THREE ESSAYS ON THE UNINTENDED CONSEQUENCES
OF SOCIAL POLICIES

Felipe A. Lozano-Rojas

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements for the degree
Doctor of Philosophy
in the O’Neill School of Public and Environmental Affairs,
Indiana University
June - 2020
Accepted by the Graduate Faculty, Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Doctoral Committee

Maureen Pirog, PhD (Co-Chair)

Justin Ross, PhD (Co-Chair)

Kosali Simon, PhD

Bradley Heim, PhD

March 6, 2020
To my mother, my siblings, and my family. To my advisors, my committee, and all of my mentors. Thanks a lot for all your help and support.
Felipe A. Lozano-Rojas

THREE ESSAYS ON THE UNINTENDED CONSEQUENCES OF SOCIAL POLICIES

The core of this dissertation is the understanding of social policy, under the limitations established by resource and context constraints. On the three essays included here, the main concern is the interaction between the goals if social policy and the means by which societies, through a level of government, finance, fund, or advance those policies. Different levels of government have to pick among a series of policy levers to deal with budget or context constraints that the institutional framework provides them with. The spirit of this work is to inform the discussions surrounding those policies beyond the immediate areas that they intend to affect. The approach I take here evaluates policies on their unintended effects. Given complexities in policy design and implementation, I have conducted empirical policy analysis with the intention of understanding some of the factors that separate implementation from the ideal design and materialize the analysis into policy recommendations.

Maureen Pirog, PhD (Co-Chair)

Justin Ross, PhD (Co-Chair)

Kosali Simon, PhD

Bradley Heim, PhD
Contents

Dedication iii

Abstract iv

1 Introduction 1

1.1 Piece-meal Study of Social Policies 7

Bibliography 11

2 A matter of design in Soda Taxes: Tax sugar instead of volume 12

2.1 Introduction 13

2.2 Literature Review 16

2.3 Drinks with Different Sugar Concentration: A Model 19

2.3.1 Tax effects 22

2.4 Data 27

2.4.1 Nielsen Data 27

2.4.2 Nutrition Facts 28

2.5 Methodology: The Effect of the Policy 29

2.6 Main Results 30

2.6.1 Parallel Trends and Control Group Selection 32

2.6.2 Changes in SSB Consumption and Prices per Unit 34

2.6.3 Effects Over Time 36

2.6.4 A Direct Test for Leakage and SSB Relative Price Changes 38

2.6.5 Robustness Exercises 39

2.7 Discussion and Conclusion 43

Bibliography 46
# Cash Subsidies to improve Student Loan programs for Low-Income Students

## 3.1 Introduction

## 3.2 Literature Review

### 3.2.1 Evaluations to the ACCES Loan Program

## 3.3 Policy Background

### 3.3.1 The ACCES-PACES loan program for higher education

### 3.3.2 Additional Cash Subsidies

## 3.4 Methodology and Data

### 3.4.1 Methodology

### 3.4.2 Data

## 3.5 Main Results

### 3.5.1 Assessment of the RDD Assumptions

## 3.6 Results on Main Academic Outcomes

### 3.6.1 Impact on Dropout Rates

### 3.6.2 Impact on First Year Dropout Rates

### 3.6.3 Impact on Graduation Rates

### 3.6.4 Heterogeneous Effects Across Geographical Areas

## 3.7 Discussion and Conclusions

# Consumer Incidence in Sales Tax Holidays: Evidence from Tennessee

## 4.1 Introduction

## 4.2 Nielsen Data, State Sales Tax Holidays, and Empirical Strategy

### 4.2.1 Estimation Strategy
Chapter 1

Introduction

At its core, this dissertation is meant to expand our understanding of social policy, under the limitations established by resource and context constraints. Accordingly, the essays included here consider the interaction between the goals of social policy and the means through which societies—through different levels of government—finance, fund, or advance those policies. In most circumstances, governments at the international, central/federal, state, or local level will have to pick among a series of policy levers and also deal with the budget constraints presented by their institutional framework. The spirit of this work is to inform the discussions about what we can and have learned from implementing specific social policies beyond the immediate areas that they intend to affect. However, as evidenced by this introduction and each individual essay, my approach puts a special emphasis on both policies’ intentions and their unintended outcomes.

Given the complexity of human interactions in policy design and implementation, I have conducted empirical policy analysis with the intention of understanding the factors that separate implementation from ideal design, learning from the particular experiences I analyze, and materializing my findings into generalizable policy recommendations.

Social policy is designed with vulnerable people in mind. Most social policy is designed on the grounds of redistribution fairness, but in some cases it may also be designed in response to corrective market failures that lead to outcomes that society might consider unfair and want the government to correct. Whenever a policymaker or a politician tries to cater to constituents to motivate and justify social policy, they will bring the voice of the populations they want to affect the most through these policies. While their motives and their explanations, might sound fair, their capacity to develop a comprehensive, centered, top-down approach to solve
problems rooted in inequality is limited to the extent that unique individuals or frameworks can process information, and adapt to new circumstances in order to avoid restrictions. In this sense, a *one-size fits all* policy seems hard to attain.

Traditional welfare economics has provided a theoretical framework to establish the trade-offs of interventions that have to some extent shaped the mindsets of academics, policy advisers, and policymakers for most of the 20th century. While we have aimed to understand complex social phenomena, we have also simplified our modeling of theoretical constructs. As a thought exercise, this simplification has proven enlightening in many cases, but it has also fallen short in the face of myriad possible imaginative responses from a public trying to adapt to new circumstances.

The debate between James Buchanan and Richard Musgrave in the late 90’s well illustrates the criticism to which welfare economics theory has been exposed. Buchanan famously questioned the theory’s reach and affect on policy, pointing out that welfare economists, pursuing an unattainable agenda, ignored the human processes required to meet the ends of government intervention (Buchanan and Musgrave, 1999). While welfare economics analysis is grounded on methodological individualism, where the government is a cohesive entity with an unique preference function, Buchanan manifested particular distrust of the possibility of creating overarching top-down policies that cater to the preferences of a benevolent despot. Policies, that at the end, could go in unexpected or even opposite ways relative to the individuals the policies intend to affect. The impossibility of aggregating social preferences and the limited scope of our understanding to create technical or even democratic consensus, mine the usefulness of the welfare economics ideal, and render their abstraction as an oversimplification of public matters. The current thought of political economy, combining classical liberal thought with insights from governance and public administration, state that *normative individualism* and *skepticism regarding preference aggregation* distinguish governance under this approach from
economic policy, which they calls classical liberalism (Aligica et al., 2019).

Such criticism, according to my reading, does not mean that learning is impossible. Quite the contrary: they invite us, researchers, to analyze the contexts in which policies do (or do not) attain their desired objectives without unintentionally introducing undesired consequences. To quote Dr. Peter Boetke on a recent visit to Indiana University, “demand curves are downward trending”; building from this almost irrefutable finding, researchers have unveiled several other facts about human behavior and how we react to market and government interventions.

The discussion easily transcends academic considerations, as the policies under question, after all, affect the lives of millions of people around the globe. The work of William Easterly is illustrative of how technocratic design policy makes its way through top-down initiatives; in his account, policies pushed by multilateral organizations in particular often end up hurting the very people they were meant to help Easterly (2002, 2017, 2014). Easterly presents several examples in which top-down policy design has lead to catastrophic events. In his conclusion, he invites to conduct careful analysis of piecemeal policies, suggesting that researchers and government advisers should not dictate policy from a position of superiority, but rather should describe policies’ contexts and intended outcomes, so that affected individuals and policy-makers in different contexts can make their own decisions. It is my belief that, constantly performing empirical analysis of policies and analyzing stakeholders’ behavioral responses are other means of listening to those who are affected. This allows policy analysts to better and dynamically understand the limits of their original theory, learn from those limits, and go back to the drawing board to improve blueprints.

My approach moves from what Aligica et al. (2019) call expert-driven governance—in which researchers see the government as a benevolent despot with a simple welfare function to optimize—toward a polycentric world of multiple decision makers in a multiplicity of contexts. This approach assumes that decision makers, embedded in their political contexts, want to im-
prove their surroundings, and that empirical research can provide them with tools for making informed decisions, rather than mandates. Decades ago, Buchanan painted a similar portrait of the scientist: as he writes, the “role of the social scientists (...) is not [one] of improving anything directly; instead, it is that of explaining behavior of a certain sort which, only remotely and indirectly, can lead to improvements in the political process itself” (Aligica et al., 2019, p.29).

Easterly’s work not only highlights the role of scientists in informing policy, but also the importance of empiricism to policy analysis. In fact, his criticism does not question the role of government in policy design. He is concerned instead with the difficulty of understanding the realities faced by researchers, and a lack of commitment to policy evaluation on the part of donors and multilateral agencies (Easterly, 2003). Easterly suggests that we still advance knowledge when we evaluate carefully and consider research design when unveiling causal estimates. This amplifies the reach of out pre-existing theories as policy analysis has the potential to provide context – under what circumstances our understanding is right or not.

Easterly is probably one of the most eloquent critics of development and welfare economics’ top-down approach. But in advocating for the rightful place of empiricism in the research agendas of policy analysts and economists alike, the so-called “credibility revolution” has perhaps most shaped the contemporary discussions out of which this dissertation’s research exercises have evolved. Angrist and Pischke (2010) argues that the emphasis on empirical research design that clears the argument for causality between policies and their effects has lead to much greater credibility for microeconomics: “such studies offer a powerful method for deriving results that are defensible both in the seminar room and in a legislative hearing” (Angrist and Pischke, 2010). This attention to empirical research design posits clinical experimental trials as the gold standard for causal inference–without treating them the only means of conducting research, as quasi-experimental design has the power and potential to unveil causal estimates,
It is worth mentioning what the state of analysis looked like before the credibility revolution—and in some approaches to public administration and macroeconomics, what it still looks like today. Naïve regression analysis and theory calibration of policy-invariant theoretical parameters were, and have been, the main empirical tools starting in the 70’s; due to increased computing capacity over the past forty years, their statistical and modeling complexity has steadily increased over time. However, (Angrist and Pischke, 2010) undermines the credibility of several studies of the time: from his perspective, empiricism is not about unveiling causal relationships so much as it is about calibrating the best fit to some available data, a process that obscures what we assume about fundamental relationships across variables when we interpret these results. Angrist and Pischke argue, indeed, that the results of such studies are subject to change whenever the data changes (Angrist and Pischke, 2010).

Some findings in Statistics have also helped the advance of research design, as today we have a less dogmatic interpretation of regression analysis and are concerned less with functional forms: the “linear models that constitute the workhorse of contemporary empirical practice usually turn out to be remarkably robust” (Angrist and Pischke, 2010, p.12). Furthermore, an explosion in the quantity and availability of data has made it possible to exploit quasi-experimental designs in many more settings, as means of capturing data have exploited and increased the means of estimating policy effects and theoretical parameters. Angrist and Pischke (2010) note this, as have Einav and Levin (2014), more recently. As I will discuss in each of the chapters of the present dissertation, I include careful discussions of the validity of estimates, but I also make use of forms of Big Data that have allowed me to address my research questions. In particular I use census reach administrative registrars, barcode consumption databases, and webscrapping from multiple online stores and other online sources.

I would like to argue that the works presented in this dissertation, all concerning different
areas of social policy, attest to the importance of several contemporary practices that constitute the state-of-the-art in policy analysis. The first is related to being careful with the interpretation of theory. Theory may help us understand the phenomena under question, but it might also cloud the very mechanisms that make the ideal unfeasible. This happens as theory describes a limited amount of relationships among variables, whereas in reality there can be a myriad, and why reality drifts away from theoretical predictions might come from those additional, not described relationships. In the work developed here, I continually challenge theory’s tendency to simplify. I work against the assumption that we can access the will of a society from theoretical optimization algorithms; I suggest that it is impossible to forecast and plan for all potential stakeholders’ responses. In fact, as the title of this dissertation states, in this work I explore the unintended consequences of social policy that are not captured by theory. The second contemporary practice I refer to involves paying particular attention to the validity of the estimates, by means of setting causal estimates as the mean research objective; in doing so, I aim to build credible and reproducible, but also contextual, estimates that can inform theory and policy-making alike.

The evaluation of these unintended consequences, from my perspective, is one of the missing connections between the idealized world of theory, simplified as it is, and the real world we live in. Thus, to some extent the study of unintended consequences is the study, one step at a time, of why we cannot reliably aggregate outcomes, of why government fails, and of how we can better approximate ideals and prevent government failure.

In what follows, this introduction will make a simple case for the study of particular complexities that raise questions about potential unintended consequences. It closes by describing the dissertation’s motivating question and the design of its essays.
1.1 Piece-meal Study of Social Policies

Before proceeding any further with the introduction of individual analyses, we need to define social policy and explain its motivations. A summary of the field can be found on the web page for the London School of Economics’s Department of Social Policy, where Dr. Platt writes:

“Social policy is concerned with the ways societies across the world meet human needs for security, education, work, health and wellbeing. Social policy addresses how states and societies respond to global challenges of social, demographic and economic change, and of poverty, migration and globalisation. Social policy analyses the different roles of: national governments, the family, civil society, the market, and international organisations in providing services and support across the life course from childhood to old age. These services and support include child and family support, schooling and education, housing and neighbourhood renewal, income maintenance and poverty reduction, unemployment support and training, pensions, health and social care. Social policy aims to identify and find ways of reducing inequalities in access to services and support between social groups defined by socio-economic status, race, ethnicity, migration status, gender, sexual orientation, disability and age, and between countries” Platt (2019).

More often than not, social policies are designed to target low-income populations. But as Dr. Platt states, welfare inequality is rooted in diverse policy areas and social dimensions. In other words, poverty and vulnerability are not necessarily determined uniquely by income disparities; there are multiple dimensions from which inequality may arise. Indeed, race, the urban-rural divide, and other regional differences are just some of the considerations relevant to policy design. Context-specific historical processes determine which are subject to more attention by policymakers. For instance, the salience of race conflict in U.S. history has led to much research and many policy settings in which race is explored. This, however, is not in case in most other countries.

Sometimes institution design unintentionally perpetuates the marginalization of social groups. This should be recognized. Even when the effects of social policies on some groups are intended, their affects on those groups their design had not considered may be opposite, deepening the marginalization of these groups. These considerations should be raised at the
moment of evaluation. And this is only one abstract margin that we can think of without careful analysis of individual policies.

On grounds of Pareto fairness, complete policy evaluations should include additional unintended margins, the effects of a policy change on both, intended and unintended outcomes. The intended population targets, and unintended population targets, or similarly, not only intended but also unintended outcomes. In both cases, we expect that outcomes for intended population targets improve, while those for the unintended populations at least do not worsen. Otherwise, we will be describing a trade-off that makes the policy in question more costly.

Even when we are aware of the existence of these consequences, dealing with them adds to the complexity of policy design in a world of second bests. But improvements are not necessarily going to come from a clean slate policy design; they could arise marginally by means of redesigning improvements on the current state of policy. In that case, dealing with unintended consequences opens up possibilities for further exploration. With this in mind, this dissertation evaluates three separate policies and the disparity of their impacts across intended and unintended behavioral responses.

Below, I present the abstracts of the three articles that constitute the present document. They can also be found in every chapter. I hope that after consideration of each essay, the committee members agree that my work has successfully advanced knowledge and contributed to discussions about the role and reach of government in the areas under discussion.


Soda taxes are meant to improve public health outcomes insofar as they curb consumption of sugary beverages and thus sugar intake. This research is the first to get closer to the aim of the policy by evaluating the effect of the soda tax introduced in Berkeley in 2015, focusing on its effect on sugar consumption. I conduct this analysis on a unique dataset that merges retailer
point-of-sale (POS) data with matched Nutrition Facts information for sugar content. Due to the vast range of sugar densities across beverages, this setting provides information about how a tax levied at volume, rather than sugar content, affects demand for taxable items. I develop a theoretical model that accounts for how a tax such as the one in Berkeley, incentivizes consumers to substitute toward drinks with higher sugar concentration, as unitary sugar prices among taxed drinks are affected differently. I evaluate these hypotheses for Berkeley relative to comparable locations and assess the effects of soda taxes on consumption and price by volume and sugar content. I find that sugar intake fell as a response to the tax, but at a lower rate compared to volume consumption. Following the introduction of the tax, in other words, the unitary sugar price exhibited a lower impact in more concentrated sugar beverages. This indicates that the policy incentivized consumers to buy drinks with higher sugar concentrations among tax beverages, and that a levy at the sugar content level would be more effective.

**Second Essay - Using Cash Subsidies to Enhance Student Loan Programs for Low-Income Students**

Every year, governments allocate considerable amounts of resources to support students’ access to and performance in higher education. Grants and loans are directed to those students for whom college would seem unaffordable otherwise. These programs are justified on the grounds that they help students access, persist in, and graduate from institutions of higher education. However, considering the time it takes to observe academic or labor market performance, the literature evaluating student loans and its additional embedded programs is limited. This article attempts to fill this gap. Using a Regression Discontinuity Design (RDD), I evaluate a cash subsidy program embedded into the Colombian government’s main student loan program. I exploit the allocation rule of these additional cash subsidies to evaluate the effect of loan transfers on dropout and graduation rates, as well as on early default rates. My
results indicate that cash subsidies have substantial and significant effects on dropout rates at the cutoff point. For graduation rates, the effects are substantial, but barely significant. And with respect to the default rates, cash subsidy beneficiaries tend to have lower rates at the cutoff, though this result is not statistically significant. However, the measure of the subsidy effect on the latter two outcomes is affected by a statistical power loss at the proximity of the cutoff scores, a limitation of the RDD.

Third Essay - Consumer Incidence in Sales Tax Holidays: Evidence from Tennessee

Policymakers frequently employ tax holidays to stimulate spending and transfer state revenues to favored groups. Perhaps the most common form of this policy in the United States is the “Back to School” tax holiday, where in the fall, states waive the retail sales tax on selected school supplies. Surprisingly, empirical research on this policy has found that tax savings are overshifted to households. This paper argues that tax holiday dates should be assumed to be selected endogenously so as to minimize forgone revenue and maximize consumer incidence by targeting periods where prices are lowest. I consider causal evidence in a natural experiment out of Tennessee, where legislation for spring sales tax holidays on school supplies in 2006 and 2007 were reauthorized as part of a broader series of tax cuts. We conclude from the evidence of these Tennessee events that households, on average, receive the full tax savings offered by these programs. However, consumer incidence is heterogeneous with some retailers recapturing the tax savings with higher pre-tax prices.
Bibliography


Easterly, W. (2017). The white man’s burden: why the West’s efforts to aid the rest have done so much ill and so little good. Tantor Media.


Chapter 2

A matter of design in Soda Taxes: Tax sugar instead of volume

Abstract ∗, †

Soda taxes are meant to improve public health outcomes insofar as they curb consumption of sugary beverages and thus sugar intake. This research is the first to get closer to the aim of the policy by evaluating the effect of the soda tax introduced in Berkeley in 2015, focusing on its effect on sugar consumption. I conduct this analysis on a unique dataset that merges retailer point-of-sale (POS) data with matched Nutrition Facts information for sugar content. Due to the vast range of sugar densities across beverages, this setting provides information about how a tax levied at volume, rather than sugar content, affects demand for taxable items. I develop a theoretical model that accounts for how a tax such as the one in Berkeley, incentivizes consumers to substitute toward drinks with higher sugar concentration, as unitary sugar prices among taxed drinks are affected differently. I evaluate these hypotheses for Berkeley relative to comparable locations and assess the effects of soda taxes on consumption and price by volume and sugar content. I find that sugar intake fell as a response to the tax, but at a lower rate compared to volume consumption. Following the introduction of the tax, in other words, the unitary sugar price exhibited a lower impact in more concentrated sugar beverages. This indicates that the policy incentivized consumers to buy drinks with higher sugar concentrations among tax beverages, and that a levy at the sugar content level would be more effective.

∗The author appreciates comments and suggestions from Justin Ross, Kosali Simon, Coady Wing, Seth Freedman, Denvil Duncan, Alex Hollingsworth, John Stavick, and Brad Heim from Indiana University; Kate Smith and participants of the 2018 NTA meetings; and Catherine Wen, John Cawley, and participants of the 2019 ASHEcon Conference.

†Nielsen’s copyright of data: Copyright © 2020 The Nielsen Company (US), LLC. All Rights Reserved. The conclusions drawn from the Nielsen data are those of the researchers and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.
2.1 Introduction

Politicians propose taxes on sugar-sweetened beverages (SSB) in hopes of curbing sugar consumption from SSB, and so improving public health outcomes. In this regard, a soda tax\footnote{In this article the terms soda tax and SSB tax are used interchangeably.} is a sin tax, as it is levied on goods considered socially undesirable. The success of such taxes depends on their ability to raise prices, which are passed through to consumers; those who are most affected by the health concerns derived from over-consumption of high caloric foods are expected to decrease their consumption in turn. In this article, I explore how SSB taxes on volume miss the main item for SSBs: sugar content.

SSB taxes are at least partially motivated by research causally linking SSB consumption to the prevalence of diabetes and obesity (Han and Powell, 2013; Ogden et al., 2014; Colchero et al., 2016; Roache and Gostin, 2017) and by interest in broadening revenue sources (Hines, 2007). Prior to the 2015 introduction of an SSB tax in Berkeley (CA), most evidence for its efficacy came from international contexts (Berardi et al., 2016; Colchero et al., 2016), where soda tax introduction has been tied to reductions in consumption; or by measuring how changes in sales taxes, in general, have been accompanied by changes in soda consumption (Fletcher et al., 2010). More recent research has evaluated this policy by studying implementation of the SSB tax in Berkeley for its effects on SSB consumption and prices (in liquid ounces or unit consumption) (Falbe et al., 2016; Silver et al., 2017; Lee et al., 2019). Other cities have followed Berkeley, and as of 2020, eight cities effectively collect soda taxes; Washington D.C., however, is the only city to collect ad valorem rather than per liquid ounce of SSB.

Sodas are the main contributor to excessive caloric consumption (Allcott et al., 2019c), help explain the level of health expenditures across all levels of government and the desire to introduce soda taxes. In 2014, U.S. local governments spent $88.5 billion on non-capital hospital functions and $45.5 billion on public health services, which in total represented about
8% of overall local expenditures. State governments spent nearly the same amount at $66.9 billion and $44.2 billion for hospital and health, respectively. By comparison, state and local governments spent a combined $220.7 billion on police, corrections, and judicial administration during the same period: nearly $25 billion less than health and non-capital hospital outlays.\(^2\)

Meaningful impacts on individual consumption choices might represent long-run cost savings on items that are otherwise difficult for public decision-makers to influence. Some studies have addressed the extent to which soda taxes could lower these social costs and find that to fully cover the externality, taxes should be higher than the Berkeley tax: between 1.26¢ to 2.82¢ per liquid ounce (Allcott et al., 2019b; Grummon et al., 2019).

