UNIVERSITY OF CALIFORNIA

Los Angeles

Essays in Public and Health Economics

A thesis submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy
in Economics

by

Maria Victoria Barone

2023

ABSTRACT OF THE DISSERTATION

Essays in Public and Health Economics

by

Maria Victoria Barone

Doctor of Philosophy in Economics
University of California, Los Angeles, 2023
Professor Maurizio Mazzocco, Chair

This dissertation studies the design and effectiveness of public policies that aim to improve population health. This dissertation is divided into three chapters. In chapter one, I study the design of an optimal paid sick leave system. To do so, I combine individual-level data on paid sick leave claims from Chile with a sick pay insurance model. First, I show that workers respond to the monetary incentives induced by the benefit scheme, and their behavior varies with the day they fall sick. I use these patterns to inform and estimate a model of sick pay insurance. In the model, risk-averse workers face health shocks and decide how many days to be on leave. Workers are insured by a risk-neutral social planner who chooses the optimal contract to maximize social welfare. I leverage the estimated model to derive the optimal sick pay contract and estimate the welfare gains from its implementation. Relative to the current system, the optimal system provides more insurance for short-term sickness and less insurance, i.e., lower replacement rates, for longer sickness spells. Workers are willing to give up 1.53% of their earnings to be insured under the optimal policy.

The second chapter focuses on the origins and effects of the opioid crisis. Drawing on unsealed records from litigation against Purdue Pharma, I uncover rich quasi-exogenous variation in the marketing of OxyContin, to causally connect supply-side factors to the

origin of this epidemic. My results indicate a strong causal link between Purdue Pharma's promotional targeting and future increases in prescription opioids. The rise in access to potent prescription opioids is responsible for a dramatic increase in opioid mortality, declines in the quality of life, increases in fertility, and deterioration of birth outcomes.

The third chapter examines the effect of non-price interventions on smoking exploiting the 2011 Argentinean anti-smoking national law. I interact state-level legislation with the national law to identify the effect of incorporating graphic tobacco warnings and implementing clean indoor air policies on smoking prevalence and cigarette consumption. I explore whether alcohol and tobacco are consumed as complements or substitutes to assess the side effects of tobacco policies.

The dissertation of Maria Victoria Barone is approved.

Michela Giorcelli

Daniel Haanwinckel Junqueira

Martin Hackmann

Adriana Lleras-Muney

Manisha Shah

Maurizio Mazzocco, Committee Chair

University of California, Los Angeles

2023

To my mother, my deepest source of inspiration.

To my sisters, my biggest support team.

To my father, my favorite economist.

Contents

List of Figures

List of Tables

ACKNOWLEDGMENTS

I am most grateful to Maurizio Mazzocco for his exceptional guidance, helpful advice, continual support, and dedication. He has fueled my intellectual curiosity and taught me how to think about economic questions. I am confident the lessons learned will inspire and shape my future work and I aim to guide my prospective students the way he has shown me. I am incredibly grateful to Daniel Haanwinckel for extensive discussions on the research that constitutes the first chapter of this dissertation. This project would not be the same without his encouragement. I am incredibly grateful to Martin Hackmann for his exceptional insights, advice, and generous support. My discussions with him have profoundly shaped how I think about health economics. I am incredibly grateful to Adriana Lleras-Muney and Michela Giorcelli for their invaluable advice, continuing support, and being fantastic female role models. I aim to inspire my future students in the way they inspire me. I thank Manisha Shah and Laura Wherry for her support during my doctorate. I also thank Bernando Silveira from whom my research benefited hugely. I am grateful to Moshe Buchinsky for his support and encouragement during my first year in the program.

I thank my friend Carolina Arteaga for her continual support, partnership, encouragement, and for being a great co-author to the second chapter of this dissertation. I thank Fernanda Rojas Ampuero and Kirill Ponomarev for their full support during the program. They exemplify hard work and have been my role models since the start. I thank Ariadna Jou, Diana Flores-Pellegrina, Joaquin Serrano, and Mariano Palleja for their extraordinary support especially during my last year in the program. I thank Nicole Gorton for the many reminders that you only need a place to start. I thank Jacob Kohlhepp for extensive talks and invaluable insight on the first chapter of this dissertation. I thank Alvaro Boitier, Julian Batista, Augusto Ospital, and Lorenzo Cativelli for being great study partners and for their exceptional insights and encouragement. Their friendship made my time in graduate school a memorable life experience. I thank Matias Morales for his unconditional support,

encouragement, and the many shared milestones during my time in the program.

This research has benefited from the Lewis L. Clarke Fellowship Fund, the Stan Ornstein Summer Mentored Research Fellowship Program, and funding from the UC Network on Child Health, Poverty, and Public Policy. I benefited from resources provided by the California Center for Population Research at UCLA (CCPR), which receives core support (P2C-HD041022, NICHD), and the Wagner School of Public Service.

VITA

Maria Victoria Barone

Education

University of California, Los Angeles

M.A. in Economics 2017-2018

Universidad Nacional de Cordoba

B.A. in Economics 2008-2012

Research in Progress

"On the Design of Paid Sick Leave: A Structural Approach" (*Job Market Paper*)

"A Manufactured Tragedy: The Origins and Deep Ripples of the Opioid Epidemic" (with Carolina Arteaga)

"Tobacco Consumption Habits in Argentina. Evidence from a New Regulation"

Relevant Positions

University of California Los Angeles (UCLA)

Graduate Student Researcher California Policy Lab 2018-2019

Inter-American Development Bank (IDB)

Research Assistant in the Department of Research 2016-2017

Fellowships, Honors, and Awards

Dissertation Year Fellowship, Graduate Division, UCLA. 2022

Best Paper Award, Albert Family Proseminar in Applied Microeconomics, UCLA. 2022

Best Teaching Assistant Award, UCLA. 2021

Remote Learning Teaching Assistance Award, UCLA. 2021

Summer Mentored Research Fellowship, Department of Economics, UCLA. 2021

Best Paper Award, Albert Family Proseminar in Applied Microeconomics, UCLA. 2020

Best Teaching Assistant Award, UCLA. 2020

Remote Learning Teaching Assistance Award, UCLA. 2020

Graduate Summer Research Mentorship Program, Graduate Division, UCLA. 2019

Fellowship, Department of Economics, UCLA. 2017-2021

Research funds and grants

Lewis L. Clarke Fellowship Fund, UCLA. 2021

Graduate Student Training Grant. UC Network on Child Health, Poverty, and Policy. 2020

Social Sciences and Humanities Research Council Insight Grant. 2020-2021

University of Toronto Mississauga Research and Scholarly Activity Fund. 2020-2021

Presentations

Notre Dame, Wisconsin School of Business, Pompeu Fabra, Universidad de Chile, 2023
PUC Chile, FGV-EPGE, U de los Andes-Chile, Insper, ITAM (Econ), Universidad Andres Bello, ZEW, ASHEcon.

Wagner School of Public Service, WEAI Graduate Student Workshop, 2022
ASHEcon, Young Economist Symposium (YES), Washington University in St. Louis.

Applied Micro Student Lunch at NYU Econ; Washington University in St. Louis 2021
Tobacco Online Policy Seminar (TOPS).

Graduate Student Retreat, Center for Poverty Research, UC Davis; ASHEcon Annual Conference (poster). 2020

Introduction

The first chapter of this dissertation studies the optimal design of social insurance programs. These programs offer valuable protection against a broad range of risks that could be detrimental to individuals' well-being, such as health deterioration that limits one's ability to work. In particular, paid sick leave provides income replacement for workers who suffer from short-term impairments caused by non-work-related sickness (e.g., common flu). If adequately designed, paid sick leave can be greatly beneficial. It allows workers to meet their personal health needs and smooth consumption. In *"On the Design of Paid Sick Leave: A Structural Approach"*, I ask: What is the optimal paid sick leave system? To answer this question, I combine individual-level data on paid sick leave claims with a model of sick pay insurance provision. Then, I leverage the estimated model to derive the optimal sick pay contract and estimate the welfare gains from its implementation.

This research exploits a unique dataset that includes administrative records of the Chilean paid sick leave system. I observe the universe of workers insured by the government eligible to file a sick leave claim between 2015 and 2019 and their utilization of sick leave benefits. This database includes rich demographic information at the worker and sick leave claim levels. In particular, I observe the exact beginning and end dates and the primary diagnosis related to a sick leave claim at the International Classification of Diseases' four-digit level. Importantly, this paper is the first to use such rich administrative data on sick leave claims. This dataset offers two advantages. First, the data are less prone to measurement error present in the

survey data used by other papers. The use of survey data raises the usual measurement error issues with self-reported recall data and prevents researchers from distinguishing the incidence of absences from their length. Second, the data allow me to study daily leave-taking behavior and estimate individual demand for sick pay.

I model the incentive problem created by paid sick leave provision. In the model, a risk-averse expected utility–maximizing worker chooses her sick leave utilization and labor supply. This choice is a function of the generosity of sick pay and the worker's preferences for the valuation of time outside work and propensity to engage in moral hazard behavior. A risk-neutral social planner offers a sick pay contract to maximize aggregate welfare. The optimal level of benefits depends on workers' behavioral responses, the costs of sick work, workers' risk preferences, and production losses. I quantify these elements by exploiting the data on paid sick leave utilization linked to medical assessments of recovery times. This model incorporates two unique features. First, in contrast to the previous literature, the proposed framework focuses on how sickness affects the absence behavior of employed individuals. Second, the model allows workers' behavior to vary with the day of the week that a sick leave claim is filed. Previous literature has shown that the temptation for shirking varies with the day of the week. I explicitly incorporate this mechanism in the model of workers' behavior.

The main empirical challenge is to recover the underlying distribution of health and workers' preferences. I propose a novel strategy to overcome this challenge. First, I combine individual-level data on sick pay utilization with detailed medical assessments of recovery times associated with each health condition. This allows me to construct the underlying distribution of health without imposing parametric assumptions on this distribution. Second, I exploit variation in the incentives for shirking behavior induced by the day of the week when workers fall sick and file sick leave claims to estimate workers' preference parameters.

The main result of this research is to determine the optimal paid sick leave system—the

system that maximizes aggregate welfare—and the corresponding replacement rates. The optimal policy has three main features. First, it offers a partial replacement for claims of up to three days. This shift increases the utility of workers who would not take sick leave under the current system but do under the optimal policy. Second, the optimal policy eliminates the discontinuity feature in the current system. Doing so curbs the cost of the behavioral responses to the program incentives. Third, the average replacement rate is increasing. I estimate that workers are willing to give up 1.53% of their earnings to be insured under the optimal policy.

Risky health behaviors such as smoking, drinking alcohol, and drugs' use and abuse are major sources of preventable deaths. The second part of this dissertation contributes to understanding the costs and consequences of these behaviors. In addition, it provides evidence of the effectiveness of government interventions to curb the consumption of addictive goods.

Opioid misuse is one of the most substantial and long-lasting public health crises faced by the United States. In the second chapter of this dissertation, *"A Manufactured Tragedy: The Origins and Deep Ripples of the Opioid Epidemic"* (joint with Carolina Arteaga), we shed new light on the origins and effects of the opioid crisis, emphasizing the critical role played by the pharmaceutical industry. Understanding the origins of this epidemic and its consequences is challenging since the variation in the use of prescription opioids across geographies and over time is not random. To overcome this challenge, we exploit detailed features of the initial marketing of prescription opioids, which we obtained from unsealed court records drawn from litigation against Purdue Pharma, the manufacturer of OxyContin—a prescription opioid at the center of the epidemic. Drawing on these records, we document a strong causal link between Purdue Pharma's promotional targeting and future increases in the level of prescription opioids. We then exploit this variation to quantify the epidemic's effects.

We find three main results. First, we estimate that easier access to potent prescription

opioids was responsible for a dramatic increase in mortality. Increasing the supply of opioids from the 25th to the 75th percentile caused an 89% increase in prescription opioid deaths, corresponding to approximately 200,000 deaths. Second, we document far-reaching declines in quality of life: for example, the number of individuals participating in social safety net programs, such as SNAP, increased with opioid supply. Finally, while the epidemic primarily affected individuals in early adulthood through mid-life, it has potential costs beyond the generation directly affected by it. We study the intergenerational impacts of the opioid epidemic and estimate a 10% increase in fertility rates driven entirely by increases in nonmarital births and concentrated among women aged 25–29.

Governments can preserve population health by implementing policies that directly restrict the availability of unhealthy products or raise the time costs of using them. In particular, tobacco regulation has been a major component of health policy in both developed and developing countries. The third chapter of this dissertation, *"Tobacco Consumption Habits in Argentina: Causal Evidence from a New Regulation"* (reject & resubmit at *Journal of Health Economics*) assesses the effectiveness of antismoking nonprice policies—such as place-based bans and warning labels—in low- and middle-income countries, where 80% of tobacco smokers worldwide live. Unlike other studies on the effects of nonprice policies in the context of middle- and low-income countries that have relied on cross-country or before-and-after comparisons, my paper employs an identification strategy that provides causal estimates of the impacts of this policy on smoking, drinking, and health outcomes.

I study the case of Argentina's law that banned smoking in public spaces—e.g., bars and restaurants—and mandated graphic tobacco warnings in 2011. To estimate the causal effect of this regulation, I exploit state-level differences in the strength of the regulation of tobacco products before the national law was passed. States with more lenient restrictions were more exposed to the effects of the new regulation than states with stricter regulations. I compare smoking and drinking behavior and health outcomes of individuals in lenient states

(treated states) with those of individuals in strict states (comparison states). I obtain two key results. First, I document that the antitobacco regulation effectively reduced smoking participation. Second, I document that the rate of hospital discharges due to diagnoses of chronic obstructive pulmonary disease decreased in the short run, suggesting an improvement in population health as a result of stricter regulation of tobacco products.

Chapter 1

On the Design of Paid Sick Leave: A Structural Approach

1. Introduction

Social insurance programs offer valuable protection against a broad range of risks that could be detrimental to individuals' well-being, such as health deterioration that limits one's ability to work. In particular, paid sick leave provides income replacement for workers who suffer from short-term impairments caused by non-work-related sickness (e.g., common flu).[1] If adequately designed, paid sick leave can be greatly beneficial. It allows workers to meet their personal health needs and smooth consumption. Nonetheless, the availability of sick pay could induce workers to request more sick days than should be assigned based on their health. This response could partially offset the welfare gains of sick pay. This raises the question: What is the optimal paid sick leave system?

[1] While closely related to worker's compensation programs, which provide income replacement and medical benefits in case of work-related sickness, and disability insurance programs, which provide income replacement in case of permanent or long-term impairments to working ability, paid sick leave programs offer protection against the risk of contracting a disease that impairs workers for a short period and has a foreseeable recovery.

The main contribution of this paper is to answer this question. To do so, I proceed in three steps. First, I use detailed data on paid sick leave claims to document how workers' characteristics and the institutional details of the sick pay system—e.g., the presence of deductibles—determine sick pay utilization. Second, I propose and estimate a model of sick pay provision. This exercise gives me key inputs to derive the optimal paid sick leave system: the value of risk protection, the costs of insurance provision in terms of moral hazard, the production cost of time off, and the underlying distribution of health shocks.[2] Finally, I use these estimates to determine the replacement rates that characterize the optimal sick paid system. I find that the optimal system features a low replacement rate for short claims, i.e., up to three days long claims are partially insured, with most of the cost on the worker side. Longer sickness spells are covered at a higher rate and the replacement rate is increasing with sick leave duration. I estimate that workers are willing to give up 1.53% of their earnings to be insured under the optimal policy.

I study this question in Chile, which is an ideal setting for this research for several reasons. First, it has a comprehensive paid sick leave system that covers all workers and features only one plan designed by the central government. Thus, workers do not choose their sick leave coverage.[3] The Chilean system provides no coverage for sickness lasting three days or fewer. This nonpayable period works like a deductible that resets with every new sick leave spell and is a common mechanism among sick leave programs.[4] Starting on the fourth day, there is full coverage of each missed day, i.e., the replacement rate is one. If the sick leave spans 11 days or more, the nonpayable period is reimbursed; this implies that the average replacement

[2]In this paper, moral hazard refers to the responsiveness of workers' demand for sick leave to changes in the generosity of sick leave benefits.

[3]This alleviates adverse selection concerns. If workers could choose their sick pay coverage, we could expect individuals with preferences for more absences to self-select into plans with more generous provisions. The presence of adverse selection would result in an upward bias in the estimates of the moral hazard responses.

[4]These resettable deductibles are similar to those used in automobile or homeowners insurance: Separate deductibles apply to each loss. Many European paid sick leave systems have a similar deductible. See Marie and Castello (2022) on the case of Spain and Pollak (2017) on the French experience. Table A.20 summarizes sick pay systems for a sample of OECD countries with available data.

rate varies with the duration of a claim and jumps discretely at 11 days.[5]

The second advantage of this setting is that Chile has greatly detailed administrative data. I observe the universe of workers insured by the government eligible to file a sick leave claim between 2015 and 2019 and their utilization of sick leave benefits. This group accounts for about 70% of the Chilean workforce. This database includes rich demographic information at the worker and sick-leave claim levels. In particular, I observe the exact beginning and end dates and the primary diagnosis related to a sick leave claim at the International Classification of Diseases 10th revision (ICD-10) four-digit level. I combine the claims data with medical assessments from the Peruvian Handbook of Recovery Times (EsSalud, 2014). This handbook specifies the average recovery times for 2,763 unique disease codes at the ICD-10 four-digit level. These recommendations are adjusted based on workers' gender, age, and occupation.

Exploiting these data, I document three facts that provide qualitative motivation for the model and serve as quantitative targets in the estimation. First, workers' sick leave claim utilization varies with age and occupation. For example, on average, workers aged between 55 and 64 use an extra 2.72 days per year relative to their younger counterparts. Similarly, compared to white-collar workers, workers in blue-collar occupations use, on average, 1.15 more days. These patterns could reflect differences in the underlying distribution of health shocks and differences on their preferences. In the estimation of the model, I allow the distribution of health shocks to vary with age and occupation.

Second, I provide evidence that workers respond to the financial incentives induced by the benefit scheme. Specifically, I test whether workers bunch around the discontinuity in the replacement rate. To do so, I construct an underlying distribution of recovery times, leveraging the handbook's recommendation of how many rest days a worker needs to recover from a disease. I compare this distribution with the observed distribution of requested days

[5]Panel (a) of Figure 1.1 presents days paid as a function of days on leave for claims of different duration. Panel (b) shows the average replacement rate.

and estimate that 11-day-long sick leave claims are 4.55 percentage points more likely than what the underlying distribution of health predicts. I rely on this empirical fact to assess the model performance and find that the proposed model can reproduce the excess mass observed at 11 days.

Third, I show that workers respond to nonmonetary shifts in the temptation to extend their leaves. To do so, I exploit the data on exact dates when sick leave claims are filed. I argue that the temptation to extend a sick leave claim varies with the day of the week when a worker falls sick. For example, the incentives to file a two-day-long sick leave claim on a Thursday differ from those to file a two-day claim on a Tuesday. I consider the following exercise: I fix the duration of a sick leave claim and inspect the share of claims filed on each day of the week. I find an excess mass on combinations of days of the week and durations that allow the worker to extend her leave through the weekend. I refer to such combinations of durations and start days as "weekend-streak combinations". I document that workers are, on average, 12.33% more likely to file a weekend-streak claim than to file a sick leave claim of the same duration on any other day of the week. To capture this empirical regularity, the model allows workers' behavior to vary with the day of the week of a sick leave claim.

I use this evidence to develop and estimate a model of sick pay provision. The model has two agents: the workers and a social planner. Workers are risk-averse expected utility maximizers and choose their sick leave utilization. When choosing the demand for sick days, the worker trades off the utility of time off with the consumption loss from taking up sick leave. The latter depends on sick leave generosity. The former is a function of two key parameters: (i) workers' relative valuation of time outside work to recover from health shocks or engage in leisure and (ii) workers' propensity to overstate their sickness. Utility varies with the day of the week when a sick leave claim is filed to capture the empirical fact that the temptation to extend leaves varies at this level. In this setting, sick pay coverage lowers the cost of a day away from work and (weakly) increases sick pay utilization. This is

the *traditional* moral hazard effect of insurance provision.

Sickness negatively impact workers' productivity. This reduction is justified by two forces. First, a *pure* productivity effect: a worker is less productive when she is sick relative to her healthy state. Second, individuals may be contagious at work, creating a negative externality. In the model, I capture this by allowing workers' productivity to vary with their health status. If the wage rate does not depend on the health-related productivity, this reduced productivity gives rise to a production externality.[6] Some workers find optimal to show up to work sick to avoid the consumption losses associated with an absence. Additionally, some workers find optimal to skip work even if their productivity is "high enough", e.g. some workers return to work too late. This production externality affects the main trade-off faced by the planner.

The optimal contract balances the benefits of risk protection with the cost associated with moral hazard and production losses. A more generous sick pay scheme would increase workers' well-being by offering more risk protection. Workers would respond to this policy change increasing sick pay utilization—moral hazard response. This response increase the pool of workers taking sick leave—more absences—and reduces production externalities—less individuals working sick. Thus, moral hazard is not necessarily welfare decreasing. The optimal level of benefits depends on risk preferences, workers' behavioral responses, production losses, and the distribution of risks. The principal empirical focus of this paper is to quantify these elements.

To determine the optimal system, I proceed in two steps. The first step concerns workers' sick pay utilization choices. In this step, I recover a vector of preference parameters—time valuation and compliance costs—from workers' observed leave-claiming behavior. The second

[6] I assume that wages are not conditioned on workers' health-related productivity. This assumption is consistent with an information asymmetry between the firm and the worker—the worker knows her health state but the firm cannot observe it—and with contracting frictions—even if the firm could observe workers' health, it might not be able to offer a contract that induces worker to report their health state. The focus of this paper is on the consequences of this production externalities and not to model the reason why wages are independent of health-related productivity.

step I find the replacement rates that maximize total welfare. The validity of this approach relies on the fact that the worker's problem can be viewed as a two-stage problem. Once the health shock is realized, workers optimally choose their sick pay utilization. Risk preferences do not affect the utilization decision since the uncertainty has been resolved when this decision is made. Thus, the focus of the first step. Nonetheless, workers' expected utility depends on their risk preferences: more risk-averse workers would prefer a contract with more coverage. This is accounted for in the second step. The main strength of this approach is that I only rely on workers' observed decisions to estimate the model and do not need to assume that the current sick pay plan is optimal.

I estimate the model of workers' behavior by the simulated method of moments (SMM). The main empirical challenge is to disentangle the underlying distribution of health from the distribution of workers' preferences. To overcome this challenge, I build the underlying distribution of health exploiting the Peruvian Handbook of Recovery Times recommendations and the observed diagnoses. This approach has two advantages: (i) it provides an objective measure of recovery times constructed outside the structure of the Chilean system, and (ii) it does not impose parametric assumptions on this distribution. The empirical distribution of health states incorporates observed heterogeneity across workers: I allow for variation in age and occupation. That is, the same diagnosis has age- and occupation-specific associated recovery times. On average, older workers and workers employed in blue-collar occupations are assigned longer recovery times. In estimating the model, I also allow an arbitrary correlation between health states and workers' income to capture that wealthier workers tend to have better health and could require shorter absences.

To recover the distribution of workers' preference parameters, I employ the day of the week when a sick leave claim is filed as a quasi-exogenous shifter of the temptation to extend sick leave claims as the main source of variation. First, the excess of weekend-streak sick leave claims informs how workers' utility from a sick leave claim of the same duration varies

with the day of the week on which the claim is filed. Second, I consider workers with similar characteristics and the same *assigned* recovery time—i.e., I hold workers' health, age, and occupation fixed—and compare their demand for sick pay across days of the week. I start by computing the share of claims filed for a duration that matches the assigned recovery time and compare this figure with the share of claims filed for an extra day. This difference is informative on how costly it is for individuals to ask for an extra day of leave. I restrict this comparison to claims filed for a combination of durations and days of the week representing a weekend streak. For example, I compare the share of two-day-long claims filed on a Thursday with that of three-day-long claims filed on a Wednesday for workers assigned two days of recovery. The larger the difference is, the more costly it is for workers to ask for an extra day of leave. Comparing these shares keeps the incentives for extending sick leave claims fixed since every combination implies that workers would be on leave through the weekend. These comparisons inform the distribution of compliance costs. Lastly, to learn the distribution of workers' time valuation, I construct the ratio of leisure to (a measure of) consumption from the claims data.

The estimation of the model incorporates observed heterogeneity across workers in the valuation of time outside work and the propensity for overstating sickness. Heterogeneity across the valuation of time outside work reflects variation in opportunity costs from missing work to recover from a disease—e.g., due to workers' role in the firm—or variation in tastes for leisure relative to consumption. Heterogeneity in the propensity for moral hazard behavior reflects variation in workers' preferences over behaving as expected or revealing their "true" health status. Additionally, the model of workers' behavior allows for heterogeneity in how workers perceive their sickness. That is, workers who suffer the same health shock can be affected by it differently. While the parameter that governs this perception is not separately identified, the derivation of the optimal policy does not require its identification. This derivation rather relies on workers' responses to the incentives generated by the provision of

sick pay.

The estimated model provides a good fit for the targeted and nontargeted moments. I exploit the discontinuity at 11 days—a nontargeted moment—to assess the model performance. The model predicts that if a worker realizes a health state just under 11 days, she will take advantage of the proximity to the full-coverage region and fake her type to gain full coverage. The proposed model can reproduce the excess mass at 11 days quite well. I estimate that, in the data, the 11-day duration accumulates an additional 4.50% mass than its neighbors. Using the model-simulated sample, I estimate an additional 4.03%.

I use the estimated model to derive the optimal sick pay policy, i.e., to determine the replacement rates that maximize aggregate welfare. The optimal policy differs from the current system in three key ways. First, it offers partial replacement, with an average replacement rate of 0.36, for claims of up to three days. This shift increases the utility of sick workers who would not take sick leave under the current system but do under the optimal policy. At the same time, partial coverage constrains moral hazard since most of the cost of those absences is faced by workers.

Second, the optimal policy eliminates the discontinuity at 11 days and exhibits a higher average replacement rate between 4 and 10 days. Doing so curbs the cost of the behavioral responses to the program incentives and provides more risk protection. Implementing the optimal scheme would shift the distribution of sick leave duration: workers would be more likely to file sick leave claims of between 8 and 10 days and less likely to file claims for 11 days relative to the corresponding probabilities under the current Chilean system.

Third, the optimal policy does not offer full replacement for sick leave claims longer than 11 days. The average replacement rate is increasing, as in the current system, but it is less generous for longer claims. Taken together, these changes in the replacement rate reflect that the workers value a contract that offers more protection for shorter claims to smooth consumption across different health states. I estimate that workers are willing to give up

1.53% of their earnings to be insured under the optimal policy.

This paper contributes to several areas of the economics literature. First, it contributes to a large body of literature on public insurance programs. This literature has modeled the trade-offs between protection against risk and moral hazard present in unemployment risks (Hopenhayn and Nicolini, 1997; Chetty, 2008; Hendren, 2017), disability and retirement risks (Gruber, 2000; Low and Pistaferri, 2015), healthcare risks (Cutler and Zeckhauser, 2000; Einav, Finkelstein and Cullen, 2010; Handel, Hendel and Whinston, 2015; Ho and Lee, 2020; Marone and Sabety, 2022) , and work-related injuries (Powell and Seabury, 2018; Cabral and Dillender, 2020). This paper contributes to this literature by being the first study to propose a theoretical framework for designing the provision of paid sick leave and quantifying the welfare gains from its implementation.

The closest programs to paid sick leave are disability insurance and workers' compensation. These three programs condition benefits on a difficult-to-verify state: the true impairment of working ability due to health deterioration. Thus, the trade-offs considered in the design of disability insurance and workers' compensation are relevant to the design of paid sick leave. Nonetheless, disability insurance and workers' compensation programs target specific groups of workers—elderly people and workers especially vulnerable to accidents—and provide protection for different set of health shocks—more permanent and more severe shocks. The nature of the health shock insured by paid sick leave provision—short spells of non–work-related illness—implies that virtually every worker could benefit from the risk protection under the program and demand paid sick leave.

This paper is close to Maclean, Pichler and Ziebarth (2020), who evaluate the labor market effects of sick pay mandates in the United States and extend the Baily–Chetty framework of optimal social insurance to assess the welfare consequences of mandating sick pay. Their framework allows researchers to study the effects of policies that vary the share of employees eligible for the benefit. This paper differs in two critical dimensions. To begin

with, I propose a structural approach to conducting welfare analysis. This approach does not rely on the assumption that policy changes are marginal.[7] Relaxing this assumption is important since it allows the optimal policy to differ freely from the actually implemented policy, i.e., it allows for non-marginal policy changes.

Second, this paper relates to the empirical literature on sick pay insurance. Exploiting arguably exogenous variations, the literature has documented a positive response of sick pay utilization to increases in benefit levels (Johansson and Palme, 2005; Ziebarth, 2013; De Paola, Scoppa and Pupo, 2014; Ziebarth and Karlsson, 2014; Pollak, 2017;Böckerman, Kanninen and Suoniemi, 2018; Cronin, Harris and Ziebarth, 2022; Marie and Castello, 2022). This paper goes beyond workers' responses to policy changes and proposes a framework and an empirical strategy to quantify the welfare effects of these policy changes.

Additionally, this paper is the first to use administrative data on sick leave claims at the individual level.[8] These data allow me to study daily leave-taking behavior and estimate an individual demand for sick pay. In addition, these data are less prone to measurement error. Many papers have used survey questions that ask respondents how many days of work they have missed due to illness in a reference period. The use of survey data raises the usual measurement error issues with self-reported recall data and prevents researchers from distinguishing the incidence of absences from their length. Observing the length of absences is a crucial input for quantifying moral hazard responses, as workers could extend their absences to obtain more sick pay.

This paper also contributes to the literature on variation in leave claiming behavior across days of the week. Card and McCall (1996) and Campolieti and Hyatt (2006) provide

[7]For the validity of the sufficient statistics approach, the analyzed policy changes should be infinitesimal or at least close enough to infinitesimal for first-order approximations to be precise Kleven, 2021.

[8]Cronin, Harris and Ziebarth (2022) constructs a similar dataset for the Scott County School District (SCSD) in Kentucky, which allows a detailed study of teachers' use of paid sick leave. While the data structure is similar to the one used in this paper, I observe sick leave utilization regardless of workers' occupations. Marie and Castello (2022) also exploit administrative data for Spain, though their data are at the spell rather than the individual level. The data in this paper capture the sick leave choices of 70% of the universe of Chilean workers.

evidence of a "Monday effect"—which refers to a spike in back injury and sprain claims on Mondays—among workers' compensation claimants. Thoursie (2004) shows that Swedish men were likelier to call in sick the day after popular skiing competitions were broadcast at night during the Winter Olympics in Calgary. Implementing a similar test, Cronin, Harris and Ziebarth (2022) document that teachers in Kentucky are not more likely to use sick leave while Keeneland is in session, on Mondays following Super Bowls, or on days when the University of Kentucky men's basketball team plays in the NCAA tournament. I provide new evidence regarding workers' behavior across weekdays by exploiting the exact dates of sick leave spans.

This paper proceeds as follows. Section 2. describes the empirical setting that I study and the data. Section 3. presents the theoretical framework and discusses the optimal design of a paid sick leave system. Section 4. presents the empirical implementation of the model. Section 5. presents the model estimates and main results. Section 6. discusses the optimal policy. Section 7. concludes.

2. Background and Data

In this section, I discuss the Chilean healthcare and sickness insurance systems, focusing on the institutional features relevant to my analysis. I then present the data and patterns in the data that motivate my modeling choices.

2.1 The Chilean Health Insurance System

In Chile, healthcare insurance providers serve two functions: (i) to offer healthcare insurance contracts and (ii) to administer the paid sick leave system. The healthcare insurance system is composed of a government-run healthcare insurance provider and a handful of private

insurers.[9] Workers are mandated to purchase health insurance, allocating at least 7% of their salary to a healthcare plan offered by an insurer of their choice.

The government-run healthcare insurance provider offers four plans, whose eligibility is based on monthly salary and household composition. The lowest-tier plan provides coverage for individuals with no income at no cost in public system hospitals. As income increases, workers qualify for a higher-tier plan. This plan provides healthcare coverage in public system hospitals with low copays and access to private healthcare institutions with high copays.[10] Private insurance companies, on the other hand, provide tiered plans with financially vertically differentiated coverage levels—similar to the Gold, Silver, and Bronze plans offered by Affordable Care Act exchanges in the US. The plans offered by private insurance companies allow beneficiaries to obtain healthcare from private healthcare institutions, which provide a higher quality of care than public institutions.

The mandatory contribution would allow workers to enrolled in a plan offered by the government-run healthcare insurance with no need to make voluntary contributions.[11] To select one of the private providers, workers might need to contribute a higher proportion of their salary to qualify for the healthcare plan of their choice.[12]

In 2017, 73% of workers enrolled in plans offered by the government-run healthcare insurance system; the remaining 27% enrolled in plans offered by one of the private providers (see Panel A of Table A.19). Workers enrolled in the government-run plans have observable characteristics that would predict that they are more costly to insure: they are older, more

[9]These are called FONASA and ISAPRES, respectively, for their names in Spanish (*Fondo Nacional de Salud* and *Instituciones de Salud Previsional*).

[10]Plans are indexed by letters, where A is the lowest-tier plan and D the highest-tier plan. The highest-tier plan has a 20% copay in public system hospitals and vouchers to use healthcare providers who participate in the plan's network at a discounted price.

[11]Workers would be enrolled in one of the four plans based on their monthly salary and household composition. For example, a single worker who earns USD$693 a month—the median salary in 2017—and chooses the government-run insurance system will be enrolled in the highest-tier plan and cannot choose any of the lower-tier plans with lower copays.

[12]Plans offered by private insurers are highly regulated. These insurers can set prices based on observable characteristics—including age and (until April 2020) sex—and risk factors.

likely to be women, and have lower salaries.

The second function of healthcare insurance providers is to *administer* the paid sick leave system. Insurers are in charge of receiving and screening sick leave claims and disbursing sick leave benefits. Insurers cannot design sick pay plans and must follow the rules set by the central government regarding eligibility criteria and benefits. Nonetheless, there are differences in how each provider applies these rules in practice. For example, panel B of Table A.19 shows that the rejection rate by private insurers is almost three times that of the government-run insurer. I argue that these differences in leniency are suggestive evidence that private insurers might have different motives—such as minimizing sick leave payments—when screening sick leave claims. My empirical analysis focuses sick pay utilization of workers enrolled in the government-run health insurance system—they represent about 73% of all Chilean workers. The main reason for this choice is that this paper focuses on the provision of paid sick leave as a social insurance system, which is closer to the behavior of the government-run healthcare provider.

2.2 The Chilean Paid Sick Leave System

The Chilean paid sick leave system gives employees the right to call in sick and receive sick pay due to short-term, non–work-related sickness—e.g., the common flu or back pain.[13] Workers can use sick leave to meet their own health needs but not to care for family members. The eligibility criteria for claiming paid sick leave requires that workers (i) have been enrolled in the social security system for six months and (ii) have made contributions to the health insurance system for three months. The paid sick leave system is financed through these mandatory contributions. That is, the 7% of workers salary that is contributed towards the healthcare system. Between 2015 and 2019, the paid sick leave portion of the system was

[13]The sickness insurance system aims to provide risk protection from impairments to working ability that are temporary and from which full recovery is foreseeable. A separate program provides disability insurance to workers in case of permanent impairments to working ability.

financed with 2.6% of the contributions (see table A.19).

Upon falling sick, workers must obtain a physician's certification of their sickness stating the primary diagnosis and the number of days that the physician considers the worker will need to recover from the disease. This certificate is necessary to justify the absence from work and must be requested regardless of the duration of the sick leave claim. This certificate is reviewed by an insurance office, which decides whether the sick leave claim is (i) approved with no changes, (ii) approved with a reduction in its length, or (iii) denied.

Sick leave payments are a function of sick leave duration subject to a maximum salary.[14] For workers with a salary below the maximum, benefits are computed as follows: The benefit scheme exhibits a nonpayable period of 3 days; i.e., the replacement rate for the first three days of a sick leave spell is zero.[15] This nonpayable period works like a deductible that resets for every new sick leave span.[16] Starting on the fourth day, there is full coverage of each additional missed day—i.e., the replacement rate is one. If the sick leave lasts 11 days or more, the nonpayable period is reimbursed. That is, claims with an 11-day or longer duration are fully covered.[17] Panel (a) of Figure 1.1 presents days paid as a function of days on leave for claims of different duration. Reimbursement of the nonpayable period after 11 days implies that the average replacement rate jumps discretely at 11 days, and it is nonconstant (see Panel (b) of Figure 1.1).

Chile constitutes an ideal setting to study the design of optimal sick pay systems for several reasons. First, its paid sick leave system is comprehensive, covers all workers, and features only one plan designed by the central government. This implies that workers do

[14] In my sample, less than 1% of workers earn above this threshold; see Figure A.27. I exclude these workers from the analysis.

[15] Cid (2006) documents that the origin of the 3-day nonpayable period dates to a regulation implemented in 1952 that aimed to prevent abusive behavior, which has not been revised since.

[16] These resettable deductibles are similar to those used in automobile or homeowners insurance: Separate deductibles apply to each loss.

[17] If a worker files two (or more) consecutive claims, they are treated as one claim for the computation of benefits. To be consistent, I treat these claims as one in the analysis. Appendix Table A.22 presents counts and summary statistics of sick leave claims and sick leave spells.

not choose their sick leave plan and alleviates adverse selection concerns. If workers could choose their sick pay coverage, we could expect sicker individuals to choose plans offering more generous insurance coverage. While this mechanism could be at play in the choice of healthcare insurance provider, conditional on this decision, sicker and healthier individuals face the same sick pay coverage.

The design of the Chilean system is similar to many European paid sick leave systems that use resettable deductibles. Table A.20 compares the design on 22 countries. Twelve of the systems are "bracket systems". These are characterized by (i) a first bracket with a low or zero replacement rate and (ii) two or three brackets with a higher replacement rates. Specifically, in 9 of these 12 countries the replacement rate for the first days is zero.

The second advantage of this setting is that Chile has greatly detailed administrative data which I describe in the next section.

2.3 Data

I exploit unique administrative data on sick leave claims matched to enrollment data for workers insured by the government-run healthcare system. These restricted-access data were provided directly by the government-run healthcare insurance office and cover the period 2015–2019.

The enrollment dataset covers the universe of individuals enrolled in government-run healthcare insurance regardless of whether they have filed a sick leave claim. I observe individuals' demographic and economic characteristics: sex, age, annual earnings, and health indicators for chronic conditions.[18] Additionally, I observe individuals' residence postal codes and the health insurance plan assigned to them. The latter allows me to exclude individuals enrolled in the lowest-tier plan from the analysis, as these individuals are not active in the

[18]These conditions are cerebral vascular accident, Alzheimer's, juvenile arthritis, rheumatoid arthritis, bronchial asthma, lung cancer, diabetes, chronic obstructive pulmonary disease, chronic kidney disease, arterial hypertension, acute myocardial infarction, leukemia, lymphoma, multiple myeloma, and HIV.

labor market.

In the claim dataset, I observe detailed information about each sick leave claim: start and end dates, prescribed days on leave, primary diagnosis (coded following the ICD-10), physician identifiers, and amount received for paid sick leave. I also observe the occupation in which the worker is employed at the moment of filling the sick leave claim.

I combine these data with medical assessments from the Peruvian Handbook of Recovery Times. I rely on these assessments to construct the underlying distribution of health; this is a key input in characterizing workers responses and in the estimation of the model. The Peruvian Handbook of Recovery Times (EsSalud, 2014) specifies the average recovery times for 2,763 unique disease codes at the ICD-10 four-digit level. Crucially, these recommendations are adjusted based on workers' sex, age, and occupation. Table A.21 provides an example of the average recovery times for three common diagnoses—lumbago with sciatica, common cold, and infectious gastroenteritis—and the correction factors proposed by the handbook. The main advantage of exploiting this external source of data is that it provides an objective measure of recovery times constructed outside the structure of the Chilean system. That is, it is not affected by the brackets used in the paid sick leave benefit function.

Based on these three sources, I construct a claim-level dataset with detailed information on workers' demographic characteristics and leave-taking behavior and the average recovery time. My primary measure of leave-taking behavior is the duration of a sick leave claim filed on a given day of the week. I assign a benchmark recovery time to each sick leave claim exploiting the disease code and the suggestions of the Peruvian Handbook of Recovery Times. I construct these measures at the sick leave spell level; i.e., I consider consecutive claims as one claim. Thus, the unit of analysis is the same as the one used to compute sick leave benefits. Appendix Table A.22 presents counts and summary statistics of sick leave claims and spells.

To construct the analysis sample, I impose two restrictions. First, the estimation sample

includes claims from private-sector male workers aged 25 to 64. This is a demographic group with high labor market participation rates. Although women's sick leave–taking behavior is of high interest for the design of sick leave programs, women have much lower participation rates than men. For example, in Chile, women's labor force participation rates are more than 20 percentage points lower—52.6% for women at the beginning of the sample period and 73.2% for men. Thus, a model of sick leave–taking behavior that explains women's choices would also require incorporating their decision to participate in the labor market. Abstracting from the participation decision simplifies the model estimation.

Second, the estimation sample includes claims associated with a subset of diseases. I exclude mental-health diagnoses because their filing process is more cumbersome than the one for non–mental health claims.[19] That is, in this paper, I focus on non–mental health sick leave claims. Among non–mental health diagnoses, I exclude diagnoses for which it is hard to assign a recovery time. For example, I exclude claims with codes corresponding to neoplasms. Table A.23 lists the conditions included in the analysis and the share of claims recorded under each diagnosis. 2.3 provides additional details. The final sample includes 90.19% of all non–mental health sick leave claims.

2.4 Descriptive Evidence

Summary statistics. Table 1.1 presents summary statistics for all the workers in the sample and for those who used sick pay during 2017. I split the last group based on the type of disease and duration of the sick leave claim. Almost 20% of Chilean workers filed a non–mental health-related sick leave claim in 2017. The average worker in the sample is 44 years old, and the average worker who used sick pay is approximately the same age (column 1 vs. column 2). Nonetheless, the average claimant has a higher salary than the average worker, and this difference is statistically and economically significant.

[19]For example, these claims must be certified by a psychiatrist and require a comprehensive medical assessment at the time of filing.

To better clarify the differences between workers who filed sick leave claims and those who did not, Table 1.2 presents characteristics of workers who used sick leave benefits based on the duration of their claims. I group workers who filed (i) at least one claim with a duration of up to 3 days, (ii) at least one claim with a duration of between 4 and 10 days, and (iii) at least one claim with a duration of 11 days or longer. Shorter sick leave claims are associated with younger workers with higher average wages, who are also less likely to have chronic conditions. This pattern is compatible with the 3-day waiting period, reducing the likelihood that lower-earning workers file a sick leave claim. Additionally, the association between workers' age and prevalence of chronic conditions is consistent with older workers experiencing more severe conditions than their younger counterparts.

Sick leave duration. Figure 1.2 shows the distribution of sick leave claims of up to 29 days. Two main patterns characterize the distribution of days on leave. First, approximately 26.54% of sick leave claims have a duration of up to 3 days. This provides evidence that workers are completing the filling process to justify their absences even when not paid for them. Sick leave claims lasting between 4 and 10 days account for 41.06% of claims. Thus, 32.40% of claims have a duration between 11 and 29 days.[20]

Second, some durations accumulate more mass than others. For example, three days on leave is the most common duration representing 15.54% of claims, follow by five and seven days (accounting for 13.56% and 13.64%, respectively). This pattern could be explained by the underlying distribution of recovery times or behavioral responses to the incentives provided by the sick leave benefit scheme; disentangling these is one of the papers' aims.

Figure 1.3 compares the underlying distribution of recovery times with the observed distribution of days on leave. It shows that three rest days are the most recommended

[20]Claims filed for 30 days are used either by workers with illness requiring a longer recovery or those transitioning to disability insurance. I do not have access to data that would allow me to differentiate between these outcomes. Thus, my analysis focus on claims of up to 29 days. This restriction approximates the universe of workers that suffer conditions with foreseeable recovery.

recovery time, consistent with 3-days-long sick leave claims being the most frequent duration. In contrast, there is a broader gap when comparing the share of claims with five and seven days as suggested recovery time and the observed claims of such duration. This pattern is consistent with physicians being more likely to write recovery times that correspond to a workweek—five days—or a calendar week and multiples of these. Panel (b) of Figure 1.3 shows the ratio of the difference between the share of sick leave claims for a given duration and the share implied by the underlying distribution to the latter. This figure illustrates that the greater gaps are at 5, 7, and 11 days.

Finally, the excess of mass or bunching at 11 days coincides with the most significant jump in the average replacement rate: starting at 11-day-long claims, workers are fully reimbursed for the time off. This jump incentivizes workers to extend their leaves to enter the "full" insurance region. Panel (b) of Figure 1.3 shows missing mass in durations just below the eleven-day jump: eight, nine, and ten days. I estimate that 11-day-long sick leave claims are 4.55 percentage points more likely than what the underlying distribution of health predicts. I interpret these patterns as suggestive evidence that workers respond to the discontinuity in the replacement rate at 11 days.[21]

Sick leave duration by workers' characteristics. Figure A.31 shows the histogram of sick leave claim duration by worker characteristics. I group workers into eight groups or bins defined based on age and occupation type: blue-collar and white-collar occupations.[22] Conditional on worker occupation, older workers require a higher proportion of long sick leave claims. Their distribution of sick leave claims is shifted toward the right relative to

[21] Missing mass at eight, nine, and ten days could also be explained by the rounding at seven days. In the model, I allow for a rounding process that rounds up (down) sick leave claims with durations in three days neighbor.

[22] A blue-collar worker refers to an individual who performs manual labor. For example, operators, assemblers, and laborers are considered blue-collar workers. A white-collar worker refers to an individual who performs professional, desk, managerial or administrative work. For example, sales representatives are considered white-collar workers. Table A.25 details the occupations classified as blue-collar and white-collar.

the distribution for younger workers (comparing across rows in Figure A.31). This pattern is consistent with workers requiring more time to recover from the same conditions as they age and with older workers suffering more severe underlying conditions.

Comparisons across occupations for workers in the same age group indicate that claims made by blue-collar workers are longer on average, with a smaller share of claims of up to 3 days. This comparison suggests that differences in the underlying distribution of health could be correlated with occupation type. Motivated by these results, I allow the underlying distribution of health to vary with workers' age and occupation in the estimation of the model.

Workers' behavior by day of the week. Incentives to take time off vary with the day of the week when a worker falls sick. For example, the incentives to file a two-day-long sick leave claim on a Thursday differ from the incentives to file a two-day-long claim on a Tuesday. The first combination implies four continuous days on leave while the second combination implies two days. I refer to the first case as a "weekend-streak combination".

Figure 1.5 shows the share of sick leave claims filed on each day of the week. For each day of the week, I compute the share of sick leave claims, indexed by j, of duration s that are filed that day day . That is:

$$\text{share}_s^{day} = \frac{\sum_j \mathbb{1}\{dow_j = day, s_j = s\}}{\sum_j \mathbb{1}\{dow_j = day\}} \ .$$

Consider 1-day-long sick leave claims: the share of claims filed on Friday is about three times higher than the share of claims filed on any other day of the week (see Panel (a) of Figure 1.5). Similarly, two-days long claims are more likely to be filed on a Thursday than any other day of the week.[23] This pattern is present for one- to five-day-long claims. Crucially, when

[23]In the data, less than 6% of sick leave claims are filed on weekends. Thus, in the rest of the paper I focus on claims filed between Monday and Friday.

inspecting 7-day-long claims, the share is constant across days of the week.[24] I document that workers are 12% more likely to file a weekend-streak claim than a claim of comparable duration on any other day of the week. To capture this empirical regularity, the model allows workers' behavior to vary with the day of the week of a sick leave claim.

3. Theoretical Framework

In this section, I present the model of paid sick leave provision that I use to derive the optimal sick leave insurance contract. First, I model the choices of an expected utility–maximizing worker who faces uncertainty about her health and her ability to work and outline a definition of moral hazard that applies to this setting. Second, I describe how workers' choices and provision of sick pay affect production. Third, I discuss the social planner's problem. In the rest of this section, I omit i subscripts to simplify notation and present the baseline version of the model. I later describe how individuals might vary across (i) their distribution of health shocks, (ii) preferences over time outside work, and (iii) preferences over extending absences.

3.1 Workers

Workers are subject to a stochastic health shock (θ, dow), drawn from a distribution $G(\theta, dow)$, where θ represents the number of days that a worker is sick and dow indicates the day of the week when a worker falls sick. I assume that θ is discrete and bounded between zero and M and that higher values of θ are associated with longer sickness spells.[25] The sickness distribution $G(\theta)$ accumulates positive mass in the no-sickness realization; i.e., the value of

[24] Claims of durations longer than six days exhibit a similar pattern. I use seven days as a reference point since the share of these claims in the data is greater than the share of 6-day-long claims. Appendix Figure A.29 presents the distribution of the share of sick leave claims by day of the week for claims with a duration of between 8 and 15 days, pooled in 2-day groups.

[25] The sickness level is bounded to capture the fact that paid sick leave insurance aims to provide risk protection from impairments to working ability when full recovery is foreseeable. I focus on sick leave claims for up to 30 days in the empirical application.

zero for θ corresponds to the healthy state.

Sick pay utilization. Upon the realization of the health shock (θ, dow), the worker decides her sick pay utilization to maximize her utility. I assume the worker derives utility over consumption (c) and time outside of work (s), given her budget constraint. The budget constraint is $c = w(M - s) + wB(s)$, where w is the daily wage rate, M is the number of workable days in a month that the worker takes as given and $B(s)$ represents the sick pay transfer function.[26] I assume that $B(s)$ is a piece-wise linear function, with marginal replacement rates (b_j) constant for sick leave claims in a duration bracket $[\underline{s}, \bar{s}]$, and that it is a non-decreasing function of s. The worker's utility takes the following form:

$$u(s; \phi, f, w, B, \theta, dow) = w(M - s) + wB(s) + \phi(s_l(s; dow) - \theta) - \phi f(s - \theta) + \phi q \ \mathbb{1}\{\text{weekend}\}.$$

The last three terms represent utility from time outside work. The preference parameter ϕ reflects the opportunity cost of time away from work relative to the time allocated to consumption. The term $(s_l(s; dow) - \theta)$ captures the utility cost of working while sick $(s_l < \theta)$, and the gains from taking time off when not sick $(s_l > \theta)$. In this expression, s_l indicates business days, a function of total days on leave and the day of the week a sick leave claim starts.[27]

The function $f(s - \theta)$ equals zero if the difference between s and θ is non-positive and it takes positive values otherwise. This implies that there is no cost for the worker to file a sick leave claim for the duration of her health shock, but there is a cost of filing a claim for a duration above her health shock. Thus, the compliance costs function $f(s - \theta)$ is increasing in $(s - \theta)$. These costs are motivated on the risks and efforts associated with extending absences above the time needed for recovery. For example, if a worker is caught in violation

[26] For individuals working full-time, M is the number of workdays in a month.

[27] For example, two days long sick leave claim that starts on a Monday represents two business days away from work, while a sick leave claim that starts on Friday implies one day away from work. See Appendix Table A.29

of *mandatory* rest, the claim can be denied—and not paid by the insurance company—and his reputation could suffer—impacting future promotions or increasing the likelihood of being fired.[28] The compliance costs function can also reflect the effort that the worker could exert to find a physician who would sign off on a longer leave. These mechanisms are captured in reduced form; i.e., I do not model the specific action that workers take to extend their absences.

