their own number of children because it is assumed that if they want to have their own children then they do not offer themselves in the surrogate market. It will also be investigated whether divorce is high among commissioning parents because marriage is usually needed for surrogacy contracts to be enforceable under state legislation, and after reproduction, the union may be less binding and end in more frequent divorce. The predictions of the stylized model are confronted with the data and found to hold.

3.2 What is surrogacy?

The commissioning parent or parents, sometimes called the social parents, may arrange a surrogate pregnancy because of female infertility or other medical issues which make pregnancy or delivery impossible, risky or otherwise undesirable. Surrogacy is the only mechanism by which a commissioning parent(s) can genetically overcome infertility. This is the main difference between surrogacy and adoption. A surrogacy contract implies a woman being pregnant on behalf of the commissioning

parent(s). The surrogate mother may be the child's genetic mother (traditional surrogacy) or she may be genetically unrelated to the child (gestational surrogacy). In traditional surrogacy, the surrogate mother's own egg is fertilized by artificial insemination or by direct sexual intercourse, and this is the most common and inexpensive type of surrogacy. Because traditional surrogacy can be performed with little or no medical assistance, there are no statistics available on this type of surrogacy. Alternatively, in gestational surrogacy, the surrogate mother receives a fertilized oocyte through in vitro fertilization techniques. Data on this can be found because this type of surrogacy can be estimated from the Centers for Disease Control and Prevention Success Rate National Summary and Fertility Clinic Report, which collects data for most assisted reproductive technologies (ART). This type of surrogacy depends heavily on technology and is much more expensive. Quotes for an all-included surrogacy surge easily to $100,000, as indicated below. Many commissioning couples choose this type of surrogacy when full family heritage and continuation of the bloodline is desired or because the existence of the legal framework that regulates the surrogate contracts ensures their names on the birth certificate. They pay more but receive the service of enforceable legal contracts. This suggests that, in a similar way to midwives, surrogate mothers may become popular in the following years until an artificial uterus becomes available. Notwithstanding this, surrogacy, as is the case with transplants, bears a force deterring its application, namely, repulsion.

Few activities have been more closely related to motherhood than pregnancy, a strictly non-market activity or a not substitutable time-input of the household production function. Surrogacy has started to challenge this. For a married couple the opening of a new market, i.e. a surrogate market, can signify gains from trade. If the couple's wife earns more from her work or she prefers not to go through the physiological changes associated with pregnancy, she can decide not to get pregnant and be a genetic and social mother, i.e. not a gestational one. To buy uterine services in a way implies that childbearing may be less of a woman's specialization, eroding the idea of marriage as a "long-term" contract to protect a wife specializing

in childbearing from abandonment and other adversities (Becker, 1973).

In his seminal work on marriage Becker (1973) states that "The obvious explanation for marriages between men and women lies in the desire to raise own children...". It is interesting that Becker adds, "Sexual gratification, cleaning, feeding, and other services can be purchased, but not *own* children." (Italics in original). In vitro fertilization has proved that fertilization can be outsourced without sex. Surrogacy has proved that pregnancy also can be purchased to bear one's *own* children. As divorce has permitted a covert polygamy, facilitating access to multiple partners in stages, surrogacy is starting to permit a kind of reproductive trade and liberalization as it uncouples the sexual and the reproductive functions.

There are many ways of conceiving of marriage. According to Family Economics (Becker, 1973), marriage is a contract tailored for women, granting them their husbands' long-term commitment in exchange for the women's specialization, at least in pregnancy and childbearing. Note that in modern families, there is less specialization because wage rates for women have increased substantially, and the number of children born per family has also declined. According to Evolutionary Biology (Trivers, 1972), marriage is a contract tailored for men, granting them the presumption of paternity. Note that this is being challenged by DNA testing. To reproduce, men need only to fertilize a woman's egg,[3] while women tend to invest more heavily in the reproduction process through a larger sex cell, pregnancy, childbearing and child nutrition and care.

There are some important facts involved in the comparison of the benefits and costs of marriage, surrogacy and divorce. Table 3.2 summarizes the arguments given in the previous paragraphs and adds an estimated cost for each choice from the author's own calculations, plus prices gathered by the author on-line from reported

[3]The oocyte is larger than the spermatozoon. This key regularity has been acknowledged in Evolutionary Biology since the initial works of Trivers (1972).

sources. Marriage costs approximately $25,000 and divorce costs $20,000 compared to the $100,000 cost of an average all-included surrogacy. Moreover, Espenshade and Calhoun (1986) calculates a per child expenditure of $312,000 from birth to age 18 in 2015 for a middle-class American family with two children. The mother's opportunity cost averages $62,400. These per child expenses account for roughly a quarter of the overall expenditure of the couple before the child enters college. This compares with the cited overall cost of entering a surrogacy contract, that is $100,000. "Adding a surrogacy" then implies an increase in the cost of a child of 32% on average, and it potentially indicates a release of part of the $62,400 of female input resources.

3.3 A literature review

Because surrogacy is a controversial theme, it is not surprising that it has been addressed more frequently in political, moral philosophical, sociological and anthropological literature than in that relating to economics. The former disciplines include examinations of nearly every possible position on the issue, while in Family Economics it is indeed a new topic.

The founding paper in what is called the law-and-economics of surrogacy is by Richard A. Posner (1989), who famously argues in favour of enforcing contracts for surrogate motherhood. "Even if there were no shortage of babies for adoption, there would be a demand for surrogate motherhood. People (a biologist would say their genes) desire genetic continuity, and surrogacy enables the couple to satisfy this desire", concludes Posner. Epstein (1995) also makes the case for full enforcement of surrogacy contracts. Also notable is a statement by Friedlander (1995), who says that if we take the enforcement of surrogacy contracts further, we could also begin enforcing contracts for prostitution. She continues by saying, "Surrogacy, then, may be a kind of "demerit good", one we—or at least I—view instinctively as harmful

regardless of what the individuals participating in the transaction decide. Society need not prohibit these goods, but may merely tax or otherwise regulate them to make them less attractive" .

These are allegations that follow the well-known Baby M case. As mentioned above, Baby M was a child born under a surrogacy arrangement which was legally confronted in the first American court case on the validity of surrogacy contracts. The commissioning parents were William and Elizabeth Stern while the surrogate mother was Mary Beth Whitehead. Mary Beth Whitehead was inseminated with William Stern's sperm corresponding to a traditional, not a gestational surrogacy. After the birth, Mary Beth Whitehead did not want to relinquish her parental rights over Baby M and decided not to give up the baby. The Sterns sued to be considered the child's legal parents as the surrogacy contract had established. The court's decision was to void the surrogacy contract and recognize as legal parents both genetic parents, William Stern and Mary Beth Whitehead. Custody was given to William Stern. During the case the news coverage of the case skyrocketed with most of the public opinion backing the surrogate and biological mother of Baby M, but the subsequent coverage of surrogacy has been modest (Markens, 2007).[4]

There are many books on surrogacy covering a variety of topics and positions, from compelling stories of women who were able to have babies through surrogacy to studies of the race and class impact of gestational surrogacy in a global market: Markens (2007), Ragoné (1994), Field (1988), Gostin (1990), Griswold (2006), Twine (2011). Field (1988) and Markens (2007) review the legal issues surrounding surrogacy. In particular, these works analyse when and to what extent the legal statutes on surrogacy have changed in recent years.

Elizabeth Kane—the pseudonym of the first legal surrogate in the US: Justin, her

[4]See Figure 1, "Coverage of surrogacy in the *New York Times, Los Angeles Times* and *Washington Post*", p. 21.

baby, was born on November 9th 1980—rapidly converted herself into an authorized voice on the subject. Her initial comments on surrogacy were strikingly favourable, such as "If a woman has a legal and moral right to her own body, isn't she free to exercise that right in any way she deems fit? Doesn't she have the legal right to rent her vagina or her uterus, or to sell her one-of-a-kind and irreplaceable heart—or unborn child?" But Elizabeth soon changed her mind and became an acid critic of surrogacy, joining Janice Raymond, Patricia Foster, Mary Beth Whitehead, Gena Corea and others in the National Coalition Against Surrogacy. In 1987 she released a press statement saying that "Surrogate parenting is an emotional mine field and I have become a statistic" and that "Today I can no longer explain to my children why I felt justified in exchanging their brother for a $10,000 check... Today the child I've sold does not know I exist". In an autobiographic book (Kane, 1988), she confessed her real name was coincidentally also Mary Beth—it seems that the order of the day was that *We are all Mary Beth!*—saying comfortingly to Whitehead "So, you fell in love with your baby, Mary Beth, and no one told you it would happen. Well, so did I. So do we all!" and condemning "a contract stating the surrogate gets $10,000 for the delivery of a healthy child, nothing for a *defective* child, and she must pay the father $25,000 if she decides to keep the child. Why should the life of a child suddenly go up in value if the birth mother keeps it, as opposed to the child's father raising it?" (Italics in original). Gena Corea asked metaphorically "Are women human beings or are we reproductive meat?". In her book, Corea (1988), she emphasizes that as buttocks, breasts and vaginas are sold for sex, so ovaries, wombs and eggs are sold for reproduction, and that supply and demand are pervasive in the sex and reproduction markets: Black wombs are demanded for White eggs when pharmacrats (the seditious patriarchal medicine) are in charge. Corea reveals that a patriarchal society understands women as machines, breeding machines used to reproduce the "stock of humans", a form of eugenic capital ready to produce and reproduce the best and to refrain from producing and reproducing the worst. Women have historically been seen as the breeders. So women-breeders have been consistently recognized as a production factory for the reproduction of humans, as just

another form of farm animal, with mothers being conceptualized simply as machines. Phyllis Chesler (1988) makes a thorough analysis of the Baby M case. She describes how father and mother surrogates are different. While the first are ejaculatory, transient, painless, riskless, orgasmic, a matter of five-minute masturbation, the second are painful, risky, permanent, nurturing, a matter of nine-month pregnancy and childbirth. Chesler (1988) mentions the most well-known surrogacy endeavours as SPA (Surrogate Parenting Associates in Louisville, Kentucky, run by lawyer Katie Brophy and Dr. Richard Levin, who contracted the first legal surrogate, Elizabeth Kane), CSP (Center for Surrogate Parenting, Inc., in Los Angeles, California, run by lawyer Bill Handel, who claims he rejects nineteen out of twenty surrogate applicants) and ICNY (Infertility Center of New York run by the surrogacy *factotum* and superstar lawyer Noel Keane, from whom she presents the full surrogacy contract, signed by Whitehead and Stern, whose highlights are as follows: "1. MARY BETH WHITEHEAD, Surrogate, represents that she is capable of conceiving children. MARY BETH WHITEHEAD understands and agrees that in the best interest of the child, she will not form or attempt to form a parent-child relationship with any child or children she may conceive, carry to term and give birth to, pursuant to the provisions of this Agreement, and shall freely surrender custody to WILLIAM STERN, Natural Father, immediately upon birth of the child; and terminate all parental rights to said child pursuant to this Agreement."; "That the consideration for this Agreement, which is compensation for services and expenses, and in no way is to be construed as a fee for termination of parental rights or a payment in exchange for a consent to surrender the child for adoption"; "MARY BETH WHITEHEAD shall be artificially inseminated with the semen of WILLIAM STERN by a physician."; "MARY BETH WHITEHEAD, Surrogate, and RICHARD WHITEHEAD, her husband, agree to surrender custody of the child to WILLIAM STERN, Natural Father, immediately upon birth, acknowledging that it is the intent of this Agreement in the best interests of the child to do so; as well as institute and cooperate in proceedings to terminate their respective parental rights to said child,"; "B) The consideration to be paid to MARY BETH WHITEHEAD, Surrogate, shall be de-

posited with the Infertility Center of New York..., the representative of WILLIAM STERN, at the time of the signing of this Agreement, and held in escrow until completion of the duties and obligations of MARY BETH WHITEHEAD, Surrogate,"; "5. MARY BETH WHITEHEAD, Surrogate, and RICHARD WHITEHEAD, her husband, understand and agree to assume all risks, including the risk of death, which are incidental to conception, pregnancy, childbirth, including but not limited to, postpartum complications. A copy of said possible risks and/or complications is attached hereto and made a part hereof"; and "7. MARY BETH WHITEHEAD, Surrogate, and RICHARD WHITEHEAD, her husband, hereby agree that it is the exclusive and sole right of WILLIAM STERN, Natural Father, to name said child").

Janice Raymond (1994) also adopts a feminist voice, arguing that all the new reproductive technologies violate the integrity of a woman's body in dangerous, destructive, debilitating, and demeaning ways, being the result of just another form of medical violence against women. She adds that reproductive surrogate contracts create a traffic in women's bodies. She underlines what she thinks to be faulty uses of language: using the terms 'surrogate' or 'substitute' mothers, for those she argues are real mothers, and using the term 'fathers' for those she indicates are only ejaculatory sperm sources. She is a strong disclaimer of procreative liberty and a critic of the happy surrogate presented by the reproductive technology industry's advertising, and the functional eschatology of utopian eternal life wielded by a patriarchal society that exploits women as wombs. Indeed, she insists on the striking resemblance between men buying prostitution in brothels and reproductive services in surrogacy arrangements. In what she calls the "spermatic economy", men unilaterally decide to have children and abuse women into reproduction. She conveys the message that for a woman there is no moral or teleological urgency to get pregnant and deliver babies for a patriarchal society: men and doctors use the sheer pornography of women's bodies in the form of photos of artificial inseminations to justify and sell the new reproductive technologies of a hegemonic male chauvinism.

In an overwhelming book, Harvard Profesor Martha A. Field (1988) depicts a sombre surrogacy and makes the case against enforcement of surrogate contracts. She stands for it as a legal practice only as a strategy for curbing illegal markets. She makes it clear that in her opinion surrogacy overall has a negative welfare impact on society, then and on behalf of public policy, contracts should be deemed as void based on babyselling prohibition. She also makes the case for always giving the surrogate mother a window to change her mind. Some of her many arguments are as follows: Whitehead receiving less than half the minimum wage, only $1.57 an hour; Whitehead recognizing the time inconsistency—in choosing to deliver the baby in advance but to keep it after the birth—of surrogacy contracts: "I signed on an egg. I didn't sign on a baby girl, a clone of my other little girl"; the fact that a surrogate may seem closer to a reproductive prostitute so society should ban the use of her body for this overly detrimental practice; that there is a Saks Fifth Avenue price tag for an intelligent and attractive surrogate, while a K-Mart price tag exists for a dull-looking one. Field (1988) concludes that sales and purchases should be made illegal, but donation should be allowed when biological material is involved, as in both organ donation and surrogacy. The book starts with two deeply moving real life quotations: one from a surrogate that decided to change her mind and who was brought to court by a commissioning couple just to learn that the commissioning mother was formerly a man who had had a sex change operation. And the poignant case of a microcephaly baby abandoned as a result of the split between the commissioning parents and the indolence of the natural mother: left alone in limbo at a hospital with non-antibacterial treatment with only the public to find help for the baby. But the story doesn't end here. All of the involved went on the Phil Donahue show where a paternity test was publicly released, revealing that the husband of the surrogate was the genetic father. Accordingly to Field (1988), this is another human drama surrogacy arrangements can provoke. Another touching example comes from the poor Frenchwomen who often sold not only their bodies, but even their teeth to be put into wealthy mouths. Most of the persuasive arguments are carefully crafted for traditional surrogacy, but Professor Field reaffirms her arguments,

extending them also to gestational surrogacies in subsequent publications as Field (1993) when establishing that the gestational mother is still the nurturer, the birth mother and, unlike her genetic counterpart, the only one ready to breastfeed the baby.

Noel P. Keane, the father of US surrogacy and the most famous baby-broker dismissed in Keane and Breo (1981) the bone of contention over him arranging for the White and wealthy to exploit the poor and Black . The founder of many Infertility Centers around the country, Keane died prematurely at 58 of melanoma having arranged more than six hundred surrogacies in his legal practice since 1976—by far more surrogacy than anyone else in the US—including the controversial match between Stern and Whitehead. An American rags to riches, self-made man story, *People* magazine wrote of him: "By devising elaborate contracts and pulling together a supply of surrogates sufficient to meet the demand, Keane has revolutionized the production of babies just as surely as that earlier son of Dearborn, Henry Ford, revolutionized the production of automobiles", Kunen (1987).

A surrogate interviewed by Ragoné (1994) said: "It's a gift of love. I have always been a really giving person, and it's the ultimate way to give. I've always had babies so easily. It's the ultimate gift of love". This gift category suppresses any consideration of money payments for surrogacy and produces a social link between the parties. Money is no longer on the table, what it is believed to be pure and simple love that takes its place. Another woman says: "...surrogacy sounded so interesting and fun. The money wasn't enough to be pregnant for nine months". And further: "I'm not doing it for the money. Take the money: That wouldn't stop me. It wouldn't stop the majority". Finally, another surrogate explained that "What's 10,000 bucks? You can't even buy a car...Money wasn't important. I possibly would have done it just for expenses especially for the people I did it for. My father would have given me the money not to do it".

Capron and Radin (1990) think that commercial surrogacy should be banned and only altruistic surrogacy should be permitted and the relationship between commissioning parents and surrogates should be treated not as a contract for services, but as an adoption arrangement similar to the New Jersey Supreme Court's ruling. Banning a market for children may help to protect women from exploitation or from becoming "breeding stock" like farm animals. That is, if reproductive capabilities were removed from a private and personal sphere and turned into blatant commerce. Defenders of surrogacy, Capron and Radin (1990) argue that a surrogacy arrangement must be understood under Family Law to suppress any comparison to "baby selling" with a hedonic regression with personal attributes such as sex, eye color, predicted IQ and athletic ability priced at a dollar value by the "babies market". Every child, whether it was sold or not, would receive this hedonic valuation. For instance, they claim the position of the Sterns was deleterious in Baby M's case, where they claimed the parents had bought the ovum—as previous donor semen sales were permitted by jurisprudence—from Mary Beth Whitehead, whose womb was subsequently hired for the gestation of the child. But the New Jersey Supreme Court ruled against the contract, saying that according to state laws on adoption there cannot be a contract previous to the birth and naming of the child, i.e. there cannot be adoptions in advance. Capron and Radin (1990) also put forward the view that paying for babies will have the effect of introducing a"commodification" of reproduction, treating babies as trivial objects to be disposed of or resold at will, as a magazine, a blouse or even a puppet.

Some also argue, see Robertson (1990), that the right of procreative liberty should be levied as a constitutional right emanating from the right of privacy applied both to coital and to non-coital reproduction, which should permit surrogacy. So the biological parents will be entitled to make use of contracts as a way of enforcing this defended constitutional right. This right can be understood as the personal human capability of causing, but also avoiding, procreation. In fact, to avoid procreation any competent woman is able to trigger an abortion up to a viable date

and any competent person is able to use contraception. On the side of causing procreation, Robertson (1990) advocates that surrogacy should also be permitted. He adds that it is a negative-right, since it obligates the state not to interfere with reproductive activities among consenting adults and physicians, but does not force the state to finance them. In another book, Robertson (1994), builds the concept of collaborative reproduction or when someone else other than one's partner provides the gametes or gestation for reproduction: sperm, egg or embryo donor, or surrogate motherhood. Notwithstanding this fact, this is problematic because a third party—a donor or a surrogate—is introduced into the usual situation of two-party parenthood, separating or deconstructing the traditional genetic, gestational and social unity of reproduction. Banerjee and Basu (2006) explain that in India surrogates are not the genetic mothers but share "blood and milk" —fluids—with the baby. An opposite statement should be to call surrogacy a heterologous harvesting of uterine tissue for ectopic motherhood.

In the economics literature, the issue of surrogacy is almost completely absent. I can only cite Gershoni and Low (2015), Banerjee and Basu (2006), Banerjee and Basu (2009), Banerjee (2013), Pelzman (2013) and Zil (2006). This last author examines the factors that affect the market for surrogate mother contracts in the United States, with a particular emphasis on the role of compensation. It has been argued that compensation for altruistic gifts (such as bearing children for another person) may decrease supply because the members of society devalue a service if it is not freely given, or the quality of the service may decrease if donors find it more advantageous to withhold information in favour of receiving a monetary benefit. This is an unpublished paper but represents a starting contribution to the somewhat inchoate literature on surrogacy.

One related article is a web publication by the Council for Responsible Genetics, authorized by Magdalena Gugucheva (2010). This publication is a review of recent topics on surrogacy in the US, including an initial (and in a way also precarious)

attempt to use the Centers for Disease Control and Prevention Clinic Table Data.

3.4 Data set

Data on surrogacy are elusive. However, in the US, there is an invaluable source of information: the US Department of Health and Human Services Centers for Disease Control and Prevention Clinic Table Data, which has been collected since 1995 based on a 1992 law that requires each fertility clinic to report some key data of their production function and overall performance to the Centers for Disease Control. With such data, the Centers for Disease Control publish a well-known yearly report and the Clinic Table Data on ART. ART includes all fertility treatments in which both egg and sperm are handled. In general, ART procedures involve surgically removing eggs from a woman's ovaries, combining them with sperm in the laboratory, and returning them to the woman's body or donating them to another woman. They do *not* include treatments in which only sperm are handled, such as artificial insemination or intrauterine insemination, or procedures in which a woman uses drugs to stimulate egg production without the intention of having the eggs retrieved. Some of the main types of ART include in vitro fertilization (IVF), gamete intrafallopian transfer (GIFT) and zygote intrafallopian transfer (ZIFT). ART is often categorized according to whether the procedure uses a woman's own eggs (non-donor) or eggs from another woman (donor) and according to whether the embryos used were newly fertilized (fresh) or previously fertilized, frozen, and then thawed (frozen). This survey includes data at the clinical level, such as the number of ART cycles started, success rates of fertility clinics and types of ART used. In 2001, a new question was included in this survey regarding the percentage of fresh non-donor cycles that used a gestational carrier, which can be used to back calculate the number of surrogacies and the number of ART-clinics performing gestational carrier services.

I created a variable for surrogacy using the gestational rate, which is the percentage of gestational carrier reported for all the fresh non-donor cycles started by each

ART clinic by state.[5] This variable shows data on surrogacy at the *state* level for the very first time. Together, these variables constitute what I call the Surrogacy Data Set. It is important to note that before this paper was initiated, nothing comparable with this data set existed in the literature. As a matter of fact, the Surrogacy Data Set is an important contribution of this paper and can be made available from this author upon request.

To further clarify what is in my sample and what is not, a spermatozoon, an ovule and a uterus are still needed for child birth. Of course surrogacy plays a relevant role in homosexual couples seeking fertility counselling. But the data has been apparently tabulated for heterosexual couples seeking infertility help in the clinics which report to the Centers for Disease Control and Prevention. Each gestational carrier is then related to a female patient with uterine infertility. Although my data are most likely only from heterosexual couples and my models are built for heterosexual couples as a result, I hope in the future surrogacy data for homosexual couples will be also generated and assessed independently.