There is a disconnect, however, between the design of these policies and the goal of reducing sugar intake. Since the policies are often structured on a liquid content basis, there is an untaxed margin on sugar density that incentivizes consumers to switch to SSBs with greater sugar concentrations. This article provides evidence on this matter. Motor fuel taxes motivated by the negative externalities of carbon are analogous and well-discussed, yet such taxes are less precise instruments than taxing the carbon content of fuels directly, because consumers may switch fuel types. Even without this substitution in SSB consumption, if the reduction in consumption happens predominantly for low-sugar SSBs, the policy’s effect would be somewhat diminished, as was found earlier for tobacco taxes (Adda and Cornaglia, 2006, 2013). The other sin taxes with which governments have experimented in the past do not exhibit such a pattern.

In the first part of this article, I present a theoretical model for differentiated products using a simple Hotelling model on consumption. The model extends from Kind et al. (2013) and is applied to SSBs differentiated by sugar density to illustrate how different soda tax designs influence consumer behavior. Soda taxes on volume affect SSB unitary sugar prices differently across the sugar density spectrum and thus affect demand differently. This is intuitive: keeping

---

\(^2\)Data source: Census of Government Finance, U.S. Census Bureau.

\(^3\)The Berkeley Soda Tax charges 1¢/Lq(oz) at the distribution level on beverages with added sugars.
volume constant, the tax per unit of sugar is higher for SSBs with lower sugar content. In the main set of results, this paper disentangles the effect of SSB liquid ounce consumption from the consumption of sugar in grams. I find that SSB consumption in liquid ounces falls by more than sugar intake from the same set of beverages. I find a decrease of 15.6% in SSB volume consumption, but only 13.4% in SSB sugar content. This difference constitutes a leakage of 14.1% in sugar intake reduction caused by the design of the policy. I also find that pre-tax sales prices of SSB barely rose by 1.27% on a per ounce basis, which represents a pass-through of only 7%; there were no changes in per gram of sugar prices in average. This explains the mechanism through which consumers are incentivized: a change in the relative price of low vs. high sugar content beverages makes the former more expensive.

This article makes several contributions. This is the first article to evaluate the Berkeley soda tax using Nielsen Data, following 11 stores in Berkeley and over 2000 in the control localities, with information from a set of more than 2,500 UPCs. In this regard, previous studies using scanner information were limited to a handful of stores at best, and not more than 20 UPCs. Under this comparison, our evaluation of the policy will better approximate the real effects of the introduction of a soda tax in Berkeley. Furthermore, this article provides a theoretical framework for analyzing the design of a tax levied at the volume level without relying on the size of the externality to demonstrate that such a tax distorts consumer preferences in favor of SSBs with higher sugar concentrations. Additionally, I carefully construct control groups that are not tainted by the policy, as most available studies estimate effects from comparisons with neighboring localities affected by cross-border shopping and different levels of economic integration. Finally, I focus on how the evidence from Berkeley can inform efforts to influence consumer behavior using the tax code, but I stress that as societies, we need to better understand the patterns of demand to design better policies.

The article is structured as follows: after this introduction, I provide a brief review of
the literature evaluating soda taxes. Section 2.3 develops the theoretical model, and Section 2.4 presents the data sources. Section 2.5 presents my identification strategy and discusses how I conduct the statistical analysis. The Main Results section presents my estimation of the policy’s effects and addresses several potential challenges to the validity of my findings. Finally, Section 2.7 discusses these results and their policy implications and concludes.

2.2 Literature Review

Previous empirical evidence from retail sales taxes (Poterba, 1996) has shown that consumption taxation has a near one-to-one pass-through to consumers. However, in the case of alcohol and tobacco taxes, research has found that while consumption decreases, that decrease is accompanied by over-shifting, in which after-tax prices increase by more than what would theoretically be expected by the introduction of the tax (Hines, 2007; Dutkowsky and Sullivan, 2014; Shrestha, 2015; Shrestha and Markowitz, 2016). Why this empirical anomaly is observed is unknown, however, this has not been the case in soda taxes so far.

Previous sin tax studies in the United States focus on consumption and incidence of taxes on tobacco and alcohol (Fletcher, 2012; Dutkowsky and Sullivan, 2014; Shrestha, 2015; Shrestha and Markowitz, 2016). Before the Berkeley tax, most of the direct evidence on the effect of soda taxes came from international contexts. For instance, Colcher o et al. (2016) analyze Mexico’s soda tax, and Berardi et al. (2016) evaluate the soda tax in France. In general, this early evidence finds that the introduction of a soda tax leads to a decrease in consumption as prices rise, with a complete or near-complete pass-through of the tax burden to consumers. More recently, Vall-Castelló and Lopez-Casasnovas (2020) have evaluated the Catalonia soda tax, which has differentiated brackets for sugar content. They find that consumption falls even more so for beverages associated with higher tax rates, as this article suggests.

Several theoretical studies motivate the introduction of sin taxes under settings in which
rational behavior is limited by self-control problems in consumption (O'Donoghue and Rabin, 2003, 2006; Haavio and Kotakorpi, 2011; Allcott et al., 2019b; Grummon et al., 2019). These studies establish models in which consumers vary in their degree of self-control, as measured by the parameters in present-biased preferences that discount inter-temporal consumption. They provide theoretical justification for the introduction of sin taxes, in which governments try to help citizens contain their self-control problems. They also provide an explanation for the high elasticities implied by the empirical literature. The aforementioned studies, as well as Cawley et al. (2019b), also consider the externality in excessive caloric intake as treatment for diseases paid for by resources from a third-party pool: either taxpayers or insurance.

In the U.S., the Berkeley soda tax of 1¢ per liquid ounce, enacted in 2015\(^4\), has been the subject of several analyses. Falbe et al. (2015) document that three months after the implementation of the policy, 67% of the tax was passed to soda consumers, and 47% to other SSB consumers on average. They conduct a study following price tag data across 71 stores in and outside of Berkeley. Falbe et al. (2016) conduct a cross-sectional survey comparison before and after the introduction of the policy and find that SSB consumption measured in liquid content decreased 21% in the city of Berkeley after the introduction of the tax, while it increased 4% in neighboring cities. Cawley and Frisvold (2017) find a pass-through of 43% and document how proximity to the border affects the effectiveness of the tax. Using POS retail information, Silver et al. (2017) find that volume consumption of SSBs fell 9.6% on average, whereas the pass-through to consumers varied depending on the kind of retailer and product. Their reported pass-through for scanner data was 67%, whereas for price tags it was close to or greater than 100% in some cases. Finally, Lee et al. (2019) finds a self-reported decrease in consumption of up to 44%, based on the reported times per day that survey respondents

\(^4\)The tax in Berkeley applies to beverages that have added sugars, including juices that are not fresh, sodas, and flavored waters. Low-calorie beverages without sugar content, fresh natural juice, and dairy products were exempted. A full description of the policy can be found at this link: www.healthyberkeley.com/about-berkeleys-tax-ordinance
drank soda.

Studies from other cities that have enacted soda taxes corroborate the correlation of increased prices and reduced SSB consumption with even more success. With higher rates, pass-through and consumption effects are more pronounced. Cawley et al. (2018c,b) find nearly 100% pass-through for Philadelphia, and Cawley et al. (2018a) finds 78% in Boulder; both are cities with higher tax rates than Berkeley. On quantity consumed, Cawley et al. (2019a); Zhong et al. (2018), and Roberto et al. (2019), find that consumption of SSBs in Philadelphia decreased by between 30 and 50%.

With the exception of Cawley et al. (2019a), no study has investigated soda taxes’ effectiveness in reducing sugar intake, or if the populations affected are also populations at risk of contracting non-communicable diseases caused by caloric over-consumption. To date, the most complete empirical evaluation is Cawley et al. (2019a), which reports how Philadelphia’s soda tax affects consumption among children and minorities. They collect self-reported information on sugar intake and do not find detectable impacts for children, nor consistent effects for demographic groups considered at risk.

Other studies have pointed to the design problems inherent to taxes based on liquid content. Grummon et al. (2019) note how a tax levied on liquid content does not account for the extent of the externality arising from the sugar content of SSBs. Cawley et al. (2019b) follow a similar argument to recommend a tax based on sugar content instead. In a parallel case, a similar outcome was found in tobacco taxation, where consumers shift toward cigarettes with higher nicotine concentration when facing higher cigarette taxes (Adda and Cornaglia, 2006, 2013). Their modeling is not directly applicable to SSBs, however, as the trade-off between nicotine and smoking requires consumers to always have a preference against higher nicotine concentration, which they call smoking intensity. In the case of sodas, on the other hand, the beverages’ sweetness cannot be modeled with a negative component to its demand. Thus, a
model of product differentiation based on sugar densities, as the one developed here, is more appropriate for understanding consumers' behavioral responses. Interestingly, the opposite is found in the alcohol literature, as the price elasticity of demand is more negative for spirits with higher alcohol content (Wagenaar et al., 2009).

This study contributes to the literature in several ways. First, it more directly addresses health concerns by being the first to study sugar content, which society has deemed as harmful, in comparison to the beverages that the policy affects. This article informs how the design of a soda tax might have an unintended consequence, as there is a margin that consumers can exploit to avoid the tax, at least partially. Theoretically, I prove that beyond the externalities of SSBs, taxes designed on volume provide incentives to consume SSBs with higher sugar concentrations. Empirically, I find that the tax reduces sugar intake, but with some substitution toward SSBs with higher sugar concentrations: I show that the reduction in sugar intake from SSBs is lower relative to the decrease in SSB consumption in terms of liquid quantity. This article is also among the few to use POS data, which includes billions of transaction reports for SSB purchases. This makes it possible to think carefully about group selection and not rely on neighboring localities, which are more likely to be affected by cross-border shopping and different levels of economic integration, a problem exhibited in most available studies comparing Berkeley to neighboring localities. Finally, this article also contributes to the literature on nutritional inequality, considering that SSBs caloric excessive contribution tends to be a problem affecting disproportionately the most vulnerable (Allcott et al., 2019a) and that any effort to improve its efficacy might render positive effects over the most affected populations.

2.3 Drinks with Different Sugar Concentration: A Model

In this section, following Kind et al. (2013), I develop a simple Hotelling model of product differentiation to analyze consumer responses. Assume that consumers make their decisions
to consume SSB in two stages: first, they decide the budget allocated to SSB as independent of different goods. This outcome can be attained by assuming any demand-separating utility function with constant elasticity of substitution. Once consumers have decided the share to be spent on SSBs, they make the decision of what SSB they will take based on their preferences for sugar density and flavor. Firms offer beverages with specific sugar concentrations. For simplicity of argument, I assume that all beverages come in the same size, so I can concentrate on the per-volume price, which has been found to increase after the introduction of taxes levied at liquid content.

This structure allows for separating the overall decrease in consumption of SSBs, found in the literature and corroborated in this article, from heterogeneous effects given differences in SSBs as a result of their sugar density (and more generally, their flavor). In other words, this model captures the fact that the market shrinks, but also what happens across different taxed beverages as a result of the tax design, in either the volume or sugar concentration dimensions.

Equation 2.1 expresses consumer cost functions. Unitary sugar prices from SSB are represented as $P_H$ and $P_L$, whereas sugar concentration for beverages $L$ and $H$ are $S_L$ and $S_H$. To illustrate the effect of the taxes designed in terms of sugar or liquid content, I assume that one of these beverages has a higher concentration of sugar per can ($S_H \geq S_L$).

$$C_i = f_i(x - S_i) + P_i$$

$$C_i = \alpha_i(x - S_i)^2 + P_i$$

(2.1)

While there is supply across the continuum of SSB sugar densities, I provide evidence that across and within beverage categories, producers tend to offer beverages that bunch in specific sugar concentrations. For instance, sodas with added sugars are offered in cans with 40 grams of sugar per 12 liquid ounce can or very close to this threshold, with 70% of transactions happening in densities of 40+/−3 grams. Sports drinks bunch in products with less than 10
grams per can, or around 20 grams of sugar, whereas beverages from fruit with added sugars include products bunching at less than 10 grams, 20 grams, 32 grams, 40 grams and 45 grams. Figure 2.1 expresses this differentiation based on sugar density, underlining the points of the distribution at which transactions tend to concentrate.

Flavors offer a wide range of product differentiation that is not accounted for in this model. The model exclusively allows for differentiation according to sugar density. Concentrating on the sugar variable allows me to track the unintended effect of a tax levied on liquid content. While simplifying with respect to other dimensions, sugar or caloric content is the most important considering that SSBs are targeted for their caloric contribution to health concerns. To select the beverage closest to their preferences, consumers are faced with a cost function associated with both price and the differential between their preferred sugar density ($x$) and that of the beverage ($S_i$) as expressed in Equation 2.1, where $i = L, H$. This specification expands on Kind et al. (2013), as it relaxes the assumption that cost functions should be equal.
when consumers consume more or less sugar relative to their preferences. Usually the cost function is assumed to be linear or quadratic in its distance to the preferred good to attain closed-form solutions. I partially follow this assumption, as consumers might not be sensitive to these differences when they are rather small, but as distance to their preferred flavor increases, flavor differences become more salient to consumers and they become more aware of the mismatch.

To model different cost functions for consumers who have to consume more (or less) sugar relative to their preferences, I index the parameter $\alpha$ by each available beverage option—though in the continuum of available options, the cost function for a given beverage could be different depending on whether consumers are located to the left or right of the sugar density continuum. Figure 2.2 shows the group of consumers that delimit demand for a given sugar density beverage in a continuum of sugar density. That group of consumers is characterized by indifference toward the two surrounding options.

Figure 2.2: Market share of beverages with different sugar densities

2.3.1 Tax effects

I assume a world with two beverages, $L$ and $H$, considering taxes formulated in three different ways: first, a tax that exclusively affects the beverage with the highest sugar concentration, $H$; second, a tax design on the basis of sugar that affects all beverages' unitary sugar prices
uniformly; and finally, a tax charged on liquid that affects the unitary prices of sugar differently.

Figure 2.3 illustrates the three scenarios under the assumption that initial sugar unitary prices are equal, simplifying the figure as shown below; this simplification does not alter the direction of the results. The figure and scenarios are described in more detail below, but they all have a no-tax framework as a starting point. Consumers with preference for density \( x_0 \) are indifferent between the two drinks as shown by point A in all panels of Figure 2.3.

Figure 2.3: Effects of taxes over beverages with different sugar concentration

A Tax on Sugar for High-Sugar Beverages

The set of consumers who are indifferent between the two beverages without a tax satisfy \( C_L = C_H \) and have preferred density \( x_0 \), as given by Equation 2.2. Hence, all the consumers with preferences in the range \((S_L, x_0)\) will consume beverage \( L \). If consumers’ preferences fall in \((x_0, S_H)\), they will consume beverage \( H \). Moving forward, the set of consumers that are indifferent between consuming the two available beverages when a tax is levied only on high-density sugar beverages also satisfies cost equality as expressed in Equation 2.3; these
consumers also have a preference for sugar density equal to $x_1$.

$$\alpha_L(x_0 - S_L)^2 + P_L = \alpha_H(x_0 - S_H)^2 + P_H \quad (2.2)$$
$$\alpha_L(x_1 - S_L)^2 + P_L = \alpha_H(x_1 - S_H)^2 + P_H + t \quad (2.3)$$

My interest lies in what happens with $x$ after the introduction of the tax in Berkeley. Subtracting Equation 2.2 from Equation 2.3, an expression for $(x_1 - x_0)$ identifies the change in demand between the two beverages as given by Equation 2.4. The first relevant outcome is that initial unitary sugar prices do not matter, as they cancel when subtracting Equation 2.2 from Equation 2.3. Furthermore, $x_1 - x_0 > 0$, as shown in Appendix A.1. Thus, the tax affects demand for $H$ negatively, as consumers substitute toward $L$. Additionally, those who continue consuming $H$ can now afford less of the drink. Panel (A) of Figure 2.3 represents this movement with red dashed lines. The price of $H$ increases from $P_S$ to $P_S + t_1$, and the preference distance cost for $H$ moves parallel; this displaces the new indifference condition from A to B.

$$x_1 - x_0 = \frac{t}{(\alpha_L - \alpha_H)(x_1 + x_0) + 2(\alpha_H S_H - \alpha_L S_L)} \quad (2.4)$$

This framework would fit the effect of the tax relative to SSB substitutes such as water or other untaxed beverages, as found by the previous literature (Falbe et al., 2016; Silver et al., 2017; Cawley et al., 2018b; Zhong et al., 2018; Cawley et al., 2019a).
A Tax on Sugar for All SSBs

To separate from the previous case, the set of consumers indifferent following the introduction of such a tax is labeled $x_2$, and its indifference condition is presented in Equation 2.5.

$$
\alpha_L(x_2 - S_L)^2 + P_L + t = \alpha_H(x_2 - S_H)^2 + P_H + t 
$$

(2.5)

Subtracting Equation 2.2 from Equation 2.5, I obtain an expression where $x_2 = x_0$. Market shares for $L$ and $H$ remain the same. However, overall SSB consumption decreases, as prices have increased while budget remains the same. This would be the ideal situation provided that the tax is structured by sugar (or caloric) content. Graphically, Panel B of Figure 2.3 shows how the prices of $H$ and $L$ increase to $P_S + t_2$, and the preference distance costs for $H$ and $L$ move in a parallel fashion. This displaces the new indifference condition from A to B. Unitary sugar prices have increased, so every individual is able to afford less of each beverage. However, individuals consume the same kind of beverage they consumed prior to taxation, just in smaller quantities.

A Tax on Volume for All SSBs

This setting of a tax on volume for all SSBs is the one that pertains to Berkeley and to most of the soda taxes implemented so far with few exceptions (i.e. UK or Catalonia). While the tax is constant at the volume level, unitary sugar prices are affected differently across beverages. The greater sugar content a beverage has, the lower the effect of a tax on volume over its sugar price: $t_H < t_L$. To illustrate, imagine a can of soda with 40 grams of sugar and a can of another beverage with 20 grams of sugar. In the context of Berkeley’s SSB tax, each can would be taxed 12¢. Sugar unitary prices, on the other hand, are different for each beverage:
the former is taxed $t_H = 12¢/40$ grams and the latter $t_L = 12¢/20$ grams.

$$\alpha_L(x_3 - S_L)^2 + P_L + t_L = \alpha_H(x_2 - S_H)^2 + P_H + t_H$$

(2.6)

The indifference condition is given by Equation 2.6, which is equivalent to Equations 2.3 and 2.5 in the previous subsections, but with the introduction of a differential tax for each beverage. Subtracting 2.2 from 2.6, the expression for the demand change is given by Equation 2.7 and depends exclusively on the differential across taxes charged to both goods considering that $t_H - t_L$, the common denominator found previously, is always positive.

$$x_3 - x_0 = \frac{t_H - t_L}{(\alpha_L - \alpha_H)(x_3 + x_0) + 2(\alpha_H S_H - \alpha_L S_L)}$$

(2.7)

As $t_H - t_L < 0$, a tax on volume affects sugar prices heterogeneously and will make the price of sugar more expensive in SSBs with low (relative to high) sugar concentration. Panel C on Figure 2.3 illustrates this movement. All consumers between $x_3$ and $x_0$, who previously consumed beverage L, will now consume beverage H. The unintended effect of the policy is this push toward SSBs with higher sugar concentrations.

Incorporating the two-step decision, prices of all SSBs are rising, so SSB consumption will decrease both in volume and in sugar intake, since all beverages are less affordable. However, given that some consumers will now prefer beverages with higher sugar concentrations, the decrease in sugar intake should be offset to some extent; this will happen to the extent that the change in after-price taxes materializes with a differential between the unitary price of liquid and the unitary price of sugar. In this article, I call this offsetting leakage resulting from the design of the policy.

The current focus of the literature evaluating soda taxes is on consumer behavior. However, this framework can be extended to analyze producer responses. At first, Berkeley’s SSB tax
will shift consumers toward untaxed beverages. But between taxed beverages, producers have incentives to change their formulas toward higher sugar concentrations. A better design could make use of this result by levying the tax at SSB sugar concentration, incentivizing producers to change their formulas. If introduced gradually, such a measure—where $S_H > S_L$—could protect industry profits, as companies would be incentivized to reduce sugar inputs to protect market shares. This has been the case with the United Kingdom’s soda tax, where tax changes are based on sugar content and where some producers have announced changes in sugar content for certain beverages (Dewey, 2018). Still, this potential effect remains understudied.

2.4 Data

2.4.1 Nielsen Data

I use information from the Nielsen\(^5\) Retail Scanner Data (Retail Scanner). The dataset captures information from scanner-based points-of-sale in participating stores across the U.S. I pull product categories that contain UPC codes affected by Berkeley’s SSB tax: soft drinks, canned or bottled beverages, juices, and carbonated and non-carbonated beverages. Whether they are affected by the policy depends on their added sugar content, as some of these beverages contain naturally occurring sugars, which are not affected by the tax.

The data has some limitations. First, it only provides information about purchases from retailers, and does not include expenditures at other types of establishments. Furthermore, households might consume foods they did not purchase or purchase foods they do not ultimately consume. However, this is a reasonably representative approximation considering the volume that Nielsen reports\(^6\) and the share of at-home consumption of soft drinks in volume,\(^7\)

---

\(^5\)Results in this chapter are calculated (or derived) based on data from The Nielsen Company (US), LLC, and marketing databases provided by the Kiilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the Nielsen data are those of the researchers and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

\(^6\)According to the AEA, Nielsen accounts for more than half of total sales at grocery and food stores in the U.S. and about a third of sales at mass merchandisers. Information available at: [www.aeaweb.org](http://www.aeaweb.org)
Table 2.1: Data Sources Information

<table>
<thead>
<tr>
<th></th>
<th>Observations (Mlls)</th>
<th>UPCs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Retail</td>
<td>No.</td>
</tr>
<tr>
<td>Total</td>
<td>2,428.6</td>
<td>60,455</td>
</tr>
<tr>
<td>Control Brands</td>
<td>518.1</td>
<td>15,243</td>
</tr>
<tr>
<td>Identifiable</td>
<td>1,910.1</td>
<td>45,312</td>
</tr>
<tr>
<td>From Identifiable UPCs successfully scrapped</td>
<td>1,350.5</td>
<td>2,580</td>
</tr>
<tr>
<td>Western U.S. sample</td>
<td>417.6</td>
<td>1,655</td>
</tr>
</tbody>
</table>

Note: Observations from the time span July 2013 to Dec 2017.

which includes retail data accounting for more than 70% of the market\(^7\).