The term $\phi q \ \mathbb{1}\{\text{weekend}\}$ captures the extra utility that a worker derives when the sick leave claim allows her to not return to work until after the weekend. The indicator variable $\mathbb{1}\{\text{weekend}\}$ is defined as follows[29]:

$$\begin{aligned}
\mathbb{1}\{\text{weekend}\} &= 1 \text{ if } dow = \text{Monday and } \ s = 5 \\
&= 1 \text{ if } dow = \text{Tuesday and } \ s = 4 \\
&= 1 \text{ if } dow = \text{Wednesday and } \ s = 3 \\
&= 1 \text{ if } dow = \text{Thursday and } \ s = 2 \\
&= 1 \text{ if } dow = \text{Friday and } \ s = 1 \\
&= 0 \text{ otherwise.}
\end{aligned}$$

Given this specification of the utility function, workers choose sick pay utilization by trading off the cost of a day away from working $w(1 - B'(s_c))$ with its net gain. This net gain depends on the day of the week and the duration of the claim. An additional day on leave beyond the worker's sickness level (i) lowers utility by increasing the compliance cost term in $\phi f'(s - \theta)$, (ii) increases utility in ϕ if s_l increases by a unit, i.e., if $s_l'(s) = 1$, and

[28] In practice, the insurance office screens claims and audit that workers are resting, e.g., are at home during working hours—over the duration of the leave.

[29] The definition of the indicator variable $\mathbb{1}\{\text{weekend}\}$ considers only the extra utility from sick leave claims with a duration of up to 5 days. A more general definition would be to have $\mathbb{1}\{\text{weekend}\}$ equal one for each sick leave claim that ends on a Friday and assume different values of q for the first and second weekends. I argue that the extra utility from the first weekend is more salient in a worker's decision when filing a sick leave claim. The data presented in Figures 1.5 and A.29 support this assumption.

(iii) increases utility in q if the sick leave claim ends on a Friday. Thus, the net gain of time away from work is given by the term $\phi\left[s_l' - f'(s-\theta) + q\mathbb{1}\{\text{weekend}\}\right]$. The optimal sick pay utilization is $s^*(\phi, f, w, B, \theta, dow) = \text{argmax}u(s; \phi, f, w, B, \theta, dow)$.

Moral Hazard. Insurance provision lowers the marginal cost of sick leave—by lowering the cost of a day away from work—weakly increasing sick pay utilization. That is, $s^*(\cdot)$ is nondecreasing in the sick leave benefits function $B(s)$. In my model, moral hazard refers to the responsiveness of the sick leave demand to varying the generosity of sick pay. The magnitude of this response depends on workers' valuation of time off (ϕ), their preferences for behaving as expected (summarized by the parameters of function f), their wages (w), and generosity of the paid sick leave contract summarized by the function B.

To formalize this, consider two alternative sick pay contracts B^0 and B^1. The contracts may specify different replacement rates given a particular piece-wise linear function or may differ in the shape of the function itself. For example, both contracts could feature a three-day-long deductible but differ in the marginal replacement rate for claims longer than three days. Alternatively, contract B^0 could have a three-days-long deductible, and contract B^1 could have a constant replacement rate. Thus, moral hazard is the change in the demand for sick leave (Δs) when the worker is shifted from contract B^0 to contract B^1:

$$\Delta s = s^*(\phi, f, w, B^1, \theta, dow) - s^*(\phi, f, w, B^0, \theta, dow) \ .$$

This definition follows the conventional use of the term moral hazard. In the health-care insurance literature, the term is used to capture the notion that insurance coverage, by lowering the marginal cost of care to the individual, may increase healthcare use (see Pauly, 1968; Cutler and Zeckhauser, 2000; Einav et al., 2013; Einav and Finkelstein, 2018). Put another way, moral hazard refers to the responsiveness of consumer demand for health-

care to the price she has to pay for it. In the context of paid sick leave contracts, workers (consumers) demand time off (healthcare) by considering the share of wages that is forgone with an absence (price). Thus, moral hazard refers to the responsiveness of workers to the replacement rate. The literature on paid sick leave refers to this responsiveness as moral hazard as well (see Johansson and Palme, 2005 and Ziebarth and Karlsson, 2010).[30]

Optimal utilization under linear contracts and quadratic penalties. To facilitate intuition, I impose the following functional form assumptions: (i) a linear benefit scheme $B(s) = bs$, where $b \in [0, 1]$, and (ii) a quadratic compliance cost function. Under these assumptions, and not taking into account the day of the week when a sick leave claim is filed, the worker's utility function is:

$$u(s; \phi, \kappa, w, b, \theta) = w(M - s) + wbs + \phi(s - \theta) - \phi\frac{1}{2\kappa}(s - \theta)^2 \times \mathbb{1}\{(s - \theta) > 0\}.$$

Thus, the optimal choice of sick leave duration from the worker's perspective, conditional on $s > \theta$, is:

$$s^*(\phi, \kappa, w, b, \theta) = \theta + \kappa\left(1 - \frac{w}{\phi}(1 - b)\right) .$$

In the case of full coverage ($b = 1$), the worker optimally chooses $s^* = \theta + \kappa$. This case is presented in Panel (a) of Figure A.30; for strictly positive values of κ, sick pay utilization is above the worker's health state. Lower compliance costs (higher κ) are associated with greater deviations from the worker's health status. The nonpaid sick leave contract ($b = 0$) is presented in Panel (b) of Figure A.30. If the worker's valuation of time outside work ϕ is

[30]This definition of moral hazard refers to "ex post moral hazard"; i.e., how workers respond to the generosity of sick pay. It abstracts from "ex ante moral hazard"; i.e., actions that workers can take to prevent deterioration of their health. Understanding how these actions are shaped by the generosity of the sick leave system is above the scope of this paper.

greater than the wage rate, the worker would optimally choose to claim longer sick leaves.

Panels (c) and (d) of Figure A.30 consider the case of partial coverage—i.e., a strictly positive replacement rate less than one $b \in (0, 1)$ —for different values of κ and ϕ.[31] All else equal, a greater valuation of time outside work (a higher ϕ) is associated with a longer sick leave claim. Similarly, lower compliance costs (higher κ) are associated with longer sick leave claims. That is, moral hazard is increasing in the valuation of time outside work and decreasing in compliance costs. Additionally, given the worker's preferences, the previous expression shows that the duration of a sick leave claim is increasing in the replacement rate b.

Scope of the model of workers' behavior. The propose model of workers' behavior aims to illustrate how illness affects the absence behavior of employed individuals. Thus, I make some simplifying assumptions to keep the model tractable and the estimation feasible. I discussed some of the most salient assumptions and its consequences next.

The model abstracts from the behavior of "when" to file a sick leave claim—that is, in the model, a worker cannot choose the day when she files a claim. Nonetheless, the model allows for strategic behavior in the duration margin of a sick leave claim. While both margins play a role in the worker's choices, incorporating a filing-day choice in the model would require detailed data on when a worker falls sick and when she files a claim. Absent such data, I assume that workers claim on the day when they fall sick.

The model also abstracts from the interaction between workers and physicians. In reality, physicians write sick leave claims with (partial) information about workers' health. The main implication of not modeling this interaction is that the compliance cost function does not disentangle workers' and physicians' risks and costs. For example, suppose physicians want to avoid facing the cost of being caught signing off an excessively long claim. In that case, compliance costs will be higher, and the duration of sick leave would closely reflect

[31] In this case, workers care about the effective valuation of time, i.e., $\frac{w(1-b)}{\phi}$.

workers' health state.[32]

Expected utility. Ex ante, the worker aims to maximize her expected utility, taken over the distribution of health shocks $G(\theta, dow)$. I assume that the worker is risk averse with a von Neumann–Morgenstern (vNM) utility function of the constant relative risk aversion (CRRA) type: $v(y) = y^{1-\gamma}/(1-\gamma)$, where y corresponds to the realized utility $u^*(\theta, dow)$. Thus, expected utility is given by

$$U = \mathrm{E}\left[v(u^*(\theta, dow))\right] = \int v(u^*(\theta, dow))\ dG(\theta, dow)\ .$$

This utility maximization problem can be viewed as a two-stage problem (Einav et al., 2013). Once the health shock is realized, the uncertainty is resolved, workers aim to maximize the contribution of the *state* utility $u^*(\theta, dow)$ to their expected utility $\mathrm{E}_{\theta,dow}\left[v(u^*(\theta, dow))\right]$ by optimally choosing the duration of a claim s.[33] Put another way, given the health shock, risk preferences become irrelevant. That is, risk aversion does not affect workers' decision over sick pay utilization, and *all else equal*, variation in the utilization of paid leave across workers reflects variation in their preference parameters (ϕ, f and q).

For the rest of the section, it is helpful to consider an economy populated by I workers and let i index workers. Thus, $s^{i*}(\phi^i, f^i, w^i, B, \theta^i, dow^i)$ represents the optimal sick pay utilization choice of worker i when insured in contract B, and $U^i(\theta^i, dow^i)$ represents her expected utility.

[32]Physicians are subject to screenings and penalties for fraudulent sick leave prescriptions. Policies aiming to correct physician behavior would use instruments not included in my model, e.g., the mentioned penalties. This paper focuses on the level of generosity of the system and the trade-off between risk protection and *workers* moral hazard responses.

[33]In this model, the only source of uncertainty the worker faces is over her health status, which is the risk the planner seeks to insure.

3.2 Production

Changes in the generosity of sick pay induce two costs. First, as discussed in the previous section, the cost of moral hazard that accrues to the government (the insurer). This is the typical cost modeled in the design of social insurance programs. Second, sick leave insurance provision, by changing labor supply, affects production. In this section, I propose a stylized version of a model of a firm to capture the production losses (or gains) of changes in the sick pay policy.

Sickness is detrimental to production for two reasons. First, sickness might impair individuals' ability to perform their work, i.e., a worker is less productive when sick. Second, if a worker is contagious and infects her coworkers, further production losses could arise: either by direct absences or by reducing other workers' productivity.

Let ν represent the worker's productivity on a day she is sick. I assume that ν also incorporates any potentially negative effects on coworkers. Thus, $\nu \in (-\inf, 1]$. If $\nu = 1$, the worker is equally productive when healthy and sick and is not contagious, and any value of ν below one implies a productivity loss from sickness.

The firm pays worker i a daily wage w^i independent of her realized health-related productivity. I assume that the firm does not observe workers' health state. Thus, the posted wage does not depend on it. Given the wage rate, labor costs are a function of total days worked by each individual. On the other hand, revenue is a function of days worked and the worker's health state. Let $d_{healthy}$ denote the number of days that i works and is healthy, and let d^i_{sick} the number of days that i works sick. These are a function of the optimal sick pay utilization s^{i*}. Using these definitions, the profits generated by worker i are:

$$\pi^i(s^{i*}) = g(d_{healthy}(s^{i*}) + \nu d_{sick}(s^{i*})) - w^i(d_{healthy}(s^{i*}) + d_{sick}(s^{i*}))\ ,$$

where the price of the good is normalized to one. The production function $g(\cdot)$ exhibits diminishing returns on effective days worked. Total profits equal the sum of individual workers' profit: $\Pi = \sum_{i=1}^{I} \pi^i(s^{i*})$.

In this setting, the firm pays w^i to a worker regardless of whether she showed up sick or healthy, i.e., wages are not conditioned on workers' health-related productivity.[34] Thus, from the point of view of the firm, it would be optimal that worker i stays home every time her marginal productivity when sick ($\nu g'(s^{i*})$) is below w^i. Nonetheless, worker i chooses her labor supply trading of the net pay from working ($w^i(1 - B'(s))$) with her marginal utility of leisure—the wedge between the wage rate and marginal productivity when sick drives the production externalities.

3.3 Social Planner

In this section, I derive the optimal sick pay contract. A maintain assumption in this paper is that the planner offers one contract and only observes the duration of sick leave claims. Thus, the replacement rates could depend only on this dimension.[35]

To facilitate intuition, I start with a stylized version of the model to illustrate the main trade-off the planner faces when designing the optimal policy. Then, I relax these assump-

[34]This assumption is consistent with an information asymmetry between the firm and the worker—the worker knows her health state but the firm cannot observe it—and with contracting frictions—even if the firm could observe workers' health, it might not be able to offer a contract that induces worker to report their health state. Nonetheless, this paper focuses on the consequences of this production externalities and does not explicitly model its source.

[35]Understanding whether it would be optimal to offer more than one contract is beyond the scope of this paper. From a theoretical standpoint, the key condition determining whether the optimal menu features vertical choice is whether consumers with higher willingness to pay have a higher efficient level of coverage (Marone and Sabety, 2022). In practice, almost all paid sick leave systems feature one contract, see Table A.20.

tions and present the full-fledged model of insurance provision.

The textbook case

Consider two simplifying assumptions: (i) the planner offers a linear contract, i.e., $B(s) = bs$, and (ii) workers are risk-neutral. This first assumption allows us to focus on one policy parameter: the replacement rate level b. The second assumption allows us to ignore the value of risk protection and focus on the role of moral hazard responses in the design of sick pay.

Worker i chooses her sick pay utilization by comparing the cost of an absence ($(1-b)w^i$) with the net gain from an extra day off (u_s^i). The former depends on the wage rate w^i and the replacement rate b. The marginal utility of an extra day off depends on the health shock and workers' preferences as derived in the previous section:

$$u_s^i = \phi^i \left[\frac{\partial s_l^i}{\partial s^i} - \frac{\partial f^i(s^i - \theta^i)}{\partial s^i} + q^i \mathbb{1}\{\text{weekend}\} \right] .$$

Thus, when $u_s^i \leq w^i(1-b)$, the individual chooses to go to work, else she takes a day off.

No externalities. Assume that sickness does not affect workers' productivity. That is, worker i is (i) equally productive when healthy and when sick, and (ii) sickness is not contagious. In equilibrium, the firm is willing to pay worker i her marginal product. Thus w_i equals the marginal productivity ν^i.[36] Given this wage rate, worker i efficiently self-selects into working or not working depending on u_s^i.

Panel (a) of Figure 1.4 graphically shows the effect of sick pay provision on workers' choices and welfare. This figure puts together the worker and the firm problems. The

[36] To keep the exposition as simple as possible, I let ν represent the marginal productivity when sick. This is equivalent to assuming that $g'(s) = 1$ in section 3.2. In the no externalities case, ν^i represents workers productivity when sick or healthy.

vertical axis shows the marginal utility of a day off (u_s^i) and her marginal productivity ν^i. The horizontal axis corresponds to the daily wage rate w^i. Thus, at the 45° line, wages equal marginal productivity $w^i = \nu^i$ for different productivity levels. Additionally, at the 45° line, the marginal utility of an extra day off equals its cost—the forgone wage. Thus, absent of sick pay, workers efficiently sort into working ($u_s^i \leq w^i$) and not working ($u_s^i > w^i$).

The provision of sick pay distorts workers' incentives by lowering the cost of absences and induces some individuals to take a day off. The blue shaded area represents *inefficient absenteeism*.[37] It comprises the pool of individuals with marginal utility for time off below their marginal product (ν^i) that takes a day off induced by sick pay provision. This is the pool of workers such that: $u_s^i \geq w^i(1-b)$ and $u_s^i \leq w^i$. If the replacement rate increases from b to b', more workers are induced to call in sick. This is the traditional moral hazard response in insurance provision applied to the sick pay setting: an increase in the replacement rate increases inefficient absenteeism, i.e., workers' behavioral responses reduce welfare.

Production externalities. Let us consider the more interesting case where sickness affects workers' productivity. Note that workers' decisions remain the same: given the wage rate w^i, worker i chooses to go to work if $u_s^i \leq w^i(1-b)$. Nonetheless, when sickness is detrimental to workers' productivity, wages no longer reflect their marginal product when sick.

The firm pays worker i a daily wage w^i independent of her health-related productivity.[38] From the point of view of the firm, it would be optimal that worker i stays home if her marginal productivity when sick (ν^i) is below the wage rate w^i. From a welfare standpoint, optimal employment trades off the productivity from working sick ν^i—which incorporates any potential adverse effect of worker i on her coworkers—against the value of leisure u_s^i. That is, it would be efficient that workers who value time off more than their productivity

[37] This terminology is close to the one proposed by Pichler and Ziebarth (2017).

[38] As discussed in the previous section, this assumption is consistent with an information asymmetry between the firm and the worker and with contracting frictions. However, this paper abstracts on modeling the determination of wages.

do not work when sick. Nonetheless, workers' trade-off abstracts from the reduction in productivity, and production externalities arise.

Panel (b) of Figure 1.4 provides a graphical illustration. I define four regions based on the relation between wages, productivity and the marginal value of a day off. The wage rate w^i is presented in the horizontal axis, and the value of a day off u_s^i in the vertical axis. As in Panel (a), at the 45° line, the marginal utility of an extra day off equals its (private) cost, i.e., the wage rate. In contrast with the previous case, in this figure, I consider one level of productivity ν^i given by the horizontal line labeled health-related productivity. This emphasizes that wages (horizontal axis) do not vary with the health-related productivity level ν.[39]

First, consider the top left area of Panel (b). It corresponds to the pool of individuals who do not work and for whom this is efficient, given their productivity. I refer to this pool of workers as involved in *efficient absenteeism*. That is, workers in this pool have the following relation between their valuation of time off, their productivity when sick, and wages:

$$\text{Efficient absenteeism:} \quad \underbrace{u_s^i \geq w^i(1-b)}_{\substack{\text{Worker's trade-off} \\ \text{do not work}}} \quad \text{and} \quad \underbrace{u_s^i \geq \nu^i}_{\substack{\text{Optimal employment} \\ \text{do not work}}}$$

The bottom right area shows the *efficient presenteeism* case: a pool of individuals who do work ($u_s^i \leq w^i(1-b)$) and for whom this is the efficient response ($u_s^i \leq \nu^i$).

The two darker areas show *inefficient absenteeism* and *inefficient presenteeism*. These are the situations where the workers' choices do not coincide with what the planner would find optimal in terms of employment. Inefficient absenteeism was first described in Panel (a). It refers to the pool of workers who find it optimal to be absent when it would be efficient that they work based on their productivity. Inefficient presenteeism, on the other hand, refers to the situation where workers' value of time off is below the cost of the absence,

[39] Appendix Figure A.28 presents each relevant trade-off separately in the absence of insurance provision.

so they work. However, that valuation is below their marginal productivity when sick:

$$\text{Inefficient presenteeism:} \quad \underbrace{u_s^i \leq w^i(1-b)}_{\substack{\text{Worker's trade-off} \\ \text{work}}} \quad \text{and} \quad \underbrace{u_s^i \leq \nu^i}_{\substack{\text{Optimal employment} \\ \text{do not work}}}$$

What is the effect of an increase in the replacement rate? Consider the case where b increases to b'; this shifts the slope of the effective wage function $w^i(1-b')$. The welfare effects of moral hazard responses has two components. On the one hand, a higher replacement rate induces absences from workers with a relatively low value of time off; it increases inefficient absenteeism. This is the same response as in the no externalities case. On the other hand, a higher replacement rate reduces inefficient presenteeism: workers with relatively low productivity take time off ($u_s^i > \nu^i$).

This exercise illustrates the main differences between the design of sick pay insurance and other health-related insurance programs (e.g., healthcare insurance, disability insurance). The presence of production externalities changes the welfare effect of moral hazard. A higher replacement rate induces workers who would take up "too little" sick pay to take time off efficiently. Thus, workers' behavioral responses do not necessarily make insurance provision more expensive.

The full-fledged model

The previous discussion made some important simplifying assumption. In this section, I relax these assumptions. First, I consider that the benefit function is piece-wise linear ($B(s)$). Second, I assume that workers are risk averse. Risk-averse individuals gain utility from insurance, because it lowers the uncertainty they face. Thus, the optimal design needs to incorporate the value of insurance provision. Aggregate welfare can be written as follows:

$$W(B(s)) = \sum_{i}^{I} \omega^i U^i(B(s); \theta^i, dow^i),$$

where ω^i represents the Pareto weight assigned to worker i and U^i represents her expected utility. To incorporate the production side to the welfare maximization problem, I assume that every worker i obtains the same share of total profits $\pi = \frac{\Pi}{N}$. Thus, workers' expected utility is given by: $U^i(B(s); \theta^i, dow^i) = \mathrm{E}_{\theta, dow}\left[v(u^*(\theta, dow) + \pi)\right]$.

The social planner chooses $B(s)$ to maximize the sum of individual welfare:

$$\max_{B(s)} \quad W(B(s)) = \sum_{i}^{I} \omega^i U^i(B(s); \theta^i, dow^i) \;\; \text{s.t.} \;\; \sum_{i}^{I} s^{i*} B(s) \leq S \;,$$

where S represents the allocated funds to cover the cost of the sick pay system. This constraint allows comparisons across policies that have the same cost. It is not feasible to solve this problem without restricting attention to a sub-set of the universe of piece-wise linear functions $B(s)$. The set of contracts that I consider are the ones characterized by the following transfer function:

$$\begin{aligned} B(s) &= b_1 \; s && \text{for } s \in [1, \underline{s}] \\ &= b_2 \; (s - \underline{s}) && \text{for } s \in (\underline{s}, \bar{s}] \\ &= b_3 \; s && \text{for } s > \bar{s}. \end{aligned}$$

This function features a deductible of $\underline{s}$ days. Absent moral hazard, the optimal sick leave contract would be one featuring full coverage for any duration, i.e., $B_s = b \;\; \forall \;\; s$. It would be socially optimal for all workers to be fully insured against health risks since their leave duration would equal their health state. In the presence of moral hazard, the optimal contract features some level of incompleteness of coverage to deter unjustified leave taking

and minimize the production cost associated with unjustified absences. I empirically derive such contract. To do so, I rely on (i) estimates of workers preferences, which allows me to quantify the value of risk protection and the cost of behavioral responses; (ii) estimates of the underlying distribution of health; and (iii) estimates of the production costs associated with sick pay provision. I discuss how I estimate these objects in the next sections.

4. Model Estimation and Identification Discussion

In this section I start discussing the parametric assumptions and the procedure that I follow to estimate the model of workers' behavior. Then I present heuristic arguments for the variation I use to identify each one of the parameters of workers' preferences.

4.1 Model Estimation: Workers' Behavior

I represent the theoretical model fully in terms of parameters to estimate. I assume the following utility function:

$$\begin{aligned} u(s^i;\phi^i,\boldsymbol{\kappa}^i,w^i,B,\theta^i,\alpha^i) =& w^i(M-s^i)+w^iB(s^i)+\phi^i(s_l^i(s^i;dow^i)-\theta^i)+\phi^i q\ \mathbb{1}\{\text{weekend}\} \\ &-\phi^i\Big[\kappa_0^i(s^i-\theta^i)^2\mathbb{1}\{s^i-\theta^i>0\}+\sum_{j=1}^{3}\kappa_j\mathbb{1}\{s^i-\theta^i=j\}\Big]\ , \end{aligned}$$

where i indexes workers. The compliance cost function takes a flexible functional form: it allows for quadratic penalties and specific costs of deviating one, two, or three days. Additionally, I assume that the transfer function is captured by a piece-wise linear function with day brackets corresponding to those currently implemented in the Chilean system.

Preference parameters. The preference parameter ϕ^i governs the valuation of time outside work. I assume that $\ln(\phi^i)$ is drawn from a normal distribution with mean μ_ϕ and variance

σ^2_ϕ such that

$$\ln(\phi^i) \sim N(\mu_\phi, \sigma^2_\phi) \ .$$

Heterogeneity in the valuation of time outside work reflects variation in the opportunity costs of missing work to recover from a disease or in tastes for leisure relative to consumption. The term ϕq captures the extra utility that a worker derives when the sick leave claim has a weekend-streak duration. I assume that all of the variation in this term is governed by the parameter ϕ; thus, q does not vary across workers and is constant across sick leave duration. That is, the value of a weekend streak combination varies across workers reflecting their valuation of time off.

Variation in the compliance cost parameter κ_0 reflects variation in workers' preferences over behaving as expected or revealing their "true" health status. Additionally, job characteristics can justify variation in κ_0. For example, if a coworker can easily perform a worker's job, workers might face high compliance costs and ask for time outside work that closely follows their health status. Alternatively, high risks of extending a sick leave claim above ones' health would be captured by a high compliance cost. I capture both of these mechanisms in a reduced-form manner. I assume that $\ln(\kappa^i_0)$ follows a normal distribution with mean and variance μ_{κ_0} and $\sigma^2_{\kappa_0}$:

$$\ln(\kappa^i_0) \sim N(\mu_{\kappa_0}, \sigma^2_{\kappa_0}) \ .$$

I interpret κ_1, κ_2, and κ_3 as shifters of the compliance cost of deviating for one, two, and three days, respectively. For example, the value of the compliance cost function for an extra day off is: $f^i_1 = f^i(s^i = \theta^i + 1; \theta^i) = \kappa^i_0 + \kappa_1$. Thus, heterogeneity in κ_0 implies that the cost of deviating for one day varies across individuals—a similar argument applies for two- and three-day-long deviations. Nonetheless, this specification assumes that workers with high

(low) compliance costs face a high (low) cost of deviating for one, two, three, four, or any number of days.

Distribution of health states (θ, dow). Worker i draws a health shock (θ^i, dow^i) from the distribution $G(\theta^i, dow^i) = P(\theta^i = m, dow^i = day|\ X)$. The vector X is a vector of observable characteristics: age and occupation.

I use the Peruvian Handbook of Recovery Times to assign the average number of days that a worker would need to recover from the condition reported in the sick leave claim data, i.e., θ^i. I use these average days as an input to construct the underlying distribution of health $G(\theta^i)$, as discussed in Section 2.. This approach allows me to construct an underlying distribution of health without imposing parametric assumptions.

I observe the day when a sick leave claim is filed from the data. When estimating the model, I assume that workers fall sick on the day that they start a recovery spell, i.e., dow^i is the first day of the sick leave claim filed by worker i. Relaxing this assumption would require an additional source of data that distinguishes between the day that a worker falls sick and the day that she starts an absence spell or files a sick leave claim. Absent such data, I use the starting day of a recovery span as the day when the worker falls sick. I assume that workers file sick leave claims from Monday to Friday and that their work schedule is precisely Monday to Friday. In the data, less than 6% of sick leave claims are filed on weekends. Additionally, 83% of Chilean workers have a regular work schedule (Aguayo Ormeño, 2019).

Heterogeneity in health states. The model of workers' behavior allows for heterogeneity in how workers suffer a health shock. To see this, let the parameter α reflect how sickness affects a worker: workers with a higher α benefit more from time outside work. Consider two workers such that worker a is more affected by the symptoms of any disease than worker b: $\alpha^a > \alpha^b$. For example, workers a and b fall sick with the common cold on Monday, i.e., they suffer the same health shock, but the realization that a gets is worse and she would

need more time to recover. The proposed model implies that worker a would file long sick leave claims: $s^{*a} > s^{*b}$. Nonetheless, the perception parameter (α) is not identified. To see this, note that the utility of worker i explicitly accounting for her perception of a shock is:

$$u(s^i;\phi^i,\boldsymbol{\kappa}^i,w^i,b^i,\theta^i,\alpha^i) = w^i(M-s^i) + w^iB(s^i) + \tilde{\phi}^i \;\; \alpha^i(s_l^i(s^i;dow^i) - \theta^i) + \tilde{\phi}^i q \;\mathbb{1}\{\text{weekend}\} \\ - \tilde{\phi}^i\Big[\kappa_0^i(s^i-\theta^i)^2\mathbb{1}\{s^i-\theta^i>0\} + \sum_{j=1}^{3}\kappa_j\mathbb{1}\{s^i-\theta^i=j\}\Big] \; .$$

The parameter α^i is not separately identified from ϕ^i, i.e., $\tilde{\phi}^i\alpha^i$ is observational equivalent to ϕ^i. Nonetheless, it is not necessary to separately identify α^i to derive the optimal sick pay policy. What matters for the optimal design of the policy are workers' responses to the incentives generated by the provision of sick pay. These responses are a function of the parameters $\phi, q, \kappa_0, \kappa_1, \kappa_2$, and κ_3 which are identified.

Rounding and measurement error. I include two additional mechanisms when estimating the model to capture the behavior of physicians who prescribe sick leave claims in a reduced-form way.[40] First, I allow the duration of sick leave claims assigned to a worker to differ from the one optimally chosen by the worker. This discrepancy allows the model to accommodate (i) informational frictions between a worker and a physician and (ii) observed sick leaves with a combination of duration and day of the week that the model does not predict. I assume that the duration of sick leave claims is measured with an additive error that has a mean zero and is uncorrelated with the "true" sickness level. That is, I assume that given the optimal sick leave duration s^*, the physician prescribes $\tilde{s}$:

$$\tilde{s} = s^* + \delta \; ,$$

[40] While explicitly modeling physician behavior is relevant for the design of paid sick leave, the lack of available data on physicians' characteristics limits our ability to address the question empirically.

where δ is a mean-zero random variable with support $[-3, 3]$. With probability p_{me}, it takes the values one or negative one; i.e., it shifts the duration of a sick leave claim by one day. With probability p_{me}^2, the duration of the sick leave claim is shifted two days. That is, δ takes the values two or negative two. And a similar argument works for p_{me}^3.

Second, I adjust the sick leave duration to consider the rounding or heaping observed in the data. I interpret this pattern as coming from physicians being more likely to prescribe rest for a number of days that is a multiple of seven. I assume that with some probability p_7, a sick leave claim of duration m is rounded up (down) to 7 days.

Estimation procedure. I estimate a vector of ten parameters: $\Lambda = \{q, \mu_\phi, \sigma_\phi^2, \mu_{\kappa_0}, \sigma_{\kappa_0}^2, \kappa_1, \kappa_2, \kappa_3, p_{me}, p_7\}$. For this estimation, I select informative moments from the sick leave claims data and use the SMM to estimate the vector of parameters that minimize the criterion function. Let $G(\Lambda)$ represent the vector of simulated moments and G^E their empirical counterpart. I aim to find the vector of parameters Λ that minimizes the squared distance between the simulated moments and the moments computed from the data:

$$\min_{\Lambda} \sum_{t=1}^{10} \left(\frac{G_t(\Lambda) - G^E}{G^E} \right)^2 .$$

To compute the simulated moments, I draw a representative sample of the data. This sample consists of a vector of wages, recovery times, and days of the week. The sample is stratified at the workers' group level.[41] The main strength of this approach is that it does not impose parametric assumptions on the distributions of wages and health shocks. Put another way, this strategy allows for arbitrary correlation between the health shocks and workers' wages to capture two empirical facts: (i) as discussed in 2.4, the duration of

[41] Table A.11 verifies balance in terms of workers' characteristics and sick leave utilization between the sample drawn for estimation of the model and the sample used to document workers' behaviors and compute data moments.

days on leave varies with income; and (ii) diagnosis prevalence changes with the age and occupation of workers.

In the estimation, I exploit workers' responses to the incentives created by sick insurance provision. That is, the estimation procedure relies only on workers' observed decisions and does not require imposing optimality of the current policy. Put another way, in the estimation of the model, I do not assume that the current policy is the optimal one, I only need to assume that workers are utility maximizers. This result relies on the fact that the worker's problem can be viewed as a two-stage problem. Once the health shock is realized, workers optimally choose their sick pay utilization. Neither risk preferences nor production effects affect workers' utility. Risk preferences do not affect the utilization decision since the uncertainty is resolved once the health shock is realized. The production effects are not internalized by workers: they take the share of profits as given and not depending on their labor supply decision.[42]

4.2 Moments and Identification

Even though the parameters are jointly estimated, below I provide a heuristic discussion of the most relevant moment for each parameter.

Weekend-streak utility (q). The term $\phi q 1\{weekend\}$ captures the extra utility that a worker derives when the interaction of the sick leave claim duration and day of the week implies a streak of days off work that includes the weekend, which I term a weekend-streak combination. To identify q, I exploit variation across days of the week on which a sick leave claim of duration s is filed. That is, I rely on the fact that the temptation to extend a sick leave claim varies between days of the week. For example, a 2-day-long sick leave claim is more attractive on a Thursday than on a Tuesday. Figure 1.5 illustrates this variation.

[42] I assume that each worker is infinitesimal and her labor supply does not impact profits.

The identification of q relies on the difference between the share of 1- to 5-day-long sick leave claims filed on a weekend-streak day and the share of 1- to 5-day-long claims filed any other day of the week. I pool all the weekend-streak combinations to compute the average share of claims on those days and compare it with the average share of claims made during the rest of the week. The model requires a higher q to rationalize the data if a larger difference is observed. This comparison relies on the idea that the share of sick leave claims of duration s on a non–weekend-streak day is a good counterfactual to estimate the effect of filing a sick leave claim of duration s on a weekend-streak day. The last panel of Figure 1.5 shows this moment graphically, and Table A.30 presents detailed computations.

Compliance cost function $(\mu_{\kappa_0}, \sigma^2_{\kappa_0}, \kappa_1, \kappa_2, \kappa_3)$: I exploit variation across days of the week and sick leave claims duration conditional on workers' health to inform the distribution of compliance costs. I consider the pool of workers with similar characteristics and the same *assigned* recovery time—i.e., I hold fixed workers' health, age, and occupation—and compare their demand for sick pay across days of the week. For each day of the week and assigned recovery time, I compute the share of sick leave claims, indexed by j, of duration s filed by workers with health θ:

$$\text{share}_{s,\theta}^{day} = \frac{\sum_j \mathbb{1}\{dow_j = day, s_j = s, \theta_j = x\}}{\sum_j \mathbb{1}\{dow_j = day, \theta_j = x\}} \ ,$$

where the denominator counts the number of sick leave claims filed on day of the week *day* with primary diagnoses that would require x days of leave and the numerator counts how many of these claims have duration s. For example, the share of workers with a 1-day-long health shock on a Friday who ask for a one-day-long leave is given by

$$\text{share}_{1,1}^{Friday} = \frac{\sum_j 1\{dow_j = Friday, s_j = 1, \theta_j = 1\}}{\sum_j 1\{dow_j = Friday, \theta_j = 1\}} \ .$$

Figure 1.6 illustrates this computation for sick leave claims with a health shock that requires a 1-day-long recovery. I start by computing the share of claims filed for a duration that matches the assigned recovery time on a weekend-streak day and compare this share with the share of claims filed for an extra day on a weekend-streak day. That is, I compare Panel (a) vs. Panel (b) of Figure 1.6. This difference is informative on how costly it is for individuals to ask for an extra day of leave. I restrict this comparison to claims filed for a combination of duration and day of the week representing a weekend streak. This conditioning keeps the incentives for extending a sick leave claim fixed. That is, every combination implies that workers would be on leave through the weekend. These are the darker columns in Figure 1.6. Panel (c) of Figure 1.6 shows how costly it would be to ask for two extra days of leave. Panel (f) summarizes the probabilities of not asking for extra days of leave, asking for one extra day of leave, asking for two extra days of leave, and asking for up to four extra days of leave conditional on filing a sick leave claim on a weekend-streak day.

I perform these comparisons for sick leave claims with diagnoses assigned one, two, and three days of rest (see Figures A.34 and A.35). To inform the distribution of compliance costs, I compute the average share of claims with a given deviation. These shares are presented in Panel (a) of Figure 1.7. The pattern in the data suggest that a one-day-long deviation is not too costly relative to truth-telling while two-day deviations are more costly, as reflected by the lower share of sick leave claims in the third column of this graph.

Value of time off work (μ_ϕ, σ^2_ϕ): The parameter ϕ captures the taste for leisure relative to the taste for consumption. It can therefore be identified by the average ratio of leisure to consumption. I leverage data on wages, duration of sick leave claims, and sick pay to compute this ratio. I compute consumption as the net earnings in a month using data on wages and sick pay, this is the consumption measure implied by the model. To compute leisure, I use

the number of days that a worker is on leave. For worker i, this ratio is computed as follows:

$$LC^i = \frac{\text{leisure}^i}{\text{consumption}^i} = \frac{1}{N^i} \sum_m \frac{w_m^i \times \textit{Days on leave}_m^i}{w_m^i \times \textit{Days worked}_m^i + \textit{Sick pay}_m^i},$$

where m indexes the month of the year. N^i is the number of months in the year in which worker i used at least one sick leave claim. The numerator estimates worker i's valuation of leisure in month m, and the denominator estimates her consumption in month m. Thus, the ratio LC^i is the average relative valuation of leisure for individual i. Figure A.36 shows the distribution of LC^i; the mean and standard deviation of this distribution inform the distribution of ϕ, which I assume to be log-normal with mean μ_ϕ and standard deviation σ_ϕ.

Rounding and measurement error. I use the difference between the share of 5-day-long sick leave claims filed on a Monday relative to the share of claims filed on a Tuesday, conditional on health shocks with a 1-day recovery, to pin down the success probability of the measurement error term δ (see Panel (e) of Figure 1.6). Given the share of claims filed on a Monday, a smaller difference implies that more sick leave claims have been moved away from the most-profitable duration. That is, the smaller the difference, the more likely it is that the observed duration is not the optimal one in terms of workers' utility. To inform the probability of a sick leave claim being rounded to duration that is a multiple of seven, I use the share of seven-day-long sick leave claims.

5. Results: Workers' Behavior

5.1 Parameter Estimates

Table 1.3 presents the values of the estimated parameters. I use the spikes in the share of claims filed on weekend-streak days relative to non-weekend-streak days to identify the

parameter governing the utility that workers derive from sick leave claims that end on a Friday (q). I estimate that, *all else equal* the utility of a worker increases in 0.79 for filing a sick leave claim weekend-streak combination.

Conditional on their health shocks, I exploit the share of sick leave claims observed on weekend-streak days to identify the parameters of the compliance cost function. Panel (b) of Figure 1.7 compares targeted moments from the data and a model-simulated sample. The model matches the distribution of compliance costs—i.e., the cost of reporting the *true* health shock—reasonable well with $\mu_{\kappa_0} = 0.82$ and $\sigma_{\kappa_0} = 1.77$.

I use the distribution of the ratio of leisure to consumption to identify the distribution of values of time outside work, i.e., the distribution of the parameter ϕ. I estimate that, on average, workers value time off work, either to recover from disease or to engage in leisure, about 45% more than their wages. To put this estimate into context, consider that 26.54% of sick leave claims involve non-paid time off, and a total of 67.60% of claims involve partial paid for workers—, i.e., 67.60% of claims have a duration of up to 10 days for which the replacement rate is less than one.

5.2 Workers' responses

Exploiting these estimates, I document how sick leave taking behavior varies with changes in the replacement rate. First, I document how workers' behavior changes if the jump at 11 days is reduced. I consider three alternative systems that keep the marginal replacement rates in each bracket fixed—the slope of each payment function is one for claims above 4 days—and reduce the size of the jump at 11 days. Panel (a) of Figure 1.9 presents the alternative payment schemes, each scheme reduces the jump at 11 days in one day, the less generous alternative features no discontinuity. Panel (b) shows the share of sick leave claims filed under each alternative system.

Two main patterns arise. First, the mass at 11 days decreases monotonically as the jump

at 11 days is reduced. Second, the share of sick leave claims filed for eight, nine, and ten days increases as the discontinuity decreases. This result suggests that workers who would extend their time off to 11 days to enter the "full insurance" region, find this behavior less attractive as the jump at 11 days decreases. I estimate that, on average, the share of 11-days-long sick leave claims declines in 0.1538 percent for each day that is not reimbursed. That is, when move from the current system to the one that pays 10 days out of 11 days, the share of claims filed for 11 days decreases in 0.1417 percent. Similarly, when move from the current system to the one that pays 9 out of 11 days, the share of claims decreases in 0.1565 percent, on average, per day.

5.3 Model Fit

Matched moments. With the estimated parameters, the model matches the most relevant moments, presented in Table 1.4. There are, on average, 12.33% more sick leave claims on weekend-streak days. The share generated by the model is very close: on average, I estimate 14.64% more sick leave claims on weekend-streak days relative to non–weekend-streak days.

Panel (b) of Figure 1.7 compares the share of sick leave claims with non, one, and up to three-days long deviation implied by the data and by a model-simulated sample. I overestimate the share of claims with non deviations—i.e., claims with duration equal to the assigned diagnosis. I slightly underestimate the share of claims for a day above the assigned diagnosis. Nonetheless, the model replicates the decay in the share of sick leave claims with positive deviations quite well. For example, a two-day deviation is more costly than a one-day deviation, as reflected by the lower share of sick leave claims in this category.

The distribution of the ratio of leisure to consumption is assumed log-normal. Under this assumption, the mean generated by the model is slightly higher than the observed in the data, while the variance, on the other hand, is very close.

Specification Tests. I test how well the model matches data moments not used in the estimation. Figure 1.8 compares the share of claims filed for a duration of 8 to 13 days from the data and a model-simulated sample. The model captures the main pattern observed in the data: sick leave claims spike at 11 days, with lower mass at 8, 9, and 10 days. In particular, using the measure of heaping proposed by Roberts and Brewer (2001), I estimate that, in the data, the 11-day duration accumulates an additional 4.50% mass than its neighbors. Using the model-simulated sample, I estimate an additional 4.03% mass relative to its neighbors.[43] The derivation of the optimal policy requires an estimate of the moral hazard costs associated with this discontinuity, thus the importance of a precise estimate of workers' responses to this feature of the paid sick leave contract.

Additionally, I construct the demand for days on leave as a function of the duration of the health shock. For each duration, I compute how many days; on average, workers request to be on leave. Figure A.37 compares the average days on leave from the data and a model-simulated sample. This figure tests the model's ability to replicate workers' sick leave utilization choices and provides evidence that the model can replicate workers' responses to different health shocks. It is important that the model performs well on this dimension since the derivation of the optimal policy relies on estimates of workers' responses to changes in the paid sick leave policy.

I also propose an out-of-sample exercise exploiting data not used in the estimation of the model. Using data on sick leave claims filed in 2019, I compute the vector of moments used to estimate preference parameters and the share of claims with a duration in the neighborhood of 11 days. I compare these moments with their model-simulated counterparts to test the model performance. To obtain the latter, I simulate the model based on a representative sample drawn from the 2019 data and the estimated vector of preference parameters.

[43] Roberts and Brewer (2001) proposes the following measure: $h_z = f(z) - \frac{f(z-1)+f(z+1)}{2}$, where z corresponds to 11 days, and $f(\cdot)$ indicates the frequency of sick leave claims with duration z. Thus, h_z gives the difference between a duration frequency and the average of the frequencies of the two immediately neighboring duration. It indicates how much a duration sticks out from the pattern suggested by its neighbors.

Table 1.5 presents the results of this exercise. The results suggest that the model performs reasonably well out of the sample: preference parameters and the share of sick leave claims of selected durations are comparable in magnitude. Additionally, the model reproduces (i) the decay of the share of sick leave claims with positive deviations and (ii) the excess mass at 11 days.

Robustness Checks: Moments' computation. The estimation of the model relies the computation of moments using data on all sick leave claims filed during the 2017. I ask whether the main moments are affected by restricting the sample to specific times of the year, e.g., Winter.

Weekend-streak utility (q). In figures A.38 to A.41, I compute this moment restricting the sample to claims filed during each quarter of the year. These figures show the same qualitative pattern than Figure 1.5 providing evidence that claims from a particular time of the year, e.g., winter do not drive the patterns in the data.

Compliance cost function ($\mu_{\kappa_0}, \sigma^2_{\kappa_0}, \kappa_1, \kappa_2, \kappa_3$): Figure A.42 proposes a similar exercise. I show the distribution of share of sick leave claims with none and positive deviations for each quarter of the year. This figure tells a similar story: the share of sick leave claims with no or one day deviations are almost the same, and the share of sick leave claims corresponding to longer deviations decreases monotonically.

6. The Optimal Sick Pay Contract: Derivation and Counterfactuals

I use the estimated model of workers' behavior to determine the sick paid leave system that maximizes aggregate welfare. In this section, first I present the set of assumptions that I impose when solving the social planner maximization problem. Second, I discuss what the

optimal system is and compare it with the current system. Finally, I present counterfactual analyses.

6.1 Solving the Social Planner's Problem

I consider solutions of the social planners problem in the set of piece-wise linear contracts with three-brackets. That is, when solving the welfare maximization problem I aim to find the marginal replacement rate b within a sick leave duration bracket $[\underline{s}, \bar{s}]$ that maximizes welfare given the budget constraint. I restrict attention to contracts where $\underline{s}$ equals three days. That is, I constraint the solution to those contracts that reproduce the bracket systems summarized in Table A.20.[44] Motivated by the features of such contracts, I also assume that transfers are nondecreasing, i.e., $B(s+1) \geq B(s)$. Additionally, I constrain the system to be at most as generous as the full-coverage case: $B(s) \leq s$.

The derivation of the optimal contract relies on the estimates of workers' preference parameters discussed in the previous section and requires an estimate of risk preferences and the parameters of the production function. In the baseline estimates, I consider the case where the pareto weights are the same for all workers.

Risk Aversion. Identification of γ would require, for instance, variation in plan choices across workers. Nonetheless, the Chilean paid sick leave system does not offer choice over sick pay plans. Absent this variation, I calibrate γ using results from the literature. I assume that $\gamma = 2$ and present results with two alternative specifications that allow for preference heterogeneity.

Production costs. To quantify the cost of the production losses associated to the changes

[44] Using the proposed framework, one could allow $\underline{s}$ to be a choice variable. Nonetheless, adding the brackets' limit as an additional choice variable increases the dimensionallity of the problem quickly. For example, if we consider the set of contracts with two brackets, and allow bracket's limits to take any value between 2 and 30 days there are 406 contracts. Then, for each contract one should find the vector of replacement rates that maximize welfare.

in the sick pay policy, I need an estimate of the toll of sickness on workers' productivity, i.e., an estimate of ν. Unfortunately, my data do not allow a direct estimate. Thus, I rely on the estimate proposed by Maestas, Mullen and Rennane (2021). This estimate is based on the American Working Conditions Survey (AWCS), which asks a nationally representative sample of U.S. adults to estimate their reduced work productivity when working sick. In the main estimates of the optimal policy, I calibrate $\nu = 0.77$ and consider sensitivity checks regarding this assumption.

6.2 The Optimal Paid Sick Pay Contract

Figure 1.10 presents the total payment function implied by the optimal sick pay contract. The total payment function could be interpreted as the monetary payment a worker with wage $w = 1$ would receive for a sick leave claim with the indicated duration. The optimal policy differs from the current system in three key ways. First, it offers partial replacement, with an average replacement rate of 0.36, for claims of up to three days. This shift increases the utility of workers who would not take sick leave under the current system but do under the optimal policy. At the same time, partial coverage constraints moral hazard since most of the cost of those absences is faced by workers.

Second, the optimal policy eliminates the discontinuity at 11 days and exhibits a higher average replacement rate between 4 and 10 days. This feature curbs the cost of the behavioral responses to the program incentives and provides more risk protection. Implementing the optimal scheme would shift the distribution of sick leave duration relative to the distribution of claims under the current Chilean system: workers would be more likely to file sick leave claims between 8 and 10 days and less likely to file claims for 11 days.

Third, the optimal policy does not offer full replacement for sick leave claims longer than 11 days. The average replacement rate is increasing, as in the current system, but is less generous for longer claims. Taken together, these changes in the replacement rate reflect

that workers value a contract that offers more protection for shorter claims to smooth consumption across different health states. I estimate that workers are willing to give up 1.53% of their earnings to be insured under the optimal policy.

Changes in compliance cost function. In this section, I examine how the optimal sick pay policy changes when workers' are more (less) reluctant to extend sick leave claims. First, *all else equal*, I reduce the cost of filing a sick leave claim longer than the health state (θ). For example, the cost of filing a sick leave claim for an extra day (f_1) is given by:

$$f_1 = f(s = \theta + 1; \theta) = \kappa_0^i + \kappa_1 \ .$$

I use the estimates of $(\mu_{\kappa_0}, \sigma_{\kappa_0}, \kappa_1, \kappa_2, \kappa_3)$ and construct a new distribution of compliance costs by shifting the mean of κ_j such that $E(\kappa_j) = \mu_{\kappa_j} \times (1 + \varepsilon)$ for $\varepsilon = 0.10$. The second exercise considers the case where $\varepsilon = -0.10$.

Using these counterfactual compliance costs distributions, I first show workers' choices assuming that they are insured under the Chilean paid sick leave system. Panel (a) of Figure 1.11 shows that when compliance costs are higher, the average number of days on leave more closely reflects workers' health state. In contrast, when compliance costs are low, workers' ask for longer claims, given their health state. That is, in the scenario with low compliance costs, the average duration of a sick leave claim for a given health shock is longer.

Panel (b) of Figure 1.11 presents the optimal policy for higher compliance costs and the policy for lower compliance costs and compares it to that in the benchmark case. This exercise provides two main lessons. First, when workers use sick leave claims that closely reflect recovery times, their duration is shorter and the optimal contract is more generous. That is, the optimal contract offers more coverage for all sick leave claims using the same budget as the baseline policy. Financing a higher level of coverage is possible due to shorter

sick leave claims—this is a mechanical effect—and smaller production losses. Second, when workers' are more prompt to extending sick leave claims, the optimal policy aims to contain these responses by lowering coverage for all durations. The new policy is bellow the baseline case. This reduction is more marked for longer sick leave claims. This result indicates that the savings from providing less coverage to longer claims outweighs its utility costs—a reduction in coverage lowers the utility value of sick pay provision.

7. Conclusions

This paper addresses a relevant but poorly understood question in the provision of social insurance: What is the optimal paid sick leave system? I answer this question by combining a unique dataset on sick pay utilization and a model of insurance provision. I start by providing descriptive evidence of the main determinants of workers' behavior. I show three main empirical facts (i) workers' sick leave claim utilization varies with age and occupation; (ii) workers respond to the discontinuity in the replacement rate bunching at 11 days; (iii) workers respond to nonmonetary shifts in the temptation to extend their time off through the weekend.

Based on these facts, I develop a model of sick pay provision. The model gives three main insights. First, workers demand sick pay by trading off the utility cost of working while sick with the consumption loss from missing work when taking sick leave. The provision of sick pay lowers the cost of absences, increasing sick pay utilization. This trade-off governs the moral hazard cost of insurance provision. Second, sick leave insurance could generate production externalities arising from extended absences and workers showing up sick (when their productivity is lower). Third, the model provides intuition on the trade-off faced by the social planner (the insurer). The optimal policy balances the benefits of risk protection with the cost associated with moral hazard and production losses.

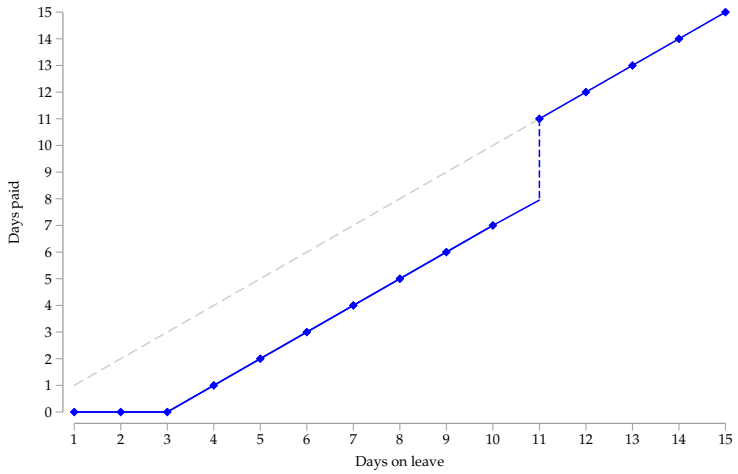
I use the estimated model of workers' behavior to determine the sick paid leave system that maximizes aggregate welfare. I limit attention to those payment schemes of the piecewise linear family. The optimal policy differs from the current system in three key ways. First, it offers partial replacement, with an average replacement rate of 0.36, for claims of up to three days. Second, the optimal policy eliminates the discontinuity at 11 days and exhibits a higher average replacement rate between 4 and 10 days. Doing so curbs the cost of the behavioral responses to the program incentives and provides more risk protection. Third, the optimal policy does not offer full replacement for sick leave claims longer than 11 days. The average replacement rate is increasing, as in the current system, but it is less generous for longer claims. I estimate that workers are willing to give up 1.53% of their earnings to be insured under the optimal policy.

