THREE ESSAYS ON APPLIED MICROECONOMICS

A Dissertation

by

MARÍA DEL SOCORRO PADILLA ROMO

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Jason Lindo
Committee Members,	Mark Hoekstra
	Jonathan Meer
	Lori Taylor
Head of Department,	Timothy Gronberg

May 2017

Major Subject: Economics

ABSTRACT

This dissertation introduces three essays on the effects of different public policies on crime, education, and labor outcomes using quasi-experimental research designs. These policies include targeting high-ranked members of criminal organizations to fighting organized crime and extending the school day.

In the first essay “Kingpin Approaches to Fighting Crime and Community Violence: Evidence from Mexico’s Drug War," a joint work with Jason Lindo, we consider the effects of the kingpin strategy, an approach to fighting organized crime in which law-enforcement efforts focus on capturing the leaders of criminal organizations, on community violence in the context of Mexico’s drug war. Newly constructed historical data on drug-trafficking organizations’ areas of operation at the municipality level and monthly homicide data allow us to control for a rich set of fixed effects and to leverage variation in the timing of kingpin captures to estimate their effects. This analysis indicates that kingpin captures cause large and sustained increases to the homicide rate in the municipality of capture and smaller but significant effects on other municipalities where the kingpin’s organization has a presence, supporting the notion that removing kingpins can have destabilizing effects throughout an organization that are accompanied by escalations in violence.

In the second essay “The Short and Long Run Effects of Full-Time Schools on Academic Performance," I study the effect of extending the school day on student achievement in Mexico, where more than 23,000 schools have extended their school day from 4.5 to 8 hours since 2007. I use the variation in the timing with which schools extended their school day to estimate the impact of this intervention on students’ math and reading test scores. I find evidence that extending the school day does not affect student achievement the year of adoption; however it improves math and reading test scores by 5 percent of a

standard deviation one year after adoption and the effect grows over time to 15 percent of a standard deviation four years after adoption. I also find that the effects are more pronounced in schools located in high-poverty communities and for students in lower grade levels.

In the third essay "The Effect of Children's Time in School on Mothers' Labor Supply: Evidence from Mexico's Full-Time Schools Program," a joint work with Francisco Cabrera-Hernández, we examine the effect of the time children spend in school on female labor supply. In particular, we investigate the degree to which extending the school day by three and a half hours, in elementary schools, affects labor force participation, the number of weekly hours worked, and the monthly earnings of females with elementary-school-age children. To do so, we exploit within-individual variation in access to full-time schools and a rotating panel of households that contains individual-level data on labor outcomes and sociodemographic characteristics. Results from long-difference models show that extending the school day increases mothers' labor supply at the extensive and intensive margins, increasing mothers' labor force participation by 7 percentage points and the number of weekly hours worked by 2.4. Moreover, these increases are accompanied by a 47 percent increase in monthly earnings.

A mis papás: María Inés y José Asunción

Con mucho cariño les dedico estos cinco años de trabajo

ACKNOWLEDGMENTS

I am forever grateful to my committee chair, Jason Lindo, for his continuous guidance, support, advice, and encouragement throughout my graduate program. Thank you for the countless hours of meetings and for helping me to become the economist that I am today.

I am grateful to the members of my dissertation committee: Mark Hoekstra, Jonathan Meer, and Lori Taylor for their invaluable research advice. I would also like to thank all the members of the applied microeconomics group for their comments and constructive criticism on this research.

I want to thank my professors from the University of Guadalajara, Willy Cortez and Mauricio Ramírez, for supporting and encouraging me to follow the academic path. Thank you for believing in me and for not letting me give up on my goal. I also want to thank my friend, José María González, for all of his trips to the Ministry of Education in Mexico City to collect part of the data used in this dissertation. Thank you for your help and your pesos.

Thanks also go to my dear friend, Francisco Escobedo, for always being there for me whether listening to my boring economics fun facts, commenting on my research, or keeping me sane during tough times. Thank you for helping me to survive graduate school and get through all the rough patches.

Lastly, I want to thank my beloved family, María Inés, José Asunción, Jorge, Efren, José, Gerardo, Heriberto, and Antonio, for being my pillar of strength and for teaching me that when life is tough I need to be tougher and work hard to achieve my goals. Thank you for your endless love, support, and encouragement throughout my entire life.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Professors Jason Lindo, Mark Hoekstra, and Jonathan Meer of the Department of Economics and Professor Lori Taylor of The Bush School of Government and Public Service.

The analyses depicted in Section 2 were conducted in part by Jason Lindo of the Department of Economics and the analyses depicted in Section 4 were conducted in part by Francisco Cabrera-Hernandez of the Centre for Research and Teaching in Economics.

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate study was partially supported by a fellowship from the Private Enterprise Research Center at Texas A&M University.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Policy interventions are aimed at achieving certain goals, however, they might not produce the expected outcomes or they might cause unintended consequences. When evaluating the success of policy interventions, it is very important to know both their intended and unintended consequences in order to be able to determine their cost-effectiveness. To this end, in this dissertation, I study the causal effects of two policy interventions on crime, education, and labor outcomes using quasi-experimental research designs. These policies include targeting high-ranked members of criminal organizations and extending the school day in developing countries.

Section 2 studies the causal effects of targeting high-ranked members of criminal organizations, also known as the kingpin strategy, on community violence in the context of the war on drugs in Mexico. We focus on homicides, in particular, which have been shown to have far reaching consequences for Mexican communities. For example, recent papers have documented that this form of violence has deleterious effects on economic conditions (Velasquez 2015; Montoya 2016), human capital accumulation (Brown and Velasquez 2016), and infant health (Brown 2016). Proponents of the kingpin strategy argue that removing a leader weakens an organization through its effect on its connections, its reputation, and by creating disarray in the ranks below, and that this may in turn reduce the organization's level of criminal activity. Detractors, however, point out that this strategy may increase violence as lower ranked members maneuver to succeed the eliminated leader and rival groups attempt to exploit the weakened state of the organization. Given sound logic underlying both types of arguments, there is a clear need for empirical research on the subject. We find that the capture of a drug-trafficking-organization (DTO) leader in a municipality increases its homicide rate by 61 percent in the six months following the capture and that this effect is highly persistent into subsequent periods. Consistent with the notion that the kingpin strategy causes widespread destabilization throughout an organization, we also find significant effects (of the same sign but smaller in magnitude) on other municipalities where a captured leader's DTO has a presence. Moreover, we find evidence of spatial displacement as captures appear to reduce the homicide rate for municipalities that neighbor a municipality of capture but where the captured leader's DTO does not have a presence.

Section 3 analyzes the causal effects of extending the school day, in primary education, on academic performance. Policies extending the school day have become more and more popular during the last two decades. The motivation behind these policies is that more time in the classroom should translate into better learning outcomes. However, there is an ongoing debate in the education literature over whether lengthening the school day in fact improves academic performance. On the one hand, those who expect large positive effects argue that longer school days provide students with more time for learning, a deeper coverage of the academic curricula, and more time for any given task. On the other hand, those expecting small or even negative effects argue that students tend to waste time, decrease effort, get exhausted, and have less time for informal learning and other recreational activities (Patall et. al., 2010). Then, it is very important to document both the sign and the magnitude of this type of intervention. The main results indicate that extending the school day does not affect test scores the year of adoption; however, it increases students' math and reading test scores by 14.7 percent and 15.6 percent of a standard deviation, respectively, four year following adoption. These results also suggest that the first year of implementation is crucial for teachers and students to learn how to utilize the additional time more efficiently.

Section 4 considers the causal effects of extending the school day on mothers' labor supply. Women's labor supply in developing countries heavily depends on their fertility decisions, and specifically, on how much they are time- and budget- constrained due to childrearing and the alternative costs of childcare institutions. Thus, the availability and affordability of childcare centers are important for women to increase their labor market participation while their children are still growing up. Our main results indicate that longer school-days increase mothers' labor supply at the extensive and intensive margins, increasing mothers' labor force participation by 7 percentage points and the number of worked hours per week by 2.4. Moreover, these increases in labor supply are accompanied by a 47 percent average increase in earnings. Policies extending the school day have the potential not only to improve children's welfare and school outcomes, but also mothers' labor supply and the available income at home, improving overall welfare, especially, for the most vulnerable sectors of the population.

2. KINGPIN APPROACHES TO FIGHTING CRIME AND COMMUNITY VIOLENCE: EVIDENCE FROM MEXICO'S DRUG WAR

2.1 Introduction

The two main reasons for waging war on drugs are to reduce societal costs associated with drug abuse and to reduce societal costs associated with the drug trade. The former includes effects on health, productivity, violent behavior, and broader impacts on health care and public assistance programs. The latter includes violence involved with the enforcement of contracts and turf battles, corruption, and activity in related "industries" that are detrimental to welfare including protection rackets, human smuggling, kidnapping, prostitution, weapons trafficking, theft, etc.[1] Naturally, the relative importance of these costs depends on many factors, including the types of drugs involved, the level and spatial distribution of demand, and the organization of the supply network.[2] Correspondingly, there is significant heterogeneity in the approaches that have been used to wage war on drugs. Demand-side approaches take the form of prevention efforts, treatment for abusers, and increases in the cost of abuse through enforcement efforts and punishment. Supply-side approaches, on the other hand, focus on disrupting operations by way of confiscation of drugs and guns, targeting precursors, and arresting and punishing those involved in the drug trade. Given resource constraints and the potential for unintended consequences, policy-makers have to consider which of these policies to use and how intensely to use them, highlighting the importance of understanding their costs and benefits. Towards this end, this paper considers the effects of a particular supply-side approach that has played a

[1] See Miron and Zwiebel (1995), Miron (1999), and Owens (2014) for in-depth discussions of the manner in which black markets can promote violence. Miron (1999) and Owens (2014) present empirical evidence of such effects in analyses of homicides caused by prohibition in the United States.

[2] For example, the societal costs associated with the drug trade are most important in areas heavily involved in the illegal production and distribution of drugs to be consumed elsewhere.

prominent role in Mexico's drug war—the targeting of high-ranked members of criminal organizations, also known as the "kingpin strategy"—on community violence. We focus on homicides, in particular, which have been shown to have far reaching consequences for Mexican communities. For example, recent papers have documented that this form of violence has deleterious effects on economic conditions (Velasquez 2015; Montoya 2016), human capital accumulation (Brown and Velasquez 2016), and infant health (Brown 2016).

To put this study into context, it is important to note that most of the existing research in this area focuses on the effects of drug-related interventions on drug abuse in "downstream markets." For example, researchers have shown that the Taliban stamping out poppy production reduced heroin use in Australia (Weatherburn et *al.* 2003), that the effect of Plan Colombia on the supply of Cocaine to the United States was relatively small (Mejía and Restrepo 2013), that reductions in methamphetamine availability in the United States in the mid-1990s reduced drug-related harms (Cunningham and Liu 2003; Dobkin and Nicosia 2009; Cunningham and Finlay 2013), that U.S. state laws limiting the availability of Pseudoephedrine have not changed methamphetamine consumption (Dobkin, Nicosia, and Weinberg 2013) nor have graphic advertising campaigns (Anderson 2010; Anderson and Elsea 2015), and that substance-abuse treatment availability reduces mortality and violent crime (Swensen 2015; Bondurant, Lindo, and Swensen 2016). Less is known about the causal effects of "upstream interventions" on "upstream communities," i.e., the effects of interventions on outcomes in areas where production, distribution, and their associated costs are most relevant. In work closely related to our study, Dell (2015) shows that drug-trade crackdowns in Mexico driven by close PAN mayoral victories increase the number of drug-trade-related homicides. Consistent with prior studies highlighting how drug-related interventions can and have shifted the spatial distribution of the drug trade in Afghanistan (Clemens 2008, 2013a, 2013b), Dell demonstrates that crackdowns increase homicides in the municipalities where the efforts take place and that they also increase homicides in other municipalities to which trafficking is likely to be diverted.[3]

This paper contributes to this literature by focusing explicitly on the effects of the kingpin strategy, which has featured prominently in Mexico's war on drugs and is one of the hypothesized mechanisms underlying Dell's results. Proponents of the kingpin strategy argue that removing a leader weakens an organization through its effect on its connections, its reputation, and by creating disarray in the lower ranks, and that this may in turn reduce the organization's level of criminal activity. Detractors, however, point out that this strategy may increase violence as lower ranked members maneuver to succeed the eliminated leader and rival groups attempt to exploit the weakened state of the organization. Given sound logic underlying arguments in favor of and against the kingpin strategy, there is a clear need for empirical research on the subject. That said, there are two main empirical challenges to estimating the effect of the kingpin strategy that are difficult to overcome. First, policies targeting organized crime are almost always multifaceted, involving the simultaneous use of various strategies. Mexico's war on drugs is no exception—it also involved various approaches implemented at various times with varying degrees of intensity, which we discuss in greater detail in the next section. The second main challenge is that the capture of a kingpin is fairly rare because, by definition, they are small in number. As a result, establishing compelling evidence on the effect of eliminating kingpins in some sense requires a series of case studies.

This study attempts to overcome these challenges by exploiting variation in the timing with which different Mexican DTOs first had their leaders captured during Mexico's drug war and by using a newly constructed data set on the geographic distribution of DTOs

[3]In related work, Mejía and Restrepo (2013) estimate the causal effect of the drug trade on violence using variation in the prominence of the drug-trade in Colombian municipalities based on land suitability for coca cultivation. Also, Angrist and Kugler (2008) show that exogenous shocks to coca prices increase violence in rural Colombian districts as groups fight over additional rents.

over time. We focus on municipalities where these major captures occurred, neighboring municipalities where the captured kingpin's DTO had a presence, non-neighboring municipalities where the captured kingpin's DTO had a presence, and neighboring municipalities where the captured kingpin's DTO did not have a presence. Municipalities without any DTO presence serve as a comparison group. This approach allows us to abstract away from the effects of broader policies and shocks (at the national and/or state level) and to conduct several ancillary analyses to guide our interpretation of the results.

We find that the capture of a drug-trafficking-organization (DTO) leader in a municipality increases its homicide rate by 61% in the six months following the capture and that this effect is highly persistent into subsequent periods. Consistent with the notion that the kingpin strategy causes widespread destabilization throughout an organization, we also find significant effects (of the same sign but smaller in magnitude) on other municipalities where a captured leader's DTO has a presence. Moreover, we find evidence of spatial displacement as captures appear to reduce the homicide rate for municipalities that neighbor a municipality of capture but where the captured leader's DTO does not have a presence. These estimates can explain 31.8 percent of the increase in homicides in Mexico between 2006 and 2010.

Several pieces of evidence support a causal interpretation of these main results. First, homicide rates in the municipalities of interest and in the comparison group track one another closely prior to captures. That this is the case despite the fact that the war on drugs began well before any of the captures we consider suggests that the empirical strategy can separately identify the effects of kingpin captures in the broader context of the war on drugs. We also show that the main results are driven by effects on the individuals most likely to be directly involved in the drug trade: males and, more specifically, working-age males. In an additional effort to show that the main results are not simply reflecting an increase in propensities to engage in violence that coincides with captures in the relevant municipalities, we demonstrate that domestic violence and infant mortality do not respond to these events in any systematic way that could explain the effects on homicides. Lastly, we present evidence that operations themselves do not increase homicides in an analysis of the first major operations of the war on drugs.

The most closely related study to our paper is Calderón et al. (2015). Though that paper also considers the effects of kingpin captures during Mexico's war on drugs on homicides, our paper differs in several critical ways. First, whereas we demonstrate that we have identified a good comparison group for the municipalities we define as being "affected by kingpin captures," municipalities they define as being affected by captures of leaders exhibit a different trend from the synthetic control they use for comparison.[4] Second, their empirical strategy analyzes *all* kingpin captures whereas we analyze the first kingpin captures for each DTO. This distinction is important because the first kingpin captures are plausibly exogenous, as we demonstrate in our empirical analysis, whereas subsequent kingpin captures are not because it is likely that the initial captures contribute to future captures. Stevens (1997) has demonstrated the importance of this sort of consideration for analyses of job displacements where approaches that evaluate the effects of all such events drastically understate the true effects on workers earnings. Third, our use of detailed data on the geographic distribution of DTOs over time allows us to analyze spillover effects of kingpin captures across the DTO, whereas they focus on spillover effects on neighboring municipalities. The importance of this difference is underscored by our finding that the effects on more-distant municipalities account for 30 percent of the effect on homicides. Fourth, we estimate how many homicides in total were caused by the regime change involving the capture of kingpins, and the degree to which it explains the dramatic increase in homicides in Mexico since 2006. Fifth, we present evidence of the (non-)effects of military operations which, without our paper, would stand out as a

[4] See their Figure 3.

major confounder for all other studies using the Mexican case study to try to understand the effects of the kingpin strategy.

The remainder of the paper is organized as follows. In the next section, we provide background on Mexico's drug war, including a discussion of the events that precipitated it, and the relevant DTOs. We then discuss our data and empirical strategy in sections 2.3 and 2.4, respectively. Section 2.5 presents a graphical analysis, the main results, and supporting analyses. Lastly, Section 2.6 discusses the results and concludes.

2.2 Background

2.2.1 Drug-trafficking in Mexico

In many ways, Mexico is ideally situated for producing and trafficking drugs. In addition to having a climate that allows for the growth of a diverse set of drugs, it shares its Northern border with the world's biggest consumer of drugs, the United States (CIA 2010).[5] Drug trafficking has also been able to thrive in Mexico as a result of corruption and weak enforcement of the law. The first DTOs were protected by the government, which designated the areas in which each DTO would carry out its illegal activities. In the 1980s, former police officer Miguel Ángel Félix Gallardo—together with Rafael Caro Quintero and Ernesto Fonseca Carrillo—founded the first Mexican Cartel: the Guadalajara Cartel.[6] After the incarceration of his partners in 1985, Félix Gallardo kept a low profile and decided to divide up the areas in which he operated.[7] According to Grayson (2013), the government and the DTOs had unwritten agreements that "DTO leaders respected the territories of competitors and had to obtain *crossing rights* before traversing their turfs...criminal

[5] The United States is listed in The Wold Factbook as the world's largest consumer of cocaine, Colombian heroin, and Mexican heroin and marijuana.

[6] In addition with his connections with the Mexican government, Félix Gallardo was the first Mexican drug trafficker to make connections with Colombian cartels, particularly he established a solid relation with Pablo Escobar (leader of the Medellín Cartel).

[7] Joaquín Guzmán Loera and Ismael Zambada García were given the pacific coast area, the Arellano Félix brothers received the Tijuana corridor, the Carrillo Fuentes family got the Ciudad Juárez corridor, and Juan García Abrejo received the Matamoros corridor.

organization[s] did not sell drugs in Mexico, least of all to children...and prosecutors and judges would turn a blind eye to cooperative criminals."

In the 1990s, however, the environment became less stable as Guadalajara's DTO splintered into four separate DTOs[8] and the Institutional Revolutionary Party (PRI) lost political power (Astorga and Shirk 2010). Morales (2011) describes the late 1990s and early 2000s as a period in which the DTOs became more independent, going from a regimen of political subordination to one of direct confrontation to dispute the control of territory. In late 2005, a new DTO—La Familia—was established in the state of Michoacán followed by a wave of violence.[9] At the beginning of the war on drugs there were five DTOs (or alliances of DTOs), Sinaloa/Beltrán-Leyva, Gulf, Tijuana, La Familia, and Juárez.

2.2.2 The War on Drugs

As shown in Panel A of Figure A.1, the homicide rate in Michoacán grew dramatically between 2005 and 2006. That said, the national homicide rate continued to be extremely stable at 0.8 per 100,000 residents per month (Figure A.1, Panel B). Nonetheless, eleven days after the beginning of his term, the newly elected President Felipe Calderón declared war on the DTOs on December 11, 2006, citing the increase in violence in Michoacán as intolerable. While intellectuals highlighted his desire to have a significant reform associated with his presidency and the fact that he was born and raised in Michoacán, his stated reasons for initiating the war was a concern "about the growth of drugs-related violence and the existence of criminal groups trying to take over control of entire regions."[10] Calderón's strategy mainly consisted in a frontal attack led by members of the army, the navy, and the federal police seeking the eradication of crops, the confiscation of drugs and

[8] After the arrest of Félix Gallardo in 1989 and his transfer to a the maximum security prison La Palma in Mexico state, the leaders of the designated areas became independent and founded the second generation of cartels (Sinaloa, Tijuana, Juárez, and Gulf).

[9] La Familia DTO is the metamorphosis of La Empresa which was a former branch of the Gulf Cartel.

[10] Financial Times interview, conducted January 17, 2007.

guns, and the incarceration or killing of high ranked drug traffickers (the kingpin strategy). The first operation took place in Michoacán on December 11, 2006, where more than 5000 army and federal police elements were deployed, and subsequent operations followed in other parts of the country.

Mexico's war on drugs was initially viewed as a great success. As shown in Figure A.2, plotting data from 2001 to 2010, the national homicide rate dropped sharply in January 2007. The homicide rate jumped back up to 0.72 in March—not quite to its earlier level—and then held steady for the following 9 months. Then, at the beginning of 2008 in a clear break from trend, the homicide rate started to climb. It would continue to climb for several years, reaching a level 150% higher than the pre-drug-war rate at the end of 2010.

This dramatic increase in violence in Mexico has drawn the attention of researchers from different disciplines trying to explain its causes—most attribute this increase in violence to Calderón's war on drugs. Different researchers have focused on the role of the deployment of federal troops all across the country (Escalante 2011, Merino 2011), the expiration of the U. S. Federal Assault Weapons Ban in 2004 (Chicoine 2011, Dube et al.. 2012), the increase of cocaine seizures in Colombia (Castillo et al. 2012, Mejía and Restrepo 2013), and the increased effort to enforce law initiated by the National Action Party (PAN) mayors (Dell 2015).

Our research is motivated by the observation that the escalation of violence began in January 2008, which was the month in which the first cartel leader was captured during the war on drugs (Alfredo Beltrán Leyva). Naturally, many other things were going on in Mexico and around the world at the same time, necessitating a more rigorous consideration to be able to draw any strong conclusions about the effects of Mexico's kingpin strategy. In order to conduct such an analysis, we make use of newly constructed data on the geographic distribution of DTOs over time—in conjunction with several other data sets—to consider the first captures of kingpins associated with each of the five DTOs in operation at the beginning of the war on drugs. These data and the associated identification strategy are described in the next sections.

2.3 Data

Our analysis brings together data from several sources that ultimately yields a data set at the municipality-month level, spanning January 2001 through December 2010.[11] Our primary outcome variable is based on monthly homicides at the municipality level, constructed using the universe of death certificates from the vital statistics of the National Institute of Statistics and Geography (INEGI).[12] In order to put these data into per capita rates, we use estimated municipality population counts from the National Council of Population (CONAPO) and El Colegio de México (COLMEX), which are based on projections from the Census of Population and Housing. While we note that drug-related homicides are available from December 2006 to October 2011, we do not use these data out of concern for the endogeneity of homicides being classified as "drug related" or "not drug related."[13]

Our information on kingpin captures are from a compendium of press releases of the Army (SEDENA), the Navy (SEMAR), and the Office of the Attorney General (PGR). While these press releases contain a wealth of additional information, we focus on the timing of the first capture of a leader or lieutenant from each of the DTOs during the war on drugs. To put into perspective the types of kingpins we are considering, as the name

[11] San Ignacio Cerro Gordo and Tulum, which were created during this timespan, are not included in our analysis.