Retail Scanner information is nevertheless a substantial improvement over self-reported information, as reporting from individual stores comes directly from their POS, and the transfer of information comes at minimum cost to the participating stores. Therefore, accuracy of reporting certain items is not an issue. Furthermore, each of the stores in a locality encompasses more information than any individual followed by consumer surveys or panels of households in any previous study. I exploit the overlapping geography of the 3-digit zip code that contains Berkeley to identify treated stores and compare them with alternative control groups to derive results.

2.4.2 Nutrition Facts

I obtained sugar information by scraping Nutritional Facts tables from the USDA Food Central dataset\(^8\) and from online stores that provide that information. From the UPCs provided in the Nielsen data, I constructed the commercial UPC by including the verification number and any remaining zeros in order to obtain complete 12-digit UPCs. Ultimately, I found sugar content information for 2,580 UPCs that feed 71% of the reports from the retailers’ information.

The resulting dataset contains the following product groups containing carbonated beverages, juices, and other non-carbonated soft drinks, filtering diet beverages, which are are

\(^7\)Statista report on Soft Drinks market trends in the U.S. available at statista.com
\(^8\)Available at https://fdc.nal.usda.gov/
easily identifiable. Table 2.1 illustrates the construction of the analysis data. There are 2.4 billion transaction reports including SSBs from 60 thousand UPCs. Discarding transactions of Control Brands, which are unidentifiable in the USDA/online store scrape, 79% of the transactions and 75% of the UPCs are identifiable. The scrape successfully recovers Nutrition Facts for 2,580 UPCs, 6% of listed UPCs. However, these UPCs represent 71% of the reports, with 1.35 billion transactions. The final dataset contains observations concentrated in the Western U.S., which keeps 31% of the dataset and 65% of the identified UPCs, as not all are sold across the nation.

2.5 Methodology: The Effect of the Policy

Previous studies approach causal effect estimation of soda taxes by introducing a control group in a difference-in-difference setting. I start by corroborating their results in my sample of stores by analyzing how outcomes vary after introduction of the policy. I track liquid quantities and unitary prices, but more importantly, sugar intake from SSBs. To identify changes in consumption attributable to the policy, I estimate the following model:

\[ y_{st} = \beta(D_s \cdot Post_t) + \eta_t + \eta_s + \epsilon_{st} \]  

The policy’s effect in a given month on store-level sales of SSB log quantities of liquid and sugar \( y_{st} \) is captured by the coefficient \( \beta \), as it states the difference between the treatment \( (D_s = 1 \text{ for Berkeley stores}) \) and control groups after the introduction of the policy \( (Post_t = 1 \text{ for months after March 2015}) \). As mentioned before, during the policy’s introduction, there was a transition period in which not all stores implemented the policy. Those stores are not included. To capture any time-invariant difference across stores or geographic locations, store

---

9To protect retailers’ identities, Nielsen modifies the UPCs from retailers’ brands; in the brand description they are called Control Brands. There are no Control Brand transactions within Berkeley for the modules used in this analysis.
\( (\eta_s) \) fixed effects are used: they fully capture the status of treatment and control groups. I also use period fixed effects \((\eta_t)\) to account for any seasonal effects of period-specific changes in SSB sales quantities. Standard errors are clustered at the store level, the unit of analysis.

The comparison between control and treatment groups provides the identifying variation. Thus, I consider six different groups: the first considers neighboring localities to Berkeley; the second considers the remaining 3-digit zip codes in California. A separate control group includes other 3-digit zip codes that are different from those neighboring Berkeley. A fourth disregards stores in the Bay Area. The last two controls are constructed with neighboring states, which are one or two states away. By separating neighbor groups, I can control for any cross-border effects between the localities, other confounders from economic integration, or even effects that originate in the level of public attention given to the matter. I evaluate the quality of the controls by testing the common pre-trend assumption in a full event model that is a more general specification relative to Equation 2.8; here, I allow the effect of the policy, the interaction of the treatment, and time to vary in each single month. However, considering that everyday life in the San Francisco Bay Area is considerably intertwined, I build a final control group with all zip codes in California outside of the Bay Area. Accordingly, I estimate:

\[
y_{st} = \sum_{t \neq 2014IV} (\beta_t(D_s \cdot \eta_t)) + \eta_t + \eta_s + \epsilon_{st} \tag{2.9}
\]

Equation 2.9 allows for testing the common pre-trend assumption and enables me to characterize the timing of the effects after the introduction of the soda tax when I estimate treatment effects for different periods.

### 2.6 Main Results

In this section I present evidence that a tax levied on volume generates leakage and that the reduction in volume consumption is always smaller than the reduction in sugar intake from
SSBs, the actual target. Using Berkeley stores during 2015 and 2017 as treatment units, I calculate treatment effects of the soda tax’s introduction on quantities and prices; from those treatment effects I in turn calculate a measure of the leakage caused by the tax design based on the ratio of change in liquid consumption to the change in sugar intake caused by the policy.

First, I introduce the different control groups considered and select one to present additional results. However, the selection of control group does not alter the main results as shown in the robustness sub-section. I select the best control based on two characteristics: one, the mean squared deviation from the parallel trend in the pre-policy period and two, on the absence of contagious effects from the policy’s introduction. Next, I present the main results in quantities and prices. In the preferred specification, sugar content decreases by about 13.4% while liquid consumption falls around 15.6%, indicating a shift in consumption toward SSBs where sugar is more concentrated. This leakage of 14.1% also implies an increase of 2.2% in the average sugar density among taxed SSBs. On the other hand, on the mechanism leading to changes in consumption, I find that prices per liquid ounce increased marginally by 1.27%, with a pass-through of only 7%, whereas prices per gram of sugar did not increase, or in some cases even decreased. Next, I present how these effects have changed over time and follow by developing a direct test for leakage, which involves bootstrapping the ratio of the change in liquid ounces to the change in grams of sugar from SSBs. I then present the results for regressions on the average density of sold beverages. The final subsection presents a series of robustness checks.

The evidence I present here suggests that the policy achieved the goal of reducing sugar intake from SSBs. However, reduction in sugar intake does not parallel reductions in liquid content, indicating that consumers react to soda taxes by shifting their demand toward SSBs with higher sugar concentrations. This is consistent with the theoretical model presented in Section 2.3 and with the results from price changes, as price per liquid ounce rose following soda tax introduction, whereas price per unit of sugar did not change, at least during the first
three quarters following the introduction of the soda tax.

2.6.1 Parallel Trends and Control Group Selection

The most important assumptions in a difference in difference setting are the common trend assumption and the strict exogeneity assumptions (Wing et al., 2018). This implies that in the absence of the policy in question, the outcome series for control states will maintain its previous behavior. However, once the treatment affects control groups, this is no longer the case. While the current literature on soda taxes focuses on the evaluation of parallel trend assumptions, in many instances main results are drawn from a comparison with stores or individuals from neighboring localities, whose behavior and environment are affected by the tax introduced by their neighbor. Cross-border shopping has been documented in the previous literature; moreover the attention brought by the media and public forum discussions might increase awareness that might lead to altered behavior, a matter less understood. People in neighboring markets can either stockpile SSBs in anticipation of higher prices due to tightened demand in their markets, or reduce their consumption as health concerns are now more salient. The bottom line is that transactions in neighboring markets are tainted by the policy in Berkeley, so they provide for an inferior counterfactual.

This creates a trade-off for the ideal counterfactual: the closer to the locality in which the policy has been implemented, the better to resemble its untreated behavior. However, it would be desirable to avoid any influence caused by the introduction of the policy in Berkeley over the counterfactual localities. Accordingly, I evaluate the parallel trends assumption for several potential control groups in terms of their mean squared deviations in the pre-policy period following the literature on synthetic controls (Abadie et al., 2010; Cavallo et al., 2013), but also select a group where localities are not affected by the tax introduced in Berkeley. Graphs are organized depending on the geographical proximity of the control group to Berkeley.
Figure 2.4: Event Studies - Sugar-sweetened beverages Log Liquid Ounces for Different Control Groups

Note: 2014 - IV used as reference and not included on the graph.
Figure 2.4 presents the event studies exercise for SSB volume as described in Equation 2.9. Left to right, the top panel presents figures for neighboring localities, the remaining California localities, and California without neighboring localities. The bottom panel shows the graphs for California without the Bay Area localities, neighboring states to California, and Western states up to two states away. The figures also report the sum of squared deviations in the period before the policy introduction, which reflects how the distance across treatment and control varies in relation to the reference period. All the groups presented comply with the parallel trend assumption, and judging by the sum of squared deviations, the closer they are to Berkeley, the better they resemble its pre-trends. However, neighboring localities are also affected by the policy. Accordingly, the neighboring states control is the preferred specification, as it is mostly certain, in relative terms, that its localities are not affected by the policy. This selection holds across all the series analyzed in this article. The respective parallel trends for the different control groups are included in Appendix A.2.

Event studies figures also reflect the effects’ experience of heterogeneity over time. Liquid and sugar reductions are consistent over time, although depending on the control group, there is some return to pre-policy levels. Price effects at either liquid or sugar levels do not show consistent positive pass-through, although the last quarters of 2014 and 2015 exhibit consistently higher prices in several specifications. Across all series, with the exception of sugar density per liquid ounce, confidence intervals widen. I will analyze this result in more detail in the following subsections.

2.6.2 Changes in SSB Consumption and Prices per Unit

Table 2.2 presents the difference in difference estimation of treatment effects for 2015 and 2016 using transactions in stores from neighboring states to California as a counterfactual. SSB consumption decreases by 15.6%, which is in the range of previous findings (Falbe et al.,
This means that each month following the introduction of the tax the average store in the Nielsen data has sold 1,840 fewer standard 20-liquid-ounce SSB bottles. On the other hand, sugar intake from SSBs only decreased by 13.4%, or 90 Kgs. The relationship between these two estimates constitutes the main result of this article, as this difference is a leakage resulting from the policy design of levying the tax at the volume level. The ratio between the change in sugar and volume consumption caused by the introduction of the soda tax is 85.9%. If there was no leakage, this ratio should be close to 100%. This gap of 14.1% is the measure of leakage I use here.

Table 2.2: Difference in Difference Treatment Effects - Control: Neighboring States

<table>
<thead>
<tr>
<th>ln(Quantities)</th>
<th>ln(Prices)</th>
<th>Sg. Density</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lq.oz</td>
<td>Sg. gr</td>
<td>cents/lq.oz</td>
</tr>
<tr>
<td>-------------</td>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>$D_{st} = 1 \cdot P = 1$</td>
<td>-0.156***</td>
<td>-0.134***</td>
</tr>
<tr>
<td>(0.0381)</td>
<td>(0.0427)</td>
<td>(0.0158)</td>
</tr>
<tr>
<td>Baseline - $D_{s,2014}$</td>
<td>236K lq.oz</td>
<td>0.68 Tons</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.985</td>
<td>0.985</td>
</tr>
<tr>
<td>N Clusters</td>
<td>2,354</td>
<td>2,354</td>
</tr>
<tr>
<td>Transact. Reports</td>
<td>114.0 Mill</td>
<td>114.0 Mill</td>
</tr>
</tbody>
</table>

Standard errors clustered at the store level in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Changes in prices per unit of liquid and sugar are not statistically significant $^{10}$. The increase of 1.27% in the price per ounce constitutes a pass-through of only 7%. This is substantially lower than what other studies have found for Berkeley. For instance, Cawley and Frisvold (2017) document a pass-through of 43% on average across different products’ price tags, and Silver et al. (2017) document pass-through levels of 69% and 46% for small SSB packages and 2 liter bottles, respectively, without including bulk packaging prevalent in mass merchandisers or chain grocery retailers. Those calculations are different than the approximation here, which relies on actual transactions instead of price tags. Hence, price tags may increase across the board, but the final average unitary price might not change because consumers react strategically to the tax and shift their consumption from higher- to

$^{10}$The tax is levied at the distribution level, so all shelf and transaction prices are tax inclusive
lower-price SSBs, for instance by buying bulk packages of 24 or 36 units rather than individual cans or bottles. This would make the final transaction price per ounce similar to pre tax levels even if prices are increasing across the board. The mechanism remains outside of the scope of this article and can also be considered to improve the policy design; what matters for the argument, however, is that in all specifications, the change in prices is higher for liquid content than for sugar content, making SSBs with higher sugar concentration more expensive as required by the modeling in Section 2.3. Regardless of what the control is, or what years are included, this relationship remains.

Finally, and reinforcing the hypothesis that the tax creates an incentive to consume beverages with higher sugar concentrations, Table 2.2 presents the effect of the tax over the average sugar density of sold SSBs. After the introduction of the tax, the sugar density of the average purchased drink is 2.2% higher than the average pre-tax. It represents additional sales of 4.5 Kgs of sugar per month in the average store in Berkeley on the remaining SSB sales. If the tax was levied at the sugar content instead, this mismatch most likely would not occur.

2.6.3 Effects Over Time

Figure 2.5 shows how the policy effect on the analyzed outcomes has changed over time from 2015 to 2017, using a modified version of Equation 2.9. The top graphs in the figure present quantities and prices at the bottom. Relative to the entire pre-policy period, each year after the introduction of the policy presents a fairly persistent reduction ranging in SSB volume from 15% to 16% in monthly sales. A similar situation is found in sugar intake from SSBs, where the reduction ranges between 13% and 14%. In 2017, for both series, the policy’s effect is slightly smaller and the standard errors are considerably wider, to the point that the change in sugar intake in 2017 is not statistically significant. With regard to price changes, the price per ounce of SSB had positive increases in 2015 and 2016 relative to the entire pre-policy
period, reaching a change of up to 2.3% in 2015 and 2% in 2016 (pass-throughs of 12.6% and 10.8%, respectively). In 2017, however, this price returned to the pre-policy level. This does not necessarily mean that individual UPC prices remain unchanged, but considering all offsetting substitutions, the final price per liquid ounce has not changed: consumers who can afford to do so may be partially avoiding the tax by purchasing bulk versions of the beverages they prefer, or they may be substituting in any other pattern that partially allows them to avoid the price hikes. The tax effect over the SSB price per gram of sugar has never been positive on average, and following SSB volume prices, the effect during 2017 also decreased substantially.
2.6.4 A Direct Test for Leakage and SSB Relative Price Changes

The measures developed for leakage and relative price change rely on the comparison between two estimated parameters in each case. This subsection presents a semi-parametric approach to formally test the hypotheses comparing those parameters. Separately bootstrapping treatment stores 1000 times and comparing their SSB sales to sales from bootstrapped control stores, I construct a distribution for the measure of leakage and for the difference in unitary liquid and sugar prices. The left graph in Figure 2.6 presents the distribution of the ratio of the change in SSB volume consumption to the change in SSB sugar intake. The actual ratio of 85.9% (with a leakage measure of 14.1%) is represented by the bold dashed line, and percentiles 1 and 99 are represented by the thin dashed lines. Leakage is observed in more than 99% of cases, with over 90% of iterations exhibiting a leakage greater than 10%. The test for the comparison across changes in price is presented in the right panel of Figure 2.6. In this case, I evaluate the fact that price changes in SSBs by grams of sugar are always lower than price changes in SSBs by liquid ounce. Again, percentiles 1 and 99 are represented by the thin dashed lines, and the thick dashed line indicates the observed difference. In more than 99% of iterations, the observed difference is positive.

The distributions generated provide further evidence that the policy incentivized consumers
to purchase beverages with higher sugar concentrations because it changed relative prices
between sugar and volume, making beverages with lower sugar concentrations more expensive.
In the next subsection I address some potential threats to the validity of the findings.

2.6.5 Robustness Exercises

Inference for Small Number of Groups Treated

In Subsections 2.6.1 to 2.6.4, standard errors are clustered at the store level, as they represent
the level at which prices are set at independent or chain stores (DellaVigna and Gentzkow,
2019). I also ran the previous set of results with standard errors clustered at the 3-digit zip
code, the geography level reported in the data, and the geography level at which the policy
is implemented. Under this clustering level the standard errors shrink, making the results
presented in the previous sections more conservative. On the other hand, the parallel trends
assumption for different control groups exhibits more violations relative to the figures presented
here in Subsection 2.6.1 and in Appendix A.2. However, the best comparison is still supplied
by the neighboring states.

Given that the data has only one locality as treated with 11 stores, I further consider
the standard errors (Conley and Taber, 2011; Bertrand et al., 2004). Accordingly, evaluating
placebo introduction of soda taxes in different localities than Berkeley, this subsection assesses
if the captured effect in the presented regressions differs from that represented by the distrib-
ution of placebo effects. Figure 2.7 presents the distribution of placebo effects among the
different localities encompassed by three-digit zip codes in the control group. The thin dashed
line encompasses the 95% confidence interval, and the thick dashed line points to the actual
estimated effect. For quantities, the changes caused by the policy are different than those
exhibited randomly in control localities. On the other hand, in the case of unitary prices, nei-
ther volume nor sugar exhibit changes different than those produced by chance in the control
In this regard, of the outcomes followed in this article, the average sugar density exhibits the greatest difference with respect to those observed by chance in the control localities. Figure 2.8 presents the results for the placebo distribution of effects over the average sugar density sold in stores. The observed change in density is greater than any placebo effect observed in the control localities.

Control Group Selection

One potential threat to the validity of the argument is that the results presented previously are sensitive to the selection of the control group. While control group selection affects some of the parameters, the relationship is robust, SSB volume consumption always decreases at a
higher rate than SSB sugar intake, and the difference in unitary price changes between volume and sugar are always positive, so the change in price per ounce is always higher than the change in price per gram of sugar.

Table 2.3 provides evidence of this for all control groups considered in in Subsection 2.6.1 and in Appendix A.2. Comparisons made with other localities in California lead to measures of leakage exceeding 20%. A similar situation occurs when focusing on prices. When California localities are the counterfactual, the average unitary price change for both volume and sugar from SSBs is negative, which underscores the importance of offsetting substitutions in the final consumption versus price tag changes as pointed out before. On the other hand, when compared to localities outside of California, the change in volume prices is positive, while the change in sugar prices remains negative. Regardless of what counterfactual is used, the difference between the two is always positive, so when an exercise points to a negative change in volume prices, sugar prices are even more negative.
Table 2.4: Difference in Difference Treatment Effects Prices - Different Controls

<table>
<thead>
<tr>
<th></th>
<th>Neighbors</th>
<th>Other CA</th>
<th>CA - No Neighbors</th>
<th>CA - No Bay Area</th>
<th>Neighbor States</th>
<th>Western States</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) ln(Lq.oz.)</td>
<td>-0.0117</td>
<td>-0.0096</td>
<td>-0.0874</td>
<td>0.0127</td>
<td>0.0098</td>
<td></td>
</tr>
<tr>
<td>(2) ln(Sg.gr.)</td>
<td>-0.0265</td>
<td>-0.0264</td>
<td>-0.0260</td>
<td>-0.0996</td>
<td>-0.0096</td>
<td></td>
</tr>
<tr>
<td>(1)-(2)</td>
<td>0.0148</td>
<td>0.0168</td>
<td>0.0170</td>
<td>0.0173</td>
<td>0.0223</td>
<td>0.0194</td>
</tr>
</tbody>
</table>

* p < 0.1, ** p < 0.05, *** p < 0.01

A final word with respect to robustness exercises and sample selection. Something that might concern the reader about the results presented above may involve the selection of goods in the analysis sample as not all the UPCs in the Nielsen data are identifiable. In this regard, it is worth mentioning again that there are no transactions in Berkeley of beverages under control brands. Furthermore, to my knowledge, this evaluation encompasses the broadest possible selection of UPCs, with more than 1,655 in the analysis sample and more than 417 million transactions in the western U.S. I performed the first draw of Nutrition Facts information on online stores\textsuperscript{11} that provide this information. I conducted the second draw after the launch of the USDA Food Data Central; the data used here combines the two draws, giving prevalence to the information provided by the USDA in case a UPC is repeated. The three exercises from two allegedly independent sources, lead to almost identical results, with slight variations in quantity levels and price changes, but robust findings on the leakage described here and on the relative price change that makes low-sugar SSBs more expensive\textsuperscript{12}. The average UPC described in the final dataset is 53.6 oz with 156 grams of sugar for a density of 2.96 grams of sugar per liquid ounce. Sugar concentrations on the other two rounds are almost identical: 2.89 grams per ounce for the scrape of online stores and 3.05 grams per ounce for the USDA-provided information.

\textsuperscript{11}I conducted Python self-implemented webscrapes on disruptit.com and drew some manual scrapes from the Google store and ShopRite

\textsuperscript{12}Results are available upon request.
2.7 Discussion and Conclusion

The central finding of this paper is that the SSB tax in Berkeley was successful in curbing volume consumption, but because of its design, it was less successful in curbing actual sugar intake from SSBs. The one cent per liquid ounce tax introduced by Berkeley barely increased the price per liquid ounce but did not increase the price per gram of SSB sugar content at all or only indirectly. If the aim of a policy is to reduce sugar intake, then taxes ideally should be levied on sugar content rather than volume. This better aligns the policy with the motivating medical research that links diseases like obesity and diabetes with sugar consumption (Ogden et al., 2014; Han and Powell, 2013). A tax levied at sugar concentration would not distort relative prices of low-sugar beverages, which in the current setting become more expensive relative to beverages with higher sugar concentrations.

This article contributes to the understanding of how policy can influence people’s behavior using the tax code. Previous findings suggested that setting taxes in higher levels of government, like the state or federal level, would minimize cross-border shopping. Analogously, if as societies we are interested in the most effective ways of influencing people’s behavior, policy design needs to be closest to the problems we want to address. In this case, setting a levy on sugar content would make soda taxes more effective in curbing sugar intake.

The results here are complementary to the current literature that finds decreases in consumption and increases in prices after the introduction of the tax in Berkeley. With that caveat, this study finds lower effects over both quantities and prices than previously reported. An evaluation of the policy over a more representative sample that focuses on actual transactions instead of price tags, allows us to observe that consumers change their bundle of SSBs consumed to strategically avoid the soda tax. A broader sample demonstrates that the consumption effects are lower than previously reported. Furthermore, consumer reactions also lead to a situation in which individual UPCs might increase their prices, but by means of
substituting across UPCs (not necessarily across SSBs), the relative unitary prices in final transactions do not exhibit much of an increase in terms of sugar content. In this article, I focus on the substitution across taxed beverages with different sugar concentrations, as that is the main health concern that the policy aims to tackle. However, there might be another myriad of strategical behavioral responses, some of which the policy might consider in its design. Though they escape the scope of this analysis, some potential margins of substitution that could be subject of further investigation in this regard are bulk packaging and different kinds of sugar sweeteners, among others.

This article demonstrates that the design of the Berkeley soda tax erodes the effect of the policy by close to 15% (and up to 20%) relative to volume consumption. This paper also introduces an additional theoretical motivation for sugar levies, as the current design changes the relative pricing of sugar across SSBs in favor of high sugar concentration beverages. This creates an incentive within the taxed beverages for consumers to drink SSBs with higher sugar concentrations.