The empirical application of this paper exploits the Chilean context but the insights are informative in other contexts and more generally for the discussion on sick pay policy design. Many paid sick leave systems use the replacement rate as the relevant policy parameter. This paper provides a framework to study and quantify the main trade offs that arise when considering changing this rate.

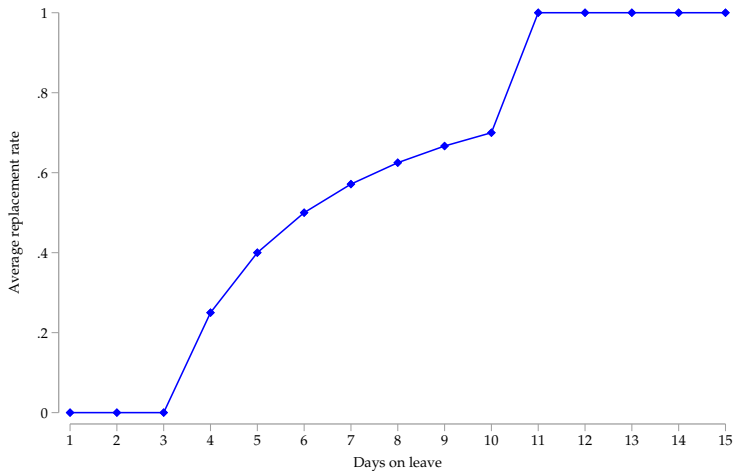
8. Figures

Figure 1.1: Chilean paid sick leave system: benefits computation

(a) Days paid as a function of days on leave

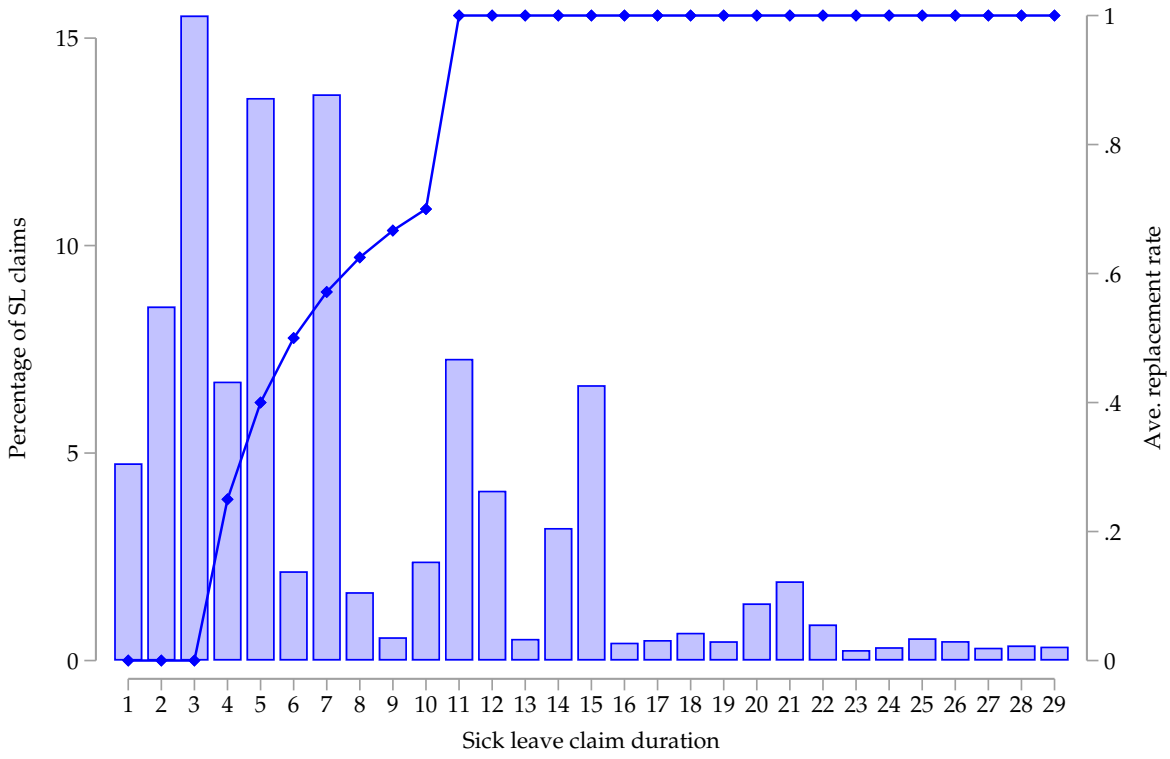


(b) Average replacement rate



Notes: This figure shows the paid sick leave benefit scheme for private-sector employees. Panel (a) shows the number of days paid as a function of days on leave. The replacement rate for the first three days of a sick leave spell is zero. Starting on the fourth day, there is full coverage of each missed day—i.e., the replacement rate is one. If the sick leave lasts 11 days or more, the nonpayable period is reimbursed. Panel (b) shows the average replacement rate, i.e., the ratio between the number of days paid and the number of days on leave. This figure is referenced in Section 2.2.

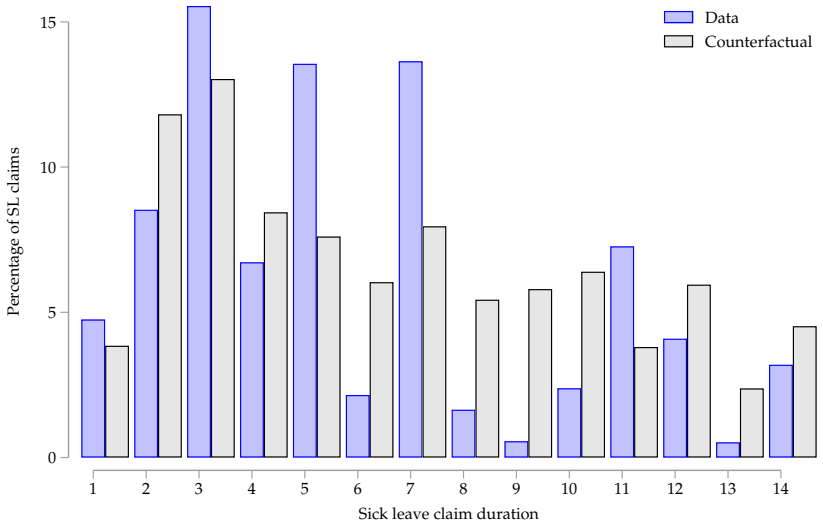
Figure 1.2: Duration of sick leave claims: Private-sector male workers



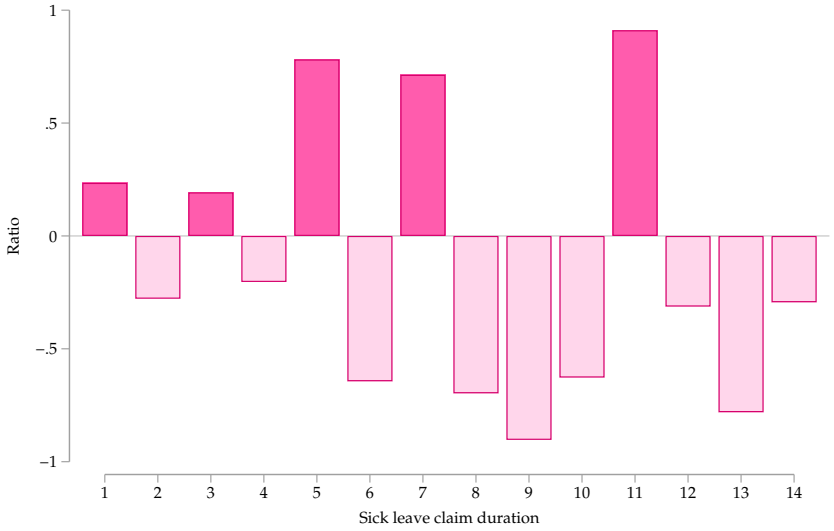
Notes: This figure shows the distribution of the duration of sick leave claims made by male workers on the left-hand-side vertical axis and the average replacement rate on the right-hand-side vertical axis. The figure includes only sick leave claims of up to 29 days; these represent 89% of all claims. This figure is referenced in Section 2.4.

Figure 1.3: Sick leave duration: data and counterfactual distribution

(a) Distributions: data vs. counterfactual



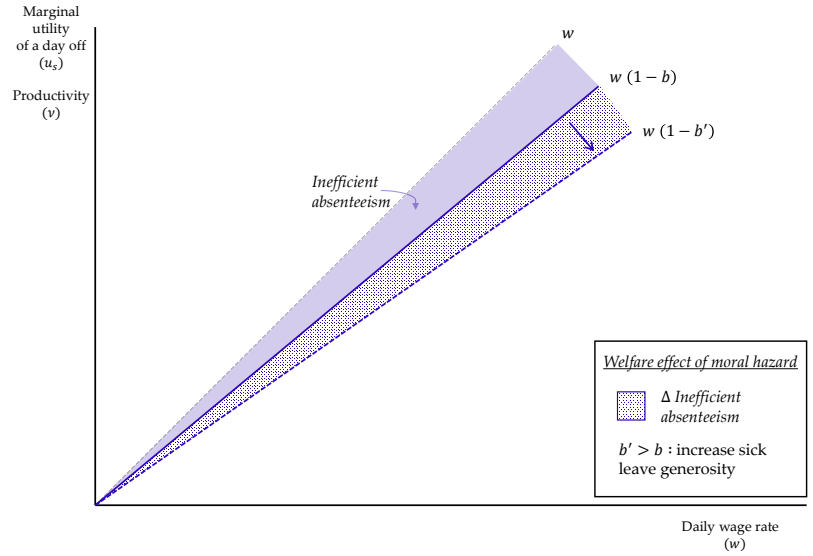
(b) Proportional differences



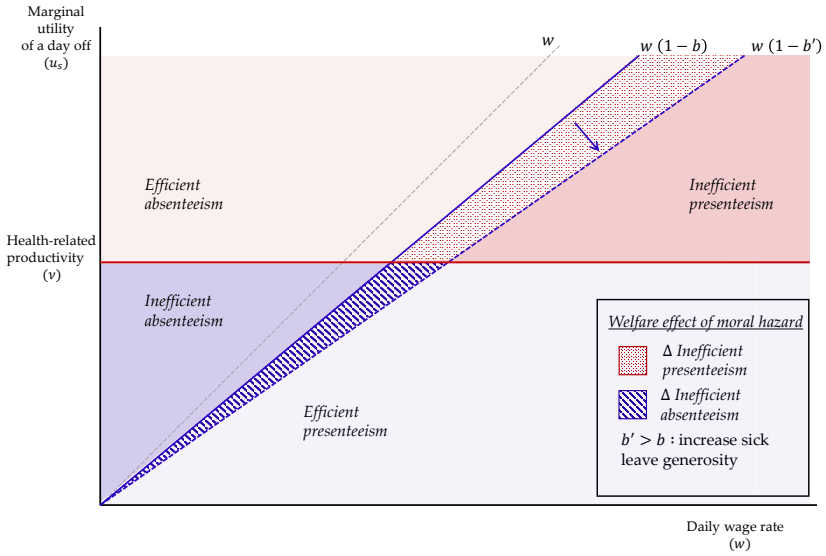
Notes: This figure shows the distribution of days on leave coming from the data, as shown in Figure 1.2, and the counterfactual distribution of days on leave. The latter is constructed assigning to each sick leave the recovery time suggested by the Peruvian Handbook of Recovery Times, adjusted by worker age and occupation. This figure is referenced in Section 2.4.

Figure 1.4: The welfare effects of moral hazard

(a) No externalities case



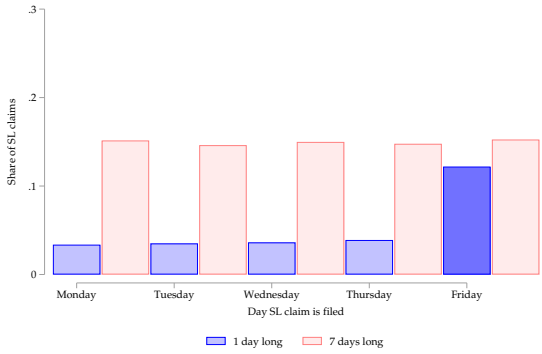
(b) Externalities case



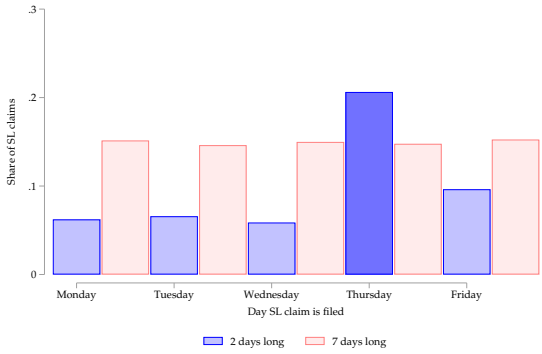
Notes: Panel (a) shows the effect of an increase in the replacement rate on absences in the no production externalities case. The provision of sick benefits gives rise to inefficient absenteeism: a pool of individuals with marginal utility for time off below their marginal product (ν^i) takes a day off. An increase in the replacement rate accentuates this response increasing inefficient absenteeism. Panel (b) presents the case where there are production externalities ($\nu < 1$). The relation between wages, productivity, and the marginal value of a day off defines four regions. The top left area corresponds to the pool of individuals who do not work ($u_s > w(1-b)$) and for whom this is efficient given their productivity ($u_s > \nu$). I refer to this pool of workers as involved in *efficient absenteeism*. The bottom right area shows the opposite situation: a pool of individuals who do work ($u_s < w(1-b)$) and for whom this is the efficient response ($u_s < \nu$). I refer to this pool of workers as involved in *efficient presenteeism*. The other two (darker) areas show inefficient absences ($u_s > w(1-b)$ and $u_s < \nu$) and inefficient work ($u_s < w(1-b)$ and $u_s > v$). This figure is referenced in Section 3.3.

Figure 1.5: Days of the week and sick leave claim duration

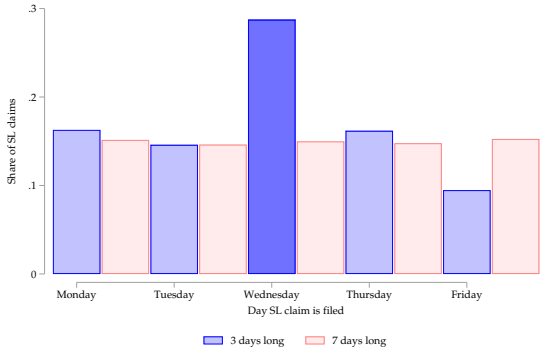
(a) One day



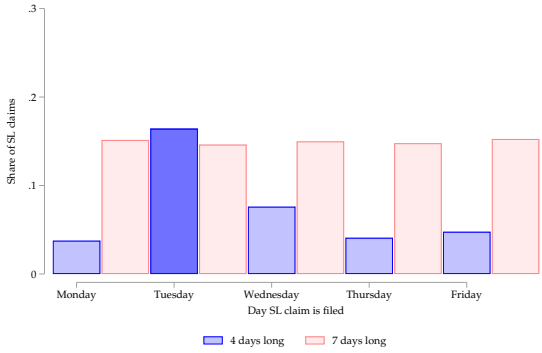
(b) Two days



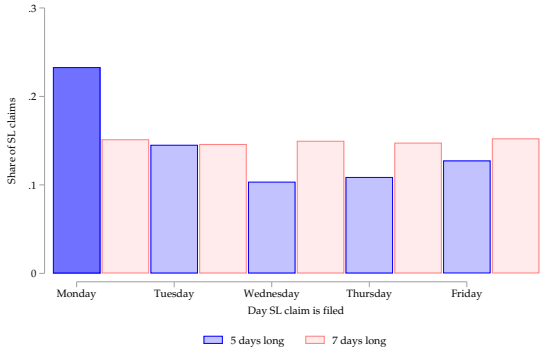
(c) Three days



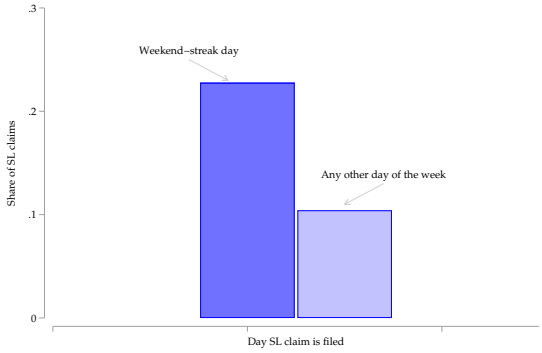
(d) Four days



(e) Five days



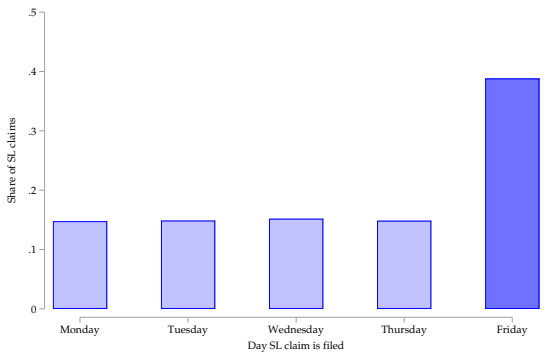
(f) Aggregate across days of the week



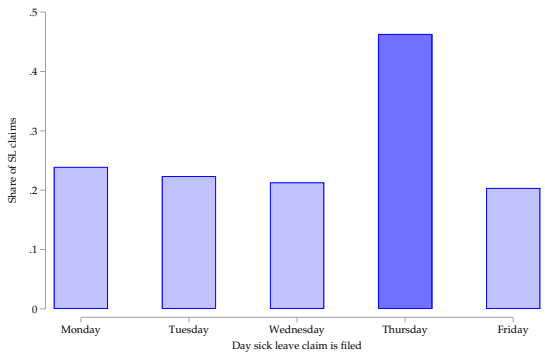
Notes: Panels (a) to (e) show the share of sick leave claims with duration s and the share of seven-day-long sick leave claims filed on each day of the week. Panel (f) aggregates across durations and days of the week: The first bar—labeled "weekend streak"—averages the share of one- to five-day-long sick leave claims that end on a Friday and are filed on any day of the week (for example, one-day-long claims filed on a Friday, two-day-long claims filed on a Thursday, and so on). The second bar—labeled "non–weekend streak"—averages the share one- to five-day-long sick leave claims filed on any other day of the week (for example, two-day-long claims filed on a Friday). Table A.30 reports the estimated shares and moments. This figure is referenced in Sections 3.1 and 4.2.

Figure 1.6: Identification of compliance cost function: Sick leave claims by duration and day of the week
(one-day recovery time)

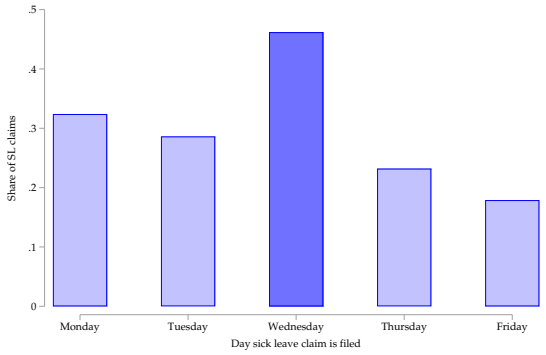
(a) Deviation = 0, duration = 1



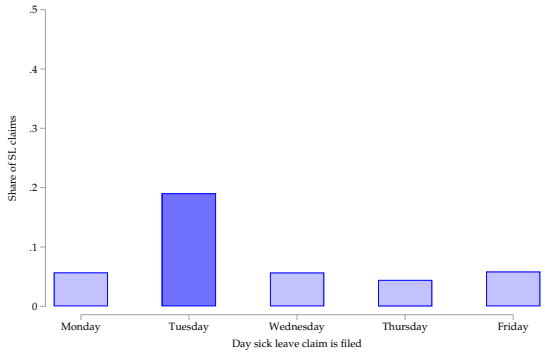
(b) Deviation = 1, duration = 2



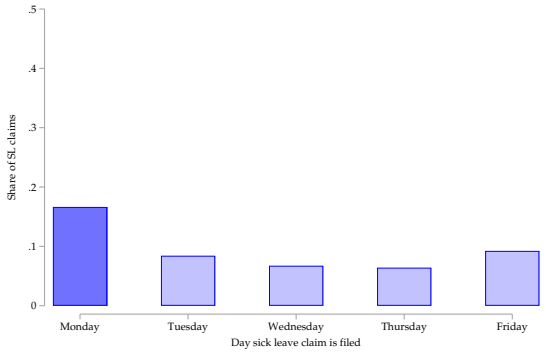
(c) Deviation = 2, duration = 3



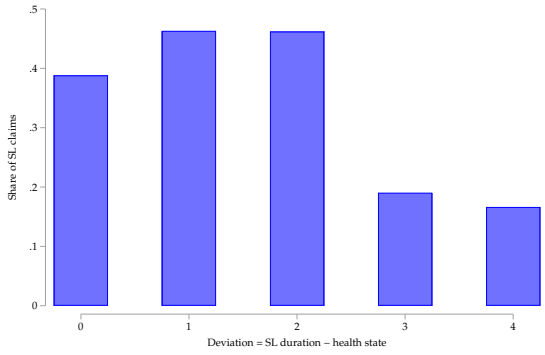
(d) Deviation = 3, duration = 4



(e) Deviation = 4, duration = 5



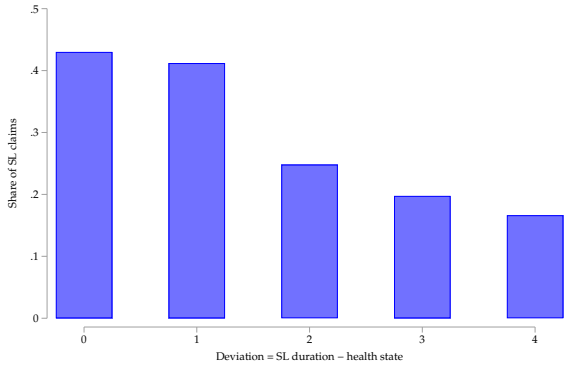
(f) Aggregate across days of the week



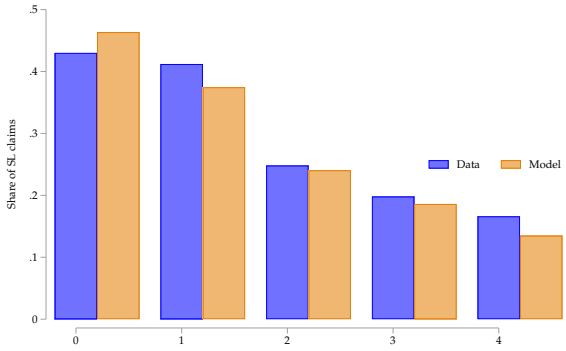
Notes: Panels (a) to (e) show the share of sick leave claims with duration s for workers whose main diagnosis would imply a health state of 1 day on leave. Panel (f) aggregates the share of sick leave claims across days of the week, including only weekend-streak combinations; e.g., from Panel (a), I consider only the share for Friday. This figure is referenced in Section 4.2.

Figure 1.7: Compliance cost function

(a) Share of sick leave claims as a function of deviation

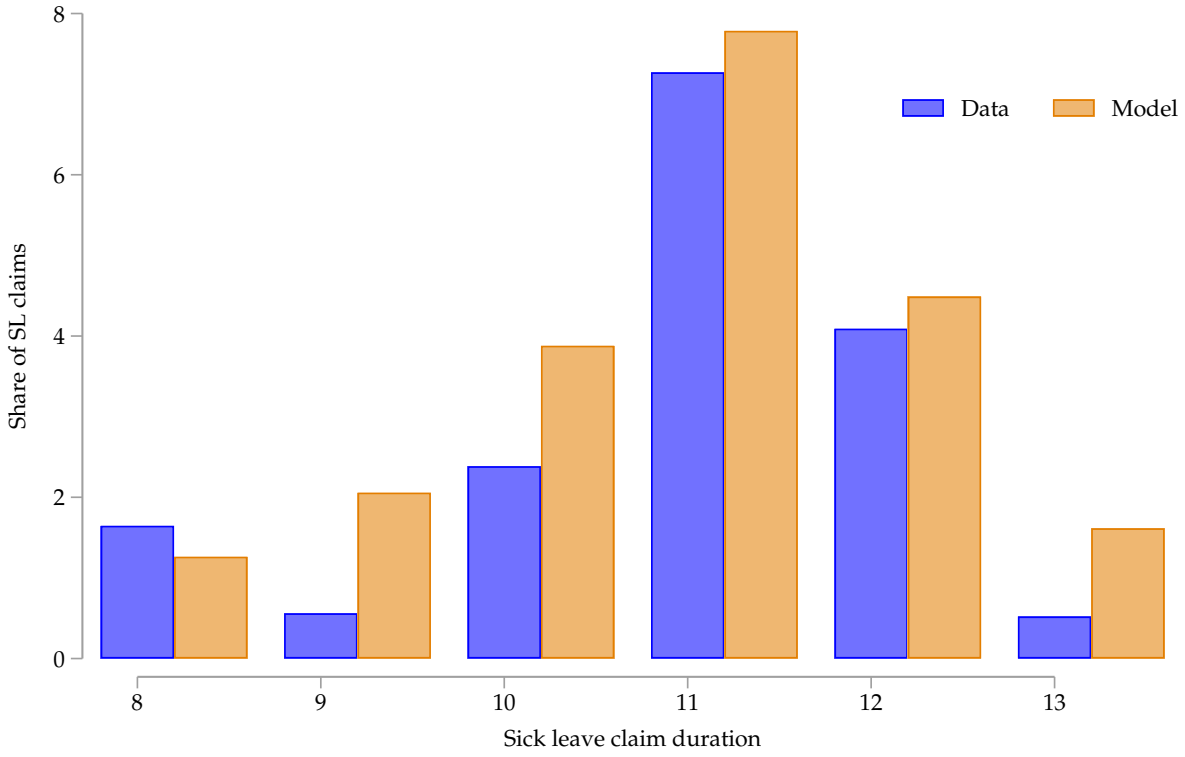


(b) Model fit: Data vs. model moments



Notes: Panel (a) shows the average share of sick leave claims with deviations between 0 and 4 days. The average is computed over sick leave claims with primary diagnoses requiring 1, 2 or 3 days of rest filed on weekend-streak days. Each column is the weighted average of the probability for each health state. Panel (b) replicates this figure and adds the moments computed from the simulated data. This figure is referenced in Section 4.2.

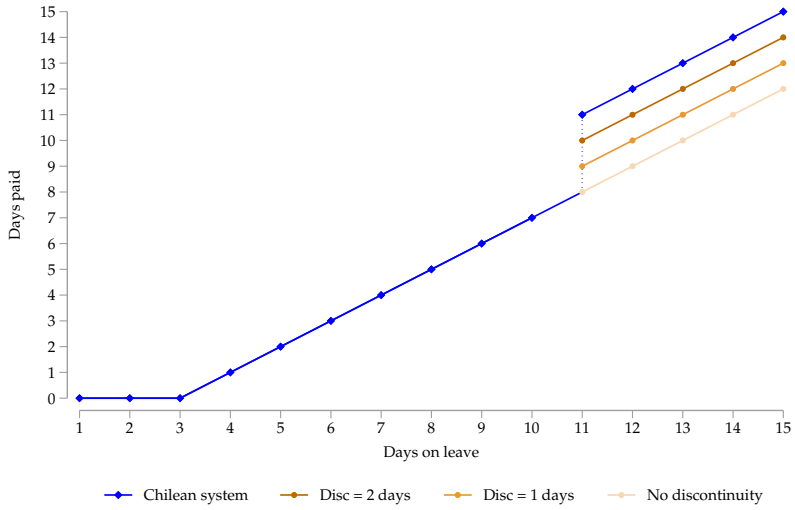
Figure 1.8: Model's fit: Distribution of sick leave claims with durations around 11 days



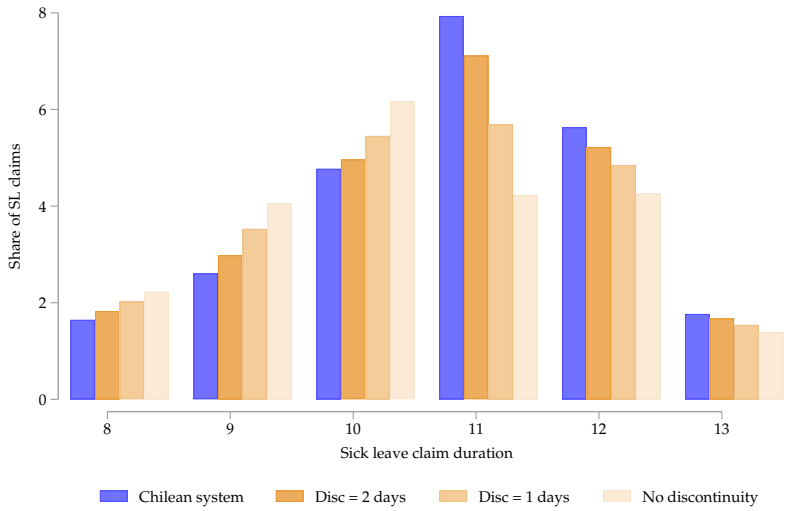
Notes: This figure compares the distribution of sick leave claims with a duration in the neighborhood of 11 days from the data and a model-simulated sample. Using the measure of heaping proposed by Roberts J. (2001) , I estimate that, in the data, the 11-day duration accumulates an additional 4.50% mass than its neighbors. Using the model-simulated sample, I estimate an additional 4.03% mass relative to its neighbors. This measure approximates how much a duration 'sticks out' from the pattern suggested by its neighbors. Roberts J. (2001) proposes: $h_z = f(z) - \frac{f(z-1)+f(z+1)}{2}$, where z corresponds to 11 days, and $f(\cdot)$ indicates the frequency of sick leave claims with duration z. Thus, h_z gives the difference between a duration frequency and the average of the frequencies of the two immediately neighboring duration. This figure is referenced in Section 5.3.

Figure 1.9: Workers' behavior: Changes in the discontinuity at 11 days

(a) Days paid as a function of days on leave

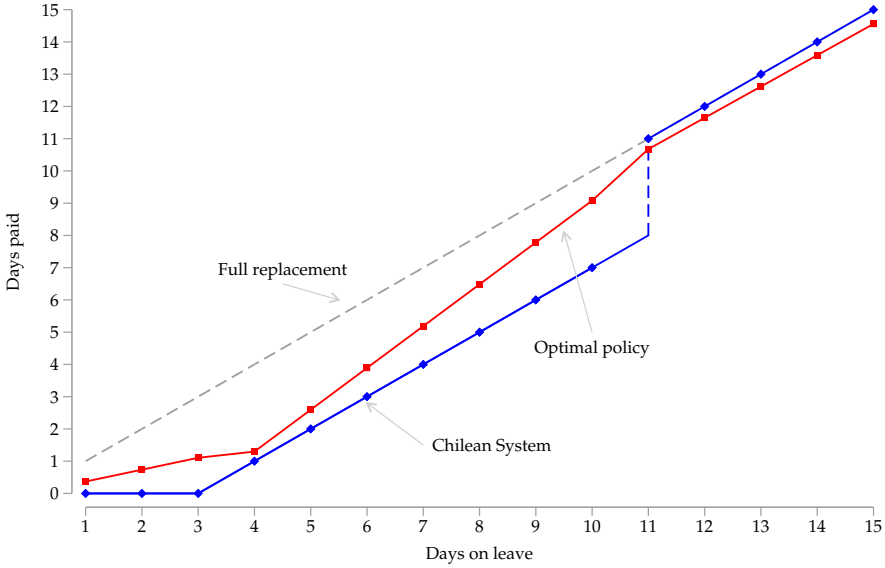


(b) Share of SLC with durations around 11 days



Notes: This figure shows workers responses to changes in the paid sick leave benefit scheme, relative to the current Chilean system. I consider three alternative systems that keep the marginal replacement rates in each bracket fixed—the slope of each function is one for claims above 4 days—and reduce the size of the jump at 11 days. Panel (a) presents the alternative payment schemes and Panel (b) shows the share of SLC filed under each alternative payment scheme. This figure is referenced in Section 5.1.

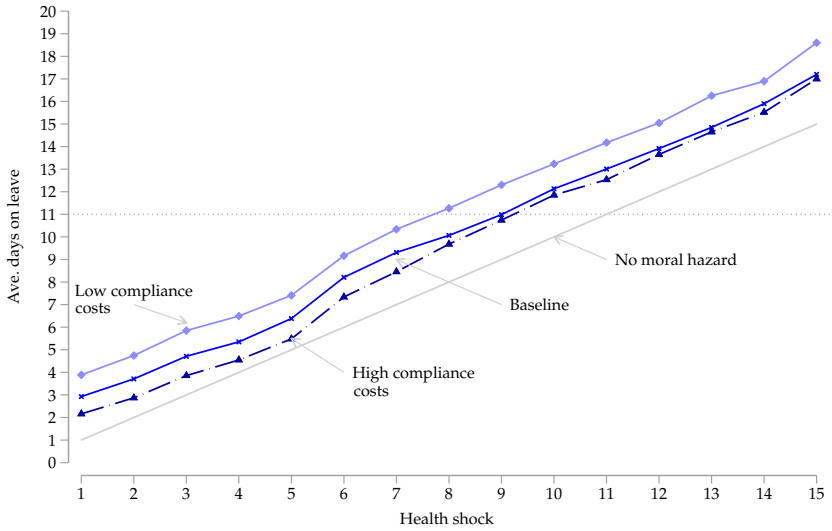
Figure 1.10: The Optimal Sick Pay System



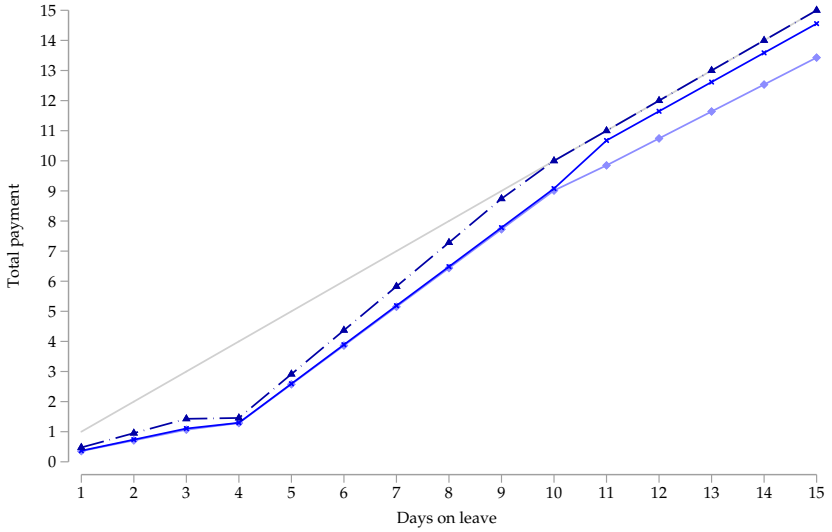
Notes: This figure compares the Chilean system with the optimal paid sick leave system. It presents the total payment function under each contract. This function could be interpreted as the monetary payment a worker with wage $w = 1$ would receive for a sick leave claim with the duration indicated in the horizontal axis. That is, this presents sick pay as a function of the duration of a claim for a unitary wage. This figure is referenced in Section 5.1.

Figure 1.11: Counterfactual exercise: Changes in compliance cost

(a) Workers' responses: Ave. number of days on leave



(b) Optimal paid sick leave contract



Notes: This figure presents the optimal policy under alternative distributions of compliance costs. Panel (a) summarizes workers behavior. It shows the average duration of sick leave claims for the estimated distribution of compliance costs and for two alternative distributions. To construct the high compliance costs distribution, I increase the average cost of an extra day off by 10%. Similarly, the low compliance costs distribution is constructed by decreasing the average cost of an extra day off by 10%. Panel (b) shows the optimal policy under each of these distributions. This figure is referenced in Section 5.1.

9. Tables

Table 1.1: Summary statistics: all workers and workers who use sick leave insurance

	All workers	Workers who had used SL benefits		
		Any	Included conditions	
			All	Up to 30 days
	(1)	(2)	(3)	(4)
Age				
Mean	43.94	43.41	43.30	42.24
Share of workers aged (%)				
25 - 34 years old	26.35	28.90	29.14	32.11
35 - 44 years old	24.48	24.10	24.34	25.35
45 - 54 years old	26.70	24.71	24.66	23.73
55 - 64 years old	22.47	22.28	21.86	18.81
Income (monthly USD)				
Mean	772.00	904.70	909.42	918.02
Standard deviation	367.27	388.79	389.99	390.03
25th percentile	484.45	587.51	591.79	601.77
Median	682.15	829.84	835.53	845.74
75th percentile	997.97	1,146.82	1,152.48	1,161.29
90th percentile	1,328.04	1,483.17	1,489.20	1,496.41
Region (%)				
Central	34.97	40.76	41.22	41.92
Mining intensive	8.96	8.49	8.32	7.62
Health - chronic conditions (%)				
Hypertension	12.90	16.12	15.96	13.93
Diabetes	6.04	7.95	7.51	6.19
Share of workers (%)	100	18.50	17.12	13.78
Observations	1,916,138	354,469	328,053	263,951

Notes: This table presents summary statistics for all male workers in the sample (column 1) and for workers who have used sick leave benefits in the past year based on the conditions and duration of sick leave claims (columns 2 to 4). The sample includes private and public sector employees age 25 to 64 years old. Income statistics are based on the winsorized distribution where the lowest and highest 5% of the income values are excluded. Sick leave claims of up to 30 days account for 95% of all claims filed in a year. This table is referenced in Section 2.4.

Table 1.2: Summary statistics: workers who use sick leave insurance by duration.

	All	Sick leave claims duration		
		1 to 3 days	4 to 10 days	11 to 29 days
	(1)	(2)	(3)	(4)
Age				
Mean	42.00	39.40	41.79	43.97
Share of workers aged (%)				
25 - 34 years old	33.07	41.74	33.74	26.38
35 - 44 years old	25.11	26.22	25.23	24.56
45 - 54 years old	23.50	19.82	23.15	26.10
55 - 64 years old	18.32	12.23	17.88	22.96
Income (monthly USD)				
Mean	870.30	953.50	852.48	849.98
Standard deviation	367.30	386.19	353.57	358.76
25th percentile	569.20	645.54	564.03	552.85
Median	797.88	892.11	782.44	776.38
75th percentile	1,103.05	1,201.53	1,075.22	1,072.83
90th percentile	1,418.81	1,523.37	1,380.31	1,390.51
Region (%)				
Central	48.35	59.10	48.05	43.78
Mining intensive	6.01	4.09	5.30	7.89
Health - chronic conditions (%)				
Hypertension	14.34	11.46	14.48	17.17
Diabetes	6.22	4.64	6.21	7.92
Share of workers (%)	100	32.00	49.76	38.08
Observations	177,531	56,803	88,345	67,599

Notes: This table presents summary statistics for all male workers who have used sick leave insurance in the past year. Column (1) presents characteristics of all workers who have filed at least one claim with duration of up to 30 days for conditions included in the analysis of this paper (see Table A.23 for more details). Columns (2) to (4) present characteristics of workers by duration of the sick leave claims filed. Columns (2) to (4) are not exhaustive, that is, workers can be included in more than one category based on the claims they have filed. This table is referenced in Section 2.3.

Table 1.3: Parameter Estimates

Parameter	Description	Value	Std. error
Preferences parameters			
q	Weekend-streak utility	0.7894	0.2110
$\tilde{\mu}_\phi$	Value of time off relative to consumption, mean	56.8930	26.8567
$\tilde{\sigma}_\phi$	Value of time off relative to consumption, std. dev.	39.8246	13.1373
$\tilde{\mu}_{\kappa_0}$	Compliance costs, mean	0.8290	0.2320
$\tilde{\sigma}_{\kappa_0}$	Compliance costs, standard deviation	1.7714	0.2616
κ_1	Cost of one day deviation	0.3594	0.0960
κ_2	Cost of two days deviation	0.3217	0.0998
κ_3	Cost of three days deviation	0.0041	0.0021
Measurement error			
p_{me}	Prob. physician assigns one day more (less) than asked	0.2992	0.0430
Rounding			
p_7	Prob. SL duration is round to the closest multiple of 7	0.5746	0.1751

Notes: This table presents the estimated parameters for the model of workers' behavior. I assume that the value of time off relative to consumption is distributed log-normal, i.e., $\ln(\phi) \sim N(\mu_\phi, \sigma^2_\phi)$. Thus, $\phi \sim \text{Lognormal}(\tilde{\mu}_\phi, \tilde{\sigma}^2_\phi)$. I report $(\tilde{\mu}_\phi, \tilde{\sigma}_\phi)$. Similarly, I assume that the compliance cost parameter κ_0 follows a log-normal distribution and I report $(\tilde{\mu}_{\kappa_0}, \tilde{\sigma}_{\kappa_0})$ which are the moments that characterize: $\kappa_0 \sim \text{Lognormal}(\tilde{\mu}_{\kappa_0}, \tilde{\sigma}_{\kappa_0})$. The standard errors are based on 200 bootstrap simulations. This table is referenced in Section 5..

Table 1.4: Moments used in the estimation.

Moments	Data (1)	Model (2)
Preferences parameters		
Weekend streak utility		
Weekend streak days relative to non-streak days	0.1233	0.1546
Compliance costs		
Sh. of SLC with 0 day deviation	0.4300	0.4635
Sh. of SLC with 1 day deviation	0.4120	0.3745
Sh. of SLC with 2 days deviation	0.2480	0.2401
Sh. of SLC with 3 days deviation	0.1975	0.1858
Sh. of SLC with 4 days deviation	0.1663	0.1371
Value of time outside work		
Mean time outside work to consumption ratio	0.2972	0.3342
SD time outside work to consumption ratio	0.1962	0.1965
Measurement error		
Sh. Monday SLC - sh Tuesday SLC conditional to 5-days-long and a day of recovery	0.0822	0.0802
Rounding		
Share of 7-days-long claims	0.1364	0.1050

Notes: This table presents the moments used to estimate the model's parameter. Column 2 reports the data moments. Column 3 reports simulated moments.This table is referenced in Section 5.3.

Table 1.5: Out-of-sample: selected moments from 2019 data and simulated counterparts

Moments	Data (1)	Model (2)
Preference parameters		
Weekend streak utility		
Weekend streak days relative to non-streak days	0.1158	0.1428
Compliance costs		
Sh. of SLC with 0 day deviation	0.4385	0.4611
Sh. of SLC with 1 day deviation	0.4223	0.3819
Sh. of SLC with 2 days deviation	0.2293	0.2330
Sh. of SLC with 3 days deviation	0.1843	0.1683
Sh. of SLC with 4 days deviation	0.1650	0.1232
Value of leisure		
Mean leisure to consumption ratio	0.2993	0.3297
SD leisure to consumption ratio	0.1985	0.1980
Share of SLC - selected durations		
7 days	0.1305	0.1053
8 days	0.0162	0.0162
9 days	0.0057	0.0262
10 days	0.0211	0.0461
11 days	0.0834	0.0939
12 days	0.0404	0.0535
13 days	0.0051	0.0172

Notes: This table presents results from an out-of-sample test of the model performance. Column (1) is constructed using data on sick leave claims filed in 2019. The moments included correspond to those used for the estimation of preference parameters and the shares of claims with duration in the neighborhood of 11 days. To test the performance of the model, I compare these moments with their model-simulated counterparts. These are presented in column (2). To construct column (2) I simulate the model based on a representative sample drawn from the 2019 data and the estimated vector of preference parameters. This table is referenced in Section 5.3.

Chapter 2

A Manufactured Tragedy: The Origins and Deep Ripples of the Opioid Epidemic

1. Introduction

Over the past two decades, mortality from opioid overdoses in the United States has increased at an alarming rate, claiming the lives of over 300,000 individuals (CDC, 2021). These tragic losses coincide with an increase in the level of prescription opioids per capita, which grew 232% from 1997 to 2018 (DEA, 2020). Prescription opioids are highly addictive. Even at medically prescribed doses and within short periods of time, they can lead to physiological dependence, with users experiencing tolerance and withdrawal (Sharma et al., 2016; Hah et al., 2017). The potential effects of the rise in the supply of prescription opioids stretch beyond the increase in overdose deaths and include transitions to the use of illegal opioids such as heroin and fentanyl or declines in one's ability to work, recover from illness, and care for children, among other daily activities (Alpert, Powell and Pacula, 2018; Lynch et al.,

2018; Meinhofer and Angleró-Díaz, 2019; Buckles, Evans and Lieber, 2022).

Understanding how we arrived at this situation and the extent of the consequences of the opioid crisis is challenging, given the non-random variation in the use of prescription opioids across geographies and over time (Ruhm, 2019; Currie and Schwandt, 2021). On the one hand, deteriorating socioeconomic conditions in certain geographic areas could cause an increase in the demand for opioid pain killers and also explain the subsequent decline in the same areas, which would lead to negatively biased estimates (Carpenter, McClellan and Rees, 2017; Case and Deaton, 2017; Hungerman, Giles and Oostrom, 2022). On the other hand, the origin of the epidemic coincides with dramatic supply-side changes such as aggressive marketing of prescription opioids, a shift in physician prescribing attitudes, and an increase in the availability of potent opioids.[1] It has been documented that this increase is positively linked to access to healthcare and the number of physicians per capita (Finkelstein, Gentzkow and Williams, 2018). As a result, areas with higher access to opioids are positively selected on these variables, which could, in turn, attenuate the estimates of the effects of the epidemic.

In this paper, we uncovered rich geographic quasi-exogenous variation in the level of prescription opioids to credibly link the pharmaceutical industry to the origin and unfolding of the opioid epidemic. We also use this variation to present causal evidence of the epidemic's tragic consequences for the wellbeing of adults and its intergenerational effects. Our approach exploits detailed features of the initial marketing of prescription opioids, which we obtained from unsealed court records drawn from litigation against Purdue Pharma, the manufacturer of OxyContin —a prescription opioid at the center of the epidemic.[2] Those records show OxyContin was initially promoted to the cancer pain market with the plan

[1] See, for example, Fernandez and Zejcirovic (2018); Alpert et al. (2022); Eichmeyer and Zhang (2020); Schnell and Currie (2018); Finkelstein, Gentzkow and Williams (2018) and Miloucheva (2021); among others.

[2] These court documents are from cases 07-CI-01303 Commonwealth of Kentucky v. Purdue Pharma and case CJ-2017-816 State of Oklahoma v. Purdue Pharma et al; C.A.No. 1884-cv-01808 Commonwealth of Massachusetts v. Purdue Pharma et al. and Case no. 17-md-2804 (N.D. Ohio).

to quickly expand to the much larger non-cancer pain market. This targeting implied that non-cancer physicians and patients in high-cancer areas were first exposed to OxyContin promotion and gained access to potent prescription opioids to treat moderate and chronic pain. Purdue Pharma's later strategy to disproportionally target top prescribers—those in the highest deciles of the opioid dispensing distribution—meant that those initial targets always received more marketing, even in the expansion phase of OxyContin, when the attention was not on the cancer market. Furthermore, Purdue's successful strategy paved the way for the widespread promotion of opioids beyond the cancer market. Other pharmaceutical companies in the market seized this opportunity, and closely emulated Purdue's marketing. Drawing on these insights, we exploit the geographic variation in cancer mortality in the mid-nineties as a proxy for the cancer pain market served by Purdue Pharma and other pharmaceutical companies to assess the role of supply-side factors in the unfolding of the opioid epidemic. We then use this variation as an instrument for the supply of prescription opioids, allowing us to estimate its effects on the wellbeing of adults and its intergenerational effects.

We collect data from multiple sources and construct a panel of commuting zones covering the United States from 1989 to 2018.[3] We use data from the Drug Enforcement Agency (DEA) on the distribution of controlled substances to measure the level of prescription opioids at the commuting zone level. We measure adult wellbeing as health and social assistance need, using data on mortality from opioids and other causes from the National Vital Statistics System and data on beneficiaries of social safety net programs—namely, the Supplemental Nutrition Assistance Program (SNAP), Supplemental Security Income (SSI)—and Social Security Disability Insurance (SSDI) program. To capture the intergenerational effects of the epidemic, we exploit linked data on births and maternal outcomes.

[3]Commuting zones are geographic areas defined to capture local economic markets. They encompass all metropolitan and nonmetropolitan areas in the US. While less granular than counties, commuting zones are much more granular than states (Tolbert and Sizer, 1996).

We start by showing the link between Purdue's marketing targets when introducing OxyContin and the future growth in prescription opioids. Specifically, we estimate a strong positive relationship between higher cancer mortality in the mid-nineties and the rise in prescription opioids after the launch of OxyContin. Commuting zones with the highest cancer incidence at the time of the launch of OxyContin—those at the 95th percentile—received 1.96 more doses of opioids per capita relative to the 5th percentile, accounting for 64% of the growth in prescription opioids from 1999 to 2018.

Turning to the effects of the epidemic on the wellbeing of adults and its intergenerational effects, we find three key results. In terms of opioid-related mortality, increasing the supply of opioids from the 25th to the 75th percentile caused an 89% increase in prescription opioids deaths and a 39% increase in deaths from all opioids. This corresponds to approximately 200,000 deaths. These deaths are concentrated in young and middle-aged adults, with no effects on those 55 and older. We do not find effects of the rise in opioid supply on non-opioid deaths of despair (such as suicide and alcohol-related deaths) or other causes of death.[4] Second, the opioid epidemic had important health and social assistance effects beyond overdose mortality. The number of individuals participating in social safety net programs increased: A move from the 25th to the 75th percentile of prescriptions corresponds to a 57% increase in the share of SNAP recipients and a 38% increase in the share of the population receiving SSI. We also find a 76% increase in the share of the population receiving SSDI. Third, we document large intergenerational effects. We estimate a 10% increase in fertility rates, driven entirely by increases in non-marital births. We find declines in pregnancy duration of 0.24 weeks, a reduction in birth weight of 0.7%, and a worsening of APGAR scores by 0.9%.[5] We estimate that there was no effect on infant mortality, but

[4]Our measure of deaths of despair follows Case and Deaton (2017)'s definition but excludes drug overdose deaths; these are counted in the prescription opioid and all-opioid death categories. More details on these definitions are provided in Section 3.

[5]The APGAR score is a measure of the physical condition of a newborn infant. It is obtained by adding points (2, 1, or 0) for heart rate, respiratory effort, muscle tone, response to stimulation, and skin coloration; a score of 10 represents the best possible condition.

we find an increase in the APGAR score of infants who died in the first year, meaning that healthier infants died. Taken together, these results point to a general deterioration in the lives of adults with serious consequences for the health of their children.