How do my results compare to others in the literature? First, increased access to IVF acts as an insurance for age-related decrease in women's fertility. Gershoni and Low (2015) studied the case of Israel, which started a free IV programme in 1994, positioning the country at the vanguard of fertility related treatments. Israeli women responded to the policy intervention by marrying later, completing their university education and pursuing postgraduate qualifications. The results are that women are more likely to marry later by a third of a year, are 3% more likely to complete college education, 4% more likely to finish graduate school, and that marriage is postponed to older ages—over 30 years. Also, the first surrogacy contract model is developed

[5]Some clinics report "<1%". Because the national number of surrogacies is known, I add all integer percentages of surrogacy cycles and split the remainder evenly in the "<1%" cycles, allocating surrogacy cycles to each of these "<1%" clinics. Because I know the total national number of surrogacies and each of the full integer surrogacy cycle percentages, it is not a problematic assumption to proceed in this manner. On the contrary, confirming that assuming a figure strictly between 0%-1% permits to match the national figure with the incipient state figures represents indeed a robustness check for the Surrogacy Data Set.

by Banerjee and Basu (2006) and Banerjee and Basu (2009). In the model, commissioning parents faces heterogeneous surrogates. High type surrogates produce better outcomes but also have higher outside options. They show that enforceable contracts are needed to achieve better outcomes in the moral hazard problem regarding the surrogate actions. Since the health of the child depends strongly on the level of care provided by the surrogate, non-contractibility leads to a worse outcome for the child and commissioning parents. They also show that making surrogacy contracts unenforceable may swing commissioning parents towards low types, who need to be compensated less for inferior outside options. In a later paper and using a moral hazard model, Banerjee (2013) shows that altruistic surrogacy is optimal just in the case that the surrogate is increasingly altruistic and has lower outside options. In contrast, if higher outside options are present, commercial surrogacy is optimal. He also analyses the effects of social ignominy on the equilibrium outcome, showing that for low outside options social ignominy causes contracts to be more altruistic, while for higher outside options, it causes contracts to be more commercial. Pelzman (2013) also explores gestational surrogacy contracts as outsourcing services and explores some contractual arrangements. He also questions whether surrogacies in India should be regulated by the Agreement on Trade Related Aspects of Intellectual Property Rights.

3.5 Surrogacy legislation

3.5.1 Why legislation matters for surrogacy

Surrogacy legislation is crucial for the existence of surrogacy. Figure 3.3 presents the consecutive geographic representation of the US surrogacy legislation by state. Underlying surrogacy legislation is a continuous dispute among pressure groups wanting to adopt permissive legislation and opponents wanting to completely restrict the practice. Before the Baby M case, just a few states had previously enacted legislation, but many more adopted a position after the surrogacy national debate and media exposure that followed the case. There are other states that had not

established a legal framework for surrogacy until recently. A significant example, in which there was a clear choice in favour of surrogacy, is the case of the Illinois 2004 Surrogate Act. Moreover, in most cases there is no state legislation; rather, case-by-case court judgements determine the outcomes of this conflict. Up-to-date state legislation can be found on the Creative Family Connections LLC website (CreFamCon, 2015) and in Darra L. Hofman (2009)'s key article, "Mama's Baby, Daddy's Maybe: A State-by-State Survey of Surrogacy Laws and Their Disparate Gender Impact". To create one hundred and four dummy variables spanning thirteen years in eight categories of legislation, I have used these sources in conjunction with the academic on-line version of LexisNexis to verify the enacting year of statutory legislations and court decisions. I have established six categories: *criminalized, unenforceable, probably unenforceable, uncertain, probably enforceable* and *enforceable*. Furthermore, I have added two more categories: *uncompensated* (for states in which only uncompensated or altruistic surrogacies are allowed) and *time to change mind* (for states in which there is a window of three to five days for surrogate mothers to change their mind). The process of creating these variables is documented in detail in Tables 3.3 and 3.4. Finally, I have grouped the six legislation categories into negative laws (criminalized and unenforceable), ambiguous (uncertain and probably enforceable) and positive laws (enforceable), omitting probably unenforceable ones. Figures 3.4a and 3.4b show surrogacy and surrogacy rate by legislation. It is clear that more stringent legislation reduces surrogacy and surrogacy rate, while more permissive legislation increases them. In Figure 3.5, I show the first stage (surrogacy rate by legislation) and the reduced form (marriage rate by legislation) of my subsequent IV strategy. This visual approach suggests that legislation is correlated with surrogacy rate and that legislation is also correlated with marriage rate.

Surrogacy is an alternative method of fecundation and substitute pregnancy. The main difficulty implied by surrogacy is the cession of parental rights for the birth mother in favour of the commissioning couple which requires contracts and legal courts willing to enforce those contracts (not only is paternity at risk, maternity

is also at risk because the surrogate mother can threaten to keep the baby). This framework suggests that official adoption is the crucial step in making surrogacy viable. Without legal binding, the practice of surrogacy is severely restricted because the birth mother has all of the power in the negotiation, whereas the commissioning couple has none. It is of particular interest to assess the extent to which permissive legislation can increase the incidence of surrogacy. If a particular state passes a law that enforces the adoption of the child by the commissioning parent(s), can one expect to observe an increase in the practice of surrogate motherhood? Thus, one of the economic questions of interest can be formulated as follows: *Do newly passed permissive or restrictive legislations or favourable or unfavourable court decisions affect the rates of surrogacy for each state in the US?*

I acknowledge the argument that state and time policy legislation may be endogenous. It is possible that state legislation could respond to anticipated changes in surrogacy rates. It is also the case that states that change surrogacy policies could also be taking other actions that change the dependent variables (marriage, divorce and fertility). Both simultaneity issues can confound the results of the chapter and can be potentially an identification threat. Besley and Case (2000)'s critique of the use of state spatial and temporal source of variation should be taken seriously. It has been long recognised that this source of variation in laws afforded by a federal system is promising in order to uncover and estimate the effect of government policies on economic outcomes. Nevertheless, if state policies are purposeful actions, determined by economic and political conditions within the state, you may need to control for these variables if unbiased estimates are sought. Indeed, in the chapter I study the political process and control for the "colour" of the state legislation and other determinants of surrogacy legislation to deal with the endogeneity of time-varying state level policies. So using a panel data with fixed effects and controlling for the political process I am able to estimate a 2SLS with the external variation of the surrogacy legislation policy as an instrument to uncover the causal effect of surrogacy on demographic variables and to address the issue of the endogeneity of

surrogacy legislations.

Two other interesting aspects of legislation are that many legislations require that the commissioning couple be married, that is that marriage is compulsory if you are going to enter a surrogacy contract and you want your name to be listed on the birth certificate; and that either court decisions or statutory laws regarding surrogacy establish specific performance of contracts and not only damages when breaches occur (see also discussion in the subsequent section). A surrogate mother that does not want to relinquish all of her rights to the child she is carrying is obliged to do so at the moment of birth, and she is not granted the right to compensate the commissioning parents, i.e. she cannot pay for the costs of surrogacy plus compensation. She has to deliver the child to fulfil the contract in full.

3.5.2 An example of a permissive legislation

As an example of how a legislation can abruptly transform the number of surrogate pregnancies in a state, I will present the adoption of permissive legislation by the state of Illinois. I will use this example to test extensively the hypothesis of legislation shaping surrogacy in a synthetic control exercise. As you can see from the following legal excerpt, surrogacy is defined, permitted and enforced with clear rules for its procedure. First, the excerpt shows the rights of parentage with the surrogate mother relinquishing parental rights to the commissioning parents. Then, established established, is the eligibility which regulates the conditions that surrogate mothers and commissioning parents must fulfil. So, according to the Illinois Compiled Statutes (ILCS), Chapter 750 Families, 47/ Gestational Surrogacy Act:

...

(750 ILCS 47/5)

Sec. 5. Purpose. The purpose of this Act is to establish consistent standards and procedural safeguards for the protection of all parties involved in a gestational surrogacy contract in this State and to confirm the legal status of children born as a result of these contracts. These standards and safeguards are meant to facilitate

the use of this type of reproductive contract in accord with the public policy of this State.

(Source: P.A. 93-921, eff. 1-1-05.)

...

(750 ILCS 47/15)

Sec. 15. Rights of Parentage.

(a) Except as provided in this Act, the woman who gives birth to a child is presumed to be the mother of that child for purposes of State law.

(b) In the case of a gestational surrogacy satisfying the requirements set forth in subsection (d) of this Section:

(1) the intended[6] mother shall be the mother of the child for purposes of State law immediately upon the birth of the child;

(2) the intended father shall be the father of the child for purposes of State law immediately upon the birth of the child;

(3) the child shall be considered the legitimate child of the intended parent or parents for purposes of State law immediately upon the birth of the child;

(4) parental rights shall vest in the intended parent or parents immediately upon the birth of the child;

(5) sole custody of the child shall rest with the intended parent or parents immediately upon the birth of the child; and

(6) neither the gestational surrogate nor her husband, if any, shall be the parents of the child for purposes of State law immediately upon the birth of the child.

...

(Source: P.A. 93-921, eff. 1-1-05.)

...

(750 ILCS 47/20)

Sec. 20. Eligibility.

(a) A gestational surrogate shall be deemed to have satisfied the requirements of this

[6]or commissioning

Act if she has met the following requirements at the time the gestational surrogacy contract is executed:

(1) she is at least 21 years of age;

(2) she has given birth to at least one child;

(3) she has completed a medical evaluation;

(4) she has completed a mental health evaluation;

(5) she has undergone legal consultation with independent legal counsel regarding the terms of the gestational surrogacy contract and the potential legal consequences of the gestational surrogacy; and

(6) she has obtained a health insurance policy that covers major medical treatments and hospitalization and the health insurance policy has a term that extends throughout the duration of the expected pregnancy and for 8 weeks after the birth of the child; provided, however, that the policy may be procured by the intended parents on behalf of the gestational surrogate pursuant to the gestational surrogacy contract.

(b)The intended parent or parents shall be deemed to have satisfied the requirements of this Act if he, she, or they have met the following requirements at the time the gestational surrogacy contract is executed:

(1) he, she, or they contribute at least one of the gametes resulting in a pre-embryo that the gestational surrogate will attempt to carry to term;

(2) he, she, or they have a medical need for the gestational surrogacy as evidenced by a qualified physician's affidavit attached to the gestational surrogacy contract and as required by the Illinois Parentage Act of 1984;

(3) he, she, or they have completed a mental health evaluation; and

(4) he, she, or they have undergone legal consultation with independent legal counsel regarding the terms of the gestational surrogacy contract and the potential legal consequences of the gestational surrogacy.

(Source: P.A. 93-921, eff. 1-1-05.)

According to an article by Gitlins (2015) in the American Academy of Marriage

Lawyers website, Illinois will become a magnet for surrogate pregnancy because of the relatively simple procedure for obtaining a birth certificate, the fact that the medical procedure should be performed in Illinois, and that the parent-child relationship can be legally established before birth. In her own words, "Though other states have codified gestational or traditional surrogacy procedures, none has made obtaining a birth certificate as easy as the Illinois statute. New Hampshire, Texas, Virginia and Florida all have statutes regulating surrogacy. Neither Florida nor New Hampshire have pre-birth procedures whereby the intended parents may be listed on the original birth certificate of the child....".

3.6 The microeconomics of the surrogacy decision

The microeconomics of the surrogacy decision builds on some stylized facts: (1) Some infertile women will use surrogacy to have children. (2) Married couples where the woman is infertile can use surrogacy, but this is costly and not always successful, so married couples where the woman is infertile who enter a surrogacy contract have fewer children on average than married couples where the woman is fertile. (3) While fertile women can have out-of-wedlock children, infertile women cannot because many states require married couples to enter a surrogacy contract. In other words, infertile women cannot have out-of-wedlock children as a result of surrogacy legislation requirements. (4) Because infertile women are compelled to get married in order to access surrogacy contracts, these convenience marriages have a higher probability of divorce. (5) Gross but anecdotal evidence suggests that the women in the Surrogacy Data Set are already married infertile women or infertile women who get married as a result of the legally prerequisite in place by the surrogacy legislation. (6) No homosexual couples are included in the Surrogacy Data Set, so even if they are an important group consuming surrogacy, data have not been released for them and the model is not designed to represent them. In fact, in the general model marriage can be substituted out by surrogacy as a new institution to assign

paternity rights and divorce can be lowered if newly fertile and happy marriages take place and do not split up. Nevertheless, in my model with only heterosexual couples marriage is likely to increase as stated and divorce may increase or decrease, depending on which effect is larger: the newly and enduring happy marriage or the breaking of a marriage of convenience consecrated only for surrogacy legislation requirements. To sum up, my model is fitted for a conservative or soft surrogacy open just to fertile couples with uterine infertility. Still it is appealing to do further research for a liberal or hard surrogacy open to singles, fertile women, gays and lesbian as data will become available. Results are likely to be opposed, depending on which hard or soft type of surrogacy arrangement is or will become under study.

To support these stylized facts I present a cost-benefit analysis of marriage, surrogacy and divorce (See again Table 3.2). From the Table, it is clear when the three occur together or when they occur separately. Infertile couples mainly enter surrogacy arrangements to overcome infertility conditions as stated in the table. For these couples, it is in fact expected that marriage and surrogacy, but not divorce, will occur. For compulsory marriages required by surrogacy laws (as underlined previously, many state laws confer validity on surrogacy contracts signed only by married commissioning parents) marriage, surrogacy and, this time also divorce, are expected to occur together. For high earning females who enter surrogacy arrangements to keep their salaries and not be subjected to (unpaid) maternity leave as stated in the table, it is expected that marriage, surrogacy and divorce will occur together. For parents motivated by continuation of the bloodline, a positive effect on surrogacy and ambiguous effects on marriage and divorce are expected. For women that do not want to risk the adverse effects of pregnancy (whether physical, mental or aesthetic), marriage and surrogacy are expected to occur together and the effect of divorce is expected to be ambiguous. Finally, for single persons who just want to be listed on the birth certificate and give or receive child support, as stated in the table, a negative effect on marriage and a positive effect on both surrogacy and divorce are expected. Gross statistics suggest that infertile couples are the most

prevalent of these categories and that single persons are the least.

A key point on surrogacy analysis is the degree of irreversibility of marriage—there is costly reversibility through divorce—and the option value to wait for a better match. The theory on this comes from the classic Dixit and Pindyck (1994) and an excellent example in education is Hogan and Walker (2007). Using option theory they study the education decisions of individuals when there is uncertainty concerning the returns of education. They model an irreversible decision of leaving the education investment, showing that high returns on education will cause individuals to postpone exercising their option to leave the human capital accumulation. In addition, riskier options somehow surprisingly implies more time spent investing, independent of the risk preference of each agent. This is explained by the irreversibility and the option to delay the investment developed in the models of Dixit and Pindyck (1994). Burdett and Coles (1999) use a search model that includes the option value of waiting for better offers in what it is called the reservation value. This implies that for every offer under the reservation value, it is worth to keep waiting for new offers and better matches and to postpone marriage. While for an above the reservation value offer, it is worth taking it, marrying and stop searching.

Each fertility treatment operates as an insurance promoting riskier fertility behaviours. As a result, the time to marriage and the time to the first child are increased. Divorce rates should generally decrease because a fertility success is more common and adults stop searching for new fertile partners. Surrogacy is nevertheless a special case, where state regulations instigate marriage, but also post surrogacy divorce.

In my model there are two kinds of contracts: a surrogacy contract between commissioning parents and the surrogate which is permeated by moral hazard issues—the surrogate has hidden actions, and the marriage contract between a man and a low fertility woman which is characterized by adverse selection—the woman

has hidden information, her being low fertility. Men are the principal because they offer a take-it-or-leave it contract of marriage to their consort. The contract is simple: divorce if not fertile, marriage forever if fertile. Following this argument surrogacy works as an insurance, so risk behaviour increases and marriages is delayed and also the divorce rate is lower. Even if delayed, marriage should be at its all-time high because low woman enter the marriage market. Divorce, nevertheless, should also increase because many new marriages (low woman marriages) are short-lasting because they are motivated only by the married requirement of surrogacy laws.

A damage measure should be high enough to induce performance in a mutually self-enforced contract. But a self-enforcement contract could also stipulate a specific performance clause[7] if it requires a party to perform a specific act: the surrogate relinquishing her paternal rights. It is a better alternative to awarding damages. Of course, high damages for not relinquishing paternal rights is a straightforward means of delivering a similar outcome. But life is—especially ex-ante—priceless, so specific performance is better. Baby M's contract was not complete because it did not stipulate a penalty for every contingency, such as the surrogate and biological mother deciding not to relinquish of her rights over the baby. As I said, a sufficiently high damages remedy for breaching the contract could suffice, but entitling specific performance is a more reasonable provision for this contract. In a gestational surrogacy the actions are clear: The surrogate mother has to let herself be inseminated and carry, gestate and bear the baby, and after birth she must relinquish her paternal rights and deliver the baby. While the intended or commissioning parents have to deliver their fertilized sex cells to the surrogate and have to accept the baby

[7]Clauses that convey the parties' intent that the court award specific performance in the event of breaches, e.g. "Each of the parties hereto agree that irreparable damage would occur if any provision of this Surrogate Contract were not performed in accordance with the terms hereof. The parties hereto recognize and agree that money damages may be insufficient to compensate the commissioning parents for breaches by the surrogate, consequently, that the equitable remedy of specific performance of the terms hereof will be available in the event of any such breach and that the parties shall be entitled to enforce specifically the performance of the terms and provisions hereof in any state court".

after the birth, all the actions can be compensated with money, but because of their supreme value in terms of life and paternity rights, who keeps the baby and who is/are going to act as parents is of great importance. These last actions/goods are not easily compensated. They are life-or-death events so whenever the possibility to honour the contract is still possible, i.e. the child is alive for instance, the child should be given to the one/both entitled to it. As Thomas and Worrall (1988) show, in a self-enforcing contract neither party has an incentive to renege, i.e. these contracts must offset any short-term gain from reneging with greater long-term benefits from compliance. So at the core of these self-enforcing contacts is how the conflict between risk-sharing and the given self-enforcement is solved. As a matter of fact, informal contracts where the surrogacy contract are not enforceable, but if they are criminalized then there is no self-enforcement possible . But an infertile couple who contract to gestate a baby will never be satisfied with anything less than the baby. Money will not suffice. In many unique goods contracts—houses, there is a presumption of specific performance. So infertile couples will file for this remedy for the breach of contract if the surrogate is not willing to deliver the baby. But at the same time, babies are not real property so courts can apply Family Laws and treat the matter as a custody case, giving custody to the surrogate mother and genetic mother and the genetic father as in Baby M's case. Irrational threats, such as the purported nonsensical threat to kill the baby or the threat to kidnapp him/her are a kind of emotional—non pecuniary—threat, difficult to evaluate for the other party but also for the courts. So being neither possible nor easy to rule out, these threats are non-credible ones, but modifications of contract can still resort to them, see Ben-Shahar and Bar-Gill (2003). For surrogacy contracts where emotional issues flourish easily, validity, enforcement and renegotiation of contracts can all be extremely complex and difficult to solve and settle, and may need case by case careful examination and judgement.

Furthermore, an interesting discussion sparks from the fact that before the fertilized eggs are implanted in the uterus of the surrogate, the contract between both

parts is similar to producing goods or services, producing the baby or providing gestational services. Then the remedy of damages for breach of contract is preferred. For instance, the surrogate will need to pay a penalty for giving up her obligations in the contract. But after the embryo is inside the surrogate body the contract is similar to conveying (i.e. "To transfer ownership of or title to") existing goods or other properties, that is the baby and the parental rights. Of course, the commissioning parents cannot buy the baby, because how can they buy something that nobody but him/herself owns? But the surrogate can transfer her paternal rights to the commissioning parents. Then the remedy of specific performance is more adequate for breach of contracts. See Shavell (2006) for an interesting and detailed analysis of this issue.

Surrogacy as a new institution for producing children may replace the traditional institution of marriage. Even single men can acquire parental rights without even having a sexual partner. Same sex couples can overcome the so far impossible sexual mix with a personal bloodline continuation. Again, marriage can be under fire. But in my sample I have only heterosexual couples and heterosexual marriage rates. Within this sample, marriage should instead be boosted because with surrogacy low fertility women become insured and start participating in the marriage market. It may happen that in the population, the substitution effect is stronger, but in my sample the insurance effect is the only visible one. At the same time, divorce may or may not increase. Even in my sample, I have two competing effects: for some marriages the gift of reproduction acts as a love potion and divorce plunges. At the same time, because many more marriages of convenience are needed for surrogacy to work the first time, many more divorces have been made possible. So in my sample it is an empirical question for which both opposite effects prevail.

Furthermore, taking the predictions of the previous discussion on the microeconomics of surrogacy: option value, insurance, adverse-selection and moral hazard, self-enforcement contracts, specific performance, irrational/emotional threats and

the discussion on cost and benefits at the beginning of the section, the following hypotheses can be formulated for my restricted sample:

1. *Hypothesis 1, surrogacy increases marriage*;

2. *Hypothesis 2, surrogacy may increase or decrease divorce.*

I emphasize again that these hypotheses are for my restricted model. For instance, in a general model marriages are likely to decrease as discussed earlier. Additionally, couples entering a surrogacy contract face very high costs of procreation, causing a substitution away from highly fertile couples, and because of an intrinsic reduced fertility rate amongst "newly fertile" infertile women, and because of its increase in surrogacy marriage crowding out fertile women's marriages, there is also an effect in births and out-of-wedlock births, and thus two more hypotheses can be added:

3. *Hypothesis 3, surrogacy causes a decrease in births*;

4. *Hypothesis 4, surrogacy causes a decrease in out-of-wedlock births.*

Note that Hypothesis 3 and Hypothesis 4 are related. The former implies a decrease in married and unmarried births, while the latter implies a decrease at least in unmarried (out-of-wedlock) births. This chapter will rigorously test these four hypotheses. Finally, for a graphic intuition of some of these results see Figures 3.6 (More marriages and fewer married births after surrogacy are introduced. When surrogacy is introduced infertile women who were out of the marriage market are now getting married, increasing the married population. In the Figure you can see that now there are more marriages; in fact now there are seven marriages and before there were just six, but there are fewer married births because new surrogacy marriages crowd out more fertile marriages) and 3.7 (Fewer unmarried births after surrogacy is introduced. Surrogacy increases marriages so unmarried births are replaced by married ones. In summary, more marriages, fewer births and fewer out-of-wedlock births). In both Figures, each pair of hollow circles is a marriage between a man

and a fertile woman; each pair of grey filled circles is a marriage between a man and an infertile woman who have entered a surrogacy contract. Each leg is a child, and fertile couples have two legs so an average childbirth rate of 2, while surrogacy marriages have only one leg, meaning so that their overall childbirth rate is lower, at 1 child per marriage (just think of the cost of surrogacy, which makes it difficult to have multiple children).

3.7 Identification strategy

This chapter seeks to answer two questions. The first is how surrogacy or the surrogacy rate is influenced by surrogacy legislation. The population model or data-generating process is as follows:

$$Surrogacy_{st} = \boldsymbol{Legislation}'_{\boldsymbol{st}}\boldsymbol{\beta_1} + \boldsymbol{Controls}'_{\boldsymbol{st}}\boldsymbol{\zeta_1} + \boldsymbol{State}'_{\boldsymbol{s}}\boldsymbol{\eta_1} + \boldsymbol{Year}'_{\boldsymbol{t}}\boldsymbol{\gamma_1} + \varepsilon_{1st} \tag{3.7.1}$$

This equation relies on the assumption that legislation[8] is as good as random after controlling for fixed effects (state and year) and a set of fertility, labour market, political process, educational attainment and demographics controls. Notice that legislation is not solely exogenous after selecting on observables, but legislation *timing* is, which makes it as good as random after selecting on observables.