These results are robust to a number of challenges. First, leakage and relative price changes hold across any control group analyzed. Furthermore, I present different exercises for inference, including the construction of placebo treatment effects, to see if the Berkeley tax differs from pure chance variation. Finally, the entire set of results is available for three draws of UPCs according to how the databases were created. All three exercises point in the same direction: quantities of volume decreased by more than SSB sugar intake, the tax changed the relative price of low versus high sugar content beverages, and final taxed beverages had higher sugar densities.

This is not the first article to refer to the volume levy as a problem of tax design; several authors have suggested that the tax should be designed on the basis of the externality it causes (Cawley et al., 2019b; Grummon et al., 2019; Allcott et al., 2019c). However, this is the first
article that demonstrates its implications empirically and shows that the decrease in sugar content is smaller than the overall consumption change.

Accordingly, some countries have started designing their soda taxes with a schedule that differentiates across different sugar concentrations. For instance, the UK and Catalonia have done so. The tax could potentially be designed to incentivize producers to change their formulations and move the whole industry in tandem, minimizing the market impact of the tax.
Bibliography


Chapter 3

Cash Subsidies to improve Student Loan programs for Low-Income Students

Abstract *

Every year, governments allocate considerable amounts of resources to support students’ access and performance in higher education. Grants and loans are directed to those students for whom college would seem unaffordable otherwise. These programs are justified on the grounds of how they help students access, persist and graduate from higher education. However, considering the time it takes to observe academic or labor market performance, the literature evaluating student loans and its additional embedded programs is scarce.

This article wants to contribute filling this gap. Using a Regression Discontinuity Design (RDD), I evaluate a cash subsidy program embedded into the main governmental student loan program in Colombia. I exploit the allocation rule of these additional cash subsidies, to evaluate the effect of these transfers over dropout and graduation rates, and also over early default rates. My results indicate that the cash subsidies have substantial and significant effects on dropout rates at the cutoff point. Over graduation rates the effects are substantial, but barely significant. With respect to the default rates, cash subsidy beneficiaries tend to have lower rates at the cutoff, but this result is not statistically significant. However, the measure of the subsidy effect on the latter two outcomes is affected by the statistical power loss at the proximity of the cutoff scores, a limitation of the RDD.

---

*The author thanks the World Bank’s Education team in Bogota, the Ministry of Education of the Republic of Colombia, and Icetex for access to data and for the valuable comments. Finally, I appreciate the comments made by Seth Friedman and the participants of the 2018 ASPS Conference.
3.1 Introduction

Skill-biased technological change has been a global phenomenon occurring during the last 50 years. Katz and Murphy (1992), found evidence of how this change reflects on the different subsectors’ labor demand and the wages they pay. These changes have deep implications on the return on investments in education. Production automation makes unskilled labor disappear, and this has made this type of labor relatively abundant depressing its wages. The opposite occurs with skilled labor, the technicians and professionals that create, interpret and implement technological change. Skilled workers are the ones who can translate technological change into productivity increases and thus enjoy higher wages. In addition to these phenomena, repetitive processes can also be externalized and unskilled labor competes with foreign workers for jobs that cannot be automated, but can be performed easily in any place in the world.

Governments and the societies they represent have two options. The first is to defend the jobs for unskilled laborers with low productivity, and be left behind. The second is to try to adapt to the knowledge society where technological change is a constant and thus, prepare their citizens as much as possible for a world in which capital is highly mobile and jobs with high-productivity and high wages require technical abilities and a dynamic and constant preparation. Additionally, economic growth induced by technological change has produced a degree of inequality that concerns societies and policy makers since it deepens income and wealth gaps between citizens.

To promote prosperity and also improve the situation of lack of opportunities for the most vulnerable, governments are trying to develop policies that ensure access to higher education. The government’s role is justified in terms of efficiency and equity. One of the values of education is to prepare citizens to develop the necessary abilities required by the labor market. It also finds justification in the fact that some population segments have experienced historical or socioeconomic disadvantages that does not allow them to take advantage in equal terms.
of a unified access and retention policy relative to privileged citizens (OECD and BIRF/WB, 2013). These individuals’ talent is lost to the extent that they cannot develop their potential when they are unable to access higher education or must dropout due to a lack of resources (Friedman, 1955; Becker, 1994; Palacios, 2004).

The proper functioning of the higher education system, in general, is one of the goals of these policies and in fact, governments are moving from initiatives that only help access to, toward policies that help student access, persist, as well as successfully graduate from higher education (Bowen et al., 2009; Bettinger et al., 2013). The reason for this is that, at the end, most of the increase in wages and productivity seem to be associated to degree attainment (Jaeger and Page, 1996), although human capital accumulation also plays a role, at least in the Colombian context (Arteaga, 2017). To achieve the proper functioning, we are referring to the fact that policies must be evaluated frequently to guarantee that the public resources are achieving the desired objective, and also to discover which mechanisms determine what works better in different circumstances. Unfortunately, the academic literature offers few evaluations that provide a causal interpretation on the effect of the student loan programs, on the different academic or labor performance outcomes that we would expect student loans to have an effect on theoretically (Dynarski, 2015; Solis, 2017; Marx and Turner, 2019a,b). On the other hand, student aid programs that reduce the costs of attending education have been evaluated mainly in terms of access, with some evaluations on long-term variables such as persistence (Page and Scott-Clayton, 2016). These policies do not tend to be analyzed jointly because they tend to be designed separately.

These types of evaluations would inform the decision-making in the distribution of scarce resources in the different instruments that governments have to invest in higher education. If policies are designed that help students according to their financial needs, the use of different instruments can help obtain better results and a better use of resources. For example, if
most low-income students desert, offering them a subsidy can improve the possibility of a
debt repayment that otherwise seems unpayable. The final question is an empirical matter
that is not only limited to student loans and subsidies, but to all the strategies that support
low-income student access and retention in higher education. Additional subsidies, emotional
or academic support programs, improvements in sources of information to motivate changes
in behavior, among others; all these policies overlap and interact with different implications
for the beneficiaries (Page et al., 2017).

In this article, I concentrate on the instruments of student aid, and in particular on cash
subsidies that low-income students receive when they take out an educational loan in Colombia.
Cash subsidies help students that receive them reduce their cost of entry and retention in
higher education and student loans remove additional financial barriers to entry (Page and
Scott-Clayton, 2016). In particular for our case, regarding the achievement of the desired
academic outcomes and sustainability of loan programs, if implementing an income based
repayment scheme proves institutionally costly, a mixture of loans and grants could improve
loan affordability for low income students.

Taking advantage of the assignment mechanism of these subsidies, which uses the SISBEN
III score (socio-economic characterization tool in Colombia), we can take advantage of the
exogenous variation to use the Regression Discontinuity Design (RDD) (Angrist and Pischke,
2008). We evaluate the cash subsidies impact, local to the allocation cutoff scores, on dropout,
graduation, and default rates by the start of the repayment period. Thus, this article con-
tributes to the literature on student aid by evaluating cash subsidies, included as an additional
aid to student loans.

The results indicate that the cash subsidies for beneficiaries of student loans have a sub-
stantial and significant effect on dropout rates, local to the allocation cutoff points. For the
2011 and 2012 cohorts, we found that at these points, those students eligible for a subsidy on
average, have dropout rates 6.6 percentage points (ppts) lower than those students that have higher SISBEN scores. When estimating the effect on beneficiaries (LATE), the reduction was of 23.5 ppts. If we consider that the dropout rate for those not eligible for the subsidy at the cutoff score is 39%, the cash subsidy decreased dropout rates to less than half (16.5%) at the cutoff point. This reduction is over 90% for first-year dropout rates. On graduation rates, the effect is marked and statistically significant at a confidence level of 10% in our favorite specification, with an increase of 60% on the graduation rate for subsidy beneficiaries. With respect to default rates, subsidy beneficiaries tend to have lower default rates, but the result is not significant. Measuring the effects on the latter two outcomes is affected by the loss of statistical power, first as repayment and graduating cohorts are less numerous, but also given the sample reduction in the proximity of the cutoff point, a limitation of the RDD (Page et al., 2017).

After to this introduction, the article has the following structure: the following section present the literature review of the most recent studies on measuring the impact on student loan program and on auxiliary programs that reduce the cost of attending higher education. Next, the framework on which the educational loan policy in Colombia was created, including the introduction of the cash subsidies. Section 4 presents the methodology of the analysis and the data used. Section 5 present the main results of the impact of the cash subsidy on dropout, graduation, and default rates. The last section presents the main conclusions and a discussion on the results in a general context.

3.2 Literature Review

The motivation for government intervention in the provision and financing of higher education has been theoretically documented (Friedman, 1955; Becker, 1994; Palacios, 2004). In the introduction we mentioned some reasons of efficiency of the government’s function in providing
mechanisms that remove financial barriers to access higher education and the economic literature has produced numerous articles discussing the existence, measurement, and mitigation of this type of barriers (Carneiro and Heckman, 2002; Cameron and Taber, 2004; Lochner and Monge-Naranjo, 2011). Another theoretical motivation of efficiency is the existence of externalities (Palacios, 2004; Page and Scott-Clayton, 2016) or even the existence of another type of barriers, such as information barriers in the process of choosing and processing the relevant options to make the decision of attending and how to attend higher education (Page and Scott-Clayton, 2016; Solis, 2017; Marx and Turner, 2019b).

Page and Scott-Clayton (2016) review the empirical evidence on the causal impact of the policies that seek to reduce the cost of attending higher education and mention several articles that have analyzed the role of subsidies, grants, and discounts on tuition or fees to enroll in a higher education program. The evidence on the effect of this type of policies is conclusive and abundant: reducing the cost of attending a higher education institution causes an increase in enrollment of the beneficiaries of this type of policy (Bound and Turner, 2002; Dynarski, 2003; Stanley, 2003; Abraham and Clark, 2006; Kane, 2007). A smaller amount of studies refers to the casual link between a decrease in the costs of attending higher education and variables of academic results like retention or successful graduation; however, there is also evidence that reducing the cost of attending higher education causes improvement in retention and potential graduation of affected students (Castleman and Long, 2016; Page et al., 2017; Goldrick-Rab et al., 2012). All these studies have reviewed the effect of the policies in experimental environments, with random assignment, or taking advantage some source of exogenous variation.

Analysis on the largest subsidy program of the United States, the Pell Grants program, has yielded more varied evidence. Some studies found that their introduction has not led to better results in retention and graduation (Hansen, 1983; Kane, 1996), while others, more recently,
found positive effects (Seftor and Turner, 2002; Bettinger, 2004). Among the explanation found more recently to explain this mixed evidence, are those related to the complexity of the application process and the eligibility conditions, which leave many eligible students without access to Pell Grants (Dynarski and Scott–Clayton, 2006; Bettinger et al., 2012). This type of studies has additionally been complemented by the analysis of more complex strategic results like the role of the availability of Pell Grants in the reduction of the provision of financial aid by higher education institutions (Turner, 2017; Goldrick-Rab et al., 2012). Similarly, it is expected that this availability of resources in terms of incidence has an additional effect on tuition costs (Long, 2004).

More recently, however, experimental evidence on the US has illustrated how the interrelation of loans and grants can actually foster student achievement. While grants can have the potential to increase students achievement in a ratio of 2 to 1 (Marx and Turner, 2019a), offering loans increases the availability of resources, while constraining them, by limiting resources, hamper student achievement (Marx and Turner, 2019b). Marx and Turner (2018) find that grants crowd out student loans and Turner (2017) find that out of each dollar of student loans, 12¢ are passed-trough to institutions in the form of higher tuition for grant eligible students. This set of studies are the ones that closer resemble this article, and we provide further evidence on this interrelation between student loans and grants.

Before these set of studies, the causal evidence on student loan impacts is much more limited in spite of the amount of resources that are annually invested in them by different governments (Dynarski, 2015; Page and Scott-Clayton, 2016). Among the few articles we find Solis (2017), who found that for Chile, student loans cause an increase in enrollment rates, and in the persistence in higher education with a differentiated and greater effect for the most vulnerable students. This study uses the assignment rules based on students’ sociodemographic profiles and on the cutoff score obtained in the higher education admission test using a RDD.
Solis (2017) found that loans increase in 20 ppts the enrollment rates at the cutoff point, a 100% increase relative to the level of the non-eligible, as well as increase in 16 ppts persistence rates after two years, a 53% increase over the level of ineligible students. Similarly, another study in the South African context, Gurgand et al. (2011), evaluates the effect of the aids in the form of student loans at the cutoff point of a credit score for assignment, like those used by the traditional banking sector, and it also found that student loans can increase in 20 ppts the enrollment rates at the cutoff point.

Among the evidence for the United States is the study of Dynarski (2002), which found inconclusive evidence on the effect of the expansion of the loan programs at the beginning of the 90s in the increase in higher education tuition. Two other studies have taken advantage of the variation in the availability of federal assistance in Community Colleges and found that those institutions that offer federal assistance exhibit higher enrollment and retention rates (Page and Scott-Clayton, 2016; Dunlop, 2013; Wiederspan, 2016).1

In this article, we analyze the combination of two types of policies that seek to help students, especially low-income students, in access to, retention in, and successful graduation from higher education: Cash subsidies, which reduce the cost of attending higher education, and student loans, which seek to remove the financial barriers students face. From that viewpoint, this evaluation seeks to contribute to the literature with the knowledge of the causal impact of cash subsidies conditional on receiving a student loan.

3.2.1 Evaluations to the ACCES Loan Program

The ACCES Loan Program has had a series of impact evaluations where the most noteworthy was the one conducted by Melguizo et al. (2016). This evaluation found significant effects of loan availability for eligible applicants in an increase in higher education enrollment, in

---

1For more information on these studies, Page and Scott-Clayton (2016) provides detailed information about the experimental designs, programs, and variables analyzed by each one of them.
a reduction in dropout rates, and in the number of approved credit-hours. As reference, the evaluation used the cutoff score of eligibility to access the loan determined by Icetex for each administrative region (departamento) of the country based on the score of the secondary education exit exam, known as ICFES or SABER 11.

Within the framework of other evaluations conducted over the ACCES loan program, Sanchez and Velasco (2014), in a working paper, apply the same methodology to find the casual impacts of the loan in labor variables. The study found that educational loan beneficiaries receive higher wages, and it attributes this result to longer job searches after exiting, which allows loan beneficiaries to find better jobs by waiting for better offers.

Evaluations prior to these studies did not include causal inference mechanisms, they were conducted on a non-experimental basis, thus they did not rigorously correct the self-selection problem, nor were they exposed to peer review publication processes.

With respect to the cash subsidies that accompany these loans, this is the first evaluation for them and the only one with the potential of arriving to causal estimates. In this article we show that a cash subsidy, small in terms of the loan\(^2\), can have a substantial impact on student achievement, and that fine-tuning of the program requires the close assessment of how loans and cash subsidies affect student achievement jointly.

3.3 Policy Background

3.3.1 The ACCES-PACES loan program for higher education

In the year 2002, the loan program Quality Access to Higher Education - ACCES (Acceso Con Calidad a la Educación Superior) was born, at a time when Colombia was exiting a strong economic recession characterized by low growth rates that left deep economic and social consequences. The difficult economic situation increased the opportunity cost of education that

\(^2\)The periodic cash subsidy represents 10\% off the periodic loan disbursement.
led to the marginalization of many youth from their studies due to the pressure of generating income for their households or because they did not have any financing options. These series of circumstances translated into the stagnation of the enrollment rates in higher education (Cárdenas, 2003). The ACCES program was developed as a strategy that allowed stimulating access to higher education, especially in among the disadvantaged population.

With the ACCES project, the country had the first long-term subsidized line of credit with an exclusively social focus. In the ACCES credit line, 124,532 new loans were granted in the period 2003-2007 (Icetex, 2013). More than 80% of the loans were assigned among students from socioeconomic strata 1 and 2.

Once the first stage of the ACCES project ended, a series of reforms guaranteed Icetex’s independence from the central government’s budget in order to offer further growth possibilities. Seeking independent growth, Icetex proceeded with the implementation of another stage of the ACCES loan program, where loan financing would be covered by Icetex resources, and the award of additional subsidies would be covered by the Central Government. During the execution of this second stage of ACCES, from 2008 until 2012, a total of 210,314 undergraduate ACCES loans were awarded. Between 2003 and 2014, within the framework of the ACCES credit line, Icetex gave financial aid to more than 450 thousand students.

As of 2015, Icetex unified all its undergraduate lines within a new program denominated “You Choose” (Tú Eliges), in which students can choose to a certain degree the payment plan that best suits them with different repayment combinations during their study period. Students eligible for the credit lines with interest rates subsidies and with access to cash subsidies are still the differentiated part of a project denominated PACES, which financed more than 23,000 students in 2017 (Icetex, 2018).

---

3Socioeconomic strata 1 and 2 concentrate approximately the first two income quintiles (OECD and BIRF/WB, 2013).
4The law 1002 of 2005 bestowed Icetex with the status of financial institution of a special nature and made it independent from the Government of Colombia’s General Budget.
5Author’s calculations based on Icetex annual reports available in Spanish at www.icetex.gov.co
### 3.3.2 Additional Cash Subsidies

During the execution of the second stage of ACCES, Icetex complemented the loan policy with the implementation of subsidies directed to tuition or as a cash transfer in 2007. From 2009 onwards, the subsidies started to be awarded exclusively as cash subsidies. Initially, the subsidy amount was set at the minimum monthly wage at the time, and its indexation was linked to the country’s general price index, the Consumer Price Index (IPC, for its Spanish acronym) published by DANE (The Colombian Statistical Administrative Direction). Currently, the subsidy is worth COP 798 thousand or around USD 275 at current exchange rates. The subsidy’s beneficiaries were defined according to vulnerability characteristics, such as their condition of violence displaced victims resulting from the country’s internal armed conflict; their presence in other state social programs, like Red Unidos; and also, the allocation of the subsidy was done through SISBEN 6, which is the tool that by law the different Colombian governmental agencies must use to award subsidies.

Before 2011, SISBEN classified survey respondents in one of seven levels. During this first phase, approximately 60% of the beneficiaries of the ACCES loan received subsidies (Icetex, 2013). As of 2011, SISBEN changed its classification methodology 7, assigning a score to each survey respondent, which represents a continuum over the survey respondents’ poverty level according to its geographic residence. The scores are comparable within three separate geographic units, the 14 main cities, other urban areas, and the rural sector. As of 2011, Icetex has published on its website the SISBEN cutoff score in force for cash subsidy assignment. This reduced the number of cash subsidy beneficiaries in the ACCES line below 40%. The score set in 2011 sought to include loan beneficiaries from the first three income quintiles and as of 2015

---

6SISBEN - System for Selecting Beneficiaries of Social Programs. Information system that classifies survey respondents according to their relative socio-demographic level. It is the Colombian state’s legal tool to award subsidies managed by the National Planning Department.

7The classification in levels was known as SISBEN II and the new scoring classification is known as SISBEN III.
over three quarters of the Colombian population had been surveyed and had a SISBEN score\textsuperscript{8}. Figure 3.1 describes the different cutoff scores published by Icetex for different moments in time between 2011 and 2017. This is the policy we evaluate in this article. In this evaluation, we use this potentially exogenous variable generated by this rule to implement the RDD. Our evaluation covers the period from 2011 to 201 considering that for these beneficiaries the SISBEN III data exists, as well as the results variables reported by SPADIES.

The third version of SISBEN was referred as an improvement by both the government and by institutions that have evaluated it (OECD and BIRF/WB, 2013). Still, the quality of the survey depends on the independence of the surveyor with respect to the potential beneficiary and there have been reports of manipulation on how scores are assigned by politicians and in remote areas (Clavijo, 2017; DNP, 2015). While the intention of the current article is not to evaluate SISBEN, the findings suggest that as a discriminatory system, the sociodemographic classification was better in the main cities (as opposed to other urban or rural areas), and in 2012 relative to 2011.

### 3.4 Methodology and Data

#### 3.4.1 Methodology

To determine the causal effect of a policy like these cash subsidies, it is necessary to consider the endogeneity problem created by the characteristics of the beneficiaries of the treatment in relation to non-beneficiaries. If we do a simple comparison between the data in the results of the two groups, beneficiaries and non-beneficiaries, the difference cannot be directly attributed to the policy. The foregoing is because many other student characteristics could be explaining the differences. The most obvious of these is family income, students who do not receive a subsidy have greater resources and this can affect the results. Including controls in regression analysis

\textsuperscript{8}Information as of 2015 referenced at www.dnp.gov.co
Figure 3.1: SISBEN Scores across time for Different Geographic Areas

### 2017

<table>
<thead>
<tr>
<th>No. del área</th>
<th>Área</th>
<th>Puntaje Mínimo</th>
<th>Puntaje máximo</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14 Ciudades, Son las 14 principales ciudades sin sus áreas metropolitanas, Bogotá, Medellín, Cali, Barranquilla, Cartagena, Cúcuta, Bucaramanga, Ibagué, Pereira, Villavicencio, Pasto, Montería, Manizales y Santa Marta.</td>
<td>0</td>
<td>30.39</td>
</tr>
<tr>
<td>2</td>
<td>Resto Urbano: es la zona urbana diferente a las 14 principales ciudades, los centros poblados y la zona rural dispersa de las 14 principales ciudades</td>
<td>0</td>
<td>30.73</td>
</tr>
<tr>
<td>3</td>
<td>Rural</td>
<td>0</td>
<td>22.19</td>
</tr>
</tbody>
</table>

### 2015 ACUERDO 013 (30.04.15)

<table>
<thead>
<tr>
<th>No. del área</th>
<th>Área</th>
<th>Puntaje Mínimo</th>
<th>Puntaje máximo</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>En las 14 principales ciudades sin sus áreas metropolitanas, Bogotá, Medellín, Cali, Barranquilla, Cartagena, Cúcuta, Bucaramanga, Ibagué, Pereira, Villavicencio, Pasto, Montería, Manizales y Santa Marta.</td>
<td>0</td>
<td>54</td>
</tr>
<tr>
<td>2</td>
<td>Resto Urbano: es la zona urbana diferente a las 14 principales ciudades, es decir, los centros poblados y la zona rural dispersa de estas.</td>
<td>0</td>
<td>52.72</td>
</tr>
<tr>
<td>3</td>
<td>Rural</td>
<td>0</td>
<td>34.79</td>
</tr>
</tbody>
</table>

### 2013 ACUERDO 009 (18.04.13)

<table>
<thead>
<tr>
<th>No.</th>
<th>Área</th>
<th>Puntaje Mínimo</th>
<th>Puntaje Máximo</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14 Ciudades, Son las 14 principales ciudades sin sus áreas metropolitanas, Bogotá, Medellín, Cali, Barranquilla, Cartagena, Cúcuta, Bucaramanga, Ibagué, Pereira, Villavicencio, Pasto, Montería, Manizales y Santa Marta.</td>
<td>0</td>
<td>57.21</td>
</tr>
<tr>
<td>2</td>
<td>Resto Urbano: es la zona urbana diferente a las 14 principales ciudades, los centros poblados y la zona rural dispersa de las 14 principales ciudades.</td>
<td>0</td>
<td>56.32</td>
</tr>
<tr>
<td>3</td>
<td>Rural</td>
<td>0</td>
<td>40.75</td>
</tr>
</tbody>
</table>

### 2011: ACUERDO 017 (25.05.2011)

<table>
<thead>
<tr>
<th>No.</th>
<th>Área</th>
<th>Puntaje Mínimo</th>
<th>Puntaje Máximo</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14 ciudades principales sin sus áreas metropolitanas: Bogotá, Medellín, Cali, Barranquilla, Cartagena, Cúcuta, Bucaramanga, Ibagué, Pereira, Villavicencio, Pasto, Montería, Manizales y Santa Marta.</td>
<td>0</td>
<td>52.86</td>
</tr>
<tr>
<td>2</td>
<td>Zona urbana</td>
<td>0</td>
<td>50.45</td>
</tr>
<tr>
<td>3</td>
<td>Rural</td>
<td>0</td>
<td>40.75</td>
</tr>
</tbody>
</table>

The header on top of each block represents the corresponding Icetex ruling (Acuerdo)
is not enough: as we incorporate more controls in a regression we eliminate the variation from the variable that has been included, but also from the variable of interest, plus we don’t know. In addition, the information gathered does not have all the information about the students’ characteristics that can determine the results in their academic performance as we could always imagine ways in which unobserved motivation, ability or knowledge accumulation can drive our results.