[12] Less than one percent of death certificates with homicide as the presumed cause of death are missing the municipality of occurrence. These observations are not used in our analysis.

[13] In particular, we might be especially concerned that events related to the war on drugs would heighten attention to drug-related violence and thus increase the propensity for homicides to be classified as drug related. Alternatively, a desire to influence the public perception regarding the success of the war on drugs could cause a reduction in the probability that homicides are classified as drug related. As we are interested in violence and not in the way that violence is classified, we believe it prudent to use an approach that abstracts away from such issues though we acknowledge that similar biases could arise if events related to the war on drugs affect the probability that deaths are classified as homicides as opposed to being due to other causes.

implies, leaders are at the very highest level of the DTO. Lieutenants are immediately below leaders in the DTO organization. As a practical matter, we classify an event as a capture of a DTO leader when a press release indicates that the individual was a head (or one of the heads) of a DTO and identify an event as a capture of a DTO lieutenant when a press release indicates that the individual was a leader of a DTO in some state or region. While these press releases also allow us to identify the capture of lower-level kingpins, such as plaza bosses who control a single municipality, our analysis of such captures (not shown) suggested that they are not as convincingly exogenous as the first captures of higher-level kingpins. As such, we do not consider such events and our estimates can thus be interpreted as identifying the effects of high-level kingpin captures.

As shown in Table A.1, there is significant variation in the timing with which high-level kingpins were captured for the five DTOs in operation at the beginning of the war on drugs. The first took place on August 29th, 2007—eight months after the war on drugs began—when Juan Carlos de la Cruz Reyna, a lieutenant in the Gulf DTO was captured. The other four DTOs (Sinaloa-Beltrán-Leyva, Tijuana, Juárez, and La Familia) first had top level leaders captured during the war on drugs at various times between January and December of 2008.[14] Juan Carlos de la Cruz Reyna was considered a main link between the Gulf DTO and to Colombian DTOs; he was responsible for receiving shipments of drugs in Tampico and Northern Veracruz and transferring them to the border areas of Matamoros and Nuevo Laredo, from where they were smuggled into the United States. Alfredo Beltrán Leyva (captured January 21, 2008) was considered one of the main leaders of the Beltrán-Leyva DTO; he directed operations in the states of Sinaloa, Sonora, Chihuahua, Durango, Jalisco, and Nayarit, and was in charge of the two assassin groups known as "Los Pelones" and "Los Güeros." Pedro Sanchez Arras (captured May 13, 2008) of the Juárez DTO was considered one of the top lieutenants in the organization and directed operations

[14]Sinaloa and Beltrán-Leyva DTOs were allied before the drug war commenced.

in Chihuahua. Eduardo Arellano Félix (captured October 25, 2008) led the Tijuana DTO with his nephew, Luis Fernando Sanchez Arellano. Alberto Espinoza Barrón (captured December 29, 2008) was a lieutenant in the La Familia DTO; he coordinated the receipt of drugs from South America at the Port of Lazaro Cardenas and the subsequent trafficking to the United States.[15]

We use newly available historical data on the municipalities of operation for each DTO, the construction of which is described in detail in Coscia and Rios (2012). Briefly, the data was constructed using a MOGO (Making Order using Google as an Oracle) framework for selecting the most reliable subset of web information to collect information on relationships between sets of entities (DTOs and municipalities in this case). It uses indexed web content (e.g., online newspapers and blogs) and various queries to identify DTOs' areas of operation at the municipality level between 1990 and 2010.[16] To avoid concerns about endogeneity, we define areas of operation using only data before the war on drugs began (2004–2006).[17] Moreover, we take a conservative approach and specify that a DTO had a presence in a municipality if the municipality was an "area of operation" for the DTO in any of these three years. Figure A.3 maps out the distribution of the DTOs based on this definition. One important takeaway from this figure—which we exploit in our empirical analysis—is that a large share of Mexico has no DTO presence (or a DTO presence that is too weak or inactive to be picked up using Coscia and Rios' approach). Table A.1 reports the number of municipalities that are associated with each DTO presence and the fraction of the total Mexican population residing in these municipalities. These measures of influence are consistent with the notion that "the Gulf Cartel was considered the most powerful

[15] Details on Juan Carlos de la Cruz Reyna are from government press releases and from a newspaper report in La Jornada on August 30, 2007. Details on all other captured leaders rely solely on government press releases.

[16] Such data was previously only available to the research community at the state level.

[17] Although Coscia and Rios (2012) report areas of operation for years prior to 2004, they note that this information is less reliable.

drug-trafficking organization in Mexico" at the beginning of the war on drugs (Stewart and Posey 2009).

2.4 Empirical Strategy

While we begin our analysis of the effects of the kingpin captures homicides with a series of graphical comparisons, our main results are based on a generalized difference-in-differences approach that allows the effects of captures to depend on (1) the association between the municipality and the capture and (2) the amount of time that has elapsed since the capture. Specifically, we estimate a model that considers the effects of captures on four types of municipalities: municipalities where the capture takes place, neighboring municipalities where the captured kingpin's DTO had a presence, non-neighboring municipalities where the captured kingpin's DTO had a presence, and neighboring municipalities where the captured kingpin's DTO did not have a presence. To consider the degree to which the short, medium, and long-run effects may differ, we allow these effects to vary from 0-5 months after a capture, 6-11 months after a capture, and 12+ months after a capture. More formally, the regression model is:

$$\ln H_{mt} = \sum_{type=1}^{4} (\text{Tr05}_{mt}^{type}\delta_{05}^{type} + \text{Tr611}_{mt}^{type}\delta_{611}^{type} + \text{Tr12}_{mt}^{type}\delta_{12}^{type}) + \alpha_m + \gamma_t + X_{mt}\beta + u_{mt} \quad (2.1)$$

where $\ln H_{mt}$ is the natural log of the homicide rate in municipality m at time t; the indicator variables and parameters in the sum capture the four different types of treatment and their effects over time; α_m are municipality fixed effects; γ_t are month-by-year fixed effects; X_{mt} can include time-varying municipality controls; and u_{mt} is an error term.[18] As

[18]We add one to the homicide count to avoid missing values. We acknowledge that a count data model, such as the fixed effects Poisson model, would be a natural alternative that would not require this ad hoc approach to avoiding missing values. However, we have not been able to get such models to converge due to the size of our data set. Another alternative would be to consider the homicide rate as the outcome instead of the log of the homicide rate. This approach, however, would impose the assumption that treatment effects on the homicide rate is the same magnitude for municipalities with high and low pre-treatment homicide rates and can thus lead to negative predicted levels. In contrast, focusing on the log of the homicide rate imposes the more natural assumption that treatment effects are proportional to pre-treatment homicide rates.

such, the estimated effects (δ) are identified by comparing changes in violence over time among municipalities that have been "treated" (according to our definition) to the changes observed over time in other municipalities, where the latter are comprised of municipalities that are not linked to any DTO and those that are only linked to DTOs that have yet to have had a captured kingpin.

This approach allows us to avoid biases that would otherwise be introduced by fixed differences across municipalities and by the effects of any shocks or interventions that are common across municipalities. The fact that we have municipalities associated with different DTOs who have kingpins captured at different times and we also have municipalities without any DTO presence allows us to additionally control for the effects of the war on drugs that are common to municipalities with a DTO presence, which we accomplish by including indicator variables for 0–5, 6–11, and 12+ months after the beginning of the war interacted with an indicator for the presence of a DTO in the municipality. We can also control for additional spatial heterogeneity by including state-by-year-by-month fixed effects in the model. In doing so, our estimates are based on comparing changes in outcomes observed over time in municipalities affected by kingpin captures to the changes observed in other municipalities in the same state.

As described in the previous section, our analysis of "kingpin" captures focuses on DTO leaders and lieutenants, i.e., those at the very top level of the organization and those who control a state or region. We further restrict attention to the *first* capture of a kingpin for each DTO during the war on drugs. We do so out of concern for the endogeneity that would be introduced when the capture of a kingpin affects homicides while also increasing the probability of the capture of subsequent kingpins. By focusing instead on the effects of an initial capture, our estimates will reflect the effect of an exogenous kingpin capture inclusive of any effects that are driven by subsequent captures.

We note that standard-error estimation is not straightforward in this context. While we are evaluating a panel of municipalities, there may be reason to cluster standard-error estimates at some higher level(s) because different municipalities may have correlated shocks to outcomes not captured by our model. In some sense, because the source of variation is at the DTO level, it may be preferable to allow the errors to be correlated across municipalities when they share the presence of the same DTO. However, with only five DTOs, this would lead to problems associated with too few clusters. As a compromise, we instead cluster on DTO-combinations, which leverages the fact that there are some municipalities where two, three or four DTOs have a presence.[19] We additionally cluster on states to allow for some spatial correlation in the errors that might occur naturally or through policies implemented at the state level, following the approach to multi-way clustering described in Cameron, Gelbach, and Miller (2011).[20]

2.5 Results

2.5.1 Graphical Evidence of the Effects of Kingpin Captures

Before presenting the results of the econometric analysis described above, in this section we present graphical evidence. To begin, Figure A.4 plots the homicide rate over time separately for municipalities with a DTO presence before the war on drugs and those that did not have such a presence. This figure shows that municipalities with a DTO presence had higher—but not much higher—homicide rates than municipalities without a DTO presence in the six years leading up to the war on drugs. Moreover, they tracked one another quite closely. Perhaps most importantly, they even tracked one another after the beginning of the war on drugs—both dipping immediately before returning to close to their earlier levels—which provides support for using municipalities without a DTO presence as a meaningful comparison group for the purpose of attempting to separate the

[19] 2,084 municipalities have no DTO presence, 208 have one, 89 have two, 55 have three, and only 18 have four.

[20] This approach leads to somewhat more conservative standard-error estimates than clustering only on states, only on DTO combinations, or only on municipalities.

effects of kingpin captures from the effects of other aspects of the war on drugs. Twelve months after the beginning of the war on drugs, however, the two series began to diverge from one another in a dramatic way. While the capture of Alfredo Beltrán Leyva, leader of the Beltrán Leyva Cartel, would appear to be the most salient event to happen around this time that would disproportionately affect municipalities with a DTO presence, we cannot rule out other explanations such as a lagged effect of earlier aspects of the war on drugs. One explanation that we *can* rule out is that the war on drugs did not begin in earnest until this time—several major operations took place in 2007 which lead to the seizure of 48,042 Kg of cocaine, 2,213,427 Kg of marijuana, and 317 Kg of heroin, significantly more than the amounts seized in the subsequent years.[21]

Across the four panels of Figure A.5, we present graphs that more closely correspond to our regression analysis, which exploits variation in the timing with which kingpins from different DTOs were first captured and separately considers municipalities of capture, neighboring municipalities where the captured kingpin's DTO had a presence, non-neighboring municipalities where the captured kingpin's DTO had a presence, and neighboring municipalities where the captured kingpin's DTO did not have a presence. In particular, each panel shows the average difference between homicide rates in the municipalities of interest and the other municipalities in their states that have no DTO presence and do not neighbor a municipality of capture.

Before discussing the effects implied by these graphs, we note that each demonstrates a constant difference in the homicide rates of the municipalities of interest and their comparison municipalities prior to a capture, providing support for the common trends assumption underlying the difference-in-differences approach. Moreover, because all of the captures considered took place at least one year after the war on drugs was initiated, this suggests that the initial activities related to the drug war had similar effects on the munic-

[21]Third Calderón's Government Report (Tercer Informe de Gobierno, 2009).

ipalities of interest and their comparison municipalities, lending support to the idea that the difference-in-differences approach can identify the effects of kingpin captures in the broader context of the war on drugs.

In terms of the differences from comparison municipalities *following* captures, Panel A shows an immediate spike in the homicide rates in municipalities where the captures occurred. This difference appears to come back down 6–12 months after the capture—though not close to the pre-capture difference—before diverging again in a manner that suggests large long-run effects. Panel B focuses on the municipalities that neighbor these municipalities of capture where a captured kingpin's DTO also had a presence. For these municipalities, we see little evidence that homicide rates change (relative to comparison municipalities) following a capture. Panel C focuses on more distant (non-neighboring) municipalities where a captured kingpin's DTO had a presence. This panel indicates that such counties experience a steady increase in homicide rates (relative to comparison municipalities) following the kingpin capture. Panel D, which focuses on municipalities that do not have a DTO presence but which neighbor a municipality where a capture occurred, suggests a modest *decline* in homicide rates following a capture.

As a whole, the evidence shown in figures A.4 and A.5 supports the notion that kingpin captures escalate violence, particularly in the municipalities of capture and non-neighboring municipalities where a captured leader's DTO has a presence. It is less clear whether captures have effects on the other municipalities of interest, which we consider further in the regression-based analysis below.

2.5.2 Regression-based Evidence of the Effects of Kingpin Captures

Columns 1 through 3 of Table A.2 present our main results, based on the generalized difference-in-differences model represented by Equation 2.1. In particular, these columns show the estimated effects of a kingpin capture over time for the municipalities where a capture occurred, neighboring municipalities where the captured kingpin's DTO had a presence, non-neighboring municipalities where the captured kingpin's DTO had a presence, and neighboring municipalities where the captured kingpin's DTO did not have a presence. The estimates are based on models that control for municipality fixed effects and month-by-year fixed effects. Column 2 additionally controls for state-by-year-by-month fixed effects to address concerns that captures may be correlated with other state-level policy initiatives and/or shocks while Column 3 further adds controls for the effects of the war on drugs that are common to municipalities with a DTO presence by including variables for 0–5, 6–11, and 12+ months after the beginning of the war interacted with an indicator for the presence of a DTO in the municipality. Across these three columns, we note that the estimates are somewhat sensitive to the inclusion of state-by-year-by-month fixed effects and that the estimates are nearly identical but less precise when we additionally control for the effects of the war on drugs that are common to municipalities with a DTO presence.

Regardless of the exact specification, the estimates indicate significant effects of kingpin captures and considerable heterogeneity. In particular, the estimates reflect an immediate and sustained effect on the homicide rate in a municipality of capture of approximately 60%.[22] Due to relatively large standard error estimates, we can neither reject no effect or reject large effects on homicide rates in municipalities where the captured kingpin's DTO had a presence that neighbored the municipality of capture. The estimated effects on non-neighboring municipalities where the captured kingpin's DTO had a presence indicate little-to-no effect in the short run and a significant effect 12+ months following capture that implies that captures increase homicides 13% for these municipalities. The estimated effects are routinely negative for municipalities neighboring the municipality of

[22] As the outcome is the log of the homicide rate, the percent effects are calculated by exponentiating the coefficient estimate—in this case 0.476—and subtracting one.

a capture where the captured kingpin's DTO does not have a presence, suggesting that kingpin captures lead to a spatial displacement of violence.

Columns 4 and 5 of Table A.2 assess the validity of the research design by considering whether homicide rates deviate from their expected levels (based on their pre-capture levels and the changes observed in comparison municipalities in the same state) *prior to a kingpin capture* in any of these types of municipalities. These estimates are routinely close to zero and are never statistically significant, which provides support for a causal interpretation of the estimates discussed above.

2.5.3 Further Analyses Supporting a Causal Interpretation of the Main Results

In the same spirit as our analysis verifying that there are no "effects" before a kingpin capture occurs, which would otherwise suggest that our regression model is picking up something other than the effects of kingpin captures, in Table A.3 we separately consider the estimated effects on male homicide rates, female homicide rates, rates of domestic violence, and infant mortality using our preferred model. Whereas gender-specific homicide rates and infant mortality rates are constructed using the data described in Section 2.3, rates of domestic violence are constructed using administrative records of individuals arrested for the crime of domestic violence from Estadśticas Judiciales en Material Penal de INEGI. Because these data are only available beginning in 2003, our analysis of domestic violence spans 2003–2010 in contrast to all of our other analyses which span 2001–2010.[23]

The estimates in Table A.3 provide further support for a causal interpretation of our main results as they indicate: (i) the effects on overall homicides are largely driven by male homicides, which is consistent with gender differences in participation in the drug trade; (ii) there are no significant effects on domestic violence, which provides reassuring evidence that the main results are not driven by idiosyncratic shocks to levels of violence

[23]There are also fewer observations used in the analysis of infant mortality than in other analyses, because the outcome is undefined for cells in which the relevant population is zero.

coinciding with captures; and (iii) there are no systematic effects on infant mortality to suggest that the main results are driven by compositional changes towards a higher-risk population in the affected municipalities.[24]

Table A.4 presents evidence along similar lines, considering effects on homicide rates for males of different age groups.[25] These estimates indicate that the effects on males are driven by those between the ages of 15 and 44, mirroring participation rates in drug trafficking (Fairlie 2002; Vilalta and Martínez 2012). Moreover, the estimated effects on homicides rates for younger and older males tend to be close to zero and not statistically significant at conventional levels.

Though our main results are able to control for national and state-level policies and shocks common across areas in addition to those common to municipalities with a DTO presence through the inclusion of fixed effects, a potential concern with the empirical strategy is that it might conflate the effects of kingpin captures with the effects of military operations more broadly. We were able to speak to this issue above by showing that the municipalities of interest and their comparison municipalities tracked one another before the first captures took place, even after the war on drugs began. In order to further speak to this issue, Figure A.6 considers each of the eight major state (or multi-state) operations of the war on drugs in the timeframe spanned by our data.[26] In particular, each panel restricts attention to the state(s) of the operation and separately plots the homicide rate for municipalities with and without a DTO presence. Collectively, these eight panels indicate that the major operations of the drug war did not precipitate increases in homicides in

[24] Interestingly, the estimates do indicate significantly elevated rates of infant mortality following a capture in municipalities that neighbor a municipality of capture and have the same DTO presence. That said, these estimated effects on infant mortality do not line up with the estimated effects on homicides, which are not statistically significant and are neither routinely negative nor positive.

[25] The observations are not constant across columns as the outcome is undefined for cells in which the relevant population is zero.

[26] The beginning dates for these operations are based on information from the fifth Calderón's Government Report (Quinto Informe de Gobierno, 2011).

municipalities with a DTO presence relative to those without a DTO presence.

Figure A.7 also focuses on homicide rates as they relate to major operations of the war on drugs but instead considers the four major operations that focused on a single municipality or a small number of municipalities: the Marlin-Culiacán-Navolato Operation, the Laguna Segura Operation, the Tijuana Operation, and the Juárez Operation. This figure shows that all municipalities that were the target of an operation saw dramatic rises in their homicide rates at some point in time. More relevant to the validity of our identification strategy, there appears to be no consistent link between operations of the war on drugs and these rises—some of these municipalities saw their homicide rates begin to rise before an operation, some after, and some at around the same time. Instead, the spikes shown in panels (a) and (c) correspond to the capture of leaders which occurred in the municipality; the growth in the homicide rate in Laguna Segura is more gradual. While the rise in the homicide rate in the municipality of Juárez preceded the Juárez Operation, we note that our identification strategy does not rely on the conditional exogeneity of the timing of major operations but instead relies on the conditional exogenity of the timing of the captures considered. Moreover, the municipality of Juárez would have been affected by the Sierra Madre-Chihuahua Operation and the capture of the Juárez DTO lieutenant Pedro Sánchez Arras, both of which preceded the Juárez Operation.

As a whole, our analysis of state- and municipality-level operations suggests that military operations do not themselves lead to discernible changes in homicide rates. This is consistent with our earlier consideration of homicide rates in the months between the beginning of the war on drugs and the months in which kingpin captures took place, providing further support for a causal interpretation of our main results.

Table A.5 offers an additional check on the main results by considering the sensitivity of the estimates to the exclusion of any given DTO. In particular, across the columns of the table we report results systematically excluding from the analysis municipalities where the

Sinaloa-Beltrán-Leyva cartels have a presence (Column 2), where the Tijuana Cartel has a presence (Column 3), where the Gulf Cartel has a presence (Column 4), where the Juárez Cartel has a presence (Column 5), and where the Familia Cartel has a presence (Column 6), respectively. This analysis is motivated by the notion that we should be less confident in the results if they are driven by municipalities associated with any one particular DTO. The estimates are most sensitive to the exclusion of municipalities where the Gulf DTO has a presence, which is perhaps not surprising in light of the fact that the the Gulf DTO spans the most municipalities and the municipalities with the largest populations (as shown in Table A.1). That said, the estimated effects are actually larger when these municipalities are excluded from the analysis and thus the estimates guide us to the same conclusion regardless of whether any one DTO is excluded from the analysis—kingpin captures lead to large and immediate increases in the homicide rates for municipalities where captures occur and this effect is quite persistent; there are spillover effects onto non-neighboring municipalities where the captured kingpin's DTO has a presence in the long run; and there appear to be effects in the opposite direction for neighboring municipalities where the kingpin's DTO does not have a presence.

2.6 Discussion and Conclusion

In the preceding sections, we have estimated the effects of the first kingpin captures during Mexico's war on drugs for the DTOs that were in operation prior to the war. Newly available data on DTOs' areas of operation at the municipality level over time and monthly data on homicides allow us to control for a rich set of fixed effects and to leverage variation in the timing of kingpin captures to consider the effects on homicides in the area of capture itself in addition to other areas where the kingpin's DTO has a presence. The results of this analysis indicate that kingpin captures have large and sustained effects on the homicide rate in the municipality of capture and smaller but significant effects on other municipalities where the kingpin's DTO has a presence, supporting the notion that the kingpin strategy can have destabilizing effects throughout an organization while highlighting that this does not imply a reduction in violence. That being said, kingpin captures do appear to reduce homicides for municipalities neighboring a municipality of capture where the captured kingpin's DTO does not have a presence.

These estimates offer a new lens through which we can view the dramatic increase in violence in Mexico since the beginning of the war on drugs. In particular, our estimates suggest that the kingpin captures we consider led to an additional 4,934 homicides between 2007 and 2010, or approximately 7.2 percent of the homicides over that period of time. Roughly 30 percent of these additional homicides are due to spillover effects onto non-neighboring municipalities where a captured kingpin's DTO has a presence. In total, the effects of the kingpin captures we consider can explain 31.8 percent of the increase in homicides between 2006 and 2010.[27] An important caveat to these figures is that we use an imperfect measure of DTOs' areas of operation (based on the MOGO approach described above) and that misclassification would serve to bias our estimates towards zero—as such, they may be best thought of as estimates of the lower bound of the true effects.

While our estimates indicate that Mexico's use of the kingpin strategy caused significant increases in homicides, it is important to note that its war on drugs had several objectives beyond reducing violence, including the establishing the rule of law, that need to be considered in evaluating the policy. Moreover, it remains possible that the kingpin strategy could reduce violence in the long-run in ways that have yet to be seen.

[27]These numbers were calculated using the regression coefficients corresponding to Column 3 of Table A.2. In particular, they are calculated by comparing the predicted number of homicides based on the regression model under the true values of all regressors and the predicted number of homicides with all treatment variables set to zero. These calculations indicate that the capture of kingpins caused an increase in homicides of 49 in 2007, 891 in 2008, 1,864 in 2009 and 2,130 in 2010.