This question is also the first stage of a two stage procedure to answer the second question of interest in this chapter. The second population model studies how vital statistics such as marriage are shaped by surrogacy using the following data-

[8] As six categories: criminalized, unenforceable, probably unenforceable, uncertain, probably enforceable and enforceable, or as three grouped variables: negative laws (criminalized and unenforceable), ambiguous laws (uncertain and probably enforceable) and positive laws (enforceable); probably unenforceable is omitted.

generating process:

$$Vital\ statistics_{st} = \beta_2 Surrogacy_{st} + \boldsymbol{Controls'_{st}\zeta_2} + \boldsymbol{State'_s\eta_2} + \boldsymbol{Year'_t\gamma_2} + \varepsilon_{2st} \tag{3.7.2}$$

The equation requires a similar assumption, and because it is the main equation of the chapter, I will be very specific in addressing it. $Vital\ statistics_{st}$ are state s and year t values for marriage rate, births, out-of-wedlock births and divorce. $\boldsymbol{State'_s}$ is an unobserved state level fixed effect that includes all time-invariant state characteristics, such as attitudes towards values, religious feelings, unobserved state demographics and labour market characteristics. $\boldsymbol{Year'_t}$ is an unobserved time level fixed effect that comprises all nation-invariant year characteristics, such as nation-wide press reports, public awareness or discussion on surrogacy, economic shocks, terrorist attacks, and other yearly national distressful issues. ε_{2st} is a state-by-year error. The main threat to identification comes from endogeneity or the fact that surrogacy may be correlated with time and state fixed effect and with the error term.[9] To give a flavour of the presence of this identification problem, I will discuss correlation with moral values, which I argue are in the error term of the vital statistics equation. Conservative values tend to forbid surrogacy and at the same time increase marriage trends. They can also shape the state numbers of surrogacy contracts. Therefore, a proxy for moral values is used as a control: party affiliation in the political process. Of course there can be other confounders, so a reasonable identification strategy will look for an external shock to disentangle the effect of surrogacy on marriage and other vital statistics. Legislation is defensibly this needed external shock. The legislation process is complex and unpredictable but there is

[9] Recall that controlling for fixed effects is needed if there is correlation between the regressors and the *unobserved effects*: $\boldsymbol{State'_s}$ and $\boldsymbol{Year'_t}$. An instrumental variable is needed on top of that when the regressors are correlated with the *idiosyncratic error*: $\varepsilon_{\mathrm{st}}$.

sorting in surrogacy legislation: states with uncertain legislation are localized to the centre, while states with enforceable legislation are localized to the periphery. Nevertheless, as a balancing test shows, the *timing* in which legislation is introduced in each state is arguably random after controlling for key covariates and year and state fixed effects. This clearly indicates that legislation instruments can help to address causality of surrogacy on vital statistics. Nudging surrogacy variation with legislation variation permits the explanation of vital statistics variation under a full causality interpretation. In summary, an instrumental variable procedure in which surrogacy rate is instrumented by surrogacy legislation results in a promising and defensible procedure to explore the causal relation between surrogacy rates and vital statistics.

A balancing test (see Table 3.7) to test the suitability of this IV procedure shows the unconditional regressions of the predicted endogenous variable (from a regression of surrogacy rate on the full-set of instruments: negative, ambiguous and positive laws) on each of the different 7 controls and 29 unobservables. The sign of the coefficients of these relations suggests that states with higher levels of surrogacy are characterized by more parents who are infertile, of Black or Hispanic origin, with higher labour market characteristics, affiliated with the Democratic party and holders of a Bachelor's degree. This dependence vanishes when controlling for a full set of state and year fixed effect (in only 2 controls and 5 unobservables there is still a slight correlation). Then, after adding the 7 controls all the correlation disappears (1 control and 1 unobservable are still marginally correlated). These balancing regressions suggest that legislation instruments are as good as random after controlling for fixed effects and covariates, i.e. that are exogenous, not correlated with the error term that includes the omitted variable (moral attitudes), with other controls (race or party affiliation) and fixed effects (unobserved demographics or religious feelings) that explain the outcome (marriage rates); and that they are independent of the outcomes (marriage rates), and treatments (surrogacy) and can be used as the causal variable of this study's population model. Multiple figures with controls

are gathered in Figure 3.8.

3.8 Results

To answer the first question of the chapter, the first stages of the IV procedure are appropriate. There are eight IV regressions in Tables 3.8, 3.9, 3.10. All of them show a negative coefficient when negative laws are enforced, not significantly different from zero when ambiguous laws are enacted and a positive coefficient when positive laws are upheld. Regarding only the regression with marriage rate as the dependent variable with the controls and state and year fixed effects, the results show the step-like coefficients going from an average of a significant value of -2.521 surrogacies in states when surrogacy is banned by negative laws, to a non-significant value of 0.008 surrogacies in states when laws are ambiguous, and an increase with a value of 0.876 (not significant in this regression) when surrogacy is backed by positive laws. These coefficients show increases and decreases of surrogacy with respect to the omitted, probably unenforceable category. In summary, restrictive legislation reduces and permissive legislation adds surrogacies to the tally of each state. Additionally, although not shown, having time to change mind slightly reduces the surrogacy rate, and being uncompensated slightly increases the surrogacy rate. However, both variables are marginally significant.

Marriage and divorce rates express how many marriages and divorces happened annually for every 1,000 people in the overall population. The data in this paper shows that the overall average for marriage rates is 7.31, meaning that approximately 7 new marriages happened annually for every 1,000 people in each of the 52 states and the overall average for divorce rates is 3.65, which means that, annually, almost 4 additional divorces are consummated each year for every 1,000 people in each state.

IV regressions of the effect of surrogacy rate on marriage rates, births, out-of-wedlock births and divorce rates are performed. Surrogacy rates is an endogenous

variable instrumented by legislation dummies as previously explained. Table 3.8 shows two OLS regressions of marriage rate on surrogacy rate with state and year fixed effect with and without controls. Selected controls are: Black percentage of female population ages 15-39, Hispanic origin percentage of female population ages 15-39, GDP (per capita), party affiliation, Bachelor's degree, Black percentage and Hispanic origin percentage. For these specifications, bad controls are omitted, such as the following outcomes: births, birth rate, fertility rate, total fertility rate, teenage birth rate, women's weekly earnings and the proportion of women's weekly earnings to men's weekly earnings. The coefficients are negatively biased for OLS. After switching to a full IV regression, they are significant and higher. It is interesting to attempt to develop an explanation of the negative bias of columns (1)-(2). For simplicity, suppose that the surrogacy rate (treatment) is D, a dummy variable that indicates whether surrogacy is not practiced, D=0, or practiced, D=1, in the state. Take Michigan as an example of the former and California as an example of the latter. To have a negative bias, the control state, Michigan, needs to have more marriages, outcome $Y = Y^0|D = 0$, when not facing surrogacy than California in the non-observed potential outcome of also not facing surrogacy, $Y^0|D = 1$. This can be explained by moral attitudes. California, a state open to surrogacy, is likely to be a "liberal" state so it is also likely to have a lower marriage rate in the "what-if" condition of not practicing surrogacy. In other words, in an omitted variable bias setting where corr(variable of interest/surrogacy rate, omitted variable/moral attitudes) is negative while corr(dependent variable/marriage rate, omitted variable/moral attitudes) is positive, the selection bias is negative, causing downward bias. This selection bias disappears when the IV strategy, controls and fixed effects are used. At the same time, a LATE interpretation of the results is recommended. The extra marriages exhibited are for the compliers, those couples that in a legal environment of criminalized surrogacy do not enter a surrogacy contract but that, in a legal environment of enforceable surrogacy, do enter a surrogacy contract; that is, compliers are the ones who voluntarily enter a surrogacy contract (the observed treatment) when nudged by the instrument (a surrogacy friendly legislation, akin

to a voucher for entering a surrogacy contract). This variation in surrogacy provoked by the instrument legislation will be the core causality input for the observed outcome effect, namely variation in the marriage rate. The local average treatment effect is only for the compliers who happen to be numerous in the sample, arguably by the importance of the rule of law. Defiers ($D^{Leg=crim} = surr, D^{Leg=enf} = \overline{surr}$), always takers ($D^{Leg=crim} = surr, D^{Leg=enf} = surr$) and never takers ($D^{Leg=crim} = \overline{surr}, D^{Leg=enf} = \overline{surr}$) do not matter. For the reason given previously, the preferred regression is the IV procedure. The main specification produces a 0.007**, which represents 6.8 marriages per surrogacy[10] or a standardized effect of 5.38% over the marriage standard deviation (see Table 3.11) assuming 41.3 additional surrogacies for each additional point in surrogacy rate.[11]

Table 3.9 show births and out-of-wedlock births results. In the main specification, births are reduced -2.1** per surrogacy. This can be explained by the direct effect of more surrogacy-related marriages, which are less fertile than regular marriages. This accounts for just 1.63% of the standardized effect on the birth standard deviation (see Table 3.11), so the effect of surrogacy on the overall birth rate is tenuous. An out-of-wedlock birth effect of -1.6** per surrogacy or 3.16% of the standardized effect was also found.

Finally, Table 3.10 presents the divorce rate results. Again, the OLS regressions are negatively biased. The main specification, i.e. IV with the selected controls and fixed effects, reveals a significant 0.003** effect of surrogacy on divorce. Assuming an average population of 290 million and 1,058,500 new divorces each year in the

[10]The results are the same if the total number of marriages regressed, with the IV setting, on total number of surrogacies and controlling for total population. This is not surprising because these are only transformations, or changes in variable dimensions.

[11]To go from 0.007*** to 281.1***, first take the US population 290 million and multiply by the prevalence of marriages, 7.34 per 1,000 habitants, and divide the result by 53 states which produces the number of 40,162 marriages per state. Multiplying it by 0.007** results in 281.1** marriages per surrogacy rate. Moreover, an average of 4,137,796 births each year imply that an additional point in surrogacy rate equals to effective 41.3 surrogacies. This permits the transformation of the 281.1 marriages per surrogacy rate into the 6.8 marriages per surrogacy as 281.1/41.3. In a few words, the coefficient 0.007** in the marriage on surrogacy rate regression corresponds to 6.8** additional marriages per surrogacy or 281.1** additional marriages per surrogacy rate.

entire US territory, implies an addition of 0.97 divorces per surrogacy or a 4.21% standardized effect over the divorce standard deviation (see Table 3.11).

In summary, the number of marriages increases because people decide to get married even when selected infertile conditions are present. Surrogacy is a new technology in the production function of children. Thus, the pool of marriageable mates increases. Births are slightly reduced because newly formed surrogacy related marriages are less fertile than regular marriages. Out-of-wedlock births are also reduced because there are more marriages and because couples are, on average, less fertile. Divorces increase because some of the marriages promoted by the availability of surrogacy are less close after marriage has been consummated and children have been born and the legal requirement to get married in order to enter a surrogacy contract is no longer binding. The results of this study align with the predictions of the proposed model: with the emergence of surrogacy, men who love an infertile woman can now marry and also secure fertility. Also, couples must marry to enter a surrogacy contract due to legal regulations. These two mixed incentives promote more marriages (both groups marry more), but they also promote more divorces (couples who married just because of legal regulations are more prone to divorce following conception). In the data, the institutional restrictions, the model and the results, these two economic regularity are ubiquitous.

Lastly, Table 3.12 shows results for other outcomes: birthrate, fertility rate, total fertility rate and teenage birthrate. In the IV model with fixed effects and controls all are slightly negative. These results are also according to the predictions of the model: lower births (birthrate) and lower fertility (fertility rate, total fertility rate, teenage birth).

In order to test the results, six robustness checks were used with the marriage rate IV regression with controls and fixed effects as benchmark (Table 3.13). In the first two, results similar to the benchmark were expected because additions do not

challenge the benchmark; they only test its integrity. However, in the last four, only noise without any correlation was expected, because the benchmark is disrupted. To begin with, the additional control of absolute number of organ donations (total, live or dead) for each state and period are added. Each organ donation related variable is a proxy to the unobserved moral values and attitudes towards body organ promiscuity common in surrogacy and organ donation. The results show a positive, significant and sizeable coefficient which substantially backs the effect of surrogacy rate on marriage rate. Note that these organ donation variables are left in the unobserved only because of missing observations (there were 82 missing observations in total). If I had obtained access to the full set of observations, I would certainly have used the variables as another proxy for moral values. The result is 0.009** according to what was expected. I also attempt to use a lag variable to increase the predetermination of the controls, and again the results are strong and according to what was expected (0.011**). Then there are the four robustness checks where the benchmark is severely disrupted. First, surrogacy rate is replaced by fertility variables grouped as lowly correlated with surrogacy: surrogacy clinics (the number of clinics that use gestational carriers) and fresh non donor (only ART fertility treatment when gestational carriers are used, in 1% of the cycles).[12] As expected, no significance is found in the coefficient of either variable. Then fertility variables not correlated with surrogacy are tested: non surrogacy clinics (clinics that do not use gestational carriers in their fertility treatments), clinics (either surrogacy and non-surrogacy clinics) and fresh donor (another ART fertility treatment in which, at least in this sample, non-gestational carries are reported). The coefficients are all non-significant. This is reassuring because it shows that only the surrogacy rate has a sizeable effect on marriage rates while these placebo variables do not. Next, the population of each state and year, which is a demographic variable, is used to try to explain marriage. Again, no effect is found. Finally, the last robustness check changes the dependent variable to the marriage rate by state in the years 1990, 1995,

[12]In fact, the vast majority of the fertility treatments do not use a gestational carrier even if a clinic that uses them or a treatment where they are used is selected. This explains why these apparently related variables are indeed scarcely correlated with surrogacy

1999 and 2000 using the surrogacy rate and control variables for the years 2001 to 2004. So for surrogacy the same years were used as for the main specification, but for married the lagged periods were used. No effect is found again, which means that the surrogacy rate just explains the marriage rate for the same year. Using surrogacy rates of posterior periods as a placebo treatment delivered results with no significance at all.

3.9 Synthetic control

Regarding the initial results of the impact of the enactment of the Illinois Gestational Surrogacy Act, Table 3.14 shows that both the extensive margin (the number of clinics) and the intensive margin (the percentage of total cycles each clinic devotes to gestational surrogacy) increased when the law was effectively introduced on January 1st 2005. New clinics entered the market, some specialized exclusively in the surrogacy business. There seems to be a permanent effect of the law in both dimensions. Figure 3.9 also documents how the Illinois statute positively impacts the extensive margin. Also presented are Illinois' controls in Figure 3.10, which shows a trough in births, birth rates, fertility rate, total fertility rate, teenage birth rate, White births, Black births and Hispanic births, just after the onset of Illinois' surrogate legislation, which is what the paper's model predicts.

To test whether the Illinois Gestational Surrogacy Act effectively increased the surrogacies within that state after its enactment on January 1st 2005, a synthetic control procedure is implemented following Abadie et al. (2011).

First, Illinois is the treated state, and the donor pool consists of 28 remaining states; 23 states were discarded due to restrictive or permissive legislation being introduced during the window of the data (2001-2013). The excluded states are Alabama, California, Connecticut, Delaware, Idaho, Maryland, Massachusetts, Minnesota, New Mexico, New York, North Carolina, North Dakota, Ohio, Oklahoma,

Pennsylvania, Rhode Island, South Carolina, Tennessee, Texas, Utah, Washington, West Virginia and Wisconsin, and the list is consistent with the black numbers in Table 3.3 which refer to legislations introduced *during* the period. Table 3.15 reproduces state weights in the synthetic Illinois.

Second, 15 predictors of the surrogacies (the dependent variable) are measured: surrogacy clinics, overall fertility clinics and non-surrogacy clinics; the legislation dummies as criminalized, unenforceable, probably unenforceable, uncertain, probably enforceable and enforceable; fertility indicators as births, birth rate and fertility rate; and demographics of birth such as White birth percentage, Black birth percentage and Hispanic birth percentage. Table 3.16 shows the means of surrogacy predictor, while Table 3.17 reproduces predictor weights in the synthetic Illinois.

Third, Figure 3.11 exhibits the result of the trends in surrogacy in Illinois versus the synthetic Illinois. The figure shows an important increase after the beginning of 2005, with a peak of nearly 80 (a fourfold increase) surrogacies in 2010. At the end of the sample there is a downward trend that future research should explore and that may add a transient aspect of the effect. Similar results are shown in Figure 3.12 which is a gap graph depicting the gap between the number of surrogacies in Illinois and the synthetic Illinois before and after the enacting of the Illinois Gestational Surrogacy Act. Again, the results seem to favour an increase in surrogacies *caused* with the introduction of permissive legislation.

Finally, Figure 3.13 shows three placebo tests: the first one, Figure 3.13a is a placebo-in-outcome where the synthetic control code is run again but instead of using surrogacies as the dependent variables, a fake variable is used (non-surrogacy clinics). No effect is found. Next, Figure 3.13b is a placebo-in-region which repeats the procedure, but instead of Illinois, Alaska is the treated unit. Again, no effect is found. Last, Figure 3.13c is a placebo-in-time where the date of the introduction of the legislation is changed to a 2010 placebo legislation. Again, no effect is found.

To conclude, I present a permutation test which brings statistical rigour to the procedure by evaluating the significance of the estimates, posing the question of whether the results could be driven entirely by chance using the mean square prediction error (MSPE). What I found is a reasonable rejection of mere chance.

3.10 Conclusion

This chapter collects, for the very first time, a complete Surrogacy Data Set over thirteen years at the state level where the US Data is from the Centers for Disease Control and Prevention and reveals the incidence of surrogacy by state.

One of the goals of this chapter is, using this Surrogacy Data Set, to document how permissive and restrictive legislation have effectively regulated the number of surrogacies during the decade and for each US state. This shows the results of restrictive legislation, which criminalizes surrogacy or makes surrogacy contracts unenforceable and reduces the number of surrogacies, as well as the results of permissive legislation, which makes surrogacy contracts enforceable and increases the number of surrogacies. If more data becomes available, further studies should consider whether to isolate cross-border surrogacy.

Furthermore, evidence shows that surrogacy affects vital statistics (marriage, births, out-of-wedlock births and divorce). Using an IV procedure that exploits variation in surrogacy predicted by changes in legislation, the data shows that one additional surrogacy increases the number of marriages by $\approx$7 for each state, which means a 5.38% standardized effect, which implies that the effect of increasing marriages for low fertility women and the effect of compulsory marriage in surrogacy legislation are larger than the effect of substituting the institution of marriage with the new social and biologic institution of surrogacy. For this to happen it is crucial that my sample does not include same sex marriages or single persons wanting

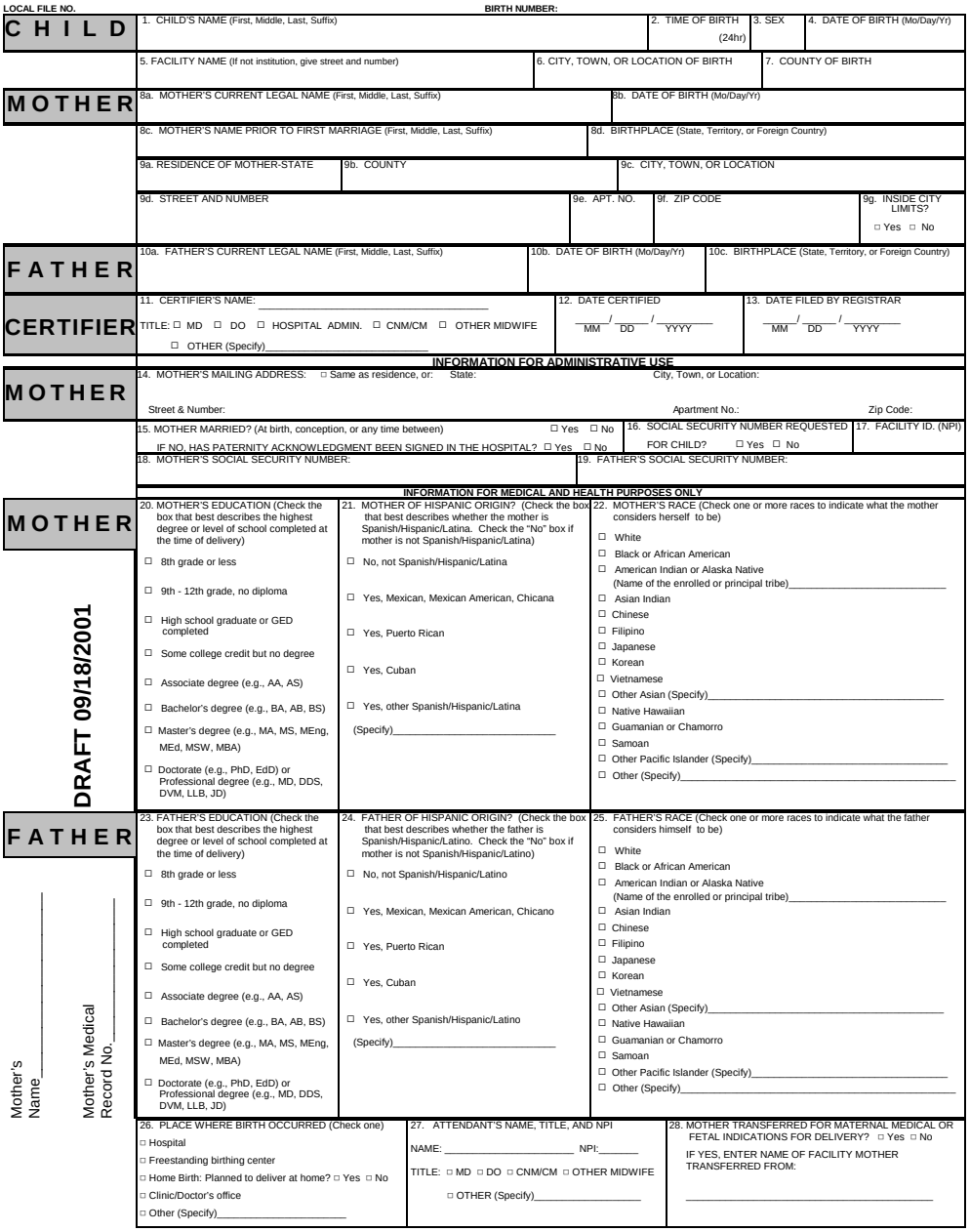
to have children without marrying. Moreover, there is evidence that an additional surrogacy increases divorces by ≈1 for each state, or 4.21% of a standardized effect, which implies that in my sample the effect of marriages of convenience as a consequence of surrogacy legislation requirement to get married for the commissioning parents over the insurance on low fertility marriages, and that it only marginally reduces births and out-of-wedlock births. Thus, the most significant contribution delivered by this paper is that the main effect of surrogacy on the vital statistics of the country is seen through marriage.

Moreover, these results align with the predictions of a model I developed, showing how the introduction of surrogacy as a new fertility technology can change the demographics of a given US state.

I tested my results with advanced econometric techniques, including a balancing test and six robustness check exercises, to show not only correlations but also causal relationships from legislation to surrogacy and from surrogacy to vital statistics.

Finally, it is of interest to underline that my data—and then my study—is inclined to what I have called a soft or conservative surrogacy where married heterosexual couples and only uterine infertility are permitted. Results confirm what the theoretical discussion predicted: more marriages, divorce and fewer births. At the same time, if surrogacy is extended to single persons, fertile women, gays and lesbians, the predictions of the model are at least theoretically tilted to less marriage and probably less divorce, with births being unclear. I deeply hope this study will trigger new data to become available to uncover the effect on surrogacy of key family vital statistics also in the hard form of surrogacy.