Our identification strategy requires that in the absence of OxyContin's marketing, outcomes in areas with higher cancer mortality would have exhibited the same *trends* as in areas with lower cancer mortality (Goldsmith-Pinkham, Sorkin and Swift, 2020). To test this identifying assumption, we use an event-study approach and investigate the possible presence of differential pre-trends. We do not find any evidence of a relationship between mid-nineties cancer mortality and growth in any of our outcome variables before the launch of OxyContin. In contrast, reduced-form event-study graphs show that for the period soon after the introduction of OxyContin, our instrument predicts higher opioid mortality, a higher share of population receiving SNAP and disability, and higher fertility rates. In addition, we document that areas with higher cancer mortality were not on a differential trend with respect to other socio-economic variables that affect outcomes such as education, income, or other health variables.[6] That is not to say that the variation in cancer mortality across space is randomly distributed. In fact, we find strong demographic predictors of cancer, such as the share of the population over 65 and the Hispanic population share. What is needed to establish a causal interpretation—and what we provide evidence to substantiate—is that areas with high and low cancer mortality were on the same trends in terms of the focal health and economic outcomes.

Further, we propose two placebo exercises to probe the validity of our instrument. First, we show that mid-1990s mortality rates from other causes, such as cerebrovascular disease mortality, are not predictive of the future prescription opioids distribution. In a second placebo exercise, we relate cancer mortality in 1989–1990—the first year of our data—to the evolution of the outcomes of interest before the launch of OxyContin. That is, we test if there

[6]For example, we find that commuting zones with high and low cancer mortality were on the same trend regarding suicide mortality and the share of employment in manufacturing and mining industries.

is a relationship between lagged cancer mortality and the growth of our outcomes outside the analysis period. We find no evidence of such a link. Both of these exercises suggest that the connection between cancer mortality and the prescription opioid distribution is not driven by other underlying health trends, but by the link created by Purdue Pharma through the marketing of OxyContin. Finally, our results are not driven by differential exposure to Chinese import competition, the 2001 economic recession or unemployment at the time of the introduction of OxyContin.

The contribution of this paper is twofold. First, we provide new evidence that links the origin and unfolding of the opioid epidemic to the marketing strategy of OxyContin. While a large literature documents the role of supply-side factors as contributing forces to the epidemic, its *origins* remain understudied.[7] We build on the work of Alpert et al. (2022), who use state-level variation in the regulation regarding the prescription of opioids. They show that five states with early versions of prescription drug monitoring programs (PDMPs), or triplicate prescriptions, received less marketing from Purdue Pharma and reported lower levels of prescription opioids and fewer overdose deaths.[8] In this paper, we exploit richer commuting zone-level variation in the initial marketing strategy of OxyContin to provide new evidence linking its launch to the origin and unfolding of the opioid epidemic.[9] Crucially, this variation allows us to account for important confounders at the state and year level, and serves as a rich instrument for future work on the epidemic.[10] In Appendix 1.3 we quantify

[7]Eichmeyer and Zhang (2020) and Schnell and Currie (2018): physicians; Powell, Pacula and Taylor (2020): access to healthcare; and Fernandez and Zejcirovic (2018) and Miloucheva (2021): pharmaceutical promotions, among others.

[8]These early versions of PDMPs were often referred to as "triplicate" programs. Our reading of Purdue and other pharma industry and academic documents suggests that the industry's perception of what constitutes a "triplicate" program could differ from the designation used by Alpert et al. (2022). Appendix 1.3 examines this evidence and extends the analysis in Alpert et al. (2022).

[9]Previous literature often relies on changes in the access to prescription opioids induced by the adoption of state-level policies, e.g., PDMPs, to indirectly assess the impact of the opioid epidemic on a broad set of outcomes. See Meara et al. (2016), Buchmueller and Carey (2018*a*), Evans, Harris and Kessler (2020), Ziedan and Kaestner (2020), and Gihleb, Giuntella and Zhang (2022). There is, however, debate on what constitutes an operational or mandatory PDMP; the definitions vary across the literature, making it difficult to leverage this variation to estimate the effects of the opioids epidemic.

[10]During this period, there was relevant state-level variation in response to the opioid epidemic, such as the implementation of Prescription Drug Monitoring Programs (PDMP), the regulation of "pill mill"

the gains in power from this empirical strategy and source of variation.

Second, this paper is the first to document the direct effects of the epidemic on important health and social assistance beyond overdose mortality. Mortality from opioids is only one of the many social costs associated with drug use. In 2019, an estimated 10.1 million people in the US aged 12 or older misused opioids (SAMHSA, 2020). These numbers are orders of magnitude larger than the number of deaths. We document the effects on the demand for disability benefits and on participation in one of the largest antipoverty programs in the United States, SNAP, which has not been studied before. Our work is related to that of Savych, Neumark and Lea (2019), who find evidence that an increase in long-term opioid prescribing behavior leads to a considerably longer duration of temporary disability, and to the work of Park and Powell (2021), who document that the rise in access to and consumption of illicit opioids such as heroin and fentanyl increased disability applications by 7%. Finally, we also document the intergenerational impacts of the opioid epidemic. The epidemic has primarily affected individuals in early adulthood through mid-life, with potential costs beyond the generation directly affected. Heil et al. (2011) and Caudillo and Villarreal (2021) document a positive correlation between opioid use and unintended pregnancies and between opioid overdose deaths and non-marital births. We provide the first causal estimates of the effects on fertility and the first direct effects on birth outcomes.

clinics, and the availability of naloxone. The term “pill mill" is typically used to describe a doctor, clinic, or pharmacy that prescribes or dispenses controlled prescription drugs inappropriately (Malbran, 2007). Naloxone is a drug that can reverse an opioid overdose if administered quickly. The level of naloxone access varies by state and over time. Between 2001 and 2017, every US state passed a law that facilitates the widespread distribution and use of naloxone (Doleac and Mukherjee, 2019).

2. Background: The Marketing of OxyContin and the Opioid Epidemic

In 1996, Purdue Pharma introduced OxyContin to the market. When patented, OxyContin was described as a controlled-release oxycodone compound that substantially reduces the time and resources needed to titrate patients who require pain relief on opioid analgesics (Oshlack et al., 1996).[11] Two key technological innovations are responsible for its success. First, its long-acting formula provided 12 hours of continuous pain relief, an improvement over the standard practice of pain relief every 6-8 hours. Second, it is a single-agent narcotic, so there is no ceiling on the amount of oxycodone per tablet.[12] Both of these factors significantly increased patients' access to potent doses of opioids and augmented the risk of dependency and use disorder. For example, Percocet was the most common oxycodone product on the market before 1996 and was mostly sold in the form of 2.5 mg of oxycodone per tablet. In contrast, the most common forms of OxyContin were 20 mg and 40 mg tablets of oxycodone, while 80 mg and 160 mg tablets were also available. Furthermore, OxyContin users rapidly learned that crushing or dissolving the pill causes the oxycodone to be delivered all at once—instead of the slow release over 12 hours—which induces strong euphoric effects.

Prior to the introduction of OxyContin, pain management focused on cancer and end-of-life pain treatment. Patients who suffered from debilitating chronic pain but who do not have a terminal illness were excluded from long-term therapy with opioids, based on care providers' fears of the risk of severe addiction (Melzack, 1990). In this context, MS Contin, a drug also produced by Purdue Pharma, was the gold standard for cancer pain treatment. OxyContin's development was in response to the generic competition Purdue Pharma expected to face when MS Contin's patent protection expired in 1996. In their

[11] Oxycodone is a semisynthetic opioid that is 50% more potent than morphine and is prescribed for acute pain management.

[12] Other oxycodone products on the market were a combination of oxycodone and ibuprofen or acetaminophen, and the toxicity of the former sets a limit on the amount of active ingredients in the formula.

words: *"Because a bioequivalent AB rated generic control-release morphine sulfate is expected to be available sometime during the later part of 1996, one of the primary objectives is to switch patients who would have been started on MS Contin onto OxyContin as quickly as possible"* (OxyContin Launch Plan, September 1995).

OxyContin was intended to take over MS Contin's market and gain ground in the much larger non-cancer pain treatment market, in which opioids were almost absent. Nonetheless, establishing the use of OxyContin for moderate and chronic pain was not an easy task; it was clear to Purdue that they were going to face pushback when expanding to the non-cancer market. Specifically, based on physicians' focus groups in 1995, Purdue concluded that *"there is not the same level of enthusiasm toward this drug for use in non-cancer pain as we identified in cancer pain"* (Purdue Pharma, 1995). The two main barriers Purdue Pharma faced were (i) the fear and stigma related to the use of opioids for non-terminal or non-cancer pain and (ii) the administrative barriers physicians and pharmacies had to overcome to prescribe and sell Schedule II drugs.

To overcome these obstacles, Purdue deployed a comprehensive marketing strategy based on three main pillars. First, to effectively change physician prescribing behaviors, Purdue Pharma implemented an aggressive marketing plan that pushed the message of an untreated pain epidemic that affected millions of Americans on a daily basis. Pain was introduced as the fifth vital sign, with the goal of encouraging the standardized evaluation and treatment of pain symptoms (Jones et al., 2018). This messaging also included misleading statements—for instance, that opioid addiction rates were lower than 1% and that oxycodone was weaker than morphine when it is 50% more potent.[13]

Second, OxyContin was promoted directly to physicians by the largest and highest-paid

[13] *"We are well aware of the view held by many physicians that oxycodone is weaker than morphine. We all know that this is the result of their association of oxycodone with less serious pain syndromes. This association arises from their extensive experience with and use of oxycodone combinations to treat pain arising from a diverse set of causes, some serious, but most less serious. This 'personality' of oxycodone is an integral part of the 'personality' of OxyContin."* Exhibit 11 from Richard Sackler's—chairman and president of Purdue Pharma—deposition, August 28, 2015.

sales force in the industry.[14] The continuous promotion of OxyContin through advertisements, gifts, and promoted medical literature was delivered through repeated visits and calls to physicians. Furthermore, the marketing team carefully tracked physician prescription habits to optimize and personalize their detailing messages.[15] These promotional efforts quickly translated into a growing number of prescription from OxyContin (Figure A.1).

Third, Purdue focused its initial marketing efforts on the physicians and pharmacies who faced less stigma around opioids and who knew how to navigate the paperwork related to the distribution of Schedule II drugs: Those in the cancer pain market. On repeated occasions Purdue states clearly that: *"OxyContin Tablets will be targeted at the cancer pain Market."* (OxyContin Team Meeting, April 1994). *"OxyContin primary market positioning will be for cancer pain."* (OxyContin Team Meeting, March 1995). *"At the time of launch, OxyContin will be marketed for cancer pain."* (OxyContin Launch Plan, September 1995). This, however, was only intended as their entering path to the larger non-cancer pain market. Purdue explicitly stated that:

> *"The use of OxyContin in cancer patients, initiated by their oncologists and then referred back to FPs/GPs/IMs, will result in a comfort that will enable the expansion of use in chronic non-malignant pain patients also seen by the family practice specialists"* (OxyContin Launch Plan, September 1995).

That is, Purdue exploited its previously established network of cancer patients and their physicians to introduce its newest product to the broader pain market. This strategy also solved additional logistical problems related to the sales of Schedule II drugs, such as OxyContin. At the time of launch, only about half of the pharmacies in the country had the

[14] The average sales representative's annual salary of $55,000, was complemented by annual bonuses that averaged $71,500, with a range of $15,000 to nearly $240,000 (Van Zee, 2009).

[15] From 1996 to 2000, Purdue increased its total physician call list from approximately 33,400 to approximately 70,500 physicians; United States General Accounting Office (2003). See Figure A.2 for details on the targeting of top deciles prescribers by Purdue Pharma.

paperwork required to sell Schedule II drugs, and because *"pharmacists are generally reluctant to stock Class II opioids"*, Purdue decided that their *"initial targets will be the 25,000 stores who stock MS Contin"*, where there was no additional paperwork or training required for pharmacies to stock OxyContin.

Purdue's marketing strategy succeeded in making the use of highly addictive opioids standard practice in the treatment of moderate and chronic pain for a wide range of conditions. By 2003, nearly half of all physicians prescribing OxyContin were primary care physicians (Van Zee, 2009). This strategy also opened the door for other pharmaceutical companies to promote their prescription opioids beyond the cancer market following Purdue's leadership. These companies—Janssen, Endo, Cephalon-Teva, Actavis, Insys, and Mallinckrodt—who are also part of dozens of lawsuits for their role in the opioid epidemic, closely shadowed OxyContin's marketing intending to grow by reducing OxyContin's market share: *"Success means increasing Duragesic share at the expense of OxyContin"* (Sales Force Memorandum, 2001, Janssen Exhibit S0510, State of Oklahoma v. Purdue Pharma et al.).[16]

Finally, Purdue's later strategy to promote only to top opioid prescribing physicians, those in the highest three deciles of the distribution (Figure A.2), meant that areas with high initial promotion as a result of the cancer market focus, also observed higher future promotion when Purdue's plan shifted to the broader pain market.[17] This created a path dependency that made initial targets always relevant even when the distribution of opioids expanded.

For our identification purposes, Purdue's marketing strategy means that areas with a higher incidence of cancer at the time of the launch of OxyContin received a disproportionate amount of marketing and prescriptions for OxyContin and other opioids. In practice, this created a spillover in high-cancer communities from cancer patients, to non-cancer patients,

[16]Duragesic is a fentanyl patch manufactured by Janssen, the pharmaceutical company of Johnson & Johnson.

[17]Other pharmaceutical companies also follow this strategy. For example, Janssen referred to *high decile prescribers* as their *highmost important customers* in a Sales Force Memorandum for Duragesic in 2001.

through their common physicians. In this context, the ideal instrument is a measure of the cancer market Purdue Pharma was serving with MS Contin prior to the introduction of OxyContin. Hypothetically, there are multiple ways to proxy this market. One is to use mid-nineties MS Contin prescription rates as this was Purdue's gateway to the non-cancer pain market. However, for the period of analysis, these data are not available at the county or commuting zone level. Alternatively, we could leverage the State Drug Utilization Data (SDUD), which reports the number of prescriptions paid by Medicaid agencies at the state level. This dataset does not allow us to exploit within-state variation. However, it is useful to document that, at the state level, there is a strong relationship between MS Contin prescription rates and mid-nineties cancer mortality prior to the launch of OxyContin (see Figure A.3).[18]

We proxy the market served by Purdue Pharma using cancer mortality between 1994 and 1996. This variable is available at the county level and is accurately and consistently measured throughout the period. Additionally, it has a close connection to the rates of cancer patients who are using opioid painkillers to manage cancer pain (e.g., MS Contin), especially in the later stages of cancer treatment.[19] This instrument allows for the identification of the role of supply in the origin of the crisis and to estimate the causal effect of the opioid epidemic on a broad range of outcomes.

[18]From reading court litigation's documents, we know that at that time, Purdue had access to extremely granular prescription drugs data through a firm called IMS (later called Xponent and today called IQVIA). We have contacted IQVIA to inquire about these data, and they stated they do not keep any historical data records. A plausible alternative instrument is the number of oncologists per capita. This measure, however, is far too concentrated in the largest commuting zones.

[19]An additional measure of cancer incidence is the rate of cancer patients in the population. Unfortunately, incidence measures reported by the CDC and the Surveillance, Epidemiology, and End Results (SEER) program are aggregated at the state level. Importantly, the two measures are highly correlated: the correlation coefficient is 0.88.

3. Data and Summary Statistics

3.1 Prescription Opioids

We digitize historical records from the Automation of Reports and Consolidated Orders System (ARCOS) of the Drug Enforcement Administration (DEA). These reports contain the distribution records of all Schedule II substances by active ingredient (e.g., oxycodone, hydrocodone, and morphine) at the 3-digit ZIP code level—the smallest geographic unit available—from 1997 to 2018.[20] We construct a geographic crosswalk from 3-digit ZIP codes to commuting zones using *Geocorr* (a geographic correspondence engine) powered by the Missouri Census Data Center. Our main independent variable is grams of prescription opioids per capita at the commuting-zone level; this corresponds to the sum of oxycodone, codeine, morphine, fentanyl, hydrocodone, hydromorphone, and meperidine in morphine-equivalent mg. The group of drugs included in the ARCOS changes over time—e.g., to account for changes in the classification of an ingredient. Nonetheless, we focus on a set of prescription opioids that can be tracked consistently over the period of analysis. We report all ARCOS measures in morphine-equivalent doses, equal to 60 morphine-equivalent mg.

The first panel of Table 2.1 presents summary statistics of shipments of all prescription opioids and the three main controlled substances: oxycodone, hydrocodone, and morphine. Oxycodone represents around half of all prescription opioids shipments, and the average commuting zone receives 3.15 doses per capita in a given year. This number masks substantial geographical variation. While some commuting zones received no doses, others report as much as 51.31 oxycodone doses per capita in a given year, Map 2.1 shows this variation. Figure A.4 shows the rapid growth of prescription opioids over time and the dominant role of oxycodone in such growth. Additional summary statistics on opioids shipment over time

[20] ARCOS system data are available online from 2000 to the first half of 2022. We retrieved and digitized the reports up to 2018, the last year of our sample. For periods before 2000, we used the WayBack Machine application to access reports for 1997 to 1999. These data is available here.

are presented in Table A.1.

3.2 Cancer Mortality

To proxy the cancer market served by Purdue Pharma at the time of OxyContin's launch we construct the average cancer mortality rate between 1994 and 1996 at the commuting zone level using a restricted-access version of the Detailed Multiple Cause of Death (MCOD) files.[21] These files record every death in the US along with the county of residence, the underlying cause of death, and up to 20 additional causes and thus represent a census of deaths in the US. The 1989-1998 data use ICD-9 codes to categorize the cause of death, and the 1999-2018 data use ICD-10 codes.[22] Map 2.2 shows large variation in average cancer mortality in 1994 and 1996.

3.3 Outcome measures and control variables

Opioid mortality. We construct two main measures of opioid-related deaths: prescription opioids and all opioid deaths. The prescription opioids category captures deaths whose underlying cause is substances usually found in prescription painkillers such as hydrocodone, methadone, morphine, and oxycodone, among others.[23] We also consider a broader measure of opioid-related deaths, in which we include deaths from heroin and synthetic opioids; e.g., fentanyl.[24] Map 2.3 shows this geographical variation; mortality rates vary from no deaths to as many as 106 per 100,000 residents in the most affected commuting zones.[25]

[21] We also consider age-adjusted cancer mortality and test if our results are sensitive to any of the years used as our baseline cancer mortality measure. We find very similar and strong first-stage estimates across these alternative measures, see Section 6.6.1

[22] We construct cancer deaths as those from malignant neoplasms (codes 140-208 in ICD-9 data and C00-C97 in ICD-10 data) and in situ neoplasms, benign neoplasms and neoplasms of uncertain or unknown behavior (codes 210-239 in ICD-9 data and D00-D48 in ICD-10 data).

[23] We use identification codes T40.2 and T40.3 to specify prescription-opioid-related overdoses in the ICD-10 data and codes 965.00, 965.02, 965.09, E850.1, and E850.2 in the ICD-9 data.

[24] We use identification codes T40.0-T40.4, X42, X62, and Y12 to count deaths from any opioid in the ICD10-data and codes 965.00, 965.01, 965.02, 965.09, E850.0, E850.1, and E850.2 in the ICD-9 data.

[25] The CDC reports that the transition from the ICD-9 to ICD-10 resulted in a small increase in poison-related deaths of 2% (Warner et al., 2011). Appendix Figure A.5 shows the time series for the US for these

Deaths of despair. We also study how the opioids epidemic affected other deaths of despair. Case and Deaton (2015) define deaths of despair as deaths by drug and alcohol poisonings, suicide, and chronic liver diseases and cirrhosis. Our measure of deaths of despair does not include drug poisonings as these are counted in prescription and any opioids deaths separatedly in our analysis. That is, our measure is limited to deaths from suicide, chronic liver disease, cirrhosis, and poisonings that are attributable to alcohol—these deaths amount to, on average, 79% of the deaths studied by Case and Deaton (2017).[26]

Demand for social insurance and welfare programs. We construct a measure of SNAP benefit recipiency rates at the commuting-zone level, using data from the Food and Nutrition Service of the Department of Agriculture. In particular, we use data on county-level participation in the month of January for all years spanning 1989-2018, focusing on beneficiaries of Food Stamps (FSP) and Electronic Benefit Transfers (EBT) in the context of the program. We then aggregate the county-level counts to compute the share of beneficiaries in the population at the commuting-zone level.[27] We include two measures of disability benefits recipiency, constructed as the share of the population that receives Supplemental Security Income (SSI) and who is blind or disabled, and the share of the population 18 to 65 that receives Social Security Disability Insurance (SSDI). Information on the total number of SSI recipients in each county is based on SSI Annual Statistical Reports and Old Age, Survivors and Disability Insurance (OASDI) reports prepared by the National Social Security Administration, which we aggregate at the commuting-zone level.[28]

---

two measures.

[26] We use identification codes K70, K73-74 to count deaths from alcoholic liver diseases and cirrhosis in the ICD10-data and codes 571.0 – 571.4 and 57109 in the ICD-9 data. We count deaths from suicide using codes X60-84 and Y87.0 in the ICD10-data and codes E950-E959 in the ICD-9 data. Deaths from alcohol poising are counted using codes X45 and Y15 in the ICD10-data and codes E850-E858, E860, and E980.1 in the ICD-9 data.

[27] When information at the local level is not available, we impute the state-level share of SNAP recipients. Table A.14 shows the result for the sample of commuting zones that do not require state-level imputation. Our results are not sensitive to this sample restriction.

[28] We observe the number of beneficiaries at a given point in time but do not observe the number of beneficiaries entering or exiting the programs. Thus, we cannot speak to the question of whether a change in the stock is due to people entering more quickly or receiving benefits for a longer time.

Birth outcomes and fertility. Data on birth outcomes come from the Linked Birth and Infant Death Data of the National Vitals Statistic System (NVSS). The microdata for each year between 1995 and 2018 include the deaths of all infants born in that calendar year for which the death certificate can be linked to a birth certificate and all births occurring in a given calendar year.[29] We construct infant mortality as the ratio of infant deaths to live births in a given calendar year. The Linked Birth and Infant Death Data also include data on the infant's condition at birth, such as weight and length of gestation. The main measures of infant health we compute from the birth files are the commuting-zone-level (i) average birth weight for all live births, (ii) share of low-birth-weight newborns, (iii) share of preterm births, (iv) APGAR score of all births, (v) APGAR score of deceased infants, and (vi) median pregnancy duration. Finally, we use the birth files to compute the average fertility rate at the commuting-zone level, defined as the ratio of the number of single pregnancies to the female population aged 15 to 44 years old.[30,31]

Demographic controls. Data on population counts comes from the Survey of Epidemiology and End Results (SEER) which reports population at the county level and by age, race, sex, and Hispanic origin. We use these data to construct the denominators for adult mortality rate measures, e.g., opioid mortality and aggregate mortality. Denominators for infant mortality rate come from the "Denominator File" provided by the NVSS.

In sum, we build a data set at the commuting-zone level, covering the period from 1989 to 2018 for our outcome variables and the instrument. We choose commuting-zones as our unit of observation since it is the geographic space that captures ones economic life—which often

[29] At least 98% of deaths are linked to their corresponding birth certificate. This figure varies by year; e.g., in 2018, 99.3% of all infant deaths were successfully linked, while in 1998, 98.4% of death records were linked.

[30] We follow the CDC's definition to compute the aggregate or general fertility rate. In additional results, we also present fertility rates for other age breakdowns.

[31] Data for the period 1989-1994 come from the Natality Birth Files. These files provide demographic and health data for all births occurring during the calendar year. We use these data to construct infant mortality rates, birth weight, fertility rate, and APGAR scores measures for the analysis in Section 4.5.2

spans beyond county borders—and the access to the local market for prescription opioids.[32,33] ARCOS data are available since 1997, so analyses using this measure are restricted to a later period.[34] We restrict our sample to areas with more than 25,000 residents. This represents 99.8% of all opioid deaths and 99.3% of the total population. Our final dataset is a balanced panel of 590 commuting zones and consists of 17,110 observations.

4. Empirical Strategy

The level of prescription opioids in a given place and time is an equilibrium object determined by supply and demand factors. Supply factors, such as the density of the healthcare network, and demand factors, such as the incidence of pain in the population, affect the level of prescription opioids and may also affect the evolution of our outcome variables. Table 2.2 shows that the distribution of opioids is not random across space, but rather is related to the demographic composition of the commuting zone and its economic performance. A greater share of the white population and higher median income at the commuting-zone level have a positive correlation with prescription opioids per capita; the share of the Hispanic population, the employment rate, and the demand for social insurance have a negative correlation with the opioid supply.[35] This is in line with Finkelstein, Gentzkow and Williams (2018), who estimate that areas with more physicians per capita, higher levels of income and education, lower Medicare spending per capita, and higher scores on a healthcare quality index have higher opioid abuse rates.

[32] We will miss prescription opioid use from those who are willing to cross commuting-zone lines to obtain opioids prescriptions, nonetheless the literature suggests that this is a rare behavior (Buchmueller and Carey, 2018*a*).

[33] We use the crosswalks developed by Autor and Dorn (2013) to go from county-level to commuting-zone-level aggregates. Some commuting zones cross state borders. When this happens, the commuting zone is assigned to the state where the higher share of the zone's population is located. This criterion helps to preserve the strong within-cluster and weak between-cluster commuting ties.

[34] Table A.2 presents summary statistics for the pre-Oxycontin launch period.

[35] We also find a small negative correlation between the share of employment in the manufacturing industry and opioid prescription rates.

To identify the effect of the opioid epidemic we use an instrumental variable strategy that exploits rich geographical variation in the promotional efforts for OxyContin and other prescription opioids as an exogenous source of variation in the future inflows of opioids. We estimate the causal effects of the supply of prescription opioids via the following equations, which are run over our sample of commuting zones for the period 1997-2018:

First Stage:

$$\Delta \ Presc. \ Opioids_{ct} = \alpha_1 + \phi \ CancerMR_{ct_0} + \alpha \ \Delta \ X_{ct} + \gamma_{st} + \upsilon_{ct} \ . \tag{2.1}$$

Second Stage:

$$\Delta \ y_{ct} = \tau_1 + \beta \ \Delta \widehat{Presc. \ Opioids}_{ct} + \tau \ \Delta \ X_{ct} + \lambda_{st} + \varepsilon_{ct} \ , \tag{2.2}$$

where c indexes commuting zones, t indexes years, s indexes states, and t_0 is defined as the average of the pre-OxyContin period. The operator Δ works as follows: For any random variable W_{ct}, ΔW_{ct} equals the difference $W_{ct} - W_{ct_0}$; we refer to this operation as the long-change of variable W_{ct}. Regarding Equation (2.1), $Presc. \ Opioids_{ct}$ corresponds to doses of opioids per capita shipped to commuting zone c in year t and $CancerMR_{ct_0}$ is the cancer mortality rate in commuting zone c in 1994-1996 (t_0). In Equation (2.2), y_{ct} refers to one of our outcomes of interest, e.g., a measure of opioid-related mortality. Both equations include a vector $\Delta \ X_{ct}$ that represents the long-changes in the time-varying control variables. The control variables included are contemporaneous cancer mortality, share of the population over 66, share of the population 18-65, share of the population under 1 year, shares of the white and black populations, share of females, and share of Hispanic population.

We add state times year fixed effects represented by the term γ_{st} (and λ_{st} in the second-stage equation). These fixed effects control for the variation in outcomes over time that is

common to all commuting zones within state s, and purge the variation in the supply of prescription opioids that results from a change in state-level policies—such as the implementation of a PDMP, access to naloxone, and regulation of "pill mills". These policy changes were quite common, for example, between 2007 and 2013, 17 states implemented some version of a PDMP (Buchmueller and Carey, 2018*a*). Between 2001 and 2017, every US state passed a law that facilitates the widespread distribution and use of naloxone (Doleac and Mukherjee, 2019). Since our exogenous variation is at the commuting-zone level, we cannot include commuting-zone fixed effects in the regression. However, by expressing our variable in changes, we can partially absorb some of the variation that is specific to the commuting zone. We cluster standard errors at the commuting-zone level.

We examine how changes in the supply of prescription opioids relate to the initial cancer mortality rate—our measure of the market initially targeted by pharmaceutical companies. Thus, ϕ captures the growth in the supply of prescription opioids per capita for an additional percentile point in cancer mortality. The parameter of interest β captures the causal effect of an increase in one dose of opioids per capita relative to the baseline year on the change in opioid mortality rate (and other outcomes of interest). That is, for a unit increase in the supply of prescription opioids relative to the period 1994-1996, the mortality rate from prescription opioids (and any other *outcome*) changes in β units relative to the pre-OxyContin's launch period. For the IV estimator of β to be consistent, the cancer mortality rate in the baseline period must be (i) strongly correlated with the opioid supply, and (ii) uncorrelated with the error term in the second-stage equation (Equation 2.2). Evidence supporting our strategy was first presented in Section II, in which we discussed Purdue Pharma's marketing strategy and its rationale for focusing on the cancer market as the place to start and expand from. Next, we provide empirical evidence to support this strategy and assess threats to the validity of the instrument.

4.1 Does cancer mortality in the mid-1990s predict growth in the supply of prescription opioids?

We start by providing graphical evidence in Figure 2.4. We divide commuting zones into quartiles according to their level of cancer mortality before the launch of OxyContin and trace the evolution of all prescription opioids, oxycodone, hydrocodone, and morphine in these communities. Panel (a) of Figure 2.4 shows the evolution of the aggregate of prescription opioids per capita in commuting zones in the bottom and top quartiles of cancer mortality in 1994-1996, as well as the evolution of oxycodone—the active ingredient of OxyContin, which accounts for the largest share of this growth.[36] It is clear from the graph that communities with high rates of cancer experienced a much larger influx of prescribed oxycodone (solid orange line) than low-cancer communities (dashed orange line), even though the two groups started the period with a comparable prevalence of oxycodone. Specifically, between 1997 and 2010, areas in the highest quartile of cancer incidence saw an increase in oxycodone gm per capita of 2,900%, and areas in the lowest quartile experienced a growth that was one-third of that, even though the incidence of cancer varied equally across the two groups, as shown in Figure A.6.

Table 2.3 shows the results of the first-stage regression defined in Equation 2.1. Column 1 is a bivariate regression of prescription opioids per capita on cancer mortality at t_0. Columns to the right add time-varying controls and different specifications of fixed effects. Our preferred specification is the one in column 5, in which we control for state-times-year fixed effects and our covariates. Across all specifications, there is a positive and strong relationship between cancer rates in the mid-1990s and the change in opioids per capita. A one-unit (one-standard-deviation) increase in 1994-1996 cancer mortality increases the change in prescription opioids per capita relative to 1997 by 1.1 (0.13 standard deviation).[37]

[36] In Appendix Figure A.7 we present the analogous analysis, but we split the data based on 8 octiles of cancer mortality and observe the same pattern.

[37] We present weak-instrument-robust inference and follow Andrews, Stock and Sun (2019) recommenda-

To put this figure in context, a change from a commuting zone in the 5th percentile of the cancer distribution to the 95th percentile explains 64% of the increase in opioids relative to the base period. We show the strength of this relationship graphically in panel (b) of Figure 2.4 where we plot the first stage coefficients by year. We find that starting in 1998, the second year of the opioids data, and until 2018, the last year in our data, there is a positive and statistically significant relationship between cancer rates and prescription opioids per capita. However, due to the limited data availability on the illicit opioid market, our first stage analysis underestimates the impact of the initial marketing of opioids on the overall level of both legal and illegal opioid use. Previous research has established a strong causal link between prescription opioid use and illegal opioid use (Alpert, Powell and Pacula, 2018; Evans, Lieber and Power, 2019).

Our hypothesis is that the connection between pre-OxyContin cancer rates and future opioid inflows is generated by the marketing efforts from Purdue and other pharmaceutical companies. Unfortunately, most of the data that could test this hypothesis is still confidential. Here, we perform two exercises that provide evidence in this direction, one using public data and the second, using new unsealed records we digitized. First, we examine pharmaceutical marketing in 2013–2018 using the CMS Open Payments database. These data report visits and payments from pharmaceutical manufacturers to physicians related to promoting specific drugs, including payments for meals, travel, and gifts. Panels A and B of Figure 2.5 show that, even 17 years after the introduction of OxyContin, the share of visits and the share of payments to promote opioids relative to all other drugs was higher in high cancer

tions and present the effective first-stage F statistic proposed by Montiel Olea and Pflueger (2013) to assess the instrument's strength. In the rest of this paper, we refer to this as the *effective F-stat*. The value of the F-statistic testing the null hypothesis that the instrument is equal to zero in the first stage is always greater than 10, suggesting that we can reject the null hypothesis. Nonetheless, Lee et al. (2020) suggest that this standard practice of relying on the first-stage F exceeding some threshold (e.g., 10) delivers tests of incorrect size. Thus, to assess the statistical significance of our estimates, we (i) compute the "tF 0.05 standard error" proposed by Lee et al. (2020), which inflates the usual standard errors to take into account the strength of the first stage, and (ii) present *p-values* based on Anderson-Rubin Test (Anderson, Rubin et al., 1949). Based on Lee et al. (2020), we use a correction factor of $\frac{2.75}{1.96} = 1.4031$ to compute the "tF 0.05 standard error." To facilitate its interpretation, we present the *t-statistic* computed with the corrected standard errors. This *t-statistic* should be compared with a critical value of 1.96 to assess the null hypothesis.

commuting zones. Commuting zones at the top quartile of the cancer distribution in the mid-nineties, relative to commuting zones in the bottom quartile of the cancer distribution, received on average 22% more opioids-related visits, and the share of payments was 83% higher. We interpret this as a measure of persistent effect of the initial targeting of cancer areas by pharmaceutical companies.

Second, records on all sales representatives' visits to promote OxyContin in Massachusetts have been released, as part of recent litigation, from May 2007 to December 2018 (Figure A.8). We digitized these data for 2007 to 2011 and created aggregate measures at the county level of the number of visits per 1,000, and the number of targets–either physicians or pharmacists, per 1,000. Panels C and D of Figure 2.5 show scatter plots of these variables on the y-axis and mid-nineties cancer mortality on the x-axis. Both of these measures show a positive relationship between cancer mortality at the time of launch and persistent future marketing in those areas. This persistence of the initial targeting is consistent with the marketing strategy discussed in the internal documents and supports our identification strategy.

4.2 Exogeneity and exclusion restriction: Is cancer mortality in the mid-1990s directly related to our outcome variables?

Variation in mid-1990s cancer mortality across locations is not random; rather, it depends on demographic, environmental and socioeconomic variables. In Table A.3 we find that cancer mortality is: strongly related to share of the population over 65, negatively associated with the share of Hispanic population, and positively associated with mortality from other causes of death. There is not, however, a cross-sectional correlation with our outcome variables: opioid mortality, shares in SNAP and disability, infant mortality rate, or fertility. Nonetheless, the validity of our identification strategy does not require that cancer be randomly distributed across areas, but rather that in the absence of OxyContin marketing, areas with

higher cancer mortality in the pre-OxyContin period (t_0) exhibit the same *trend* as areas with lower cancer mortality in t_0 in terms of our outcome variables (Goldsmith-Pinkham, Sorkin and Swift, 2020).

We provide evidence to support this assumption in four ways. First, we estimate reduced-form type regressions where we interact our instrument with year dummy variables to directly test for the presence of pre-trends, i.e., we estimate an event-study version of the reduced form relationship between the outcome variables and our instrument. For each outcome variable we consider the following specification, which is run over a balanced sample of commuting zones for the years 1989 to 2018:

$$\Delta\ y_{ct} = \alpha_1 + \sum_{\tau=1989}^{2018} \phi_\tau\ CancerMR_{ct_0}\mathbf{1}(Year = \tau) + \alpha\ \Delta\ X_{ct} + \gamma_{st} + \upsilon_{ct}\ , \qquad (2.3)$$

where Δ is the long change operator, y_{ct} is the outcome of interest, and X_{ct} is a vector of time-varying control variables defined previously. $CancerMR_{ct_0}$ is the cancer mortality rate in commuting zone c at time t_0 and it is interacted with a full set of year fixed effects index by τ. In this specification, the coefficients for the pre-OxyContin period; i.e., ϕ_{1989}, ϕ_{1990}, to ϕ_{1995}, test whether the outcome of interest y_{ct} in higher and lower cancer mortality areas followed similar trends before OxyContin was introduced to the market in 1996. This research design allows us to control for state-specific trends and state-level policy changes which were common during this period that directly affected the supply of opioids—e.g., the implementation of PDMP, the regulation of "pill mill" clinics, and the availability of naloxone—and also the evolution of our outcome variables—e.g., welfare reform and child support policies.

Figures 2.6, 2.7, and 2.8 show the results of this estimation on main outcomes of interest.[38] We find that areas with higher cancer mortality in the mid-nineties were not on a

[38] Appendix Figures A.13, A.15, and A.17 complement this analysis.

differential trend along: opioid-related mortality, despair mortality, infant mortality, birth weight, fertility, or share of population using SNAP.[39] There is no evidence of pre-trends, i.e., the estimated coefficients for the pre-OxyContin period are jointly statistically indistinguishable from zero. After the introduction of OxyContin in 1996, strong patterns appear, and mid-nineties cancer mortality starts to predict opioid-related mortality, demand for SNAP, and increased fertility.

Second, there is no evidence of a systemic relationship between lagged cancer mortality and past or future overall health trends and despair. We show that young adults entirely drive the excess opioid mortality we estimate, and opioid mortality for adults over 55 years old does not increase (see Figure 2.9). This supports the argument that our results are not driven by underlying health conditions since the population over 55 would be the closest to the population that drives the variation in our instrument and that instead, what we observe is a spillover from the cancer population to the younger and healthier population, through the introduction of OxyContin in those markets. We also report event study estimates for suicide mortality and overall 75+ mortality —excluding cancer (Figure 2.10). Additionally, we find no evidence of pre-trends for suicide and overall mortality prior to or after the introduction of OxyContin. Finally, in Figure A.9 we document that high cancer places were not on a differential trend along health behaviors such as smoking.

Third, we provide evidence that higher cancer mortality places were not on a differential trend along economic outcomes. To do so, we perform an out-of-sample dynamic reduced-form analysis in our pre-period. That is, we run Equation 2.3 over a sample of commuting zones for the years 1989 to 1995 and estimate if lagged cancer mortality—average cancer mortality rate in 1989 and 1990, the first years in our data—predicts our outcome variables. We present the results of this analysis in Figure 2.11 and Figure A.10. These results demonstrate that before the introduction of OxyContin there is no relationship between our

[39]Data on SSDI and SSI are not available at the county level before 1996 so we can not conduct this exercise for such outcomes.

outcome measures and lagged cancer mortality—the estimated coefficients are statistically indistinguishable from zero. In Appendix Figure A.11, we test for a relationship between the share of employment in the manufacturing and mining industry and cancer mortality before the launch of OxyContin and find no evidence of a differential trends in these variables.

Finally, for variables such as income per capita, educational attainment, or our outcome variables SSI and SSDI rates, for which we do not have yearly data for 1989-1995, we test whether lagged cancer mortality in 1989 and 1990, predicts changes in these variables, using a cross-sectional reduced form analysis. Table 2.4 presents the results of this exercise. In column 1, we find no evidence of a relationship between cancer incidence and relevant economic indicators, and similarly in column 2, which presents this analysis for our outcome variables, including SSI and SSDI, we do not find any relationship. Taken together these results suggest that in the absence of OxyContin marketing, areas with higher cancer mortality exhibit the same *trends* as areas with lower cancer mortality in terms of our outcome variables and additional socio-economic measures.

5. Results

5.1 Effects on Opioid-related Mortality

We start by inspecting the raw data; in panel (a) of Figure 2.6 we split commuting zones based on the cancer mortality distribution and document that early in the 2000s, a wedge starts to appear between high- and low-cancer-incidence groups, and by the end of the sample opioid mortality in high-cancer areas is 75% higher.[40] Second, following the reduced-form approach from Equation 2.3, we estimate that after the launch of OxyContin a strong relationship emerges between mid-nineties cancer mortality and opioid-related mortality as

[40]In Appendix Figure A.12 we present the analogous analysis, but we split the data based on 8 octiles of cancer mortality and observe the same pattern.

shown in Panel (b) of Figure 2.6.[41]

Next we take Equations 2.1 and 2.2 to the data. Commuting zones with the highest cancer incidence at the time of OxyContin's launch received 64% more opioids per capita than their counterparts—i.e., the 95th percentile relative to the 5th percentile. Using this increase as an exogenous change, we estimate that an additional dose of prescription opioids per capita caused an increase in prescription opioid mortality of 0.0068 points and in all opioid mortality of 0.0065 points. The estimates presented in columns 3 and 6 of Table 2.5 are statistically significant using t-ratio inference, Anderson-Rubin weak instrument robust inference, and the tF procedure suggested by Lee et al. (2020). Our results imply that when doses per capita increase from the 25th to the 75th percentile—i.e., a 5.02 dose increase—mortality from prescription opioids increases by 88.6% and all opioid mortality increases by 39.3%.[42]

The ordinary least squares (OLS) estimates (columns 1 and 4 of Table 2.5) differ considerably from the IV estimates. We find a positive correlation between opioid supply and opioid mortality rate, but the difference in magnitude between the OLS and the IV estimates suggests that the former suffers from downward bias. Put another way, by looking at the correlation between opioid supply and opioid deaths, we would underestimate the role of the supply of prescription opioids in explaining the rise in mortality. The negative bias in the OLS estimates is consistent with commuting zones that receive a disproportionate amount of marketing being positively selected on observable characteristics: Areas initially targeted by OxyContin campaigns had better access to healthcare and a larger number of physicians per capita, which served as OxyContin initial network.

Heterogeneous effects. The excess opioid-related mortality induced by the marketing of prescription opioids is by and large coming from young and middle-aged adults and at the begging of the epidemic, is driven mainly by white adults. In Figure 2.9, we present the

[41] In Appendix Figure A.13 we replicate this analysis for any opioid mortality and document similar patterns.

[42] The standard deviation of the distribution of prescription opioids per capita between 1997-2018 is 4.34, thus a change from the 25th to the 75th percentile in such distribution represents 1.15 of a standard deviation.

interactive-reduced-form analysis for three age groups and in Figure A.14 we split the data by race and gender. The analysis by age shows (i) no evidence of pre-trends on opioid mortality for any of these groups, and (ii) opioid mortality increases that are concentrated among individuals aged less than 55 years old. Furthermore, different from the trends in prescription opioids mortality, for the case of any opioid mortality—which adds deaths from heroin and fentanyl—the effects are persistent even in the last years of our sample, for those under 55. Additionally, we find the epidemic affected men and women similarly. When splitting the data by race, we find that estimates for whites are positive and statistically significant starting soon after the launch of OxyContin. For non-whites it takes around a decade for estimates to be positive and statistically significant.

The opioid crisis can be viewed as having occurred in three waves (Maclean et al., 2020). In the first prescription opioids was the main source of overdose deaths, in the second was heroin, and in the third fentanyl. Panel A of Table A.4 reproduces the estimates of the first stage for different starting and ending years, and Panel B replicates the main instrumental variables regressions. We find a strong first-stage relationship between mid-nineties cancer mortality and the supply of prescription opioids in all the stages of the crisis. In terms of the effects, our results suggest that the increase in the supply of prescription opioids had a stronger impact on opioid-related mortality in the first wave of the epidemic. However, the effects are positive and statistically significant across all periods.

5.2 Adult Wellbeing and Intergenerational Effects

In this section, we study whether the access to potent opioids has deteriorated the wellbeing of adults by looking at overall mortality, the demand for social insurance and welfare programs. We then turn to the intergenerational effects of the epidemic.

Other mortality measures. We ask whether the dramatic increase in opioid supply affected all-cause mortality, excluding cancer deaths. These results are presented in Table 2.6. We

find no relationship between overall mortality and the increase in prescription opioids. To put this result into context, note that at their peak in 2017, opioid-related deaths accounted for 1.8% of all deaths. The introduction of effective pain medication could have improved the quality of life of individuals with high incidence of pain and low risks of addiction, and those potential improvements could translate into health indicators. To asses if there is any indication of such improvements we estimate our reduced-form exercise on mortality for those 75 and older, but find no evidence of any effects in mortality (Figure 2.10 Panel b).

Case and Deaton (2017) document a dramatic decline in life expectancy for white non-Hispanic Americans, which is mostly driven by deaths from despair such as drug overdoses, suicides, and alcohol-related liver mortality, and point to a possible connection to the opioid epidemic. We explore this connection studying the effects of opioid supply on *non-opioid-related* deaths of despair. In Appendix Figure A.15 we show that there is only a weak link between the increase in the supply of opioids and deaths of despair during the last wave of the opioid epidemic that is driven by alcohol-related deaths. We estimate a positive but small increase in deaths from alcoholic liver diseases and cirrhosis significant only at the 10% level and no effect on suicides, see Table 2.6. The category alcoholic liver diseases includes causes of deaths such as hepatitis and related conditions, that may be directly affected by opioid use (Ruhm, 2021), so it is possible this small effect is directly driven by opioid use. Furthermore, the fact that we do not find an effect on suicides (Panel a), suggests that first, there are no pre-trends between our instrument and this measure of despair, and second, that the marketing of OxyContin did not trigger further despair, when measured as suicides. Similarly, we do not find evidence of changes in smoking rates (Figure A.9).

Beneficiaries of social insurance and welfare programs. Addiction to and misuse of prescription opioids could deteriorate one's health, reduce one's productive capacity and put one at risk of permanently reducing ones attachment to the economy. We document a tight link between the opioid epidemic and an increase in disability-programs beneficiaries. These

results are presented in Table 2.7. We find positive and significant effects for measures of both disability programs. A change from the 25th to the 75th percentile in the growth of opioids per capita caused a 38% increase in the share of the population receiving SSI and a 76% increase in the share receiving SSDI.[43] We do not have data for SSI or SSDI claims in the pre-period, but we can estimate our reduced-form event-study after the introduction of OxyContin. Figure A.16 shows the strong positive pattern between mid-nineties cancer mortality and disability claims.

SNAP is designed to act as a safety net for low-income families. In our context, the share of SNAP beneficiaries in the population is a useful proxy for deteriorating economic conditions. We find a positive effect on the share of SNAP beneficiaries: Our estimates suggest that a change from the 25th to the 75th percentile in the growth of oxycodone per capita caused a 57% increase in the share of the population enrolled in SNAP. This effect is comparable to an increase of 3.78 percentage points in the unemployment rate (Ganong and Liebman, 2018).[44] Figure 2.7 shows the dynamic evolution of these effects. These results point to a substantial worsening of economic conditions. The effects we observe on SSI, SSDI, and SNAP are particularly strong during the third wave of the epidemic, when the incidence of illicit drug use, such as of heroin and fentanyl, increased (Table A.5).

Fertility and birth outcomes. An important feature of the crisis is that it has primarily impacted people in early adulthood through mid-life with potential costs beyond the generation directly affected by it. One in five pregnant women filled a prescription for opioids from 2000 to 2007 (Desai et al., 2014); and between 2008 and 2012, 39% of women of reproductive age covered by Medicaid obtained a prescription for opioids. These figures, joint with the

[43] SSDI uses 1996 data as the baseline year, and SSI uses 1998 as the baseline year.

[44] The receipt of benefits from multiple programs is not uncommon. SNAP administrative data from 2011 indicate that 20% of SNAP households received SSI benefits and 22% received Social Security benefits (see, for example, Strayer et al., 2012). We claim that our estimated effect on SNAP applications cannot be entirely driven by dual applicants. Under the assumption that 20% of SNAP recipients are also SSI recipients, the lower bound for the effect on SNAP recipiency rate is 15.6% (0.20×78). Our estimated effect is well above this figure, suggesting that the average effect on SNAP applications is also driven by low-income workers.

staggering increase in the incidence of neonatal abstinence syndrome (Patrick et al., 2015) raise concerns about the risks and consequences of opioid misuse in this population.[45] Terplan et al. (2015) document that the higher rates of unwanted pregnancies in the population of women who take opioids is mostly driven by the lack of adherence to contraception. Additionally, opioid use early in pregnancy, often before women know they are pregnant, can increase the risk for some birth defects and other poor pregnancy outcomes, such as preterm birth or low birth weight (Ailes et al., 2015).

We find that a 25th-to-75th percentile increase in opioids increases fertility by 10% (Table 2.9). This result is entirely explained by non-marital births as we can see in column 1 of Appendix Table A.6.[46] By age, we find that most of the increase in fertility is coming from women 25 to 29 years old, which compensates a decline in fertility for those over 35 years old. To the best of our knowledge, we are the first to document a causal rise in fertility as a result of the opioid epidemic. This is in line with cross-sectional estimates of the higher risk of unintended pregnancies for women with opioid use disorder (Stone et al., 2020), and with work that documents access barriers to contraceptives for women with substance use disorders (Rinehart et al., 2021).

Regarding birth outcomes, we find evidence that a 25th-to-75th-percentile increase in the supply of prescription opioids decreases birth weight by 0.7%, deteriorates APGAR scores by 1% relative to its mean value, and reduces median pregnancy duration by 0.63% which translates to a reduction in the median length of pregnancy of 0.24 weeks. We also estimate increases in the incidence of preterm births and the share of low-weight births, but these are not statistically significant. We find an increase in the APGAR score of infants who died in the first year, which means that healthier infants died. However, in aggregate terms we do

[45] Neonatal abstinence syndrome is a result of the sudden discontinuation of fetal exposure to medicine or drugs that were used or misused by the mother during pregnancy.

[46] We compute non-marital fertility as the ratio between births to unmarried mothers and the female population aged 15 to 44 years old. We do this, as data on the population of unmarried females does not exist at the year–commuting-zone level.

not find any increase in the infant mortality rate.

Our estimated declines in birth weight are not small in magnitude. For a reference, Almond, Hoynes and Schanzenbach (2011) estimate an increase in birth weight of 0.5% as a result of the roll-out of food stamps among participants, i.e., a treatment on the treated estimates.[47] Hoynes, Miller and Simon (2015) find a 0.3% increase in birth weight from the expansion of the Earned Income Tax Credit (EITC). This is particularly relevant in light of evidence on the importance of birth weight and health at birth for future health, schooling, and earnings (Behrman and Rosenzweig, 2004).

In summary, our results suggest that the opioid epidemic lead to important increases in fertility, driven by young and unmarried mothers. While not affecting directly the infant mortality rate, the epidemic worsened birth outcomes through reductions in pregnancy duration and infant health at birth. In 24 states and the District of Columbia, the use of any illegal substance during pregnancy constitutes child abuse, and can lead to foster care placement. Nonetheless, Eichmeyer and Kent (2021) document that treatment for opioid use disorder increases in the year after childbirth, and that the timing of this increase is consistent with pregnancy triggering treatment for a pre-existing disorder. Using, state-level data, Buckles, Evans and Lieber (2022) document that greater exposure to the crisis increases the likelihood that a child's mother or father is absent from the household and it increases the likelihood that he or she lives in a household headed by a grandparent. Unfortunately, after multiple efforts we were not able to access foster care records with county or commuting zone identifiers.[48] Future work is needed to quantify the effect of the opioid crisis on foster care placements, and to assess the future outcomes for these children.

[47] Estimates are higher for black participants: between 0.4% and 1.4%.

[48] The Adoption and Foster Care Analysis and Reporting System (AFCARS) provides case level data, but county identifiers are only available for counties with more than a 1,000 cases.

5.3 Complier Analysis

Our instrumental variable estimates identify the causal effects of the increased supply of opioids for commuting zones where marketing was more aggressive because of higher cancer mortality in the mid-nineties. Variation in underlying characteristics across commuting zones could have made such opioids' marketing more or less successful. In Table 2.10 we characterize compliers based on observable characteristics before OxyContin's launch.