3. THE SHORT AND LONG RUN EFFECTS OF FULL-TIME SCHOOLS ON ACADEMIC PERFORMANCE

3.1 Introduction

In many developing countries the demand for schooling so overextends the available resources that schools offer their services in two or three shifts throughout the day. As capacity catches up with need, policy-makers can redirect their focus to the quality of education that students receive. Some countries have tried to do so by affecting different inputs of the education production function, such as teachers' quality, schools' resources, class sizes, and instructional time. As countries continue to consider making these sorts of investments, it is important to understand their costs and benefits, in addition to understanding which sectors of the population could benefit the most from these policies. To this end, this paper studies the effect of extending the school day on student achievement in Mexico, where more than 23,000 schools extended their school day by three and a half hours between 2007 and 2015.

Policies extending the school day have become more and more popular during the last two decades.[1] The motivation behind these policies is that more time in the classroom should translate into better learning outcomes. However, there is an ongoing debate in the education literature over whether lengthening the school day in fact improves academic performance. On the one hand, those who expect large positive effects argue that longer school days provide students with more time for learning, a deeper coverage of the academic curricula, and more time for any given task. On the other hand, those expecting small or even negative effects argue that students tend to waste time, decrease effort, get exhausted, and have less time for informal learning and other recreational activities (Patall

[1] For example, Argentina, Brazil, Chile, Colombia, Dominican Republic, El Salvador, and Uruguay extended their school day between 1996 and 2011 (Holland et al., 2015).

et. al., 2010). Given the sound logic underlying both types of arguments, this is a clear area in which empirical evidence is necessary in order to document the sign and the magnitude of the effects of the policy in question.

To offer evidence on the effect of longer school days on student academic performance, I take advantage of a natural experiment and a rich dataset on student performance and school characteristics. In particular, I analyze a natural experiment in Mexico where more than 23,000 schools[2] have extended their school day from 4.5 hours (half-time schools) to 8 hours (full-time schools) since 2007. My empirical strategy exploits variation in the timing with which each school extended its school day to estimate the impact of this intervention on students' reading and math test scores. I use a rich dataset on student achievement at the student-level as well as schools' and teachers' characteristics that allows me (i) to identify the short- and long-run effects of extending the school day on students' test scores; (ii) to provide evidence that absent the extension of the school day, test scores of students in full-time schools and half-time schools follow the same trends; (iii) to identify heterogeneous effects across subjects, across grade levels, and across different types of communities; (iv) to estimate the effect of the intervention on school-level enrollment, gender composition of the classes, class sizes, and some measures of teacher quality, addressing the potential of students and teachers systematically sorting themselves into full-time schools based on their observed and unobserved characteristics; and (v) to provide lower and upper bounds for the estimates in light of the effects on enrollment by making extreme assumptions about changes in composition induced by the intervention.

The existing literature on extended school day programs in developing countries suggests that extending the school day has improved academic performance in countries like Uruguay (Cerdan-Infantes and Vermeersch, 2007), Chile (Bellei, 2009), and Colombia (Hincapie, 2013). Using difference-in-differences and propensity score matching method-

[2]This number roughly represents a 26 percent of all primary schools in Mexico.

ologies, they estimate effects on student achievement that range between 0 and 12.6 percent of a standard deviation. However, these studies suffer from some limitations that this paper is able to avoid. First, due to the fact that they have little or no pre-intervention data, these studies are not able to provide much evidence in favor of the common trends assumption needed for their difference-in-differences estimates to be valid. Second, these studies evaluate shorter-run effects on student achievement.[3] Finally, these studies focus on the effects for older children or for students in high-poverty areas. Therefore, a more comprehensive understanding of the effects over time, on younger children and on children from different types of backgrounds, is important for thinking about where these sorts of policies should be targeted in the future.[4]

More closely related with this study, Cabrera-Hernandez (2015) also considers the effects of extending the school day on academic performance in Mexico. He uses difference-in-differences and propensity score matching methodologies that compare school-average math and reading test scores before and after the intervention. Whereas he uses aggregated school-level data, I use student-level data. Moreover, I estimate positive effects on enrollment and evaluate the degree to which this might bias the estimated effects on achievement. Another important difference is that I consider heterogeneous effects of the intervention for students of different grade levels, allowing me to differentiate the effects for younger versus older children and to consider the degree to which the extension the school day has more beneficial effects over time.[5]

The main results indicate that extending the school day does not affect test scores

[3]For example, Bellei (2009) estimates the effects of a 2-year exposure to extended school days on tenth graders' math and reading test scores.

[4]Heckman (2008) shows that the rate of return to investment in human capital is the highest for early childhood interventions and for programs targeted at disadvantaged children who are less likely to receive parental investments to compensate for shortcomings of the public schools.

[5]In related work, Agüero and Beleche (2013) study the effect of extending the school *year* in Mexico. They find that a 10-day increase in the length of the school year increases academic performance between 4 and 7 percent of a standard deviation.

the year of adoption; however, it increases students' math and reading test scores by 4.9 percent and 4.1 percent of a standard deviation, respectively, a year following adoption. These effects grow to 14.7 percent in math and 15.6 percent in reading four years after the adoption of the program, highlighting the fact that educational interventions that have no immediate effects may have very large effects in later years. In addition, I show that the first year of implementation seems to be crucial for schools and teachers to learn how to utilize the additional time more efficiently. Moreover, the estimated effects are greater for students in schools located in high poverty communities and for students in lower grade levels.

In an effort to show that the main results are not driven by changes in other factors of the education production function or changes in the composition of students,[6] I estimate the effects of extending the school day on enrollment, the gender composition of the classes, class sizes, the number of teachers, and measures of teachers' quality. These results indicate that the intervention does not affect the gender composition of classes, the number of teachers, or teachers' quality. However, I find positive and significant effects on class sizes and enrollment.

The increase in class size is generally found to negatively affect academic performance (Krueger, 1999; and Urquiola, 2006), so the estimated effects of extending the school day more than offset any negative effects of increased class sizes.

The enrollment effects suggest that it is possible that the composition of students at schools changes after the intervention. That is, full-time schools might be more attractive to high (low) ability students. To account for these possibilities, I use a modified version of Lee's (2009) trimming procedure for constructing bounds. I provide two sets of lower and

[6]For example, researchers have shown that improving teachers quality (e.g., Rockoff, 2004; Rivkin et al., 2005; Santibañez, 2006; and Chetty et al. 2014), increasing the proportion of female students (e.g., Hoxby, 2000; and Lavy and Schlosser, 2011), attending single-gender classes (Lee et al., 2014; and Eisenkopf et al. 2015), and reducing class sizes (e.g., Krueger, 1999; Angrist and Lavy, 1999; Hoxby, 2000; and Urquiola, 2006) increased academic performance in both developed and developing countries.

upper bounds for the treatment effects by making extreme assumptions about the change in composition of students. These results indicate that even if it is the highest performers who switch schools, the estimated effects are positive and statistically significant.

The rest of the paper is organized as follows. Section 3.2 provides institutional background on the Full-Time Schools program in Mexico; Section 3.3 describes the data; Section 3.4 describes the identification strategy and the bounding procedure; Section 3.5 presents graphical and empirical results; and Section 3.6 concludes.

3.2 Full-Time Schools Program in Mexico

The FTS program, which began in 2007 extended the school day in public schools of basic education.[7] The goal of the program is to improve children's rights to equitable and inclusive, quality education. In doing so, full-time schools promote a deeper learning of the academic curricula, including English as a second language, and the use of technology as a learning tool, as well as practical information on healthy living, regular physical activity, and appreciation for art and culture.

Schools that participate in the program increase their school day from 4.5 to 8 hours during the same 200 days as other public schools, and increase their instructional time from 800 to 1200 hours per academic year.[8] Since 2007, more than 23,000 schools have adopted the FTS program, reaching 3,463,041 students.

The federal government disperses FTS funds to the states and the Federal District (Mexico City). These funds are used to supplement teachers' and principals' wages for the extended schedule, to arrange and equip school areas, and to support the provision of food in schools (DOF, 2012). Although the FTS program targets all public schools of basic education, preference in funding is given to those schools that operate in only one shift, do

[7]Basic education in Mexico includes preschool (P1-P3), elementary school (1st-6th grades), and junior high school (7th-9th grades).

[8]These increases in instructional time are as follow: from 240 to 320 in Spanish, from 200 to 280 in math, and from 360 to 600 in other subjects for 3rd to 6th grade students.

not share the building with another institution, have a teacher per class, are located in urban areas, and have poor educational outcomes.[9] The states' authorities decide which schools will participate in the program based on the aforementioned schools' characteristics and their budget constraint, although the selection process is not clearly defined.[10]

Table A.6 shows the annual federal budget for the program for each academic year along with the number of participant schools and the number of students enrolled in participating schools.[11] In the first academic year, 500 schools received the funds and expanded their school day. By the 2014-2015 academic year, the program had reached 23,182 schools. Figure A.8 displays the staggered implementation of the program across Mexico. Note that by the 2012-2013 academic year all states in Mexico have full-time schools.

3.3 Data

The two datasets used in this study come from the Ministry of Education (SEP) in Mexico and the National Population Council (CONAPO). The primary dataset contains information at the student level that covers the 2007-2008 to 2012-2013 academic years. This dataset includes all public-school students' math and reading test scores, and grade level. Test scores are based on annual student level math and reading test scores from the National Assessment of Academic Achievement in Schools (ENLACE). ENLACE is a standardized test from the National Education System that is conducted in all public and private schools of basic education in Mexico.[12] The test covers math, reading, and a third subject that rotates over time.[13] It is composed of approximately 50 multiple-choice

[9]Every academic year, the Ministry of Education discloses the targeted schools for that year in the program's operating rules.

[10]I will discuss, in Section 3.4, how this fact might be a threat to the validity of the main results.

[11]These amounts are in addition to the budget previously allocated for the operation of schools in the states.

[12]In elementary school, ENLACE is administered to students in grades three to six.

[13]Science in 2008, Civic and Ethical Formation in 2009, History in 2010, Geography in 2011, Science in 2012, and Civic and Ethical Formation in 2013.

questions for each subject, and the scores range from 200 to 800.[14] The main advantage of ENLACE test scores is that they are intended to be comparable across students and over time, and the test is administered by people unrelated to the class.[15] The test is conducted during eight sessions of 45 minutes each over two days. For this reason, some students do not answer both reading and math section or do not respond to all the questions for one of the subjects. Additionally, after the tests are conducted, they are verified in an automatized system to check whether the results are reliable.[16] I drop from the main analysis test scores of students that did not answer at least half of the questions in a given subject or students with test scores classified as unreliable by the automatized system.[17]

In addition to test scores, the ENLACE dataset also provides information on the location of each school. I use this information to match each school to the poverty index of the locality[18] where the school is located. This allows me to consider heterogeneous effects of the impact of the FTS program on test scores for students living in high- and low- poverty areas.[19] The poverty index is estimated by CONAPO as a measure of social exclusion, and it is composed of localities' deficiencies in terms of education, housing, population, and households' income. To avoid concerns about endogeneity, I use the poverty index of the

[14]I normalize test scores across the entire sample for each exam (grade level, subject, and year) to have mean zero and variance one.

[15]Although ENLACE test scores are available since 2005/06, I only include in my sample test scores from 2007/08 to 2012/13 because, as shown in Figures A.10 and A.11, there are inconsistencies in the distribution of test scores for the first two years of application. However, including those years in the analysis does not affect the main results.

[16]The automatized system uses the K-index and Scrutiny methods, which are based on similar incorrect response patterns, to detect answer copying. In case of detecting any irregularity, it is recorded in the individual reports for the student.

[17]In Table A.12, I show evidence that the fraction of students with such test scores is not affected by the FTS program.

[18]A locality is a generic territorial division for a population center with its own identity. It can be small in size and population (village) or large and highly populated (city). The union of several localities form a municipality. The National Institute of Statistics and Geography keeps control of the list of localities in Mexico.

[19]Students living in high- and low- poverty areas are defined as those living in a locality above and bellow the median poverty index.

localities in 2005, which is prior to the introduction of the FTS program.[20]

The treatment variable is an indicator of the year in which each of the schools adopted the FTS program. The information on if and when each school adopted the program comes from the Ministry of Education. The schools that were incorporated to the program in 2007/08 are dropped from the main analysis because they already had an extended school day, making a before-and-after comparison impossible.

The secondary dataset contains information at the school level on schools' and classes' characteristics; it is based on school-level census data from *Estadísticas 911* collected by the Ministry of education at the beginning of every academic year. This dataset contains information on total enrollment, enrollment by grade, gender composition of the classes, average class sizes, number of teachers, and the fraction of teachers with some graduate education for each academic year.

3.4 Identification Strategy

3.4.1 Estimation Method

To identify the effect of longer school days on student achievement, I exploit variation in the timing with which schools adopted the program between the 2008-2009 and 2012-2013 academic years. The main results are based on a difference-in-differences research design. Formally, I estimate the effects of extending the school day using the following regression model:

$$TS_{isgt} = \sum_{k=0}^{4} \delta_k FTS_{st-k} + \alpha_s + \gamma_t + X_{sgt}\beta + u_{isgt} \tag{3.1}$$

where TS_{isgt} is the math or reading test score of student i in school s, grade g, year t; FTS_{st-k} is an indicator variable that takes the value of one for full-time schools k years

[20] In particular, it is possible that the poverty index will decrease in localities with a larger share of students enrolled in full-time schools.

after they adopt the program and zero otherwise; γ_t are year fixed effects; X_{sgt} includes time varying school-grade controls; and u_{istg} is an error term. The coefficient of interest is δ_k, which measures the average treatment effect of the FTS program k years after its implementation.

The identifying assumption is that in the absence of the FTS program, students in adopting schools would have experienced changes in achievement similar to those in non-adopting schools. I address the validity of this assumption in two ways. First, I look for graphical evidence to see whether test scores of students in half-time and full-time schools' do not diverge *prior* to treatment. Second, I test this assumption formally by including indicators from one year and two years *prior* to treatment in Equation 3.1; if the estimated coefficients for the leading indicators are close to zero and statistically insignificant, it provides support for the validity of the identifying assumption.

Although treatment is at the school level, there might be shocks not captured by this model that are correlated across students, such as unobserved state-level education policies or expenditures, or tests with different levels of difficulty each academic year. In order to account for potential within state and within grade-by-year correlation of the errors, I estimate robust standard errors that are two-way clustered at the state and grade-by-year level using the procedure described in Cameron, Gelbach, and Miller (2011).

3.4.2 Bounding Estimates

One potential threat to my ability to identify the causal effect of extending the school day on academic performance is that parents might be motivated to switch their children from half-time to full-time schools. If such decisions are independent of students' abilities, there would not be a concern; however, if it is high (low) performing students who systematically switch from half-time to full-time schools, the estimated coefficients would be biased upwards (downwards). Ideally, if I were able to follow students over time, I could control for their unobserved characteristics by including individual fixed effects. However, given that the data do not include unique student identifiers, this is not feasible.

In order to asses the sample selection problem that occurs when treatment affects sample attrition, Lee (2009) proposes an intuitive trimming procedure that bounds the average treatment effects by making extreme assumptions about missing observations. Lee identifies the excess number of individuals who are likely to select into treatment, and then he trims the upper and lower tails of the outcome distribution by this number of individuals. That is, he obtains the worst-case and best-case scenarios for the estimated effects. In the spirit of Lee (2009), I construct bounds based on the opposing assumptions that the students that switch from half-time to full-time schools are either the highest or the lowest performing students. However, since I observe the universe of test scores, I do not trim the tails as Lee suggests. Instead, I reassign the students at the extreme end of the full-time schools' test score distribution to a half-time school. Statistically, reassigning the highest performing full-time students from full-time schools to half-time schools represents the extreme assumption that the highest performing students systematically select into full-time schools. Imagining that full-time schools' highest performing students had instead remained in half-time schools and reestimating the coefficients yields a lower bound of the effects of full-time schools. Conversely, the extreme assumption that the lowest performing students systematically select into full-time schools yields an upper bound to the effect of the extended school day.

The validity of this approach relies on a monotonicity assumption (i.e. the flow of students is only in one direction, from half-time to full-time schools). This assumption seems plausible in this context for three reasons. First, 76 percent of full-time schools' principals saw increases in school applications. Second, 93 percent of parents with children in full-time schools are satisfied with the schools that their children attend. Finally, 91 percent of parents with children in full-time schools perceive that the FTS program improves children's education outcomes (CONEVAL, 2013).

3.5 Results

3.5.1 Main Results

In this section, I examine whether the FTS program improves students' academic performance. Figure A.9 panels (a) and (b) show the average difference in math and reading test scores, respectively, between students in full-time schools and students in half-time schools from 2007-2008 to 2012-2013. These panels indicate that extending the school day increases student-average math and reading test scores. Moreover, test scores of students in full-time and half-time schools have a similar trajectory *prior* to the implementation of the program; this suggests that math and reading test scores of students in full-time schools would not have increased relative to those in half-time schools in the absence of the program.

Table A.7 panels A and B show the regression-based estimates for math and reading test scores, respectively. Column 1 shows the estimated effects based on the difference-in-differences model described by Equation 3.1. Estimates in Column 2 additionally control for grade-specific unobserved characteristics that are constant over time. In Column 3, I also control for possible heterogeneity in the different ENLACE tests for each grade and year by including grade-by-year fixed effects. In Column 4, I additionally control for school-by-grade fixed effects. In Column 5, I present the preferred specification which additionally controls for other programs in which the schools participate that might affect students' performance, these include the Quality Schools Program and the Secure School Program.[21] Finally, in columns 6 and 7, I perform a formal statistical test to check whether

[21] The purpose of the Quality Schools Program is to decentralize the decision making process, to improve infrastructure, and to provide school supplies in public schools of basic education (see Skoufias and Shapiro, 2006 for a deeper discussion of the program). The purpose of the Secure Schools Program is to provide schools with technical and financial supports to prevent addiction, delinquency and violence in schools of basic education (SEP, 2007).

test scores of students in schools that adopted the program diverge from the scores of students in schools that did not adopt the program *prior* to the adoption. I do so by including indicator variables for one year and two years prior to the adoption of the FTS program using the preferred specification.

Estimates in Panel A indicate that extending the school day has no significant effect on students' math test scores the year of adoption. However, it increases math performance by 5 percent of a standard deviation a year after the adoption, 7.7 percent two years after, 12.4 percent three years after, and 14.7 percent four years after.[22] Additionally, the estimated coefficients for one year and two years prior to treatment, in columns 6 and 7, are close to zero and not statistically significant. This provides evidence in favor of the validity of the identifying assumption.

The estimates in Panel B indicate a similar pattern of effects on reading scores. Longer school days do not affect students' reading test scores the year of adoption; however, it increases reading performance by 4.1 percent of a standard deviation one year after the adoption, 6.8 percent two years after, 11.6 percent three years after, and 15.6 percent four years after.[23] In addition, the estimated coefficients for the indicators of one year and two years prior to treatment, in columns 6 and 7, are not statistically significant. This suggests that reading test score trends did not diverge from expectations before the introduction of the program.

Even though both graphical- and regression- based evidence suggest that extending the school day is effective in increasing students' academic performance, it is important to note that these apparent effects could be driven by (i) changes in the composition of students at schools, or (ii) changes in other factors of the education production function after schools adopt the program. Regarding the composition of students, we might be

[22]Note that the inclusion of grade and grade-by-year fixed effects do not affect the estimates; I do not expect them to change because I normalize test scores for each test (i.e. subject, grade, and year).

[23]This effect represents more than two thirds of a year of schooling (Carlsson et al., 2015).

concerned that longer school days might signal a better quality of education, so motivated parents might switch their children to schools with longer school days. Also, parents might prefer to keep their troublesome children busy at school but not their well-behaved ones, or they might prefer to keep their boys for longer time at school but not their girls. Regarding other factors of the education production function, full-time schools might attract more (less) qualified teachers, or they might attract more teachers and, therefore, have smaller classes. In any of these scenarios, the estimated effects might be driven by these changes rather than by the extension of the school day.

In consideration of these issues, Table A.8 shows the estimated effects of the FTS program on enrollment, the fraction of students who are male, class sizes, the number of teachers, and the fraction of teachers with some graduate education. While the FTS program has no significant effects on the gender composition, the number of teachers, or on the fraction of teachers with some graduate education, it *does* increase both enrollment and class size. Specifically, the estimates in Column 1 indicate that extending the school day increases student enrollment by 4.9 percent the year of adoption, and the estimated effect grows to 7.9 percent four years after adopting the program. Additionally, Column 3 shows that the FTS program causes an increase of approximately one student per class. Considering that increasing class sizes reduces academic performance in both developed and developing countries (Krueger, 1999; and Urquiola, 2006), the change in class size is almost certainly not exaggerating the positive effect of extending the school day. Therefore, I concentrate my efforts on understanding how students switching schools might affect the main results.

3.5.2 Bounded Treatment Effects

The estimated effects of the FTS Program on enrollment suggest that it is possible that the composition of students at schools changes after the intervention. To address the potential bias that could possible be caused by students switching schools, I consider two approaches to constructing bounds using the procedure described in Section 3.4.

To begin, I consider the two extreme scenarios of the switching behavior to construct lower and upper bounds. First, I consider the scenario in which it is the highest performing students who are switching schools. Intuitively, this extreme scenario can be addressed by reassigning the highest performers in full-time schools to half-time schools. That is, I reassign treatment status to the top 2.3 percent of the students' math/reading test scores distribution for each test in the treated schools the year of adoption, the top 3 percent a year after, the top 3.3 percent two years after, the top 3.5 percent three years after, and the top 3.8 percent four years after.[24] I then reestimate the coefficients, obtaining the lower bounds. Second, I consider the scenario in which it is the lowest performing students who are switching schools. I do so by reassigning the lowest performers in full-time schools to half-time schools. That is, I reassign treatment status to the bottom 2.3 percent of the math/reading student-grade-school-year test scores distribution for the treated schools the year of adoption, the bottom 3 percent a year after, the bottom 3.3 percent two years after, the bottom 3.5 percent three years after, and the bottom 3.8 percent four years after. I then reestimate the coefficients to obtain the upper bounds for the effect of the FTS program on student achievement.

My second bounding approach is motivated by the fact that the effect on enrollment is immediate and the effect on student performance does not appear until a year after adoption. This evidence suggests that the initial enrollment effect (4.9 percent) does not increase test scores for schools changing to full-time schedules. To account for this possibility, I construct a second set of lower and upper bounds by reassigning treatment status to

[24]These numbers are calculated using the regression coefficients corresponding to Column 1 of Table A.8 and Equation A.3. For example, the year of adoption, 2.3%=$\left(1 - \frac{1}{1+\frac{1}{2}(0.049)}\right)100\%$. See Appendix A.2.1 for further details on these calculations.

the *remaining* proportion of students switching schools.[25] This is, I reassign treatment to the top/bottom 0.6 percent of the math and reading student-grade-school-year test scores distribution for the treated schools a year after adoption, the 1 percent two years after, the 1.1 percent three years after, and the 1.4 percent four years after.[26] I then reestimate the coefficients obtaining lower and upper bounds.