Figure 3.1: US birth certificate



U.S. STANDARD CERTIFICATE OF LIVE BIRTH

LOCAL FILE NO. BIRTH NUMBER:

CHILD

1. CHILD'S NAME (First, Middle, Last, Suffix)
2. TIME OF BIRTH (24hr)
3. SEX
4. DATE OF BIRTH (Mo/Day/Yr)
5. FACILITY NAME (If not institution, give street and number)
6. CITY, TOWN, OR LOCATION OF BIRTH
7. COUNTY OF BIRTH

MOTHER

8a. MOTHER'S CURRENT LEGAL NAME (First, Middle, Last, Suffix)
8b. DATE OF BIRTH (Mo/Day/Yr)
8c. MOTHER'S NAME PRIOR TO FIRST MARRIAGE (First, Middle, Last, Suffix)
8d. BIRTHPLACE (State, Territory, or Foreign Country)
9a. RESIDENCE OF MOTHER-STATE
9b. COUNTY
9c. CITY, TOWN, OR LOCATION
9d. STREET AND NUMBER
9e. APT. NO.
9f. ZIP CODE
9g. INSIDE CITY LIMITS? ☐ Yes ☐ No

FATHER

10a. FATHER'S CURRENT LEGAL NAME (First, Middle, Last, Suffix)
10b. DATE OF BIRTH (Mo/Day/Yr)
10c. BIRTHPLACE (State, Territory, or Foreign Country)

CERTIFIER

11. CERTIFIER'S NAME: _____
TITLE: ☐ MD ☐ DO ☐ HOSPITAL ADMIN. ☐ CNM/CM ☐ OTHER MIDWIFE
☐ OTHER (Specify)_____
12. DATE CERTIFIED _____/_____/_____ MM DD YYYY
13. DATE FILED BY REGISTRAR _____/_____/_____ MM DD YYYY

INFORMATION FOR ADMINISTRATIVE USE

MOTHER

14. MOTHER'S MAILING ADDRESS: ☐ Same as residence, or: State: City, Town, or Location:
Street & Number: Apartment No.: Zip Code:
15. MOTHER MARRIED? (At birth, conception, or any time between) ☐ Yes ☐ No
IF NO, HAS PATERNITY ACKNOWLEDGMENT BEEN SIGNED IN THE HOSPITAL? ☐ Yes ☐ No
16. SOCIAL SECURITY NUMBER REQUESTED FOR CHILD? ☐ Yes ☐ No
17. FACILITY ID. (NPI)
18. MOTHER'S SOCIAL SECURITY NUMBER:
19. FATHER'S SOCIAL SECURITY NUMBER:

INFORMATION FOR MEDICAL AND HEALTH PURPOSES ONLY

MOTHER

20. MOTHER'S EDUCATION (Check the box that best describes the highest degree or level of school completed at the time of delivery)
☐ 8th grade or less
☐ 9th - 12th grade, no diploma
☐ High school graduate or GED completed
☐ Some college credit but no degree
☐ Associate degree (e.g., AA, AS)
☐ Bachelor's degree (e.g., BA, AB, BS)
☐ Master's degree (e.g., MA, MS, MEng, MEd, MSW, MBA)
☐ Doctorate (e.g., PhD, EdD) or Professional degree (e.g., MD, DDS, DVM, LLB, JD)

21. MOTHER OF HISPANIC ORIGIN? (Check the box that best describes whether the mother is Spanish/Hispanic/Latina. Check the "No" box if mother is not Spanish/Hispanic/Latina)
☐ No, not Spanish/Hispanic/Latina
☐ Yes, Mexican, Mexican American, Chicana
☐ Yes, Puerto Rican
☐ Yes, Cuban
☐ Yes, other Spanish/Hispanic/Latina
(Specify)_____

22. MOTHER'S RACE (Check one or more races to indicate what the mother considers herself to be)
☐ White
☐ Black or African American
☐ American Indian or Alaska Native (Name of the enrolled or principal tribe)_____
☐ Asian Indian
☐ Chinese
☐ Filipino
☐ Japanese
☐ Korean
☐ Vietnamese
☐ Other Asian (Specify)_____
☐ Native Hawaiian
☐ Guamanian or Chamorro
☐ Samoan
☐ Other Pacific Islander (Specify)_____
☐ Other (Specify)_____

DRAFT 09/18/2001

FATHER

23. FATHER'S EDUCATION (Check the box that best describes the highest degree or level of school completed at the time of delivery)
☐ 8th grade or less
☐ 9th - 12th grade, no diploma
☐ High school graduate or GED completed
☐ Some college credit but no degree
☐ Associate degree (e.g., AA, AS)
☐ Bachelor's degree (e.g., BA, AB, BS)
☐ Master's degree (e.g., MA, MS, MEng, MEd, MSW, MBA)
☐ Doctorate (e.g., PhD, EdD) or Professional degree (e.g., MD, DDS, DVM, LLB, JD)

24. FATHER OF HISPANIC ORIGIN? (Check the box that best describes whether the father is Spanish/Hispanic/Latino. Check the "No" box if mother is not Spanish/Hispanic/Latino)
☐ No, not Spanish/Hispanic/Latino
☐ Yes, Mexican, Mexican American, Chicano
☐ Yes, Puerto Rican
☐ Yes, Cuban
☐ Yes, other Spanish/Hispanic/Latino
(Specify)_____

25. FATHER'S RACE (Check one or more races to indicate what the father considers himself to be)
☐ White
☐ Black or African American
☐ American Indian or Alaska Native (Name of the enrolled or principal tribe)_____
☐ Asian Indian
☐ Chinese
☐ Filipino
☐ Japanese
☐ Korean
☐ Vietnamese
☐ Other Asian (Specify)_____
☐ Native Hawaiian
☐ Guamanian or Chamorro
☐ Samoan
☐ Other Pacific Islander (Specify)_____
☐ Other (Specify)_____

Mother's Name_____ Mother's Medical Record No._____

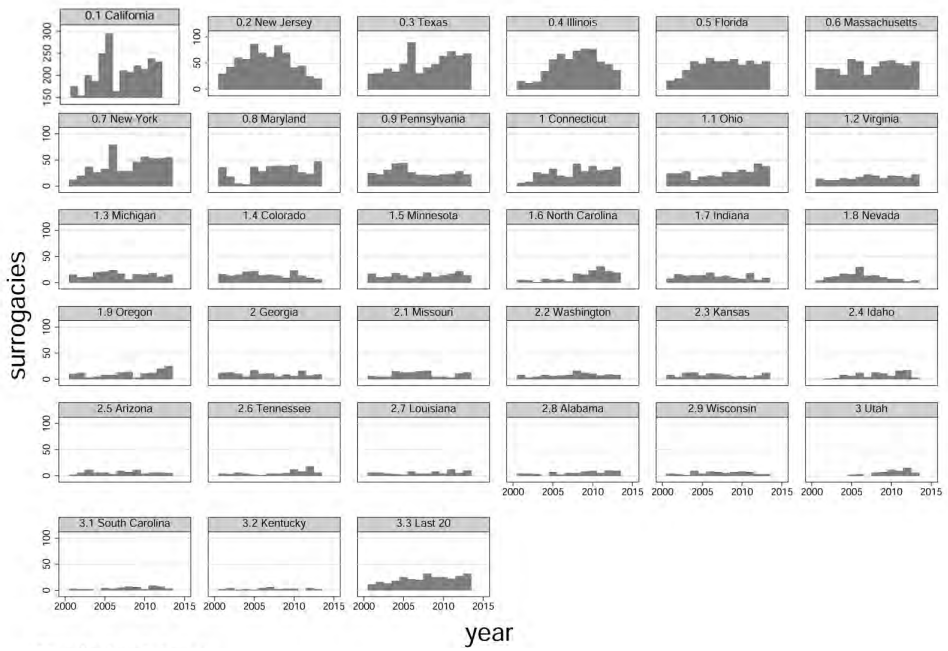
26. PLACE WHERE BIRTH OCCURRED (Check one)
☐ Hospital
☐ Freestanding birthing center
☐ Home Birth: Planned to deliver at home? ☐ Yes ☐ No
☐ Clinic/Doctor's office
☐ Other (Specify)_____

27. ATTENDANT'S NAME, TITLE, AND NPI
NAME: _____ NPI:_____
TITLE: ☐ MD ☐ DO ☐ CNM/CM ☐ OTHER MIDWIFE
☐ OTHER (Specify)_____

28. MOTHER TRANSFERRED FOR MATERNAL MEDICAL OR FETAL INDICATIONS FOR DELIVERY? ☐ Yes ☐ No
IF YES, ENTER NAME OF FACILITY MOTHER TRANSFERRED FROM:
_____

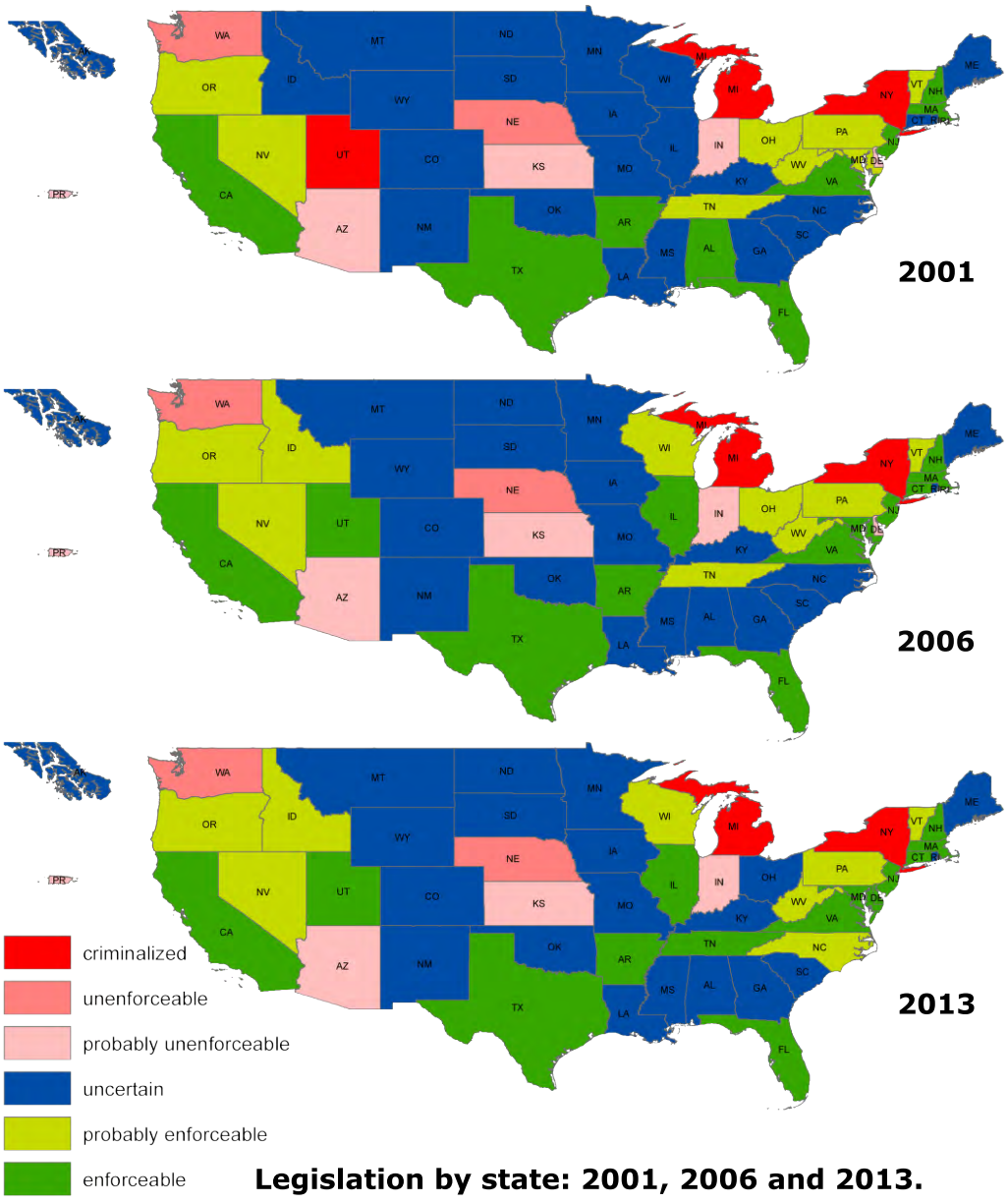
Note(1): Rev. 11/2003. **Note(2): Source:** CDC (Centers for Disease Control and Prevention, US Government). For more information see http://www.cdc.gov/nchs/nvss/vital_certificate_revisions.htm

Figure 3.2: A graphic snapshot of the Surrogacy Data Set: Surrogacies by state and sorted by average number of surrogacies



Note(1): Source: Author's calculation, obtained by collapsing the Clinic Table Data 2001-2013 from the Centers for Disease Control and Prevention, US Government. Data are available on-line at www.CDC.gov/ART/ARTReports.htm.**Note (2):** This graphic snapshot of the Surrogacy Data Set shows the evolution of the variable surrogacy. To calculate this variable by state I multiply the total fresh non-donor cycles started at each clinic by the gestational rate (the percentage of fresh non-donor cycles using a gestational carrier over the total fresh non-donor cycles started) reported by each clinic in each of the fifty states plus the District of Columbia and Puerto Rico. Before collapsing I adjust the <1% rate to match the number of surrogacies with the official national count from the Centers for Disease Control and Prevention, this is both a calculation but also a control using all available information. I repeat the process for each of the thirteen years of data. Then I sort the data by the average number of surrogacies in each state and group the twenty states with the fewest surrogacies.**Note (3):** California, the most surrogacy friendly state with an average of 206 surrogacies each year, is shown with a different scale to allow the other states to have a visible trend. All the other states including the last 20 aggregated states use the same scale.

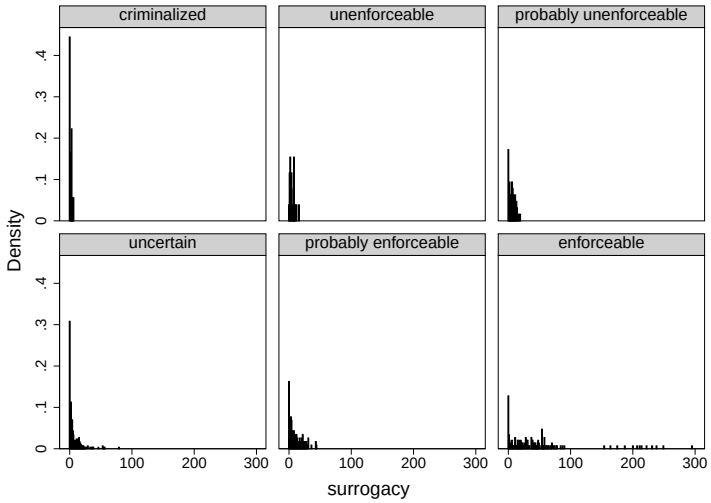
Figure 3.3: Surrogacy legislation



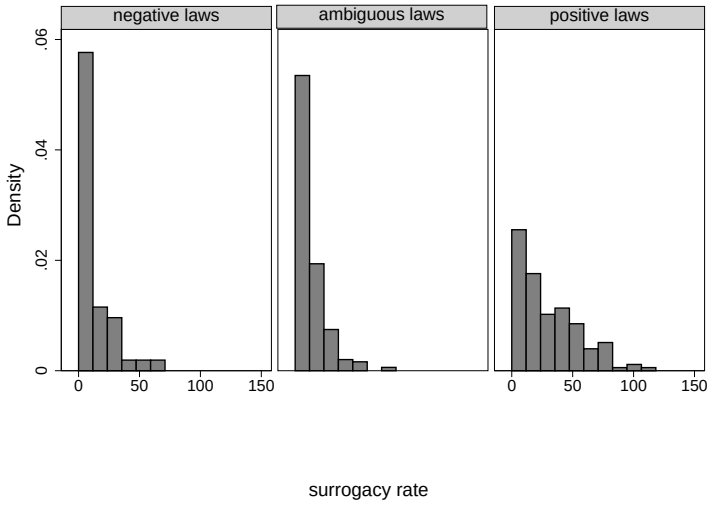
Note (1): Source: Own calculation from data on US surrogacy legislation by state for the period 2001 to 2013 in the academic on-line version of LexisNexis—www.lexisnexis.com/hottopics/lnacademic/—in conjunction with Creative Family Connections LLC website CreFamCon (2015) and Darra L. Hofman's key article, "Mama's Baby, Daddy's Maybe: A State-by-State Survey of Surrogacy Laws and Their Disparate Gender Impact" (2009). Each register verifies the enacting year of surrogacy statutory legislations and court decisions. I have established six categories: *criminalized, unenforceable, probably unenforceable, uncertain, probably enforceable,* and *enforceable.*
Note (2): These are three consecutive snapshots of geographic representations of surrogacy legislation by states in the US. Alaska is partially shown. Puerto Rico is shown disregarding its natural location. Hawaii is not shown. New York is shown as criminalized because compensated surrogacy is banned: Anyone who signs a surrogacy contract risks a fine of up to $10,000. Facilitators of these surrogacy contracts—surrogacy agencies and barristers—risk a fine and if it is a repeated offence, are guilty of felony. Nevertheless, in the Surrogacy Data Set I code the state as uncertain because a prosperous uncompensated market is permitted and in full-practice. Michigan is a similar case: While compensated surrogacy is illegal, uncompensated one is permitted.

Figure 3.4: Surrogacy/surrogacy rate by legislation

(a) Surrogacy by legislation



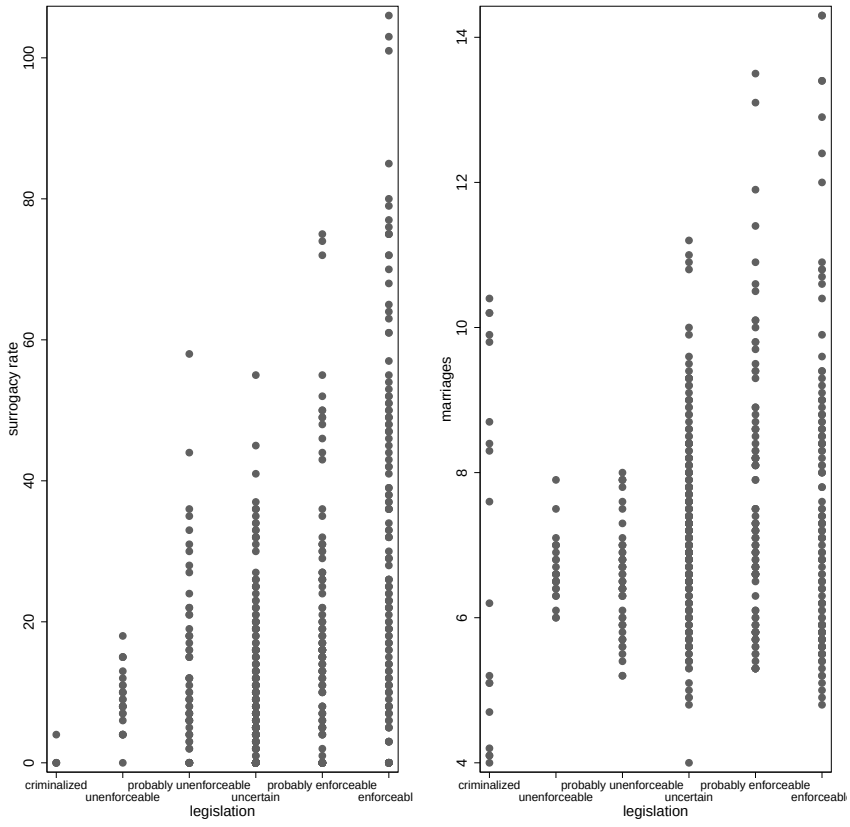
(b) Surrogacy rate by legislation



Note (1): Source: Author's calculations from CDC and LexisNexis source.

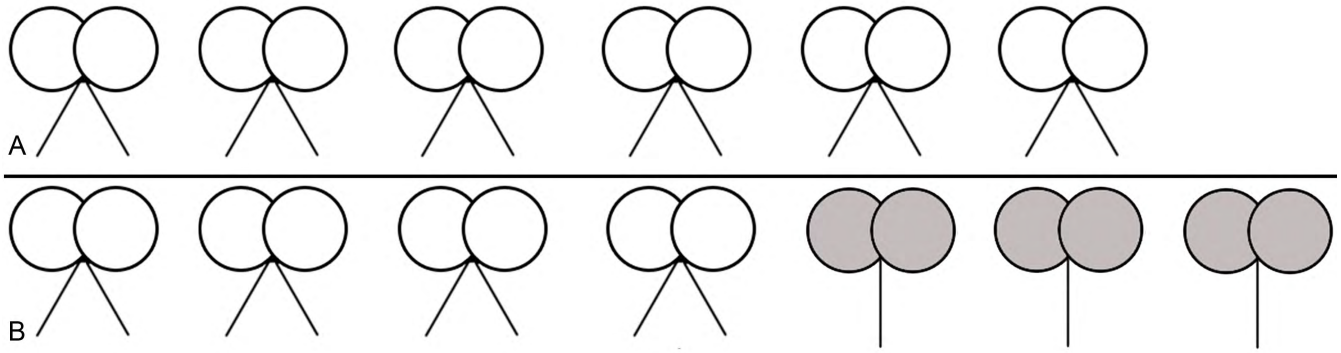
Figure 3.5: First stage and reduced form

(a) First stage (b) Reduced form



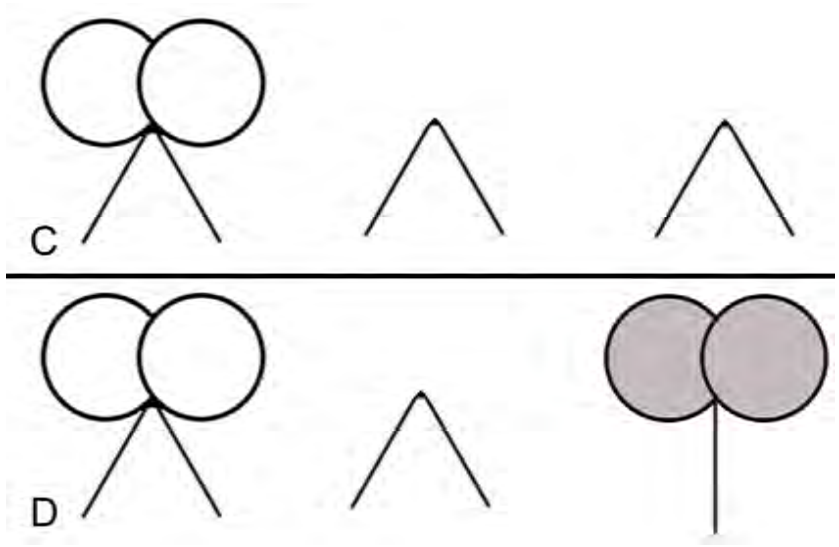
Note (1): Source: Author's calculations from CDC and LexisNexis source. Graphs show a visual approach to the identification strategy. Seen are the first stage—surrogacy rate on legislation—(DC not shown) and the reduced form—marriage rate on legislation—(NV and HI not shown) of a subsequent IV strategy.