First, we assess the strength of the first stage in different sub-samples. Column 1 shows that the positive and strong relationship between cancer rates in the mid-1990s and the change in opioids per capita is present in all the considered sub-samples. Nonetheless, there are important differences on the strength of this relationship. The first-stage is stronger in commuting zones with higher poverty rates, employment in mining, and cocaine and alcohol mortality rates, however, these difference are not statistically significant. Next, we follow Abadie (2003), to recover the fraction of compliers for different characteristics. We find that commuting zones with higher poverty rates, lower educational attainment, and a higher share of mining employment are more likely to be compliers. In terms of health outcomes, areas with a higher number of primary care physicians (PCP) per capita are overrepresented among the compliers, as is the case for places with higher smoking rates and higher cocaine and alcohol mortality rates.

6. Robustness Checks

In this section, we explore alternative explanations for our findings and test the robustness of our results. We start by presenting alternative specifications of the first stage and then test the robustness of the main results.

6.1 First Stage

Additional demographic controls. A potential concern with our choice of instrument is that mid-nineties cancer mortality may be capturing demographic variation along the age distribution. Our baseline regression already controls for the change in the share of the population over 65, but our instrument is expressed in levels, so some of this variation may still be important. We directly test this by including the share of population over 65, the size of the population over 65, and total population as additional control variables. Table A.7 shows the results of this exercise. We find that the first stage regression is as strong as in our baseline regression.

Alternative measures of our instrument. Additionally, we test the robustness of the first stage to alternative choices of instruments. Column 1 of Table A.8, replicates the first stage with age-adjusted cancer mortality, we find a very similar and strong first-stage estimates. We also test whether the relationship between future opioid distribution and mid nineties cancer mortality is sensitive to any of the years used as our baseline cancer mortality measure. Columns two to four show there is a strong first-stage for 1994, 1995 and 1996 cancer mortality. In column five we estimate a population weighted regression and find similar results.[49] As an additional robustness check, in Panel b we construct a measure of cancer mortality that exclude deaths from lung cancer, this measure is less likely to be driven by behavioral and environmental factors that could correlate with our outcome variables. As with other alternative instruments, in column one, we find a strong first stage coefficient. In columns two to four we use mid-nineties cancer mortality for those over 55, 65 and 75 respectively, and find a positive and statistically significant first stage in all cases. Finally, we inspect the dynamic relation between mid-nineties cancer mortality and opioid supply. Panel (a) of Figure A.19 presents the dynamic first-stage using cancer mortality in 1994 as

[49]Figure A.18 shows the results for our main event-study specification with population weights and the results are very similar.

the instrument and the remaining panels in this Figure show the dynamic reduced-form for the main outcomes of interest. The pattern of the estimated coefficients is equivalent to the one we obtain using the cancer 1994-1996 as instrument. Additionally, we do not find evidence of pre-trends in the relation between opioid mortality—and the main outcomes of interest—and mid-nineties cancer mortality in this alternative specification.

Mechanical Effects. A potential threat to our hypothesis is that commuting zones with higher cancer mortality would see a larger uptake of opioids from an innovation in the pain medication market, even in the absence of marketing efforts. Several facts suggest this is not the mechanism at play. First, cancer patients had access to potent opioids before the launch of OxyContin, as this was standard pain management practice. For these patients, the introduction of OxyContin represented a switch from MS Contin—the gold standard to treat cancer pain—to OxyContin. Second, our results suggest there is no evidence of misuse of opioids in the population most affected by cancer, as we find no increases in deaths from prescriptions and non prescription opioids for those over 55 years of age (Figure 2.9).

Alternative instrumental variable approach. Our instrumental variable approach is similar in spirit to a shift-share instrument. In this research design, the shares measure differential exposure to common shocks and identification is based on its exogeneity (Goldsmith-Pinkham, Sorkin and Swift, 2020). In our application, the shares are cancer rates in the mid-1990s, which capture exposure to the marketing of prescription opioids, and the shift is the national growth of Purdue Pharma's marketing or the growth in the supply of prescription opioids. Our preferred specification uses as an instrument cancer mortality before the launch of OxyContin, which highlights the fact that our only source of exogenous variation corresponds to the shares. In Appendix Table A.9, we show results using the shift-share instrument. To construct this instrument, we use the national growth rate of prescription opioids as the shift component. The results are quantitatively indistinguishable from our preferred specification. As Goldsmith-Pinkham, Sorkin and Swift (2020) point out, using a

Bartik instrument is "equivalent" to exploiting the shares as an instrument. This is because the temporal variation induced by the growth of prescription opioids is mostly absorbed by the time dimension of our state times year fixed effects.

Additional sample restrictions. We test whether the positive relationship in our first stage is driven by a given state or a group of states. Figure A.20 presents the estimate of the first stage coefficient restricting the sample to (i) all non-triplicate states, (ii) only triplicate states, and (iii) to the exclusion of all states, one at the time. We find that the positive relationship between mid-nineties cancer mortality and the supply of opioids is present in both triplicate and non-triplicate states, and is robust to the exclusion of any given state. Furthermore, the first stage is stronger in the five triplicate states defined in Alpert et al. (2022), which would be consistent with a story in which pharmaceutical companies need to be more strategic in promoting opioids in places where they face additional barriers.[50]

Alternative measures of prescription opioids. Even though Purdue Pharma was the pioneer in the use of opioids in the non-cancer pain market with OxyContin, many pharmaceutical companies—Janssen, Endo, Cephalon-Teva, Actavis, Insys, and Mallinckrodt—promoted their prescription opioids beyond the cancer market following Purdue's leadership. To asses whether the positive relationship that we observe between the total level of prescription opioids and mid-nineties cancer mortality is present across different opioids' categories, in Panel C of Table A.8, we split opioids between oxycodone—which captures the supply of Oxycontin—and the rest of opioids. We find there is a strong positive first stage for both categories.

[50]The first stage coefficient is 1.542 in the five triplicate states defined in Alpert et al. (2022), and 0.917 in the other states. This difference is not statistically different and the p-value of such test is 0.203.

6.2 Placebo checks

Are other mid-1990s mortality rates predictive of future prescription opioids per capita distribution? Our identification strategy connects mid-1990s cancer mortality to future growth in the supply of prescription opioids through the targeted marketing of Purdue Pharma. This implies that we can test the validity of our design by estimating first-stage regressions for placebo instruments—i.e., mid-1990s mortality from causes unrelated to cancer. Finding a good placebo instrument is challenging, given that the causes that underlie the incidence of cancer and other conditions, such as heart disease are not independent (Chiang, 1991 and Honoré and Lleras-Muney, 2006). As a result, there is substantial overlap across underlying causes and the correlation across measures is very high. With this caveat, in Table A.10 we show placebo instrument regressions for three mortality rates that are less likely to be affected by the previous concern: Cerebrovascular disease (CVD), transit accidents, and homicide.[51] We find that none of these measures predict future distribution of opioids (Columns 1 to 3) or change the predicted power of our instrument (Columns 4 to 6).

6.3 Alternative Definitions: Opioid Mortality and Opioid Supply

Drug overdose deaths can be hard to categorize, specially when using data that spans more than one version of the ICD codes. We construct an additional outcome measure for opioid mortality and present the results using this measure in Table A.11. This measure has the advantage that comparisons across years are less affected by changes in the ICD classification, but this comes at the cost of including a broader set of drugs as the cause of deaths.[52] Exploiting this measure, we arrive at similar conclusions: An additional dose of opioids per

[51] A good candidate for this placebo check is mortality from external causes of deaths. External causes are defined as intentional and unintentional injury and poisoning (including drug overdose). From this category, we construct measures of mortality that do not include any of our outcome measures: accidental poisoning and suicide.

[52] Drug-induced deaths category includes deaths from poisoning and medical conditions caused by the use of legal or illegal drugs, as well as deaths from poisoning due to medically prescribed and other drugs.

capita caused an increase in the drug-induced mortality rate of 0.0112 points. An increase from the 25th to the 75th percentile of prescription opioids per capita increases drug-induced mortality by 47%.

As an additional check, we use data only on Oxycodone as an alternative measure of opioid supply. We find a positive relationship between cancer mortality rates and this measure of opioid supply. In Table A.12, columns (2) and (3) we estimate that an additional dose of oxycodone per capita caused an increase in prescription opioid mortality of 91% and in all opioid mortality of 40%.

6.4 Alternative Sample Restrictions and Specifications

In our main specification, we restrict our sample to areas with more than 25,000 residents, which represents 99.8% of all opioid deaths and 99.3% of the total population. In Table A.13 we reproduce our analysis using alternative restrictions on the size of commuting zones. We arrive at analogous conclusions to the main analysis; there is a strong and positive relation between mid-nineties cancer mortality and supply of prescription opioids which translates to (i) increases in opioid-related mortality, and (ii) deteriorating economic conditions and health outcomes.

SNAP benefit recipiency rates at the commuting-zone level required imputations for some commuting zones with no available data at the local level. Table A.14 shows the result for the sample of commuting zones that do not require state-level imputation. Our results are not sensitive to this sample restriction.

Finally, in Table A.15 we expand the set of controls in our regression to include either the unemployment rate or the employment rate and we find our results are quantitatively indistinguishable.

6.5 Trade shocks & the 2001 Economic Recession

During our period of study, the US experienced significant economic changes that affected communities differentially. In October, 2000, the US Congress passed a bill granting permanent normal trade relations (PNTR) with China. This trade liberalization impact on communities is a function of the importance of the manufacturing industries for local employment, especially in industries subjected to import competition from China. Researchers have estimated the impact of this trade shock on a host of outcomes. Regions more exposed to Chinese import competition experienced relatively larger declines in employment and a greater uptake of social welfare programs (Autor and Dorn, 2013). Additionally, areas more exposed to Chinese import competition exhibit relative increases in fatal drug overdoses (Pierce and Schott, 2020).

In light of this evidence, we ask whether our results are confounded or mediated by this trade policy. To answer this, we follow the trade literature to construct two alternative measures of exposure to PNTR and then estimate our first-stage and reduced-form models controlling for these exposure measures (Pierce and Schott, 2020 and Autor and Dorn, 2013). Table A.16 columns two to four reproduce the first stage when we control for exposure to Chinese import competition. We find the results are unaffected by the inclusion of these variables. Figure A.21 replicates our main results adding the china shock measures to our event-study specification. Here as well we find our estimates do not change with this exercise. This is the result of a very low correlation between our instrument and the exposure to Chinese import competition.

The timing of some of our results overlaps with the 2001 economic recession. To asses whether the recession is mediating some of our effects, we construct a measure of exposure to the recession as the change in the unemployment rate from 2001 to 2000 at the commuting zone. Similar to the china shock, we find that our instrument and this exposure measure have

a very low correlation level (ρ=0.03), and our first stage estimate are completely unaffected (column 1 in Table A.16). More broadly, in the last three columns of Table A.16 we add controls for the unemployment rate in years 1994 to 1996, respectively, and find that our estimates do not change.

7. Policy Implications and Conclusions

This paper studies the origins and effects of the opioid epidemic, one of the most tragic episodes in recent history. To do so, we uncover novel geographical variation in the initial promotion of OxyContin that targeted the cancer patients market. We document that this initial targeting had long-term effects on the supply of prescription opioids, overdose deaths involving opioids, and a deterioration in adult wellbeing measured by the demand for SSDI, SSI, and SNAP. These effects will continue to unfold as a result of the increase in non-marital fertility and the worsening of birth outcomes.

In this paper, we sought to provide a comprehensive picture of the effects of the opioid epidemic. However, data access limitations have prevented us from speaking about important topics, such as the effects on children's well-being, foster care referrals, and the demand for and use of healthcare. We hope that future research will shed light on these subjects.

In terms of policy recommendations, we want to highlight how complex and far-reaching the effects of the opioid epidemic are and how this calls for a coordinated response from multiple policy angles. Monitoring, limiting, and restricting access to prescription opioids, which has been the main policy response, is essential, but it falls short for the needs of the affected population. Increasing access to rehabilitation treatments and programs aimed at reincorporating parents and workers into their lives should be the center of this response.

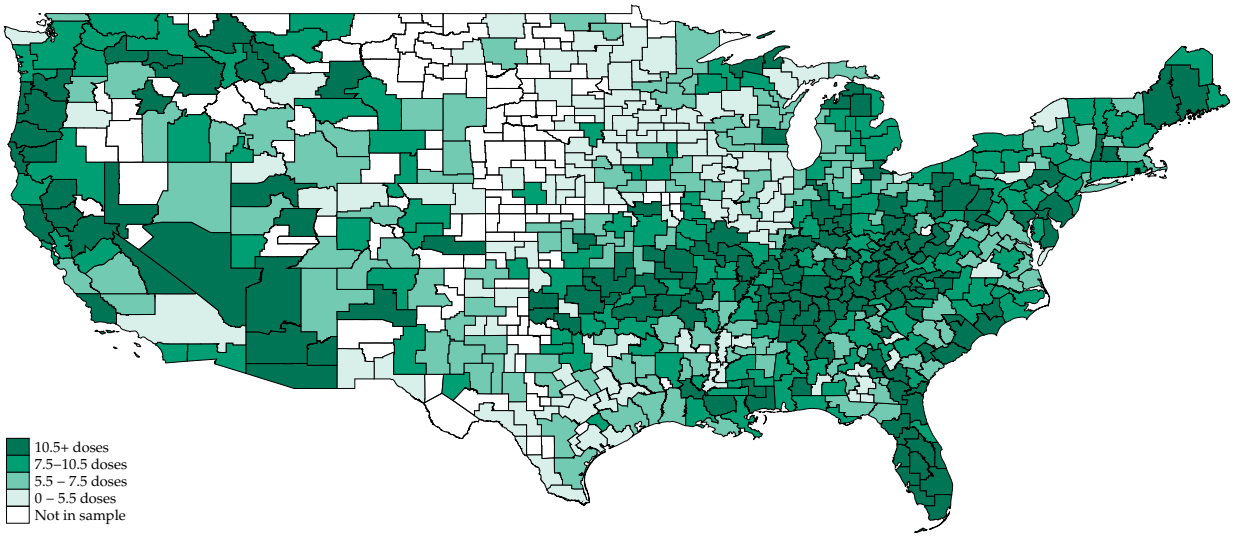
Finally, our results have direct policy implications regarding the desirability of promotional efforts of addictive drugs by pharmaceutical companies that target physicians, phar-

macies, and patients. We document the devastating consequences of aggressive and deceitful marketing of addictive drugs. Although marketing has expanded over the 25 years since the introduction of OxyContin, regulatory oversight remains relatively limited.[53] Some regulatory initiatives constitute small steps in the right direction, however, most of these initiatives are concerned with the rising costs of prescription drugs and not with the risks of abuse and addiction. More can be done to restrict the pharmaceutical promotion that carries this risk.

[53]Currently, prescription drug marketing practices in the US include direct-to-consumer and professional branded advertising, detailing visits, free drug samples, and direct physician and hospital payments (e.g., speaker fees, food, travel accommodations).

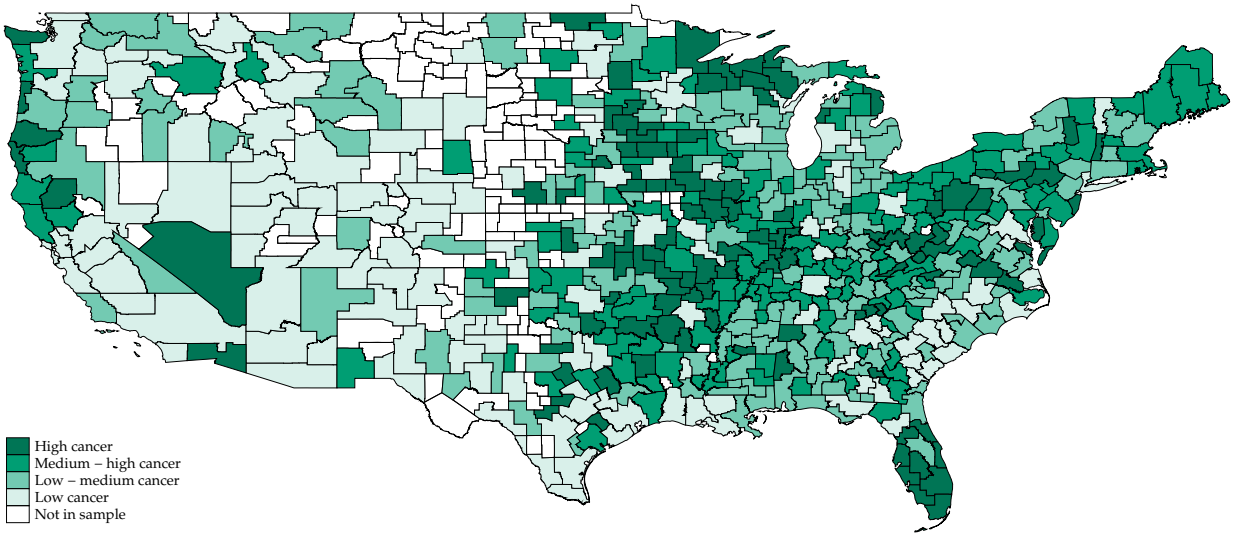
8. Figures

Figure 2.1: Prescription Opioids Distribution at the Peak of the Epidemic (2010).



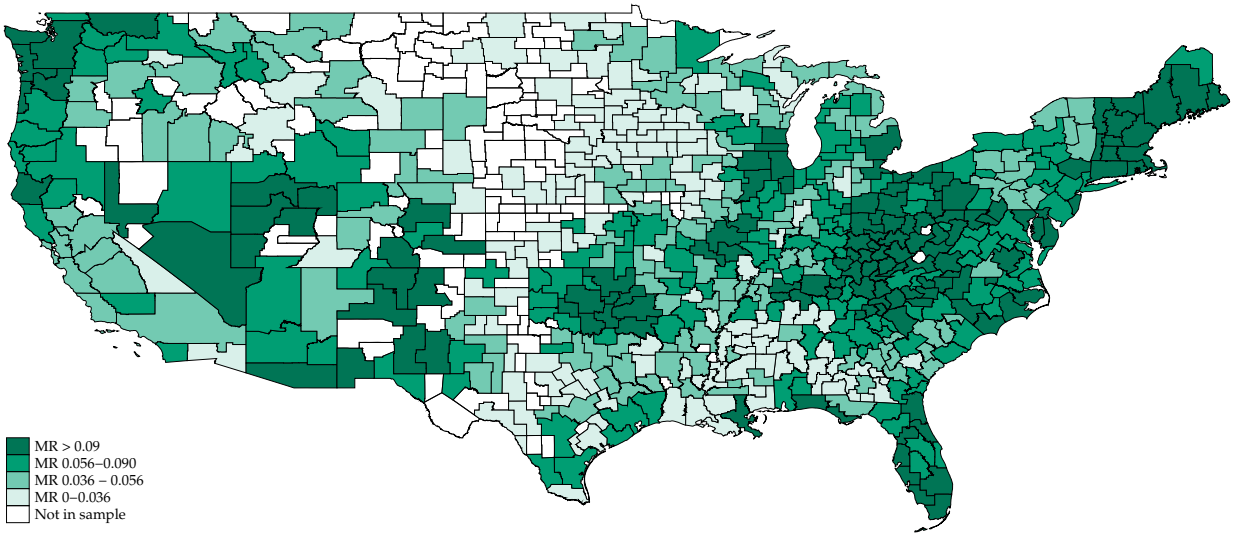
Notes: This map shows the distribution of prescription opioids at the commuting zone level in 2010, the year when the distribution of prescription opioids peaked as shown in Figure 2.4. Lighter shades indicate commuting zones with a lower prescription-opioid supply and darker shades indicate commuting zones with a higher prescription-opioid supply. Each group corresponds to one quartile of the prescription opioids distribution; i.e., each color accumulates 25% of the mass of this distribution. Commuting zones included in our sample represent 99.8% of all opioid deaths and 99.3% of the total population. This figure is referenced in Section 3.3.1

Figure 2.2: Distribution of Cancer Mortality Rates Before the OxyContin's Launch.



Notes: This map shows the cancer mortality rate at the commuting-zone level for the year 1994 - 1996, before OxyContin was introduced to the market. Lighter shades indicate commuting zones with lower cancer prevalence, while darker shades indicate commuting zones with higher cancer prevalence. Each group corresponds to one quartile of the cancer mortality distribution; i.e., each color accumulates 25% of the mass of this distribution. Commuting zones included in our sample represent 99.8% of all opioid deaths and 99.3% of the total population.

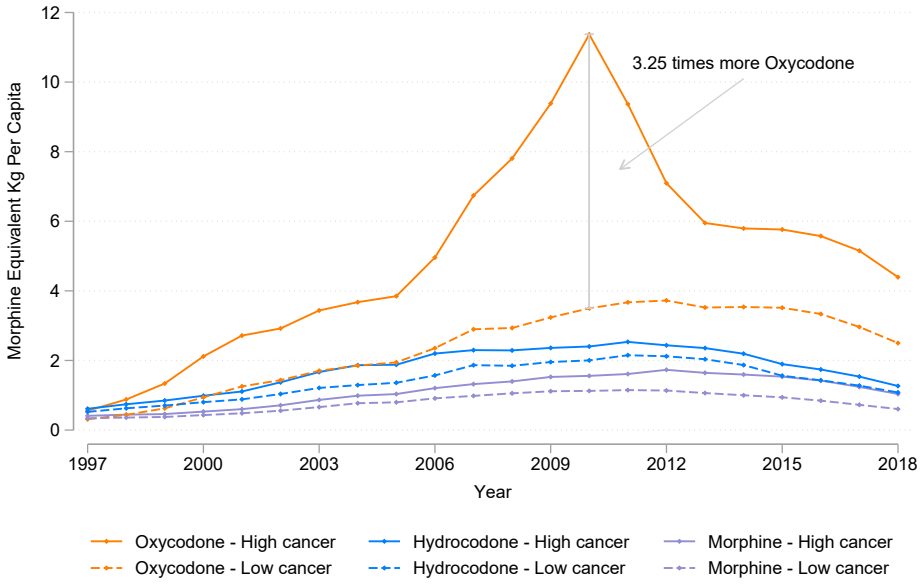
Figure 2.3: Prescriptions Opioid Mortality Rate 1999 - 2018



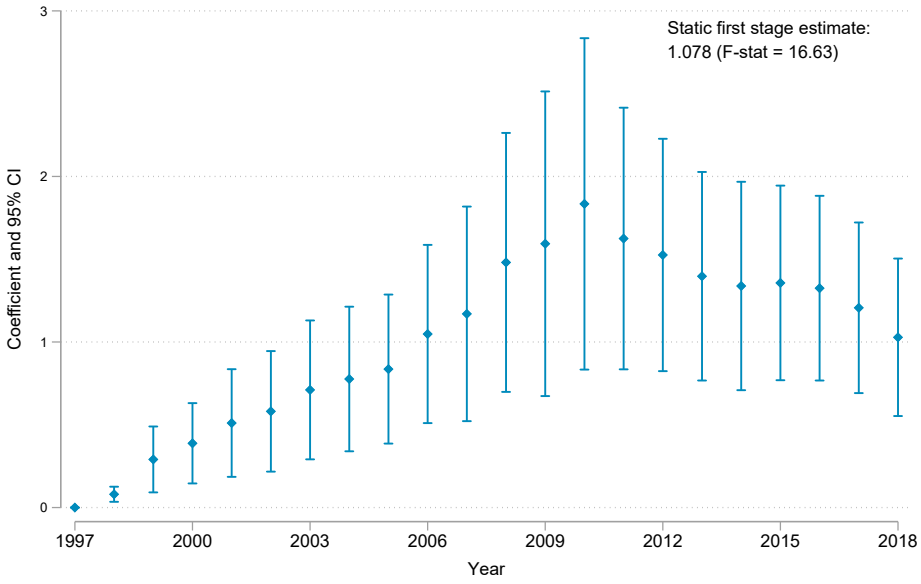
Notes: This map shows the distribution of prescription opioid mortality at the commuting zone level for the period 1999 - 2018. Lighter shades indicate commuting zones with lower opioid mortality, while darker shades indicate commuting zones with higher opioid mortality. Each group corresponds to one quartile of the opioid mortality distribution; i.e., each color accumulates 25% of the mass of this distribution. Commuting zones included in our sample represent 99.8% of all opioid deaths and 99.3% of the total population.

Figure 2.4: Prescription Opioids Distribution by Mid-nineties Cancer Prevalence

(a) Trends in High versus Low Cancer Mortality CZs

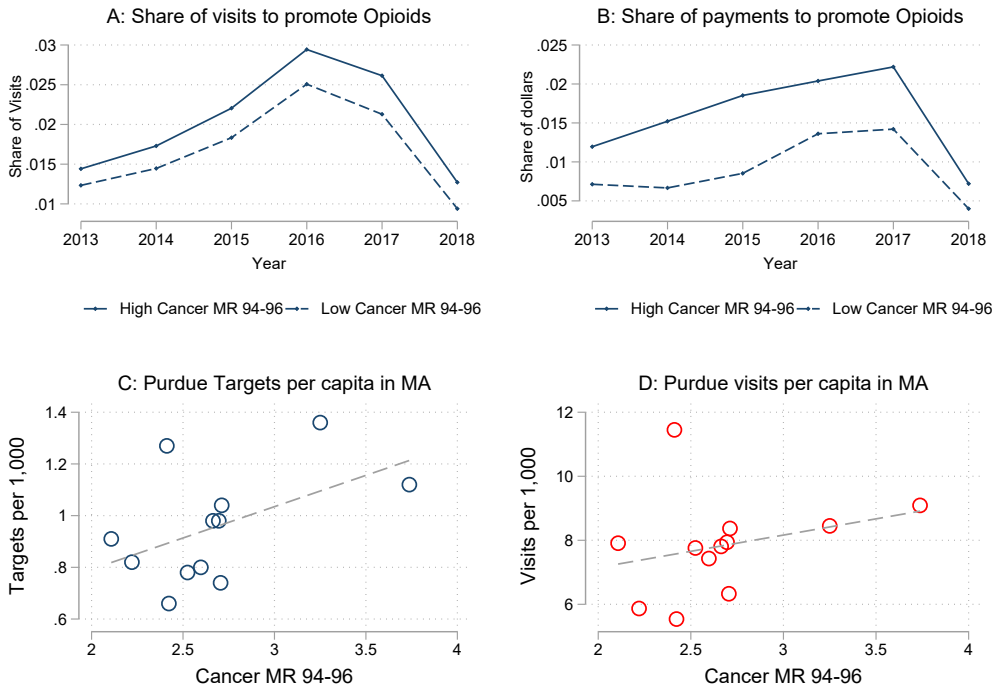


(b) Dynamic First Stage



Notes: Panel (a) shows the evolution oxycodone, hydrocodone, and morphine in commuting zones in the bottom (dashed lines) and top (solid lines) quartiles of cancer mortality before the launch of OxyContin. Oxycodone is OxyContin's active ingredient. Between 1997 and 2010, areas in the highest quartile of cancer incidence saw an increase in oxycodone grams per capita of 2,900%, while areas in the lowest quartile experienced a growth that was one-third that. All prescription opioids and oxycodone are measured in morphine-equivalent doses. Panel (b) shows estimates of the coefficients of the dynamic first stage.

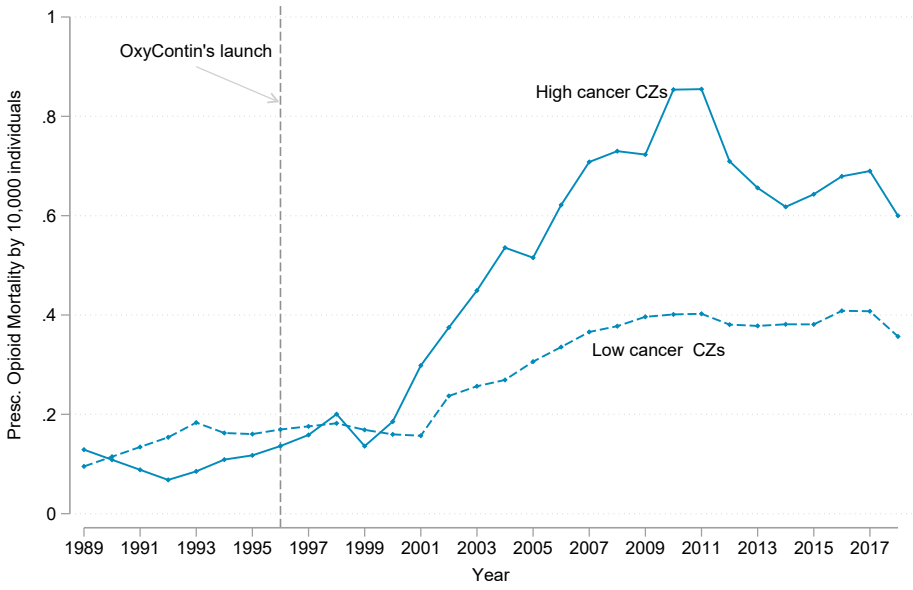
Figure 2.5: Opioids Marketing and Mid-nineties Cancer Mortality



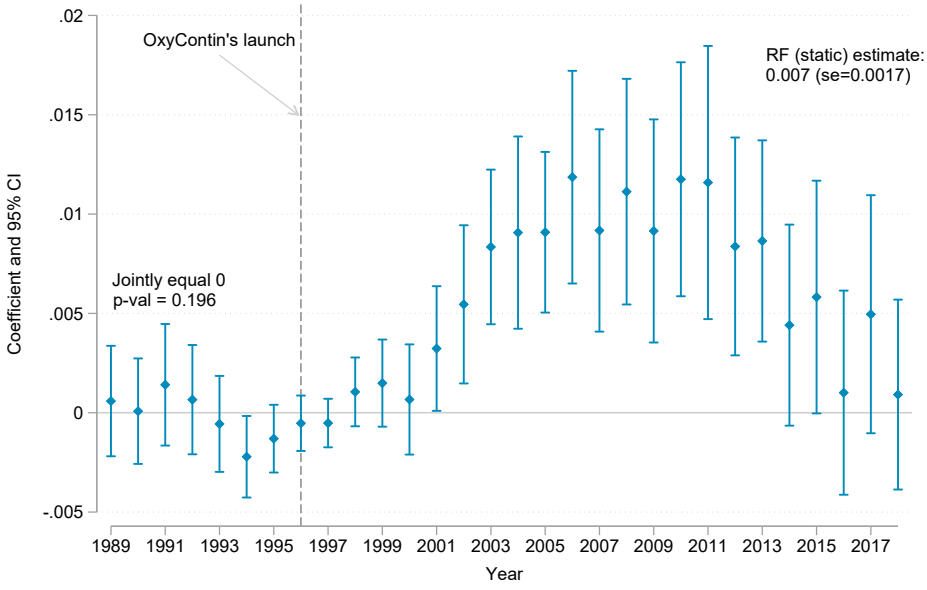
Notes: Panels A and B use data from the CMS Open Payments. High cancer corresponds to the top quartile of cancer incidence in 1994-1996, and low cancer to the bottom quartile. Panels C and D use digitized data from "Exhibit 1 - Sales Visits By Purdue In Massachusetts. COMMONWEALTH OF MASSACHUSETTS v. PURDUE PHARMA C.A. No. 1884-cv-01808"(Figure A.8) to construct county-level averages.

Figure 2.6: Effects of Mid-nineties Cancer-market Targeting on Prescription Opioid Mortality

(a) High versus Low Cancer Mortality CZs

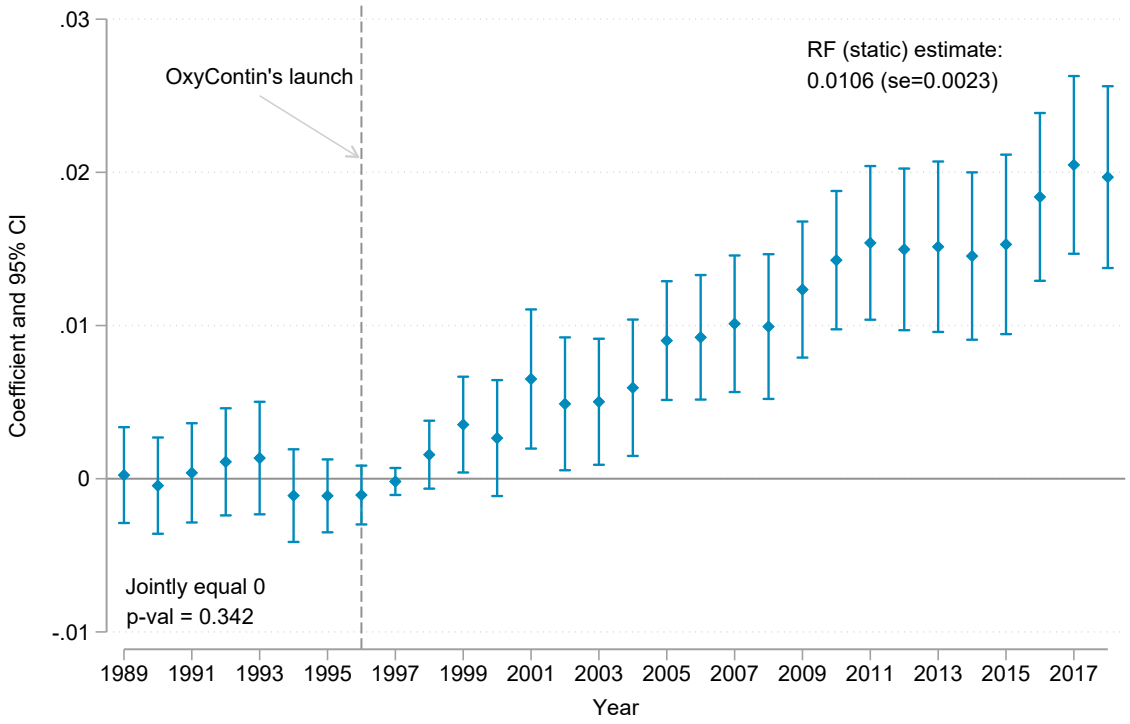


(b) Reduced Form - Event Study Approach



Notes: This figure shows the effects of the increase in prescription opioid supply in prescription opioid mortality. Panel (a) shows the raw data, early in the 2000s, a wedge starts to appear between high- and low-cancer-incidence groups, and by 2018 prescription opioid mortality in high-cancer areas is 75% higher. Panel (b) shows the dynamic reduced-form estimation. We regress prescription opioid mortality on a set of year-dummy variables interacted with our instrument—cancer mortality in 1994-1996.

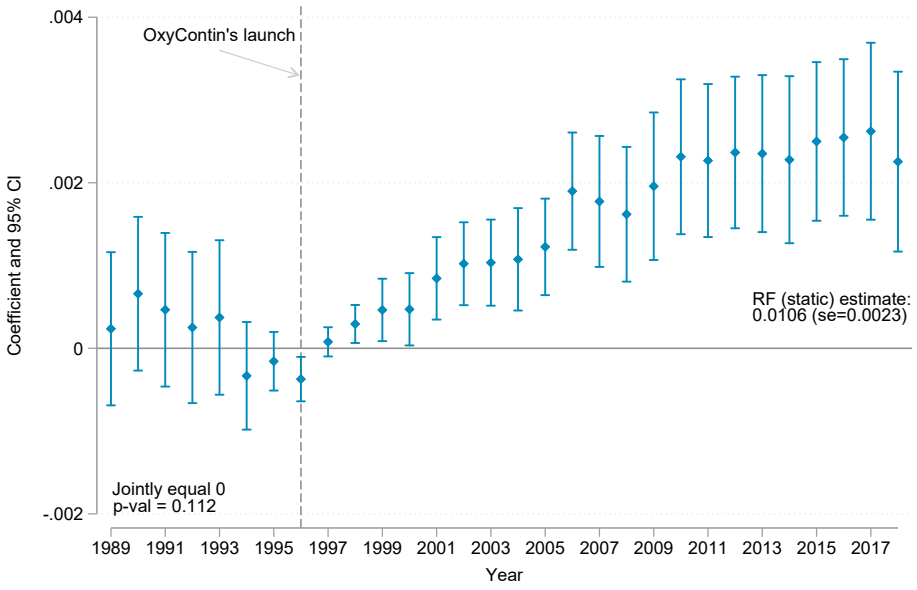
Figure 2.7: Effects of Mid-nineties Cancer-market Targeting on the Sh. of SNAP Recipients



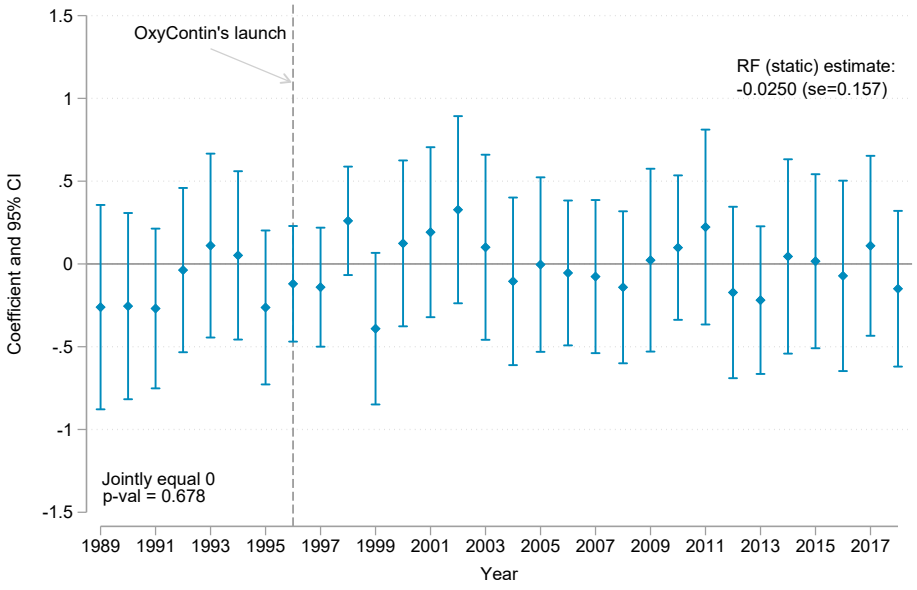
Notes: This figure shows the effects of the increase in prescription opioid supply on SNAP recipients per capita. We present the results of a dynamic reduced-form estimation were we regress SNAP claims per capita on a set of year-dummy variables interacted with our instrument—cancer mortality in 1994-1996. These coefficients corresponds to the estimate of ϕ_t in Equation 2.3. We use this specification to test for the presence of pre-trends in the relation between SNAP claims and mid-nineties cancer mortality; we do not reject the null hypothesis that the estimated coefficients before 1996 are jointly equal to zero, the p value of this test equals 0.6539.

Figure 2.8: Effects of Mid-nineties Cancer-market Targeting on Fertility Rates and Birth Outcomes

(a) Non-marital Fertility Rate



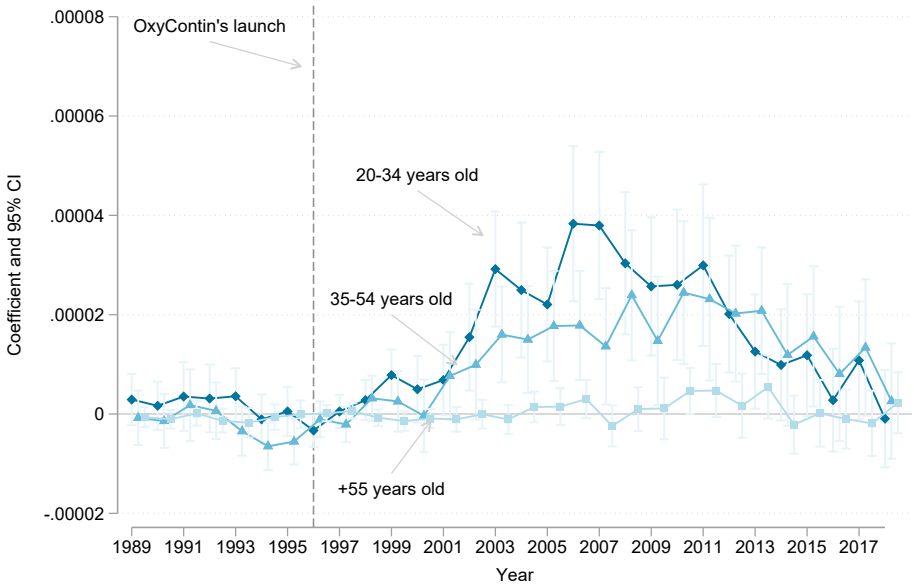
(b) Infant Mortality Rate



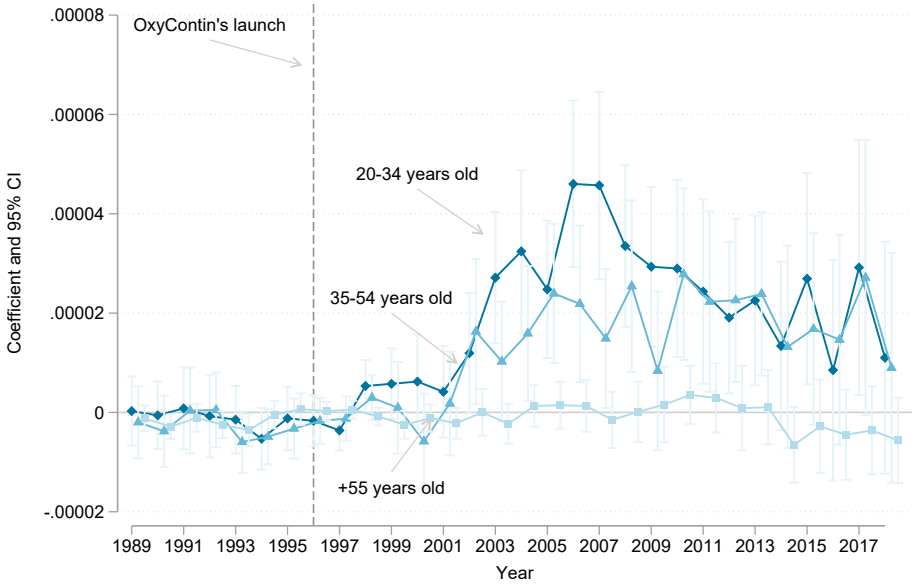
Notes: This figure shows the effects of the increase in prescription opioid supply in the fertility rate of unmarried women (panel a) and in infant mortality rate (panel b). We present the results of a dynamic reduced-form estimation were we regress these outcomes on a set of year-dummy variables interacted with our instrument—cancer mortality in 1994-1996. These coefficients corresponds to the estimate of ϕ_t in Equation 2.3.

Figure 2.9: Effects of Mid-nineties Cancer-market Targeting on Opioid Mortality by Age

(a) Prescription Opioids



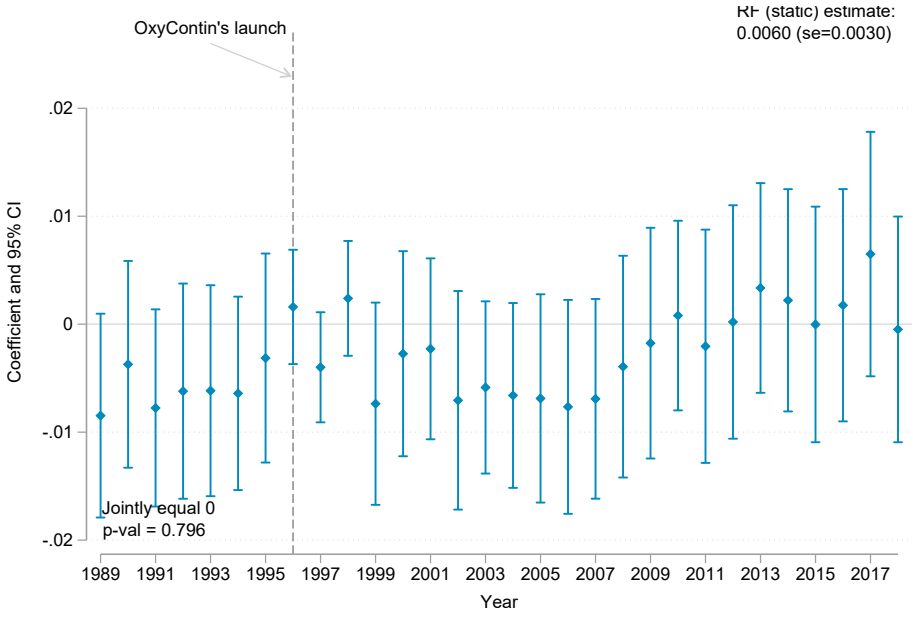
(b) Any Opioids



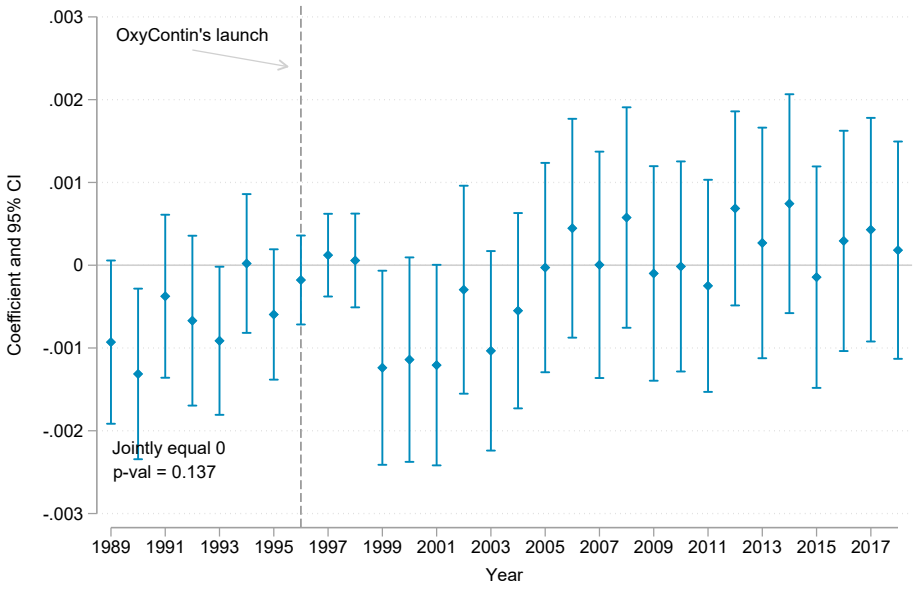
Notes: This figure shows the effects of the increase in prescription opioid supply in opioid related mortality by age group. We present the results of a dynamic reduced-form estimation were we regress these outcomes on a set of year-dummy variables interacted with our instrument—cancer mortality in 1994-1996. These coefficients corresponds to the estimate of ϕ_t in Equation 2.3.

Figure 2.10: Trends on Despair and Overall Health

(a) Suicide Mortality Rate



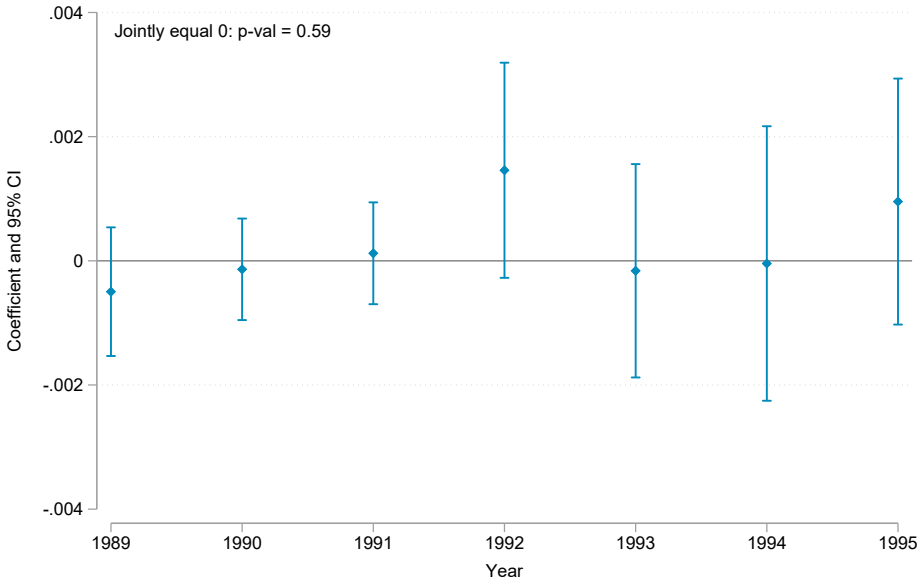
(b) Non-cancer Mortality for Adults Aged +75 Years Old.



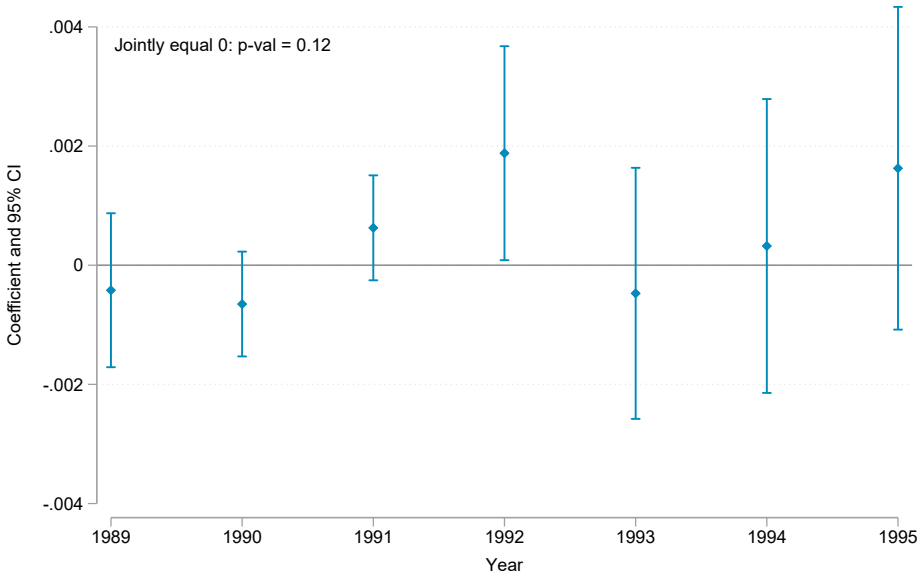
Notes: This figure shows the dynamic reduced-form relationship between suicide mortality rate (panel a) and mortality of 75-years-old and older adults (panel b) and our instrument. That is, the figure presents the results of a dynamic reduced-form estimation were we regress these outcomes on a set of year-dummy variables interacted with our instrument—cancer mortality in 1994-1996.

Figure 2.11: Robustness Check: Dynamic Reduced Form for Out-of-sample Opioid-Mortality

(a) Prescription Opioids Mortality



(b) All Opioid Mortality



Notes: This figure shows the dynamic reduced-form relationship between outcomes of interest and our instrument in a out-of-sample period. That is, we replicate our dynamic reduced-form analysis in the pre-OxyContin period. We regress each outcome on a set of year-dummy variables interacted with the out-of-sample instrument—cancer mortality in 1989 - 1990.