Table A.9 shows the estimates based on my preferred regression specification and its lower and upper bounds using the two bounding procedures described above. Column 1 shows unbounded estimates, columns 2 and 3 show the set of lower and upper bounds for the first bounding approach, and columns 4 and 5 show the set of lower and upper bounds for the second approach.

The bounds based on the first approach, in columns 2 and 3, indicate treatment effects on math test scores between 0 and 5.1 percent of a standard deviation the year of adoption, and between 5.8 and 23.4 four years after adoption. The estimated effects for reading test scores (Panel B) follow the same pattern and are similar in magnitude.

The results of the second procedure are reported in columns 4 and 5, and the estimated lower and upper bounds for the treatment effects are much tighter. I find that there is no significant effect on math performance the year of treatment, but the effects grow to between 14.4 and 14.9 percent four years after adoption. The estimated effects for reading test scores (Panel B) follow the same pattern and are similar in magnitude.

Given that the lower and upper bounds are similar in magnitude using my preferred bounding approach, in the last section I focus only on unbounded estimates.

[25] Students who switch schools after the first year of the implementation of the FTS program.

[26] These numbers are calculated using the regression coefficients corresponding to Column 1 of Table A.8 and Equation A.7. For example, the year after adoption, 0.6%=$\left(1 - \frac{1+\frac{1}{2}(0.049)}{1+\frac{1}{2}(0.062)}\right)100\%$. See Appendix A.2.1 for further details on these calculations.

3.5.3 Treatment Effect Heterogeneity

Given resource constraints, policy-makers have to consider how to efficiently allocate their budget by identifying which types of schools and students should be targeted. For this purpose, in this section, I explore the degree to which there are heterogeneous effects on test scores by grade and by the poverty level of the area where the schools are located.

3.5.3.1 Grade Level Heterogeneity

Table A.10 panels A and B separately consider the estimated effects on math and reading test scores for each grade using my preferred specification.[27] While these estimates hold important insights, some care is required in interpreting the estimates across the various rows and columns, which correspond to the estimated effects across time and across grades. In particular, the estimates need to be interpreted in light of the fact that (i) the effects of the program may grow over time as schools learn how to make use of the additional class time, and that (ii) the effects are likely to be stronger for students/grades who have had an opportunity for full-day schooling for a larger share of their schooling. In discussing the estimates reported in the table, I highlight the degree to which each of these factors is relevant.

Column 1 Panel A presents the estimated effects on math test scores for students in third grade. These results indicate that extending the school day does not have an effect on third graders' academic performance the year of adoption. However, it increases test scores by 4.9 percent of a standard deviation a year after adoption, and this effect grows to 8.7 percent two years after adoption.

In order to consider the degree to which schools become more effective at utilizing the additional class time over time, we can compare the estimated effects of newer full-time

[27] Particularly, I estimate the effects for students in grades third to sixth, which are the students for which ENLACE test scores are available.

schools (unbolded) with the estimated effects of full-time schools that have gain more experience over time (bolded). The bolded estimates indicate that the FTS program increases third graders' math test scores by approximately 14 percent of a standard deviation for students that have been enrolled in full-time schools for three years while the schools have been part of the program for four or five years. These results suggest that the first year of implementation is crucial for teachers and students to learn how to utilize the additional time more efficiently.

Similarly, Column 2 Panel A presents the estimated effects on fourth graders' math test scores. These estimates provide additional support for the notion that schools get better at using the extra class-time over time. The estimates in columns 3 and 4 show the estimated effects for fifth and sixth graders, respectively. For all grades, test score gains increase with each additional year in the FTS program, but the gains are consistently higher for younger students. In math, the gains for third graders reached 8.7 percent of a standard deviation two years after adoption, compared to 5.3 percent for sixth graders.The estimates in Panel B indicate a similar pattern for reading test scores.

3.5.3.2 Poverty Level Heterogeneity

Children from high poverty areas might be exposed to multiple factors that affect academic performance such as poor health and nutrition, or parents' absenteeism that could be alleviated by keeping and feeding them at school. Panels A and B of Table A.11 separately consider the effects on student performance for students in schools located in low and high poverty localities. In particular, columns 1 and 2 show the effects for students below and above the median poverty index of the locality where the school is located, which I refer to as low and high poverty levels. Schools in high poverty localities make statistically significant gains in test scores one year after the adoption of the FTS program, whereas test scores gains in low poverty schools do not become statistically significant until three years after adoption. For both high and low poverty schools, test score gains increase with each additional year in the FTS program, but the gains are consistently higher for high poverty schools. In math, the gains for high poverty schools reached 15.2 percent of a standard deviation four years after adoption, compared to 9.7 percent in low poverty schools. Similarly, in reading, the four-year gains in high and low poverty schools are 16.4 percent and 10.8 percent, respectively. Because test scores tend to be higher in low poverty schools, the greater gains in high poverty schools would seem to reduce the academic gap between the poorer students and the more affluent students.

In summary, the estimated effects of the extended school day differ by students' age and the poverty status of their locality. The effects are stronger for those students in lower grade levels and for students in schools located in high poverty communities.

3.6 Conclusion

In this paper, I study the causal effects of extending the school day on students' academic performance by analyzing a large-scale education program in Mexico—the Full-Time Schools Program—in which participant schools increased the length of their school day from 4.5 to 8 hours. I take advantage of the staggered rollout of the program and a novel dataset to disentangle the causal effects of extending the school day from the effects of other possible confounding factors. Because the program has an effect on enrollment, I use an intuitive bounding procedure that allows me to address the potential for selection bias caused by non-random switching into full-time schools. The results indicate that extending the school day increases both reading and math test scores. In addition, the long follow-up period allows me to demonstrate that the effects increase considerably over time. Moreover, the estimated effects are more pronounced in schools located in high poverty localities and for students in lower grade levels. This evidence suggests that as developing countries expand their capacity to deliver education, extending the school day can be an effective way to improve the quality of education they offer and it can be effective to close the education gap between the rich and the poor. A broader implication of my results is that early assessments of educational interventions may be poor predictors of longer run effects.

4. THE EFFECT OF CHILDREN'S TIME IN SCHOOL ON MOTHERS' LABOR SUPPLY: EVIDENCE FROM MEXICO'S FULL-TIME SCHOOLS PROGRAM

4.1 Introduction

Despite the growth in female labor force participation (LFP) in recent decades, female participation rates have remained lower than their male counterpart. Moreover, this gap is especially large in the developing world, where traditional gender roles assign women the primary responsibility of childrearing. As a result, women's labor supply in developing countries heavily depends on their fertility decisions, and specifically, on how much they are time- and budget- constrained due to childrearing and the alternative costs of childcare institutions. Thus, the availability and affordability of childcare centers are important for women to increase their labor market participation while their children are still growing up.[1]

Studies of the US have proposed that the absence of family-friendly policies, including parental leave and part-time work entitlements, explains 28-29 percent of the decrease in female labor force participation in the US, relative to other OECD countries, over the period from 1990 to 2010 (Blau and Kahn, 2013). Different governments around the globe have responded to the low female participation rates in the labor market with a variety of policies such as tax reliefs, child benefits, paid leaves and childcare subsidies. As different countries continue to consider these types of policies, it remains important to understand their costs and benefits. To this end, this paper studies the effect of an implicitly large childcare subsidy, through longer school-days in primary education, on mothers' labor supply at the extensive and intensive margins.

When the public provision of regulated childcare institutions is low or absent, moth-

[1]For a discussion of such influences on female labor participation in the context of the countries in the Organization for Economic Cooperation and Development (OECD), see Jaumotte (2003).

ers' chances to participate in the labor market may decrease depending on the supply and quality of the available alternatives for childcare. The option to take care of their children may range from a costly private institution with an uncertain quality to non-professional options, such as family members, close friends, and untrained babysitters (Bernal and Keane, 2011; Schady et al., 2015). In such contexts, Full-Time School (FTS) programs work as a childcare alternative provided by trained caregivers (i.e. teachers) in a controlled environment. Consequently, FTS programs have the potential to positively affect children's outcomes along with mothers' labor force participation (LFP).[2]

To provide evidence on the effects of extending the school-day in primary education on mothers' labor supply, we take advantage of a natural experiment in Mexico where the government implemented a FTS program that extended the school-day from four and a half to eight hours in primary schools (1st-6th grades) all over the country between 2007 and 2014. Our empirical strategy exploits within-individual variation in exposure to full-time schools— defined as the share of predicted FTS seats in a municipality—to estimate the effects on female labor supply at the extensive and intensive margins.

We use ten years of data collected in the National Employment and Occupation Survey in Mexico (ENOE, for its abbreviation in Spanish). ENOE is a rotating panel of households that contains information on mothers' labor force participation, number of weekly hours worked, earnings, and sociodemographic characteristics that allow us to identify the cumulative effects of longer school days on mothers' labor supply, to identify heterogeneous effects by education level and by poverty level of the locality of residence, and to provide evidence that the effects are not driven by changes in the propensity to participate in the labor force in municipalities with full-time schools.

[2]Researchers have documented positive effects of FTS programs on children's outcomes in the short- and the long-run. For example, full-time schools improve children's academic performance (Bellei, 2009; Cabrera-Hernandez, 2015; Padilla-Romo, 2015), reduce high school dropout rates (Pires and Urzua, 2010), and reduce the probability of teenage pregnancy (Kruger and Berthelon, 2009).

The existing literature has focused on evaluating the impact of childcare institutions for preschool age children (3 to 5 years old) in developed countries, showing some positive effects on mothers' labor supply.[3] However, the differences between richer and poorer countries in labor institutions and trends in female labor supply reduce the scope of such evidence to guide policies in developing countries where, additionally, mothers' low LFP is commonly attributed to cultural factors besides economic conditions.

Evidence for developing countries is scarce but shows a higher likelihood of mothers' employment after increases in childcare supply. Berlinski and Galiani (2007) estimate the effects of an 18 percent increase in preschool availability between 1994 and 2000 in Argentina and find that the likelihood of maternal employment increased between 6 and 16 percentage points depending on the model specification. Similarly, in the case of Mexico, Ángeles et al. (2011) use a time discontinuity in children's eligibility to "Estancias Infantiles," a public childcare program for 265,415 preschool-age children (0-4 years old) all over the country and find an increase of 18 percent in mothers' probability of employment and an average effect of six more hours worked per week.

Few studies have focused on children aged 6 years and older who are still in need of parents' care. Even less so in developing countries.[4] This is an important omission because many children in developing countries are in school for only a few hours a day (4-5 hours), which means they spend more time at home, potentially reducing mothers' availability for paid work. In this regard, Contreras et al. (2010) offer evidence of the

[3]General results of free preschools on female LFP in the US and Canada show no impact for single mothers with younger children and positive effects on married mothers, both at the intensive and extensive margins (Gelbach, 2002; Baker et al., 2008). Similarly, smaller but significant effects were found for childcare subsidies on mothers' labor supply in countries such as Belgium, France and the Netherlands, while no effects were found for Norway (Dujardin et al., 2015; Givord and Marbot, 2015; Bettendorf et al., 2015; Havnes and Mogstad, 2011)

[4]For example, Nemitz (2015) studies the effects of a sharp increase of more than 30 percentage points in full-time schools in Germany and finds effects close to zero. Similarly Felfe et al. (2013) find a positive effect on mothers' full-time employment, but a negative effect on fathers' employment at the intensive margin in Switzerland.

effects of a FTS Program that lengthened the school day by two hours in Chilean high schools. The authors estimate an average gain on mothers' labor force participation of 11 percentage points (equivalent to 17 percent of the baseline), despite the relatively small increase in the time of instruction and that high school pupils are older (13 to 17 years old). Furthermore, Berthelon et al. (2015) offer evidence of a more permanent effect on female participation, as the probability of staying more than six months in the Chilean labor market increased by 19 percentage points. To the best of our knowledge, evidence on the Chilean FTS Program is the only analysis of the relation between "childcare" for older children and female LFP in a developing country.

By focusing on the Mexican context, we are able to contribute to the existing knowledge on the relationship between childcare for older children and mothers' labor supply in developing countries. Moreover, evaluating the effects of a major policy change on labor supply is important for a more comprehensive understanding of how welfare in a broader sense may be improved by these sorts of policies. FTS policies have the potential not only to improve children's welfare and school outcomes, but also mother's LFP and the available income at home, improving overall welfare, especially, for the most vulnerable sectors of the population.

Our main results indicate that longer school-days increase mothers' labor supply at the extensive and intensive margins, increasing mothers' labor force participation by 7 percentage points and the number of worked hours per week by 2.4. Moreover, these increases in labor supply are accompanied by a 47 percent average increase in earnings across the population of mothers as a whole, and a 63 percent increase in earnings in high poverty areas. The greater gains in high poverty communities would seem to reduce the income gap between the rich and the poor. Overall, these results suggest that previous to the introduction of the FTS program, female LFP was certainly constrained by the absence of family friendly policies, particularly childcare institutions.

The rest of the paper proceeds as follows. Section 4.2 offers information on Mexico's childcare policies, female labor force participation and the Full-Time Schools Program. Section 4.3 presents the details of the data used for the main analysis as well as some descriptive statistics. Section 4.4 explains the main methodology to identify the effects of longer school days on mothers' labor supply. Section 4.5 presents the main results. Section 4.6 concludes.

4.2 Background

4.2.1 Female Labor Force Participation and Childcare in Mexico

In recent decades, female participation rates in Mexico have substantially increased. Diverse factors have pushed women into the labor force, including demographic and cultural shifts, the opening of the Mexican economy, a rise in the levels of formal education, the implementation of structural reforms and a series of economic crises (Orraca et al., 2016).

According to information from Mexico's population censuses, the percentage of women between 18 and 65 years of age participating in the labor force grew from 19.4% in 1970, to 24.2% in 1990 and to 42.3% in 2010. However, Mexican female LFP remains as one of the lowest among Latin-American countries with similar per-capita income. In 2015, only 44% of Mexican women participated in the labor market; this is comparable to a similar proportion in Chile but it is lower than the 53% registered in Argentina and Uruguay and the 59% observed in Brazil. Furthermore, considering the female to male LFP ratio, Mexico stands next to the lowest in the whole continent with women's labor force participation standing at only 55% that of men, above only Honduras (49%) and below the Latin American average of 66% (Martínez Gómez et al., 2013).

This low women's participation in the labor market potentially relates to the absence of family oriented programs (Staab and Gerhard, 2010). Although some childcare policies have been applied in Mexico before, such as the Federal Daycare Program for Working Mothers which subsidizes community- and home-based daycare to facilitate employment of low-income mothers,[5] the country's spending on family benefits including childcare has not changed dramatically in the last decade and it is barely above 1% of the GDP. This is the worst average of the 33 countries in the OECD, including mid-income countries such as Israel (2.4%) and Chile (1.4%).[6]

Finally, Mexico's enrollment rates in preschool (children 3 to 5 years old) are relatively high (91%) and above the OECD average of 81%. Primary school (for ages 6 to 12) is practically universal. However, all preschools and primary schools, before FTS implementation, were part-time, having daily schedules of four to five hours. This plausibly discouraged mothers' full-time participation in the labor market, especially for the 88% of mothers who have no access to full-time childcare services at any given age. The FTS program therefore offers an important potential for the analysis of changes in labor supply in a context of low public investment and low female participation.

4.2.2 The Full-Time Schools Program

The FTS program started in 2007. Its aim was to improve learning opportunities in primary education by extending the school day from four-and-a-half to eight hours. The program started in 500 schools and by the 2014-2015 academic year it had reached 23,182 schools all over Mexico. This number represents approximately 25 percent of all elementary schools in Mexico. Notably, from its inception, the FTS program identified two secondary objectives of the program: to help single mothers to participate in the labor market and to support mono-parental families (SEP, 2010, p.3). In total, the FTS program represented a public spending of approximately US$460 millions from 2007 to 2013.[7]

[5]For a thorough review of this program see (Staab and Gerhard, 2010).

[6]Data on family policies and school participation and childcare presented in this section are extracted from the OECD Family database downloaded in February 2016.

[7]For further details of the FTS program see (Cabrera-Hernandez, 2015) and (Padilla-Romo, 2015).

Schools selected into the program generally have certain characteristics. The most relevant to this study are: (i) schools have minimum infrastructure requirements (e.g. space for the construction of a kitchen and computer classrooms, sports infrastructure, and basic services such as water and electricity), (ii) schools are working in one shift either in the morning or afternoon but not both (in Mexico, approximately 40% of primary schools offer two shifts), and (iii) preferentially, schools should have been located in vulnerable areas. Nonetheless, these guidelines were only a suggestion provided by the Ministry of Education, not binding requirements, and, in the end, the states were the ones in charge of choosing the schools to be treated. We will discuss in Section 4.4 how this fact might bias our estimates.

4.3 Data

Our analysis uses survey and administrative data from the National Institute of Statistics and Geography (*Instituto Nacional de Estadística y Geografía*, INEGI), the Ministry of Education, and the National Population Council (*Consejo Nacional de Población*, CONAPO) that together brings a quarterly individual-level dataset covering the period from the first quarter of 2005 (2005:Q1) to the third quarter of 2015 (2015:Q3). Our primary outcome variables are labor force participation, number of weekly hours worked, and monthly earnings of females with elementary school-age children, while our treatment variable is the share of predicted FTS seats in a municipality at a given quarter.[8]

The labor outcomes used in our analysis are based on the National Survey of Occupation and Employment (*Encuesta Nacional de Ocupación y Empleo*, ENOE) from INEGI. ENOE is a rotating panel of households, in which each household remains in the survey

[8] To avoid concerns about endogeneity, we define the predicted number of seats in full-time schools as the average school enrollment before the FTS program began (2001-2006). Ideally, if we were able to observe schools' capacity, our treatment variable would be the share of FTS seats at a given year. However, we only observe school enrollment which, particularly in full-time schools, may be correlated with mothers' propensity to participate in the labor market.

for five consecutive quarters. That is, we observe whether household members change labor force participation status, number of weekly hours worked, or monthly earnings over five consecutive periods.[9] In addition, ENOE contains information on sociodemographic characteristics of the individuals, as well as the location of the household. These allow us to control for time-varying individual characteristics, and to match each individual with the share of FTS seats in the municipality every quarter. We also use the location information to match each mother to the poverty index of her locality of residence.[10] This allows us to consider heterogeneous effects of the extension of the school day on mothers' labor supply that reside in high- and low- poverty areas. The poverty index is estimated by CONAPO as a measure of social exclusion in the locality using information from the Census of Population and Housing on education, housing characteristics, population, and income.[11] To avoid concerns about endogeneity, we use the poverty index of the localities in 2005, which is two years prior to the extension of the school day.

The treatment variable is constructed using annual school-level census data on enrollment and participation in the FTS program from the Ministry of Education. Information on enrollment is based on *Estadísticas 911* from the Ministry of Education. To transform this information from academic years to quarters, we take the last and the first three quarters of the year. For example, the fraction of seats in full-time schools during the 2007-2008 academic year affects labor outcomes on 2007:Q4, 2008:Q1, 2008:Q2, and 2008:Q3.

Our main analysis focuses on mothers who are the household's head and on wives

[9] In ENOE, labor force participation is defined as people over 15 years old that had a job or were looking for one during the week the survey was conducted, the number of weekly hours worked is defined as the average number of hours worked by an individual in the week of the survey, and the monthly earnings is defined as the income that the employed population received for the job they held in the week of the survey.

[10] The term locality in Mexico refers to the smallest of the three levels of division (locality, municipality, and state) of the national geostatistical framework. It is a generic territorial division for a population center with its own identity. It can be small in size and population (country, or village) or large and highly populated (city). INEGI keeps control of the list of localities in Mexico.

[11] Mothers living in high- and low- poverty areas are defined as those living in a locality with poverty index above and below the sample median.

or partners over 15 years old whose children are studying elementary education, because for this group we can unambiguously match mothers to their children. Moreover, it is plausible to think that mothers are the group of females who are more affected by the policy.[12]

4.4 Identification Strategy

We estimate the effects of extending the school day on female labor outcomes using a difference-in-differences research design that uses within-individual variation in access to full-time schools, which is defined as the share of predicted FTS seats in a municipality. The logic behind this approach is that mothers living in municipalities with a high share of predicted FTS seats are in a position to benefit from the extended school day, increasing their labor supply, while females in municipalities with a low share are not. Therefore, we compare changes in labor outcomes of females with school-age children in municipalities with full-time schools to the change observed in municipalities not affected by the policy extending the school day.

Because the fraction of predicted FTS seats changes only once within the range of the data for each individual, we use only the variation from the first and fifth periods with a long-difference regression model. This specification allows us to estimate the longer-run effects of extending the school day. Our main results are based on the following model,

$$\Delta_4 Y_{imt} = \Delta_4 FTS_{mt}\delta + \gamma_t + \Delta_4 X_{imt}\beta + \Delta_4 u_{imt} \tag{4.1}$$

where Y_{imt} denotes either an indicator variable reflecting whether individual i in municipality m participated in the labor force at quarter t, the number of weekly hours worked, or the log of monthly earnings of individual i in municipality m at quarter t; FTS_{mt} is the fraction of predicted FTS seats in municipality m at quarter t; X_{imt} in-

[12]Note that, by using this approach we cannot identify the effect for females with elementary school age children that live in extended households.

clude time-varying individual controls including years of schooling, age, and age of the youngest child; γ_t are year-by-quarter fixed effects; u_{imt} is an error term; and Δ_4 denotes the 4-period difference operator (e.g., $\Delta_4 FTS_{mt} = FTS_{mt} - FTS_{mt-4}$). This long-difference regression equation allows us to control for individual specific observed and unobserved characteristics that are constant over time, as well as, nationwide time-varying shocks to mothers' labor outcomes common to all municipalities. The coefficient of interest (δ) can be interpreted as the cumulative effect of the FTS program on the change in labor outcomes over the 5-quarter period that each individual is observed, instead of the average effect, as in fixed effects models.

Additionally, in some specifications we control for state-by-year-by-quarter fixed effects where we identify δ by comparing changes in labor outcomes in municipalities with a high fraction of predicted FTS seats to the change observed in the remaining municipalities in the same state. Robust standard errors are clustered at the municipality-level to account for potential error correlations within municipalities.

The identifying assumption underlying our research design is that in the absence of the extension of the school day, changes in mothers' labor supply in municipalities with a high fraction of predicted FTS seats would have been similar to those in municipalities with a lower fraction. Even though we cannot prove that this assumption holds, we can argue that it is plausible in our setting. First, we are able to provide graphical evidence that mothers' labor outcomes in treatment and control municipalities do not diverge *prior* to the adoption of the FTS program. Second, we formally test for divergence by including lead terms of the change in fraction of predicted FTS seats a year and two years prior to treatment to Equation 4.1. Finally, we provide evidence showing that the time-varying factors that affect female labor outcomes are orthogonal to the within-municipality variation in the fraction of predicted FTS seats. In this case δ would provide the causal effects of extending the school day on female labor supply.