Figure 3.6: More marriages and fewer married births after surrogacy has been introduced



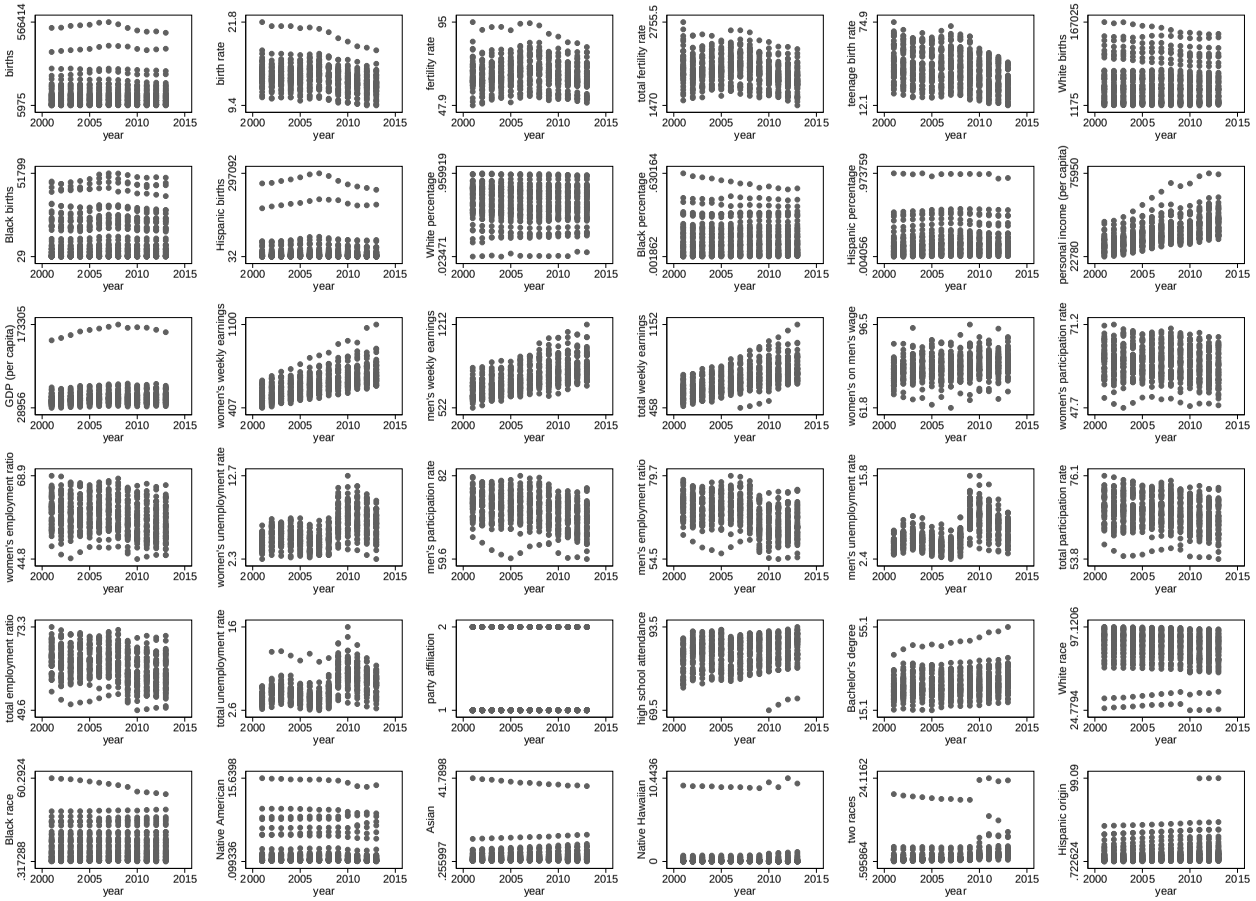
On top (A) just fertile man and fertile woman marriages. At the bottom (B) surrogacy has been introduced causing marriages to increase (now 7 and before only 6) and births to decrease (now 11 and before 12)

Figure 3.7: Fewer unmarried births after surrogacy has been introduced



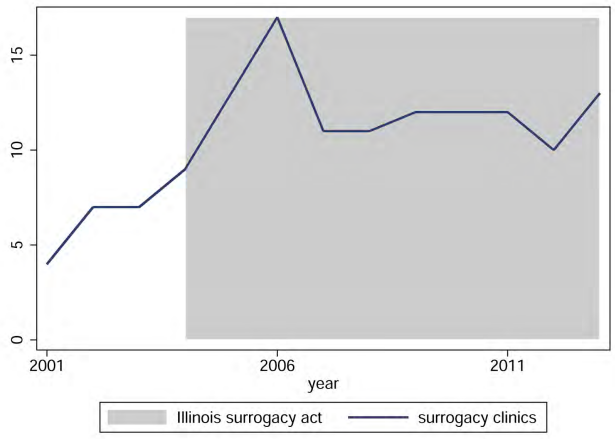
On top (C) married and unmarried children co-exists. At the bottom (D) surrogacy has been introduced causing marriages to increase (2 now, before only 1), this effect causes out-of-wedlock births to decrease (2 now, before 4)

Figure 3.8: Controls and unobservables by state for the whole US and the whole period 2001-2013



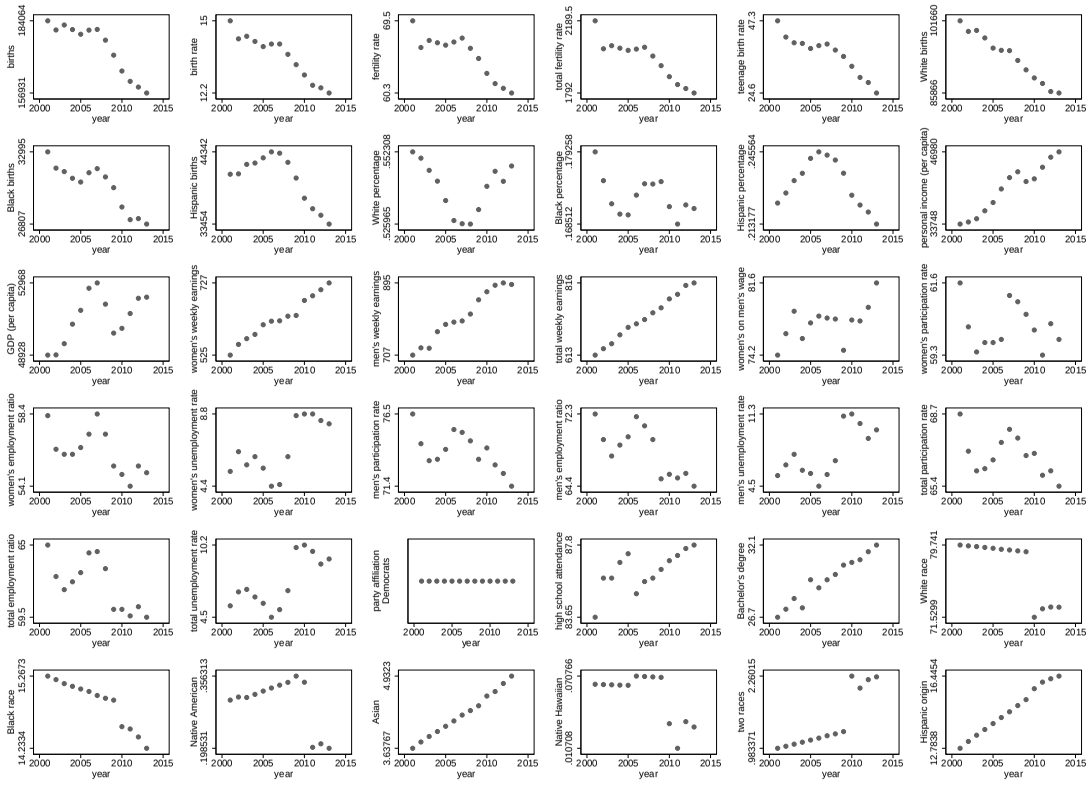
Footnote for Figure Controls and unobservables by state for the whole US and the whole period 2001-2013 Note (1): Source: Author's calculations. Data are available on-line at www.CDC.gov. Controls—graph by state for the whole US for the whole period—are births (number of live births, range 5,975-566,414), birth rate (number of live births per 1,000 of a population in a particular year, range 9.4-21.8), (general) fertility rate (number of births per 1,000 women between the ages of 15 and 44 in a particular year, range 47.9-95), total fertility rate (TFR) (number of births that a cohort of 1,000 women would have if they experienced throughout their childbearing years the same age-specific birth rates observed in a given year, range 1,470-2,755.5), teenage birthrate (number of births per 1,000 women between the ages of 15 and 19 in a particular year, range 12.1-74.9), White births (number of White live births, range 1,175-167,025), Black births (number of Black live births, range 29-51,799), Hispanic births (number of Hispanic live births, range 32-297,092), percentage of female population age 15-39 and race percentage (White(range 0.023-0.959), Black (range 0.001-0.630) and Hispanic (range 0.004-0.973)), personal income (per capita, range $22,780-$75,950), GDP (per capita, range $28,956-$173,305), women's (range $407-$1,100), men's (range $522-$1,212) and total (range $458-$1,152) weekly earnings, women's participation rate (range 47.7-71.2), employment ratio (range 44.8-68.9) and unemployment rate (range 2.3-12.7), men's participation rate (range 59.6-82), employment ratio (range 54.5-79.7) and unemployment rate (range 2.4-15.8), total participation rate (range 53.8-76.1), employment ratio (range 49.6-73.3) and unemployment rate (range 2.6-16), party affiliation (Republican=1, Democrat=2), high school attendance (range 69.5-93.5), Bachelor's degree (range 15.1-55.1), White race (range 24.77-97.12), Black race (range 0.31-60.29), Native American (range 0.09-15.63), Asian (range 0.25-41.78), Native Hawaiian (range 0-10.44), two races (range 0.59-24.11), Hispanic origin (range 0.72-99.09).

Figure 3.9: Effect of the change in surrogacy legislation in Illinois on the number of clinics



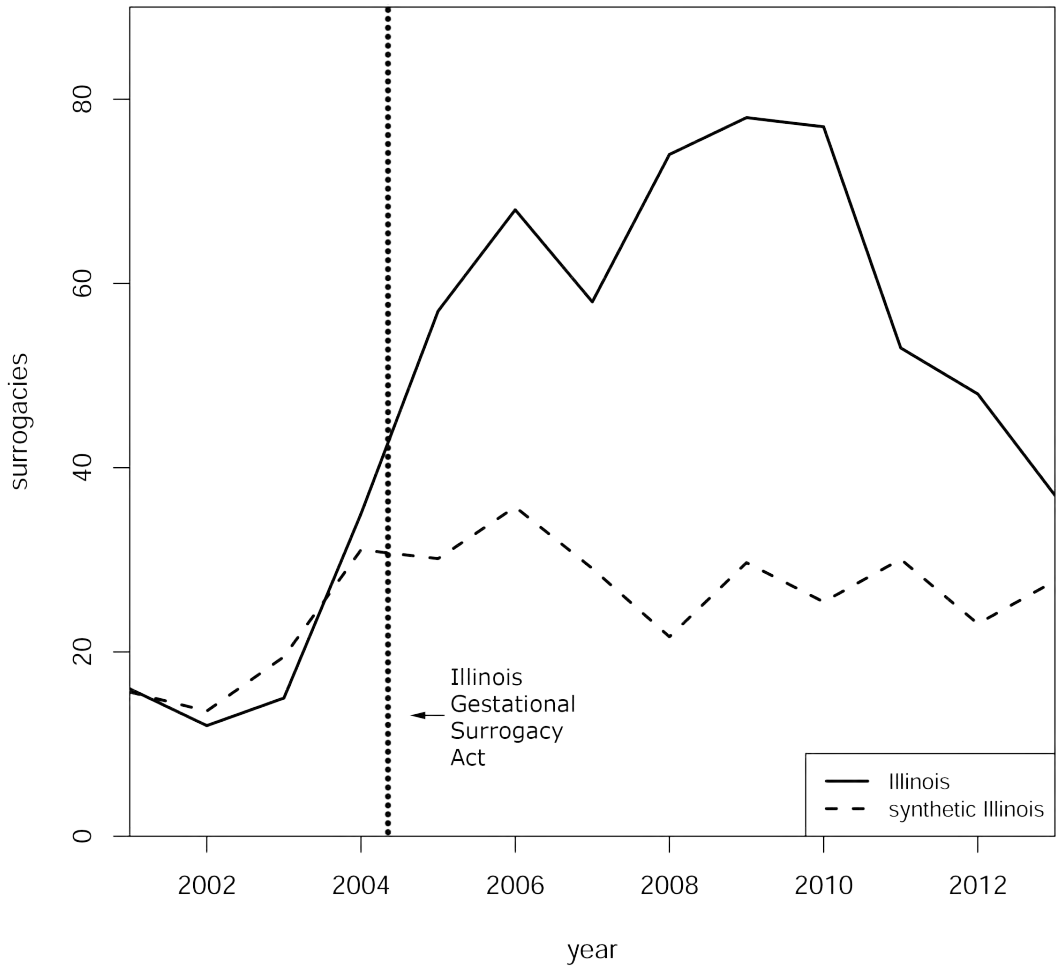
Note (1): Source: Author's calculation. The number of surrogacy clinics—extensive margin—is assessed from the Clinic Table Data of the Centers for Disease Control and Prevention, US Government from 2001-2013. Data are available on-line at www.CDC.gov/ART/ARTReports.htm. The legislation change corresponds to the enactment of the Illinois Gestational Surrogacy Act on January 1st 2005.

Figure 3.10: Controls for Illinois 2001-2013



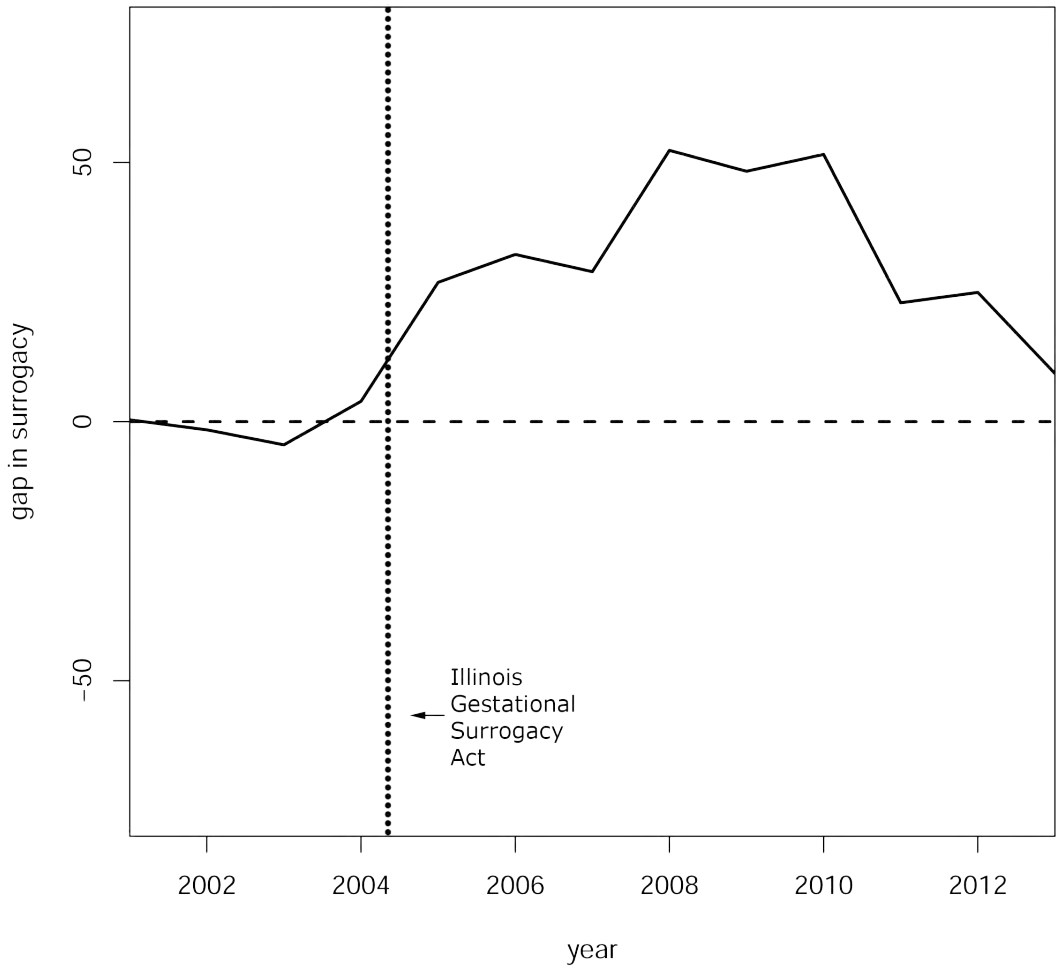
Note (1): Source: Author's calculation. The breaks are only apparent. It is just the differential zooming in each cell of the graph. For instance, the break in 2010 Native American is only one and a half decimal points. So it is of not interest.

Figure 3.11: Trends in surrogacy: Illinois vs. synthetic Illinois



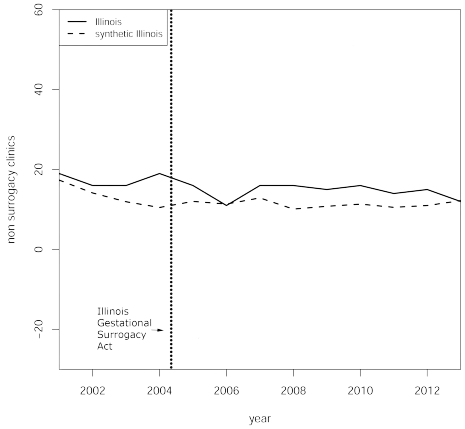
Note (1): Synthetic control: Illinois' surrogacies vs. synthethic Illinois' surrogacies. Trends in surrogacy before and after 01/01/2005 when a new permissive legislation was introduced in the state of Illinois, a statute called the Illinois Surrogate Act.

Figure 3.12: Gaps in surrogacy between Illinois and synthetic Illinois

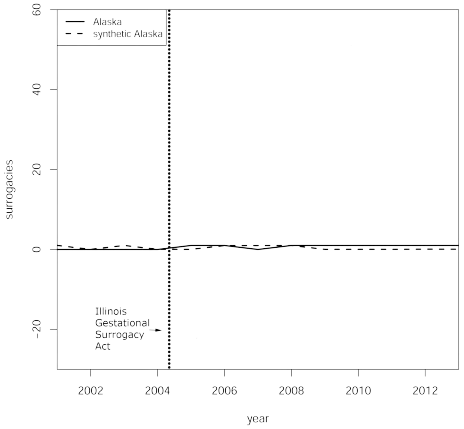


Note (1): Gaps in surrogacy between Illinois and synthetic Illinois before and after 01/01/2005 when a new permissive legislation was introduced in the state of Illinois, a statute called the Illinois Surrogate Act.

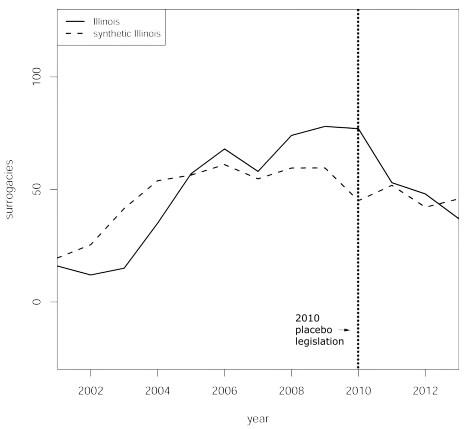
Figure 3.13: Placebo tests



(a) Placebo-in-outcome

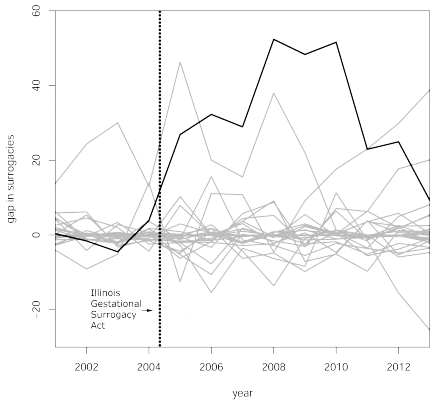


(b) Placebo-in-region

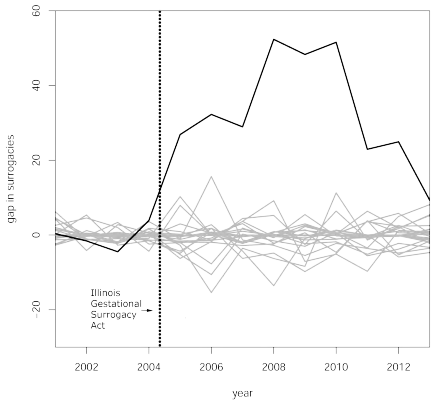


(c) Placebo-in-time

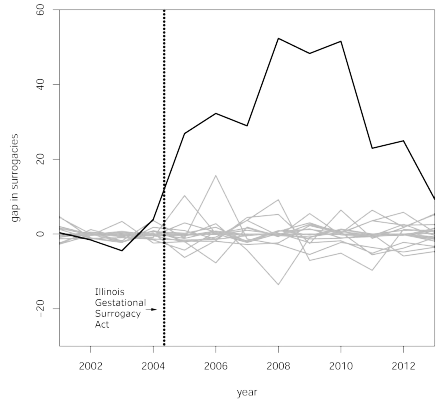
Figure 3.14: Permutation test



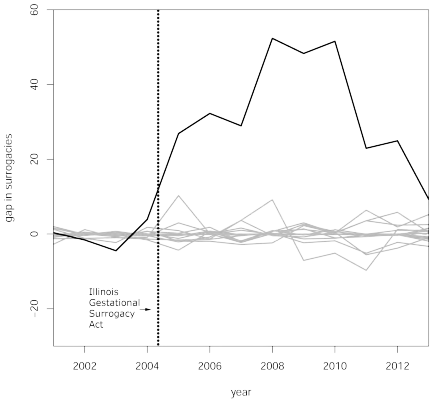
(a) All states in the donor pool



(b) State MSPE $\leq$ 2 times Illinois MSPE



(c) State MSPE $\leq$ 1 times Illinois MSPE



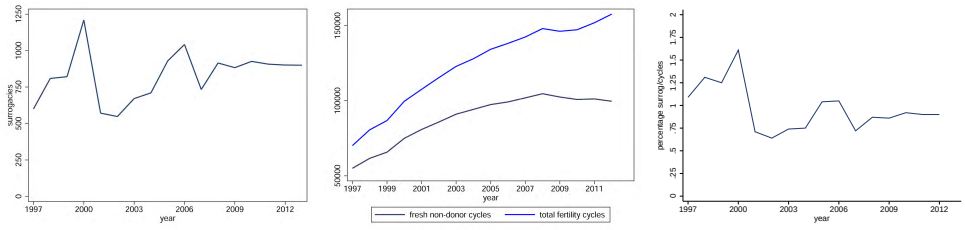
(d) State MSPE $\leq$ half times Illinois MSPE

Figure 3.15: Surrogacy basic facts at the national level

(a) Surrogacies

(b) Fresh non-donor and total fertility cycles

(c) Percentage surrog/cycles



Note (1): Source: Author's calculations. National level data from the official yearly reports of the Centers for Disease Control and Prevention, US Government from 1997-2013 (fresh non-donor and total fertility cycles only available from 1997-2012). Data are available on-line at www.CDC.gov for recent years. Previous years are only available through the Wayback Machine—archive.org/web/web.php—at the same internet address. These values represent country level data. See also the Surrogacy Data Set for state level data.