This problem can solved by the random assignment to the treatment. Groups randomly separated with control and balancing on the observable characteristics allows us to assume confidently that the non-observable characteristics will also be balanced. This type of ideal experimentation is not always possible in public policy. For example, in Colombia, the central government’s subsidies can only be awarded as an aid for vulnerable populations and that beneficiary selection problem create the problem of endogeneity in the empirical analyses of policy.

In spite of this, a quasi-experimental design exists that allows conducting an investigation with causal interpretation that uses the knowledge in the treatment assignment rules: RDD (Angrist and Pischke, 2008; Calonico et al., 2017; Cattaneo et al., 2019). An RDD takes advantage of the exogenous variation introduced by the assignment rule. The rule sets, in a relatively arbitrary manner, a cutoff score that precisely defines who is and who is not a beneficiary of the treatment. Right on the cutoff point, those individuals that are sufficiently close to the cutoff score are comparable and the assignment to them is nearly as random, provided the assignment variables is not manipulated, and therefore there is no possibility of ineligibles to self-select into the treatment. And also to the extent that there is no other characteristic that jumps up right on the cutoff score and that can explain the results. If the latter should occur, when we compare the outcome data, if any difference exists, we could not necessarily attribute the difference between the outcomes to our variable of interest, as it
could also be attributed to the other changing variable. However, even if these conditions are met, the effects found cannot be generalized to all the population, rather they are exclusive to the locality of the cutoff scores used.

In conclusion, if we consider that the composition of the characteristics that can determine the results varies slightly over the range of the assignment variable, and if any significant difference exists in the results variables at the cutoff score, this difference is caused by the treatment.

We formalize this discussion following Angrist and Pischke (2008) and Cattaneo et al. (2019) with the model below\(^9\). In a first instance, a deterministic and discontinuous assignment mechanism exists, as a function of an independent variable \(x_i\), such that:

\[
D_i = \begin{cases} 
1 & \text{if } x_i \leq x_0 \\
0 & \text{if } x_i > x_0 
\end{cases}
\] (3.1)

The treatment variable, \(D_i\), will take the value of one when the assignment variable is below a certain threshold, and it will take the value of zero when the value is above a certain threshold. If this criterion of treatment access is translated into treatment take-up, we are in the case of a Sharp RDD and equation (3.1) describes the assignment process, where all the eligible individuals receive the treatment and all the non-eligible individuals are excluded from it. When this parity is not fulfilled and some level of non-compliance happens for the eligible or for the non-eligible, we are in the case of a Fuzzy RDD. In a Fuzzy RDD the probability of access to treatment is described by:

\[
P(D_i = 1|x_i) = \begin{cases} 
g_1(x_i) & \text{if } x_i \leq x_0 \\
g_0(x_i) & \text{if } x_i > x_0 \end{cases}, \text{ where } g_1(x_i) \neq g_0(x_i)
\] (3.2)

In that case, treatment access is still determined by the exogenous variation introduced

---

\(^9\)I compute the main results in Stata 16 using \texttt{rdrobust} and \texttt{rdplot} as explained in Calonico et al. (2017) and Cattaneo et al. (2019)
by the assignment variable; however, rather than going from total absence of treatment to its total presence, there is a discontinuity in the probability of treatment. To the extent that that discontinuity exists and is strong enough, \((g_1(x_i) \neq g_0(x_i))\), the inference in the comparison of results of those eligible and those ineligible still has causal implications, but local to the beneficiaries induced to take the subsidy because of the exogenous variation in the assignment variable. Figure 3.2 presents the difference between the two types of RDD, *Sharp* and *Fuzzy*, with respect to treatment access. In the case of the cash subsidies as we previously described, the beneficiaries that satisfy the SISBEN’s cutoff scores are entitled to the subsidy, but they do not all take it. The degree of parity between eligibility and ineligibility is over 80%. However, at the cutoff score, a disparity also exists between ineligibility and taking the subsidy since 55% of ineligible individuals at the cutoff score access the subsidy. Due to this degree of non-compliance at both sides of the cutoff, the analysis corresponds to a *Fuzzy* RDD. As we will see below, it is relevant in the final take up of the treatment which allows us to conduct our statistical analysis. The assignment rule does not prevent every one from accessing the subsidy, however, it increases the transaction costs for those who require to alter their SISBEN score before being able to claim it, this is the source of our exogenous variation.

The estimation of the effects of a *Fuzzy* RDD is done in the same context of instrumental variables, with and Intent to Treat (ITT) and a Local Average Treatment Effect (LATE). ITT reflects the difference in the result variable at the cutoff score and LATE reflects the
effect caused on the beneficiaries as a result of the assignment variable. Formally, ITT is the difference in the observed outcome \((Y_i)\) of the eligible individuals close to the cutoff point \((x_0 + \Delta < x_i < x_0)\), and the ineligible individuals close to the cutoff point \((x_0 < x_i < x_0 + \Delta)\):

\[
\text{ITT} = \lim_{\Delta \to 0} \left\{ E[Y_i|x_0 - \Delta < x_i < x_0] - E[Y_i|x_0 < x_i < x_0 + \Delta] \right\} 
\] (3.3)

Analogously, LATE is the difference adjusted by the probability of treatment at both sides of the cutoff point. Formally,

\[
\text{LATE} = \lim_{\Delta \to 0} \left\{ \frac{E[Y_i|x_0 - \Delta < x_i < x_0] - E[Y_i|x_0 < x_i < x_0 + \Delta]}{P[D_i|x_0 - \Delta < x_i < x_0] - P[D_i|x_0 < x_i < x_0 + \Delta]} \right\} 
\] (3.4)

To estimate the results of equations 3.3 and 3.4 it is necessary to first determine \(\Delta\), the bandwidth on which the comparisons of the results will be made. In this article, we follow the methodology developed by Calonico et al. (2017) as described in Cattaneo et al. (2019) that uses a local linear regression to calculate the optimal bandwidth without having to make assumptions on the functional forms of the relationship between the assignment variable and the results. Once we obtain the bandwidth we obtain the coefficient of equation 3.3 through comparing the intercept coefficients of the local polynomial regressions above and below the cutoff, weighted by a triangular kernel \(^{10}\). I follow a local polynomial of order two, following the recommendation by Gelman and Imbens (2019), however the choice of polynomial does not alter the main conclusions\(^{11}\). As in any estimation of instrumental variables, in the first stage, the probabilities of treatment, access to the subsidy, are estimated and in the second stage, LATE is estimated using the initial estimation of the probability of treatment (Angrist and Pischke, 2008; Calonico et al., 2017; Cattaneo et al., 2019).

\(^{10}\)This is the \textit{robust} estimator. Cattaneo et al. (2019) show how “the local polynomial regression is simply a weighted least-squares fit”

\(^{11}\)I also considered polynomials of order 4 to allow for additional flexibility of the underlying function and no significant changes appear. Results available upon request.
3.4.2 Data

Our initial dataset counts 176,210 observations of the cohorts of Icetex loan beneficiaries for the period from 2011 to 2015. These records were merged with information from SPADIES\textsuperscript{12}, from where their state of retention and graduation was obtained.

Similarly, the information provided by Icetex included the information, at the time of applying for the loan, of SISBEN, which is used as one of the criterion to access cash subsidies. Not all loan beneficiaries have records in SISBEN, so when we only considered those beneficiaries with SISBEN records, the number of records falls to 85,191.

Additionally, for most part of the analysis presented below, we dismissed the beneficiaries of subsidies for reasons different to the SISBEN classification\textsuperscript{13}, which eliminated 4,812 records, and the beneficiaries of the Alianzas and CERES loan programs which are traditionally allocated under different conditions relative to the remaining mainstream student loans. Our final database consisted of 44,023 records for the cohorts 2011-2015 with complete information for all variables considered, and we especially highlight the 2011 and 2012 cohorts that have had enough time to graduate and dropout.

3.5 Main Results

In this section we present the numerical results of the implementation of the RDD on the cash subsidy assignment for beneficiaries of the Icetex loan using the SISBEN cutoff score as a reference, which delimits the access to these subsidies.

First, I present the analysis on the assignment variable and document the existence of the discontinuity in the probability of subsidy take-up, with non-compliance between assignment and take-up. The discontinuity created by the allocation rule is enough to estimate the

\textsuperscript{12}SPADIES – Dropout Analysis and Prevention System in Higher Education Institutions. Information and monitoring system of the Ministry of Education of the Republic of Colombia.

\textsuperscript{13}With a certain degree of variation in the different periods of our analysis, the access to the cash subsidy has been awarded to the beneficiaries dully registered as indigenous, displaced persons, demobilized, victims of violence or beneficiaries of the Red Unidos program.
equations 3.3 (ITT) and 3.4 (LATE) for the different outcome variables. The foregoing determines that the estimation procedure follows a Fuzzy RDD as described in Section 3.4.

The impact on the Dropout Rate, Dropout Rate in year one, and Graduation Rate are presented below for the cohorts 2011 and 2012 and for the general sample. Finally, the effect for different indicators of the default rate conditional on the analyzed beneficiaries’ graduation is presented.

Figure 3.3: Probability of Receiving a Subsidy \((P(D_i))\) and SISBEN Score

3.5.1 Assessment of the RDD Assumptions

Discontinuity in the Treatment Variable

The first result is the continuity test in the assignment variable. As mentioned in section 3.4, the estimation of ITT (Equation 3.3) and LATE (Equation 3.4), depends on the existence of the discontinuity and that it is significant. Figure 3.3 presents the discontinuity in the probability of receiving the cash subsidy among the loan beneficiaries, conditional on the SISBEN score for the cohorts 2011 and 2012. The left panel presents the raw data accompanied by the estimation of an unrestricted local polynomial regression, in order to allow for the flexibility in the way that I portrait the relationship among the two variables. On the right, the graph presents the local polynomial estimates within the optimal bandwidth. The degree of compliance for

\[14\] In the text, the word \textit{treatment} is constantly interchanged with cash subsidy. Treatment that we are evaluating in this case.
the eligible below the cutoff points is, on average, greater than 92.1%, and 43.8% for the ineligible above the cutoffs, and within the optimal bandwidth the compliance is 91.7% and 52.0% for eligible and ineligible accordingly. Even closer, in the limit, in the last bin of the local polynomial regression of figure two, the baseline probability of receiving the subsidy for the ineligible right at the cutoff is 62.7%, and the discontinuity in treatment with a robust coefficient is -28.5% implies that exactly at the cutoff 90.2% of the eligible receive the subsidy. Coefficient estimates are presented in Table B.1 and the description corresponds to the model without covariates (Model 1). Additional columns include additional covariates in order to sharpen the estimates and to address any additional source of bias. The bottom subsection of B.1 specifies what additional variables were considered, for the Individual characteristics, the age (linear and in its quadratic form), the gender, the cohort and weather the beneficiary was an Afro-Colombian and weather the individual was a member of the Colombian Indigenous communities. For the Academic characteristics, the variables considered included weather the diploma pursued was an AD degree or not, the SABER 11 scores (ACT/AST equivalent). For the household characteristics, the regressions include regional fixed effects and the socio-economic strata of the household.

Figure 3.4: Subsidy Disbursed ($E(\text{Subsidy})$) and SISBEN Score

Several things stand out about the graphs: the first, the jump in the probability of receiving the subsidy between ineligibles and ineligibles according to the SISBEN score, is 40 units at
the cutoff points. This a substantial discontinuity, however, the data among the ineligible exhibits a pattern of increasing likelihood of accessing the treatment as the individual score is closer to the cutoff. This result is a potential indication of the manipulation on the assignment variable since in appearance, after the process of allocation and legalization of the loans, many students seem to ask to be re-interviewed to get their SISBEN scores reviewed. We use the SISBEN score that the students have at the first moment of treatment. When the current score is used, it shows that several students had their scores revisited and became eligible, it could indicate that the review was made in order to become eligible to receive a subsidy. According to the information provided by Icetex, all cash subsidy beneficiaries to the right of the cutoff score should not receive the subsidy unless they subsequently revised their score or other conditions to access subsidies. If this is true, those students closer to the cutoff in the ineligible side, are more likely to get their scores reviewed as apparently, the cost for them to review their score is lower, and doing so grants them eligibility to the cash subsidy.

Table 3.1: First Stage: Discontinuity in the Probability of Receiving a Subsidy and on the Amount of Subsidy Received

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Robust Coefficient</td>
<td>-0.285***</td>
<td>-0.295***</td>
<td>-0.293***</td>
<td>-0.291***</td>
<td>-651.6***</td>
<td>-734.5***</td>
<td>-734.0***</td>
<td>-723.0***</td>
</tr>
<tr>
<td></td>
<td>(0.0376)</td>
<td>(0.0381)</td>
<td>(0.0371)</td>
<td>(0.0361)</td>
<td>(98.05)</td>
<td>(88.82)</td>
<td>(83.59)</td>
<td>(81.35)</td>
</tr>
<tr>
<td>N</td>
<td>19,279</td>
<td>19,192</td>
<td>19,192</td>
<td>19,192</td>
<td>19,279</td>
<td>19,192</td>
<td>19,192</td>
<td>19,192</td>
</tr>
<tr>
<td>Ineligible Baseline</td>
<td>0.627</td>
<td>0.626</td>
<td>0.626</td>
<td>0.626</td>
<td>1180.4</td>
<td>1151.3</td>
<td>1132.8</td>
<td>1129.5</td>
</tr>
<tr>
<td>Individual</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Academic</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Household</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

Coefficients from difference across local polynomial regression of order 2. Bandwidth selection method: common MSE-optimal bandwidth selector (MSERD). All coefficients are robust bias-corrected (Calonico et al., 2017; Cattaneo et al., 2019). Statistical significance reported at the 1 percent (***) , 5 percent (**), and 10 percent (*) level. Heteroskedastic standard errors in parentheses clustered at the departmental (Colombian political division) level.

Manipulating the assignment variable decreases the gap in the discontinuity, but it does not eliminate it. Therefore, the causal implications of our study are still identified, but the experiment’s strength decreases, generating more noise and higher standard errors for the estimators of the instrumental variables in LATE (Equation 3.4), which we will discuss further in the results.
Additionally, Figure 3.4 present the intensity of the treatment with the average of subsidies wired to students according to the distance to the SISBEN cutoff score. At the cutoff up to 2017, the students ineligible to receive a cash subsidy belonging to the 2011 and 2012 cohorts received USD1,180 (COP2.34 million) and the eligible students received USD1,832 (COP3.47 million). This means that at the cutoff, the eligible students received USD652 (COP1.29 million) more than the ineligible counterpart.  

McCrary Test

The assignment variable is also evaluated through the McCrary test (McCrary, 2008). This test calculates a t-statistic comparing the probability density estimation in logarithm at both sides of the cutoff point. If the distribution exhibits a jump and the difference is statistically significant, evidence exists of manipulation in the assignment variable. On the other hand, if the difference is not statistically significant, we can conclude that the distribution of the assignment variable is smooth and therefore there is no evidence of manipulation.

In this case, if the students were interested in accessing the cash subsidies and they could manipulate the assignment variable, they would revise their score to obtain one below the cutoff point. Figure 3.5 presents the estimation of the empirical distribution of the assignment variable in our data. At the cutoff point in the sample that pools all the geographical areas together, I found a difference in the logarithms of $-0.027$, with a standard error of $0.038$, reason why the distribution is slightly taller below the cutoff point. The p-value of the t-test can be rejected at a confidence level of $0.1\%$. With this difference, the hypothesis of equality in the estimation of both sides of the distribution cannot be rejected and our conclusion is that there is no evidence of manipulation in the assignment variable at this stage of the process.

This result is nuanced and Figure 3.5 shows this. For Urban Areas there is no discontinuity.

\footnote{The exchange rate used corresponds to COP 1,985/USD corresponding to the average exchange rate of the period 2011-2015.}
of any kind, which means that the identification through the assignment to a subsidy is clean. On the other hand, for the Other Urban Areas and for Rural Areas the discontinuity seems sharper which could be a degree of manipulation on the assignment variable. This is consistent with two hypotheses that refer to SISBEN as a classification system. First, the idea that far from the main cities, in which institutions are stronger, the manipulation of the assignment variable is harder and imposes a higher cost. Second, the possibility that in 2011, being one of the first years of implementing the new methodology, some student loan beneficiaries managed to get their scores cleared before. The results tend to favor the latter. In 2012, there are no discontinuities in any of the sub samples by geographic area (Figure B.4), while most of the discontinuities come from the samples in 2011 (Figure B.3).

Additionally, in Appendix B.2 Figure B.3 and Figure B.4 present the estimation of the
McCrary test for the two cohorts included in our study. For 2011, there is a discontinuity overall, driven by the student loan beneficiaries in Other Urban Areas and in Rural Areas (Bottom panel). Although there is not a significant discontinuity on the Main Urban Areas, countering the intuition, there is a small discontinuity in and individuals bunch on the non-eligible side of the cutoff (Figure B.3, Upper Right panel). I attribute this to the proximity to the highest point in the density distribution, 10 points above the eligibility cutoff points.

In conclusion, considering these results, our information is consistent with the experimental design assumptions overall, and among individual subgroups, with the exception of the Other Urban or Rural Areas during 2012. This allows us conducting an impact evaluation with causal implications on the outcome variables, with those considerations in mind.

**Continuity over Additional Variables**

In addition to the McCrary test, I measured the discontinuity at the assignment to treatment cutoff on variables that we would expect to find no effect whatsoever. If there is a discontinuity in a variable such as age, that would compromise the interpretation of the coefficients as causal. In fact, if in the analysis finds a discontinuity on age, I could not attribute the discontinuity over outcomes to the treatment only, in this case the access to both the student loan and to a cash subsidy. If the discontinuity is only in age, then the discontinuity in outcomes would be attributed to the joint effect of age and treatment.

If the analysis finds more discontinuities in variables that we would not expect to find such an effect, our estimates would be confounded and thus, not causal. However, if the analysis finds that the variables that should not have a discontinuity are actually continuous at the cutoff point, and it also finds this consistent result across several of these observables, I can safely conclude that, local to the cutoff, individuals at both sides are comparable. This result would allow to further extrapolate and assume that the assignment to treatment is exogenous.
as it balances individuals at both sides of the cutoff.

Figure 3.6: Discontinuity Test on Additional Covariates

Note: Estimates of within optimal bandwidth local polynomial regression and bin averages.

Figure 3.6 presents the discontinuity test for four different variables: Age, Saber 11 Score (SAT analogous test), the probability of come from the Capital ($P(Bogota)$), and the probability of being female. We follow the same estimation procedure as in the previous subsection. None of them exhibit a statistically significant difference on these variables at the cutoff point. Appendix B.3, presents these figures for the different geographic areas. I don’t find a particular pattern on these variables either, when we split the samples on the basis of SISBEN scores, by geographical location of the household.
3.6 Results on Main Academic Outcomes

Considering the results on the previous section, in this section, I separately present the combined results for the main academic outcomes considered: the probabilities of dropping out, dropping out during the first year, and graduating, and finally, the estimated heterogeneous coefficients by geographical area. Appendix B.4 presents the detailed graphical estimation by geographic area and Appendix B.4, presents them by cohort.

3.6.1 Impact on Dropout Rates

The first outcome that we present is the effect on the global dropout rate. Table 3.2 presents the estimations of equations 3.3, Intend-to-Treat, and 3.4, LATE, Local Average Treatment Effect. I estimate LATE following Equation 3.4, and weighting, first with respect to the probability of getting a subsidy \( P(\text{Subsidy}) \), and also with respect to the amount that the beneficiaries receive subsidies \( E(\text{Subsidy}) \). Models (1) to (4) followed the previous specifications according to the inclusion of covariates, and used the default bandwidth estimation procedure with one common mean squared error objective function (MSERD) (Calonico et al., 2017).

Overall, the probability of dropping out decreases by 11.7 percentage points at the left side of the cutoff, for the eligible students\(^\text{16}\). The LATE estimators are presented for reference, although they are imprecise, considering the probability of accessing the subsidy, the local average treatment effect is a reduction of 44.2% in the dropout rate. Considering that the baseline among the ineligible is lower than that number, at the cutoff the probability of dropping out disappears for the eligible students. Following from the interpretation of the LATE coefficient on the subsidy amount, each USD100 in subsidies for eligible students decrease the dropout rate between 1.1 and 2.1 percentage points.

Figure 3.7 presents the results graphically. The left panel presents the unrestricted data.