Another concern about the validity of the estimated effects is that females with school age children could select into or out of the municipality in quarters that increased the fraction of predicted FTS seats. For example if females with school-age children that are more likely to participate in the labor force move to municipalities with full-time schools, we would overestimate the effects of extending the school day. We address this potential selection bias problem by estimating the degree to which the change in the fraction of predicted FTS seats affects the probability of mothers staying in their municipality of residence during the first and fifth survey quarters.

4.5 Results

We begin our analysis by providing graphical evidence on the effects of extending the school day and on the identifying assumption underlying our research design. Figure A.12 panels (a) to (c) respectively show the state-by-year-by-quarter adjusted average mothers' LFP, number of weekly hours worked, and log of monthly earnings over time for municipalities with a fraction of predicted FTS seats that is in the top quartile (high-intensity of treatment) relative to those in the bottom quartile (low-intensity of treatment). While it is not easy to appreciate the size of the effects, the three panels show an increase on LFP, number of weekly hours worked, and monthly earnings for females with elementary-school-age-children in municipalities with a high-intensity of treatment (relative to municipalities with low-intensity of treatment). Furthermore, mothers' LFP, number of weekly hours worked and monthly earnings for municipalities with a high and low intensity of treatment have similar trends *prior* to the introduction of the FTS program, providing support in favor of the identifying assumption needed for the difference-in-differences estimates to be valid.

Table A.13 shows the estimated effects of extending the school day on female labor outcomes based on the long-difference model represented by Equation 4.1. Panel A shows the estimated effects on mothers' LFP, Panel B the number of weekly hours worked, and Panel C the log of monthly earnings. Particularly, estimates in Column 1 show the baseline model represented by Equation 4.1. In Column 2, we additionally control for state-by-year-by-quarter fixed effects. In Column 3, we include time-varying individual controls. Finally, in columns 4 and 5, we test for divergence prior to treatment by including the 4-period difference in fraction of predicted FTS seats one year and two years *prior* to treatment.

The long-difference estimates show the cumulative effects of going from none to all schools being full-time, which results in increases of mothers' LFP of 7 percentage points, number of weekly hours worked of 2.4, and monthly earnings of 47 percent. In addition, the coefficients for one year and two years prior to treatment are not significant and small in magnitude, providing support for our identification strategy.

It is important to note that mothers were linked to the fraction of predicted FTS seats based on their municipality of residence. Given this approach, non-random sample attrition could be a threat to identification if most of the mothers leaving the sample lived in municipalities that increase (decrease) the intensity of treatment.[13] In consideration of this potential attrition problem, we examine whether the change in the fraction of predicted FTS seats in a given quarter affects the likelihood that the mother will be in the sample during the first and last periods.

Table A.14 shows the estimated effects of the share of predicted FTS seats on an indicator variable of whether or not the mother is in the sample during the first and fifth interviews. In Column 1, we present the baseline model represented by Equation 4.1; in Column 2, we additionally control for state-by-year-by-quarter fixed effects. The long-differences estimates indicate that changes in the share of predicted seats in full-time

[13]Cano-Urbina (2016) highlights the attrition problem in the ENOE for the period 2005–2012; he finds that 84.19 percent of the individuals who started the sample are still in it during the fifth interview.

schools do not affect the probability of leaving the sample in the fifth interview. These results suggest that attrition is independent of changes in exposure to full-time schools.

4.5.1 Treatment Heterogeneity

We now explore the extent to which there are heterogeneous effects of extending the school day on mothers' labor supply. In particular, we consider heterogeneous effects by education, and by poverty level of the locality of residence, using our preferred identification strategy.

4.5.1.1 Education

Motivated by the fact that education strengthens the connection of mothers to the labor force by increasing their potential earnings, or by reducing the range for specialization within the household (Eckstein and Lifshitz, 2011), in columns 2 and 3 of Table A.15, we report separate estimates for mothers with levels of education below (0-9 years) and above (10 or more years) the sample median.[14] The estimates in columns 2 and 3 indicate that the effects of extending the school day are concentrated among low educated mothers, who are less attached to the labor market.[15] Specifically, these results indicate that going from none to all schools being full-time increases LFP by 7.8 percentage points, number of weekly hours worked by 2.5 hours, and monthly earnings by 48 percent for low educated mothers. We find no evidence of significant effects on labor outcomes for highly educated mothers.

4.5.1.2 Poverty Level

Now we estimate the effects of extending the school day on labor supply separately for mothers with residence in localities with poverty levels below (low poverty) and above

[14]Nine years of schooling translates into completed junior high school, which is also the sample median and the compulsory level of education in Mexico.

[15]In our sample, LFP for mothers with education levels between zero and nine years is 42.4 percent, compared to 63.7 percent for mothers with ten or more years of schooling.

(high poverty) the sample median. This analysis is motivated by the fact that preference in FTS funding was given to schools located in vulnerable areas. Therefore, we expect our long-difference estimated effects to be mostly driven by mothers with residence in high poverty localities, as they are more likely to live closer to full-time schools.

Table A.16 shows the estimated effects by poverty level of the locality of residence. The estimated coefficients indicate that increases in labor supply are mostly driven by mothers residing in high poverty communities, increasing LFP by 8.6 percentage points, weekly hours worked by 3.27, and monthly earnings by 63 percent in high poverty areas. We find no effects on mothers residing in low poverty localities. These results support the notion that mothers living in vulnerable areas are the most likely to be affected. Moreover, the greater gains in high poverty communities would seem to reduce the income gap between the rich and the poor.

4.5.2 Robustness Checks

In an effort to show that the main results are not driven by a simultaneous increase in the propensity to participate in the labor force due to the labor market characteristics in those municipalities with a higher share of predicted FTS seats, we explore the degree to which extending the school day differently affects labor supply of women and men with and without school age children. If we find effects on subgroups of the population that should not be affected by the policy, our results might not be valid. It is important to note that it is possible that labor outcomes for these groups might be affected by the extension of the school day; however, we argue that such effects must be second-order.

Table A.17 shows the long-difference estimated effects on labor outcomes for women and men. In particular, in Column 2 we show the effects for women without school age children; for this group of women, we find positive but smaller effects than for women with school age children (Column 1) on LFP, suggesting a spillover effect possibly driven by other family members taking care of the children after school. However, for this same group we find no significant effects on number of weekly hours worked or on labor income. In columns 3 and 4, we show the estimated effects for men with and without school age children. For both groups estimates for LFP, number of weekly hours worked, and monthly earnings are close to zero and statistically insignificant.

4.6 Conclusion

This paper examines whether the FTS program in Mexico, which substantially increased the length of the school day, increased mothers' labor supply. More broadly, it asked whether childrearing hinders women's participation in the labor market. We exploit the variation in the staggered implementation of the FTS program and the intensity of treatment across municipalities to measure the effects of the extension of the school day on mothers' labor supply at the intensive and extensive margins. Using survey and administrative data, we estimate long-difference models which exploit differences in mothers' exposures to the FTS program. Our main results document positive and statistically significant effects on mothers' labor force participation, number or weekly hours worked, and monthly earnings. Mothers with fewer years of schooling and mothers living in high poverty localities showed the strongest labor market response to the availability of full time schools for their children. This evidence suggests that longer school days can be an effective policy to increase mothers' labor market participation while their children are still growing up and that the greater gains for low income mothers would seem to reduce the income gap between the rich and the poor.

5. SUMMARY AND CONCLUSIONS

The three essays in this dissertation have looked at the causal effects of different interventions on crime, education, and labor outcomes using quasi-experimental research designs. Particularly, how these interventions affect differently the most vulnerable sectors of the population. In Chapter 2, we find that kingpin captures have large and sustained effects on the homicide rate in the municipality of capture and smaller but significant effects on other municipalities where the kingpin's DTO has a presence, supporting the notion that the kingpin strategy can have destabilizing effects throughout an organization while highlighting that this does not imply a reduction in violence. In Chapter 3, we find that extending the school day increases both reading and math test scores. Finally, in Chapter 4, we document that longer school days have positive and statistically significant effects on mothers' labor force participation, number or weekly hours worked, and monthly earnings.

All analyses use difference-in-differences research designs which allow us to disentangle the causal effects of the interventions from the effects of other possible confounding factors. The validity of the results rely on the assumptions that in the absence of the interventions, changes in outcomes in the treatment and the control groups would have been the same and that there are not other simultaneous interventions. The rich data used for the analyses allow us to show that these assumptions seem plausible in our settings. Then the effects that we find can be interpreted as causal and have many policy implications. First, while our estimates indicate that Mexico's use of the kingpin strategy caused significant increases in homicides, it is important to note that its war on drugs had several objectives beyond reducing violence, including the establishing the rule of law, that need to be considered in evaluating the policy. Second, as developing countries expand their capacity to deliver education, extending the school day can be an effective way to improve the quality of education they offer and it can be effective to close the education gap between the rich and the poor; a broader implication of these results is that early assessments of educational interventions may be poor predictors of longer run effects. Finally, longer school days can be an effective policy to increase mothers' labor supply while their children are still growing up and the greater gains for low income mothers would seem to reduce the income gap between the rich and the poor.

REFERENCES

Agüero, J. M. and Beleche, T. (2013). Test-Mex: Estimating the effects of school year length on student performance in Mexico. *Journal of Development Economics*, 103:353–361.

Anderson, D. M. (2010). Does information matter? The effect of the meth project on meth use among youths. *Journal of Health Economics*, 29:732–742.

Anderson, D. M. and Elsea, D. (2015). The meth project and teen meth use: New estimates from the national and state youth risk behavior surveys. *Health Economics*, 24:1644–1650.

Ángeles, G., Gadsden, P., Galiani, S., Gertler, P., Herrera, A., Kariger, P., and Seira, E. (2011). Evaluación de impacto del programa estancias infantiles para apoyar a madres trabajadoras. *Informe final de la evaluación de impacto. CIEE e INSP.*

Angrist, J. D. and Kugler, A. D. (2008). Rural windfall or a new resource curse? Coca, income, and civil conflict in colombia. *Review of Economics and Statistics*, 90(2):191–215.

Angrist, J. D. and Lavy, V. (1999). Using Maimonides rule to estimate the effect of class size on scholastic achievement. *The Quarterly Journal of Economics*, 114(2):533–575.

Astorga, L. and Shirk, D. A. (2010). Drug trafficking organizations and counter-drug strategies in the US-Mexican context. *Center for US-Mexican Studies*.

Baker, M., Gruber, J., and Milligan, K. (2008). Universal child care, maternal labor supply, and family well-being. *Journal of Political Economy*, 116(4):709–745.

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

Berlinski, S. and Galiani, S. (2007). The effect of a large expansion of pre-primary school facilities on preschool attendance and maternal employment. *Labour Economics*, 14(3):665–680.

Bernal, R. and Keane, M. P. (2011). Child care choices and children's cognitive achievement: The case of single mothers. *Journal of Labor Economics*, 29(3):459–512.

Berthelon, M., Kruger, D. I., and Oyarzun, M. A. (2015). The effects of longer school days on mothers' labor force participation. *Working Paper*.

Bettendorf, L. J., Jongen, E. L., and Muller, P. (2015). Childcare subsidies and labour supply—Evidence from a large Dutch reform. *Labour Economics*, 36:112–123.

Blau, F. D. and Kahn, L. M. (2013). Female labor supply: Why is the United States falling behind? *The American Economic Review*, 103(3):251–256.

Bondurant, Samuel R., J. M. L. and Swensen, I. D. (2016). Substance abuse treatment centers and local crime. *NBER Working Paper 22610*.

Brown, R. (2016). The mexican drug war and early-life health: The impact of violent crime on birth outcomes. *Working Paper*.

Brown, R. and Velasquez, A. (2016). The effect of violent crime on the human capital accumulation of young adults. *Working Paper*.

Cabrera-Hernandez, F. (2015). Does lengthening the school day increase students' academic achievement? Evidence from a natural experiment. Technical report, University of Sussex, No. 74-2015.

Calderon, G., Robles, G., Diaz-Cayeros, A., and Magaloni, B. (2015). The beheading of criminal organizations and the dynamics of violence in Mexico. *Journal of Conflict Resolution*, 59(8):1455–1485.

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2011). Robust inference with multiway clustering. *Journal of Business and Economic Statistics*, 29(2):238–249.

Cano-Urbina, J. (2016). Informal labor markets and on-the-job training: Evidence from wage data. *Economic Inquiry*, 54(1):25–43.

Cascio, E. U., Haider, S. J., and Nielsen, H. S. (2015). The effectiveness of policies that promote labor force participation of women with children: A collection of national studies. *Labour Economics*, 36:64–71.

Castillo, J. C., Mejía, D., and Restrepo, P. (2012). Illegal drug markets and violence in Mexico: The causes beyond Calderón. *Working Paper*.

Cerdan-Infantes, P. and Vermeersch, C. (2007). More time is better: An evaluation of the full time school program in Uruguay. *World Bank Policy Research Working Paper*, 1(4167).

Chetty, R., Friedman, J. N., and Rockoff, J. E. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *The American Economic Review*, 104(9):2633–2679.

Chicoine, L. (2011). Exporting the second amendment: U.S. Assault weapons and the homicide rate in Mexico. *Working Paper*.

CIA (2010). The world factbook. *Central Intelligence Agency*.

Clemens, J. (2008). Opium in Afghanistan: Prospects for the success of source country drug control policies. *Journal of Law and Economics*, 51(3):pp. 407–432.

Clemens, J. (2013a). An analysis of economic warfare. *American Economic Review*, 103(3):523–27.

Clemens, J. (2013b). Evaluating economic warfare: Lessons from efforts to suppress the Afghan opium trade. *Working Paper*.

CONEVAL (2013). Informe de la evaluación específica de desempeño 2012 - 2013. Programa Escuelas de Tiempo Completo.

CONEVAL (2015). Informe de la evaluación específica de desempeño 2014 - 2015. Programa Escuelas de Tiempo Completo.

Contreras, D., Sepúlveda, P., and Cabrera, S. (2010). The effects of lengthening the school day on female labor supply: Evidence from a quasi-experiment in Chile. *Serie Documentos de Trabajo*, 323.

Coscia, M. and Rios, V. (2012). Knowing where and how criminal organizations operate using web content. *CIKM*, pages 1412–1421.

Cunningham, J. K. and Liu, L. M. (2003). Impacts of federal ephedrine and pseudoephedrine regulations on methamphetamine-related hospital admissions. *Addiction*, 98(9):1229–1237.

Cunningham, S. and Finlay, K. (2013). Parental substance use and foster care: Evidence from two methamphetamine supply shocks. *Economic Inquiry*, 51(1):764–782.

Dell, M. (2015). Trafficking networks and the mexican drug war. *American Economic Review*, 105(6):1738–1779.

Dobkin, C. and Nicosia, N. (2009). The war on drugs: Methamphetamine, public health, and crime. *American Economic Review*, 99(1):324–349.

Dobkin, C., Nicosia, N., and Weinberg, M. (2014). Are supply-side drug control efforts effective? Evaluating OTC regulations targeting methamphetamine precursors. *Journal of Public Economics*, 120:46–61.

DOF (2012). Reglas de operación del Programa Escuelas de Tiempo Completo. Acuerdo 610. Diario Oficial de la Federación.

Dube, A., Dube, O., and García-Ponce, O. (2013). Cross-border spillover: U.S. Gun laws and violence in Mexico. *American Political Science Review*, 107(3):397–417.

Dujardin, C., Fonder, M., and Lejeune, B. (2015). Does formal child care availability for 0-3 year olds boost mothers' employment rate? Panel data based evidence from Belgium. *Working Paper*.

Eckstein, Z. and Lifshitz, O. (2011). Dynamic female labor supply. *Econometrica*, 79(6):1675–1726.

Eisenkopf, G., Hessami, Z., Fischbacher, U., and Ursprung, H. W. (2015). Academic performance and single-sex schooling: Evidence from a natural experiment in Switzerland. *Journal of Economic Behavior & Organization*, 115:123–143.

Escalante, F. (2011). Homicidios 2008-2009 la muerte tiene permiso. *Nexos*.

Fairlie, R. W. (2002). Drug dealing and legitimate self-employment. *Journal of Labor Economics*, 20(3):538–537.

Felfe, C., Lechner, M., and Thiemann, P. (2013). After-school care and parents' labor supply. *Working Paper*.

Gelbach, J. B. (2002). Public schooling for young children and maternal labor supply. *The American Economic Review*, 92(1):307–322.

Givord, P. and Marbot, C. (2015). Does the cost of child care affect female labor market participation? An evaluation of a French reform of childcare subsidies. *Labour Economics*, 36:99–111.

Grayson, G. W. (2013). *The Cartels: The Story of Mexico's Most Dangerous Criminal Organizations and Their Impact on US Security*. ABC-CLIO, Santa Barbara, CA.

Havnes, T. and Mogstad, M. (2011). Money for nothing? Universal child care and maternal employment. *Journal of Public Economics*, 95(11):1455–1465.

Heckman, J. J. (2008). Schools, skills, and synapses. *Economic Inquiry*, 46(3):289–324.

Hincapie, D. (2014). Do longer school days improve student achievement? Evidence from Colombia. Technical report, The George Washington University.

Holland, P. A., Alfaro, P., and Evans, D. (2015). Extending the school day in Latin America and the Caribbean. *World Bank Policy Research Working Paper*, 1(7309).

Hoxby, C. M. (2000). The effects of class size on student achievement: New evidence from population variation. *Quarterly Journal of Economics*, 115(4):1239–1285.

Jaumotte, F. (2003). Female labour force participation: past trends and main determinants in OECD countries. *OECD Working Paper*, 1(376).

Krueger, A. B. (1999). Experimental estimates of education production functions. *Quar-*

terly Journal of Economics, 114:497–532.

Kruger, D. I. and Berthelon, M. (2009). Delaying the bell: The effects of longer school days on adolescent motherhood in Chile. *Working Paper*.

Lavy, V. and Schlosser, A. (2011). Mechanisms and impacts of gender peer effects at school. *American Economic Journal: Applied Economics*, 3(2):1–33.

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *The Review of Economic Studies*, 76(3):1071–1102.

Lee, S., Turner, L. J., Woo, S., and Kim, K. (2014). All or nothing? The impact of school and classroom gender composition on effort and academic achievement. Technical report, National Bureau of Economic Research.

Martínez Gómez, C., Miller, T., and Saad, P. (2013). Participación laboral femenina y bono de género en América Latina. *CEPAL Working Paper*.

McKinnish, T. G. (2000). Model sensitivity in panel data analysis: Some caveats about the interpretation of fixed effects and differences estimators. *Terra*, 303:492–6770.

Mejía, D. and Restrepo, P. (2013). Bushes and bullets: Illegal cocaine markets and violence in colombia. *Documento CEDE*.

Merino, J. (2011). Los operativos conjuntos y la tasa de homicidios: Una medición. *Nexos*.

Miron, J. and Zwiebel, J. (1995). The economic case against drug prohibition. *Journal of Economic Perspectives*, 9(4):175–192.

Miron, J. A. (1999). Violence and the U.S. prohibitions of drug and alcohol. *American Law and Economics Review*, 1(1):78–114.

Montoya, E. (2016). Violence and economic disruption: Firm-level evidence from Mexico. *Working Paper*.

Morales-Oyarvide, C. (2011). La guerra contra el narcotráfico en méxico. Debilidad del Estado, orden local y fracaso de una estrategia. *Aposta*, 50.

Nemitz, J. (2015). The effect of all-day primary school programs on maternal labor supply.

Working Paper.

Orraca, P., Cabrera, F.-J., and Iriarte, G. (2016). The gender wage gap and occupational segregation in the Mexican labour market. *EconoQuantum*, 13(1):51–72.

Owens, E. G. (2011). Are underground markets really more violent? Evidence from early 20th Century America. *American Law and Economics Review*, 13:1–44.

Owens, E. G. (2014). The american temperance movement and market-based violence. *American Law and Economics Review*, 16:433–472.

Padilla-Romo, M. (2015). The short and long run effects of full-time schools on academic performance. *Working Paper*.

Patall, E. A., Cooper, H., and Allen, A. B. (2010). Extending the school day or school year a systematic review of research (1985–2009). *Review of Educational Research*, 80(3):401–436.

Phillips, B. J. (2015). How does leadership decapitation affect violence? The case of drug trafficking organizations in Mexico. *The Journal of Politics*, 77(2):pp. 324–336.

Pires, T. and Urzua, S. (2010). Longer school days, better outcomes? *Working Paper*.

Rivkin, S. G., Hanushek, E. A., and Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, 73(2):417–458.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review*, 94:247–252.

Santibañez, L. (2006). Why we should care if teachers get A's: Teacher test scores and student achievement in Mexico. *Economics of Education Review*, 25(5):510–520.

Schady, N., Expósito, A. P., Bóo, F. L., Kagan, S. L., Jalmovich, A., Hincapie, D., Flabbi, L., Cruz-Aguayo, Y., Cristia, J., and Berlinski, S. (2015). The early years: Child well-being and the role of public policy. *Inter-American Development Bank Publications (Books)*.

Schlosser, A. (2005). Public preschool and the labor supply of arab mothers: Evidence from a natural experiment. *Manuscript, The Hebrew University of Jerusalem*.

SEP (2007). Programa Nacional Escuela Segura, México: Secretaría de Educación Pública. *Secretaria de Educación Pública*.

SEP (2010). Orientaciones pedagógicas para la organización del trabajo en las escuelas de tiempo completo. Technical report, Dirección General de Desarrollo Curricular. Subsecretaría de Educación Básica.

Skoufias, E. and Shapiro, J. (2006). Evaluating the impact of Mexico's Quality Schools Program: The pitfalls of using nonexperimental data. *World Bank Policy Research Working Paper Series, Vol. 8*.

Staab, S. and Gerhard, R. (2010). Childcare service expansion in Chile and Mexico. *Working Paper*.

Stevens, A. (1997). Persistent effects of job displacement: The importance of multiple job losses. *Journal of Labor Economics*, 15(1):165–188.

Stewart, S. and Posey, A. (2009). Mexico: The war with the cartels in 2009. *Stratfor Global*.

Swensen, I. D. (2015). Substance-abuse treatment and mortality. *Journal of Public Economics*, 122:13–30.

Urquiola, M. (2006). Identifying class size effects in developing countries: Evidence from rural Bolivia. *Review of Economics and Statistics*, 88(1):171–177.

Velasquez, A. (2015). The economic burden of crime: Evidence from Mexico. *Working Paper*.

Vilalta, C. J. and Martinez, J. M. (2012). The making of narco bosses: Hard drug dealing crimes among Mexican students. *Trends in Organized Crime*, 15(1):47–63.

Weatherburn, D., Jones, C., Freeman, K., and Makkai., T. (2003). Supply control and harm reduction: Lessons from the Australian heroin 'drought'. *Addiction*, 98(1):83–91.