Table 3.1: A snapshot of the Surrogacy Data Set

	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Alabama	4	4	3	0	7	2	5	5	8	9	5	10	9
Alaska	0	0	0	0	0	0	1	0	0	0	0	0	0
Arizona	2	6	11	6	6	3	9	7	11	4	6	6	5
Arkansas	2	1	0	0	11	2	0	2	0	0	1	1	1
California	175	154	200	187	249	295	164	211	207	222	214	238	231
Colorado	16	13	15	21	22	14	15	13	9	23	13	9	6
Connecticut	6	8	26	23	33	20	17	43	28	38	30	31	37
Delaware	3	4	5	4	0	0	0	7	1	2	0	2	0
District of Columbia	0	2	0	0	0	2	3	3	3	3	6	5	4
Florida	17	21	37	54	49	60	54	54	58	47	55	48	54
Georgia	12	13	10	5	17	10	11	5	9	6	16	7	9
Hawaii	0	0	1	1	0	2	0	2	1	2	1	2	2
Idaho	0	1	2	8	6	12	4	13	11	6	16	17	3
Illinois	16	12	15	35	57	68	58	74	78	77	53	48	37
Indiana	8	16	13	14	14	19	11	13	9	7	18	5	9
Iowa	1	2	0	2	1	4	1	2	3	2	4	8	2
Kansas	8	4	12	13	7	11	10	9	5	6	3	7	12
Kentucky	2	4	1	2	1	4	6	2	3	3	0	4	2
Louisiana	6	6	4	3	2	8	4	4	8	4	12	5	10
Maine	0	0	0	0	0	0	0	0	0	0	0	0	2
Maryland	36	18	5	3	37	28	38	39	38	40	26	23	47
Massachusetts	41	39	39	28	58	54	28	44	54	56	50	46	54
Michigan	15	10	11	20	21	24	17	6	16	15	18	11	15
Minnesota	17	10	11	18	14	8	12	19	11	14	17	22	14
Mississippi	0	4	1	0	0	2	2	2	1	0	0	0	1
Missouri	6	5	5	15	13	13	15	16	5	5	3	11	13
Montana	0	0	0	0	0	0	0	0	0	0	0	0	2
Nebraska	2	2	1	4	3	4	3	2	1	1	2	0	4
Nevada	4	10	12	17	16	30	13	14	10	7	7	2	4
New Hampshire	0	0	0	0	0	1	0	0	0	0	0	0	0
New Jersey	30	43	61	58	87	70	62	84	70	42	45	25	21
New Mexico	1	0	0	3	0	1	2	1	2	2	0	0	0
New York	12	20	37	26	33	79	29	29	46	56	53	53	55
North Carolina	5	4	1	7	5	6	2	17	15	23	31	22	19
North Dakota	0	0	0	1	3	0	4	3	1	5	1	2	1
Ohio	24	24	28	11	18	20	18	27	26	31	27	43	38
Oklahoma	0	0	0	0	0	0	0	3	1	2	2	0	1
Oregon	10	12	3	5	8	8	13	14	3	11	12	20	25
Pennsylvania	25	22	31	43	44	26	21	21	20	22	23	28	22
Puerto Rico	2	0	0	0	3	2	0	1	2	0	0	1	0
Rhode Island	0	0	4	0	4	0	2	4	3	4	0	2	4
South Carolina	3	2	2	0	4	3	5	7	6	2	9	7	3
South Dakota	0	1	0	3	0	0	0	0	3	1	4	4	5
Tennessee	4	3	6	4	2	1	4	4	5	12	8	18	6
Texas	30	31	40	34	49	90	31	42	48	64	73	65	69
Utah	0	0	0	0	2	3	0	6	7	11	9	15	6
Vermont	0	0	1	0	0	1	2	0	3	1	0	0	3
Virginia	14	11	11	15	13	17	22	20	15	20	19	16	22
Washington	8	3	5	8	6	7	8	16	12	10	7	9	8
West Virginia	0	0	0	0	0	0	0	0	0	0	1	0	0
Wisconsin	4	3	1	9	4	8	7	5	7	8	7	3	3
Wyoming	0	0	0	0	0	0	0	0	0	0	0	0	0
Total	571	548	671	710	929 (3)	1042	733	915	883	926	907	901	900

Notes (1): Source: Author's calculation from the Clinic Data Table of the Centers for Disease Control and Prevention, US Government. Data are available on-line at www.CDC.gov/ART/ARTReports.htm.
Note (2) This snapshot of the Surrogacy Data Set documents the number of surrogacies by state. In particular shows the surrogacies for the years 2001-2013. To calculate this variable by state I multiply the total fresh non-donor cycles started at each clinic by the gestational rate (the percentage of fresh non-donor cycles using a gestational carrier over the total fresh non-donor cycles started) reported by each clinic in each of the fifty states plus the District of Columbia and Puerto Rico. Before collapsing I adjust the <1% rate to splice the number of surrogacies with the official national count from the Centers for Disease Control and Prevention. This is both a calculation but also a control using all available information. I repeat the process for each of the thirteen years of data. **The complete Surrogacy Data Set is available upon request from the author.**
Note (3): A clinic in Hawaii is reported to have undertaken 83 surrogacies in 2005. It seems an outlier. With the Hawaii surrogacies the count for 2005 will be the official 1,012 surrogacies and not the 929 shown here.

Table 3.2: Cost-benefit analysis.[1]

Marriage

Husband	**Wife**
Benefits	
Be in the birth certificate	Insurance for specialization in household production
Costs	
$25,000[3]	
Breadwinner[2] for wife and children	Unpaid household production

Surrogacy

Father	**Mother**
Benefits	
Be in the birth certificate. Continuation of the bloodline. (Better) substitute for adoption.	Be in the birth certificate. Overcome a female infertility. Continuation of the bloodline. (Better) substitute for adoption. Earn a salary and do not be subject to (unpaid) maternity leave. Do not risk adverse effects of pregnancy (physical, mental or esthetic). Male breadwinner for child.
Costs	
$100,000[4]	
Breadwinner for child	Lose maternal attachment

Divorce

Ex-husband	**Ex-wife**
Benefits	
Be single or build a new family	Be single or build a new family. Get children custody. Alimony and children support. Stop unpaid household production not related to the rearing of children.
Costs	
$20,000[5]	
Alimony and children support. Lose child custody. Children's stress.	Raising children on her own. Children's stress.

[1] **Source:** Compiled by author based on various U.S. sources including facts, customs, traditions, court rulings, statutory laws and expectations.
[2] Assumption of a male breadwinner is just for simplicity. Nothing fundamental changes if a dual-earner (in fact more common, ∼60%) or a female breadwinner is assumed instead.
[3] Cost of wedding, www.costofwedding.com
[4] Circle surrogacy, www.circlesurrogacy.com/costs
[5] Forbes, www.forbes.com/2006/11/07/divorce-costs-legal-biz-cx_lh_1107legaldivorce.html

Table 3.3: Surrogacy legislation state-by-state

Surrogacy Contracts are...	Criminalized	Unenforceable	Probably unenforceable	Uncertain	Probably enforceable	Enforceable
Alabama					2008	
Alaska				1989		
Arizona			1989			
Arkansas						1989
California						2005
Colorado				1987		
Connecticut						2002
Delaware			1988			2013
District of Columbia	1992					
Florida						1993
Georgia				NO		
Hawaii				NO		
Idaho					2004	
Illinois						2005
Indiana			1997			
Iowa				1989		
Kansas			1996			
Kentucky				1981		
Louisiana				1987		
Maine				NO		
Maryland					2000	2005
Massachusetts						2001
Michigan	1988					
Minnesota				2007		
Mississippi				NO		
Missouri				1997		
Montana				NO		
Nebraska		1988				
Nevada					1993	
New Hampshire						1990
New Jersey						2000
New Mexico				2005		
New York	2004					
North Carolina					2008	

Surrogacy Contracts are...	Criminalized	Unenforceable	Probably unenforceable	Uncertain	Probably enforceable	Enforceable
North Dakota				2005		
Ohio					2007	
Oklahoma				2005		
Oregon					1997	
Pennsylvania					2006	
Puerto Rico				NO		
Rhode Island				2007		
South Carolina				2003		
South Dakota				NO		
Tennessee						2009
Texas						2003
Utah	1989					2005
Vermont					1999	
Virginia						2000
Washington		2002				
West Virginia					2001	
Wisconsin					2003	
Wyoming				NO		

Note (1): Source: I take an on-line survey of Academic Lexis-Nexis—www.lexisnexis.com/hottopics/lnacademic/—in conjunction with Creative Family Connections LLC website (CreFamCon, 2015) and Darra L. Hofman (2009)'s key article, "Mama's Baby, Daddy's Maybe: A State-by-State Survey of Surrogacy Laws and Their Disparate Gender Impact" retrieving the year when the legislation change is introduced, which is the year in this table. With this information I create seventy-eight dummies, one for each year of data for surrogacy at the state level, i.e., 2001-2013, and for each category of legislation. Whenever the date of legislation is before the 2001-2013 period—coded as blue dates—the dummy takes the value of one for the whole period and for the corresponding category of legislation; similarly, the dummies *uncertain* have a value of one for the whole period when no legislation has been introduced, coded as a cyan NO. When I allocate a restrictive legislation—*criminalized, unenforceable, probably unenforceable,* and *uncertain*—I code ones before and in the year of the change in legislation and zeros afterwards. When dealing with a permissive category—*probably enforceable* and *enforceable*—I do the opposite: I code zeros before and in the year of the change in legislation and ones afterwards. This so far when I deal with statutory laws. In contrast, when dealing with common laws, I code five years around the year of the court decision using a pattern similar to that described above. When there is no legislation and then a statute or court law is passed I do not label the previous NO period.**Note (2):** Note that during the period of analysis (2001-2013) there are only two case where the legislation changes from one rather extreme category to the contrary: In Utah from 1989 surrogacy was criminalized but from 2005 it was enforceable. Delaware had surrogacies probably unenforceable, but after a 2013 statute they have become enforceable. In all other cases there was no change in the legislation category. For instance, California had an enforceable court decision in 2005: The California Supreme Court decided three companion cases that concerned lesbian couples who had reproduced via surrogacy, Elisa B. v. Superior Court, Kristine H. v. Lisa R. and K.M. v. E.G. The court held that under the Uniform Parentage Act, two women can be the legal parents of a child produced through surrogacy. Before that California had had several court decisions in the same direction: Johnson v. Calvert 1993, Myers v. Moschella 1996, and Buzzanca v. Buzzanca 1998. Besides, the 2005 court decision, just the last court decision, in 1998, has an effect in the period of analysis because of the five years around the year of the court decision rule, in this case of multiple and sequential same direction court decisions—as well, in a few similar cases—I have preferred to code the court decisions as a whole row of ones in the appropriate category of legislation. Tennessee is another case, it had court decisions in 2009, 2003, 2002, 2001, 1998, 1997, 1996, and 1995. Six of the seven has at least some effect in our period of analysis and has been coded according to the last criteria with a row of ones in the appropriate legislation category. When the court decisions are in the same direction I show only the year of the last decision for clarity and give the previous years here for completeness: Alabama 1996; California 1998, 1996, and 1993; Idaho 1986; Kansas 1982 (attorney general opinion); Massachusetts 1998; New Mexico 2001, and 1993; New York 1999 (law banning commercial surrogacy); Ohio 2001, 1999, 1994, and 1992; Oregon 1989 (attorney general opinion); Pennsylvania 1997; Tennessee 2003, 2002, 2001. 1998, 1997, 1996 and 1995; Virginia 1991; Wisconsin 2003.**Note (3):** Finally, in Minnesota in the year 2008 a Surrogacy Statute—similar to the Illinois Gestational Act—was passed by the Legislature but was then vetoed by Republican Governor Tim Pawlenty.

Table 3.4: Surrogacy legislation state-by-state: Uncompensated and time to change mind

Surrogacy Contracts are...	Uncompensated	Time to Change Mind	*Surrogacy Contracts are...*	Uncompensated	Time to Change Mind
Alabama			Montana		
Alaska			Nebraska	x	
Arizona			Nevada	x	
Arkansas			New Hampshire	x	x
California			New Jersey	x	
Colorado			New Mexico		
Connecticut			New York	x	
Delaware			North Carolina	x	
District of Columbia			North Dakota		
Florida	x	x	Ohio		
Georgia			Oklahoma	x	
Hawaii			Oregon	x	
Idaho			Pennsylvania		
Illinois			Puerto Rico		
Indiana			Rhode Island		
Iowa			South Carolina		
Kansas			South Dakota		
Kentucky	x		Tennessee		
Louisiana	x		Texas		
Maine			Utah		
Maryland	x		Vermont		
Massachusetts		x	Virginia	x	x
Michigan	x		Washington	x	
Minnesota			West Virginia		
Mississippi			Wisconsin		x
Missouri			Wyoming		

Note(1): Source: Same source as the previous table.
Note (2): In the previous table, New York is shown as criminalized because compensated surrogacy is banned: Anyone who signs a surrogacy contract risks a fine of up to $10,000. Facilitators of these surrogacy contracts—surrogacy agencies and barristers—risk a fine and if it is a repeated offence, are guilty of felony. Nevertheless, I code the state as uncertain because a prosperous uncompensated market is permitted and in full-practice. Michigan is a similar case: While compensated surrogacy is illegal, uncompensated one is permitted. This table uses the dates of the previous table and interacts each legislation category. New York has a criminalized surrogacy legislation but at the same time allows uncompensated surrogacies which explains the non zero surrogacies reported for this state by the Centers for Disease Control and Prevention. The term uncompensated means that commercial surrogacy is banned in the state and only altruistic surrogacy is allowed. The term time to change mind indicates a clause in the legislation that allows a period of three to five days during which the surrogate mother can change her mind and decide to keep the baby.

Table 3.5: Variable sources and definitions

variable	source	definition
surrogacy	CDC[a]	the cycle has used a gestational carrier
surrogacy rate	CDC	how many surrogacies happened annually per 100,000 total births in area
surrogacy clinics	CDC	clinics which use gestational carriers in at least one cycle
clinics	CDC	fertility clinics under CDC surveillance
non surrogacy clinics	CDC	clinics which do not use gestational carriers for any cycle
fresh non donor	CDC	gestational carrier cycles (1%) [g]
frozen non donor	CDC	in my sample, non gestational carrier cycles
fresh donor	CDC	in my sample, non gestational carrier cycles
frozen donor	CDC	in my sample, non gestational carrier cycles
criminalized	LexisNexis[b]	surrogacy is forbidden under state legislation (statutory law and cases)
unenforceable	LexisNexis	surrogacy contracts are not enforceable under state legislation (statutory law and cases)
probably unenforceable	LexisNexis	evidence on legislation (statutory law and cases) makes surrogacy contracts probably unenforceable
uncertain	LexisNexis	contradictory legislation (statutory law and cases) makes unclear whether surrogacy is permitted or not
probably enforceable	LexisNexis	evidence on legislation (statutory law and cases) makes surrogacy contracts probably enforceable
enforceable	LexisNexis	surrogacy contracts are enforceable, i.e. surrogacy is permitted under state legislation (statutory law and cases)
negative laws	LexisNexis	group variable (criminalized + unenforceable): laws—state legislation: statutory law and cases—are negative on surrogacy contracts

Continued on next page

variable	source	definition
ambiguous laws	LexisNexis	group variable (uncertain + probably enforceable): laws—state legislation: statutory law and cases—are ambiguous on surrogacy contracts. Probably unenforceable omitted
positive laws	LexisNexis	group variable (enforceable): laws—state legislation: statutory law and cases—are positive on surrogacy contracts
uncompensated	LexisNexis	commercial surrogacy is banned in the state and only altruistic surrogacy is allowed
time to change mind	LexisNexis	a clause in the legislation that allows a period of three to five days during which the surrogate mother can change her mind and decide to keep the baby
marriage rate	CDC	how many marriages happened annually per 1,000 total population residing in area. Population estimated as of July 1st each year[h]
out-of-wedlock	CDC	non marital fertility, births occurred outside of marriage
divorce rate	CDC	how many divorces happened annually per 1,000 total population residing in area. Population estimated as of July 1st each year[i]
births	CDC	number of live births
birthrate	CDC	number of live births per 1,000 of a population in a particular year
(general) fertility rate	CDC	number of births per 1,000 women between the ages of 15 and 44 in a particular year
total fertility rate (TFR)	CDC	number of births that a cohort of 1,000 women would have if they experienced throughout their childbearing years the same age-specific birth rates observed in a given year
teenage birthrate	CDC	number of births per 1,000 women between the ages of 15 and 19 in a particular year
White births	CDC	number of White live births
Black births	CDC	number of Black live births

Continued on next page

variable	source	definition
Hispanic births	CDC	number of Hispanic origin live births
White female percentage	CDC	percentage of female population age 15-39 and White race percentage
Black female percentage	CDC	percentage of female population age 15-39 and Black race percentage
Hispanic female percentage	CDC	percentage of female population age 15-39 and Hispanic origin percentage
personal income (per capita)	BEA[c]	per capita income received by persons [j]
GDP (per capita)	BEA	per capita real GDP by state
women's weekly wage	BLS[d]	women's median usual weekly earnings of full-time wage and salary women workers (in dollars)
men's weekly wage	BLS	men's median usual weekly earnings of full-time wage and salary men workers (in dollars)
total weekly wage	BLS	both sexes median usual weekly earnings of full-time wage and salary workers (in dollars)
women's on men's wage	BLS	women's earnings as percent of men's
women's participation rate	BLS	percentage of women's population in civilian labour force
women's employment ratio	BLS	percentage of women's population employed
women's unemployment rate	BLS	percentage of women's population unemployed
men's participation rate	BLS	percentage of men's population in civilian labour force
men's employment ratio	BLS	percentage of men's population employed
men's unemployment rate	BLS	percentage of men's population unemployed
total participation rate	BLS	percentage of total population in civilian labour force
total employment ratio	BLS	percentage of total population employed
total unemployment rate	BLS	percentage of total population unemployed
party affiliation	Voting America[e]	"colour" of most voted party in last Presidential election (2000, 2004, 2008, 2012) by state
high school attendance	Census[f]	percentage of high school graduates or more for persons 25 years old and over

Continued on next page

variable	source	definition
Bachelor's degree	Census	percentage of Bachelor's degrees or more for persons 25 years old and over
White race	Census	percentage of White race
Black race	Census	percentage of Black race
Native American	Census	percentage of Native American
Asian	Census	percentage of Asian
Native Hawaiian	Census	percentage of Native Hawaiian
two races	Census	percentage of two or more races
Hispanic origin	Census	percentage of Hispanic origin, not a single race

Table Notes

[a] CDC is the Centers for Disease Control and Prevention, www.cdc.gov

[b] LexisNexis is the Academic LexisNexis, www.lexisnexis.com/hottopics/lnacademic/ in conjunction with Creative Family Connections LLC website CreFamCon (2015) and Darra L. Hofman's key article, "Mama's Baby, Daddy's Maybe: A State-by-State Survey of Surrogacy Laws and Their Disparate Gender Impact" Hofman (2009).

[c] BEA is the Bureau of Economic Analysis, www.bea.gov

[d] BLS is Bureau of Labor Statistics, www.bls.gov

[e] Voting America is a project of the University of Richmond (Digital Scholarship Lab) with political party strength in US states maps in the Presidential elections. Retrieved October 1, 2015, from http://dsl.richmond.edu/voting/statelevel.html For 2012 election I use Dave Leip's Atlas of US Presidential Elections. Retrieved October 1, 2015, from http://uselectionatlas.org/RESULTS/national.php

[f] Census is the US Census Bureau, www.census.gov

[g] transfers can be fresh or thawed—frozen—and donor or non donor. In my data, surrogacies are related only to fresh non donor transfers/cycles (only 1% of them uses gestational carriers). Also during the procedure a woman's can use her own eggs—non-donor—or eggs from another woman—donor—and the embryos used can be newly fertilized—fresh—or previously fertilized, frozen, and then thawed—frozen.

[h] Puerto Rico with no data; Oklahoma with partial data; 2013 data from American Community Survey (American FactFinder).

[i] California, Indiana, Puerto Rico with no data; Georgia, Hawaii, Louisiana, Minnesota, Oklahoma with partial data; all the rest with no 2013 data

[j] from participation in production, plus transfer receipts from government and business, plus government interest (which is treated like a transfer receipt). It is defined as the sum of per capita wages and salaries, supplements to wages and salaries, proprietors' income with inventory valuation and capital consumption adjustments, rental income of persons with capital consumption adjustment, personal dividend income, personal interest income, and personal current transfer receipts, less contributions for government social insurance. Per capita personal income is often used as an indicator of consumers' purchasing power and of the economic well-being of the residents of an area.