9. Tables

Table 2.1: Summary Statistics for 1999-2018

	Mean	Median	SD	Min	Max	Obs.
	(1)	(2)	(3)	(4)	(5)	(6)
Opioid Prescriptions: Doses per capita						
All Prescription Opioids	6.42	5.48	4.32	0.00	57.65	11,800
Oxycodone	3.15	2.52	2.60	0.00	51.31	11,800
Hydrocodone	1.93	1.55	1.50	0.00	16.66	11,800
Morphine	0.94	0.77	0.69	0.00	10.67	11,800
Cancer Mortality per 1,000						
Cancer mortality rate 1994-1996	2.53	2.53	0.58	0.12	6.24	590
Cancer mortality rate	2.48	2.49	0.55	0.59	4.75	11,800
Opioid-related Mortality per 1,000						
Prescription opioids	0.04	0.03	0.05	0.00	1.06	11,800
Any opioids	0.07	0.05	0.07	0.00	1.22	11,800
Other Mortality Measures per 1,000						
All-cause mortality (+20 years old)	9.87	9.93	2.06	2.79	20.92	11,800
Deaths of despair	0.27	0.25	0.10	0.00	1.17	11,800
Alcoholic liver diseases and cirrhosis	0.12	0.11	0.06	0.00	0.63	11,800
Suicide	0.15	0.14	0.06	0.00	0.48	11,800
Demand for Social Services						
Share SSI	0.04	0.03	0.02	0.00	0.30	11,800
Share SSDI	0.05	0.04	0.02	0.01	0.16	11,800
Share SNAP	0.12	0.11	0.07	0.00	1.20	11,800
Infant and Fertility Outcomes						
Infant MR (per 1,000 births)	6.86	6.54	2.87	0.00	30.61	11,800
Birth weight	3,274.25	3,276.53	79.47	2,930.28	3,569.76	11,800
Share low birth weight	0.08	0.08	0.02	0.02	0.20	11,800
Share preterm	0.12	0.12	0.03	0.05	0.62	11,800
APGAR score - all infants	8.82	8.84	0.19	5.00	10.00	11,800
APGAR score - dead infants	5.62	6.00	2.28	0.00	10.00	11,460
Median gestation	38.95	39.00	0.24	35.00	40.00	11,800
Fertility rate	0.08	0.08	0.01	0.04	0.19	11,800
Fertility rate 25-29	0.13	0.12	0.02	0.05	0.27	11,800
Fertility rate - unmarried women	0.03	0.03	0.01	0.00	0.09	11,800

Notes: This table presents summary statistics for our main outcomes, measures of the prescription opioid supply, and cancer mortality incidence for the period 1999 - 2018. We leverage data from multiple sources. Prescription drugs distribution data come from the DEA. Data on opioid, cancer, birth, and fertility outcomes come from the NVSS. We use data from the Food and Nutrition Service of the Department of Agriculture and the SSA to construct demand for the Supplemental Nutrition Assistance Program (SNAP), Supplemental Security Income (SSI), and Social Security Disability Insurance (SSDI). This table is referenced in Section 3.

Table 2.2: Determinants of the Opioid Distribution in 2000

Dependent variable: Prescription opioids per capita			
	(1)		(2)
Demographics (in shares)		*Crime (in rates)*	
White	3.526***	Overall	-0.0000622
	[0.961]		[0.0000752]
Hispanic	-3.323***	Violent	0.00160***
	[0.807]		[0.000614]
Female	6.709	*Economic characteristics*	
	[9.973]	Ln income	2.517***
Aged 18-65	21.67***		[0.922]
	[4.348]	Share below poverty line	0.0521
Aged +66	6.211		[0.0625]
	[7.665]	Share employed in manufacturing	-0.0374***
Infants	-100.8*		[0.0105]
	[56.42]	Share with some college education	0.00938
Labor market			[0.0135]
Employment rate	-16.18***	*Health outcomes*	
	[6.031]	Cancer mortality rate	-0.164
Labor Force Participation	-1.805		[0.330]
	[2.493]	Infant mortality rate	-0.0117
Safety net and social insurance			[0.0199]
SSDI	48.45***	Birth weight	0.000336
	[9.821]		[0.00127]
SSI	5.740	Share preterm births	2.330
	[8.944]		[4.796]
SNAP	-1.914	Gestation	-0.200
	[3.848]		[0.396]
		Fertility rate	52.51***
			[14.07]
Mean dependent variable			2.8567
Year			2000
Observations			590

Notes: This table presents estimated coefficients from a cross-section regression of oxycodone distribution per capita on demographic characteristics, labor market outcomes, measures of social assistance demand, crime outcomes, economic characteristics, and health outcomes at the commuting-zone level. Data on economic characteristics come from county-level tabulations of Decennial Census Data. The variable *share with some college* measures the share of the population older than 25 years old who have some education at the college level or higher. Standard errors are robust to heteroskedasticity. *$p<0.10$, **$p<0.05$, *** $p<0.01$.

Table 2.3: First-stage Results

Dependent variable: Prescription opioids per capita					
	(1)	(2)	(3)	(4)	(5)
Cancer MR 94-96	0.960***	1.091***	1.061***	1.132***	1.078***
se	[0.210]	[0.222]	[0.231]	[0.258]	[0.264]
t-stat	4.571	4.914	4.593	4.388	4.083
Effective F-stat	20.894	24.147	21.096	19.254	16.630
Effect size	56.92	64.69	62.91	67.12	63.92
Controls	No	No	No	Yes	Yes
FE	No	State Year	State × Year	State Year	State × Year
Observations	11,800	11,800	11,800	11,800	11,800
Clusters	590	590	590	590	590
Adj. R^2	0.019	0.524	0.559	0.533	0.564

Notes: This table presents estimates of the first-stage equation. The dependent variable is the long change in prescription opioids per capita and it is constructed using a baseline the year 1997—the first year ARCOS data are available. Control variables in long changes are contemporaneous cancer mortality rate, share of population under 1 year old, share of population between 18 and 65, share of population over 66 years, share of Black, White, and Hispanic population, and share of female population. Effect size is computed as the explained changes in doses of prescription opioids per capita from an increase in cancer mortality that would change a commuting zone in the 5th percentile of the cancer distribution to the 95th percentile, relative to change in our period. $t-stat$ corresponds to the $t-statistic$ for the null hypothesis that the coefficient on cancer mortality rate is equal to zero. *Effective F-stat* corresponds to the the effective first-stage F statistic proposed by Montiel Olea and Pflueger (2013). Standard errors are clustered at the CZ level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$. This table is referenced in Section 4.4.1

Table 2.4: Cancer Mortality Rate: Out-of-sample Analysis

	Cancer MR 89-90 (1)		Cancer MR 89-90 (2)
Dependent variables:		*Dependent variables:*	
Income per capita	19.42	Prescription Opioids MR	-0.000795
	[62.24]		[0.000580]
Share with some college	0.0063	Any Opioids MR	-0.00101
	[0.00386]		[0.000671]
Share with high school or less	0.00257	Share SNAP	-0.000529
	[0.00420]		[0.000840]
Share working in manufacturing	0.0063	Share SSDI	-0.000523
	[0.00386]		[0.000890]
Labor Force Participation	-0.00153*	Share SSI	0.000151
	[0.000821]		[0.000345]
Employment rate	-0.000781	Infant Mortality Rate	-0.0989
	[0.000489]		[0.154]
Total crime rate	44.5	Fertility rate	-0.641
	[28.63]		[0.490]

Notes: Each coefficient corresponds to a separate regression where the dependent variable is measured as the change with respect to 1989-1990. For prescription opioids, any opioids, labor market variables, SNAP, and infant mortality rate, we run a panel regression; for the other variables, where yearly data are not available, we run one cross-sectional regression. MR stands for mortality rate. All regressions include as control variables in long changes: cancer mortality rate, share of population under 1 year, share of population between 18 and 65, share of population over 66 years old, share of Black, White, and Hispanic population, and share of female population. In panel-level regressions, standard errors are clustered at the commuting-zone level; in cross-sectional regressions, standard errors are robust to heteroskedasticity. * $p<0.10$, ** $p<0.05$, *** $p<0.01$. This table is referenced in Section 4.4.2

Table 2.5: Direct Effects on Opioid Mortality

Dependent var:	Prescription opioids MR			Any Opioid MR		
	(1)	(2)	(3)	(4)	(5)	(6)
Prescription opioids pc	0.00374***		0.00679***	0.00419***		0.00646***
	[0.00117]		[0.00200]	[0.00139]		[0.00231]
tF 0.05 se			(0.00281)			(0.00324)
t-stat using tF 0.05 se			2.3876			1.9747
AR p-value			0.0000			0.0019
Cancer MR 94-96		0.00732***			0.00697***	
		[0.00167]			[0.00229]	
Effect size (%)	49.47		88.63	25.73		39.30
Model	OLS	RF	IV	OLS	RF	IV
Observations	11,800	11,800	11,800	11,800	11,800	11,800
Clusters	590	590	590	590	590	590
Adj R^2	0.4304	0.3908		0.5368	0.5144	
Effective F-stat			16.63			16.63
Cragg-Donald Wald F-stat			358.58			358.58

Notes: Control variables in long changes are contemporaneous cancer mortality rate, share of population under 1 year old, share of population between 18 and 65, share of population over 66 years old, share of Black, White, and Hispanic population, and share of female population. All regressions include state times year fixed effects. MR stands for mortality rate. Effect size indicates the percent change in the dependent variable relative to its mean when doses of prescription opioids per capita increase from the 25th to the 75th percentile. Standard errors in square brackets are clustered at the CZ level. Using these standard errors, we report * $p<0.10$, ** $p<0.05$, *** $p<0.01$. tF 0.05 se, t-stat using tF0.05 se, and the *AR p-value* correspond to weak-instrument-robust inference procedures. This table is referenced in Section 5.5.1

Table 2.6: Effects of the Opioid Epidemic on Other Mortality Measures

Dependent var:	All cause mortality			Deaths of Despair		
	(1)	(2)	(3)	(4)	(5)	(6)
Prescription opioids pc	0.0213		0.0286	-0.000442		-0.00494
	[0.0136]		[0.0469]	[0.000732]		[0.00621]
tF 0.05 se			(0.0658)			(0.0087)
t-stat using tF 0.05 se			0.4346			-0.459
AR p-value			0.5319			0.4311
Cancer MR 94-96		0.0309			-0.00533	
		[0.0515]			[0.00699]	
Effect size (%)	3.68		4.94	-0.74		-7.39
Model	OLS	RF	IV	OLS	RF	IV

Dependent var:	Alcoholic Liver Diseases and Cirrhosis			Suicide		
	(1)	(2)	(3)	(4)	(5)	(6)
Prescription opioids pc	0.000765**		0.00552*	-0.0000460		-0.00582
	[0.000353]		[0.00292]	[0.000430]		[0.00378]
tF 0.05 se			(0.0041)			(0.0053)
t-stat using tF 0.05 se			1.3473			-1.0974
AR p-value			0.0351			0.1065
Cancer MR 94-96		0.00596**			-0.00628	
		[0.00302]			[0.00402]	
Effect size (%)	3.23		23.34	-0.16		-19.80
Model	OLS	RF	IV	OLS	RF	IV

Notes: The all-cause mortality measure excludes deaths from cancer. Deaths of despair refers to deaths from suicide, chronic liver disease, cirrhosis, and poisonings that are attributable to alcohol. Each regression is run over a sample of 11,800 observations with 590 clusters (commuting zones). Control variables in long changes are contemporaneous cancer mortality rate, share of population under 1 year old, share of population between 18 and 65, share of population over 66 years old, share of Black, White, and Hispanic population, and share of female population. All regressions include state times year fixed effects. Effect size: indicates the percent change in the respective dependent variable relative to its mean when doses of prescription opioids per capita increase from the 25th to the 75th percentile. Standard errors in square brackets are clustered at the CZ level; using these standard errors, we report * $p<0.10$, ** $p<0.05$, *** $p<0.01$. tF 0.05 se, t-stat using tF0.05 se, and the *AR p-value* correspond to weak-instrument-robust inference procedures. This table is referenced in Section 5.5.2

Table 2.7: Effects of the Opioid Epidemic on the Share of Social Insurance and Welfare Programs Recipients

Dependent var:	SSDI			SSI		
	(1)	(2)	(3)	(4)	(5)	(6)
Prescription opioids pc	0.000444***		0.00574***	0.0000137		0.00184**
	[0.0000985]		[0.00132]	[0.0000896]		[0.000841]
tF 0.05 se			(0.0018)			(0.0012)
t-stat using tF 0.05 se			3.1250			1.5612
AR p-value			0.0000			0.0107
Cancer MR 94-96		0.00619***			0.00198**	
		[0.000385]			[0.000803]	
Effect size (%)	5.36		76.39	0.27		38.43
Model	OLS	RF	IV	OLS	RF	IV

Dependent var:	SNAP		
	(1)	(2)	(3)
Prescription opioids pc	0.000144		0.00982***
	[0.000285]		[0.00299]
tF 0.05 se			(0.0041)
t-stat using tF 0.05 se			2.4134
AR p-value			0.0000
Cancer MR 94-96		0.0106***	
		[0.00227]	
Effect size (%)	0.58		56.70
Model	OLS	RF	IV

Notes: Each regression is run over a sample of 11,800 observations with 590 clusters (commuting zones). Control variables in long changes are contemporaneous cancer mortality rate, share of population under 1 year old, share of population between 18 and 65, share of population over 66 years old, share of Black, White, and Hispanic population, and share of female population. All regressions include state times year fixed effects. Effect size: indicates the percent change in the respective dependent variable relative to its mean when doses of prescription opioids per capita increase from the 25th to the 75th percentile. Standard errors in square brackets are clustered at the CZ level; using these standard errors, we report * $p<0.10$, ** $p<0.05$, *** $p<0.01$. tF 0.05 se, t-stat using tF0.05 se, and the *AR p-value* correspond to weak-instrument-robust inference procedures. This table is referenced in Section 5.5.2

Table 2.8: Effects of the Opioid Epidemic on Infant Outcomes

Dependent var:	Infant Mortality Rate			Birth Weight		
	(1)	(2)	(3)	(4)	(5)	(6)
Prescription opioids pc	0.0511**		-0.0232	-0.552*		-4.490**
	[0.0242]		[0.140]	[0.331]		[2.143]
tF 0.05 se			(0.19643)			(3.00676)
t-stat using tF 0.05 se			-0.1181			-1.4933
AR p-value			0.8678			0.0163
Cancer MR 94-96		-0.0250			-4.843**	
		[0.157]			[2.127]	
Effect size (%)	4.06		-1.84	-0.08		-0.69
Model	OLS	RF	IV	OLS	RF	IV

Dependent var:	Share low birth weight			Preterm births		
	(1)	(2)	(3)	(4)	(5)	(6)
Prescription opioids pc	0.000169*		0.000905	0.000270*		0.00141
	[0.000102]		[0.000640]	[0.000150]		[0.000937]
tF 0.05 se			(0.00090)			(0.00131)
t-stat using tF 0.05 se			1.0023			1.0649
AR p-value			0.1272			0.1126
Cancer MR 94-96		0.000976			0.00152	
		[0.000665]			[0.00100]	
Effect size (%)	0.62		5.55	0.84		5.90
Model	OLS	RF	IV	OLS	RF	IV

Dependent var:	APGAR Score - All Infants			APGAR Score - infant casualties		
	(1)	(2)	(3)	(4)	(5)	(6)
Prescription opioids pc	-0.000501		-0.0169*	0.0155		0.282*
	[0.00188]		[0.00994]	[0.0179]		[0.153]
tF 0.05 se			(0.01395)			(0.21467)
t-stat using tF 0.05 se			-1.2118			1.3137
AR p-value			0.0674			0.0383
Cancer MR 94-96		-0.0189*			0.319*	
		[0.0107]			[0.164]	
Effect size (%)	-0.03		-0.96	1.38		25.17
Model	OLS	RF	IV	OLS	RF	IV

Notes: Each regression is run over a sample of 11,800 observations with 590 clusters (commuting zones). Control variables in long changes are contemporaneous cancer mortality rate, share of population under 1 year old, share of population between 18 and 65, share of population over 66 years old, share of Black, White, and Hispanic population, and share of female population. All regressions include state times year fixed effects. Effect size: indicates the percent change in the respective dependent variable relative to its mean when doses of prescription opioids per capita increase from the 25th to the 75th percentile. Standard errors in square brackets are clustered at the CZ level; using these standard errors, we report * $p<0.10$, ** $p<0.05$, *** $p<0.01$. tF 0.05 se, t-stat using tF0.05 se, and the *AR p-value* correspond to weak-instrument-robust inference procedures. This table is referenced in Section 5.5.2

Table 2.9: Effects of the Opioid Epidemic on Fertility Outcomes

Dependent var:	Fertility rate			Gestation		
	(1)	(2)	(3)	(4)	(5)	(6)
Prescription opioids pc	0.0000665		0.00153***	-0.000164		-0.0489***
	[0.0000621]		[0.000566]	[0.00304]		[0.0186]
tF 0.05 se			(0.00079)			(0.02610)
t-stat using tF 0.05 se			1.9266			-1.8738
AR p-value			0.001			0.0011
Cancer MR 94-96		0.00165***			-0.0527***	
		[0.000482]			[0.0171]	
Effect size (%)	0.43		9.85	0.00		-0.63
Model	OLS	RF	IV	OLS	RF	IV

Notes: Each regression is run over a sample of 11,800 observations with 590 clusters (commuting zones). Control variables in long changes are contemporaneous cancer mortality rate, share of population under 1 year old, share of population between 18 and 65, share of population over 66 years old, share of Black, White, and Hispanic population, and share of female population. All regressions include state times year fixed effects. Effect size: indicates the percent change in the respective dependent variable relative to its mean when doses of prescription opioids per capita increase from the 25th to the 75th percentile. Standard errors in square brackets are clustered at the CZ level; using these standard errors, we report * $p<0.10$, ** $p<0.05$, *** $p<0.01$. tF 0.05 se, t-stat using tF0.05 se, and the *AR p-value* correspond to weak-instrument-robust inference procedures. This table is referenced in Section 5.5.2

Table 2.10: Complier Analysis

Commuting zone characteristic	First Stage (1)	$P[X = x]$ (2)	$P[X = x\|\text{Complier}]$ (3)	$\frac{P[X=x\|\text{Complier}]}{P[X=x]}$ (4)
High sh. of pop below poverty line	1.382***	0.52	0.991	1.923
	[0.499]		[0.178]	
Low sh. of pop below poverty line	0.629***	0.48	0.00863	0.018
	[0.236]		[0.178]	
High sh. of pop w/ less than HS degree	1.125***	0.51	0.964	1.883
	[0.419]		[0.210]	
Low sh. of pop w/ less than HS degree	0.855***	0.49	0.0365	0.075
	[0.322]		[0.210]	
High sh. of employment in mining	1.232***	0.50	0.931	1.856
	[0.356]		[0.235]	
Low sh. of employment in mining	0.820*	0.50	0.0694	0.139
	[0.471]		[0.235]	
High sh. of PCP per capita	1.180**	0.50	1.427	2.854
	[0.518]		[0.224]	
Low sh. of PCP per capita	1.103***	0.50	-0.427	-0.854
	[0.218]		[0.224]	
High sh. of smoking	1.012***	0.54	0.645	1.197
	[0.365]		[0.213]	
Low sh. of smoking	0.825***	0.46	0.355	0.770
	[0.298]		[0.213]	
High cocaine and alcohol MR	1.273***	0.50	1.12	2.240
	[0.440]		[0.188]	
Low cocaine and alcohol MR	0.656**	0.50	-0.12	-0.240
	[0.275]		[0.188]	

Notes: Column 1 corresponds to the first stage regression for each specific group. Column 2 is the frequency of the group in the estimation sample. Column 3 corresponds to the estimation of the characteristic in the complier sample, following Abadie (2003) this is a 2SLS regression where the dependent variable corresponds to the endogenous variable multiplied by the indicator of the group. Column 4 divides column 3 by column 2 and corresponds to the complier relative likelihood. For each of the commuting zone characteristics, we consider a commuting zone to be in the low (high) group if the value of such characteristic is below (above) the median value. Poverty, share of the population with less than a high school (HS) degree, and employment in the mining sector are measured in 1994. Primary care physicians (PCP) per capita, smoking, and cocaine and alcohol mortality rates are measured in 1996.

Chapter 3

Tobacco Consumption Habits in Argentina: Causal Evidence from a New Regulation

1. Introduction

Modifiable risky behaviors, such as smoking tobacco, are a major determinant of premature death in both developed and developing countries (Blecher, 2008; Cawley and Ruhm, 2011). The World Health Organization (WHO) estimates that tobacco use is the leading risk factor in high-income countries, accounting for 18% of deaths.[1] In middle-income countries, where 80% of smokers worldwide live, tobacco use is the second most important risk factor and is responsible for about 11% of deaths. Moreover, Goodchild, Nargis and d'Espaignet (2018) estimate that 40% of the health care expenditure due to smoking-attributable diseases occurs in developing countries, highlighting the substantial burden these countries face.

Although tobacco has long been heavily taxed, non-price policies have become increas-

[1] WHO, 2009

ingly common in the last two decades. The most frequently implemented non-price policies are place-based bans and warning labels. For instance, by 2016, at least 105 countries required graphic warnings to be printed on cigarette packages.[2] Non-price policies are particularly attractive in middle- and low-income countries, where tobacco use is concentrated among low-income households.[3]

In this paper, I examine the effect of non-price interventions on smoking, drinking, and health outcomes in the context of a middle-income country, Argentina, where smoking is responsible for 13.2% of deaths, with a direct cost equal to 0.75% of its GDP—or, equivalently, 7.5% of the 2015 federal health budget (Alcaraz et al., 2016). In 2011, the Argentinean government implemented an anti-smoking law that banned smoking in public spaces, including bars and restaurants, and in public and private workplaces; restricted cigarette sales; and regulated the inclusion of graphic tobacco warnings on cigarette packages. The Argentinean case offers a good setting for this research because (i) the policy implementation was not accompanied by changes in taxation or cigarette prices, so it represents a *pure* non-price policy, and (ii) Argentina is a federal country, like the US, which allows me to exploit state-level differences in the stringency of tobacco regulation before 2011 as a source of exogenous variation.

The contribution of this paper is threefold. First, this paper contributes to a vast literature in health economics that evaluates place-based bans and has yielded mixed results. North America-based research shows mixed impacts of smoking bans on smoking behavior (Adda and Cornaglia, 2010; Bitler, Carpenter and Zavodny, 2010; Carpenter, Postolek and

[2]In 2001, Canada became the first country in the world to introduce graphic warnings on cigarette packages and was quickly followed by many others (Azagba and Sharaf, 2012). In the US, the 2009 Tobacco Control Act required that cigarette packaging includes graphic warnings, but implementation has been delayed by legal challenges from cigarette manufacturers. For example; on March 2, 2021, the U.S. District Court for the Eastern District of Texas granted a motion by the plaintiffs in the case of R.J. Reynolds Tobacco Co. et al. v. United States Food and Drug Administration et al., No. 6:20-cv-00176, to postpone the effective date of the "Required Warnings for Cigarette Packages and Advertisements" final rule by an additional 90 days. The new effective date of the final rule is April 14, 2022.

[3]Tobacco use rises to 29% in households in the lowest income quintile versus 20% in the richest households; see Crawfurd and Nestour (2019) and Fuchs, Gonzalez Icaza and Paz (2019).

Warman, 2011; Burton, 2020), while research in Europe suggests a systematic reduction of small magnitude in smoking behavior (Buonanno and Ranzani, 2013; Pieroni et al., 2013; Sureda et al., 2014). The local context of smoking behavior could help explain differences in estimates: how popular and accepted smoking is, whether it is strongly associated with other behaviors, and whether the accessibility of smoking cessation programs among others would contribute to individuals' responses to anti-tobacco policies. Nonetheless, the effectiveness of smoking restrictions in the context of middle- and low-income countries, where 80% of tobacco smokers worldwide live, has been much less studied. My research fills this gap by analyzing the anti-tobacco law passed by the Argentinean federal government in 2011. Unlike other studies of the effects of smoking bans in the context of middle- and low- income countries that have relied on cross-country (e.g, Blecher, 2008; Abascal et al., 2012) or before-and-after comparisons (e.g., Thrasher et al., 2010), I employ an identification strategy that provides causal estimates of the impacts of this policy on smoking, drinking, and health outcomes.

Second, this paper contributes to the literature on the impacts of graphic tobacco warnings. Cigarette warning labels are well-established policies, but research based on actual consumption behavior is limited. The effectiveness of this policy has been assessed by randomized controlled laboratory experiments (e.g., Hammond, 2011) and cross-country comparisons (e.g, Blecher, 2008). The main empirical challenge in evaluating this policy is that graphic warnings have often been enacted as part of broader anti-smoking campaigns (DeCicca, Kenkel and Lovenheim, 2022). To overcome this challenge and disentangle the effect of graphic tobacco warnings and place-based bans, I combine (i) a simple framework that imposes structure on the effect of a bundled policy and (ii) additional variation among early and late adopters of place-based bans. My paper is closest to Kuehnle (2019) who uses individual-level panel data from Australia to examine the association between pictorial warnings and smoking behavior—prevalence, quitting,initiating and relapsing. Nonetheless, one limitation

of their study is that it cannot separate the effect of pictorial warnings from the effects of the introduction of Quitline reference, and mass media campaigns that accompanied the policy.

Third, I aim to advance the economic research on tobacco regulation effects on health outcomes. The main empirical challenge in measuring the effects of regulatory changes on health outcomes originates in the fact that many health conditions associated with smoking—such as, heart disease, lung cancer, and overall premature mortality—are processes that develop over long periods of time making it difficult to know the appropriate lag between policy implementation and disease onset and progression (DeCicca, Kenkel and Lovenheim, 2022). To overcome this limitation, I exploit data on hospital discharges, this approach has two advantages. On one hand, when studying hospitalization events rather than mortality, the lag required to observed health effects is shorter. On the other hand, hospitalization data is well suited to capture health gains, such as less severe symptoms associated with smoking cessation, that might not translate into mortality declines. For example, in the case of chronic obstructive pulmonary disease (COPD), it has been shown that smoking cessation and avoidance of cigarette smoke reduce the excess lung function decline (Eklund et al., 2012). [4]

Data for this paper come from three main sources. First, to estimate the effects of this non-price intervention on smoking and drinking, I use individual-level data from two national surveys for the years 2008 to 2013.[5] These data include geographic location, demographic characteristics, and cigarette and alcohol consumption, which I use to construct the smoking status and participation and drinking behaviors of smokers. Second, to study the effects on health outcomes, I use restricted-access administrative data on hospital discharges provided by the Argentinean National Center for Health Statistics, which I aggregate at the state

[4]Chronic obstructive pulmonary disease (COPD) is a chronic inflammatory lung disease that causes obstructed airflow from the lungs. Symptoms include breathing difficulty, cough, mucus (sputum) production and wheezing. The main cause of COPD is tobacco smoke and smoking accounts for up to 8 out of 10 deaths (US Department of Health, 2014)

[5]The National Survey on Risk Factors (ENFR) and the National Survey on Consumption Prevalence of Psychoactive Substances (EnPreCoSP); data are available for the years 2008, 2009, 2011, and 2013.

level. Third, I compile data on tobacco regulation from state statutes and laws, coding a total of 47 state-level regulations to characterize the regulatory environment of each state before the 2011 national law was enacted.

My identification strategy relies on state-level differences in the strength of regulation on tobacco products before the national law was passed. States with more lenient restrictions were more exposed to the effects of the new regulation than states with strict regulations. I compare the smoking and drinking behavior of individuals in lenient states (treated individuals) with that of individuals in strict states (comparison states). I show that before 2011, treated and comparison states were comparable in observable characteristics, including smoking prevalence and demographic composition. I also show that individuals in treated and comparison states were observationally equivalent before 2011.

I find that the anti-tobacco regulation effectively reduced smoking participation: The probability of being a current smoker decreased by 22% among people aged 18 to 65. On the intensive margin, the share of individuals smoking fewer than 5 cigarettes a day decreased by 0.16 percentage point, suggesting that the reduced probability of being a current smoker is driven by former smokers whose consumption was on the lower end of the distribution. But importantly, the 2011 national law has regressive implications. Non-price policies are expected to be less regressive than price policies, because they do not increase the economic burden of consumption on poorer smoker; who make up a large proportion of the smoking population in developing countries. Despite not changing the monetary cost of smoking, I find that the policy disproportionately benefited more educated and richer individuals. Specifically, an elementary school graduate is 4.37 percentage points less likely to be a current smoker, while a college graduate is 9.86 percentage points less likely after the policy implementation. A similar pattern is observed across the income gradient: an individual who belongs to the highest income quintile is 8.92 percentage points less likely to be a smoker after the national law was introduced—3 percentage points greater in absolute values than

the change for the lowest income quintile.

Understanding whether alcohol and tobacco are consumed as complements or substitutes is crucial in determining the side effects of tobacco control policies, especially when these policies target venues such as bars and restaurants where smoking and drinking are usually combined. Smoking bans in bars and restaurants represent a change in the non-price determinant of demand for alcohol consumed in bars; under the hypothesis that these goods are substitutes, an effective anti-smoking policy would backfire and lead to an increase in alcohol consumption. I find that the new tobacco regulation induced a decrease in the abusive consumption of beer (3 percentage points), wine (2.74 percentage points), and binge drinking (5.44 percentage points). The direction of these changes is consistent with a complementary relationship when it comes to the consumption of these pairs of goods. [6]

The impacts on extensive and intensive margin outcomes are reflected in slightly better health outcomes. The rate of hospital discharges due to diagnoses of chronic obstructive pulmonary disease (COPD) decreased in the short run, but I find no effects on rates of hospital discharges due to lung cancer. I interpret these results as suggestive of an improvement in population health correlated with more strict regulation of tobacco products, but due to data limitations, I cannot perform a longer term analysis of these outcomes or identify whether the results are driven by fewer individuals being diagnosed or the same number of individuals being diagnosed with less severe symptoms and requiring fewer hospitalizations.

The rest of the paper is organized as follows. Section II presents background information on tobacco regulation in Argentina. Section III describes the data. Section IV presents the research design and provides evidence supporting the identification assumption. Section V examines the effects of the national regulation on smoking, drinking, and health outcomes.

[6] I find that, in contrast to beer and wine, smoking and spirits—vodka, gin, tequila, rum, and whiskey—have a substitution relationship on consumption. One potential explanation of this result is that smokers who drink spirits are different from other smokers who abuse alcohol. I find that this result is driven by single young people (less than 25 years) with more education who are increasing their consumption of alcohol.

Section VI discusses the mechanisms through which the law operates and section VII concludes.

2. Non-price Regulation and the Argentinean National Law

Argentina's 2011 national law has three main components: (i) it bans smoking in public spaces, such as restaurants, bars, educational institutions, and public and private work sites; (ii) it bans sales in schools, hospitals, public buildings, and means of public transportation; and (iii) it regulates advertising of tobacco products and mandates the incorporation of graphic tobacco warnings. These warnings consist of short messages about smoking's health consequences and side effects displayed with a shocking picture on one side of the package (see Appendix Figure A.43 for examples). Images are designed by the National Department of Health and are updated once every 18 months in order to attenuate potential wear-out effects of graphic warnings. After the law was approved, producers had up to 6 months to incorporate the designs on their packages and could not adapt the warnings to reduce their effectiveness. [7]

Argentina is a federal country; as in the US, states have autonomy to implement different regulations. States do not have control over cigarette prices, and the taxation of tobacco-related products is exclusive to the national government. However, each state has the autonomy to regulate smoking and access to tobacco products. I compile data on tobacco regulation from state statutes and laws for each of the 23 Argentinean states and the capital city to characterize the regulatory environment of each state before the 2011 national law was enacted. Then, I use this information to construct a legislation index, which is a

[7] An advantage of this setting is that individuals are *almost* randomly assigned to graphic warnings. Cigarettes packages are mainly sold at convenience stores, where consumers cannot directly access the cigarette package and must ask the cashier to retrieve it, leaving very little room to pick which warning is printed on the package they are buying. This means that the effects I estimate are an average effect of the graphic warnings selected by the Argentinean government, but, I cannot provide evidence about which warning is more effective.

discrete and bounded index that summarizes the state-level regulations.[8] Higher values on the legislation index represent strong regulation of tobacco products. A value of 0 indicates that a given state has only banned cigarette sales to minors. A value of 1 indicates that the state has also banned consumption in public means of transportation. A value of 2 indicates that a state has banned consumption in educational and health care institutions or that a state has banned some type of advertising—e.g., event sponsorship. A value of 3 indicates that the state has banned consumption in additional venues, such as bars and restaurants. Each subsequent unit increase indicates a tightening in the regulation. The highest value of the index before 2011 was 6, which implies limiting advertising in public spaces, sport events and other venues; banning indoor smoking in several venues (bars, restaurants, educational and health care institutions); and allowing cigarette sales only in specialty or authorized shops. A value of 7 was reached by all states in 2011 when the national law was passed, which demonstrates the fact that no state had legislation as strict as the new national law. The legislation index allows for a convenient summary of the regulatory apparatus, but its weakness is that each unit increase does not reflect the same change in regulation. The data construction section in Appendix 3.3 provides details on the criteria and definitions used.

Figure 3.1 shows the regional variation across states in 2009. Darker shades on the map (higher values of the legislation index) indicate stricter regulation, and lighter shades (lower values of the legislation index) indicate less strict regulation. The 2011 law impacted the regulation of tobacco products in every state with different intensity. For example, in a state in which only sales to minors were banned (a state with a legislation index equal to 0), the national law introduced place-base bans, advertising limitations and graphic warnings. On the other hand, in a state in which tobacco was heavily regulated—such as a state with a legislation index of 6,—the main change was the incorporation of graphic warnings.

[8]This attempt to assemble data on the enactment of tobacco regulation policies is similar to the Prescription Drug Abuse Policy System, which aims to track state laws related to prescription drug abuse. Researchers have widely used these data to study the effects of such laws on opioid use and mortality (see Buchmueller and Carey, 2018*b*)

Nonetheless, every state experienced tightened regulation after introduction of the 2011 national law. The ideal research design would exploit the intensity of treatment—i.e., the difference between the value of the legislation index for state s and the maximum value of the index. This exercise would be informative regarding the effectiveness of the policy for different increments in strength of regulation. Unfortunately, the number of states and years observed is relatively small, which limits the implementation of this design. Hence, I use the legislation index to identify lenient vs strict states, and use the latter as the comparison or untreated group. In the rest of the paper, I refer to "lenient" states as those with low legislation index value (less than or equal to 3) and to the other states as "strict" or comparison states. As a robustness check, I replicate the main analysis using an alternative definition of treated and comparison states, and find similar results (see robustness Section 3.41).

3. Data and Descriptive Statistics

To measure tobacco and alcohol consumption, I use individual-level data from the National Survey on Risk Factors (ENFR) and the National Survey on Consumption Prevalence of Psychoactive Substances (EnPreCoSP). The surveys provide information on self-reported consumption of tobacco and alcohol within the last year for individuals aged 18 to 65.[9] Importantly, these survey data include state identifiers and are available for the years before and after the policy was implemented: 2008 , 2009, 2011, and 2013. I study two outcomes of smoking behavior. First, I construct extensive margin outcomes, defined as the probability that an individual is a current smoker and the probability that an individual has never smoker. I consider an individual to be a current smoker if she has ever smoked more than 100

[9]Both surveys use the same questionnaire for smoking and alcohol consumption. The main difference between the surveys is that the ENFR also asks individuals about their health, diet, and physical activity, among other behaviors, while the EnPreCoSP also asks individuals about consumption of other drugs, such as marijuana.

cigarettes and if, at the moment of the survey, she smokes every day or some days. I consider an invidual to be a never smoker if she has never smoked before or if, she has smoked less than 100 cigarettes. These definitions are close to the one used by Carpenter, Postolek and Warman (2011) with Canadian survey data. Second, I study an intensive margin measure given by the number of daily cigarettes smoked in the last month; this measure is informative regarding whether the distribution of smoked cigarettes was responsive to the national law.

I recognize two potential limitations of these survey data. First, *social smokers* could be classified as non-smokers. This would imply that my results provide a conservative estimate of the true parameters of interest, since restaurants and bars are where social smokers are more likely to smoke and drink and the policy directly bans consumption in these venues. Second, extensive- and intensive- margin measures are based on self-reported data, which are not free of measurement error; this would arise if smokers deny their habit when surveyed or incorrectly report how much they smoke. I use data on sales to provide evidence on the direction of measurement error in Table A.31. Consumption time series from sales and survey data show a similar trend and suggest that (i) individuals underreport their consumption and (ii) under-reporting is stable across the years, i.e. it did not change after the implementation of the national law.[10] Smoking behavior outcomes are measured with some error in the survey data; however, the direction of this error is consistent with my results providing a lower bound rather than being driven by changes in reporting after the policy.

To document whether cigarettes and alcohol are substitutes or complements, I construct measures of alcohol consumption among smokers. I study alcohol consumption in the last month and two measures of harmful alcohol consumption: abusive consumption and binge drinking. I define alcohol abuse as an indicator of having more than eight drinks of beer, five drinks of wine or three drinks of spirits in a given day.[11] Binge drinking is defined as

[10]Survey data captures about 58% of consumption relative to sales data for every year that is available. Unfortunately, sales data is not available at the state level so I can not reproduce my main results using these data.

[11]Spirits include vodka, gin, tequila, rum, and whiskey. These thresholds are defined following the guide-

consuming five or more drinks on a single occasion in the past 30 days, on either a weekend or a weekday.

Last, I study effects on health outcomes. I use restricted-access administrative data on hospital discharges to compute the prevalence rate of hospitalization from COPD and respiratory-system-related cancer.[12] These data come from the Argentinean National Center for Health Statistics and are provided at the hospitalization level.[13] In the data, one observation corresponds to a hospitalization event for which I observed the individual's gender, age, and the main diagnosis and a state identifier. The main limitation of these data is that I do not observe individual identifiers or their smoker status. For this reason, my analysis is only suggestive with respect to how the population's health changed after the national law and cannot distinguish whether there are fewer diagnoses or less severe ones.

Table 3.1 shows mean outcomes before the national law was enacted. Strict states (high values in the legislation index) and lenient states (low values in the legislation index) have comparable proportions of never and current smokers. But, the distribution of smoked cigarettes per day in lenient states has a heavier right tail; i.e., the proportion of more-than-10-cigarettes-per-day smokers is higher in these states. The difference in the distribution of smoked cigarettes per day—statistically significant at the 5- and 10-percent level—suggests that strict and lenient states might have been on different trends before 2011. I directly test for the presence of pre-trends in intensive-margin outcomes in Table 3.3. I find no statistically significant differences in the proportion of smokers whose consumption is between one and five cigarettes per day, or between 10 and 15 cigarettes a day. I find a small difference in the consumption group 15 to 20 cigarettes per day, which is statistically significant at a 10-percent significance level and not economically meaningful. Importantly, strict and lenient

lines from the World Health Organization (WHO) and the Center for Diseases Control (CDC).

[12]COPD is defined using ICD 10th revision codes J41, J42, J43, and J44. Respiratory-system-related cancer is defined using ICD 10th revision codes C30, C33, and C34.

[13]In Spanish, Direccion de Estadisticas e Informacion de Salud (DEIS). This agency is analogous to the U.S. National Center for Health Statistics.

states show similar prevalence rates in the diagnosis directly affected by tobacco smoking.

4. Empirical Strategy

The 2011 law impacted the regulation of tobacco products nationally, but I exploit the fact that states with more lenient restrictions were more exposed to the effects of the new regulation, by comparing more lenient with stricter states. To do so, I define a dichotomous variable ($Treat_s$) that equals 1 if the legislation index for state s is less than or equal to 3 before 2011; i.e., states with legislation strictly greater than 3 serve as comparison states.[14] I estimate the causal effect of the national regulation via the following regression, which is run over the sample of individuals in treated and comparison states two periods before and one period after the regulation was implemented:

$$y_i = \sum_{\tau=-2}^{1} \delta_\tau \left[Treat_s \cdot (Years\ After\ Treat = \tau)\right] + \beta' X_{is} + \Gamma' X_{st} + \alpha_s + \alpha_t + \varepsilon_i \ , \qquad (3.1)$$

where s indexes the state in which individual i is observed in year t,and y_i is the outcome of interest. The variable $Years\ After\ Treat = \tau$ indexes time relative to the law's implementation.[15] The variable $Treat_s$ indicates whether state s is a comparison state ($Treat_s = 0$) or a treated state ($Treat_s = 1$). In this specification, lagged coefficients (δ_{-2} and δ_{-1}) test whether the outcome of interest y in comparison and treated states followed similar trends before the policy was introduced. The coefficient for the lead period, δ_1 captures the effect of the national law 1 year after implementation. The omitted coefficient is δ_0, which corresponds to the year of the treatment. X_{is} is a vector of control variables that comprises individual-level characteristics—e.g., age, gender, educational attainment, and income. I also include state-time varying controls: private employment and population. States' fixed

[14] I present estimates of the main results defining $Treat_s = 1$ if the legislation index is strictly less than 3 in the robustness checks section (see Section 3.4) and arrive at similar conclusions.

[15] I normalize time relative to the law implementation to match the years when I observe outcome variables: i.e., $\tau = -2$ corresponds to 2008, $\tau = -1$ to 2009, $\tau = 0$ to 2011, and $\tau = 1$ to 2013.

effects (α_s) control for variation in outcomes across states that is constant over time. Time fixed effects (α_t) control for variation in outcomes over time that is common across all states. The variable ε_i is an individual error term. Standard errors are block-bootstrapped at the state level with 200 replications.[16]

My identifying assumption is that comparison states would have been on the same trend as treated states absent the national legislation. The remainder of this section provides evidence in support of the research strategy and discusses the identifying assumption.

4.1 Why Did State Governments Regulate Tobacco Consumption?

As shown in Table 3.1, treated and comparison states are comparable in terms of the outcomes of interest before the national law was enacted. Nonetheless, identifying the causal effect of the 2011 national law on tobacco consumption requires that treatment and comparison status at the state level be exogenous to outcomes of interest. In order to better understand why some states chose to regulate tobacco more than others before 2011, I regress the probability of having enacted strong regulation before 2011 on a set of covariates that include the political affiliation of the state government, an indicator of whether a state is a tobacco producer, measures of the prevalence of tobacco consumption and hospitalizations due to COPD, a measure of private employment that proxies the economic well-being of a state, and demographic characteristics. These results are presented in Table A.32. The Peronist party was responsible for enacting the 2011 national law; nonetheless, I find that the it was not more likely to regulate tobacco consumption at the state level than opposition parties. Similarly, I find that states that are tobacco producers did not have a smaller probability of being strict states, although one would expect these states to be more lenient regarding

[16]Using cluster-robust standard errors permits both error heteroskedasticity and flexible error correlation within clusters. With a small number of clusters, cluster-robust standard errors are downward biased (Cameron, Gelbach and Miller, 2008); block-bootstrap procedures yield more accurate cluster-robust inference when there are few clusters. Block-bootstrap standard-error estimation procedure maintain the autocorrelation structure within groups—states in this specific application—by keeping observations that belong to the same group together in a block, as it samples groups instead of observations.

the consumption of tobacco products. States with a higher prevalence of COPD diagnoses were 0.394% more likely to have enacted a stricter regulation; although statistically different from zero, this coefficient indicates that for the probability of enacting a strong regulation to increase in 1% the prevalence of COPD has to increase 2.53 times. Finally, states with a greater share of individuals between 15 and 25 years were more likely to have a stricter regulation. In my estimates, I control for state fixed effects, which partially purge initial differences across states in the prevalence of COPD and demographic composition.

4.2 Were Individuals in Treated and Comparison States Observationally Equivalent?

I test whether individuals in treated and comparison states were statistically indistinguishable in terms of their observed characteristics in 2008 and 2009.[17] Balancing test results are presented in Table 3.2. The estimated differences indicate that individuals are statistically indistinguishable in terms of gender, age, marital status, employment status, and educational attainment.

I find evidence that on average, individuals in comparison states are poorer than individuals in treated states. The proportion of households in the second quintile is bigger in comparison states, while the proportion of households in the third and fourth quintiles is smaller in comparison states. To account for this difference in observable characteristics, I include controls for household income in every specification, as well as age, gender, employment status, and other individual characteristics.

4.3 Did State Governments Enforce the 2011 National Law?

Before examining the effects of the policy on outcomes of interest, I provide evidence of compliance with the law using data on exposure to environmental tobacco smoke in closed

[17] I present pooled results only. Conclusions on balance on observable characteristics do not change when the analysis is done by year.

venues. In the 2009 and 2013 surveys, non-smokers were asked whether they notice someone smoking inside their home, the workplace, an educational institution, a hospital, or in bars and restaurants. With the exception of bars and restaurants, all states banned consumption in these venues before implementation of the national law. I use these data to document how exposure to tobacco smoke changed across venues before and after 2011. Figure A.44 shows that the share of non-smokers not exposed to environmental tobacco smoke increased by 5% in homes, workplaces, educational institutions and hospitals; whereas the share of non-smokers not exposed in bars and restaurants increased by 10%. These results suggest that exposure to tobacco smoke in bars and restaurants decreased more than in other venues. I argue that, since smoking in the workplace, educational institutions, and hospitals was already regulated in comparison and treatment states before the 2011 law, the differential response in restaurants is informative of law enforcement and compliance.

5. Results

My empirical analysis proceeds in three steps. First, I estimate the causal impact of the 2011 national law on smoking behavior, by examining smoking participation and intensity among smokers. These exercises provide evidence on the effectiveness of the policy. Second, to assess potential side-effects of the tobacco control policy, I ask whether alcohol and tobacco are consumed as complements or substitutes . Finally, I study health outcomes in the short run.

5.1 Effects of the 2011 National Law on Tobacco Consumption

Extensive margin outcomes. Smoking participation decreased as a result of the stricter regulation implemented in 2011. To understand how the policy affected smoking participation, I study its effects on (i) the probability of individuals' having never smoked, and (ii) the proba-

bility of individuals' being current smokers. The former increased by 4.34 percentage points. This suggests that the regulation curbed tobacco initiation: The law caused a roughly 10% increase in the probability of an individual's being a never-smoker. This result is reinforced by a reduction of 6.17 percentage points in the probability of being a current smoker, which represents a 22% reduction in this outcome. Taken together these effects suggest that the law caused a reduction on smoking participation helping individuals to quit smoking and curving smoking initiation. Figure 3.2 shows these results and Appendix Table A.33 presents point estimates and standard errors. These event-study graphs also show that comparison and treated states were on similar trends regarding cigarette consumption before the national law was passed. The estimates for $\delta_{\tau=-2}$ and $\delta_{\tau=-1}$ are not statistically distinguishable from zero, which provides additional support to the identification assumption.

Intensive Margin Outcomes. I examine whether there is a reduction in the number of daily cigarettes smoked. To measure smoking intensity, I generate bins (b) of consumption in five-unit increments up to 20 cigarettes a day, and 10-unit increments for higher quantities. These intervals are open to the left: (0,5] (5, 10] (10, 15] (15, 20] (20, 30] (30,40] and (40,50].[18] For example, b_1 is a dummy variable that takes the value of one if the number of cigarettes smoked is positive but less than or equal to five. Results are presented in Table 3.3.

I find evidence that the distribution of cigarettes smoked daily shifts to the right; i.e., it is more likely that smokers consume more than 5 cigarettes per day and less likely that they consume fewer than 5 cigarettes per day. Specifically, the proportion of smokers who consume between one and five cigarettes per day decreased by 16 percentage points, while the proportion consuming between 5 and 20 cigarettes a day increased by 13 percentage points. Notice that about 93% of the consumers have consumption in the first four bins—i.e., up to a pack per day. Thus, estimating the effects for bins above 20 cigarettes a day requires a

[18]In Table 3.3, I present the proportion of smokers for each bin in 2011. I did not include smokers who report smoking more than 50 cigarettes a day; they represent the 3% of the sample in 2011. Note that the shares of smokers in each bin add up to one and that the coefficients for each year add up to zero, since they represent the changes in accumulated mass in the distribution of smoked cigarettes.

bigger sample or sizable effects to avoid power-related limitations.

Extensive and intensive margin results indicate that the reduction in the probability of being a current smoker is driven by quitters whose consumption was less than five cigarettes per day. Thus, after the law was passed, the pool of smokers becomes more negatively selected.

Heterogeneity in Smoking Participation. Do the effects of stricter restrictions on tobacco consumption vary across populations? To tackle this question, I estimate equation (3.1) restricting the sample by age group, gender, educational attainment, and household income. Figures 3 and 4 plot estimates of the causal effect of the national law by demographic group. Tables A.34, A.35, A.36, and A.37 complement these figures and present estimates of the pre-policy coefficients which are statistically indistinguishable from zero and provide additional support for the parallel trends assumption.

The reduction in the probability of being a current smoker is driven by individuals aged 18 to 25 years old and individuals aged 40 to 55 years old. Although the estimated effects are comparable in magnitude, the mechanisms behind them are quite different. The reduction in current smokers among the younger adults is driven by an increase in non-smokers and shows how the policy curbed tobacco initiation; whereas the result among adults aged 40 to 55 is driven by an increase in former smokers. In 2013, the percentage of never-smokers aged 18 to 25 was 66.54% and this figure for adults aged 50 to 55 was 46.89%, a 20-point difference. At the same time, the percentage of current smokers was comparable in the two age groups—26.66% and 28.81%, respectively, suggesting that the group of younger non-smokers had a higher share of never-smokers, while the group of older non-smokers had a higher share of former smokers.

Adults aged 55 to 65 did not respond to the increase in the cost of smoking induced by the national law. A potential explanation is that people might find quitting more difficult the older they are and the longer they have been addicted to nicotine. The length of ad-

diction mechanically increases as consumers age if they do not quit. Using intention-to-quit data, I find that 67% of smokers aged 50 to 55 have intended to quit smoking, while only 60% of smokers aged between 55 and 60 have intended to quit. Hence, the hypothesis of discouraged smokers is a plausible explanation for the null effects estimated for these groups. Other papers have find similar results, e.g., Kuehnle (2019) find that the smoking behavior of individuals aged 50 and over was not affected by the introduction of tobacco graphic warnings.

Women are less likely to smoke than men, but they are equally responsive to increases in the cost of smoking induced by the national law. Panel (b) of Figure 3.3 presents estimated effects by age and gender and tests the null hypothesis that the effect on women of a given age is equal to the effect on men of the same age group. I do not find evidence to reject this hypothesis for any age group (see Table A.35 for the respective p-values).

Cigarette taxes would be regressive with respect to income if poorer and richer consumers smoked at the same rate, since the same price increase entails a higher economic burden for poorer consumers; this regressive effect is exacerbated when smoking prevalence is inversely related to income, as is the case in most developing countries. In contrast, non-price policies are particularly attractive in middle- and low-income countries, where tobacco use is concentrated among low-income households. These policies change the non-monetary costs of smoking and have the comparative advantage of not increasing the economic burden for consumption on poorer smokers. Nonetheless, I find that the 2011 national law has regressive implications. The causal impacts of the law on tobacco consumption are higher in absolute values for more educated and richer individuals (see Figure 3.4).

I document that an elementary school graduate is 4.37 percentage points less likely to be a current smoker after the policy, while a college graduate is 9.86 percentage points less likely to be a current smoker after the policy implementation. Also, an individual who belongs to the highest income quintile is 8.92 percentage points less likely to be a smoker after the

national law was introduced. This change is 3 percentage points higher in absolute values than the change for the lowest income quintile (a 5.61 percentage points decrease).

Robustness checks. The main results on extensive and intensive margin outcomes are robust to a change in the definition of the treatment. The baseline specification defines the treatment group as those states with a legislation index less than or equal to 3 before 2011. I consider an alternative definition and use as a treatment group those states with a legislation index strictly less than 3 before 2011. This change implies that the pool of treated states has on average a more lenient regulation. I present the results of this exercise in Appendix Tables A.42, A.43, and A.44. Regarding extensive margin outcomes, I find that (i) estimates of the coefficients for 2008 and 2009 are not statistically distinguishable from zero, and (ii) results on the probability of being a never-smoker and the probability of being a current-smoker are smaller in absolute value but still significantly different from zero. Regarding intensive margin outcomes, I find reductions in the probability of smoking up to 5, up to 10 and up to 15 cigarettes per day. These results imply that the distribution of cigarettes smoked daily shift to the right, similarly to the main results.