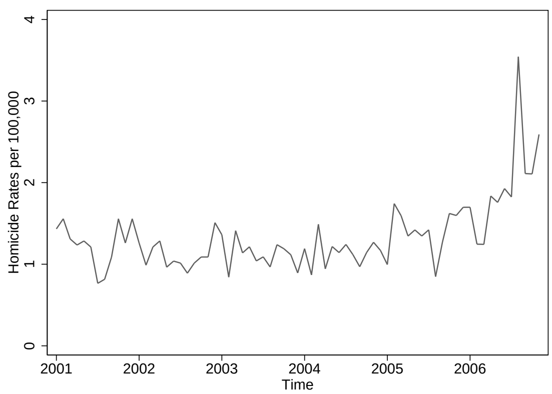
APPENDIX A

FIGURES AND TABLES

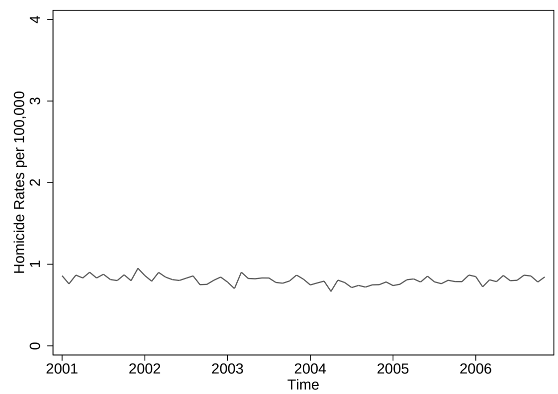
A.1 Kingpin Strategy

Figure A.1: Monthly Homicide Rates Prior the Beginning of the War on Drugs

(a) Michoacán

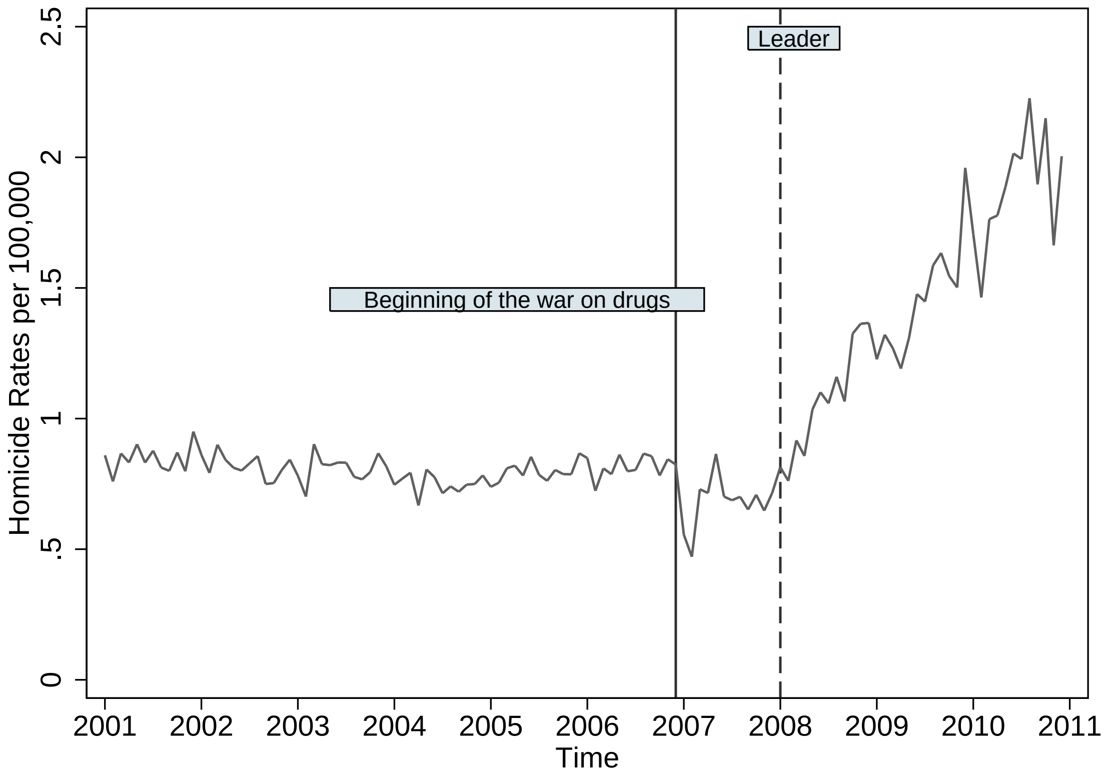


(b) National



Notes: Panel A plots the homicide rate in the state of Michoacán, President Felipe Calderón's home state, leading up to his declaring war on drugs. Panel B plots the nationwide homicide rate over the same time period. These homicide rates are calculated based on the universe of death certificates from the vital statistics of the National Institute of Statistics and Geography (INEGI) and population counts from the National Council of Population (CONAPO) and El Colegio de México (COLMEX).

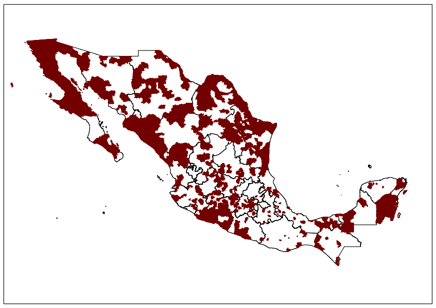
Figure A.2: National Homicide Rate



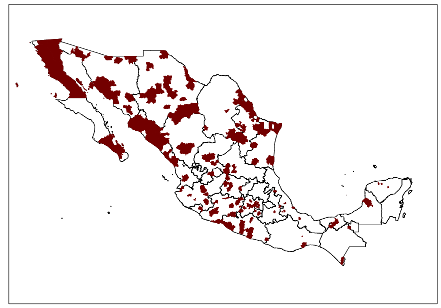
Notes: See Figure A.1. Vertical lines are drawn to highlight the beginning of the war on drugs and the first capture of a DTO leader during the war on drugs.

Figure A.3: Municipalities with DTO Presence, 2004-2006

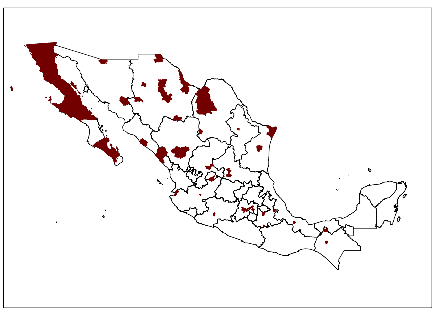
(a) Any DTO



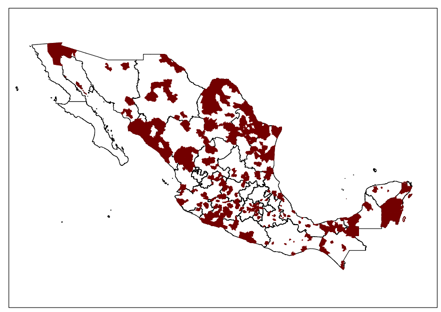
(b) Sinaloa-Beltrán-Leyva DTO



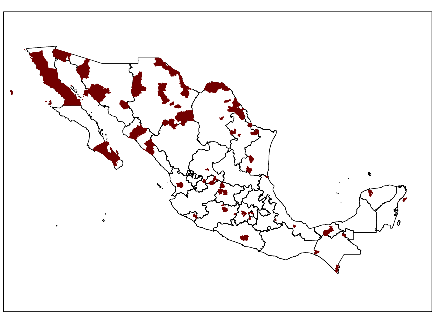
(c) Tijuana DTO



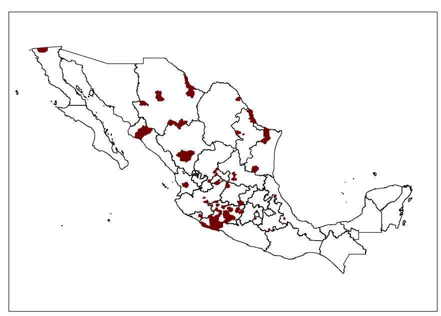
(d) Gulf DTO



(e) Juárez DTO

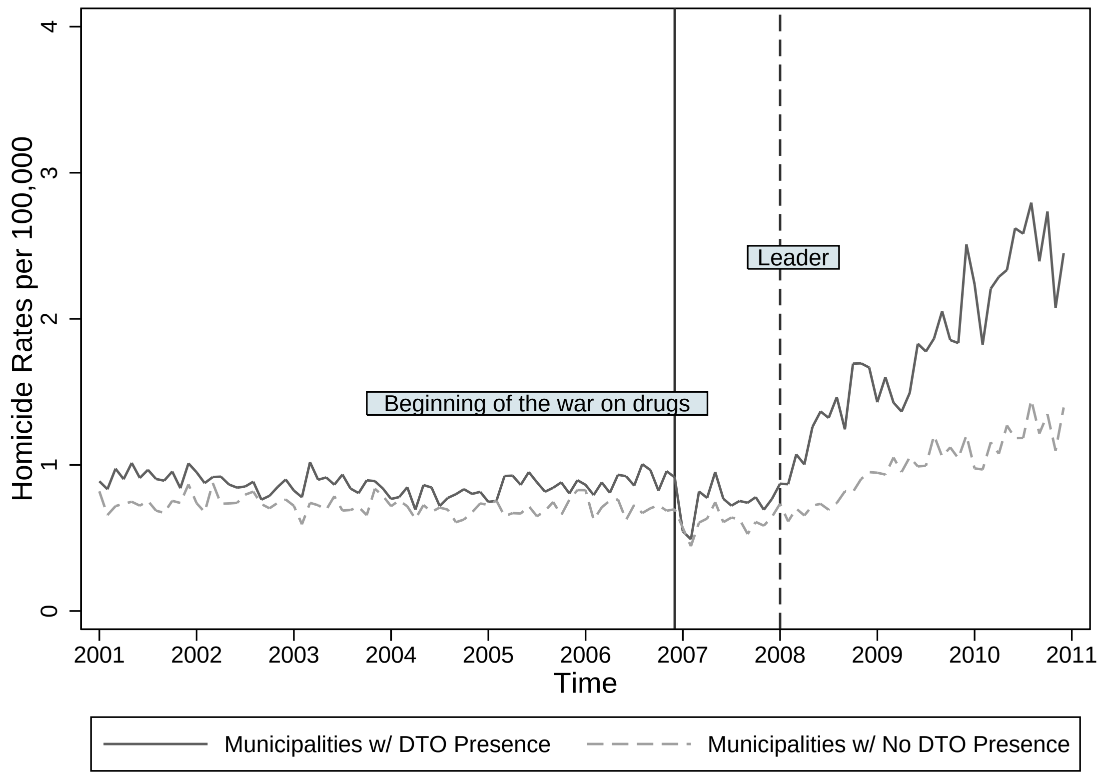


(f) La Familia DTO



Notes: Each panel shows the municipalities with the specified DTO presence prior to the war on drugs. The areas of operation for each DTO are based on Coscia and Rios (2012).

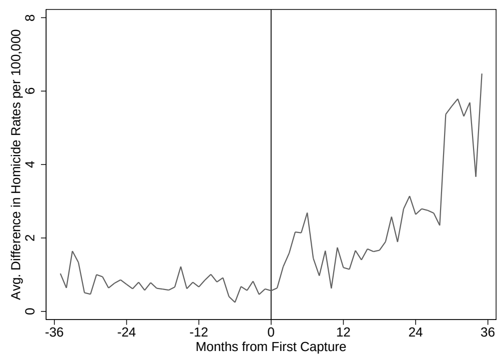
Figure A.4: Homicide Rates for Municipalities With and Without a DTO Presence



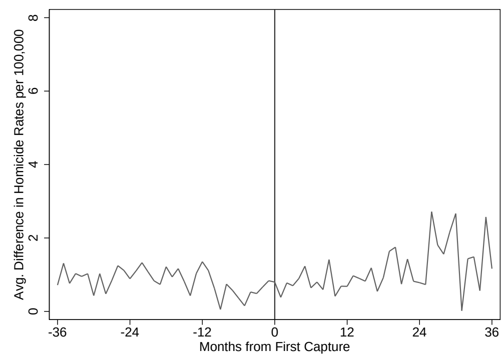
Notes: Municipalities with and without a DTO presence prior to the war on drugs are shown in Figure A.3. Vertical lines are drawn to highlight the beginning of the war on drugs and the first capture of a DTO leader during the war on drugs. Homicide rates are calculated based on the universe of death certificates from the vital statistics of the National Institute of Statistics and Geography (INEGI) and population counts from the National Council of Population (CONAPO) and El Colegio de México (COLMEX).

Figure A.5: Homicide Rates in Municipalities of Interest
Relative to Others in the Same State without a DTO Presence

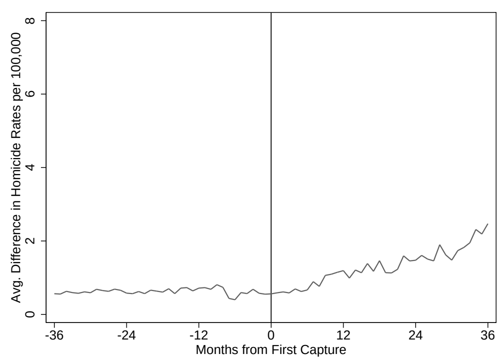
(a) Municipalities of capture



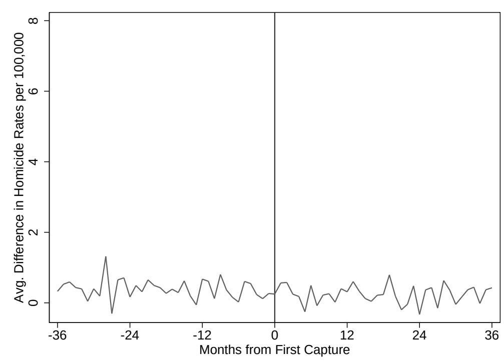
(b) Neighboring municipalities where captured leader's DTO has a presence



(c) Non-neighboring municipalities where captured leader's DTO has a presence



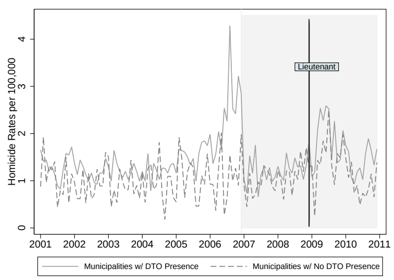
(d) Neighboring municipalities where captured leader's DTO does not have a presence



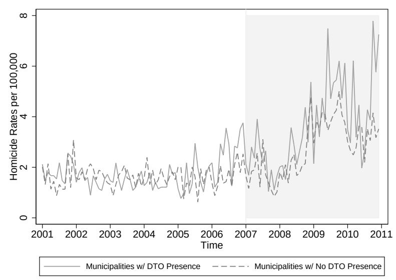
Notes: Each panel shows the average difference over time between homicide rates in the highlighted municipalities and the other municipalities in their states that have no DTO presence (and do not neighbor a municipality of capture). The time scale is adjusted to address the fact that different municipalities were affected by first captures taking place at different times—it is centered on months from such a capture. Municipalities with and without a DTO presence prior to the war on drugs are shown in Figure A.3. Homicide rates are calculated based on the universe of death certificates from the vital statistics of the National Institute of Statistics and Geography (INEGI) and population counts from the National Council of Population (CONAPO) and El Colegio de México (COLMEX).

Figure A.6: Homicide Rates for Areas Targeted in Major State-Level Operations

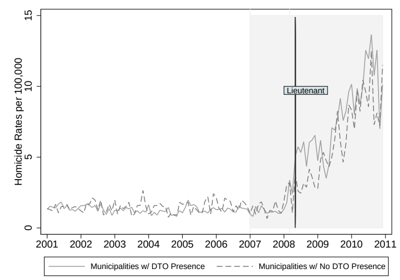
(a) Michoacán Operation



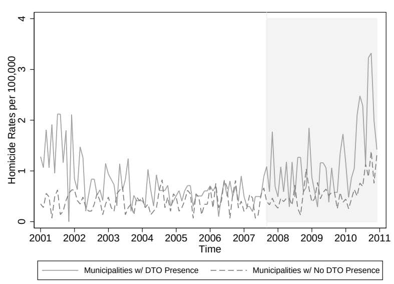
(b) Guerrero Operation



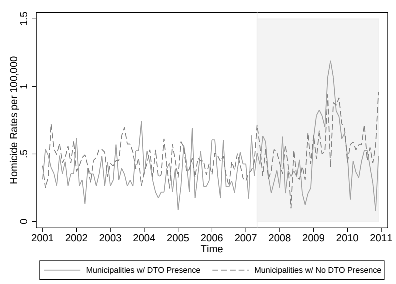
(c) Sierra Madre-Chihuahua Operation



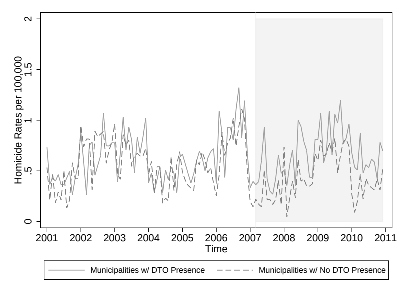
(d) San Luis Potosí Operation



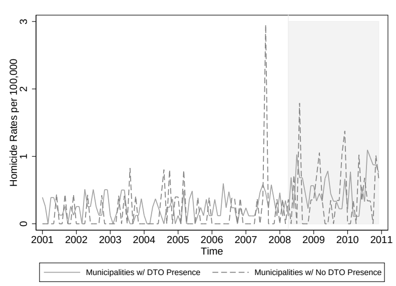
(e) Veracruz Operation



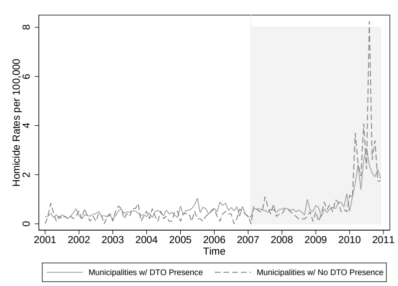
(f) Chiapas Operation



(g) Aguascalientes Operation



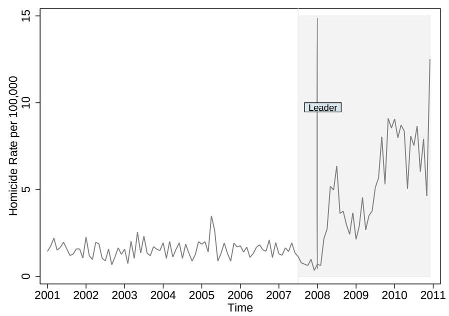
(h) Nuevo León-Tamaulipas Operation



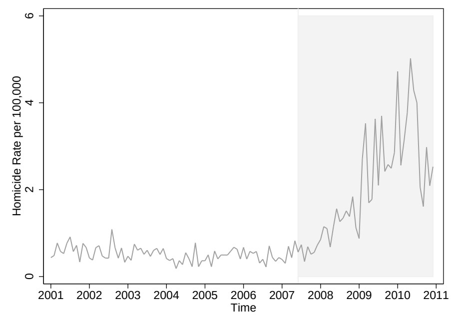
Notes: Each panel shows the homicide rates in the state(s) corresponding to the operation, with separate lines for municipalities with a DTO presence and municipalities without a DTO presence. The shaded region begins when the operation began and ends when the operation ended (where known). The Sierra Madre operation includes the states of Chihuahua, Durango, and Sinaloa. Where applicable, vertical lines show the capture of a kingpin considered in our analysis. Municipalities with and without a DTO presence prior to the war on drugs are shown in Figure A.3. Homicide rates are calculated based on the universe of death certificates from the vital statistics of the National Institute of Statistics and Geography (INEGI) and population counts from the National Council of Population (CONAPO) and El Colegio de México (COLMEX).

Figure A.7: Homicide Rates for Areas Targeted in Major Municipality-Level Operations

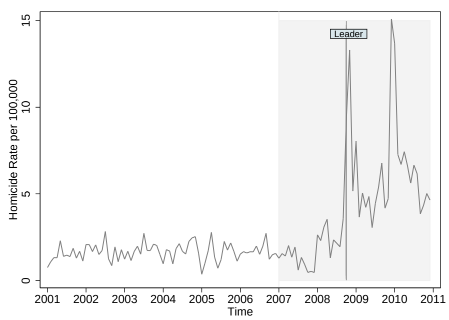
(a) Marlin-Culiacán-Navolato Operation



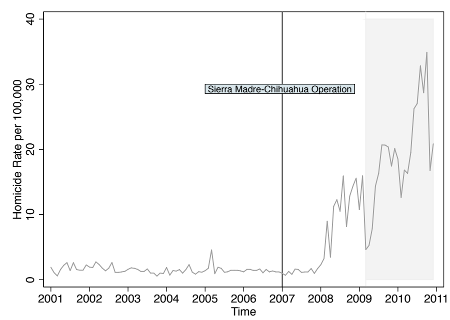
(b) Laguna Segura Operation



(c) Tijuana Operation



(d) Juárez Operation



Notes: Each panel shows the homicide rates in the municipality or municipalities corresponding to the operation. The shaded region begins when the operation began and ends when the operation ended (where known). The Marlin Operation includes the municipalities of Mazatlán and Culiacán while the Laguna Segura Operation includes the municipalities of Saltillo, Torreón, San Pedro de las Colinas, Lerdo, and Gómez Palacio. Where applicable, vertical lines show the capture of a kingpin considered in our analysis. Municipalities with and without a DTO presence prior to the war on drugs are shown in Figure A.3. Homicide rates are calculated based on the universe of death certificates from the vital statistics of the National Institute of Statistics and Geography (INEGI) and population counts from the National Council of Population (CONAPO) and El Colegio de México (COLMEX).

Table A.1: First Capture of a Kingpin For Each DTO During the War on Drugs

DTO	Name	Position	Date	Municipalities w/ DTO Presence (2004-2006)	Fraction of Population in These Municipalities
Sinaloa-Beltrán-Leyva	Alfredo Beltrán Leyva	Leader	1/21/08	166	0.36
Tijuana	Eduardo Arellano Félix	Leader	10/25/08	47	0.18
Gulf	Juan Carlos de la Cruz Reyna	Lieutenant	8/29/07	277	0.44
Juárez	Pedro Sánchez Arras	Lieutenant	5/13/08	65	0.15
La Familia	Alberto Espinoza Barrón	Lieutenant	12/29/08	68	0.09

Notes: Information of first captures is based on a compendium of press releases of the Army (SEDENA), the Navy (SEMAR), and the Office of the Attorney General (PGR). Municipalities with a DTO presence prior to the war on drugs are shown in Figure A.3. The proportion of the population is estimated based on population counts from the National Council of Population (CONAPO) and El Colegio de México (COLMEX).