Table 3.6: Overall variable summary statistics

variable	obs	mean	std. dev.	min	max
surrogacy	676	15.73	32.42	0	295
surrogacy rate	676	20.04	17.72	0	146
surrogacy clinics	675	3.76	6.06	0	48
clinics	676	8.21	11.07	0	68
non surrogacy clinics	675	4.45	5.96	0	32
fresh non donor	638	1,883	2,740	8	15,391
frozen non donor	676	441	721.25	0	6,525
fresh donor	676	182	312.22	0	2,096
frozen donor	676	110	198.51	0	1,908
criminalized	676	0.02	0.16	0	1
unenforceable	676	0.03	0.19	0	1
probably unenforceable	676	0.09	0.29	0	1
uncertain	676	0.44	0.50	0	1
probably enforceable	676	0.17	0.38	0	1
enforceable	676	0.22	0.41	0	1
negative laws	676	0.07	0.25	0	1
ambiguous laws	676	0.62	0.49	0	1
positive laws	676	0.22	0.41	0	1
uncompensated	676	0.25	0.44	0	1
time to change mind	676	0.09	0.29	0	1
marriage rate	659	7.31	2.32	4	19.6[13]
births	676	79,573	94,651	5,975	566,414
out-of-wedlock	674	30,636	36,979	1,813	221,568
divorce rate	548	3.65	0.85	1.7	7.4
birthrate	676	13.59	1.61	9.4	21.8
fertility rate	676	65.73	6.27	47.9	95
total fertility rate	676	2,002	177.76	1,470	2,755
teenage birthrate	676	38.34	12.42	12.1	74.9
White births	676	43,220	36,846	1,175	167,025

Continued on next page

[13]Nevada dropped

variable	obs	mean	std. dev.	min	max
Black births	676	11,428	13,457	29	51,799
Hispanic births	676	19,182	45,017	32	297,092
White female percentage	676	0.55	0.18	0.02	0.96
Black female percentage	676	0.15	0.09	0.001	0.63
Hispanic female percentage	676	0.23	0.18	0.004	0.97
personal income (per capita)	666	38,291	6,781	22,780	75,950
GDP (per capita)	663	47,434	8,739	28,956	71,047
women's weekly wage	663	622.66	89.25	407	1100
men's weekly wage	663	772.96	100.72	522	1212
total weekly wage	667	697.77	94.04	458	1152
women's on men's wage	663	80.54	4.55	61.8	96.5
women's participation rate	663	58.95	3.37	47.7	71.2
women's employment ratio	663	55.26	3.74	44.8	68.9
women's unemployment rate	663	6.3	1.9	2.3	12.7
men's participation rate	663	72.41	3.36	59.6	82
men's employment ratio	663	67.43	4.2	54.5	79.7
men's unemployment rate	663	6.92	2.45	2.4	15.8
total participation rate	663	65.45	3.17	53.8	76.1
total employment ratio	663	61.14	3.75	49.6	73.3
total unemployment rate	672	6.68	2.21	2.6	16
party affiliation	663	1.57	0.49	1	2
high school attendance	665	84.91	3.7	69.5	93.5
Bachelor's degree	665	27.7	4.45	15.1	55.1
White race	664	78.07	9.4	24.7	97.1
Black race	664	12.73	8.1	0.32	60.3
Native American	664	0.94	1.54	0.1	15.6
Asian	664	4.45	4.24	0.26	41.79
Native Hawaiian	664	0.17	0.63	0	10.44
two races	664	2.01	1.6	0.6	24.11
Hispanic origin	664	15.23	12.73	0.72	98.95

Table 3.7: Balancing Test

	predicted surrogacy rate											
	unconditional				after fixed effects				after controls + fixed effects			
	coefficient	std error	t-test	p-value	coefficient	std error	t-test	p-value	coefficient	std error	t-test	p-value
Controls												
Black percentage	4.499	2.297	1.959	(0.051)*	10.991	14.680	0.749	(0.454)	30.771	36.231	0.849	(0.396)
Hispanic percentage	8.379	1.680	4.988	(0.000)***	25.729	11.790	2.182	(0.029)**	0.838	20.825	0.040	(0.968)
GDP (per capita)	0.000	0.000	1.923	(0.055)*	0.000	0.000	-0.481	(0.631)	0.000	0.000	0.553	(0.580)
party affiliation	2.283	0.580	3.940	(0.000)***	-0.921	0.562	-1.639	(0.102)	-1.424	0.574	-0.479	(0.113)
Bachelor's degree	0.409	0.049	8.286	(0.000)***	0.117	0.119	0.980	(0.327)	0.206	0.136	1.513	(0.131)
Black race	0.049	0.026	1.881	(0.060)*	0.183	0.173	1.054	(0.292)	-0.189	0.420	-0.450	(0.653)
Hispanic origin	0.155	0.025	6.141	(0.000)***	0.646	0.211	3.066	(0.002)***	0.656	0.341	1.923	(0.055)*
Unobservables												
births	0.000	0.000	11.788	(0.000)***	0.000	0.000	-1.379	(0.169)	0.000	0.000	-1.150	(0.181)
birth rate	-0.091	0.161	-0.564	(0.573)	-1.056	1.238	-0.853	(0.394)	-2.667	1.958	-1.362	(0.173)
fertility rate	-0.066	0.037	-1.775	(0.076)*	-0.181	0.047	-3.882	(0.000)***	-0.014	0.254	-0.057	(0.955)
total fertility rate	-0.002	0.001	-1.487	(0.138)	-0.005	0.004	-1.257	(0.209)	0.005	0.005	1.133	(0.258)
teenage birth rate	-0.107	0.022	-4.854	(0.000)***	-0.079	0.039	-2.041	(0.042)**	-0.012	0.063	-0.192	(0.848)
White births	0.000	0.000	9.001	(0.000)***	0.000	0.000	-1.450	(0.147)	0.000	0.000	1.360	(0.174)
Black births	0.000	0.000	8.842	(0.000)***	0.000	0.000	-1.230	(0.219)	0.000	0.000	-0.241	(0.810)
Hispanic births	0.000	0.000	11.936	(0.000)***	0.000	0.000	0.157	(0.875)	0.000	0.000	1.094	(0.274)
White percentage	-6.504	1.452	-4.481	(0.000)***	-18.733	6.857	-2.732	(0.006)***	-15.759	10.288	-1.532	(0.126)
personal income (per capita)	0.000	0.000	7.827	(0.000)***	0.000	0.000	0.173	(0.862)	0.000	0.000	1.007	(0.314)
women's weekly earnings	0.021	0.003	7.515	(0.000)***	0.000	0.005	0.091	(0.927)	0.033	0.030	1.108	(0.268)
women's on men's wage	0.231	0.061	3.766	(0.000)***	-0.032	0.046	-0.703	(0.482)	-0.196	0.231	-0.851	(0.395)
men's weekly earnings	0.017	0.003	6.839	(0.000)***	0.002	0.004	0.480	(0.631)	-0.005	0.024	-0.207	(0.836)
total weekly earnings	0.020	0.003	7.437	(0.000)***	-0.002	0.005	-0.389	(0.697)	-0.036	0.116	-0.310	(0.756)
women's participation rate	-0.140	0.067	-2.093	(0.037)**	-0.066	0.101	-0.652	(0.515)	-0.190	2.033	-0.094	(0.925)
women's employment ratio	-0.135	0.062	-2.185	(0.029)**	-0.045	0.098	-0.459	(0.646)	0.555	2.199	0.252	(0.801)
women's unemployment rate	0.218	0.157	1.392	(0.164)	-0.057	0.147	-0.388	(0.698)	-1.225	1.815	-0.675	(0.500)
men's participation rate	0.156	0.073	2.145	(0.032)**	0.200	0.097	2.050	(0.041)**	0.806	1.772	0.455	(0.650)
men's employment ratio	0.097	0.051	1.902	(0.057)*	0.069	0.082	0.844	(0.399)	0.108	1.783	0.061	(0.952)
men's unemployment rate	0.022	0.124	0.182	(0.856)	0.172	0.119	1.440	(0.150)	-0.673	1.777	-0.379	(0.705)
total participation rate	-0.053	0.032	-1.656	(0.097)*	0.086	0.113	0.759	(0.448)	0.596	2.629	0.227	(0.821)
total employment ratio	-0.025	0.064	-0.387	(0.699)	0.022	0.100	0.221	(0.825)	-1.754	2.562	-0.684	(0.494)
total unemployment rate	0.040	0.012	3.333	(0.000)***	0.107	0.139	0.772	(0.440)	1.048	3.018	0.347	(0.729)
high school attendance	-0.153	0.075	-2.053	(0.040)**	-0.270	0.138	-1.957	(0.051)	-0.283	0.156	-1.816	(0.070)*
White race	-0.027	0.021	-1.240	(0.216)	-0.086	0.069	-1.253	(0.211)	0.095	0.115	0.828	(0.408)
Native American	-0.595	0.098	-6.093	(0.000)***	0.746	0.923	0.807	(0.420)	1.856	1.204	1.542	(0.124)
Asian	0.122	0.052	2.334	(0.020)**	0.656	0.375	1.749	(0.081)*	-0.175	0.524	-0.335	(0.738)
Native Hawaiian	-0.405	0.221	-1.829	(0.068)*	0.577	1.748	0.330	(0.742)	2.960	2.173	1.362	(0.174)
two races	-0.223	0.102	-2.177	(0.030)**	-0.153	0.284	-0.541	(0.589)	-0.148	0.460	-0.323	(0.747)

Note (1): Source: Author's calculations.
Note (2): Table reports coefficients, standard errors, t-test and p-values from regressions of predicted surrogacy rate on fertility, labour market, political process and demographics characteristics assorted as being controls or observed variables drawn from the unobservable in the model. The dependent variable is always the predicted surrogacy rate from a regression of surrogacy rate on the full-set of instruments: negative, ambiguous and positive laws. The first four columns do not include any additional control or fixed effect. The next four columns include state and year fixed effects. The last four columns include controls and state and year fixed effects. Regressions are OLS.
Note (3): *** p<0.01, ** p<0.05, * p<0.1

Table 3.8: Effect of surrogacy rates on marriage rates

	dependent variable marriage rate			
	OLS		IV	
causal variable	(1)	(2)	(3)	(4)
surrogacy rate	0.005**	0.000	0.019***	0.007**
	(0.002)	(0.002)	(0.003)	(0.003)
regressions include				
controls	No	Yes	No	Yes
state fe	Yes	Yes	Yes	Yes
year fe	Yes	Yes	Yes	Yes
instruments				
First stage coeff				
negative laws			-1.287	-2.521**
ambiguous laws			0.849	0.008
positive laws			1.226**	0.876
First stage F t.(F-s r.)			316.49***	215.01***
N	659	659	659	659

[1] **Source:** Author's calculation from the Clinic Data Table of the Centers for Disease Control and Prevention, US Government. Data are available on-line at www.CDC.gov/ART/ARTReports.htm. Controls are available from US Census Bureau, US Bureau of Labor Statistics and US Bureau of Economic Analysis.
[2] Columns (1-2) are OLS regressions and Columns (3-4) are IV regressions of marriage as the dependent variable and surrogacy rate as the causal variable. I use year and state fixed effect throughout the model and add iteratively controls. Selected controls are Black percentage of female population age 15-39, Hispanic origin percentage of female population age 15-39 personal, GDP (per capita), party affiliation, Bachelor's degree, Black percentage and Hispanic origin percentage. In the IV regressions surrogacy rate is instrumented by the legislation dummies: negative, ambiguous and positive laws. I implement in Stata a 2SLS. First stage coefficients and First stage tests are shown. Negative laws = criminalized + unenforceable, ambiguous laws = uncertain + probably enforceable and positive laws = enforceable. Omitted probably unenforceable.
[3] F t.=F test, F-s r.=First-stage relevance.

Table 3.9: Effect of surrogacy rates on births and out-of-wedlock births

	dependent variable births			
	OLS		IV	
causal variable	(1)	(2)	(3)	(4)
surrogacy rate	-3.91	-22.39	-20.66	-87.05**
	(18.36)	(18.35)	(30.48)	(33.53)
	dependent variable out-of-wedlock births			
	OLS		IV	
causal variable	(1)	(2)	(3)	(4)
surrogacy rate	-3.19	-8.39	-39.59	-66.44**
	(14.84)	(14.40)	(24.92)	(26.39)
regressions include				
controls	No	Yes	No	Yes
state fe	Yes	Yes	Yes	Yes
year fe	Yes	Yes	Yes	Yes
instruments				
First stage coeff				
negative laws			-1.284/-1.788**	-1.447*/-1.954**
ambiguous laws			0.853/1.075	0.006/0.223
positive laws			1.229**/1.789**	0.546/0.911
First stage F t.(F-s r.)			318.65***/355.10***	248.13***/246.47***
N	676/676	676/676	676/676	676/676

[1] **Source:** Author's calculation from the Clinic Data Table of the Centers for Disease Control and Prevention, US Government. Data are available on-line at www.CDC.gov/ART/ARTReports.htm. Controls are available from US Census Bureau, US Bureau of Labor Statistics and US Bureau of Economic Analysis.
[2] Columns (1-2) are OLS regressions and Columns (3-4) are IV regressions of births/out-of-wedlock births as the dependent variable and surrogacy rate as the causal variable. I use year and state fixed effect throughout the model and add iteratively controls. Selected controls are Black percentage of female population age 15-39, Hispanic origin percentage of female population age 15-39 personal, GDP (per capita), party affiliation, Bachelor's degree, Black percentage and Hispanic origin percentage. In the IV regressions surrogacy rate is instrumented by the legislation dummies: negative, ambiguous and positive laws. I implement in Stata a 2SLS. First stage coefficients and First stage tests are shown. Negative laws = criminalized + unenforceable, ambiguous laws = uncertain + probably enforceable and positive laws = enforceable. Omitted probably unenforceable.
[3] F t.=F test, F-s r.=First-stage relevance.

Table 3.10: Effect of surrogacy rates on divorce rates

	dependent variable divorce rate			
	OLS		IV	
causal variable	(1)	(2)	(3)	(4)
surrogacy rate	0.001	-0.000	0.005***	0.003**
	(0.001)	(0.001)	(0.002)	(0.002)
regressions include				
controls	No	Yes	No	Yes
state fe	Yes	Yes	Yes	Yes
year fe	Yes	Yes	Yes	Yes
instruments				
First stage coeff				
negative laws			-0.251	-1.324*
ambiguous laws			1.628***	0.879
positive laws			2.304***	1.279*
First stage F t.(Fs r.)			226.05***	140.21***
N	548	548	548	548

[1] **Source:** Author's calculation from the Clinic Data Table of the Centers for Disease Control and Prevention, US Government. Data are available on-line at www.CDC.gov/ART/ARTReports.htm. Controls are available from US Census Bureau, US Bureau of Labor Statistics and US Bureau of Economic Analysis.
[2] Columns (1-2) are OLS regressions and Columns (3-4) are IV regressions of divorce as the dependent variable and surrogacy rate as the causal variable. I use year and state fixed effect throughout the model and add iteratively controls. Selected controls are Black percentage of female population age 15-39, Hispanic origin percentage of female population age 15-39 personal, GDP (per capita), party affiliation, Bachelor's degree, Black percentage and Hispanic origin percentage. In the IV regressions surrogacy rate is instrumented by the legislation dummies: negative, ambiguous and positive laws. I implement in Stata a 2SLS. First stage coefficients and First stage tests are shown. Negative laws = criminalized + unenforceable, ambiguous laws = uncertain + probably enforceable and positive laws = enforceable. Omitted probably unenforceable.
[3] F t.=F test, F-s r.=First-stage relevance.

Table 3.11: Summary effects of surrogacy on selected vital statistics

	per surr. rate	per surrogacy	standardized effect
Δ marriage	+281.1**	6.8**	5.38%**
Δ divorce	+39.9**	0.97**	4.21%**
Δ births	-87.06**	-2.1**	-1.63%**
Δ out-of-wedlock	-66.44**	-1.6**	-3.16%**

[1] **Source:** Author's calculation from the Clinic Data Table of the Centers for Disease Control and Prevention, US Government. Data are available on-line at www.CDC.gov/ART/ARTReports.htm.Controls are available from US Census Bureau, US Bureau of Labor Statistics and US Bureau of Economic Analysis. [2] Magnitudes implied by the coefficients of the IV preferred regressions (Columns (4) of the previous tables) elucidating the causal relationship between surrogacy on selected vital statistics. I calculate how many more or less marriages, divorces, births and out-of-wedlock births imply an additional unit in surrogacy rate (1 additional surrogacy per 100,000 births) and per surrogacy (in my sample there are an average of 4,137,796 births each year: Then an increase in 1 additional point in surrogacy rate adds up to approximately 41.3 surrogacies). I report per 100,000 birth and per surrogacy changes. Additionally, I show the standardized effects as $\frac{\text{s.d.}_{\text{surrogacy\_rate}} * \text{coefficient}_{\text{surrogacy\_rate}}}{\text{s.d.}_{\text{vital\_statistics}}}$ (e.g. $\text{s.d.}_{\text{surrogacy\_rate}} = 17.72$; $\text{coefficient}_{\text{surrogacy\_rate}} = 0.007$; $\text{s.d.}_{\text{marriage\_rate}} = 2.32$), so that it gives the percentage of the variation of the vital statistics (marriage rate, divorce rate, births, out-of-wedlock) explained by the variation on the causal variable (surrogacy rate).

Table 3.12: Effect of surrogacy rates on other outcomes

	dependent variable			
	(1)	(2)	(3)	(4)
	birthrate	fertility rate	total fertility rate	teenage birthrate
causal variable				
surrogacy rate			OLS + year + state fe	
	0.001	-0.016*	-0.205	-0.019*
	(0.002)	(0.009)	(0.250)	(0.012)
			OLS + year + state fe + ctls	
	-0.003**	-0.024***	-0.464**	-0.016
	(0.001)	(0.007)	(0.198)	(0.010)
surrogacy rate			IV + year + state fe	
	0.004	-0.046***	-0.476	-0.05***
	(0.003)	(0.015)	(0.413)	(0.019)
			IV + year + state fe + ctls	
	-0.011***	-0.081***	-1.461***	-0.044**
	(0.002)	(0.013)	(0.365)	(0.018)
N	650	650	650	650

[1] **Source:** Author's calculation from the Clinic Data Table of the Centers for Disease Control and Prevention, US Government. Data are available on-line at www.CDC.gov/ART/ARTReports.htm. Controls are available from US Census Bureau, US Bureau of Labor Statistics and US Bureau of Economic Analysis.
[2] Rows (1-2) are OLS regressions and Rows (3-4) are IV regressions of other outcomes—birthrate, fertility rate, total fertility rate, teenage birthrate—as the dependent variable and surrogacy rate as the causal variable. I use year and state fixed effect throughout the model and add iteratively controls. Selected controls are Black percentage of female population age 15-39, Hispanic origin percentage of female population age 15-39 personal, GDP (per capita), party affiliation, Bachelor's degree, Black percentage and Hispanic origin percentage. In the IV regressions surrogacy rate is instrumented by the legislation dummies: negative, ambiguous and positive laws. I implement in Stata a 2SLS. First stage coefficients and First stage tests are shown. Negative laws = criminalized + unenforceable, ambiguous laws = uncertain + probably enforceable and positive laws = enforceable. Omitted probably unenforceable.

Table 3.13: Robustness checks

	(1) organ donation	(2) lag variables (outcomes)	(3) fertility surrogacy lowly-correlated	(4) fertility surrogacy uncorrelated	(5) population	(6) lagged marriage rates '90-'95-'99-'00
			dependent variable			
endogenous variable			marriage rate			lagged marriage rate
surrogacy rate	0.009**	0.011*				-0.109
	(0.004)	(0.005)				(0.093)
surrogacy clinics			0.011			
			(0.088)			
fresh non donor			0.000			
			(0.000)			
non surr. clinics				0.055		
				(0.114)		
clinics				0.078		
				(0.117)		
fresh donor				0.002		
				(0.002)		
population					-0.012	
					(0.009)	
N	594	624	659	659	659	201
Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1						

Note (1): Source: Author's calculations.
Note (2): In Column (1) I add organ donation to the regressors. In Column (2) I use lags as controls. In the next two columns I replace surrogacy by lowly-surrogacy-correlated variables (surrogacy clinics and fresh non donor, Column (3)) and surrogacy-non-correlated variables (non surrogacy clinics, clinics and fresh donor, Column (4)). In Column (5) instead of surrogacy I try to explain marriage by population. In Column (6) I replace the dependent variable for marriage in 1990, 1995, 1999 and 2000 using as RHV surrogacy and the controls for years 2001 to 2004. In the first two regression I expect a positive and seizable association between surrogacy and the dependent variables. In the other five regressions I do not expect any association nor significance.

Table 3.14: Clinics gestational rate (percentage). Illinois 2001-2013

	2001	2002	2003	2004 Surrogacy Act	2005	2006	2007	2008	2009	2010	2011	2012	2013
clinic 1	< 1	< 1	0	0	0	0	0	1	0	0	1	0	2
clinic 2	0	0	1	94	100	0	0	0	0	0	0	0	< 1
clinic 3	0	0	0	0	0	0	1	1	< 1	0	1	< 1	0
clinic 4	1	< 1	< 1	0	0	< 1	< 1	< 1	< 1	< 1	< 1	< 1	< 1
clinic 5	0	< 1	0	< 1	< 1	< 1	< 1	1	2	1	1	< 1	< 1
clinic 6	0	< 1	0	0	< 1	< 1	0	< 1	< 1	2	< 1	0	0
clinic 7	0	0	0	0	0	4	3	0	0	< 1	0	3	< 1
clinic 8	0	0	0	0	< 1	< 1	0	< 1	0	0	2	0	< 1
clinic 9	0	0	1	< 1	0	0	< 1	< 1	0	< 1	0	0	0
clinic 10	0	0	0	2	0	< 1	0	0	0	0	< 1	0	0
clinic 11	2	0	0	0	0	1	0	0	< 1	0	< 1	0	< 1
clinic 12	< 1	0	< 1	0	< 1	< 1	0	0	2	< 1	0	2	< 1
clinic 13	0	< 1	< 1	< 1	0	< 1	< 1	0	< 1	2	2	0	0
clinic 14	0	0	1	0	< 1	0	0	0	0	< 1	0	0	0
clinic 15	0	0	0	< 1	0	0	0	3	3	0	0	0	0
clinic 16	0	3	0	0	0	2	0	0	0	3	0	2	4
clinic 17	0	0	0	0	0	1	0	0	0	0	0	< 1	0
clinic 18	0	0	0	0	0	0	0	0	0	0	< 1	< 1	0
clinic 19	0	0	< 1	0	0	1	0	0	0	0	1	< 1	< 1
clinic 20	0	3	0	1	0	0	0	< 1	< 1	0	0	0	0
clinic 21	0	0	0	< 1	< 1	< 1	< 1	< 1	2	< 1	0	0	< 1
clinic 22	0	0	0	0	2	2	0	0	0	0	2	0	1
clinic 23	0	0	0	0	< 1	< 1	0	0	1	0	0	< 1	0
				0	0	1	< 1	2	< 1	< 1	0	0	< 1
				< 1	1	0	0	0	0	< 1	0	0	0
				0	2	< 1	0	0	0	0	0		
				0	< 1	0	0	0	0	3			
				0	0	0				0			
					2								
Number of clinics	23	23	23	28	29	28	27	27	27	28	26	25	25

Note(1): Source: Author's calculations. Clinic Table Data 2001-2013 from the Centers for Disease Control and Prevention, US Government. Data are available on-line from www.CDC.gov/ART/ARTReports.htm.

Note(2): The twenty three clinics operating in Illinois in 2001 were Rush-Copley, center For Reproductive Health, Life-Women's Health center, IVF Lincoln Park, Northwestern University, Rush center for Advanced Reproductive Care, University of Chicago Hospitals, University of Illinois at Chicago IVF Program, Watertower Women's center, Midwest Fertility center, The Hoxsey-Rinehart center, The Hoxsey-Rinehart center for Reproductive Medicine, Advanced Fertility center of Chicago, Highland Park IVF center of Chicago, Hinsdale center for Reproduction, center for Human Reproduction—Illinois, Reproductive Health Specialists, IVF1, Reena Jabamoni M.D., Oak Brook Fertility center, Lutheran General Hospital IVF Program, Advanced Reproductive center, Reproductive Health and Fertility center, Reproductive Endocrinology Associates, Seth Levrant.

Note (3): There is attrition therefore some clinics drop from the sample and after the introduction of new permissive legislation new clinics enter the market. For these reasons the initial twenty three clinics are not necessarily the final ones. After a while of initial competition some of the new entrants exit the market.

some clinics report "< 1". Because the national number of surrogacies is known, I add all integer percentage of surrogacy cycles and then split evenly the reminder in the "< 1" cycles.

Table 3.15: State weights in the synthetic Illinois

state	weight	state	weight
Alabama	-	Montana	0
Alaska	0	Nebraska	0
Arizona	0	Nevada	0
Arkansas	0	New Hampshire	0
California	-	New Jersey	0
Colorado	0	New Mexico	-
Connecticut	-	New York	-
Delaware	-	North Carolina	-
District of Columbia	0	North Dakota	-
Florida	0.326	Ohio	-
Georgia	0	Oklahoma	-
Hawaii	0	Oregon	0
Idaho	-	Pennsylvania	-
Illinois	0	Puerto Rico	0
Indiana	0	Rhode Island	-
Iowa	0	South Carolina	-
Kansas	0	South Dakota	0
Kentucky	0	Tennessee	-
Louisiana	0	Texas	-
Maine	0	Utah	-
Maryland	-	Vermont	0
Massachusetts	-	Virginia	0
Michigan	0.674	Washington	-
Minnesota	-	West Virginia	-
Mississippi	0	Wisconsin	-
Missouri	0	Wyoming	0

[1] State weights in the synthetic Illinois. Weigth symbol (-) is for 23 states discarded because they change legislation status during the period 2001-2013 so that they are better excluded from the donor pool.