I also consider two additional robustness checks. First, I test whether the results are driven by one state by estimating the main outcomes of interest in samples that drop one state at a time. I present these results in Figure A.45, which results indicate that the estimated effect is not driven by one state with particularly strong effects. Second, another concern is whether the trend in cigarette price confounds my results. I present evidence that for the period of analysis, there is no sharp changes in the price of the cheapest cigarette package (see Figure A.46). I pay special attention at the minimum price because smokers could easily substitute between brands to avoid price changes in the more expensive brands.

Policies that ban the consumption of tobacco in bars and restaurants could have important effects on a smoker's alcohol consumption, since a smoking ban could lower a smoker's utility from drinking in a bar. Also, the question of whether alcohol and tobacco are consumed as complements or substitutes is crucial for determining the side effects of tobacco control policies. Under the hypothesis that these goods are substitutes, an effective anti-smoking policy would backfire and lead to an increase in alcohol consumption, all else equal. Conversely, under the hypothesis that they are complementary goods, an effective anti-smoking policy would lead to a decrease in alcohol consumption. Most of the literature has studied the responses of alcohol consumption to a price policy by exploiting changes in cigarette prices (e.g., Decker and Schwartz, 2000; Tauchmann et al., 2013; Krauss et al., 2014; Shrestha, 2018). A notable exception is Burton (2020), who finds that smoking bans in bars and restaurants lead to an increase of one serving of alcohol a month, though these bans do not have an effect on smoking participation. In this paper, I study alcohol consumption and alcohol abuse; the latter is particularly important in assessing the potential negative side effects of tobacco control policies.

Figure 3.5 reports estimates of the marginal effects on alcohol consumption from the estimation of equation (3.1) using a probit model. Appendix Table A.38 presents the estimates of the (raw) coefficients. Marginal effects can be interpreted as the percentage-point change in the relevant outcome relative to the average level of such outcomes. Standard errors are clustered at the state level. The estimated coefficients for 2008 and 2009 are statistically indistinguishable from zero; this supports the identifying assumption of parallel trends in the consumption of alcoholic beverages between comparison and treated states.

I find no evidence of changes in the average consumption of alcohol by smokers (column 1 of Table A.38), but measures of extreme consumption did change as a result of stricter

tobacco regulation, i.e., I do find a reduction in unhealthy alcohol drinking behaviors, such as, alcohol abuse and binge drinking. In particular, cigarette smoking and the abusive consumption of beer and wine are complementary goods, i.e., the consumption of wine and beer decreases after an increase in the non-monetary cost of smoking.

Similarly, tobacco consumption and binge drinking also change in a direction that suggests a complementary relationship. The stricter regulation of tobacco consumption caused a reduction in smoking participation and a decrease in the abusive consumption of beer of 3.20 percentage points; of 3.00 percentage points for wine, and of 5.67 percentage points for binge drinking.

Tobacco and spirits—vodka, gin, tequila, rum, and whiskey—have a substitution relationship. The consumption of spirits increased by 5.98 percentage points after the stricter regulations on tobacco consumption and advertising went into effect. The average level of abusive consumption of spirits in 2011 was 3.18%, and thus the estimated effect suggests a huge increase in this risky behavior. A potential explanation of this result is that smokers who drink spirits are different from other smokers who abuse alcohol. Smokers who abuse spirits are more likely to be single young people (less than 25 years) and are more educated and wealthier than those who abuse other drinks; see appendix Table (A.39). Thus, the difference in sign of the coefficient could be attributable to young people who were induced to substitute smoking for drinking in bars and pubs. This seems to be especially true for young men; the policy effect on spirits consumption is three times larger for men than for women.

5.3 Effects of the 2011 National Law on Health Outcomes

Reduced smoking in public places could entail important health benefits on both nonsmokers and smokers. Smoking can cause lung disease by damaging the airways and the small air sacs (alveoli) in the lungs. COPD, which includes emphysema and chronic bronchitis, is one

of the most common lung diseases caused by smoking. Also, cigarette smoking causes most cases of lung cancer. Estimates from the US indicate that smoking causes about 90% of all lung cancer deaths and 80% of all deaths from COPD.[19] In the Argentinean setting, Alcaraz et al. (2016) estimate that cigarette smoking caused 75% of COPD deaths and 82% of lung cancer deaths, while 33% of other cancers could be attributed to cigarette smoking.

I estimate the effects of the clean-indoor-air policy and the introduction of graphic tobacco warnings using hospital discharge data. I construct the prevalence rate for disease c for age group a as the ratio between the number of cases with diagnoses c in age group a and the total population aged a of state s in year t times 1,000. I run the following regression over the sample of treated and comparison states for the period 2008-2014 by cause of discharge c:

$$y_a = \sum_{\tau=-3}^{4} \delta_\tau \left[Treat_s \cdot (Years\ after\ treat = \tau) \right] + \Gamma' X_{st} + \alpha_a + \alpha_s + \upsilon t + \epsilon_a, \tag{3.2}$$

where $Years\ after\ treat = \tau$ is equal to the difference between the calendar year and the year the national law was passed. The variable $Treat_s$ is defined as before. X_{st} is a vector of control variables that comprise state-level characteristics. I include age-group fixed effects (α_a), and thus δ_τ is identified by within age-group differences between lenient and strict states over time. The linear time trend t controls for changes in population health over time that are constant across states. Standard errors are block-bootstrapped at the state level with 200 replications. The coefficients can be interpreted as a change in the prevalence rate of diagnosis c when the legislation was tightened.

The left (right) panel in figure 3.6 shows the estimated effect on COPD (lung cancer) prevalence rate for the population aged 18 to 65—Table A.40 presents point estimates and standard errors. I estimate that by 2014, the prevalence of COPD decreased about 3.6

[19] See Department of Health and Human Services, *The Health Consequences of Smoking—50 Years of Progress: A Report of the Surgeon General, 2014.*

percentage points relative to 2011, while respiratory system cancer decreased by about 2.2 percentage points. However, the latter result is not statistically different from zero. The effect for each year after the national law was passed is greater in absolute value than the previous one, which suggests that the policy could have important long-term effects.[20] The results on health outcomes could be driven by fewer individuals being diagnosed with these diseases; the same number of individuals being diagnosed with less severe symptoms and requiring fewer hospitalizations, or a combination of both. Although I can not shed light on which channel is the most important, my results suggest an improvement in population health correlated with stricter regulation of tobacco products.

My results are consistent with the fact that many health conditions associated with smoking—such as, heart disease, lung cancer, and overall premature mortality—are processes that develop over long periods of time making it difficult to know the appropriate lag between policy implementation and disease onset and progression (DeCicca, Kenkel and Lovenheim, 2022). Nonetheless, in the case of chronic obstructive pulmonary disease (COPD), it has been shown that smoking cessation and avoidance of cigarette smoke reduce the excess lung function decline (Eklund et al., 2012). Thus, the reduction on hospitalization from COPD could reflect the health gains from smoking cessation and reduced exposure to cigarette smoke induce by the law. In contrast, the null effects on cancer diagnoses reflect the fact that reverting the onset of this disease requires a longer time horizon.

6. Bans or Graphic Tobacco Warnings?

The implementation of Argentina's 2011 policy features a common characteristic across countries: The government implemented two anti-tobacco policies at the same time, i.e., the clean-indoor-air policy and the mandate for graphic tobacco warnings. In particular,

[20] Unfortunately, the Argentinean National Center for Health Statistics stopped collecting hospital discharge data after 2014, so conducting a longer term analysis is not feasible.

graphic warning labels have often been enacted as part of broader anti-smoking campaigns (DeCicca, Kenkel and Lovenheim, 2022). One example is the 2001 Canadian Federal Tobacco Control Strategy, which proposed raising tobacco taxes in addition to the incorporation of graphic tobacco warnings (Health Canada, 2002).[21] Another example of joint implementation of policies is the Taiwan Tobacco Hazards Prevention Amendment Act of 2009, which extended smoke-free areas to almost all enclosed workplaces and public places, added graphic health warnings to cigarette packages, and banned tobacco advertisement (see, Chang et al., 2010).

The joint-implementation setting poses two empirical challenges to the identification of policy effects. The first challenge is that researchers cannot exploit either regional or temporal variation to disentangle the effects of each policy because these are implemented at the same time. The second challenge is the presence of spillover effects between policies. If these are nonzero, the estimated effect of the policy bundle would be a function of the effects of each policy and the spillover effects. To overcome these identification challenges, I exploit an additional source of variation in my data and impose structure on how the two policies interact. This exercise helps in understanding what mechanisms are driving the extensive margin outcomes.

Effects of place-based bans. I exploit the fact that, while tobacco graphic warnings were implemented nationally in 2011, place-based bans were adopted by some states before 2011. Thus, the comparison between states that have implemented place-based bans before the national implementation—bans early adopters—and the states that have not implemented bans until the national law—bans late adopters—provides useful variation to identify the effect of place-based bans.[22] Under the assumption that conditional on time and fixed effects

[21] Beginning in April 2001, the federal government implemented a sequence of tax hikes. The excise tax was first raised to $10.99 per carton in May 2001, and then to $12.62 by the end of 2001. In mid-2002, the federal tax was further raised to $13.86 per carton and then to $15.85 in July 2002 (Gabler and Katz, 2010).

[22] Importantly, the group of states that adopted bans before the national implementation is not the same

the two groups of states would have followed the same trend absence the 2011's federal policy; the following regression provides a causal estimate of δ^{bans}:

$$y_i = \sum_{\tau=-2}^{1} \delta_{\tau}^{bans} \left[Bans_s \cdot (Years\ after\ treat = \tau) \right] + \beta' X_{is} + \Gamma' X_{st} + \alpha_s + \alpha_t + \varphi_i, \qquad (3.3)$$

where $Bans_s$ is a dummy variable that equals 1 if the state s is a late-adopter of place-based bans; and the rest of the parameters are defined as in equation 3.1. Figure 3.7 shows the results of this estimation and replicates the main results for comparison.[23] I find no effect of place-based bans on the proportion of non-smokers, which highlights the fact that non-smokers do not experience the cost of place-based bans—i.e., non-smokers are not compelled to go outside to smoke. On the other hand, roughly 40% of the change in the proportion of current smokers can be attributed to place-based bans, suggesting that the reduction in the probability of being a current smoker is driven by both components of the policy.

Effects of graphic tobacco warnings. To construct estimates of the effects of graphic tobacco warnings, I make two additional assumptions (i) I assume that treatment effects are constant across states and time, and (ii) I assume that the effects of each policy branch are additive. Under these assumptions, the total effect of the policy (δ) can be written as a linear combination of the effects of each branch of the policy: tobacco graphic warnings (δ^{tgw}) and place-based bans (δ^{bans}):[24]

$$\delta = \delta^{tgw} + \omega\ \delta^{bans}\ , \qquad (3.4)$$

where ω is the proportion of states that have not implemented a place-based ban before the national regulation, and δ^{bans} is the effect of place-based bans, and estimates of these effects

as those states classified as strict states. These two groups differ because the legislation index has multiple dimensions. Thus, a strict state could have ban cigarette sales, advertising, and consumption in schools and other venues but not in restaurants. This state will be a control state in the regression that estimates δ but a treated state in the regression that estimates $\tilde{\delta}^{bans}$.

[23] Appendix Table A.41 presents point estimates and standard errors.

[24] See Appendix 3.5 for a detailed derivation of this equation.

can be used to construct estimates of the effect of graphic tobacco warnings. The left panel of Figure 3.8 shows that graphic tobacco warnings are an effective policy instrument to deter smoking initiation. I find that most of the effect on the probability of being a never-smoker is explained by the introduction of these warnings. The rationale behind pictorial warnings on tobacco products is to inform smokers and non-smokers about the documented health risks of smoking. When non-smokers were asked about graphic warnings, 73% reported having seen them in the last 30 days. Studies in experimental settings have documented that graphic warnings elicit emotional responses, such as fear or disgust (DeCicca, Kenkel and Lovenheim, 2022).

My results indicate that these emotional responses may be an effective tool to keep non-smokers away from cigarettes. For current smokers, around 60% of the effect can be attributed to the incorporation of tobacco graphic warnings (Figure 3.8, right panel). I estimate that graphic warnings reduce the probability of being a current smoker by 3.85 percentage points, these results are in line with Kuehnle (2019) who finds that the introduction of pictorial warnings on cigarette packages reduced smoking by around 4% within the first year of the policy in Australia.

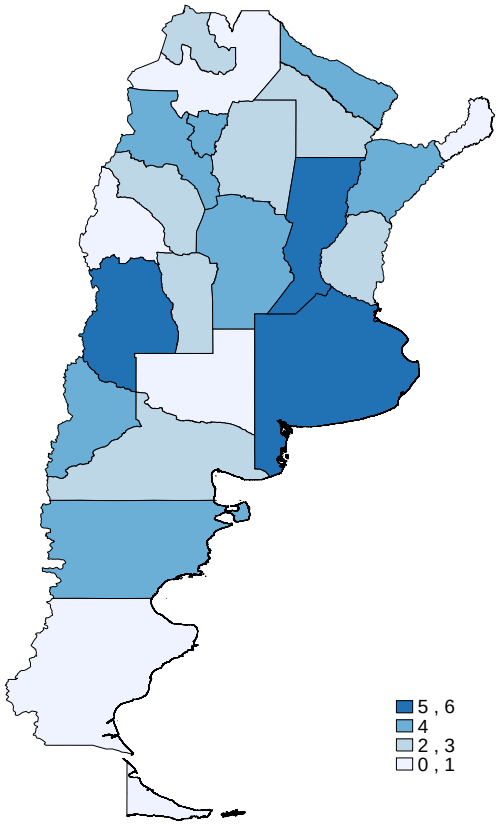
Interpretations. The previous exercise helps in understanding what mechanisms are driving the extensive margin outcomes; nonetheless, some caveats are in place. First, spillover effects may arise from complementarities in policy intervention and these effects can not be disentangled from the effect of place-based bans. If these spillover effects are positive—for example, because the effectiveness of place-based bans is enhanced by the implementation of the advertising restriction as the latter helps to raise awareness of the side effects of smoking—the estimates of δ^{bans} would provide an upper-bound estimate of the true effect. Second, the derivation of Equation 3.4 relies on the assumption that effects are additive. If this where not the case, this expression would still provide a useful decomposition of the total effect but we would need to assume that interaction terms are small.

7. Conclusions and Discussion

Smoking bans and advertising regulations have played an increasingly important role in tobacco control policies over the past 20 years. These policies are particularly attractive in middle- and low-income countries, where tobacco use is concentrated among low-income households. This paper provides evidence of the effectiveness and potential limitations of these policies by studying the Argentinean case. Argentina's 2011 national law has two main components: (i) it implements indoor smoking bans and (ii) it regulates advertising of tobacco products and mandates the incorporation of graphic tobacco warnings. I exploit regional variation in the leniency of tobacco regulation before 2011 to identify the effects of this new regulation on various outcomes. I find that the new regulations effectively curbed smoking initiation and consumption. Nonetheless, when looking at heterogeneous effects, my results suggest that the policy disproportionately benefited more educated and richer individuals, and thus potentially widens health disparities between these groups. This paper also highlights the incorporation of graphic tobacco warnings as a valuable policy instrument. Well-designed package warnings are a highly cost-effective means for increasing awareness of the smoking's effects on health and deterring individuals from smoking.

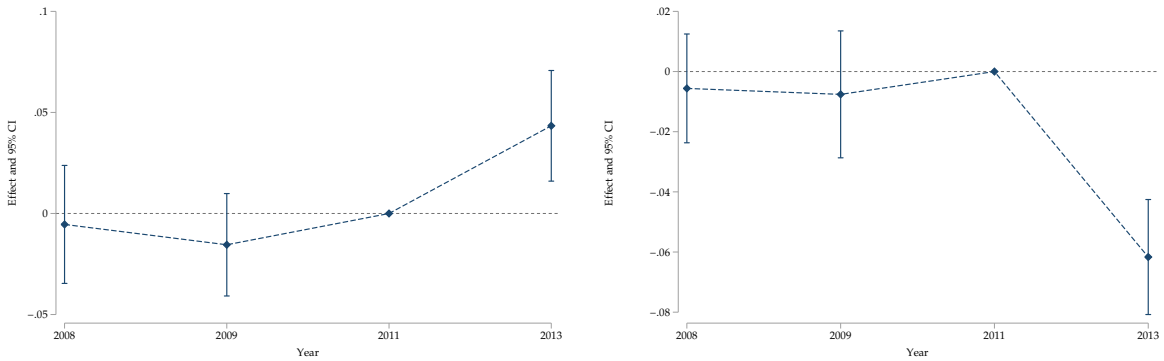
8. Figures

Figure 3.1: Legislation Index in 2009



Notes: The legislation index summarizes the strength of regulations at the state level; higher index values (darker shades) indicate stricter regulation and lighter shades indicate more lenient regulation. See Appendix 3.3 for details on construction of the legislation index.

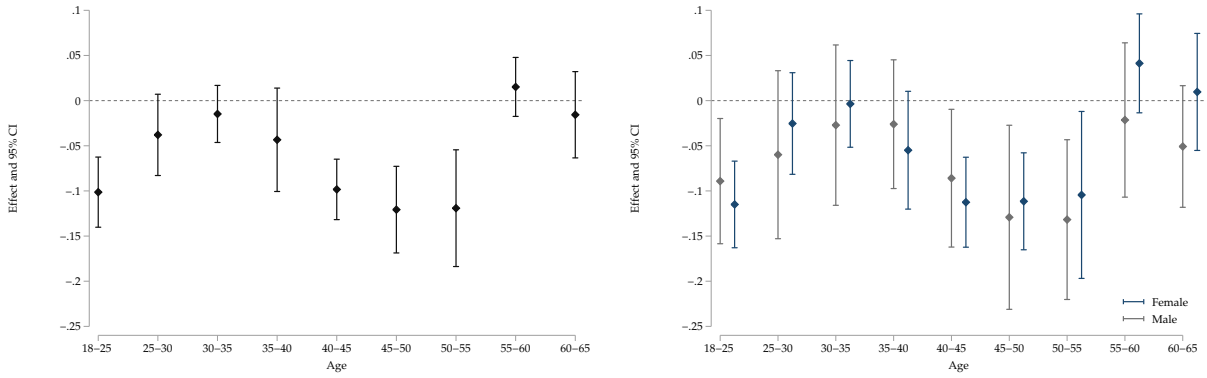
Figure 3.2: Effects on Extensive Margin Outcomes



(a) Prob. Never Smoker (b) Prob. Current Smoker

Notes: This figure presents point estimates and symmetric percentile-t confidence intervals of the causal effect of the national regulation on the probability of being a never smoker (left panel) and the probability of being a current smoker (right panel). The omitted year corresponds to 2011, the year the law was passed. Standard errors are block-bootstrapped at the state level with 200 replications.

Figure 3.3: Extensive Margin Effects by Age and Gender

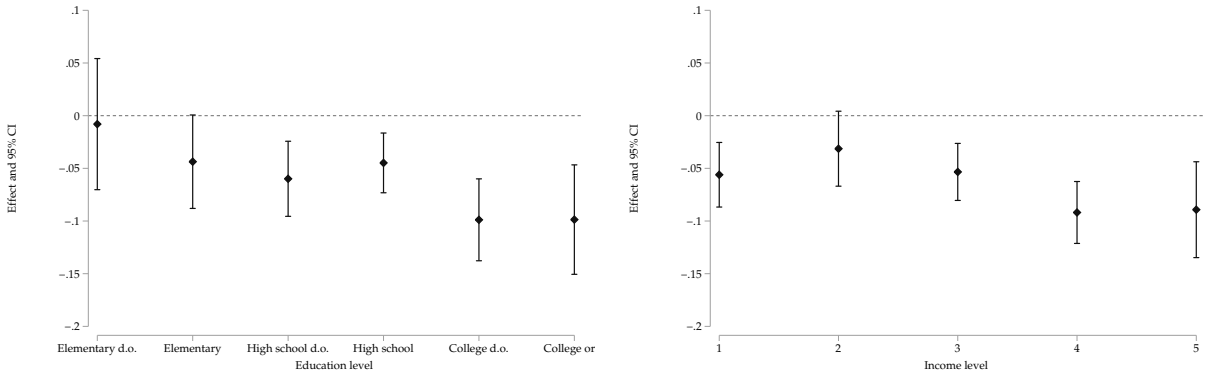


(a) Age groups

(b) Age-gender groups

Notes: This figure presents point estimates and symmetric percentile-t confidence intervals of the causal effect of the national regulation on the probability of being a smoker in 2013 by subgroups. Standard errors are block-bootstrapped at the state-level with 200 replications.

Figure 3.4: Extensive Margin Effects by Education and Income

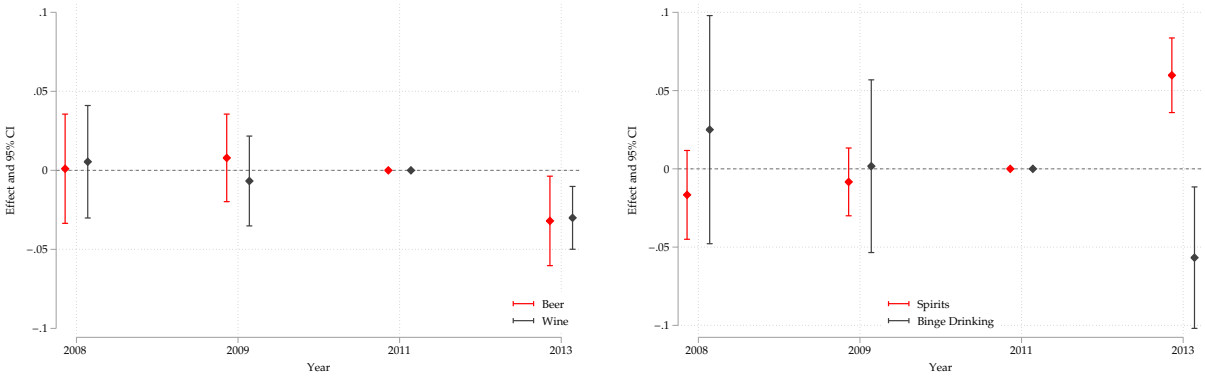


(a) Educational attainment

(b) Income

Notes: This figure presents point estimates and symmetric percentile-t confidence intervals of the causal effect of the national regulation on the probability of being a smoker in 2013 by subgroups. Standard errors are block-bootstrapped at the state-level with 200 replications. Panel (a) presents results by educational attainment, in which d.o. stands for drop-out; panel (b) presents results by household income ranking. The 1^{st} quintile corresponds to the lowest 20% of households in the income distribution. Analogously, the 5^{th} quintile corresponds to the highest 20% of households in the income distribution.

Figure 3.5: Effects on the Risky Consumption of Alcoholic Beverages

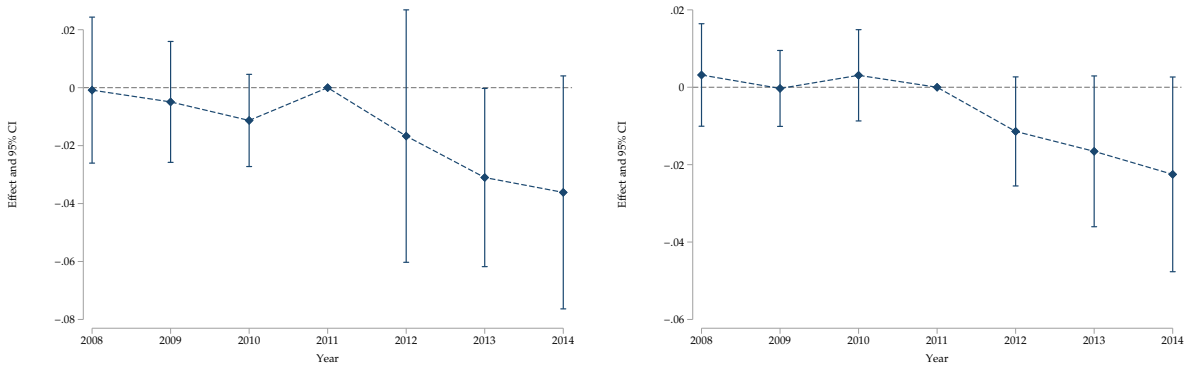


(a) Beer and Wine

(b) Spirits and Binge Drinking

Notes: This figure presents estimates of the marginal effects and confidence intervals of the 2011 national law on the probability of abusive consumption of beer, wine, and spirits and binge drinking across smokers. Abusive consumption thresholds and binge drinking are defined in the main text. The omitted category corresponds to 2011, the year the federal law was passed. Standard errors are clustered at the state level.

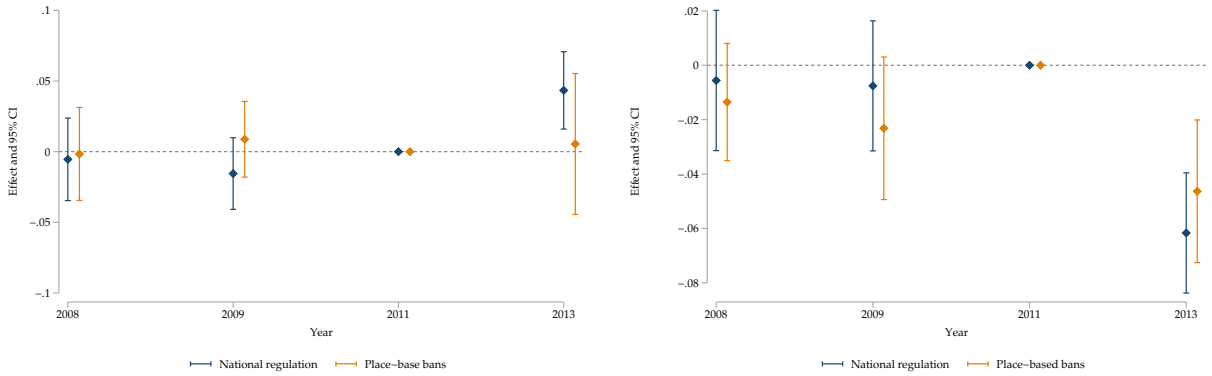
Figure 3.6: Effects on Health Outcomes



(a) COPD (b) Lung cancer

Notes: This figure presents point estimates and symmetric percentile-t confidence intervals of the causal effect of the national regulation on the prevalence rate of COPD (left panel) and lung cancer (right panel). The omitted year corresponds to 2011, the year the law was passed. Standard errors are block-bootstrapped at the state level with 200 replications.

Figure 3.7: Effects of Place-based Bans

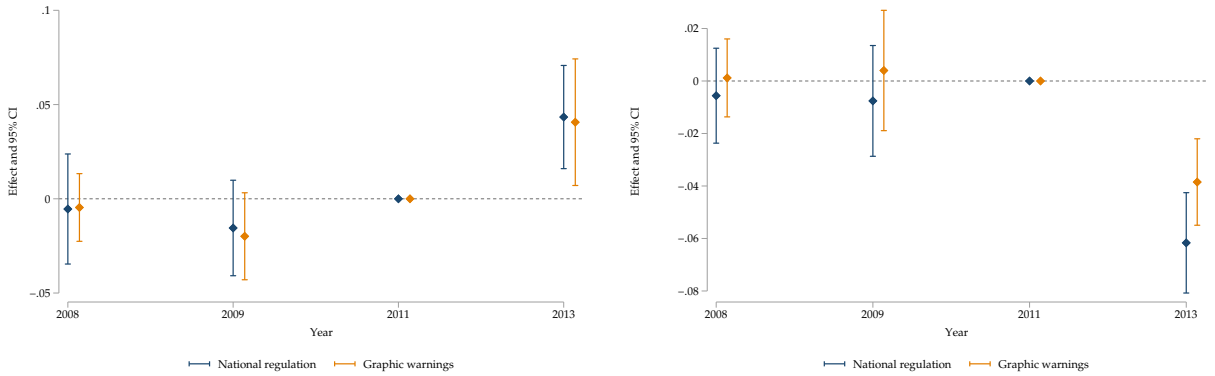


(a) Prob. never smokers

(b) Prob. current smokers

Notes: This figure presents point estimates and symmetric percentile-t confidence intervals of the causal effect of the national regulation and place-base bans on the probability of being a never smoker (left panel) and the probability of being a current smoker (right panel). The omitted year corresponds to 2011, the year the law was passed. Standard errors are block-bootstrapped at the state level with 200 replications.

Figure 3.8: Effects of Graphic Tobacco Warnings



(a) Prob. never smokers

(b) Prob. current smokers

Notes: This figure presents point estimates and symmetric percentile-t confidence intervals of the causal effect of the national regulation and tobacco graphic warnings on the probability of being a never smoker (left panel) and the probability of being a current smoker (right panel). The omitted year corresponds to 2011, the year the law was passed. Standard errors are block-bootstrapped at the state level with 200 replications.

9. Tables

Table 3.1: Summary Statistics, Pre-policy Period

Variable (mean)	All states	Strict states	Lenient states	Equality of means
Never-smokers	0.560	0.561	0.558	0.777
Current smokers	0.286	0.281	0.290	0.587
Mean cigarettes per day				
0 to 5	0.373	0.397	0.352	0.024
6 to 10	0.266	0.260	0.270	0.513
11 to 20	0.297	0.282	0.310	0.073
more than 20	0.064	0.060	0.068	0.382
Prevalence rate (per 1,000)				
COPD	0.203	0.196	0.209	0.765
Lung cancer	0.069	0.061	0.076	0.330
Number of states	24	11	13	

Notes: This table presents summary statistics for the outcomes of interest for the pre-policy period. I pooled 2008 and 2009 for consumption outcomes and 2008 to 2010 for health outcomes. The last column reports the p-value for the test of equality of means for strict and lenient states. The category more than 20 cigarettes per day includes smokers with consumption strictly greater than 20 and less than or equal to 50 cigarettes a day.

Table 3.2: Balance in Terms of Individual Characteristics in Strict versus Lenient States

Variable (mean)	All states	Strict states	Lenient states	Equality of means
Average age	37.719	37.745	37.698	0.896
Male	0.491	0.495	0.487	0.581
Young (< 25 years old)	0.236	0.239	0.234	0.749
Married or cohabitant	0.579	0.570	0.587	0.316
Employed	0.701	0.702	0.700	0.878
Educational level				
Elementary school	0.196	0.195	0.196	0.973
High school	0.237	0.234	0.239	0.598
College	0.146	0.150	0.141	0.451
Income category				
First quintile	0.134	0.131	0.136	0.605
Second quintile	0.294	0.355	0.242	0.000
Third quintile	0.214	0.199	0.227	0.038
Forth quintile	0.222	0.181	0.257	0.000
Fifth quintile	0.136	0.135	0.137	0.876
Observations	60,449	23,830	36,619	
Number of states	24	11	13	

Notes: This table presents summary statistics for demographic characteristics for the pre-policy period. I pooled 2008 and 2009. Demographic characteristics other than age indicate the proportion of individuals with a given attribute. The last column reports the p-value for the test of equality of means for strict and lenient states.

Table 3.3: Effects on Intensive Margin Outcomes

Cigarettes smoked	(0,5]	(5, 10]	(10, 15]	(15, 20]	(20, 30]	(30 , 40]	(40 , 50]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
2008	-0.0201	0.0225	-0.0006	0.0003	0.0015	-0.0042	0.0006
	(0.0293)	(0.0161)	(0.0085)	(0.0143)	(0.0077)	(0.0064)	(0.0009)
2009	-0.0052	0.0603	0.0008	-0.0379*	-0.0097*	-0.0101*	0.0018
	(0.0276)	(0.0227)	(0.0092)	(0.0201)	(0.0047)	(0.0052)	(0.0018)
2013	-0.1652***	0.0584*	0.0102*	0.0786***	0.0216	-0.0035	-0.0001
	(0.0295)	(0.0141)	(0.0087)	(0.0150)	(0.0050)	(0.0035)	(0.0009)
Mean dep. var.	0.3500	0.2708	0.0966	0.2143	0.0385	0.0272	0.0024
Observations	40,651	40,651	40,651	40,651	40,651	40,651	40,651
R-squared	0.0716	0.0108	0.0071	0.0324	0.0138	0.0181	0.0017
Correctly predicted	0.6242	0.6949	0.8627	0.7526	0.9077	0.9181	0.9398

Notes: This table presents estimated effects on cigarettes smoked in the last 30 days. The omitted category corresponds to 2011, the year the federal law was passed. All regressions include individual-level controls: age, gender, educational attainment, employment status, and income category of the household. State-time varying controls include total private employment and total population. All regressions include state and time fixed effects. Correctly predicted indicates the percentage of times the predicted outcome matches the actual outcome. Standard errors are block-bootstrapped at the state-level with 200 replications. *** $p<0.01$, ** $p<0.05$, * $p<0.10$

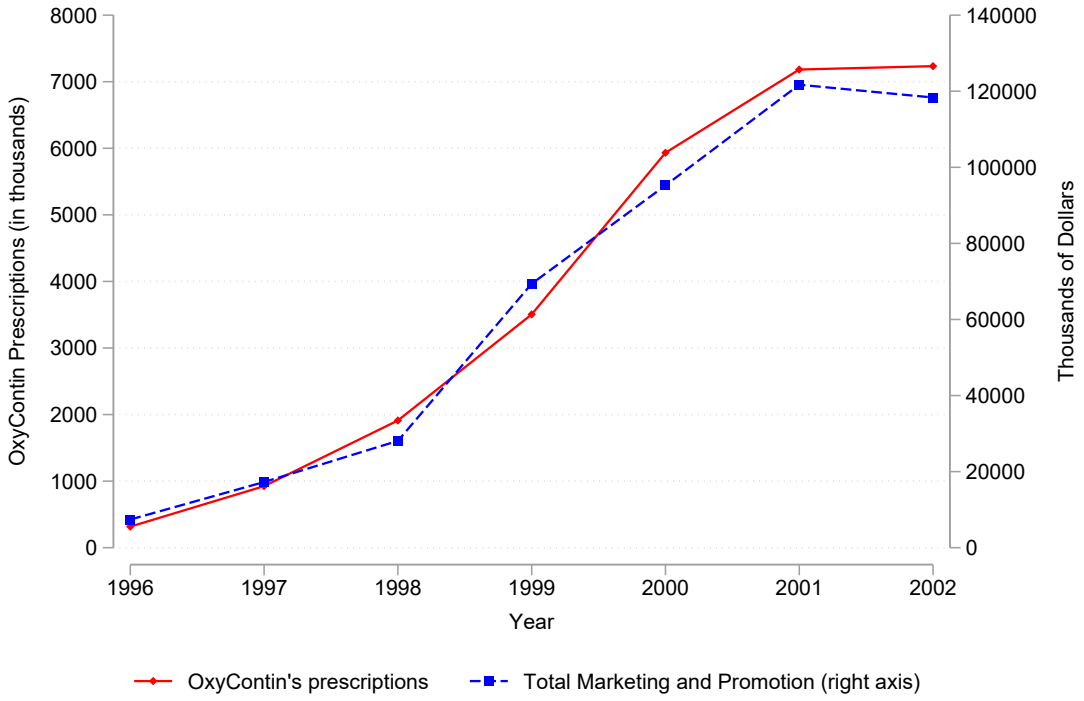
Chapter 4

Appendix and Supplementary Material

1. Appendix to "A Manufactured Tragedy: The Origins and Deep Ripples of the Opioid Epidemic"

1.1 Additional Figures

Figure A.1: OxyContin Marketing Budget and Total Prescription Sales



Notes: Author's constructions based on OxyContin Budget Plans 1998-2002 and *United States General Accounting Office (GAO). Prescription Drugs: OxyContin Abuse and Diversion and Efforts to Address the Problem: Report to Congressional Requesters. 2003.* This figure is referenced in Section 2.

Figure A.2: Purdue Pharma Budget Plan 1997: Target Audiences

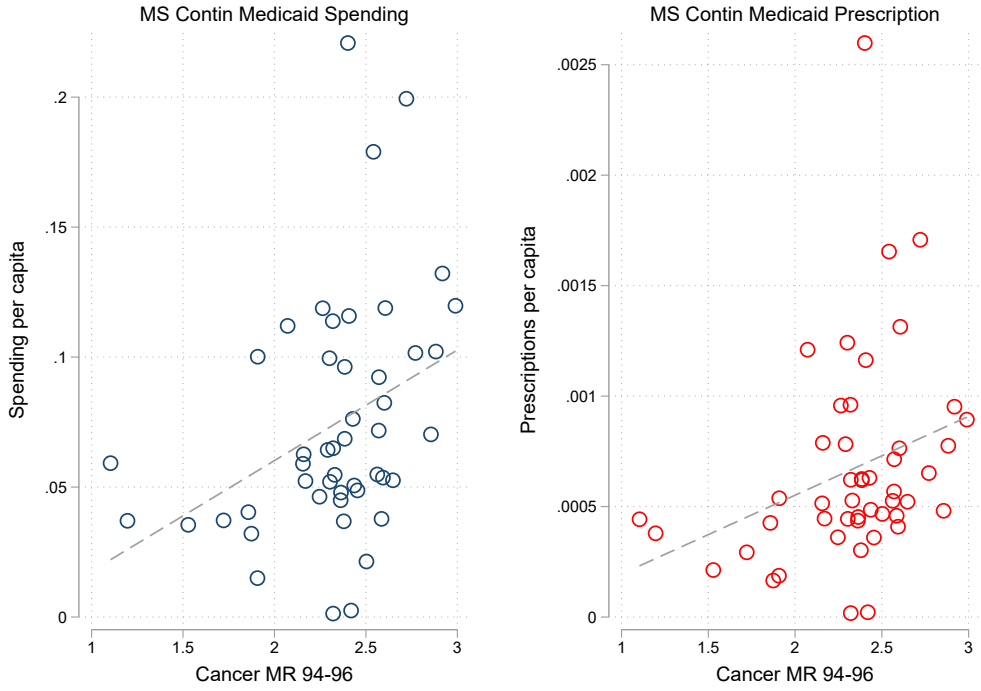
C. Target Audiences

1. Primary Audiences

Primary Audiences	Site	Targets	Comments
A. Physicians • ONCs Hem/Oncs Rad/Oncs • IMs • FP/GPs • DOs • ANS • Surg • Other	• Office and Hospital	13,000 7,600 33,000	Target List 1A Decile 8, 9, or 10 for "Strong" opioids who are also Decile 8, 9, or 10 for "Combo" opioids Target List 1B Decile 9 and 10 for "Strong" opioids only Target List #2 Decile 10 for combo only but not in Target List 1A; non-malignant market

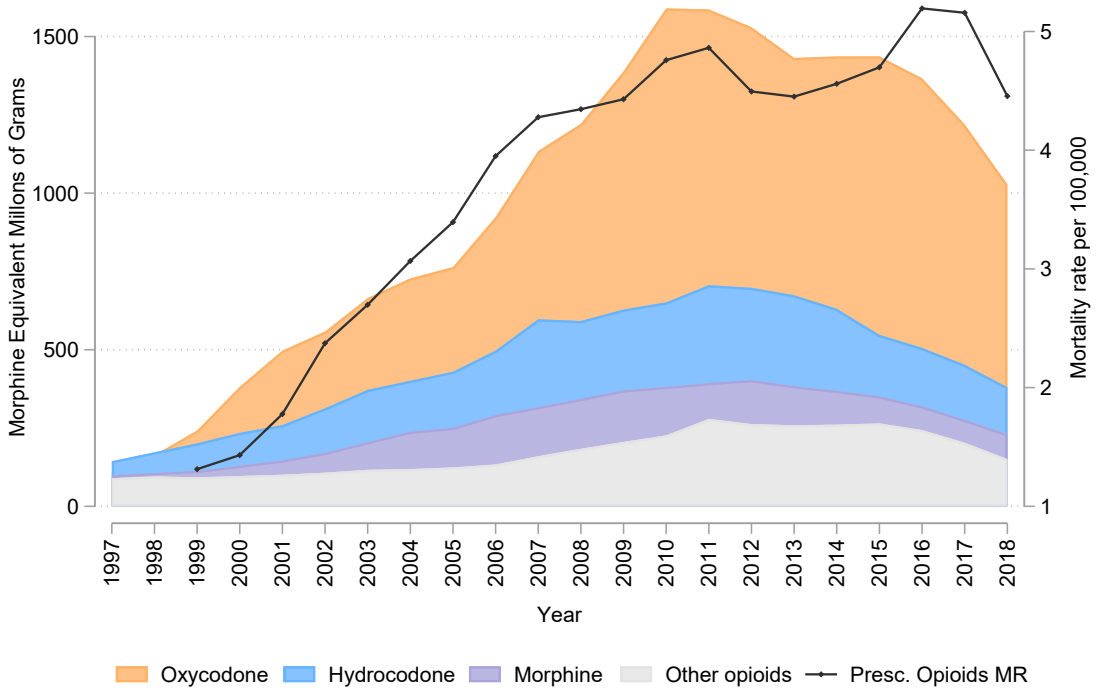
Notes: This figure is an extract of Purdue Pharma marketing plan. It shows that Purdue marketing targeted top opioid prescribers. Purdue Pharma Budget Plan 1997, p.25. This figure is referenced in Section 2.

Figure A.3: Mid-nineties Cancer Mortality Rate and 1994 MS Contin prescriptions



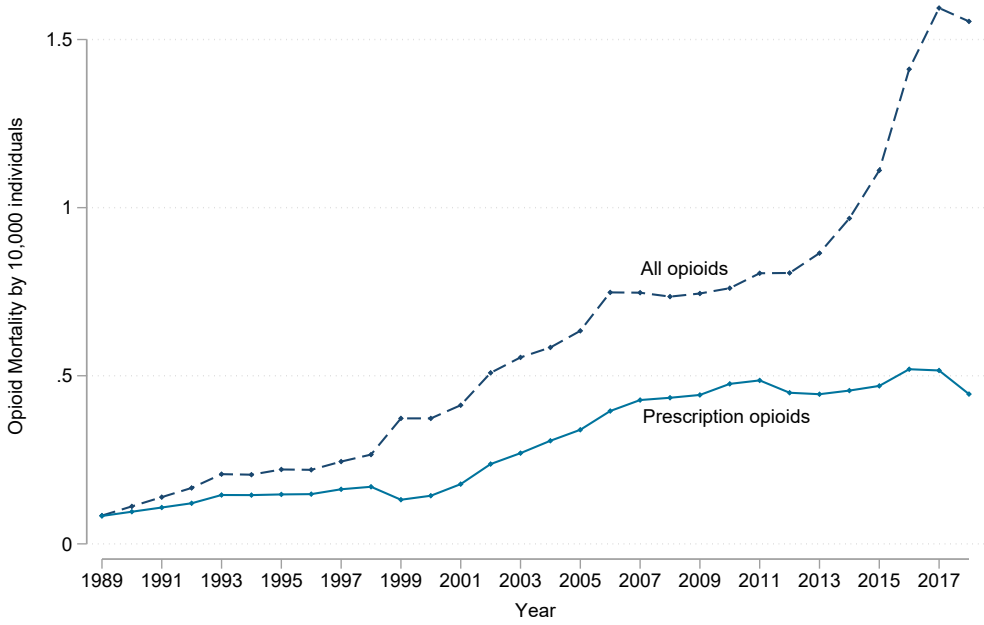
Notes: This figure shows the relationship between MS Contin prescription rates prior to the launch of OxyContin and mid-nineties cancer mortality. Source: CMS- Medicaid State Drug Utilization. This figure is referenced in Section 2.

Figure A.4: Evolution of Prescription Opioid Distribution



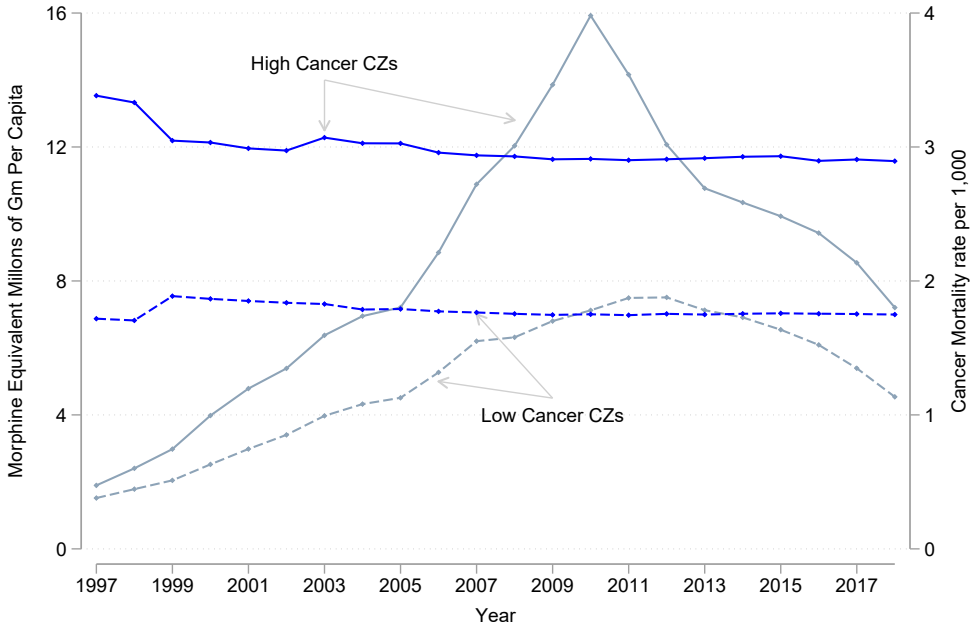
Notes: This figure shows the evolution of shipments of all prescription opioids and the three main components: oxycodone, hydrocodone and morphine. Oxycodone is the active ingredient of OxyContin. Shipments of prescription opioids are expressed in morphine-equivalent doses. Data on opioids distribution come from the ARCOS. The mortality rate (MR) from prescription opioids is constructed using data from the National Vital Statistic System and plotted in the right-hand-side axis. Details on the construction of this measure are found in 3.3 This figure is referenced in Section 3.3.1

Figure A.5: Evolution of Prescription Opioid and All Opioid Mortality Rates



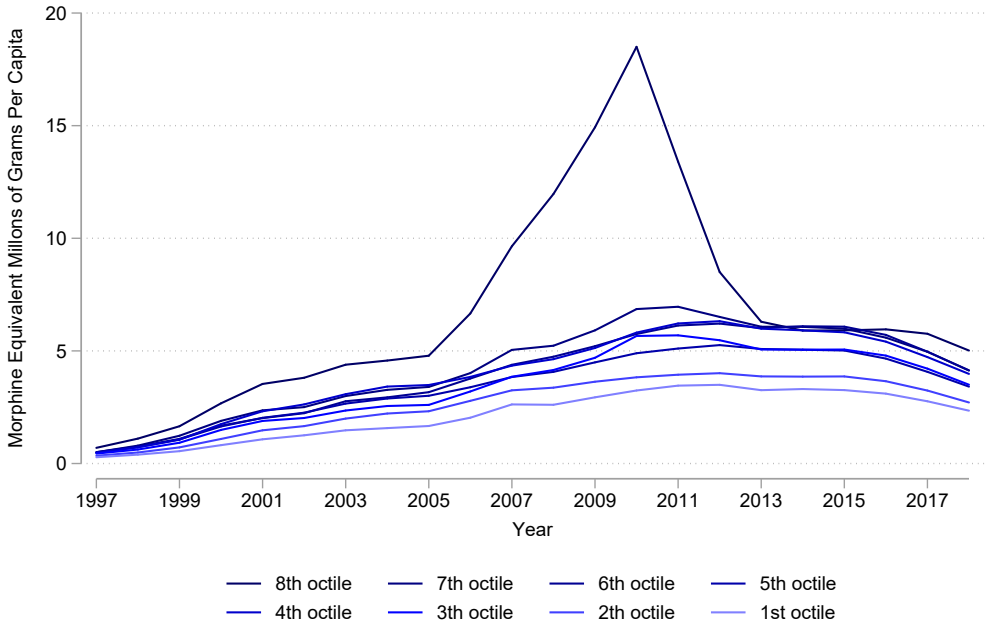
Notes: This figure shows the evolution of prescription opioid and all opioid mortality rates from 1989 to 2018. The 1989-1998 data use ICD-9 codes to categorize the cause of death, and the 1999-2018 data use ICD-10 codes. The time series show that the transition from ICD-9 to ICD-10 classifications resulted in a small increase in poison-related deaths; this is consistent with what the CDC reports (Warner et al., 2011). This figure is referenced in Section 3.3.3

Figure A.6: Evolution of Cancer Mortality and Prescription Opioid Supply



Notes: This figure shows the evolution of prescription opioids (light blue lines in the left-hand axis) and cancer mortality rates (dark-blue lines in the right-hand axis) over time for commuting zones in the top and bottom quartiles of the cancer mortality distribution. Areas in the top quartile of the cancer distribution experienced an influx of opioids that was up to 3 times larger than the one experience by areas in the bottom quartile. Changes in cancer mortality does not explain this discrepancy; trends in carcer mortality rates in these groups of commuting zones suggest that mortality was quite stable in the period. Prescription opioids is measured in morphine-equivalent mg. This figure is referenced in Section 4.4.1

Figure A.7: Evolution of Oxycodone by Octiles of the 1994-1996 Cancer Prevalence



Notes: This figure shows the evolution of oxycodone in eight groups of commuting zones. Each group is composed of those commuting zones in the *n-th* octile of the cancer mortality rate distribution before the launch of OxyContin. Darker colors indicate groups with higher cancer prevalence (e.g., the 8th octile corresponds to the series that peaked in 2010 at 19 morphine-equivalent millions of gm per capita). Lighter colors indicate groups with lower cancer prevalence. This figure is referenced in Section 4.4.1

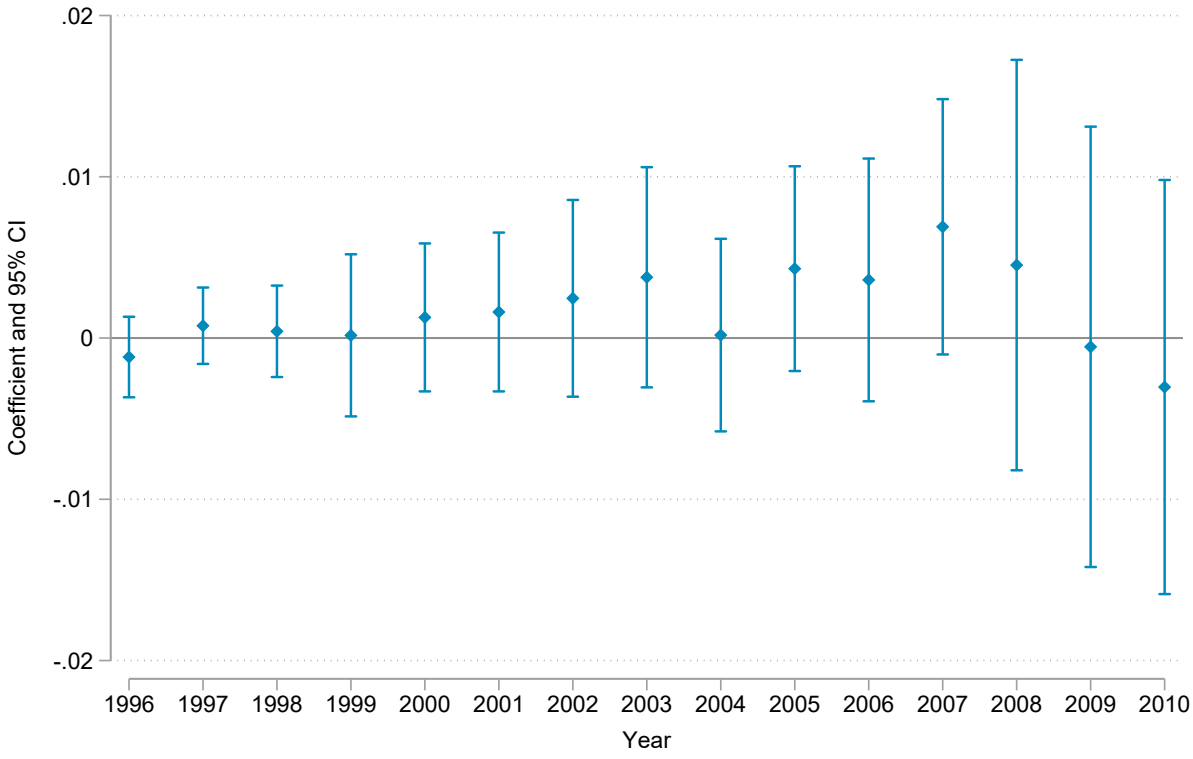
Figure A.8: Extract Exhibit 1 - Sales Visits By Purdue In Massachusetts