Table A.2: Estimated Effects of Kingpin Captures on Homicide Rates

	(1)	(2)	(3)	(4)	(5)
Municipality of capture prior 7 to 12 months					0.012
					(0.124)
Municipality of capture prior 1 to 6 months				-0.044	-0.032
				(0.305)	(0.293)
Municipality of capture after 0 to 5 months	0.694***	0.473**	0.476**	0.469**	0.481**
	(0.248)	(0.213)	(0.226)	(0.216)	(0.230)
Municipality of capture after 6 to 11 months	0.622***	0.394***	0.392***	0.386***	0.397***
	(0.138)	(0.089)	(0.093)	(0.088)	(0.098)
Municipality of capture after 12 or more months	0.816***	0.525*	0.523*	0.516*	0.528*
	(0.300)	(0.288)	(0.302)	(0.272)	(0.308)
Neighbor w/ same DTO prior 7 to 12 months					0.004
					(0.108)
Neighbor w/ same DTO prior 1 to 6 months				-0.025	-0.019
				(0.128)	(0.133)
Neighbor w/ same DTO after 0 to 5 months	0.297*	0.097	0.101	0.098	0.105
	(0.178)	(0.119)	(0.121)	(0.115)	(0.134)
Neighbor w/ same DTO after 6 to 11 months	0.182*	-0.062	-0.066	-0.069	-0.062
	(0.103)	(0.055)	(0.073)	(0.076)	(0.085)
Neighbor w/ same DTO after 12 or more months	0.212	-0.132	-0.134	-0.138	-0.131
	(0.131)	(0.142)	(0.124)	(0.134)	(0.146)
Non-neighbor w/same DTO prior 7 to 12 months					0.051
					(0.040)
Non-neighbor w/same DTO prior 1 to 6 months				-0.014	0.021
				(0.042)	(0.042)
Non-neighbor w/same DTO after 0 to 5 months	0.008	-0.002	0.008	-0.002	0.032
	(0.025)	(0.022)	(0.024)	(0.055)	(0.050)
Non-neighbor w/same DTO after 6 to 11 months	0.066	0.046	0.040	0.030	0.064
	(0.041)	(0.031)	(0.039)	(0.078)	(0.062)
Non-neighbor w/same DTO after 12 or more months	0.175***	0.122***	0.115***	0.105	0.139**
	(0.061)	(0.040)	(0.044)	(0.083)	(0.065)
Other neighbor prior 7 to 12 months					0.031
					(0.021)
Other neighbor prior 1 to 6 months				-0.001	0.001
				(0.024)	(0.025)
Other neighbor after 0 to 5 months	0.003	-0.062*	-0.060**	-0.060*	-0.059**
	(0.028)	(0.032)	(0.030)	(0.033)	(0.027)
Other neighbor after 6 to 11 months	-0.025	-0.085**	-0.085**	-0.085**	-0.083*
	(0.028)	(0.042)	(0.042)	(0.040)	(0.044)
Other neighbor after 12 or more months	-0.025	-0.098	-0.098	-0.098	-0.096
	(0.035)	(0.088)	(0.090)	(0.089)	(0.087)
N	294480	294480	294480	294480	294480
State-by-year-by-month fixed effects	no	yes	yes	yes	yes
Additional controls	no	no	yes	yes	yes

Notes: Observations are at the municipality-month level, spanning January 2001 through December 2010. All estimates include month-by-year fixed effects and municipality fixed effects. The additional controls for columns 3–5 are indicator variables for 0–5, 6–11, and 12+ months after the beginning of the war for municipalities with DTO presence. Standard-error estimates in parentheses are two-way clustered at the state and DTO-combination levels. Homicide rates are calculated based on the universe of death certificates from the vital statistics of the National Institute of Statistics and Geography (INEGI) and population counts from the National Council of Population (CONAPO) and El Colegio de México (COLMEX). Areas of DTO operation for each DTO are based on Coscia and Rios (2012) as described in the text.
* significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.3: Estimated Effects on Other Outcomes

	(1) Homicide Male	(2) Homicide Female	(3) Domestic Violence	(4) Infant Mortality
Municipality of capture after 0 to 5 months	0.460**	0.107	0.145	0.007
	(0.229)	(0.221)	(0.368)	(0.036)
Municipality of capture after 6 to 11 months	0.365***	0.223	0.168	0.046
	(0.062)	(0.198)	(0.192)	(0.042)
Municipality of capture after 12 or more months	0.496*	0.337	0.080	-0.021
	(0.290)	(0.267)	(0.272)	(0.035)
Neighbor w/ same DTO after 0 to 5 months	0.085	-0.060	-0.050	0.071
	(0.114)	(0.119)	(0.166)	(0.044)
Neighbor w/ same DTO after 6 to 11 months	-0.096	-0.094	-0.007	0.093***
	(0.071)	(0.102)	(0.141)	(0.031)
Neighbor w/ same DTO after 12 or more months	-0.159	-0.092	0.004	0.066*
	(0.145)	(0.068)	(0.092)	(0.038)
Non-neighbor w/same DTO after 0 to 5 months	0.002	0.003	-0.025	-0.023
	(0.024)	(0.016)	(0.037)	(0.018)
Non-neighbor w/same DTO after 6 to 11 months	0.038	-0.001	-0.046	-0.022
	(0.038)	(0.022)	(0.042)	(0.021)
Non-neighbor w/same DTO after 12 or more months	0.103**	0.023	-0.077*	0.007
	(0.043)	(0.024)	(0.041)	(0.022)
Other neighbor after 0 to 5 months	-0.047*	-0.020	0.033	-0.025
	(0.026)	(0.039)	(0.051)	(0.043)
Other neighbor after 6 to 11 months	-0.071	0.005	0.064	-0.056
	(0.044)	(0.015)	(0.070)	(0.046)
Other neighbor after 12 or more months	-0.082	-0.012	0.041	-0.028
	(0.087)	(0.035)	(0.038)	(0.030)
N	294480	294480	235584	294357

Notes: See Table A.2. Additionally note that all models control for municipality fixed effects, state-by-year-by-month fixed effects, and indicator variables for 0–5, 6–11, and 12+ months after the beginning of the war for municipalities with DTO presence. Domestic violence data begin in January 2003 and are based on administrative records of individuals arrested for the crime of domestic violence from Estadśticas Judiciales en Material Penal de INEGI. Infant mortality rates are calculated based on the universe of death certificates from the vital statistics of the National Institute of Statistics and Geography (INEGI) and population counts from the National Council of Population (CONAPO) and El Colegio de México (COLMEX).
* significant at 10%; ** significant at 5%; *** significant at 1%

Table A.4: Estimated Effects on Male Homicide Rates by Age

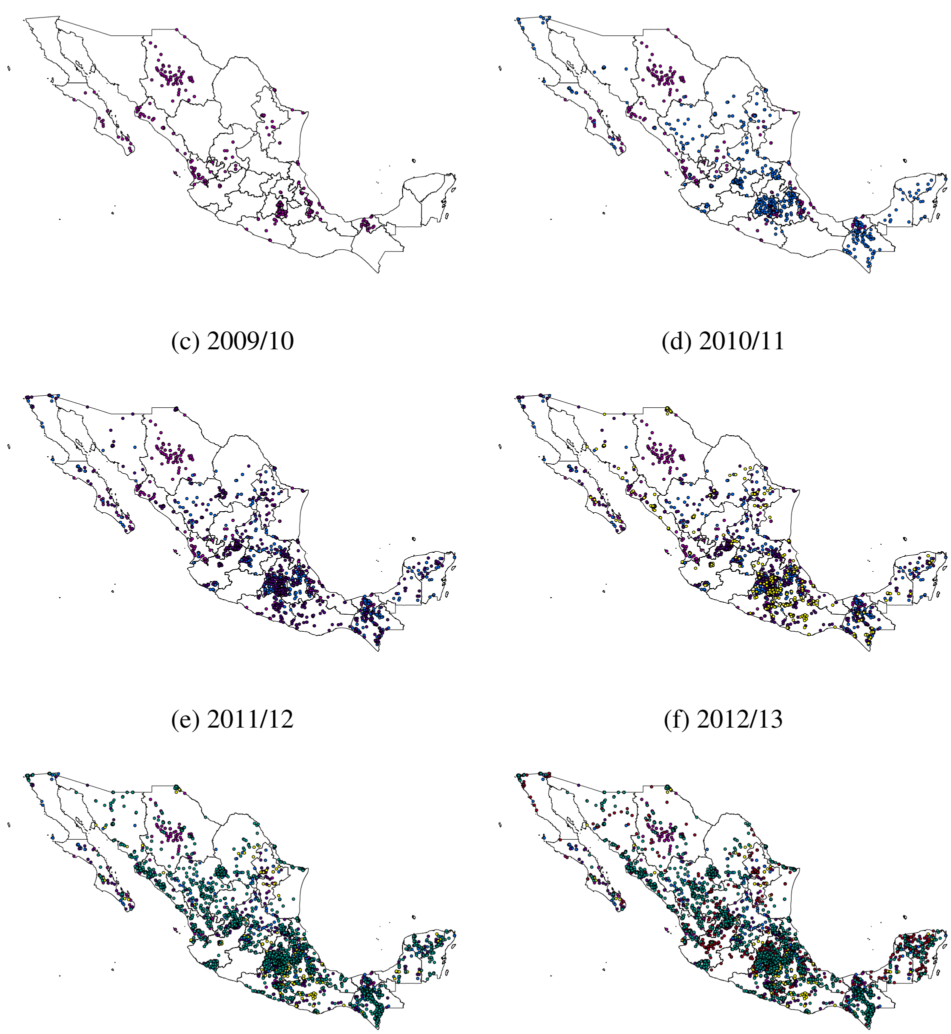
Age group:	(1) 0-14	(2) 15-29	(3) 30-44	(4) 45-59	(5) 60-74	(6) 75-89	(7) 90+
Municipality of capture after 0 to 5 months	0.085	0.474***	0.436**	0.143	0.075	0.007	-0.001
	(0.101)	(0.165)	(0.174)	(0.208)	(0.165)	(0.077)	(0.060)
Municipality of capture after 6 to 11 months	0.010	0.220	0.505***	0.196	0.076	-0.064	-0.007
	(0.052)	(0.189)	(0.131)	(0.163)	(0.057)	(0.052)	(0.058)
Municipality of capture after 12 or more months	0.008	0.527*	0.523***	0.299	0.122	-0.051	0.010
	(0.074)	(0.276)	(0.191)	(0.233)	(0.154)	(0.044)	(0.050)
Neighbor w/ same DTO after 0 to 5 months	-0.006	-0.033	0.130	-0.122	-0.060	-0.017	-0.013
	(0.044)	(0.103)	(0.082)	(0.105)	(0.054)	(0.033)	(0.022)
Neighbor w/ same DTO after 6 to 11 months	-0.046	-0.007	-0.145**	-0.137*	-0.078	-0.009	-0.008
	(0.062)	(0.083)	(0.060)	(0.080)	(0.068)	(0.032)	(0.022)
Neighbor w/ same DTO after 12 or more months	-0.009	-0.093	-0.116	-0.126	-0.064	-0.016	0.005
	(0.038)	(0.139)	(0.152)	(0.105)	(0.045)	(0.027)	(0.027)
Non-neighbor w/same DTO after 0 to 5 months	-0.007	-0.003	-0.002	0.002	-0.009	-0.006	0.009
	(0.006)	(0.026)	(0.021)	(0.025)	(0.010)	(0.007)	(0.009)
Non-neighbor w/same DTO after 6 to 11 months	-0.011	0.025	0.016	0.018	0.001	-0.007	0.011
	(0.008)	(0.036)	(0.028)	(0.032)	(0.013)	(0.009)	(0.012)
Non-neighbor w/same DTO after 12 or more months	-0.008	0.087**	0.062*	0.037	-0.010	-0.017	0.000
	(0.009)	(0.041)	(0.034)	(0.035)	(0.011)	(0.010)	(0.014)
Other neighbor after 0 to 5 months	0.020	-0.008	-0.035	-0.006	-0.007	-0.003	-0.008
	(0.013)	(0.025)	(0.041)	(0.022)	(0.013)	(0.013)	(0.009)
Other neighbor after 6 to 11 months	0.022	-0.015	-0.045	-0.002	0.001	-0.003	-0.008
	(0.015)	(0.026)	(0.036)	(0.019)	(0.013)	(0.013)	(0.011)
Other neighbor after 12 or more months	0.024	-0.042	-0.035	-0.005	-0.006	0.005	-0.002
	(0.015)	(0.066)	(0.063)	(0.025)	(0.025)	(0.015)	(0.022)
N	294480	294480	294480	294429	294480	293938	252559

Notes: See Table A.2. Additionally note that all models control for municipality fixed effects, state-by-year-by-month fixed effects, and indicator variables for 0–5, 6–11, and 12+ months after the beginning of the war for municipalities with DTO presence.
* significant at 10%; ** significant at 5%; *** significant at 1%

A.2 Full-Time Schools

Figure A.8: Full-Time Schools, 2007/08-2012/13

(a) 2007/08 (b) 2008/09 (c) 2009/10 (d) 2010/11 (e) 2011/12 (f) 2012/13



Notes: Each panel shows the geographic distribution of schools in the FTS program in a given academic year. Schools' location data comes from the Schools Information System (SNIE).

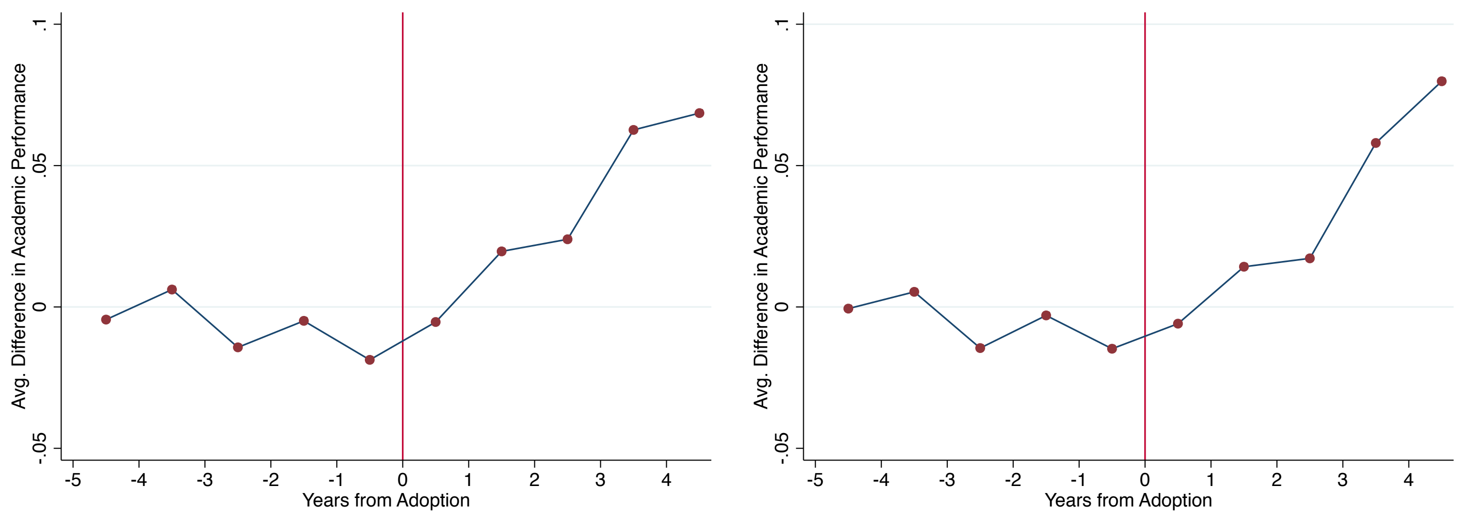
Table A.5: Sensitivity Analysis for Estimated Effects of Kingpin Captures on Homicide Rates

	(1)	(2)	(3)	(4)	(5)	(6)
DTO-controlled municipalities omitted from analysis:	none	Sinaloa-Beltrán-Leyva	Tijuana	Gulf	Juárez	Familia
Municipality of capture after 0 to 5 months	0.476**	0.529**	0.458*	0.888***	0.228	0.444
	(0.226)	(0.240)	(0.271)	(0.129)	(0.312)	(0.292)
Municipality of capture after 6 to 11 months	0.392***	0.419***	0.394***	0.573***	0.298**	0.464***
	(0.093)	(0.103)	(0.108)	(0.130)	(0.116)	(0.121)
Municipality of capture after 12 or more months	0.523*	0.517	0.440	1.253***	0.317	0.594*
	(0.302)	(0.318)	(0.353)	(0.036)	(0.428)	(0.352)
Neighbor w/ same DTO after 0 to 5 months	0.101	0.195**	0.084	0.261***	0.125	0.080
	(0.121)	(0.081)	(0.094)	(0.095)	(0.111)	(0.130)
Neighbor w/ same DTO after 6 to 11 months	-0.066	0.024	-0.080	-0.070	-0.011	-0.051
	(0.073)	(0.069)	(0.061)	(0.149)	(0.082)	(0.111)
Neighbor w/ same DTO after 12 or more months	-0.134	-0.058	-0.199	-0.058	-0.139	-0.095
	(0.125)	(0.167)	(0.134)	(0.041)	(0.181)	(0.131)
Non-neighbor w/same DTO after 0 to 5 months	0.008	0.017	0.005	0.042	-0.000	0.013
	(0.024)	(0.025)	(0.018)	(0.030)	(0.026)	(0.033)
Non-neighbor w/same DTO after 6 to 11 months	0.040	0.047	0.022	0.086	0.018	0.073
	(0.039)	(0.049)	(0.035)	(0.056)	(0.055)	(0.045)
Non-neighbor w/same DTO after 12 or more months	0.115***	0.107**	0.096**	0.149***	0.091	0.142***
	(0.044)	(0.046)	(0.039)	(0.042)	(0.063)	(0.051)
Other neighbor after 0 to 5 months	-0.060**	-0.050***	-0.046***	-0.052**	-0.038	-0.084
	(0.029)	(0.011)	(0.013)	(0.023)	(0.039)	(0.078)
Other neighbor after 6 to 11 months	-0.085**	-0.094***	-0.077**	-0.117***	-0.048	-0.095
	(0.042)	(0.023)	(0.032)	(0.025)	(0.051)	(0.143)
Other neighbor after 12 or more months	-0.098	-0.071**	-0.095	-0.124*	0.010	-0.171
	(0.090)	(0.033)	(0.077)	(0.072)	(0.036)	(0.178)
N	294480	274560	288840	260760	285960	284880
State-by-year fixed effects	yes	yes	yes	yes	yes	yes
Controls	yes	yes	yes	yes	yes	yes

Notes: See Table A.2. Additionally note that all models control for municipality fixed effects, state-by-year-by-month fixed effects, and indicator variables for 0–5, 6–11, and 12+ months after the beginning of the war for municipalities with DTO presence.
* significant at 10%; ** significant at 5%; *** significant at 1%

Figure A.9: Estimated Difference in Students' Academic Performance Before and After the Adoption the Full-Time Schools Program Between Students in Adopting and Non-Adopting Schools

(a) Math (b) Reading



Notes: Each panel shows the average difference over time between test scores of students in full-time schools and students in other public schools that did not adopt the FTS program between 2007/08 and 2012/13. The time scale is adjusted to address the fact that different schools adopt the FTS program at different times.

Table A.6: Annual Budget of the Full-Time Schools Program

Academic Year	Additional Budget (Millions of Pesos)	Number of Schools	Avg. Additional Budget per School (Thousands of Pesos)	Number of students	Avg. Additional Budget per Student (Pesos)
2007/08	NA	500	N/A	136,500	N/A
2008/09	N/A	953	N/A	192,830	N/A
2009/10	457.01	2,000	228.51	368,620	1,240
2010/11	448.36	2,273	197.25	439,231	1,021
2011/12	1,548.17	4,758	325.38	932,324	1,661
2012/13	2,508.72	6,708	373.99	1,368,022	1,834
2013/14	5,289.16	15,349	344.59	2,143,811	2,467
2014/15	10,382.86	23,182	447.88	3,463,041	2,998

Notes: Program's budget is based on the evaluation report 2014-2015 (CONEVAL, 2015). The number of schools and students in full-time schools are based on data from the Ministry of Education in Mexico. Values are expressed in constant 2012 prices.

Table A.7: Estimated Effects of Extending School Days on Student Achievement

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Math							
2 years prior							0.010
							(0.023)
1 year prior						-0.001	0.002
						(0.019)	(0.026)
Year of adoption	0.014	0.014	0.014	0.014	0.015	0.015	0.019
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.021)	(0.026)
1 year after	0.049**	0.049**	0.049**	0.049**	0.050**	0.050**	0.054*
	(0.019)	(0.019)	(0.019)	(0.020)	(0.020)	(0.023)	(0.027)
2 years after	0.075***	0.075***	0.075***	0.076***	0.077***	0.076***	0.081***
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.021)	(0.027)
3 years after	0.123***	0.123***	0.123***	0.124***	0.124***	0.123***	0.128***
	(0.022)	(0.022)	(0.022)	(0.021)	(0.021)	(0.025)	(0.031)
4 years after	0.149***	0.149***	0.149***	0.149***	0.147***	0.146***	0.150***
	(0.027)	(0.027)	(0.028)	(0.029)	(0.029)	(0.031)	(0.035)
N	42035412	42035412	42035412	42024942	42024942	42024942	42024942
Grade Fixed Effects	no	yes	yes	yes	yes	yes	yes
Grade by Year Fixed Effects	no	no	yes	yes	yes	yes	yes
School by Grade Fixed Effects	no	no	no	yes	yes	yes	yes
Other programs	no	no	no	no	yes	yes	yes
Panel B: Reading							
2 years prior							0.011
							(0.020)
1 year prior						0.002	0.006
						(0.013)	(0.020)
Year of adoption	0.014	0.014	0.014	0.014	0.014	0.015	0.019
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.018)	(0.023)
1 year after	0.041**	0.041**	0.041**	0.041**	0.041**	0.042*	0.046*
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.020)	(0.025)
2 years after	0.066***	0.066***	0.066***	0.068***	0.068***	0.069***	0.074***
	(0.019)	(0.019)	(0.019)	(0.018)	(0.018)	(0.022)	(0.026)
3 years after	0.116***	0.116***	0.116***	0.116***	0.116***	0.116***	0.122***
	(0.021)	(0.021)	(0.021)	(0.019)	(0.019)	(0.023)	(0.029)
4 years after	0.157***	0.157***	0.157***	0.157***	0.156***	0.156***	0.161***
	(0.027)	(0.027)	(0.027)	(0.025)	(0.026)	(0.027)	(0.030)
N	41873388	41873388	41873388	41862944	41862944	41862944	41862944
Grade Fixed Effects	no	yes	yes	yes	yes	yes	yes
Grade by Year Fixed Effects	no	no	yes	yes	yes	yes	yes
School by Grade Fixed Effects	no	no	no	yes	yes	yes	yes
Other programs	no	no	no	no	yes	yes	yes

Notes: Each column in each panel represents a different regression. Observations are at the student level, spanning 2007/08 though 2012/13 academic years. All estimates include school fixed effects and year fixed effects. The other programs for columns 5-7 are indicator variables for schools participating in the Quality Schools program and the Secure School Program. Estimated standard errors in parentheses are two-way clustered at the state and grade-by-year level. Student test scores from each exam are normalized to have mean zero and standard deviation one.