Table 3.16: Surrogacy predictor means

	Illinois		Average of
Variables	Real	Synthetic	28 control states
surrogacy clinics	6.75	6.771	2.107
overall fertility clinics	24.25	17.553	5.402
non surrogacy clinics	17.5	10.782	3.295
criminalized	0	0	0.036
unenforceable	0	0	0.036
probably unenforceable	0	0	0.143
uncertain	1	0.674	0.5
probably enforceable	0	0	0.107
enforceable	0	0.326	0.179
births	181,989.75	156,935.328	55,260.643
birth rate	14.475	12.895	13.764
fertility rate	67.325	62.194	64.977
White birth percentage	0.547	0.643	0.638
Black birth percentage	0.174	0.187	0.126
Hispanic birth percentage	0.23	0.122	0.149

[1] Surrogacy predictor means: Illinois real and synthetic; average of 28 control states. Legislation dummies: criminalized, unenforceable, probably unenforceable, uncertain, probably enforceable, enforceable.

Table 3.17: Predictor weights in the synthetic Illinois

predictor	weight	predictor	weight
surrogacy clinics	0.078	enforceable	0.253
overall fertility clinics	0.108	births	0
non surrogacy clinics	0.168	birth rate	0.015
criminalized	0.071	fertility rate	0.019
unenforceable	0.072	White birth percentage	0.011
probably unenforceable	0.008	Black birth percercentage	0.076
uncertain	0.037	Hispanic birth percentage	0.014
probably enforceable	0.071		

[1] Predictor weights in the synthetic Illinois. Note that all but births have positive weights. Legislation dummies: criminalized, unenforceable, probably unenforceable, uncertain, probably enforceable, enforceable.

Table 3.18: Surrogacy basic facts at the national level

year	surrogacies	fresh non-donor cycles	%	total fertility cycles
1997	600	55,002	1.09	70,147
1998	809	61,650	1.31	80,634
1999	821	65,751	1.25	86,822
2000	1210	74,957	1.61	99,629
2001	571	80,864	0.71	107,587
2002	548	85,826	0.64	115,392
2003	671	91,032	0.74	122,872
2004	710	94,242	0.75	127,977
2005	929(*)	97,442	1.04	134,260
2006	1042	99,199	1.05	138,198
2007	733	101,897	0.72	142,435
2008	915	104,673	0.87	148,055
2009	883	102,478	0.86	146,244
2010	926	100,824	0.92	147,260
2011	907	101,213	0.90	151,923
2012	901	99,665	0.90	157,662
2013	900(**)	t.b.r.	t.b.r.	t.b.r.

Note(1): Source: Author's calculations. Assisted Reproductive Technology Reports (yearly) 1997-2012 (ART reports). Available on-line from the Centers for Disease Control and Prevention's website. See note (1) in the previous table.
Note(2): A clinic in Hawaii is reported to have undertaken 83 surrogacies in 2005. It seems an outlier. With the Hawaii surrogacies the count for 2005 will be the official 1,012 surrogacies, not the amended 929(*).
Note (3): Notice that data for surrogacy are aggregated at the national level for the years 1997-2012. In contrast, I have made available the data for surrogacy at the state level for 2001-2012. Using these data I extrapolate the country total for 2013 (**) while the other figures are yet to be released (t.b.r.).

Appendix

The big picture

As I have stated in the main part of the chapter before its Surrogacy Data Set there were data on surrogacy only at the *national* level. I will go over them now. I built a reasonably long set of data on surrogacies from fresh non-donor cycles. Table 3.18 and Figure 3.15 summarize the basic facts in the industry. First, I present an exhaustive series of officially released data on surrogacies (note that national data are available from 1997 to 2013). There has been an increase of 50% in surrogacies but this increase has not been steady; it has peaks (i.e. 2000 with 1,210 cycles) and troughs (i.e. 2002 with 548 cycles, the latter roughly half of the former and less than the first point in the series). Thus for this last period, there is negative growth. Then there are the fresh non-donor cycles (only available from 1997 to 2012). Note that the increase in fresh non-donor cycles is a steady trend with a smooth change year by year, 81% increase in the whole period with an annual increase of 6.2%. Next its the turn to the overall fertility cycles (also only available from 1997 to 2012), with a steady increase of 125% in the sixteen years of the sample and an average annual increase of 9.6%. Finally, I present the percentage of surrogacies over fresh non-donor cycles (this figure has appreciable change during the period, with a peak in 2000 and a trough in 2002), but then the rate of these two variables seem to stabilize at approximately 0.9%, meaning that the increase in surrogacies after 2001 in the numerator is matched by an also ascending trend of the increasing fresh non-donor cycles in the denominator. As shown the increase in surrogacies is well behind the rate of increase in fresh non-donor and overall fertility cycles, representing a less steady technology adoption compared with overall ART technology adoption.

Conclusion

In this thesis, I have investigated the spatial aspects of education and family economics. Firstly, I measured the effect of spatially determined competition in the Chilean school voucher system. Recently, the effect of voucher schools has been under great scrutiny, and Chile is perhaps the ideal causal quasi-experiment because it is home to the longest lasting national voucher system in operation in the world. Using a similar approach to Gibbons et al. (2006), I created two spatial indices in order to measure spatially determined competition: a choice index which counts the number of schools accessible from a given municipality, and a competition index which summarizes the choice index for a given community of students. The chapter tests the hypothesis that schools which spatially compete more are also more efficient. Testing whether performance is responsive to market forces is of crucial interest. I use sophisticated econometrics techniques, including two geographical instruments (slope and ruggedness). The results show no effect of spatially determined competition on value-added. Disaggregating the results reveals that private schools do respond to market-oriented incentives, but the voucher system does not. I investigated next the potential reasons behind the underperformance of the voucher system. Schumpeterian (Schumpeter, 1934) competition *for* and *in* the market is related to these outcomes. In addition, non or slow response of parents to "poorly performing" schools and a "too low" voucher are claimed as the proposed causes of this poor functioning of the voucher system. Nonetheless, this effect is the spatial economics equivalent of the literature finding that controlling for peer effects, socio-economic background and non-spatial selection bias results in the difference between private voucher and public voucher schools either disappearing or persisting only for selected groups; see for instance Howell et al. (2001) and, for Chile, Hsieh and Urquiola (2002). I find two groups

that are better off with the spatially determined competition: the first group is pupils with families that have more income, more family education, a pc and internet at home. The second group is pupils in smaller schools with higher achievements in standardized testing and with median dispersion. This last result could be of policy interest, because a system that increases or reduces the dispersion in the composition of schools could imply large improvements or detriments in school educational productivity. According to my results, the best system does not select too much, but also does not include too much.

In the student riot chapter I exploited a police report on occupied schools in the so-called *Chilean Winter* to test the hypothesis that a decrease in attendance has a causal effect in reducing students' performance in standardized tests. The *Chilean Winter* was a social outburst of indignation in which the protagonists were pupils who decided to occupy their schools, stage repeated walkouts or simply protest on the streets. When occupying their schools, the pupils forced cancelled days of schools. Thanks to a published police report, I can identify 205 schools that were occupied during the revolts. My evidence indicates that the performance of pupils affected by missed days of school dropped nearly 5%, which is a huge number in terms of its cost in human capital.

Finally, in the chapter on maternal surrogacy, I addressed the implications of the new fertility technique in which another woman carries and gives birth to a baby for a couple who want—but usually physically cannot have—a child. Surrogacy remains relatively rare but has provoked heated and passionate public debates. This is the first quantitative spatial policy evaluation of surrogacy. Using, until now, elusive data from the Centers for Disease Control and Prevention, I back calculate for the first time the incidence of surrogacy for each US state and for the period 2001-2013. Fertility clinics must report to the Centers for Disease Control and Prevention, and in 2001 a new question on the percentage of use of gestational carriers made the survey the first detailed register of surrogacy. I then studied surrogacy legislation and classified it as criminalized, unenforceable, probably unenforceable, uncertain, probably enforceable and enforceable according to the degree that the state legislation enforces the listing of the commissioning parents in the birth certificate; a key

step for surrogacy to be viable (because otherwise the surrogate mother has all the negotiation power). I explored surrogacy as a causal variable and used changes in legislation as an instrument for surrogacy whose exogeneity I defended both theoretically and with a smooth balancing test. Then I calculated the Local Average Treatment Effect of surrogacy running an IV procedure. I found that one additional standard deviation in surrogacy rate causes an increase of 5.38% standard deviation in marriage and 4.21% standard deviation in divorce and a decrease of -1.63% standard deviation in births and -3.16% standard deviation in out-of-wedlock births. I tested the results with at least 6 robustness checks, including lag outcomes and alternative "hoax" causal variables. I also analysed the case of Illinois where a surrogacy friendly statute was introduced in 2005. I scrutinized the legislation adopted to facilitate surrogacy, and I implemented a synthetic control approach to test the causal link of this legislative innovation and the number of surrogacies. The results are reassuring, and I performed a series of placebo tests to contrast the results, including placebo-in-outcome, -in-region, and -in-time. In conclusion, my results show that surrogacy has real, meaningful and measurable effects in vital statistics which must be carefully considered by policy makers.

To sum up, in this thesis I hope to have exhibited a distinct contribution to the knowledge of education and Family Economics from an Economic Geography perspective. I also hope to have produced original evidence with the discovery of new facts and relations via my exercise of independent critical power, which is the most valuable asset of a researcher in Academia. These results encompass the unresponsiveness of the voucher system to spatially determined competition, the downward effect in performance of students who participated in student riots and the implications of maternal surrogacy on vital statistics. In doing so, I have presented a critical assessment of the relevant literature. This assessment encompassed even the literature on surrogacy, which still holds significant unexplored territory. I have carefully overviewed the methods of research as well, with strong emphasis on causal inference and with an introduction and critical discussion on the stylized facts of my chapters. I have presented my findings and an open discussion of them for each paper. In each case, the main results have been followed by an additional critical discussion of their implications, including external and internal validity and profuse robustness checks. I have also confirmed

the results with new and independent identification strategies to reassess the data from as many view points as possible to corroborate each defensible result.

Bibliography

Abadie, A., and Gardeazabal, J., "The Economic Costs of Conflict". *American Economic Review*, 93(1), 113-132, 2003.

Abadie, A., Diamond, A., and Hainmueller, J., "Synth: An R Package for Synthetic Control Methods in Comparative Case Studies". *Journal of Statistical Software*, 42(13), 1-17, 2011.

AQE, Chilean Ministry of Education's Agency for Quality in Education, "SIMCE Dataset (1999-2013)", Santiago, Chile, 1999-2013.

Angrist, J. D., and Krueger, A. B., "Does Compulsory School Attendance Affect Schooling and Earnings?". *Quarterly Journal of Economics*, 106(4), 979-1014, 1991.

Angrist, J. D., and Pischke, J-S., "Mostly Harmless Econometrics: An Empiricist's Companion", *Princeton University Press*, 2009.

Asahi, K., "The Impact of Better School Accessibility on Student Outcomes", *SERC Discussion Paper*, 156, 2014.

Banerjee, S., "Gestational Surrogacy Contracts: Altruistic or Commercial? A Contract Theoretic Approach", *The Manchester School*, 81(3), 438-460, 2013.

Banerjee, S., and Basu, S., "Rent a Womb: Surrogate Selection, Investment Incentives and Contracting", *SSNR*, 894022, 2006.

Banerjee, S., and Basu, S., "Rent a Womb: Surrogate Selection, Investment Incentives and Contracting", *Journal of Economic Behavior and Organization*, 69(3), 260-273, 2009.

Becker, G. S., "A Theory of Marriage: Part I". *The Journal of Political Economy*, 81(4), 813-846, 1973.

Ben-Shahar, O., and Bar-Gill, O., 'Threatening an Irrational Breach of Contract", *11 Supreme Court Economic Review*, 143, 2003.

Besley, T., and Case, A., "Unnatural Experiments? Estimating the Incidence of Endogenous Policies", *The Economic Journal*, 110(467), F672-F694, 2000.

Bingley, P., Jensen, V. M., and Walker, I., "The Effect of School Class Size on Post-Compulsory Education: Some Cost Benefit Analysis", *UCD Geary Institute Discussion Paper*, 200717, 2007.

Burdett, K., and Coles, M. G., "Long-Term Partnership Formation: Marriage and Employment". *The Economic Journal*, 109(456), 307-334, 1999.

Card, D., "The Impact of the Mariel Boatlift on the Miami Labour Market". *Industrial and Labor Relations Review*, 43(2), 245-257, 1990.

Card, D., and Krueger, A. B., "Minimum Wages and Employment: A Case Study of the Fast-food Industry in New Jersey and Pennsylvania". *American Economic Review*, 84(4), 772-793, 1994.

Card, D., and Lemieux, T., "Draft avoidance and college attendance: The unintended legacy of the Vietnam War". *American Economic Review Papers and Proceedings*, 91(2), 97-102, 2001.

Capron, A. M., and Radin, M. J., "Choosing Family Law over Contract Law as a Paradigm of Surrogate Motherhood", in Gostin, L., "Surrogate Motherhood: Politics and Privacy", *Indiana University Press*, 1990.

Casen, Social Observatory, Ministry of Social Development of the Chilean Government, "Casen Survey (2000-2011)", Santiago, Chile, 2000-2011.

Chesler, P., "Sacred Bond: The Legacy of Baby M", *Virago Press*, London, 1988.

Chumacero, R., Gómez, D., and Paredes, R., "I Would Walk 500 Miles (If it Paid)". *Economics of Education Review*, 30(5), 1103-1114, 2011.

Corea, G., "The Mother Machine: Reproductive Technologies from Artificial Insemination to Artificial Wombs", *The Women's Press*, 1988.

Creative Family Connections, Surrogacy law by state. Retrieved June 29, 2015.

Cullen, J. B., Jacob, B. and Levitt, S. D., "Impact of School Choice on pupil Outcomes: An Analysis of the Chicago Public Schools". *Journal of Public Economics*, 89(5-6), 729-760, 2000.

Diamond, A., Abadie, A., and Hainmueller, J., "Synthetic Control Methods for Comparative Case Studies: Estimating the Effects of California's Tobacco Control Program". *Journal of the American Statistical Association*, 105(490), 493-505, 2010.

Dixit, A. K., and Pindyck, R. S., "Investment under Uncertainty", *Princeton University Press*, New Jersey, 1994.

Edlund, L., and Korn, E., "A Theory of Prostitution". *Journal of Political Economy*, 110(1), 181-214, 2002.

Epstein, R., "The Case for Full Contractual Enforcement". *Virginia Law Review*, 81, 2305-2341, 1995.

Ermisch, J., "An Economic Analysis of the Family", *Princeton University Press*, 2003.

Espenshade, T. J., and Calhoun, C. A., "The Dollars and Cents of Parenthood". *Journal of Policy Analysis and Management*, 5(4), 813-817, 1986.

ESRI, Blog on Z-factor. Retrieved April 15, 2014.

Field, M., "Surrogate Motherhood: The Legal and Human Issues", *Harvard University Press*, 1988.

Field, M., "Surrogate Motherhood", in Eekelaar, J., and Šarčević, P., (Eds.), Parenthood in Modern Society: Legal and Social Issues for the Twenty-First Century, The International Society of Family Law, *Martinus Nijhoff Publishers*, 1993.

Friedlander, M., "A Maternalistic Approach to Surrogacy". *Virginia Law Review*, 81, 2377-2399, 1995.

Gallego, F., "Voucher-School Competition, Incentives, and Outcomes: Evidence from Chile". *PUC Working paper*, 2006.

Gaynor, M., "Competition and Quality in Hospital Markets. What do we know? What don't we know?". *Economie Publique*, 15(2), 3-40, 2004.

Gershoni, N., and Low, C., "The Impact of Extended Fertility on Education and Marriage Outcomes: Evidence from Israel's Expansion of IVF Access", *Working Paper*, 2015.

Gibbons, S., and Silva, O., "Urban Density and Pupil Attainment". *Journal of Urban Economics*, 63(2), 631-650, 2008.

Gibbons, S., Machin, S., and Silva, O., "Choice, Competition, and Pupil Achievement". *Journal of the European Economic Association*, 6(4), 912-947, 2008.

Gitlin, J., "Illinois Becomes Surrogacy Friendly", *American Academy of Marriage Lawyers*, www.aaml.org/sites/default/files/Illinois%20Becomes%20Sur-rogacy%20Friendly.pdf. Retrieved June 29, 2015.

Goodman, J., "Flaking Out: Student Absences and Snow Days as Disruption of Instructional Time", *NBER Working Paper*, 20221, 2014.

Gostin, L., "Surrogate Motherhood: Politics and Privacy", *Indiana University Press*, 1990.

Griswold, Z., "Surrogacy was the Way", *Nightengale Press*, 2006.

Gugucheva, M., "Surrogacy in America", *Council for Responsible Genetics*, www.councilforresponsiblegenetics.org. Retrieved June 29, 2015.

Hanushek, E. A., "Making Schools Work. Improving Performance and Controlling Costs", *The Brookings Institution*, Washington D.C., 1994.

Hanushek, E. A., "Evidence, Politics, and the Class Size Debate" in Mishel, L., and Rothstein, R., (Eds.), The Class Size Debate, *Economic Policy Institute*, 2002.

Hanushek, E. A., "The Failure of Input-Based Schooling Policies", *The Economic Journal*, 113(485), 64-98, 2003.

Hincapié, D., "Do longer school days improve student achievement? Evidence from Colombia", *IDB Working Paper*, 679, 2016.

Hofman, D., "Mama's Baby, Daddy's Maybe: A State-by-State Survey of Surrogacy Laws and Their Disparate Gender Impact". *William Mitchell Law Review*, 35(2), 449-468, 2009.

Hogan, V., and Walker, I., "Education Choice under Uncertainty: Implications for Public Policy", *Labour Economics*, 14, 894-912, 2007.

Howell, W. G., and Peterson, P. E., "The Education Gap: Vouchers and Urban Schoools", *Brookings Institution Press*, Revised Edition, Washington D.C., 2006.

Hoxby, C., "School Choice and School Productivity (Or Could School Choice be a Tide that Lifts All Boats?)", in Hoxby, C., (Ed.), The Economics of School Choice, *The University of Chicago Press*, Chicago, Chapter 8, 287-341, 2003.

Hsieh, C-T., and Urquiola, M., "When Schools Compete, How Do They Compete? An Assessment of Chile's Nationwide School Voucher Program". *National Center for the Study of Privatization in Education*, Occasional Paper No. 43, New York, 2002.

Kane, E., "Birth Mother. The Story of America's First Legal Surrogate Mother", *Harcourt Brace Jovanovich*, San Diego, 1988.

Keane, N. P., and Breo, D. L., "Surrogate Mother: A Revolutionary Option and New Source of Hope for Infertile Couples", *Dodd Mead*, 1981.

Krueger, A. B., "Understanding the magnitude and effect of class size on student achievement" in Mishel, L., and Rothstein, R., (Eds.), The Class Size Debate, *Economic Policy Institute*, 2002.

Kunen, J. S., "Childless Couples Seeking Surrogate Mothers Call Michigan Lawyer Noel Keane—He Delivers", *People Magazine*, March 30, 1987. Retrieved December 23, 2016.

Ladd, H., "Comment on Caroline Hoxby: School Choice and Competition: Evidence from the United States", *Swedish Economic Policy Review*, 10, 67-76, 2003.

Lewis, David (1973), "Causation", *The Journal of Philosophy*, 70(17), 556-567.

Markens, S., "Surrogate Motherhood and the Politics of Reproduction", *University of California Press*, 2007.

Maurin, E., and McNally, S., "Vive la Révolution! Long-Term Educational Returns of 1968 to the Angry pupils". *Journal of Labor Economics*, 26(1), 1-33, 2008.

Mizala, A., and Romaguera, P., "School Performance and Choice: The Chilean Experience". *The Journal of Human Resources*, 35(2), 392-417, 2000.

Morgan, S. L., and Winship, C., "Counterfactuals and Causal Inference. Methods and Principles for Social Research", *Cambridge University Press*, New York, USA. 2010.

Neyman, J., "Sur les Applications de la Theorie des Probabilites aux Experiences Agricoles: Essai des Principes". *Master's Thesis*, 1923. Excerpts reprinted in English, Statistical Science, 5(4), 465-472. Dabrowska, D. M., and Speed, T.P., Translators.

OCDE, Organization for Economic Co-operation and Development, "Education at a Glance 2008", Table B1, 2008.

Pearl, J., "Probabilistic Reasoning in Intelligent Systems. Networks of Plausible Inference". *Morgan Kaufmann Publishers*, San Mateo, California, 1988.

Pearl, J., "Causality. Models, Reasoning, and Inference", *Cambridge University Press*, Second Edition, New York, 2009.

Pelzman, J., ""Womb for Rent": International Service Trade Employing Assisted Reproduction Technologies (ART)", *Review of International Economics*, 21(3), 387-400, 2013.

Pischke, J-S., "The Impact of Length of the School Year on Student Performance and Earnings: Evidence from the German Short School Years", *The Economic Journal*, 117(523), 1216-1242, 2007.

Posner, R., "The Ethics and Economics of Enforcing Contracts of Surrogate Motherhood". *Journal of Contemporary Health Law and Policy*, 5(21), 21-31, 1989.

Propper, C., Burgess, S., and Gossage, D., "Competition and Quality: Evidence from the NHS Internal Market 1991-9". *The Economic Journal*, 118(525), 138-170, 2008.

Ragoné, H., "Surrogate Motherhood: Conception in the Heart", *Westview Press*, 1994.

Raymond, J. G., "Womens as Wombs: Reproductive Technologies and the Battle over Woman's Freedom", *Harper San Francisco*, 1993.

Robertson, J. A., "Procreative Liberty and the State's Burden of Proof in Regulating Non-coital Reproduction", in Gostin, L., "Surrogate Motherhood: Politics and Privacy", *Indiana University Press*, 1990.

Robertson, J. A., "Children of Choice: Freedom and the New Reproductive Technologies", *Princeton University Press*, New Jersey, 1994.

Rothstein, J., "Does Competition Among Public Schools Benefit Students and Taxpayers? A Comment on Hoxby (2000)", *American Economic Review*, 97(5), 2026–2037, 2007.

Rubin, D., "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies". *Journal of Educational Psychology*, 66(5), 688-701, 1974.

Shavell, S., "Specific Performance versus Damages for Breach of Contract: An Economic Analysis", *Texas Law Review*, 84(4), 831-876, 2013.

Schumpeter, J. A., "The Theory of Economic Development", *Transaction Publishers*, US/UK, 1934.

Thomas, J., and Worrall, T., "Self-Enforcing Wage Contracts", *Review of Economic Studies*, LV, 541-554, 1988.

Trivers, R., "Parental Investment and Sexual Selection", in Campbell, B., (Ed.), Sexual Selection and the Descent of Man, *Aldine de Gruyter*, Chicago, 136-179, 1972.

Twine, F., "Outsourcing the Womb. Race, Class, and Gestational Surrogacy in a Global Market", *Routledge*, 2011.

Coordinadora Nacional de Subvenciones (Vouchers) of the Chilean Ministry of Education, "Vouchers", Santiago, Chile, 2010. Retrieved December 8, 2016.

Whitehead, M. B., "A Mother's Story: The Truth About the Baby M Case", *St Martin's Press*, 1989.

Woodward, J., "Making Things Happen: A Theory of Causal Explanation", *Oxford University Press*, New York, 2003.

Zil, C., "The Effects of Compensation on the Supply of Surrogate Mothers", *Senior Essay Seminar*, Vincent Crawford, 2006.