\(^{16}\text{The negative sign of the coefficient is related to the fact that the probability of treatment increases with negative numbers of the running variable}\)
Table 3.2: Effect over the Probability of Dropping Out

<table>
<thead>
<tr>
<th></th>
<th>P(Dropout</th>
<th>Sisben III)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>ITT</td>
<td>0.117***</td>
<td>0.124***</td>
</tr>
<tr>
<td></td>
<td>(0.0405)</td>
<td>(0.0395)</td>
</tr>
<tr>
<td>LATE - P(Subsidy)</td>
<td>-0.442***</td>
<td>-0.478***</td>
</tr>
<tr>
<td></td>
<td>(0.152)</td>
<td>(0.150)</td>
</tr>
<tr>
<td>LATE - E(Subsidy)</td>
<td>-0.0164***</td>
<td>-0.0159***</td>
</tr>
<tr>
<td></td>
<td>(x USD100)</td>
<td>(0.00582)</td>
</tr>
<tr>
<td></td>
<td>(0.00497)</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>19,279</td>
<td>19,192</td>
</tr>
<tr>
<td>Bandwidth - MSERD</td>
<td>11.27</td>
<td>11.00</td>
</tr>
<tr>
<td>Ineligible Baseline</td>
<td>0.408</td>
<td>0.409</td>
</tr>
<tr>
<td>Individual</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Academic</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Household</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

Coefficients from difference across local polynomial regression of order 2. Bandwidth selection method: common MSE-optimal bandwidth selector (MSERD). All coefficients are robust bias-corrected (Calonico et al., 2017; Cattaneo et al., 2019). Statistical significance reported at the 1 percent (***) , 5 percent (**), and 10 percent (*) level. Heteroskedastic standard errors in parentheses clustered at the departmental (Colombian political division) level.

with a local polynomial of order four. One striking results is the lack of a gradient on SISBEN score on the running variable. Considering the well documented relationship between poverty and educational attainment, the fact that there is almost no gradient on dropout rates in the overall sample calls the attention. With respect to the effect of the policy, in appearance the shape of the polynomial could be referred to as determinant of the effects, however, the results I found in \textit{rdrobust} are consistent across different orders of the polynomial and across different methodology estimations\footnote{A previous version of this analysis was conducted using the methodology described by Imbens and Kalyanaraman (2012) and implemented to the stata command \textit{rd} results are very similar in that procedure}.

### 3.6.2 Impact on First Year Dropout Rates

In Colombia, at least half of the dropouts happen during the first year of attending classes (OECD and BIRF/WB, 2013). Although, some of those dropouts might not be related to financial obstacles that students face, as expectation mismatch and adaptation play an important role in the beginning of the higher education experience, it is still a subject of interest to
understand if the loans and subsidies provided by the Colombian Government can affect the decision of students to dropout at the beginning of their programs.

Table 3.3 provides the estimations of the different coefficients. The ITT coefficient oscillates between 7.5% and 8% across the different specifications. Again, I provide estimates of the LATE, though the results seem imprecise. The LATE of the subsidy indicates that being eligible to the loan decreases the the first year dropout rate in up to 30 percentage points. This is way in excess of the measured ineligible baseline, in the order of 18%. Finally, the
LATE coefficient weighted on the differential subsidy received by the eligible, indicates that for each USD100 of subsidies given to the eligible students, the dropout rates decreased in about 1.15%. Again this estimates of the LATEs would imply that the dropout probability is brought down to zero local for the eligible students local to the access to the subsidy and to the cutoff.

Graphically, in Figure 3.8 the discontinuity in the probability of dropping out during the first year can be appreciated from the full range unrestricted local polynomial. Again, it is surprising that the first year dropout rates do not exhibit a gradient on the SISBEN variable, and the results are robust to the choice of polynomial and method of estimation\textsuperscript{18}.

3.6.3 Impact on Graduation Rates

The final outcome I analyze here is the probability of graduation. The final academic objective is to help low income students to cross the finish line and set them for a improved career prospect. Graduation rates from higher education in Colombia are low and the effort that the government makes is considerable (OECD, 2016).

Table 3.4 present the robust estimation for the effect over graduations rates for the 2011 and 2012 cohorts. On graduation rates, the effect is not statistically significant, however it

\textsuperscript{18}See previous footnote
Table 3.4: Effect over the Probability of Graduation

<table>
<thead>
<tr>
<th></th>
<th>P(Graduation</th>
<th>Sisben III)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>ITT</td>
<td>-0.0526</td>
<td>-0.0426</td>
</tr>
<tr>
<td></td>
<td>(0.0354)</td>
<td>(0.0301)</td>
</tr>
<tr>
<td>LATE - P(Subsidy)</td>
<td>0.127</td>
<td>0.0627</td>
</tr>
<tr>
<td></td>
<td>(0.179)</td>
<td>(0.168)</td>
</tr>
<tr>
<td>LATE - E(Subsidy USD) (x USD100)</td>
<td>0.00827</td>
<td>0.00412</td>
</tr>
<tr>
<td></td>
<td>(0.00601)</td>
<td>(0.00531)</td>
</tr>
<tr>
<td>N</td>
<td>19,279</td>
<td>19,192</td>
</tr>
<tr>
<td>Bandwidth - MSERD</td>
<td>15.54</td>
<td>17.61</td>
</tr>
<tr>
<td>Ineligible Baseline</td>
<td>0.204</td>
<td>0.203</td>
</tr>
<tr>
<td>Individual</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Academic</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Household</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Coefficients from difference across local polynomial regression of order 2. Bandwidth selection method: common MSE-optimal bandwidth selector (MSERD). All coefficients are robust bias-corrected (Calonico et al., 2017; Cattaneo et al., 2019). Statistical significance reported at the 1 percent (***) and 5 percent (**), and 10 percent (*) level. Heteroskedastic standard errors in parentheses clustered at the departmental (Colombian political division) level.

is consistently positive. The ITT is ranges from 4.3 and up to 5.3 percentage points and the late oscillates between 6.3% to 12.7%. Considering that the ineligible baseline at the cutoff is a graduation rate of 20.4%, the effect of the subsidy represents an increase on graduation rates of up to 25% if the ITT is considered and of up to 60% when considering the LATE. The investment of USD100 on subsidies increases graduation rates among the eligible beneficiaries between 0.3 and 0.8 percentage points.

In spite of the fact that the coefficients are not statistically significant, the size of the effects is considerable. However, estimating the effect on this outcome is affected by the loss of statistical power in the proximity to the cutoff point, a limitation of RDD (Page et al., 2017).

3.6.4 Heterogeneous Effects Across Geographical Areas

In Subsection 3.5.1 and in Appendix B.2 I presented the conclusion that not all the RDD assumptions hold across all the units in which the assignment variable was split as discontinuities on the running variable for Other Urban Areas and for Rural Areas in 2011 seemed
to present some degree of sorting. In addition to this, the differences across urban and rural settings in the access to education that the Colombian society provides, also makes the case to estimate the effects of the student loans separately.

Appendix B.4 presents the graphical representation of the evaluation of discontinuities on the academic variables that so far have been discussed in this, Section 3.6, separating across the different geographic areas. Figure 3.10 presents the coefficients of the specification in Model 4 for the different outcomes and compares it across the three geographical areas. Model 4 is the most precise specification as it includes all the covariates. As mentioned before, the positive sign on the ITT reflects that the probability of dropping out increases with the SISBEN score, however eligibility has an inverse relationship as it is determined by having a low score.

The effects for the Main Cities over the probability of dropout is significant 9.9 percentage points, as well as the coefficient for Other Urban Areas. For the Rural Areas, the coefficient is larger, but the estimates are less precise. Splitting the sample in this fashion decreases the precision with which the coefficients are estimated, as all of the coefficients for the different areas exhibit wider confidence intervals relative to the exercise that includes all of them together.

In contrast, the highest decrease in the probability of dropping out during the first year is exhibited in the Main Urban Areas, with 8.9 percentage points. The decrease in the probability in the Rural Areas is also, high but the coefficients are not precisely estimated.
Finally, graduation rates which decrease with SISBEN score, thus its negative sign for the ITT, are lowest among Other Urban Areas with 4.5 percentage points and followed by Main Urban Areas. None of the geographical areas have a significant coefficient on graduation rates.

### 3.7 Discussion and Conclusions

In this study, we analyzed the effect of cash subsidies given to the most vulnerable beneficiaries of student loans. This evaluation finds the causal effects of cash subsides conditional on being a beneficiary of a student loan. From that viewpoint, it combines two of the policy tools most commonly used by governments to promote access and persistence in higher education: student loans, that partially remove financial barriers to entry, accompanied by additional financial aid in the form of subsidies, which decrease the costs of attending higher education.

Findings indicate that the cash subsidies for the student loan beneficiaries have a substan-
tial and significant effect on dropout rates, local to the cutoff points of subsidy allocation. For the 2011 and 2012 cohorts, I found that at the cutoff point, those students eligible to obtain a subsidy, on average, have dropout rates between 11.5 and 12.4 percentage points (ITT) lower than those ineligible students. When estimating the effect on beneficiaries, it is a reduction of 44 percentage points (LATE). If we consider that the dropout rate for those that are ineligible to get the subsidy is 41%, the cash subsidy reduced this rate at the cutoff point for the eligible beneficiaries, in more than 25% considering the ITT and entirely if we consider the LATE. This reduction surpasses 70% for first-year dropout rates for the ITT and again, is over 100% when considering the LATE.

On graduation rates, the effect is not statistically significant, however it is consistently positive. The ITT is between 4.3 and up to 5.3 percentage points and the late is oscillates between 6.3% to 12.7%. Considering that the ineligible baseline at the cutoff is a graduation rate of 20.4%, the effect of the subsidy represents an increase on graduation rates of up to 25% if the ITT is considered and of up to 60% when considering the LATE. Estimating the effect on this outcome is affected by the loss of statistical power in the proximity to the cutoff point, a limitation of RDD (Page et al., 2017).

Low graduation rates are noteworthy: the 2012 cohort has been in the system at least 5 years and the 2011 cohort at least 6 years, and the graduation rate is not 30% after such a period of time. Another type of analysis in SPADIES shows that the graduation rates in Colombia become stable after 7 or 8 years after enrolling in higher education. These results place a spotlight on potential measures that motivate institutions to promote graduating among its students in the initially expected time of around 4 to 5 years. The restricted access to governmental financial aid mechanisms can be used to that end, as it has been done since the introduction of the quality accreditation system, and more since 2015, when credit was restricted to high-quality accredited institutions. A greater emphasis on quality variables
that are directly related to students could be considered for those institutions that seek to grant massively access undergraduate studies.

Finally, the running variable exhibit sorting specially for 2011 outside of the Main Urban Areas. While the situation was substantially improved for 2012, it is still striking the amount of students who get their scores reviewed, apparently, in order to access the subsidies. While there is still a discontinuity in treatment, if this is a generalized case, it could be the case that the Colombian government is allocating its scarce resources among those who do not need them the most.
Bibliography


83


Icetex (2013). Informe de gestión proyecto ACCES II.


Chapter 4

Consumer Incidence in Sales Tax Holidays: Evidence from Tennessee

Justin Ross & Felipe Lozano-Rojas

Abstract *

Policymakers frequently employ tax holidays to stimulate spending and transfer state revenues to favored groups. Perhaps the most common form of this policy in the United States are “Back to School” tax holidays where states waive the retail sales tax on selected school supplies in the fall. Surprisingly, the empirical research of this policy has found that tax savings are overshifted to households. This paper argues that tax holiday dates should be assumed to be selected endogenously so as to minimize forgone revenue and maximize consumer incidence by targeting periods where prices are lowest. Causal evidence is considered in a natural experiment out of Tennessee, where legislation for a spring sales tax holiday in 2006 and 2007 for school supplies was reauthorized as part of a broader series of tax cuts. We conclude from the evidence of these Tennessee events suggest that households, on average, receive the full tax savings during these programs. However, consumer incidence is heterogeneous with some retailers recapturing the tax savings with higher pre-tax prices.

*The authors appreciate helpful comments and suggestions from Bradley Heim, Kate Yang, Thomas Spreen, Sian Mughan, JM Carroll, Todd Ely and participants of the Applied Research in Public Finance conference at Indiana University. Results in this article are calculated (or derived) based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

**Nielsen’s copyright of data: Copyright © 2020 The Nielsen Company (US), LLC. All Rights Reserved. The conclusions drawn from the Nielsen data are those of the researchers and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.
4.1 Introduction

This paper seeks to provide evidence on the gains to consumers from state sales tax holidays, which are temporary tax cuts for selected classes of goods. Temporary tax cuts and subsidies are a common form of public policy at state and federal levels intended to stimulate certain forms of economic activity and to provide tax relief to selected consumers (Phillips, 2016). Sales tax holidays applicable to certain types of goods and services (clothing, cars, guns, emergency supplies, etc.) have become annual events in several U.S. states since New York first promoted a sales tax holiday in 1997. According to the Federation of Tax Administrators, 19 states hosted sales tax holiday programs in 2018. Six of these states held multiple tax holidays during the year that followed a seasonal pattern, applying to clothing and school supplies in the fall and storm preparedness or ENERGY STAR products in the late spring before the summer heat. Public commentary on the objectives of these policies includes a variety of rationales such as inducing economic stimulus, marketing the state’s commerce, and providing a subsidy to certain types of consumers. However, the popularity of tax holidays among policymakers stands in contrast to their bipartisan condemnation by public finance experts in academia and public think tanks. Both the left-leaning Institute on Taxation and Economic Policy and right-leaning Tax Foundation regularly issue special reports criticizing these policies for violating their respective principles for taxation. Writing in State Tax Notes, professors (Hawkins and Mikesell, 2001, p.802) liken tax holidays to “a Soviet-style state-directed price reduction on items selected by the state.”

Perhaps the most consistent concern of tax holiday critics is that despite the intent of these policies to serve as a kind of welfare program, the generated resources may instead be captured by retailers with higher prices. The reasoning behind this concern is apparent from

---

1See Cole (2009) for a comprehensive history of sales tax holidays. The Federation of Tax Administrators maintains an annual database of state sales tax holiday programs and their coverage.
2Ross (2018) reviews public finance think tanks across different ideologies and documents uniform opposition to sales tax holidays, albeit with differing rationales.
the classic competitive market model, whereby the abolition of a tax eliminates the wedge between consumer and producer prices with the consequence that the producer price rises to match the consumer price. (Henchman and Malm, 2014, p.10) articulate this widely held belief:

When lawmakers create sales tax holidays, the assumption is that the benefit will be passed on to consumers in the form of lower prices. In reality, retailers often absorb those benefits for themselves.

The critique is particularly sensitive to the intent of fall back-to-school sales tax holidays, where an interest in arranging transfers to families with back-to-school expenses may be prominent.

It is far from clear, however, that this relatively straightforward theoretical result actually bears out in the data. There are two studies on sales tax holidays that address the question of incidence in state tax holidays, and interestingly both suggest that tax holidays actually result in lower rather than higher pretax prices\(^3\). The earliest evidence is reported in Harper et al. (2003), the data for which corresponds to the start of the state sales tax holiday phenomenon. Their paper provides descriptive evidence by following a basket of ten goods over a three-week period surrounding the Florida sales tax holiday in 2001 at five large department stores in the Pensacola (FL) and Mobile (AL) metropolitan statistical areas. These goods primarily consisted of clothing and apparel. If the taxes had been waived prior to the sales tax holiday, the exemption would have resulted in $125 in savings for the consumer. Based on posted pretax prices during the sales tax holiday, however, Harper et al. (2003) found only $100.06 in consumers’ savings on the same basket of goods from waived taxes during the holiday. This level was the result of the total pretax price of the basket rising by slightly more than 1 percent during the tax holiday in Florida. However, the same basket of goods in Mobile, where no tax holiday occurred, also increased in pretax price by about three percent,

---

\(^3\)There are many additional studies on other effects of tax holidays. These topics include welfare effects (Phillips, 2016), household expenditures (Agarwal et al., 2017; Mogab and Pisani, 2007), and state revenue loss (Cole, 2009)
which was a larger increase than seen in Pensacola under the holiday. While this study has the advantage of following the same products at the same retailers, it suffers from limitations in sample size, geographic coverage, and duration that prevent the use of causal investigative techniques. Furthermore, following posted prices overlooks the myriad of ways in which retailers are capable of altering their prices, including membership discounts, coupons, in-store special sales, bundling, among others. Just as retailers may decrease prices by offering these promotions, retailers can also withhold these price discrimination strategies and raise effective prices without changing posted prices.

Arguably, the best evidence on the incidence of tax holidays is Cole (2009). His study employed weekly data on the number of retail sales of desktop and laptop computers in the 48 contiguous U.S. states over 30 weeks (from May to December of 2007), during which nine sales tax holidays were implemented across the sample states. Cole’s empirical strategy employs computer model, state, and week fixed effects in estimating the effect of the sales tax holiday on the average pretax price. The estimation results indicated that the average pretax price of a computer model decreased by 0.27 percent in response to an average tax rate cut of 4.76 percent during the holiday. While the effect was not statistically significant, Cole interprets the pretax price decrease to provide evidence that the sales tax break is overshifted onto consumers.

In this paper, we argue that the case of sales tax holidays in Tennessee during the springs of 2007 and 2008 represents an unusually good quasi-experimental investigation for identifying causal effects of a sales tax holiday. Generally, the timing of most sales tax holidays is plausibly endogenous to seasonal supply and demand fluctuations. For example, “Back-to-School” tax holidays occur in late summer and early fall when retailers increase inventories and hold sales on school supply products that consumers begin demanding. This seasonality is likely the reason both tax holiday and non-tax holiday states have similarly timed peaks in the prices
and sales volumes of school supply products (see Figure 4.1). Likewise, spring tax holidays in the southern states tend to target products promoting energy savings on air conditioning products just before the arrival of summer heat. For policy makers, this seasonal timing is potentially advantageous for a number of reasons, such as increasing the number of affected households or raising the visibility of their tax cut to consumer for political gain. However, this presents a challenge for investigation, as it requires disentangling the seasonality effect from the tax holiday induced effect. In the aforementioned Tennessee cases, the state ran its first back-to-school tax holiday in August 2006. It then followed in the spring of 2007 with a series of tax amendments that provided various tax policy changes to exemptions, particularly with exemptions and credits for business purchases. These acts also briefly altered the section of the tax code that had authorized the previous fall tax holiday for school supplies, computers, and clothing apparel sold by retailers by changing the tax code section language from “first Friday of August and 11:59 p.m. the following Sunday” to instead reference April 27, 2007 through April 29, 2007. This language was duplicated again in the spring 2008 tax holiday, providing a set of off-season school supply sales tax holidays where we are plausibly isolating the effects from the sales tax holiday on prices from the confounding effects of the seasonal patterns. In this setting we find overshifting is qualitatively eliminated with minor amounts of retailer recapture; consumers save only the dollar amount of the tax waived by the holiday.

This paper employs retail scanner data to investigate the aforementioned claims on seasonality and consumer incidence for both the Tennessee cases of interest and the more routine state tax holidays\(^4\). This scanner data is a long panel of weekly product data from approximately 35,000 yearly participating retailers across the entire U.S. on 2.6 million Universal Product Codes (UPC) for a wide range of product categories. The richness of the data allows us to select specific products and follow the retailers’ tax-exclusive prices throughout each year.

\(^4\)Results in this article are calculated (or derived) based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.
Figure 4.1: Mean Per Unit Price and Total Quantity Sold for School Supplies, 2006 to 2014

Note: Non-Sales Tax Holiday states are centered in the most common weekend holding a STH on the country for a given year

on a weekly basis from 2006 to 2014. Whereas Harper et al. (2003) were limited to following ten products at five stores in a pair of metropolitan areas and Cole (2009) tracked total sales for computer models across the 48 contiguous states for 30 weeks in 2007, we follow an array of 63 holiday-eligible school supplies on a weekly basis across the contiguous 48 states for nine years, a period that includes 117 state sales tax holidays. This larger scale allows us to overcome the limitations of these earlier studies related to specificity of time, location, and product type. Ultimately, we are able to confirm the previous findings in fall back-to-school tax holidays observed in Cole (2009) and Harper et al. (2003): in fall back-to-school tax holidays retailers appear to overshift the tax savings to consumers through lower pretax price adjustments. However, in the Tennessee spring-time natural experiment, we find that some retailers return only the amount of the tax, with some heterogeneity among retailers that leads to some retailer recapture of the tax incidence. We further substantiate our interpretation that seasonality drives the overshifting result in the fall-season cases through a placebo test and examination of fall Tennessee tax holidays.

The next section provides further background on sales tax holidays and the Nielsen data.
Section 3 provides analysis of the data across all school-supply sales tax holidays, before proceeding to the natural experiment case of Tennessee in section 4. Section 5 explores heterogeneity in the main Tennessee results across product modules and retailer classifications. The paper concludes with a summary and discussion in section 6.

4.2 Nielsen Data, State Sales Tax Holidays, and Empirical Strategy

This study employs the Nielsen Marketing Retail Scanner Data set (“the Nielsen data”) from 2006 to 2014. The Nielsen data consist of weekly purchase and pricing data from participating retail stores across all 48 contiguous U.S. states. The data are weekly reports summarizing transactions occurring in retailers’ point-of-sale systems, registering when a product UPC code is scanned and capturing both price and quantity. The UPC codes identify a specific product, e.g. “BIC Correction Fluid (0.7 OZ bottle)” or “Crayola Markers (10 Pack)”, which are applicable to any store from which they can be sold. About 35,000 retailers participate in the Nielsen data collection and carry a set of Nielsen-tracked product categories that includes food, non-food grocery, health and beauty aids, and selected general merchandise. The weekly (Sunday to Saturday) data report the total quantity sold by UPC code and the information required for a per-unit price calculation that accommodates the various retailer options for packaging, bundling, promotional sales, and other discount pricing. This information is relatable to a database of store characteristics that includes information on the retailer, geographic identifier at the three-digit zip code level, and the type of distribution channel (grocery stores, drug stores, mass merchandisers, and convenience stores).

---

5The description of the product at the UPC level in the Nielsen data indicates the brand, size, packaging, and style information, but the labeling is still broad enough that these descriptions could apply to multiple specific products. The UPC code is item specific, but the Nielsen name is not.

6This is a limitation of the data as sales tax holidays happening during a weekend will span over two different weeks. This is a limitation shared by the previous research in Cole (2009), which was treated by dropping the lower-sales week of the pair. Isolating the higher sales week may have been a clearer choice in that study because computers represent relatively large expenditures. However, these patterns are not as clear in our data on school supplies, which spans over a longer period of time and over a broader range of items.