* Significant at the 10% level; ** Significant at the 5% level; *** Significant at the 1% level

Table A.8: Estimated Effects on School Composition and Other Factors of the Education Production Function

	(1) Enrollment	(2) Fraction of Male	(3) Class Size	(4) Number of Teachers	(5) Fraction of Teachers w/ Graduate Education
Year of adoption	0.049***	-0.000	0.585***	0.065	0.004
	(0.007)	(0.000)	(0.125)	(0.042)	(0.003)
1 year after	0.062***	-0.000	0.684***	0.148	0.003
	(0.009)	(0.001)	(0.206)	(0.094)	(0.004)
2 years after	0.070***	-0.002	0.633*	0.272	0.006
	(0.010)	(0.002)	(0.369)	(0.205)	(0.004)
3 years after	0.073***	-0.002	0.721**	0.236	0.006
	(0.012)	(0.003)	(0.324)	(0.188)	(0.005)
4 years after	0.079***	-0.003	0.784**	0.259	0.005
	(0.016)	(0.003)	(0.308)	(0.186)	(0.008)
N	705024	705024	705024	705024	705024

Notes: Each column represents a different regression. Observations are at the school level, spanning 2005/06 though 2012/13 academic years. All estimates include school fixed effects and year fixed effects. Estimated standard errors in parentheses are clustered at the state level.
* Significant at the 10% level; ** Significant at the 5% level; *** Significant at the 1% level

Table A.9: Estimated Effects on Student Achievement: Lower and Upper Bounds

	Estimates (1)	Lower Bound 1 (2)	Upper Bound 1 (3)	Lower Bound 2 (4)	Upper Bound 2 (5)
Panel A: Math					
Year of adoption	0.016	-0.021	0.051***	0.015	0.016
	(0.017)	(0.018)	(0.017)	(0.017)	(0.017)
1 year after	0.050**	-0.002	0.101***	0.050**	0.051**
	(0.020)	(0.020)	(0.022)	(0.020)	(0.020)
2 years after	0.077***	0.005	0.146***	0.076***	0.078***
	(0.017)	(0.019)	(0.021)	(0.017)	(0.017)
3 years after	0.124***	0.046*	0.201***	0.122***	0.125***
	(0.021)	(0.024)	(0.025)	(0.021)	(0.021)
4 years after	0.147***	0.058*	0.234***	0.144***	0.149***
	(0.029)	(0.033)	(0.031)	(0.029)	(0.029)
N	42024942	42024942	42024942	42024942	42024942
Panel A: Reading					
Year of adoption	0.014	-0.023	0.050***	0.014	0.014
	(0.015)	(0.016)	(0.016)	(0.015)	(0.015)
1 year after	0.041**	-0.011	0.093***	0.040**	0.041**
	(0.017)	(0.018)	(0.019)	(0.017)	(0.017)
2 years after	0.068***	-0.005	0.138***	0.067***	0.069***
	(0.018)	(0.021)	(0.022)	(0.018)	(0.018)
3 years after	0.116***	0.036	0.193***	0.114***	0.117***
	(0.019)	(0.023)	(0.023)	(0.019)	(0.019)
4 years after	0.157***	0.065**	0.243***	0.153***	0.158***
	(0.025)	(0.029)	(0.030)	(0.026)	(0.025)
N	41862944	41862944	41862944	41862944	41862944

Notes: Each column in each panel represents a different regression. Observations are at the student level, spanning 2007/08 though 2012/13 academic years. All estimates include school fixed effects, year fixed effects, grade fixed effects, grade-by-year fixed effects, school-by-grade fixed effects, and indicator variables for schools participating in the Quality Schools Program and the Secure School Program. Estimated standard errors in parentheses are two-way clustered at the state and grade-by-year level. Student test scores from each exam are normalized to have mean zero and standard deviation one.
* Significant at the 10% level; ** Significant at the 5% level; *** Significant at the 1% level

Table A.10: Estimated Effects on Student Achievement by Grade

	(1) 3rd grade	(2) 4th grade	(3) 5th grade	(4) 6th grade
Panel A: Math				
Year of adoption	0.009	0.008	0.035**	0.010
	(0.016)	(0.019)	(0.013)	(0.014)
1 year after	0.049**	0.040	0.052**	0.060**
	(0.018)	(0.022)	(0.014)	(0.016)
2 years after	0.087**	0.081***	0.085***	0.053*
	(0.024)	(0.019)	(0.010)	(0.022)
3 years after	**0.142*** **	0.143***	0.119***	0.092**
	(0.023)	(0.021)	(0.021)	(0.025)
4 years after	**0.141*** **	**0.214*** **	0.101***	0.132**
	(0.026)	(0.022)	(0.018)	(0.036)
N	10441478	10559490	10598414	10425560
Panel B: Reading				
Year of adoption	0.014	0.015	0.028	0.000
	(0.016)	(0.015)	(0.015)	(0.013)
1 year after	0.053**	0.038*	0.038**	0.036*
	(0.020)	(0.019)	(0.014)	(0.015)
2 years after	0.089**	0.075**	0.074***	0.033
	(0.026)	(0.020)	(0.011)	(0.022)
3 years after	**0.140*** **	0.127***	0.104***	0.093***
	(0.029)	(0.017)	(0.019)	(0.022)
4 years after	**0.166*** **	**0.192*** **	0.125***	0.139**
	(0.025)	(0.018)	(0.026)	(0.035)
N	10393256	10535282	10591644	10342762

Notes: Each column in each panel represents a different regression. Observations are at the student level, spanning 2007/08 though 2012/13 academic years. All estimates include school fixed effects, year fixed effects, and indicator variables for schools participating in the Quality Schools Program and the Secure School Program. Estimated standard errors in parentheses are two-way clustered at the state and grade-by-year level. Student test scores from each exam are normilized to have mean zero and standard deviation one. * Significant at the 10% level; ** Significant at the 5% level; *** Significant at the 1% level

Table A.11: Estimated Effects on Student Achievement by Poverty Level of the School's Locality

	(1) Low Poverty	(2) High Poverty
Panel A: Math		
Year of adoption	0.006	0.026
	(0.022)	(0.022)
1 year after	0.025	0.045**
	(0.029)	(0.022)
2 years after	0.043	0.087***
	(0.026)	(0.030)
3 years after	0.068**	0.140***
	(0.034)	(0.032)
4 years after	0.097**	0.152***
	(0.037)	(0.039)
N	19531410	19618329
Grade Fixed Effects	yes	yes
Grade by Year Fixed Effects	yes	yes
School by Grade Fixed Effects	yes	yes
Panel B: Reading		
Year of adoption	0.000	0.026
	(0.017)	(0.022)
1 year after	0.015	0.048**
	(0.023)	(0.022)
2 years after	0.040	0.071**
	(0.026)	(0.028)
3 years after	0.068**	0.137***
	(0.030)	(0.030)
4 years after	0.108***	0.164***
	(0.031)	(0.041)
N	19441446	19555486
Grade Fixed Effects	yes	yes
Grade by Year Fixed Effects	yes	yes
School by Grade Fixed Effects	yes	yes

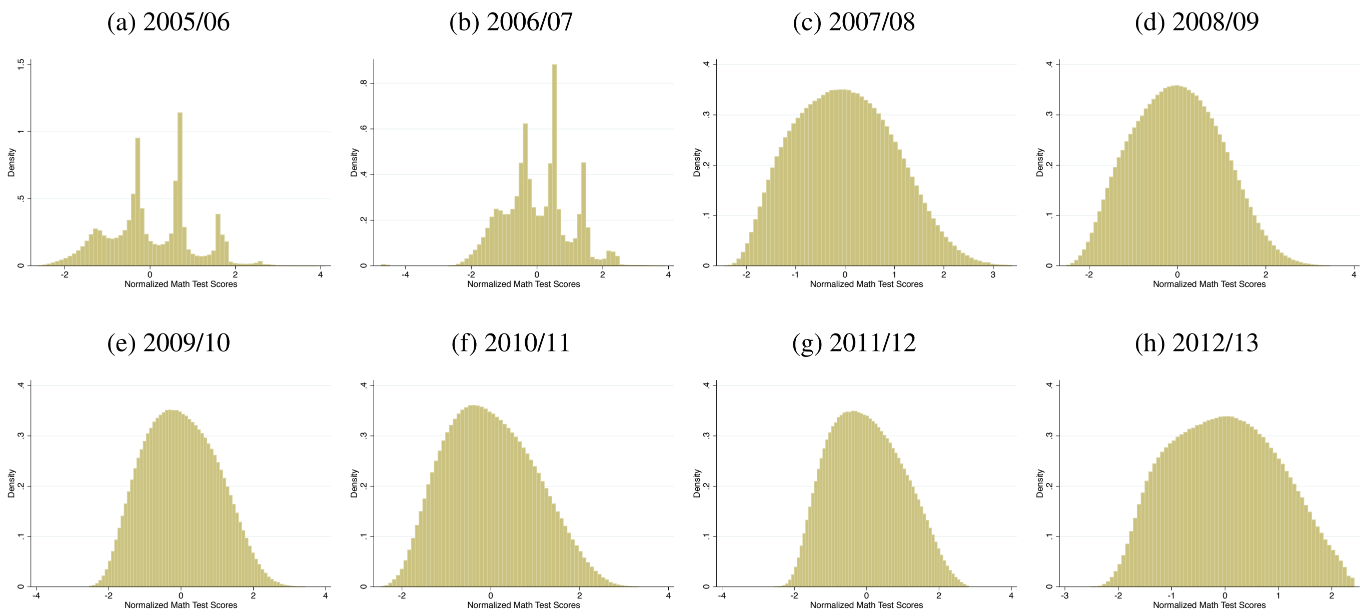
Notes: Each column in each panel represents a different regression. Observations are at the student level, spanning 2007/08 though 2012/13 academic years. All estimates include school fixed effects, year fixed effects, and indicator variables for the Quality Schools Program and the Secure School Program. Estimated standard errors in parentheses are two-way clustered at the state and grade-by-year level. Student test scores from each exam are normalized to have mean zero and standard deviation one.
* Significant at the 10% level; ** Significant at the 5% level; *** Significant at the 1% level

Table A.12: Estimated Effects on Dropped Observations

Fraction of students with:	(1) <50% Math Answers	(2) <50% Reading Answers	(3) Unreliable Test Scores
Year of adoption	-0.000	0.000	-0.001
	(0.001)	(0.001)	(0.002)
1 year after	-0.001	0.000	-0.001
	(0.001)	(0.001)	(0.002)
2 years after	-0.003*	-0.003	0.006
	(0.002)	(0.002)	(0.005)
3 years after	-0.002	-0.002	0.001
	(0.002)	(0.003)	(0.006)
4 years after	-0.003	-0.002	-0.004
	(0.002)	(0.002)	(0.005)
N	472676	472676	472676

Notes: Each column represents a different regression. Observations are at the school level, spanning 2007/08 though 2012/13 academic years. All estimates include school fixed effects and year fixed effects. Estimated standard errors in parentheses are clustered at the state level.
* Significant at the 10% level; ** Significant at the 5% level; *** Significant at the 1% level

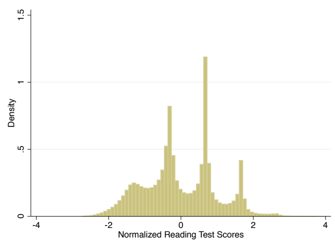
Figure A.10: Distribution of Normalized Math ENLACE Test Scores 2005-2012



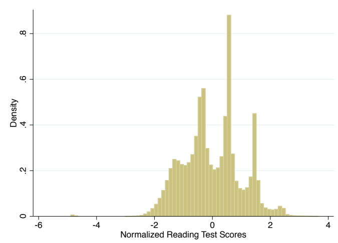
Notes: Each panel shows the test score distribution based on student level ENLACE Test Scores from the Ministry of Education.

Figure A.11: Distribution of Normalized Reading ENLACE Test Scores 2005-2012

(a) 2005/06

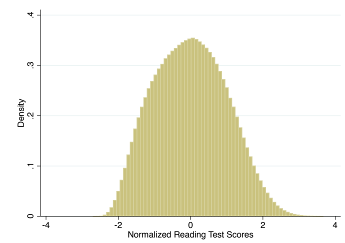


(b) 2006/07

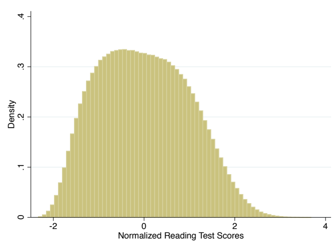


(c) 2007/08

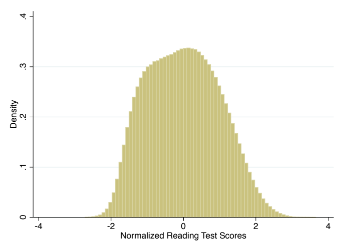
(d) 2008/09



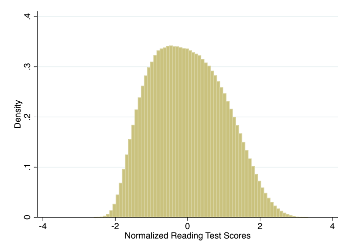
(e) 2009/10



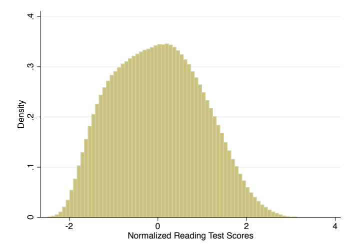
(f) 2010/11



(g) 2011/12



(h) 2012/13



Notes: Each panel shows the test score distribution based on student level ENLACE Test Scores from the Ministry of Education.

A.2.1 Enrollment Effects and Bounds

Let FTE_t and HTE_t denote the number of students enrolled in full- and half-time schools at school year t, respectively. And let $\%\Delta e$ denote the estimated effect of the FTS program on enrollment. For the first set of bounds, I find the percentage of students that will change treatment status, denoted by $\%\Delta' e$.

Assume $FTS_t = HTS_t$. Then

$$FTE_{t+1} = (1 + \frac{1}{2}\%\Delta e) \cdot FTE_t \tag{A.1}$$

$$FTE_t = (1 - \%\Delta' e) \cdot FTE_{t+1} \tag{A.2}$$

Solving equations A.1 and A.2, the percentage of students in the treatment group that will change treatment status is given by

$$\%\Delta' e = 1 - \frac{1}{1 + \frac{1}{2}\%\Delta e} \tag{A.3}$$

For the second set of bounds, I solve for remaining percentage of students that will change treatment status, denoted by $\%\Delta'' e$.

Assume $FTS_t = HTS_t$. Then

$$FTE_{t+1} = (1 + \frac{1}{2}\%\Delta_1 e) \cdot FTE_t \tag{A.4}$$

$$FTE_{t+2} = (1 + \frac{1}{2}\%\Delta_2 e) \cdot FTE_t \tag{A.5}$$

$$FTE_{t+1} = (1 - \%\Delta'' e) \cdot FTE_{t+2} \tag{A.6}$$

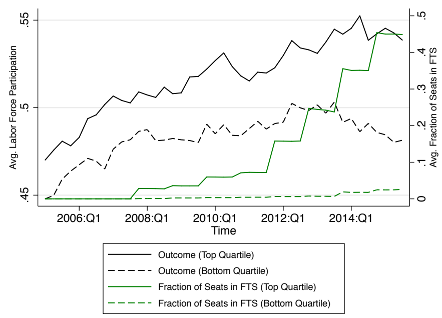
Solving equations A.4 - A.6, the fraction of students in the treatment group that will change treatment status is given by

$$\%\Delta'' e = 1 - \frac{1 + \frac{1}{2}\%\Delta_1 e}{1 + \frac{1}{2}\%\Delta_2 e} \tag{A.7}$$

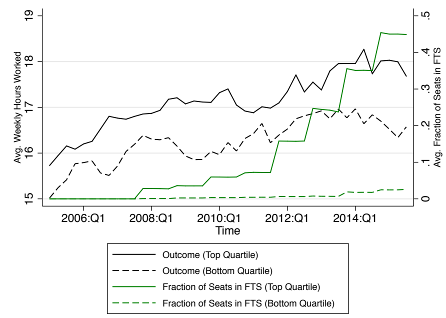
A.3 Mothers' Labor Supply

Figure A.12: Adjusted Female Labor Outcomes for Municipalities with High and Low Intensity of Treatment

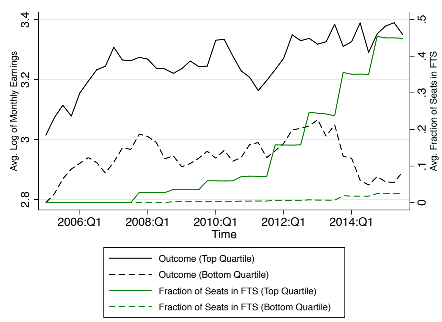
(a) Labor Force Participation



(b) Number of Weekly Hours Worked



(c) Log of Monthly Earnings



Notes: Each panel separately shows female labor outcomes for municipalities with a high (top quartile) and low (bottom quartile) average fraction of predicted FTS seats. The left axis shows the state-by-year-by-quarter adjusted average of labor outcomes and the right axis the average fraction of predicted FTS seats. Female labor outcomes are calculated based on the National Survey of Occupation and Employment (ENOE) and the fraction of predicted seats in FTS is calculated based on census data from the Ministry of Education in Mexico.

Table A.13: Long-Difference Estimated Effects of the Fraction of Seats in FTS on Female Labor Outcomes

	(1)	(2)	(3)	(4)	(5)
Panel A: Labor Force Participation					
Fraction of seats in FTS	0.043**	0.071***	0.070***	0.063**	0.064**
	(0.021)	(0.026)	(0.026)	(0.027)	(0.028)
1 Year prior				-0.019	-0.019
				(0.019)	(0.020)
2 Years prior					0.001
					(0.020)
N	166089	166089	166085	166085	166085
Panel B: Number of Weekly Hours Worked					
Fraction of seats in FTS	0.261	2.352**	2.352**	2.205**	2.190**
	(0.892)	(0.983)	(0.980)	(0.993)	(1.013)
1 Year prior				-0.398	-0.408
				(0.760)	(0.804)
2 Years prior					-0.036
					(0.726)
N	166089	166089	166085	166085	166085
Panel C: Log of Monthly Earnings					
Fraction of seats in FTS	0.338**	0.469**	0.465**	0.459**	0.445**
	(0.164)	(0.201)	(0.201)	(0.207)	(0.219)
1 Year prior				-0.015	-0.025
				(0.139)	(0.147)
2 Years prior					-0.034
					(0.152)
N	166089	166089	166085	166085	166085
State-by-time fixed effects	no	yes	yes	yes	yes
Time-variant individual controls	no	no	yes	yes	yes

Notes: Each column in each panel represents a different regression. Observations are at the individual level, spanning from 2005:Q1 to 2015:Q3. All specifications include year-by-quarter fixed effects. Estimated robust standard errors in parentheses are clustered at the municipality level. Individual controls include a quadratic function of age, a quadratic function of the age of the youngest child, and a quadratic function of the number of children.
*, **, *** Significant at the 10%, 5%, and 1% levels, respectively.

Table A.14: Long-Difference Estimated Effects of the Fraction of Seats in FTS on Attrition

	(1)	(2)
Fraction of seats in FTS	0.031	0.013
	(0.045)	(0.035)
N	328409	328409
Time fixed effects	yes	yes
State-by-time fixed effects	no	yes

Notes: Each column in each panel represents a different regression. Observations are at the individual level, spanning from 2005:Q1 to 2015:Q3. Estimated robust standard errors in parentheses are clustered at the municipality level.
*, **, *** Significant at the 10%, 5%, and 1% levels, respectively.

Table A.15: Long-Difference Estimated Effects of the Fraction of Seats in FTS on Female Labor Outcomes by Education

Years of schooling:	overall (1)	0–9 (2)	10+ (3)
Panel A: Labor Force Participation			
Fraction of seats in FTS	0.070***	0.078**	0.011
	(0.026)	(0.031)	(0.046)
N	166085	111512	54297
Mean	0.494	0.424	0.637
Panel B: Number of Weekly Hours Worked			
Fraction of seats in FTS	2.352**	2.536**	0.977
	(0.980)	(1.137)	(1.911)
N	166085	111512	54297
Mean	16.478	14.086	21.400
Panel C: Log of Monthly Earnings			
Fraction of seats in FTS	0.465**	0.479**	0.135
	(0.201)	(0.228)	(0.542)
N	166085	111512	54297
Mean	3.101	2.550	4.232

Notes: Each column in each panel represents a different regression. Observations are at the individual level, spanning from 2005:Q1 to 2015:Q3. All specifications include year-by-quarter fixed effects, state-by-year-by-quarter fixed effects, and individual controls. Estimated robust standard errors in parentheses are clustered at the municipality level. Individual controls include a quadratic function of age, a quadratic function of the age of the youngest child, and a quadratic function of the number of children. Nine years of schooling translates into completed junior high school, which is the median education level and the compulsory level of education in Mexico.
*, **, *** Significant at the 10%, 5%, and 1% levels, respectively.

Table A.16: Long-Difference Estimated Effects of the Fraction of Seats in FTS on Female Labor Outcomes by Poverty Level of the Locality of Residence

Poverty level:	overall (1)	Low Poverty (2)	High Poverty (3)
Panel A: Labor Force Participation			
Fraction of seats in FTS	0.070*** (0.026)	-0.002 (0.059)	0.086*** (0.032)
N	166085	83569	81072
Panel B: Number of Weekly Hours Worked			
Fraction of seats in FTS	2.352** (0.980)	-0.465 (2.195)	3.274*** (1.161)
N	166085	83569	81072
Panel C: Log of Monthly Earnings			
Fraction of seats in FTS	0.465** (0.201)	-0.214 (0.499)	0.630*** (0.233)
N	166085	83569	81072

Notes: Each column in each panel represents a different regression. Observations are at the individual level, spanning from 2005:Q1 to 2015:Q3. All specifications include year-by-quarter fixed effects, state-by-year-by-quarter fixed effects, and individual controls. Estimated robust standard errors in parentheses are clustered at the municipality level. Individual controls include a quadratic function of age, a quadratic function of the age of the youngest child, and a quadratic function of the number of children. Low and high poverty localities are defined as those below and above the median poverty index, respectively.

*, **, *** Significant at the 10%, 5%, and 1% levels, respectively.

Table A.17: Heterogeneous Treatment Effects on Labor Outcomes

	Women		Men	
	w/ school age children (1)	w/o school age children (2)	w/ school age children (3)	w/o school age children (4)
Panel A: Labor Force Participation				
Fraction of seats in FTS	0.070**	0.041***	-0.022	-0.000
	(0.026)	(0.014)	(0.014)	(0.013)
N	166085	456755	140883	374963
Panel B: Number of Weekly Hours Worked				
Fraction of seats in FTS	2.352**	0.930	0.604	-0.0391
	(0.980)	(0.588)	(1.443)	(0.928)
N	166085	456755	140883	374963
Panel C: Log of Monthly Earnings				
Fraction of seats in FTS	0.465**	0.130	-0.039	0.063
	(0.201)	(0.116)	(0.269)	(0.175)
N	166085	456755	140883	374963

Notes: Each column in each panel represents a different regression. Observations are at the individual level, spanning from 2005:Q1 to 2015:Q3. All specifications include year-by-quarter fixed effects, state-by-year-by-quarter fixed effects, and individual controls. Estimated robust standard errors in parentheses are clustered at the municipality level. Individual controls (only for mothers' with children) include a quadratic function of age, a quadratic function of the age of the youngest child, and a quadratic function of the number of children.
*, **, *** Significant at the 10%, 5%, and 1% levels, respectively.