7Individual store identities are confidential, however some information about the retailer is disclosed, including information about the parent company where applicable.
Table 4.1: Tax Rate Waived in States Implementing Back-to-School Sales Tax Holidays

<table>
<thead>
<tr>
<th></th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALABAMA</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>ARKANSAS</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>6</td>
<td>6</td>
<td>6.5</td>
<td>6.5</td>
</tr>
<tr>
<td>D.C.</td>
<td>5.75</td>
<td>5.75</td>
<td>5.75</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>FLORIDA</td>
<td>6</td>
<td>6</td>
<td>X</td>
<td>X</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>GEORGIA</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>X</td>
<td>X</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>ILLINOIS</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>6.25</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>LOUISIANA*</td>
<td>X</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>MASSACHUSETTS*</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>X</td>
<td>6.25</td>
<td>6.25</td>
<td>6.25</td>
<td>6.25</td>
<td>6.25</td>
</tr>
<tr>
<td>NEW MEXICO</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5.125</td>
<td>5.125</td>
<td>5.125</td>
<td>5.125</td>
<td>5.125</td>
</tr>
<tr>
<td>NORTH CAROLINA</td>
<td>4.5</td>
<td>4.25</td>
<td>4.25</td>
<td>4.25</td>
<td>5.75</td>
<td>4.75</td>
<td>4.75</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>SOUTH CAROLINA</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>TENNESSEE</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>VERMONT*</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>VIRGINIA</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4.3</td>
<td>4</td>
<td>4.3</td>
</tr>
</tbody>
</table>

X' indicates the state did not hold a back to school sales tax holiday in that year.
* Massachusetts, Louisiana and Vermont held sales tax holidays of Tangible Property Products, which included school supplies.

In using scanner data, the Nielsen reports offer the advantage of capturing the effective price paid by the consumer, which is the posted price less any of the variety of discounting strategies the retailer may have employed (e.g., coupons, sales, bundling). A disadvantage of this data is that we do not have price data on products that went unsold during the week, and consequently our product analysis is limited to high-frequency items. We selected UPC codes with at least 200,000 weekly observations in 2014 and one million weekly observations over the entire 2006 to 2014 period. These criteria provide us with over 175 million weekly observations, which is large enough that even a simple regression can take a day to complete on a high-throughput computing cluster.

Table 4.1 summarizes school-supply sales tax holidays in the contiguous U.S. from 2006 to 2014. While some states use tax holidays to apply to products other than school supplies, these policies are less common, more varied, and generally less likely to apply to the types of products appearing in the Nielsen data. For example, in addition to the school supply sales tax holiday, we used Stata 14’s multicore processor version on a system comprising 228 general-access compute nodes, where each node was equipped with two Intel Xeon E5-2650 v2 8-core processors containing 32 gigabytes of RAM.
Texas has an early spring holiday for hurricane and disaster preparedness supplies, and an early summer tax holiday for ENERGY STAR products and air conditioners. The hurricane and disaster supplies are not items that appear with enough frequency to construct long and wide panels of sales data, while ENERGY STAR products are not in Nielsen participating retailer data.

Table 4.2: Nielsen Product Modules Containing UPCs Categorized as “School Supplies”

<table>
<thead>
<tr>
<th>Product Module Code</th>
<th>Product Module Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>7630</td>
<td>School and office paper forms</td>
</tr>
<tr>
<td>7631</td>
<td>Report covers and sheet protectors</td>
</tr>
<tr>
<td>7632</td>
<td>Correction fluid and erasers</td>
</tr>
<tr>
<td>7633</td>
<td>Personal planners, binders, and folders</td>
</tr>
<tr>
<td>7634</td>
<td>Dividers, tabs, labels, and tags</td>
</tr>
<tr>
<td>7635</td>
<td>School and office storage and dispensers</td>
</tr>
<tr>
<td>7636</td>
<td>Home school and office combinations</td>
</tr>
<tr>
<td>7646</td>
<td>Adhesive note pads</td>
</tr>
<tr>
<td>7652</td>
<td>Crayons</td>
</tr>
<tr>
<td>7653</td>
<td>Pencils - colored</td>
</tr>
<tr>
<td>7659</td>
<td>Markers</td>
</tr>
<tr>
<td>7660</td>
<td>Pens &amp; pencils</td>
</tr>
</tbody>
</table>

All items are found in General Merchandise department files of the Nielsen data.

The Nielsen data do not indicate the sales taxes collected by the retailers on the products. Generally, states require retailers to report UPC codes if the retailers use UPC codes in determining taxability, but cases of states listing taxable UPC codes are rare. More commonly, retailers (with guidance from the state) interpret the legislation for both sales taxes and tax holidays to determine the status of a particular product. For the purposes of this study, we drew UPC codes from product modules that were included in the “School Supplies” group of general merchandise. Table 4.2 lists the module descriptions and associated codes. The 63 products in our dataset identified by UPC codes are found in these twelve modules, after applying the aforementioned frequency criteria that resulted in 175 million observations. Table 4.3 provides some summary statistics on weekly pre-tax prices and quantities of these products by store. Because retailers have discretion in applying the tax holiday exemptions, it is possible that we have drawn products that were not always eligible for the school-supply sales tax
holidays. However, we selected the product modules to represent a conservative list that avoids such measurement error. If we were to exclusively use this list to produce an estimate of state tax revenue lost to the holiday, we would certainly produce an underestimate due to likely missed items that stores do include as part of the holiday.

Finally, another limitation of the data is that the geographic identifier is the three-digit zip code, which does not allow us to match the data to local sales tax rates. Estimates by Drenkard and Kaeding (2016) indicate that local sales taxes account for about 19% of the combined rate in the 38 states with both taxes. For tax holidays, states vary on whether local government participation is mandatory. In a robustness check discussed later in this paper, we limit our data to a subsample with states that do not charge local sales taxes and we find almost identical results.

Table 4.3: Selected Statistics on Weekly Store Prices and School Supply Sales by Sample

<table>
<thead>
<tr>
<th>Sample</th>
<th>N</th>
<th>Unit Price (USD)</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Dev.</td>
<td>Mean</td>
</tr>
<tr>
<td>Total</td>
<td>175,089,295</td>
<td>2.43</td>
<td>4.46</td>
</tr>
<tr>
<td>STH States &amp; Year</td>
<td>52,499,327</td>
<td>2.32</td>
<td>4.12</td>
</tr>
<tr>
<td>Non-STH States &amp; Year</td>
<td>122,589,968</td>
<td>2.47</td>
<td>4.61</td>
</tr>
<tr>
<td>Tennessee</td>
<td>4,512,999</td>
<td>2.23</td>
<td>3.99</td>
</tr>
</tbody>
</table>

Note: N is the number of weekly reports in the Nielsen database. Quantity is the number of units at which the price is reported for a school supply product.

4.2.1 Estimation Strategy

The archetypical study on commodity tax incidence specifies the tax-exclusive price \( p \) received by the seller of a good to be some mark-up over marginal cost. Because a retailer must collect the ad valorem sales tax of rate \( \tau \) at the point of sale, the tax-inclusive price of \( p(1+\tau) \) determines the quantity demanded by their consumers. Consequently, the tax-exclusive price can be specified as an unknown function of both cost and tax rate. A semilog specification provides a linear approximation of the unknown function to determine the pretax price of good.
\[ \ln(p_{isw}) = \gamma \cdot \tau_{isw} + \mu_i + \mu_s + \mu_{tx} + \mu_m + \mu_Y + \epsilon_{isw}, \]  

(4.1)

where the use of a series of fixed effects in \( \mu \) is intended to capture the cost structure. Specifically, month fixed effects (\( \mu_m \)) capture seasonality, year fixed effects (\( \mu_Y \)) capture annual events, and store (\( \mu_s \)) and UPC (\( \mu_i \)) fixed effects are intended to sweep away time-invariant cost and demand influences. Additionally, this paper introduces the idea of a state-tax-regime fixed effect in (\( \mu_{tx} \)), which is defined to be a unique identifier for each observed state-sales tax rate combination that appears in the data set. The potential value of a sales tax holiday rises directly with the state’s general tax rate. This rate varies across states, and for some states also varies over time. Therefore, direct estimation of \( \gamma \) in Equation 4.1 using only state fixed effects would result in \( \tau \) offering variation from both sales tax holidays and general changes in the tax rate. Since the ultimate aim of the paper is to produce evidence only on the temporary tax cuts created by sales tax holidays, we introduce the state tax regime concept so that the only variation in \( \tau \) occurs by the holiday when the rate falls to zero, and other sales tax rate variation from other policy changes will be swept away by the tax regime fixed effects\(^{11} \).

School tax holidays waived an average tax rate of 5.4 percent over the 117 events in our data. The within-variation on the tax rate is negative as tax holidays cause rates to decline to zero. The phenomenon of retailers recapturing the tax savings through higher prices is evidenced by \( \hat{\gamma} \) revealing a negative correlation. A percentage point decrease in the sales tax rate (\( \Delta \tau = -1 \)) on the tax holiday yields a \( \hat{\gamma} \) percent change in pretax price, allowing for a direct interpretation on the share of shifting. If \( \hat{\gamma} = -1 \), for example, then the waiving of a

\(^{11}\) Using Table 4.1 to illustrate the tax regime fixed effects, Alabama represents only one single tax regime as it never changed the sales tax rates during the entire time. The same is true of the District of Columbia, Florida, Georgia, Louisiana, Illinois, Missouri, Tennessee, Texas, and Vermont. Arkansas represents two tax regimes, one for their 6 percent period (2011-2012) and one for their 6.5 percent phase (2013-2014). Other states with two tax regimes include Massachusetts, New Mexico, South Carolina, and Virginia. Only North Carolina has more than two tax regimes for this fixed effect specification.
sales tax rate of 4% would be matched with a 4% increase in pretax price and the retailers would recapture the entirety of the tax savings. Consumers receive the full tax savings or more when \( \hat{\gamma} \geq 0 \).

One potentially problematic issue with specification in Equation 4.1 in the context of this study is that monthly fixed effects might not adequately address the back-to-school season because of variation in the dates that students return to school. Even within a given state the exact dates of when students return to classes varies across school districts. However, judging from the observed pattern of consumer purchases of school supplies, targeting of a date that is beneficial to most families concerned with the return to school is probably the explanation for annual tax holidays within states shifting on a year-to-year basis. Even with these movements, year-by-year there is significant overlap in the adopted dates among states that use tax holidays for school supplies. In an effort to recenter the regression more closely with the return to school, for each state we attempt to identify with a “detrending” counter that records the days to/from the back-to-school shopping weekend, represented as in Equation 4.2:

\[
\ln(p_{isw}) = \gamma \cdot \tau_{isw} + \mu_i + \mu_s + \mu_t + \mu_m + \mu_Y + \theta_d + \epsilon_{isw}, \tag{4.2}
\]

We will present results that employ this approach as our best effort to treat the school start seasonality that we regard as a significant lurking concern as a source of bias. The next section shows this is an inadequate correction and that seasonality remains a concern, motivating our attention to the Tennessee case study.
4.3 Investigation of Seasonality in Back-to-School-Supply Sales Tax Holidays

Table 4.4 presents the results of estimating equations 4.1 and 4.2 against different counterfactuals for all back-to-school tax holidays occurring in late-summer early fall from 2006 to 2014\textsuperscript{12}. All specifications report heteroskedastic standard errors clustered by state in parentheses and include fixed effects for UPC, store, state tax regime, month, and year. The coefficients report the effect of the state’s sales tax rate falling to zero for the holiday on the pretax price per unit. The specifications in Table 4.4 differ according to their counterfactual source of variation based on sample splitting and indicators to adjust for differences in state school start dates. Specification (A) reports results for all states with no variables that would center the groups across consistent school start dates, so variation arises from differences in the magnitude of the tax rate cut caused by the holiday, the timing of the holiday relative to school start date, and the existence of a sales tax holiday. Specification (B) includes the school start date counter and is intended to detrend the data relative to the start dates. Specifications (C) and (D) include only states that hold tax holidays at some point in the time period, which potentially reduces some possible selection bias in the heterogeneity between those states inclined to have these policies and those which do not.

Table 4.4: Regression Results For Back-to-School Tax Holiday Price Effects, 2006-2014

<table>
<thead>
<tr>
<th>Control Variables</th>
<th>All States</th>
<th>Only States with Tax Holidays</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales Tax Rate</td>
<td>0.400*</td>
<td>0.392*</td>
</tr>
<tr>
<td></td>
<td>(0.224)</td>
<td>(0.224)</td>
</tr>
<tr>
<td>School Start Detrend</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>R2</td>
<td>0.664</td>
<td>0.664</td>
</tr>
<tr>
<td>N (Millions)</td>
<td>175.1</td>
<td>175.1</td>
</tr>
<tr>
<td>Number of Clusters</td>
<td>49</td>
<td>49</td>
</tr>
</tbody>
</table>

Notes: Sales Tax Rate falls to zero during tax holiday. All specifications include UPC, retailer, state tax regime, month, and year fixed effects. Statistical significance reported at the 1 percent (***) , 5 percent (**), and 10 percent (*) level. Heteroskedastic robust standard errors clustered by state reported in parentheses.

\textsuperscript{12}These results are generated with the estimator described in Correia (2016) using “reghdfe” in Stata Version 14.
The findings of Table 4.4 basically confirm the previous work indicate that pretax prices are directly correlated with the tax holiday changes in tax rates (Cole, 2009; Harper et al., 2003), which is not consistent with the conventional expectation from partial equilibrium supply and demand models of tax incidence under perfect competition. For every percentage point decrease in the states sales tax rate caused by the holiday, the pretax price decreases by approximately 0.39 to 0.54 percent, and these effects are statistically significant at ten and five percent levels. Specifications (A) and (C) in Table 4.4 are the most comparable to Cole’s (2009) main findings in modeling. This similarity suggests that Cole’s findings were not artifacts of the time period or the focus on computers, as the same result manifests for a variety of school supplies in our study over the years 2006 to 2014.

The results in Table 4.4 hold up to a battery of robustness checks that are not reported but are available upon request. For instance, reducing the time period under study to a shorter range surrounding the sales tax holiday weeks (e.g. six to ten weeks), when most of the variation is observed, produces almost identical magnitudes in effect size to Table 4.4 with statistical power declining as the time band narrows. Another possibility is that results are driven by measurement error caused by unobserved local tax rates. States vary on whether it is mandatory for local governments with sales taxes to participate in the tax holiday, and Nielsen data does not provide geographic data that allows for matching to a particular local sales tax rate, nor is information available on the local sales tax charged across sometimes overlapping municipalities, localities and districts. However, if the sample is pared down to states with no local sales taxes, the results are nearly identical to those reported in Table 4.4. This consistency suggests that the findings are not driven by correlated measurement error related to missing local sales tax rates.

---

13We run the specifications for a subsample of states with no local sales taxes (or no sales taxes), including: Connecticut, Delaware, Indiana, Kentucky, Maine, Maryland, Massachusetts, Michigan, Montana, New Hampshire, Oregon, Rhode Island and the District of Columbia. Among them, only two states implemented Back to School Sales Tax Holidays during the time span of the analysis, Massachusetts and the District of Columbia.
As previously argued, policy makers might reasonably target tax holidays to be near the dates students return to school and parents wish to shop, which also coincides with when retailers would be running their own back-to-school sales, and likewise motivated efforts for inclusion of controls for proximity to school start date. Since states overwhelmingly allow school districts some flexibility in determining their own start date, a statewide tax holiday might only serve as a focal point for narrowing the dispersion of dates over which retailers run their back-to-school specials for local consumers. A concern then raised by the results presented so far is that the specifications do not adequately capture this seasonal adjustment, and we obtain reverse causality in the results because the low prices predict sales tax holidays.

Figure 4.1 illustrates per unit prices and volumes of school supply sales in this time period for states with and without tax holidays. These figures demonstrate that there are substantial seasonal market structure changes in both holiday and non-holiday states, likely owing to changes in consumer demand patterns surrounding the end of summer and the return of fall schooling. To explore further this seasonality concern, Table 4.5 presents the results of a placebo test. In this test, we include states that never run a tax holiday, as well as states in years without a tax holiday over the 2006 to 2014 period. For example, from Table 4.1, Arkansas is included in the placebo group and coded as if they held a sales tax holiday over the 2006 to 2010 period, but excluded from the sample during the period where they actually conducted the holiday (2011 to 2014). North Carolina is similarly excluded from 2006 to 2012 but coded as a placebo holiday for 2013 and 2014 when they actually ran no such program. Seven of our 16 tax holiday states provide these placebo opportunities over the sample period. If overshifting is caused by the sales tax holiday, then these placebos should produce results closer to zero since no actual holiday occurred\(^\text{14}\). Instead, Table 4.5 demonstrates the opposite, with larger positive effect sizes of just over one percent. The difference in results from Table

\(^{14}\)To code the week of the sales tax holiday in the placebo years, we assigned to the state the same week that was most commonly observed in the years that they actually held a sales tax holiday.
4.4 to Table 4.5 might be suggestive that the tax holiday reduced the amount of overshifting that occurred, and the actual point estimates were simply swamped by the seasonality effects. Of course, this cannot be definitively concluded from the placebo test, but it is reasonable evidence that seasonality in the timing of the sales tax holidays is such that they tend to fall on weeks where prices would be low in any case and that this reverse causality is driving the results provided in Table 4.4. We now turn to our preferred case, the quasi-experimental case of Tennessee.

Table 4.5: Regression Results for Placebo Back-to-School Tax Holidays, 2006 to 2014

<table>
<thead>
<tr>
<th>Control Variables</th>
<th>All States in Placebo Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(A)</td>
</tr>
<tr>
<td>Sales Tax Rate</td>
<td>1.090***</td>
</tr>
<tr>
<td></td>
<td>(0.238)</td>
</tr>
<tr>
<td>School Start Detrend X</td>
<td>X</td>
</tr>
<tr>
<td>R²</td>
<td>0.661</td>
</tr>
<tr>
<td>N (Millions)</td>
<td>122.6</td>
</tr>
<tr>
<td>Number of Clusters</td>
<td>43</td>
</tr>
</tbody>
</table>

Notes: Sales Tax Rate falls to zero during (placebo) tax holiday. All specifications include UPC, retailer, state tax regime, month, and year fixed effects. Statistical significance reported at the 1 percent (***) and 5 percent (**), and 10 percent (*) level. Heteroskedastic robust standard errors clustered by state reported in parentheses.

4.4 Main Results: Quasi-Experimental Case of Tennessee

If the underlying problem is an endogenous targeting of low prices via deliberate timing of sales tax holidays for back-to-school sales, then a cleaner set of results might be found in tax holidays that are inclusive of these school supplies but not a part of the back-to-school season. An example of this occurrence is found in Tennessee, which added a spring sales tax holiday to its schedule in 2007 and 2008. Tennessee’s first tax holiday on school supplies was held in fall of 2006. This timing makes Tennessee a somewhat late adopter of tax holidays, but robust tax collections made the legislature willing to try to return money to taxpayers with an additional tax credits and exemptions the following spring. The legislation extended numerous tax credits
or exemptions, mostly for businesses, and in one provision stated that Tennessee Code 67-6-393 was to be amended by adding a provision that added April 27, 2007 to April 29, 2007 as eligible dates where the section would be applicable. This was the statutory code that enabled the sales tax holiday the previous spring, and while of course these holidays could not be regarded as back-to-school tax holidays, the legislation consequently specified the same set of goods for the spring holiday and were inclusive of school supplies, and we can find examples of retailers advertising these products in their marketing materials. Starting in 2009, Tennessee reverted to the more common practice of a single end-of-summer or beginning-of-fall holiday. Figure 4.3 illustrates that school supplies underwent no similar seasonal changes in Tennessee, which is in direct contrast to Figure 4.1 for the fall back-to-school holidays. Furthermore, the general price and sales trends surrounding the off-season Tennessee tax holiday are similar for both Tennessee and all other states. Tennessee’s 2007 and 2008 spring tax holidays offer the advantage of stripping away the seasonal variation. Table 4.6 presents the regression results for the Tennessee spring tax holiday. To avoid the inclusion of other school-supply tax holidays,

\[15\] The authorizing language can be found in Tennessee state legislative summaries Public Chapters 1019 (June 27, 2006), 600 (June 27, 2007), and 617 (March 11, 2008).
we limit the sample to only the first half of the calendar years in 2007 and 2008. Unlike the other tax holiday results that showed overshifting, the evidence in Table 4.6 is consistent with retailers increasing pre-tax prices in response to the tax holiday, albeit even where that finding is statistically significant the effect is qualitatively small. In column (A), comparing Tennessee to all states and including excluding a weekly time trend variable, the shifting parameter was the largest in absolute value at -0.03 and statistically significant at the 5 percent level. Since the waived tax rate in Tennessee was 7 percent, this correlation implies a 0.21% increase in pre-tax prices. Including a weekly time trend in (B) shrinks this effect further to just -0.019 that is statistically insignificant, albeit the precision is high as the 95 percent confidence interval puts the effect between 0.023 and -.060, which even at the most negative range of that interval suggests just a 0.42 percent increase in pre-tax prices. If we compare Tennessee only to other states that have tax holidays in the fall (columns C and D), the results are not substantively changed.

Table 4.6: Regression Results for TN Spring Tax Holiday, January through June of 2007 and 2008

<table>
<thead>
<tr>
<th>Control Variables</th>
<th>All States</th>
<th>Only States with Tax Holidays</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales Tax Rate</td>
<td>-0.0304**</td>
<td>-0.0195</td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.021)</td>
</tr>
<tr>
<td>Week Trend Variable</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>R2</td>
<td>0.788</td>
<td>0.798</td>
</tr>
<tr>
<td>N (Millions)</td>
<td>17.1</td>
<td>17.1</td>
</tr>
<tr>
<td>Number of Clusters</td>
<td>49</td>
<td>13</td>
</tr>
</tbody>
</table>

Notes: Sales Tax Rate falls to zero during tax holiday. All specifications include UPC, retailer, state tax regime, month, and year fixed effects. Statistical significance reported at the 1 percent (**), 5 percent (*), and 10 percent (*) level. Heteroskedastic robust standard errors clustered by state reported in parentheses.

One concern is that Tennessee might just be an across-the-board unusual case, no matter if we were looking at these spring quasi-experiments or not. Since Tennessee also conducted fall back-to-school tax holidays, we can check to see how different their case is from the full sample of fall results. Table 4.7 compares the Tennessee case for the fall back-to-school sales all other
states excluding those states with contemporaneous fall back-to-school programs (specifications A and B), as well as to states that have never held such tax holidays (specifications C and D). These results demonstrate a positive correlation between the waived sales tax rate and the pre-tax price, again suggesting overshifting. While the effect sizes are about half those in Table 4.4, Tennessee’s effect sizes in Table 4.7 have overlapping confidence intervals, so Tennessee results for the fall likely experience the same seasonality problems for identification as seen in the rest of the country in the fall season.

Table 4.7: Regression Results for Tennessee Fall Sales Tax Holiday, 2006 to 2014

<table>
<thead>
<tr>
<th>Dependent Variable: $\ln(\text{Pre-Tax Price Per Unit})$</th>
<th>All States</th>
<th>Only States with Tax Holidays</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control Variables</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sales Tax Rate</td>
<td>0.145*</td>
<td>0.117</td>
</tr>
<tr>
<td></td>
<td>(0.081)</td>
<td>(0.080)</td>
</tr>
<tr>
<td>Week Trend Variable</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>0.657</td>
<td>0.657</td>
</tr>
<tr>
<td>N (Millions)</td>
<td>117.8</td>
<td>100.8</td>
</tr>
<tr>
<td>Number of Clusters</td>
<td>44</td>
<td>34</td>
</tr>
</tbody>
</table>

Notes: Sales Tax Rate falls to zero during tax holiday. All specifications include UPC, retailer, state tax regime, month, and year fixed effects. Statistical significance reported at the 1 percent (***), 5 percent (**), and 10 percent (*) level. Heteroskedastic robust standard errors clustered by state reported in parentheses.
4.5 Heterogeneity in Tennessee Spring Results

The previous section made the case for the Tennessee spring 2007 and 2008 as the best identified causal effects, so we extend the analysis of this case to explore heterogeneity. In particular, we are interested if there are products or types of retailers that drive the results found in Table 4.6. We will consider three types of heterogeneity: product module, store chain size, and retailer type. The sample will be split and estimated for each type according to the same specification found in column A of Table 4.6 where the full sample parameter was -0.034\textsuperscript{16}.

Figure 4.4: Point Estimates and Confidence Intervals for Sales Tax Rate by Product Module

Figure 4.4 depicts the point estimate and confidence intervals after splitting the sample by product module. This grouping of UPCs by product module (i.e. Paper Forms, Erasers, binders, etc.) is defined by the data source provider (Nielsen). There are a couple of products with a substantive degree of retailer recapture of the tax, particularly for paper forms, dividers, storage folders, and notepads recaptured about 0.15 to 0.2 percent for every one percent waived. Crayons and binders demonstrated about 0.1 percent overshifting of the tax savings back to consumers. Otherwise, the results are rather precisely estimated near zero, there is not a great

\textsuperscript{16}The results of the heterogeneity are nearly identical across the alternative specifications and are available upon request.
deal of heterogeneity across product modules.

Figure 4.5 looks at heterogeneity by retailer size, as measured by the number of stores the retailer owns nationwide\(^{17}\). The chains are split according to small (less than 100 stores), moderate (100 to 3,500), and large (more than 3,500)\(^{18}\). These are admittedly ad hoc thresholds selected from observing grouping in the raw data. Interestingly, the small and large chains demonstrate some modest recapturing of the tax in the range of about -0.2 percent for every percentage point waived of the sales tax, while the moderate size chains demonstrate over-shifting of a little over 0.2 percent. So the degree of shifting does seem to be related to chain size.

Figure 4.5: Point Estimates and Confidence Intervals for Sales Tax Rate by Retailer Size

We can also split the sample according to the type of retailer, which has four categories: convenience stores, drug stores, grocery stores, and mass merchandizers\(^{19}\). Figure 4.6 shows

\(^{17}\)For clarification, 4.5 is still results for Tennessee stores, but the definition of retailer size is based on how many stores the retailer holds nationwide in 2014, the last year in our sample.

\(^{18}\)In our sample we can identify the parent to which a store belongs and classify by the size of the retailer nationwide: Small retailers with up to 100 stores nationwide account for 3% of the weekly reports and 4% of the total number of stores in our sample. In contrast, big chain retailers, with more than 3,500 stores nationwide, account for 55% of the weekly reports and 56% of the total number of stores. Within Tennessee in 2008, there are 16 parents and according to the nationwide classification small retailers represent 6% of the total weekly reports in our sample and 12% of the number of stores, whereas big retailers account for 58% of their reports and 45% of the stores in the state.

\(^{19}\)The share of stores in the full sample and the Tennessee spring sample is 3% and 2% for convenience stores, 35% and 36% for drug stores, 29% and 31% for grocery stores, and 33% and 31% for mass merchandisers.
convenience stores to nearly recapture all of the waived tax with a corresponding increase in the pre-tax price.\textsuperscript{20} Drug stores also demonstrate recapturing of the waived tax, but to a much smaller degree. In contrast, mass merchandizers demonstrate overshifting of 0.087 percent for each percent of the tax waved. One explanation for this could be that convenience and drug stores are selling to consumers when they are more price inelastic, conferring the ability to capture the tax savings. It should be noted, however, that mass merchandizers account for about one-third of all transactions in Tennessee during the sample period, whereas convenience stores accounted for just 0.13 percent. Drug stores possess the largest market share in Tennessee at 46.6 percent.

Figure 4.6: Point Estimates and Confidence Intervals for Sales Tax Rate by Retailer Type

Clearly, in looking at Figures 4.5 and 4.6, there is some intriguing heterogeneity based on characteristics of the retailer. However, it is also clear that the consumers do retain most of the tax savings even if the tax is not overshifted to them, and that this is true where they overwhelmingly do their shopping. The only case where retailers are able to recapture the majority of the tax savings is in convenience stores, where a fraction of the school supplies are sold.

\textsuperscript{20}The 95 percent confidence interval on convenience stores in Figure 4.6 is -0.725 to -0.942.
This paper advances the literature on the consumer incidence of state sales tax holidays by exploiting the unique natural experiment of Tennessee’s spring 2007 and 2008 sales tax holiday. These holidays were executed for products identical to the other school supply tax holidays but in April rather than August. We argue that this is advantageous because this case avoids important seasonality problems that occur when during typical tax holidays for school supplies that occur in the late summer and early fall. We demonstrate that this seasonality introduces a positive bias that tends to drive the conclusions towards finding an overshifting of the waived sales tax to consumers. We provide evidence that this is not a concern for the school supply tax holidays of Tennessee in spring, and in these cases we find that on average consumers receive nearly the full amount of the tax waved. In exploring the data’s heterogeneity, we find some important differences in retailer characteristics. In particular, drug and convenience stores demonstrate tax savings recapture through higher pre-tax prices, but only the convenience stores receive the majority of the tax savings and they represent an extremely small market share of the school supply retail sales. Mass merchandizers, on the other hand, do depict some amount of overshifting albeit to a much smaller degree than would be suggested by the seasonality-biased estimates from the fall tax holidays. Generally, the conclusion holds that consumer households are the primary beneficiaries of the sales tax holidays.

Although the richness of the data allowed for contributions beyond the previous research, there remain limitations that future work might address. First, the geographic identifiers on the retail stores did not allow us to tie in local sales tax rates. Second, the dataset only records the prices of items sold, which forces us to follow items with high levels of within-group frequency. A “posted” price approach could expand the number of items studied, but stores employ a variety of price discrimination tactics so that consumers may pay different effective prices, making such an investigation challenging but potentially fruitful. Another
line of inquiry could try to understand the interactions of different actors in price setting, as we see in our specifications that discover heterogeneity by retailer characteristics. Finally, exploring the behavior of different producers and retailers could inform policymakers of the best channels to use holiday-style transfers to specific consumer groups. This paper was the first of its kind to explore this form of heterogeneity, and a deeper theory-driven investigation with data that is richer on retailer characteristics may be informative to policy makers.
Bibliography


Appendix A

A matter of design in Soda Taxes: Tax Sugar instead of Liquid

A.1 Math Appendix

\[ x_1 - x_0 = \frac{t}{(\alpha_L - \alpha_H)(x_1 + x_0) + 2(\alpha_H S_H - \alpha_L S_L)} \]

Equation 2.4 determines the effect of a tax on beverage \( H \) over the demand of the two beverages. From the assumptions and definitions, it is known that: \( t > 0, S_H > x_1, x_0 > S_L \), and the sign of the expression in Equation 2.4 depends on the sign of the denominator. This proves that regardless of the choice of \( \alpha_H \) and \( \alpha_L \), a tax on \( H \) would cause an increase in the demand for the untaxed beverage, \( L \), as given by \((x_1 - x_0) > 0\). To do so, let’s consider three possible cases: \( \alpha_L \) being greater than, lesser than, or equal to \( \alpha_H < 0 \).

Case 1: \((\alpha_L - \alpha_H > 0)\)

\[
S_H \geq x_0, x_1 \rightarrow 2\alpha_H S_H \geq \alpha_H x_0 + \alpha_H x_1 \\
x_0, x_1 \leq S_L \rightarrow 2\alpha_L S_L \leq \alpha_L x_0 + \alpha_L x_1 \\
\text{Adding} \rightarrow 2\alpha_L S_L + \alpha_H x_0 + \alpha_H x_1 < 2\alpha_H S_H + \alpha_L x_0 + \alpha_L x_1
\]

\[
\text{Re-arranging} \rightarrow (\alpha_H - \alpha_L)(x_0 + x_1) < 2(\alpha_H S_H - \alpha_L S_L) \\
\text{times } (-1) \rightarrow (\alpha_L - \alpha_H)(x_0 + x_1) > 2(\alpha_H S_H - \alpha_L S_L)
\]

This last expression renders the whole denominator of Equation 2.4 positive. Given that \( t > 0 \), this case makes \((x_1 - x_0) > 0\). Thus, the market share for \( L \) increases and for \( H \) decreases, as
consumers from $x_0$ to $x_1$ migrate to $L$.

**Case 2: $(\alpha_L - \alpha_H < 0)$**

\[ \alpha_L > \alpha_H \& S_H > S_L \rightarrow \alpha_H S_H - \alpha_L S_L > 0 \]

\[ \alpha_L > \alpha_H \& S_H \geq x_0, x_1 \rightarrow 2\alpha_H S_H \geq \alpha_L (x_0 + x_1) \]

\[ \alpha_L > \alpha_H \& S_L \leq x_0, x_1 \rightarrow 2\alpha_L S_L \leq \alpha_H (x_0 + x_1) \]

Adding \[ 2\alpha_L S_L + \alpha_L (x_0 + x_1) < 2\alpha_H S_H + \alpha_H (x_0 + x_1) \]

Re-arranging \[ (\alpha_L - \alpha_H)(x_1 + x_0) < 2(\alpha_H S_H - \alpha_L S_L) \]

Considering that the last term of the denominator of Equation 2.4 is positive, the complete expression is always positive, and the result is the same as in the previous case with $(x_1 - x_0) > 0$.

**Case 3: $(\alpha_L - \alpha_H = 0)$** Again, in this case $(x_1 - x_0) > 0$. The first term of the denominator of Equation 2.4 disappears, and the second is positive given, that $\alpha_L = \alpha_H$ and $S_H > S_L$. These three cases illustrate that the result is not dependent on what values the parameter $\alpha_i$ takes.
A.2 Trends for Additional Series

Figure A.1: Event Studies - Sugar-Sweetened Beverages Log Grams of Sugar for Different Control Groups

Note: 2014 - IV used as reference and not included on the graph.
Note: 2014 - IV used as reference and not included on the graph.
Figure A.3: Event Studies - Sugar-Sweetened Beverages Log Price Per Gram of Sugar for Different Control Groups

Control: Berkeley's Neighboring localities

Control - Remaining CA

Control - CA, No neighbors

Control - Neighboring States

Control - Western States (2 borders)

Sum of Sq. Deviations Pre-Period: 0.00171

Sum of Sq. Deviations Pre-Period: 0.00151

Sum of Sq. Deviations Pre-Period: 0.00159

Sum of Sq. Deviations Pre-Period: 0.00222

Note: 2014 - IV used as reference and not included on the graph.
Figure A.4: Event Studies - Sugar-Sweetened Beverages Log of Sugar Density per Liquid Ounce for Different Control Groups

Note: 2014 - IV used as reference and not included on the graph.
A.3 Parallel Trends Clustering at the 3-Digit Zip Code Level

Figure A.5: Event Studies - Sugar-Sweetened Beverages Log Liquid Ounces Clustering at the 3-Digit Zip Code

Note: 2014 - IV used as reference and not included on the graph.
Appendix B

Cash Subsidies to improve Student Loan programs for Low-Income Students

B.1 Probability of Receiving a Cash Subsidy Across Geographies

Table B.1: First Stage: Discontinuity in the Probability of Receiving a Subsidy and on the Amount of Subsidy Received

<table>
<thead>
<tr>
<th></th>
<th>14 Main Cities</th>
<th>Other Cities</th>
<th>Rural Areas</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>P(Subsidy</td>
<td>Sisben III)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Robust Coefficient</td>
<td>-1351.8***</td>
<td>-1358.6***</td>
<td>-1336.7***</td>
</tr>
<tr>
<td></td>
<td>(169.2)</td>
<td>(159.5)</td>
<td>(144.8)</td>
</tr>
<tr>
<td>N</td>
<td>6,806</td>
<td>6,770</td>
<td>6,770</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>7.46</td>
<td>7.03</td>
<td>6.55</td>
</tr>
<tr>
<td>Ineligible Baseline</td>
<td>731</td>
<td>733</td>
<td>739</td>
</tr>
<tr>
<td>E(Subsidy (USD)</td>
<td>Sisben III)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Robust Coefficient</td>
<td>-0.162***</td>
<td>-0.175***</td>
<td>-0.175***</td>
</tr>
<tr>
<td></td>
<td>(0.0406)</td>
<td>(0.0393)</td>
<td>(0.0387)</td>
</tr>
<tr>
<td>N</td>
<td>10,827</td>
<td>10,780</td>
<td>10,780</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>16.17</td>
<td>16.08</td>
<td>15.64</td>
</tr>
<tr>
<td>Ineligible Baseline</td>
<td>1,390</td>
<td>1,389</td>
<td>1,388</td>
</tr>
</tbody>
</table>

Coefficients from difference across local polynomial regression of order 2. Bandwidth selection method: common MSE-optimal bandwidth selector (MSERD). All coefficients are robust bias-corrected (**). Statistical significance reported at the 1 percent (**), 5 percent (**), and 10 percent (*) level. Heteroskedastic standard errors in parentheses clustered at the departmental (Colombian political division) level.
Figure B.1: Probability of Receiving a Subsidy ($P(D_i)$) and SISBEN Score across Geographical Areas

**Panel A: Main Cities**

- Sample average within bin
- Polynomial fit of order 4
- Unrestricted local polynomial regression

**Panel B: Other Cities**

- Sample average within bin
- Polynomial fit of order 4
- Unrestricted local polynomial regression

**Panel C: Rural Areas**

- Sample average within bin
- Polynomial fit of order 4
- Unrestricted local polynomial regression

Optimal Bandwith = 5.52 SISBEN Units

Optimal Bandwith = 10.75 SISBEN Units

Optimal Bandwith = 14.12 SISBEN Units
Figure B.2: Subsidy Disbursed ($E(\text{Subsidy})$) and SISBEN Score

**Panel A: Main Cities**

E(Subsidy Disbursed|SISBEN Score) (Unrestricted local polynomial regression)

Optimal Bandwidth = 7.46 SISBEN Units

**Panel B: Other Cities**

E(Subsidy Disbursed|SISBEN Score) (Unrestricted local polynomial regression)

Optimal Bandwidth = 16.17 SISBEN Units

**Panel C: Rural Areas**

E(Subsidy Disbursed|SISBEN Score) (Unrestricted local polynomial regression)

Optimal Bandwidth = 11.68 SISBEN Units
B.2 McCrary test: Discontinuity in the Assignment Variable Across Geographies and Cohorts

Figure B.3: McCrary Test: Continuity of the Assignment Variable Across Geographical Area - 2011 Cohort
Figure B.4: McCrary Test: Continuity of the Assignment Variable Across Geographical Area - 2012 Cohort

Across All Areas
McCrary Test - Discontinuity on the Running Variable
2012 Cohort

Main Urban Areas
McCrary Test - Discontinuity on the Running Variable
2012 Cohort

Other Urban Areas
McCrary Test - Discontinuity on the Running Variable
2012 Cohort

Rural Areas
McCrary Test - Discontinuity on the Running Variable
2012 Cohort
B.3 Additional Variables Discontinuity Test Across Geographies

Figure B.5: Discontinuity Test on Additional Covariates - Main Cities

Note: Estimates of within optimal bandwidth local polynomial regression and bin averages.
Figure B.6: Discontinuity Test on Additional Covariates - Other Urban Areas

Note: Estimates of within optimal bandwidth local polynomial regression and bin averages.
Figure B.7: Discontinuity Test on Additional Covariates - Rural Areas

Note: Estimates of within optimal bandwidth local polynomial regression and bin averages.
B.4 Effect on Academic Outcomes across Geographical Areas

Figure B.8: Dropout Rates ($P(Dropout)$) and SISBEN Score

Panel A: Main Cities

Panel B: Other Cities

Panel C: Rural Areas

Optimal Bandwidth = 11.71 SISBEN Units

Optimal Bandwidth = 14.10 SISBEN Units

Optimal Bandwidth = 12.46 SISBEN Units

126
Figure B.9: First Year Dropout Rates ($P(Dropout \ 1^{st}\ Year)$) and SISBEN Score

**Panel A: Main Cities**

P(Dropout in 1st year|SISBEN Score)

Unrestricted local polynomial regression

Optimal Bandwith = 13.74 SISBEN Units

**Panel B: Other Cities**

P(Dropout in 1st year|SISBEN Score)

Unrestricted local polynomial regression

Optimal Bandwith = 20.08 SISBEN Units

**Panel C: Rural Areas**

P(Dropout in 1st year|SISBEN Score)

Unrestricted local polynomial regression

Optimal Bandwith = 14.81 SISBEN Units

127
Figure B.10: Graduation rates \( P(\text{Graduation}) \) and SISBEN Score

**Panel A: Main Cities**

- **Panel A.1:** Graph showing the relationship between the probability of graduation \( P(\text{Graduation}|SISBEN \text{ Score}) \) and the distance to the SISBEN III cutoff. The graph includes a polynomial fit of order 4 and an unrestricted local polynomial regression. The optimal bandwidth for the analysis is 10.76 SISBEN Units.

- **Panel A.2:** Graph showing the same relationship for other cities, with a polynomial fit of order 4 and an unrestricted local polynomial regression. The optimal bandwidth is 14.06 SISBEN Units.

- **Panel A.3:** Graph showing the relationship for rural areas, with a polynomial fit of order 4 and an unrestricted local polynomial regression. The optimal bandwidth is 12.65 SISBEN Units.

**Panel B: Other Cities**

**Panel C: Rural Areas**
**Felipe A. Lozano-Rojas, Ph.D.**

<table>
<thead>
<tr>
<th>School of Public and Environmental Affairs</th>
<th>felipelozanorojas.com</th>
</tr>
</thead>
<tbody>
<tr>
<td>1315 E 10th St., (412)</td>
<td>Skype: <a href="mailto:flozanor@iu.edu">flozanor@iu.edu</a></td>
</tr>
<tr>
<td>Bloomington, IN 47405</td>
<td>Phone: +1 (812) 929-6717</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Education</strong></th>
<th><strong>Indiana University Bloomington</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ph.D. in Public Affairs (2015 - 2020)</td>
<td><strong>Minor</strong>: Data Science</td>
</tr>
<tr>
<td><strong>Dissertation Project</strong>: Three Essays on the Unintended Consequences of Social Policies.</td>
<td></td>
</tr>
<tr>
<td><em>Essay I</em> - A matter of design in Soda Taxes: Tax sugar instead of volume</td>
<td></td>
</tr>
<tr>
<td><em>Essay II</em> - Cash Subsidies to improve Student Loan programs for Low-Income Students</td>
<td></td>
</tr>
<tr>
<td><em>Essay III</em> - Consumer incidence in sales tax holidays: Evidence from Tennessee</td>
<td></td>
</tr>
<tr>
<td><strong>Committee</strong>: Maureen Pirog, (co-chair), Justin Ross (co-chair), Kosali Simon &amp; Bradley Heim.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>National Taiwan University</strong></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th><strong>Research affiliations</strong></th>
<th><strong>SPEA, Indiana University (2015-present)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Research Assistant to Maureen Pirog, (2015-2016)</td>
<td></td>
</tr>
<tr>
<td>Research Assistant to Justin Ross &amp; Denvil Duncan (2016-2017)</td>
<td></td>
</tr>
<tr>
<td>Research Assistant to Kosali Simon (2018-2020)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>World Bank Colombia (2016-2019)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Research Consultant to Education Team lead by Pedro Cerdan-Infante.</td>
</tr>
</tbody>
</table>

|--------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|


---

**Research related to Soda Taxes**
- APPAM 40th and 41st Annual Fall Research Conference. 2018 and 2019.
- 2019 APPAM International Conference, Barcelona. 2019
- American Society of Health Economists Conference (ASHEcon). 2019
- APPAM 40th Annual Fall Research Conference. 2018.
- NTA meetings, 111th Annual Conference on Taxation. 2018.

**Research related to Sales Tax Holidays**

**Research related to Higher Education Finance**
- APPAM 38th Annual Fall Research Conference. 2016.

**Other Conferences**
- NTA meetings, 112th Annual Conference on Taxation, 2019.
- Association for Education Finance and Policy 41st Annual Conference.
Teaching

**SPEA, Indiana University**
Instructor of Record, Undergraduate Financial Management, 2017-2018-2019
Teaching Assistant, Statistics for Research in Public Policy, 2019
Official Tutor, Ph.D. Statistics for Research in Public Policy, 2018-2019

**Department of Economics, Universidad del Rosario - Colombia**
Instructor of Record, International Trade Theory, 2012-2014
Instructor of Record, Contemporary Finance Studies, 2012-2014
Instructor of Record, Financial Theory I, 2011

Professional Service & Relevant Experience

**Reviewer**
Health Affairs, International Journal of Public Sector Management

**Indiana University - Beth Israel Deaconess Medical Center (2018)**
Pro-bono designers of a new Discharged Worksheets for surgical procedures.
Team lead by Gabriel Brat, MD, and executed with Timothy Whitson, MDS; Neha Rawat, MDS and Himani Bhatt, MDS.

**Association of SPEA Ph.D. Students (2017, 2018)**
President of the Association (2019-2020)
Executive Board Member. Lead organizer Students Association’s conference.

**WEAI - 13th International Conference (2016-2017)**
Organizer - Panel Coordinator

**Icetex - Colombia Student Loan Agency**

Awards and Fellowships

**Colombia-Colciencias Doctorate Scholarship 2015-2019**
Colciencias, Bogota, Colombia. 2015-2019.

**SPEA - Indiana University PhD Fellowship 2015-2020**
SPEA - Indiana University Graduate, Bloomington, Indiana.

**SPEA Sponsored - ICPSR Fellowship Summer 2016**
SPEA - Indiana University Graduate, Bloomington, Indiana.

**ROC Taiwan - Ministry of Education Masters Scholarship**

**Best ECAES Scholarship (Higher Education Quality Exam)**
Top score at my university and 2nd nationwide among more than 3,000 graduating Economics students.
Universidad de Los Andes, Bogota, Colombia. 2004-2005.

Languages and Skills

Spanish (native), English (proficient user), Chinese (independent user)
Stata, Latex, Python, R & R-Studio, Risk-Simulator, Crystal Ball.