Commonwealth of Massachusetts v. Purdue Pharma et al.
Exhibit 1 - Sales Visits By Purdue In Massachusetts

Date	Purdue Sales Rep	Target	Address	City
5/16/2007	Raczkowski, Paula J	Belezos, Elias	164 South Street Harrington Hospital Medical Arts Bld	Southbridge
5/16/2007	Raczkowski, Paula J	Jeznach, Gary	118 Main St Rte 131	Sturbridge
5/16/2007	Raczkowski, Paula J	Welch, Heidi	118 Main Street	Sturbridge
5/16/2007	Raczkowski, Paula J	Keaney, Stephanie	100 South Street Medical Arts Bldg Suite#201	Southbridge
5/16/2007	Raczkowski, Paula J	Kereshi, Stjepan	100 South St Harrington Hospital	Southbridge
5/16/2007	Raczkowski, Paula J	Litani, Vladas	100 South St Harrington Mem Hos	Southbridge
5/16/2007	Raczkowski, Paula J	CVS Pharmacy Southbridge	380 Main St	Southbridge
5/16/2007	Mulcahy, Maurice	Boyd, Kenneth	23 Whites Path	South Yarmouth
5/16/2007	Mulcahy, Maurice	Hartley, Marie	23 Whites Path	South Yarmouth
5/16/2007	Mulcahy, Maurice	Terrill, Donna	269 Chatham Rd	Harwich
5/16/2007	Mulcahy, Maurice	Fair, Dianne	269 Chatham Rd	Harwich
5/16/2007	Mulcahy, Maurice	CVS Patriot Square-So. Dennis	Paqtriot Square/Route 134	S Dennis
5/16/2007	Mulcahy, Maurice	CVS Rt 28 S.Yarm.	976 Route 28 Yarmouth Shopping Plaza	S Yarmouth
5/16/2007	Arias, Alexander	Huang, Wynne	1 Roosevelt	Peabody
5/16/2007	Arias, Alexander	CVS - Main St [Woburn]	415 Main St	Woburn
5/16/2007	Ritter, Andrew	Kehlman, Glenn	637 Washington St	Brookline

Notes: Exhibit 1 - Sales Visits By Purdue In Massachusetts. COMMONWEALTH OF MASSACHUSETTS v.PURDUE PHARMA C.A. No. 1884-cv-01808. This figure is referenced in Section 4.4.1

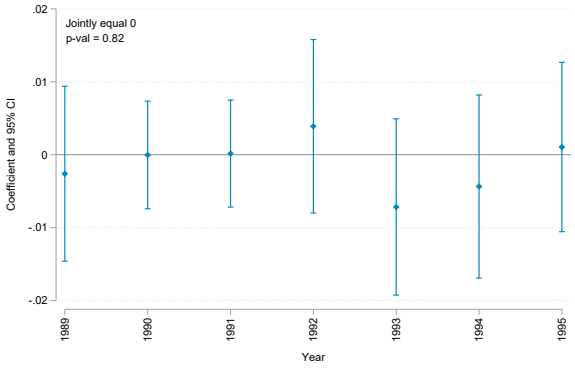
Figure A.9: Effects of Mid-nineties Cancer-market Targeting on Share of Smokers



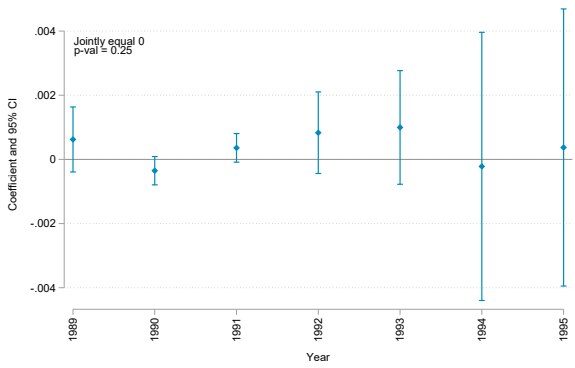
Notes: This figure shows the effects of the increase in prescription opioids supply in the share of smokers. We present the results of a dynamic reduced-form estimation were we regress the outcome on a set of year-dummy variables interacted with our instrument. We construct the share of smokers using data from the Behavioral Risk Factor Surveillance System (BRFSS). We perform the analysis up to 2010 since starting in 2011, BRFSS changed its data collection, structure, and weighting methodology. In 2011 there is an increase in the proportion of people being surveyed on cell phones and it also coincides with a rise in the percentage of respondents with unknown smoking status as documented by DeCicca, Kenkel and Lovenheim (2022). This figure is referenced in Section 4.4.2 and Section 5.5.2

Figure A.10: Dynamic Reduced Form Estimates - Out-of-sample Analysis

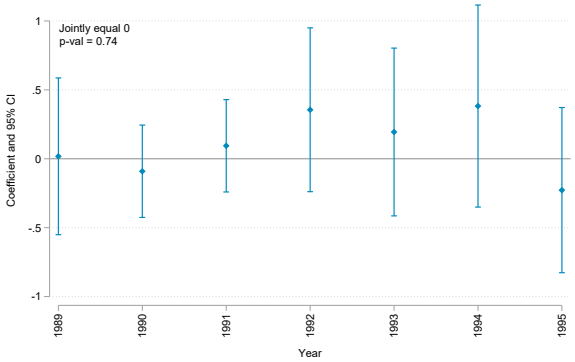
(a) Deaths of Despair



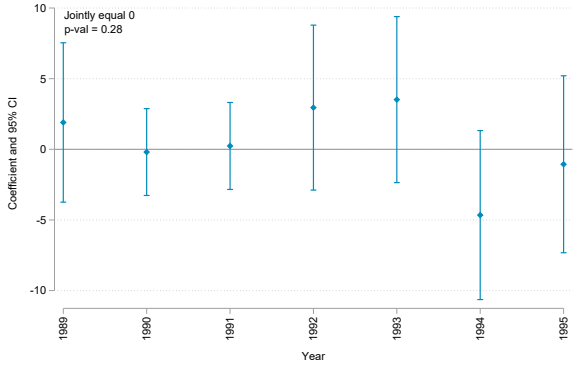
(b) SNAP



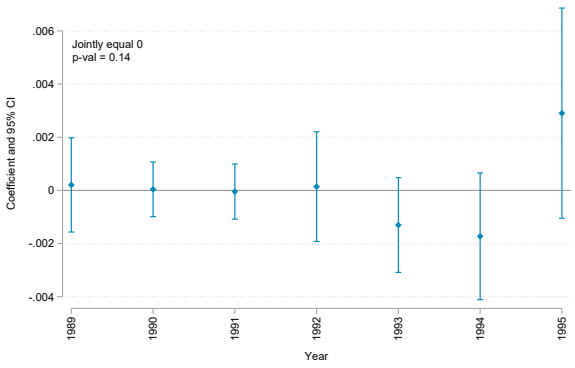
(c) Infant Mortality Rate



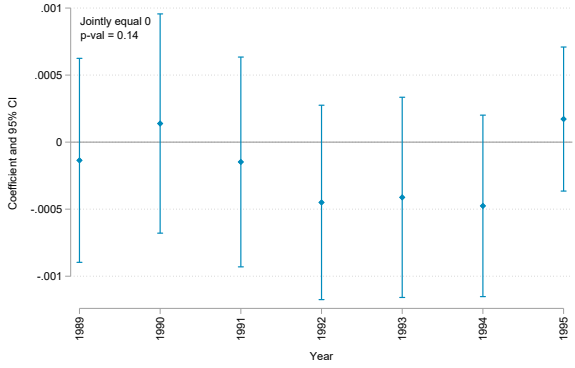
(d) Birth Weight



(e) Fertility Rate - 25 to 29 years old



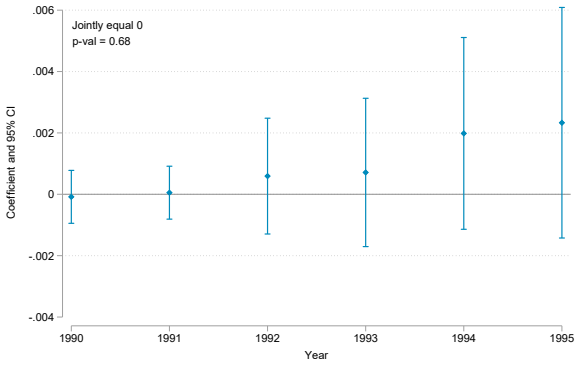
(f) Fertility Rate - unmarried woman



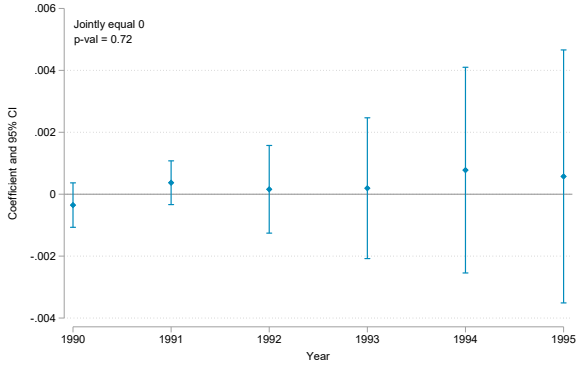
Notes: This figure shows the dynamic reduced-form relationship between outcomes of interest and our instrument in an out-of-sample period. That is, we replicate our dynamic reduced-form analysis in the pre-OxyContin period. We regress each outcome on a set of year-dummy variables interacted with the out-of-sample instrument—cancer mortality in 1989 - 1990. We use this specification to test for the presence a relationship between our outcomes and mid-nineties cancer mortality before the introduction of OxyContin. We do not reject the null hypothesis that the estimated coefficients are jointly equal to zero, the p value of these tests are presented in the figures. This figure is referenced in Section 4.4.2

Figure A.11: Dynamic Reduced Form Estimates - Out-of-sample Analysis: Labor Market Outcomes

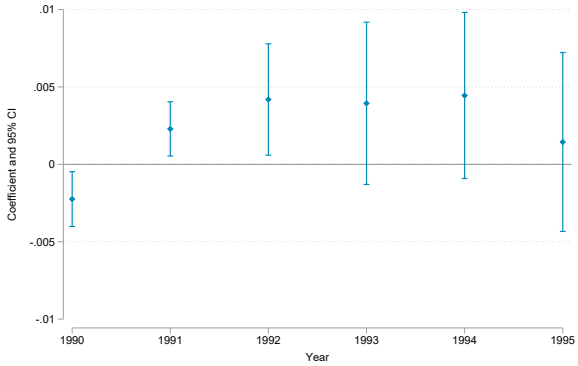
(a) Manufacturing



(b) Mining



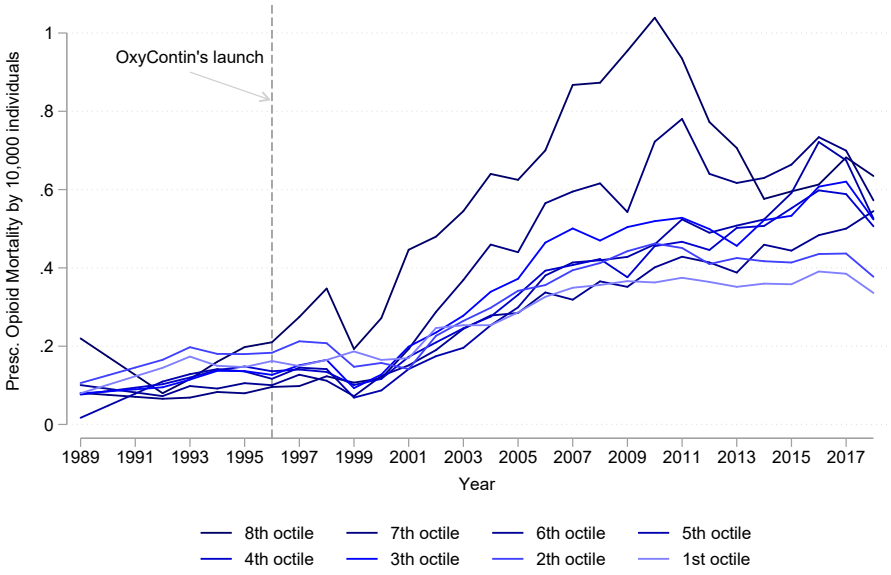
(c) Labor Force Participation



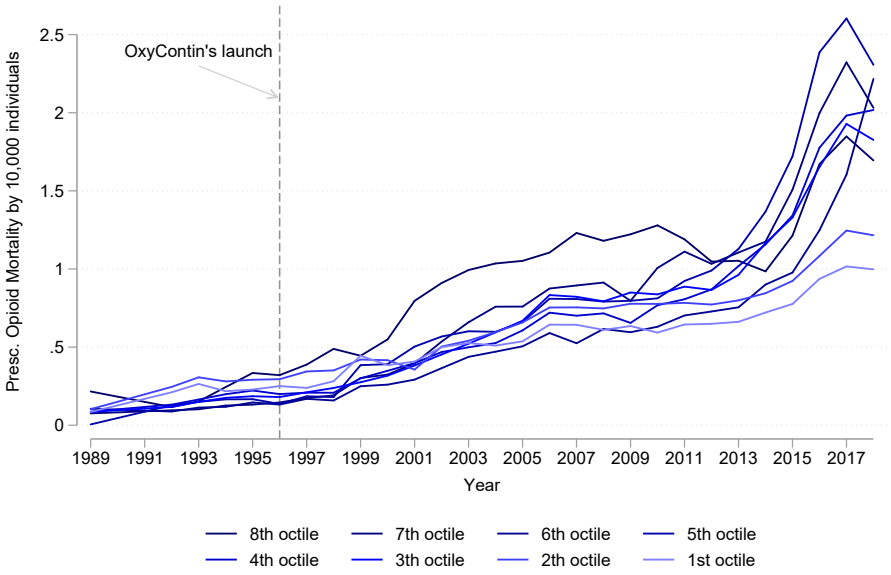
Notes: This figure shows the dynamic reduced-form relationship between the share of employment in the manufacturing and mining industries and labor force participation and our instrument in an out-of-sample period. The first year of available data is 1990. We use this specification to test for the presence a relationship between our outcomes and mid-nineties cancer mortality before the introduction of OxyContin. We do not reject the null hypothesis that the estimated coefficients are jointly equal to zero, the p value of these tests are presented in the figures. This figure is referenced in Section 4.4.2

Figure A.12: Opioid Mortality Rate by Octiles of the 1994-1996 Cancer Prevalence

(a) Prescription Opioids



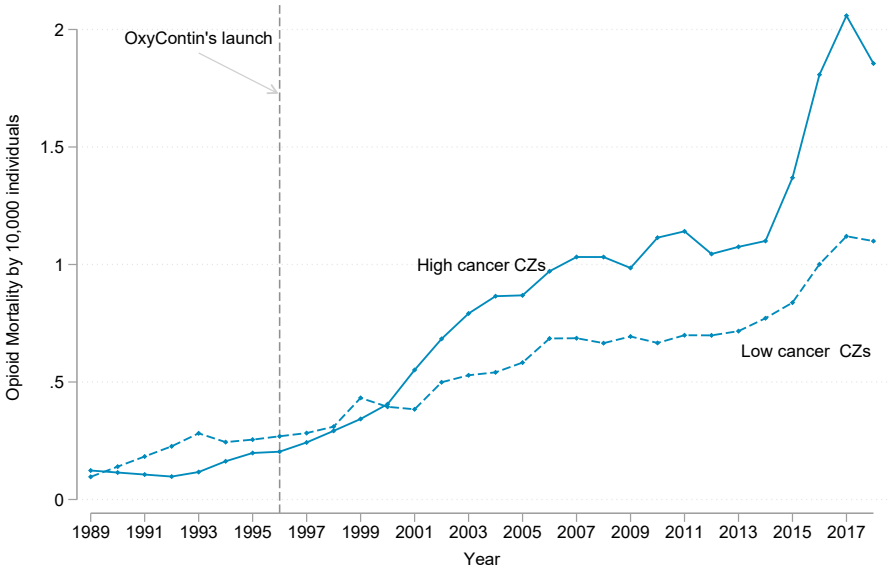
(b) All Opioids



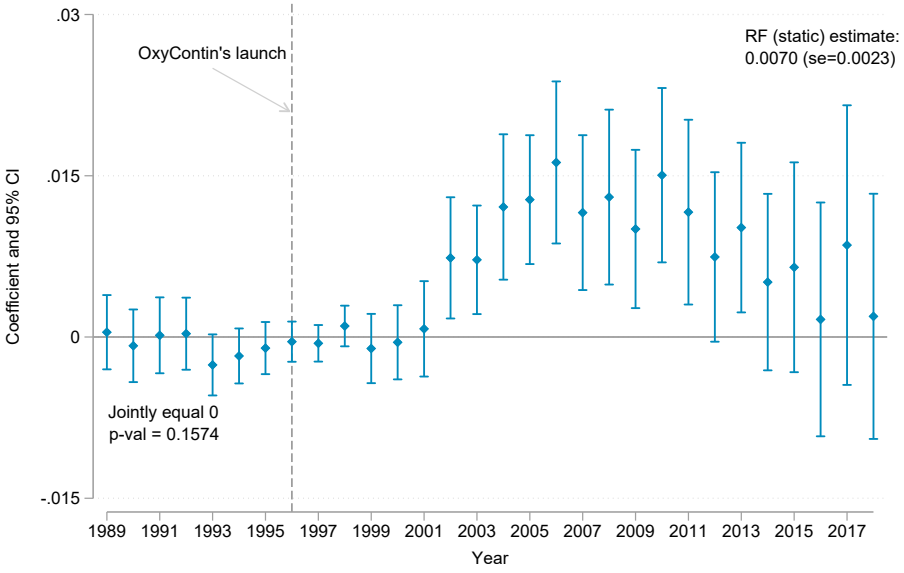
Notes: This figure shows the evolution of prescription opioid (panel a) and all opioids (panel b) mortality in eight groups of commuting zones. Each group is composed of those commuting zones in the *n-th* octile of the cancer mortality rate distribution before the launch of OxyContin. Darker colors indicate groups with higher cancer prevalence. Lighter colors indicate groups with lower cancer prevalence. This figure is referenced in Section 5.5.1

Figure A.13: Effects of Mid-nineties Cancer-market Targeting on All-Opioid Mortality

(a) High vs Low Cancer Mortality CZs



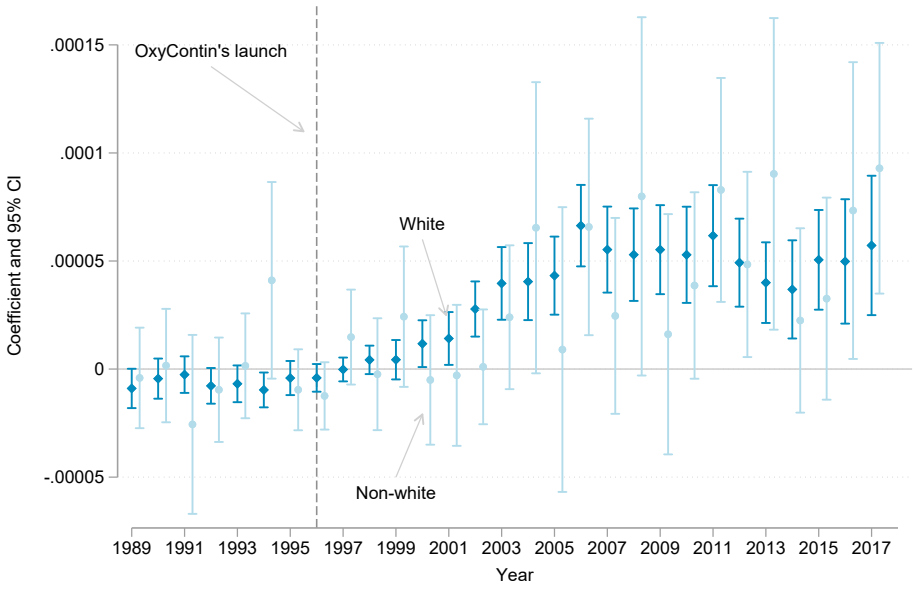
(b) Reduced Form - Event Study Approach



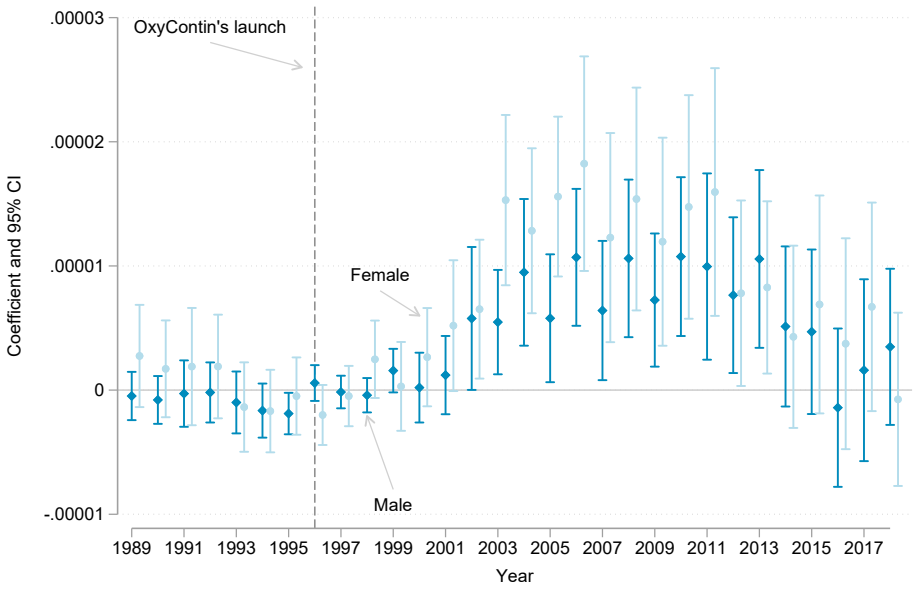
Notes: This figure shows the effects of the increase in prescription opioid supply in all-opioid mortality. Panel (a) shows the raw data, early in the 2000s, a wedge starts to appear between high- and low-cancer-incidence groups. Panel (b) shows the dynamic reduced-form estimation. We regress all-opioid mortality on a set of year-dummy variables interacted with our instrument—cancer mortality in 1994-1996. These coefficients corresponds to the estimate of ϕ_t in Equation 2.3. We use this specification to test for the presence of pre-trends in the relation between opioid mortality and mid-nineties cancer mortality; we do not reject the null hypothesis that the estimated coefficients before 1996 are jointly equal to zero, the p value of this test equals 0.1574. This figure is referenced in Section 4.4.2, in Section 5.5.1, and in Section 5.5.2

Figure A.14: Effects of Mid-nineties Cancer-market Targeting on Opioid Mortality by Race and Gender

(a) Race



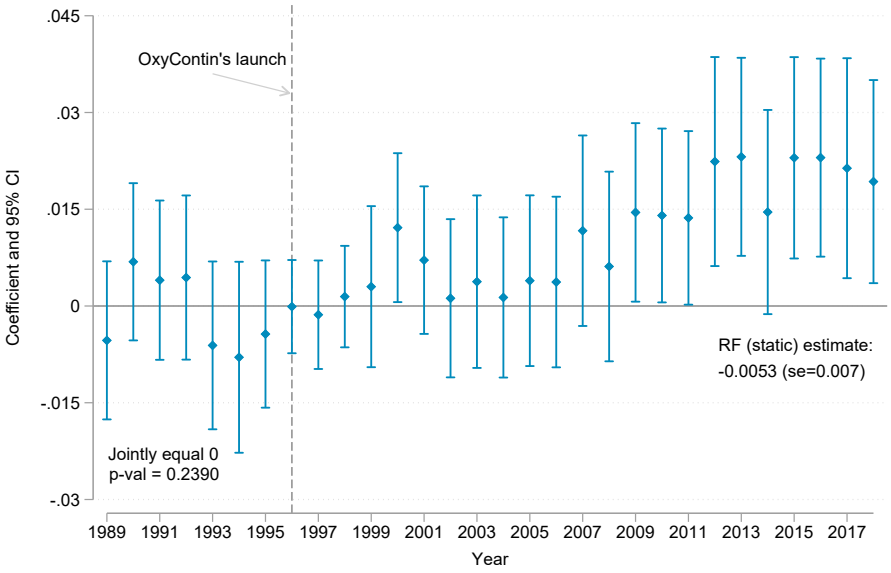
(b) Gender



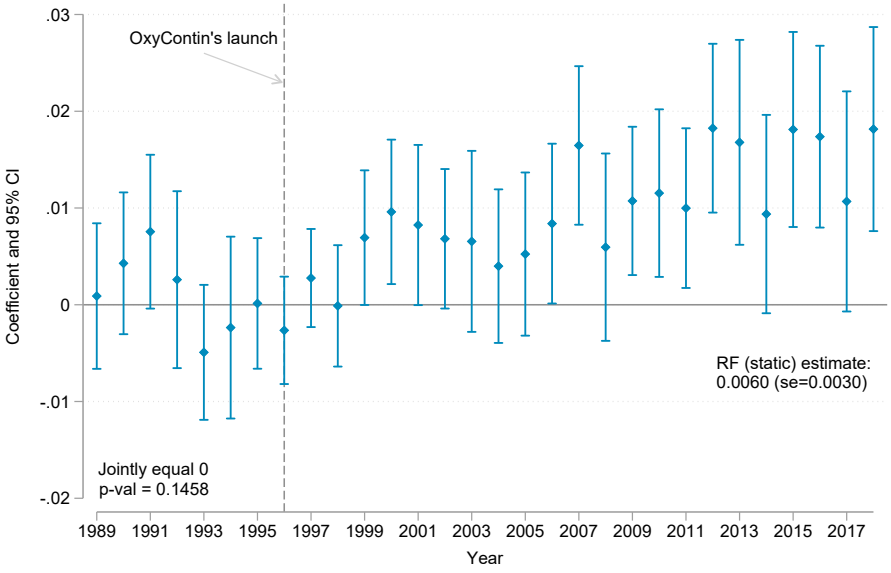
Notes: This figure shows the effects of the increase in prescription opioid supply in opioid related mortality by race group (panel a) and by gender (panel b). We present the results of a dynamic reduced-form estimation were we regress these outcomes on a set of year-dummy variables interacted with our instrument—cancer mortality in 1994-1996. These coefficients corresponds to the estimate of ϕ_t in Equation 2.3. This figure is referenced in Section 5.5.1

Figure A.15: Effects of Mid-nineties Cancer-market Targeting on Deaths of Despair

(a) Deaths of Despair (excluding opioids)



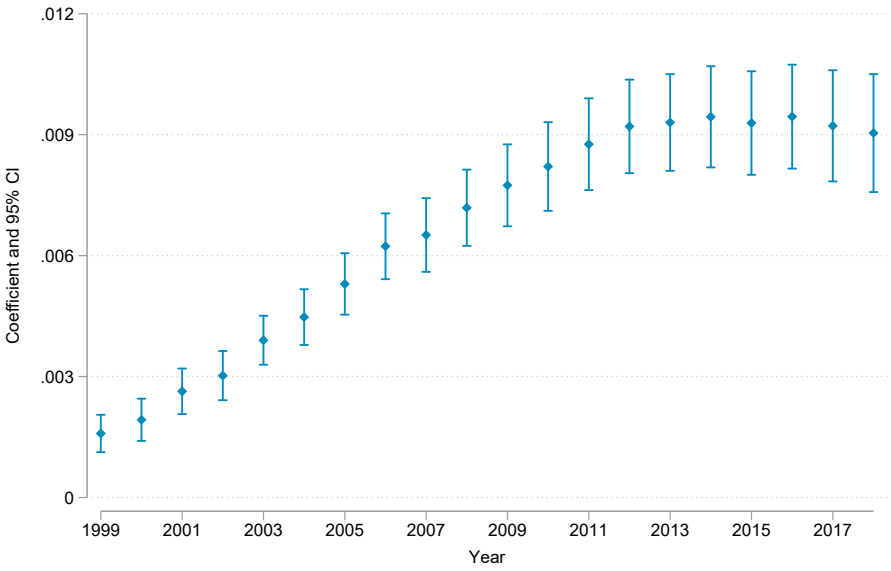
(b) Alcohol-related Deaths



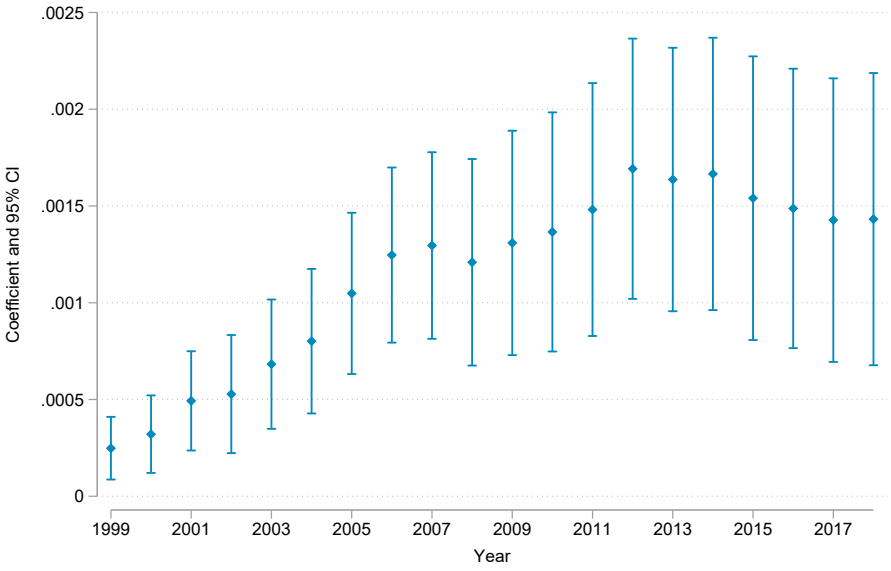
Notes: This figure shows the effects of the increase in prescription opioids supply in various measures of deaths of despair; excluding deaths which cause is opioid poisoning. We present the results of a dynamic reduced-form estimation were we regress these outcomes on a set of year-dummy variables interacted with our instrument. We test for the presence of pre-trends in the relation between deaths of despair and mid-nineties cancer mortality and do not reject the null hypothesis that the estimated coefficients before 1996 are jointly equal to zero. This figure is referenced in Section 4.4.2, in Section 5.5.1, and in Section 5.5.2

Figure A.16: Effects of Mid-nineties Cancer-market Targeting on Disability Claims

(a) SSDI



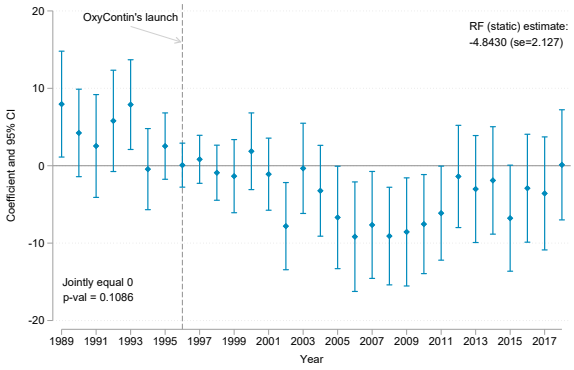
(b) SSI



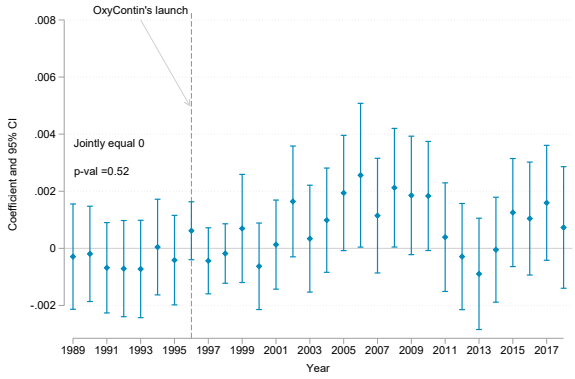
Notes: We present the results of a dynamic reduced-form estimation were we regress these outcomes on a set of year-dummy variables interacted with our instrument—cancer mortality in 1994-1996. These coefficients corresponds to the estimate of ϕ_t in Equation 2.3. This figure is referenced in Section 5.5.2

Figure A.17: Effects of Mid-nineties Cancer-market Targeting on Birth Outcomes

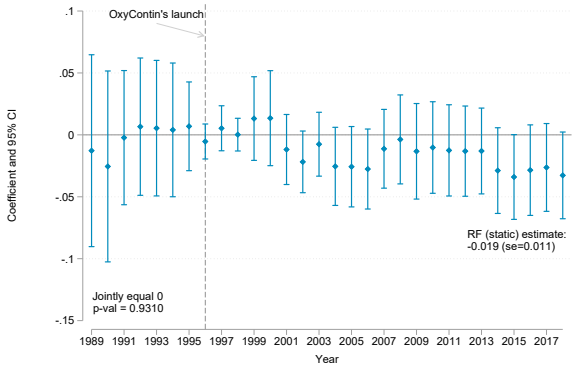
(a) Birth Weight



(b) Share infants born with low weight



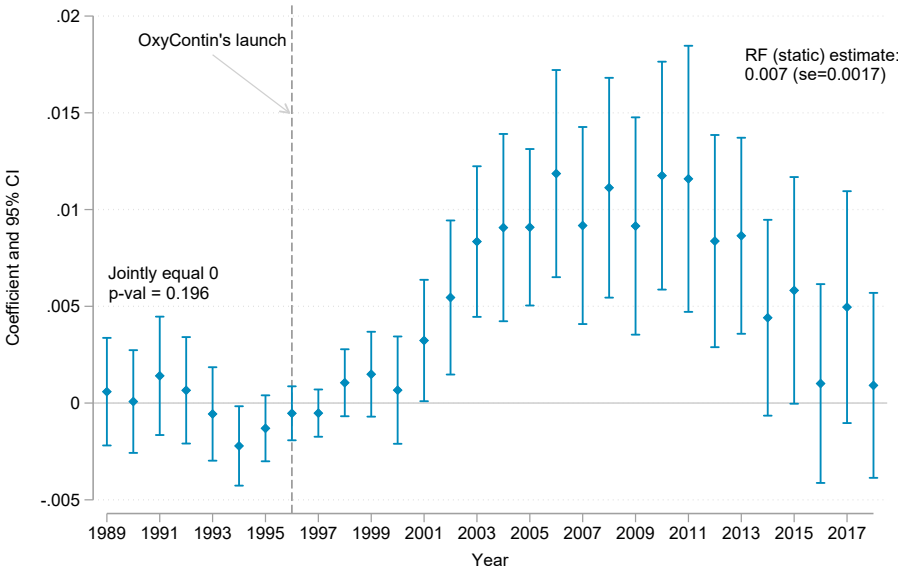
(c) APGAR Score



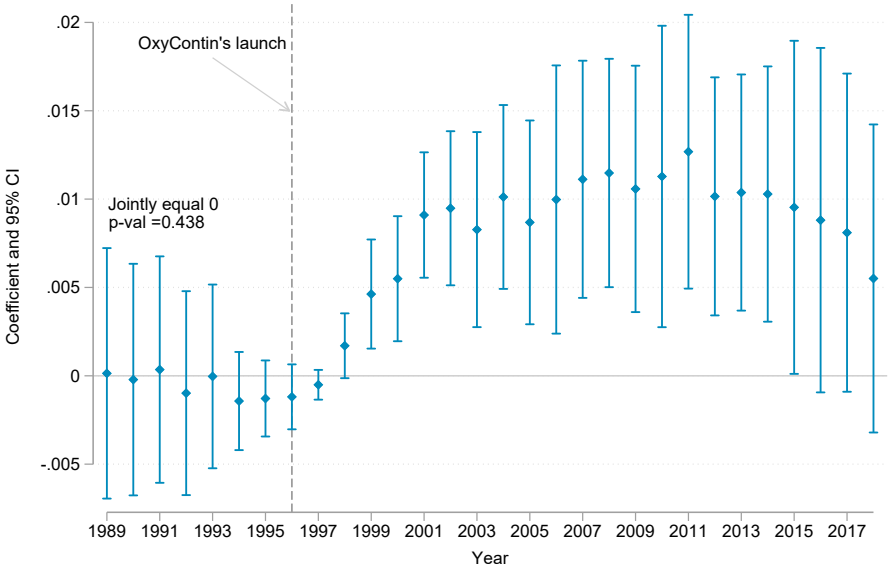
Notes: This figure shows the effects of the increase in prescription opioids supply in birth weight (panel a), share of infants with low birth weight (panel b) and in APGAR score (panel c). The APGAR score is a measure of the physical condition of a newborn infant. It varies from 0 to 10, a score of 10 represents the best possible condition. We present the results of a dynamic reduced-form estimation were we regress these outcomes on a set of year-dummy variables interacted with our instrument—cancer mortality in 1994-1996 (ϕ_t in Equation 2.3). We use this specification to test for the presence of pre-trends in the relation between infant outcomes and mid-nineties cancer mortality; we do not reject the null hypothesis that the estimated coefficients before 1996 are jointly equal to zero, the p values are presented in the figures. This figure is referenced in Section 4.4.2, in Section 5.5.1, and in Section 5.5.2

Figure A.18: Effects of Mid-nineties Cancer-market Targeting on Prescription Opioid Mortality

(a) Baseline specification

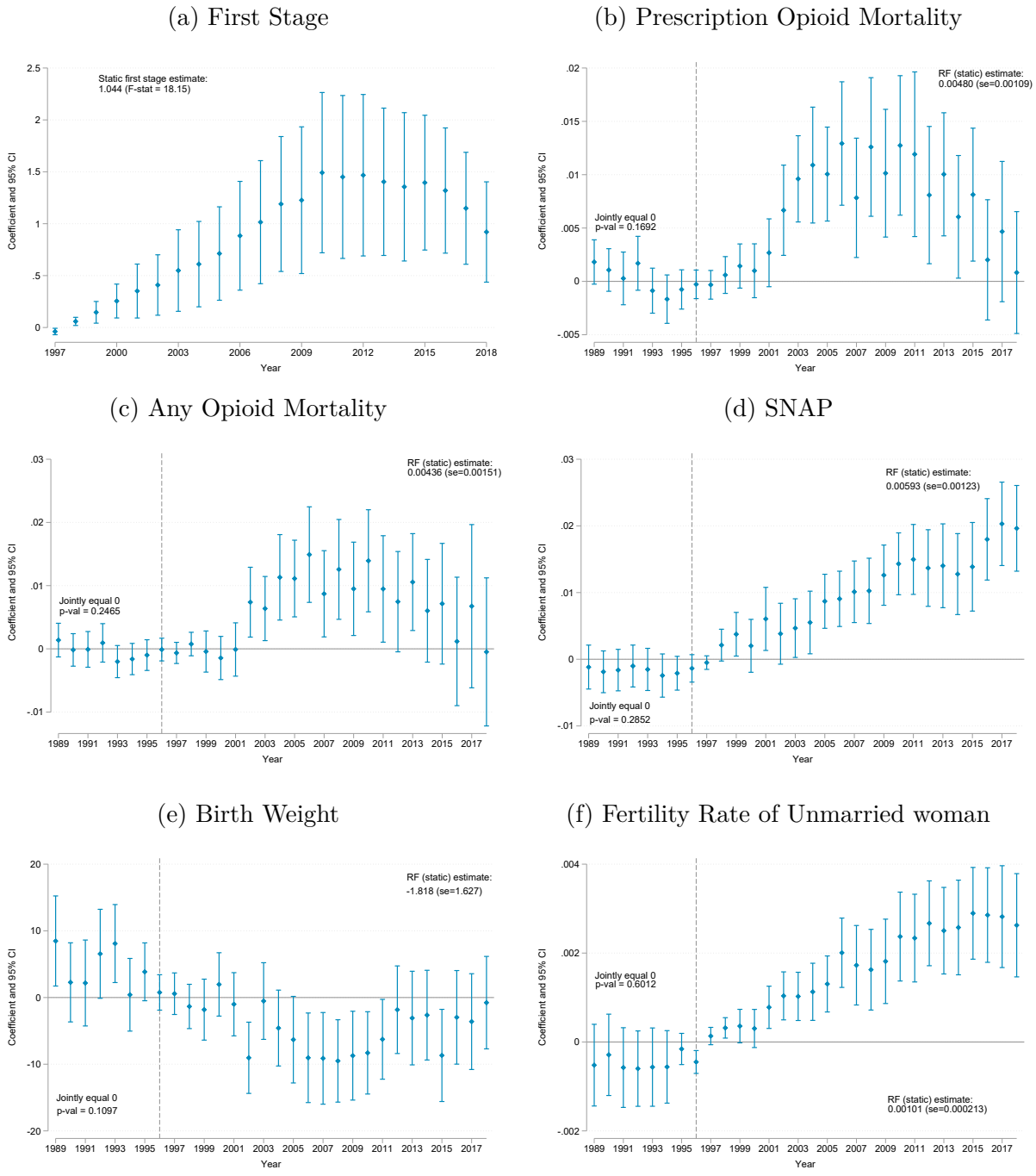


(b) Population weighted Specification



Notes: This figure shows the effects of the increase in prescription opioid supply in prescription opioid mortality. Panel (a) reproduces our main analysis (see Panel (b) in Figure 2.6) and Panel (b) reproduces this analysis using population weights in the estimation of the event study coefficients. This figure is referenced in Section 6.6.1

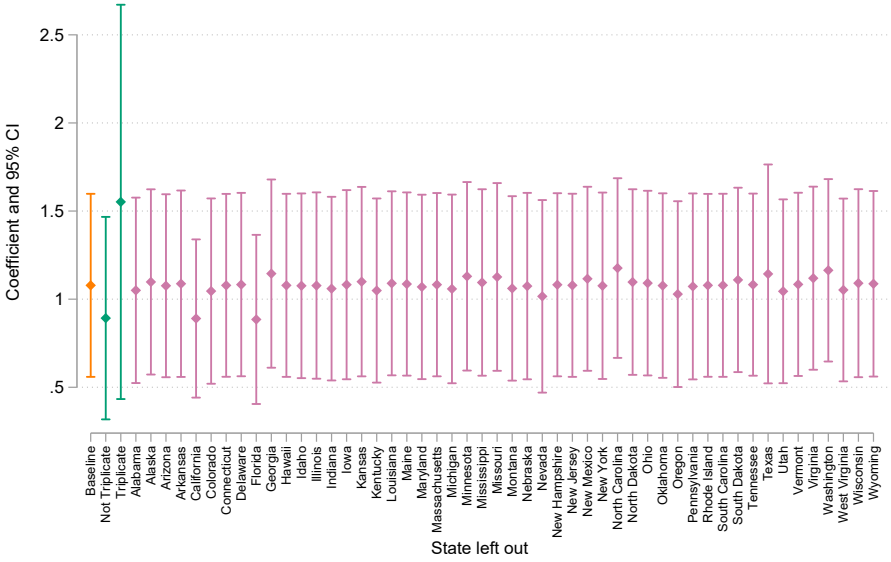
Figure A.19: Dynamic First Stage and Reduced Form Estimates - Alternative specification



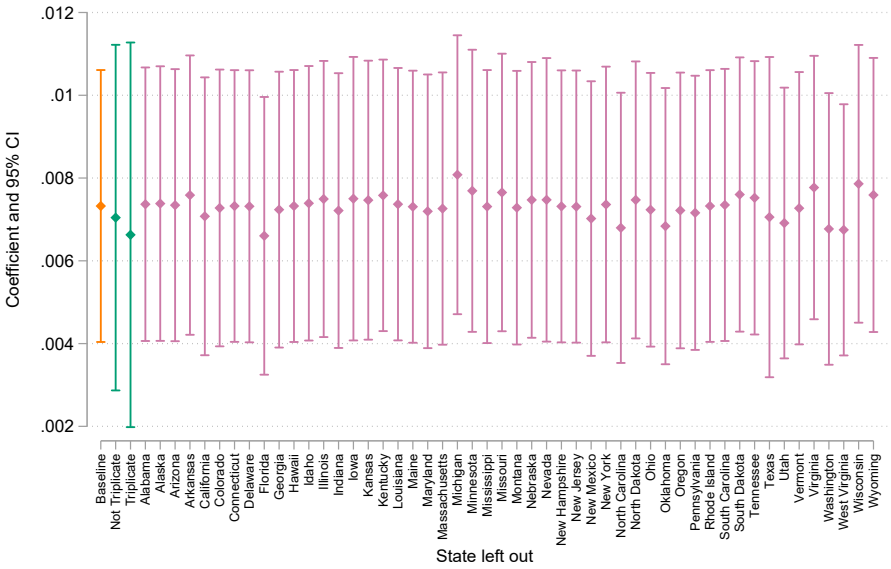
Notes: This figure shows the dynamic first stage (panel a) and reduced-form (panels b to f) relations between outcomes of interest and an alternative instrument: cancer mortality rate in 1994. That is, we regress the outcomes of interest on a set of year-dummy variables interacted with cancer mortality in 1994. This figure provides a robustness check for our preferred specification which uses cancer mortality in 1994-1996 as an instrument. We do not find evidence for the presence of pre-trends in the relation between opioid mortality—and other outcomes of interest—and mid-nineties cancer mortality in this alternative specification. We test if the estimated coefficients before 1996 are jointly equal to zero and do not reject the null hypotheses, the p values are reported in each panel. This figure is referenced in Section 6.6.1

Figure A.20: Robustness check: Leave-ones-out estimates

(a) First-stage coefficient estimates



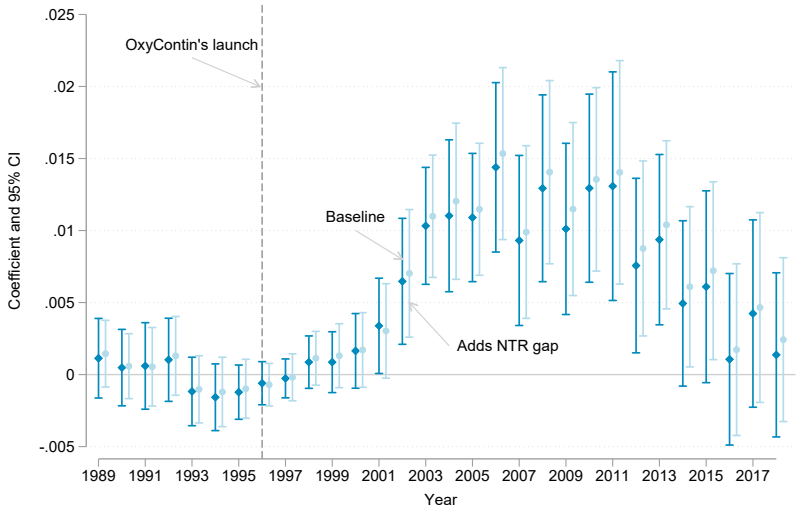
(b) Prescription opioids: Reduced-form coefficient estimates



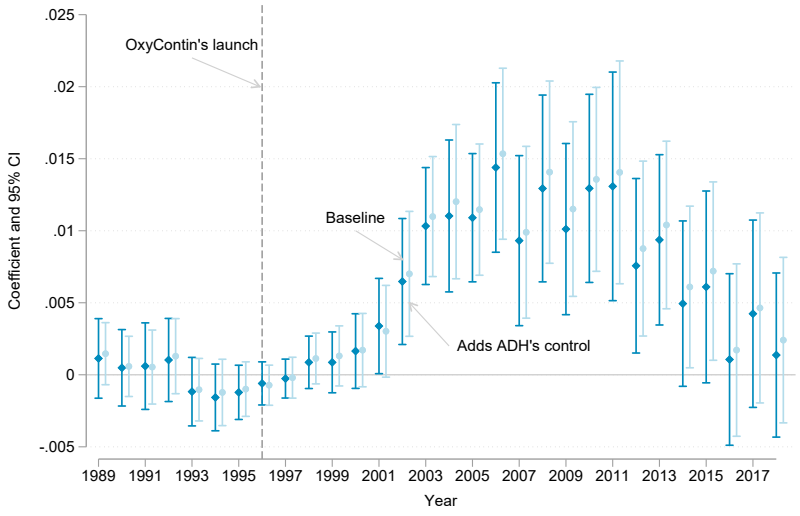
Notes: Panel (a) of this figure reports the estimated coefficient ϕ of the first stage equation (Equation 2.1) and the corresponding 95% confidence interval. Panel (b) of this figure reports the estimated reduced-form coefficient. The first coefficient and confidence interval of each graph replicate the main result result—see column 5 of Table 2.3 and column 2 of Table 2.5. Each of the subsequent coefficients are computed by excluding all commuting zones in the state or group of states indicated on the horizontal axis (thus, the x-axis label). Triplicate states are: California, Idaho, Illinois, New York, and Texas. Not triplicate group excludes all these 5 states. This figure is referenced in Section 6.6.1

Figure A.21: Robustness check: Control for exposure to permanent normal trade relations to China ("China shock")

(a) NTR gap (Pierce and Schott, 2020)



(b) Change in Chinese import exposure (ADH, 2013)



Notes: This figure presents the baseline dynamic reduced-form estimates and the dynamic reduced-form estimates when we control for exposure to permanent normal trade relations (PNTR) to China—termed the China shock in the trade literature. In October, 2000, the US Congress passed a bill granting permanent normal trade relations to China, a trade liberalization that granted China imports access to normal trade relations (NTR) tariff rates. This trade liberalization differentially exposed US regions to increased import competition from China via their industry structure. We test whether results on opioid mortality are driven by this differential exposure. First, we follow Pierce and Schott (2020) and construct a measure of exposure to trade liberalization as the difference between the non-NTR rates to which tariffs could have risen prior to PNTR and the NTR rates that were locked in by the change in policy. A higher NTR gap indicates a larger trade liberalization after the passage of PNTR.

Table A.1: Additional Summary Statistics: Opioid Prescriptions, doses per capita

	Mean	**Median**	**SD**	**Min**	**Max**	**Observations**
1997						
All opioids	1.49	1.40	0.67	0.04	7.64	590
Oxycodone	0.35	0.32	0.21	0.01	1.76	590
Hydrocodone	0.55	0.49	0.34	0.01	2.73	590
Morphine	0.31	0.29	0.17	0.01	1.89	590
2007						
All opioids	7.03	6.24	4.01	0.22	36.24	590
Oxycodone	3.26	2.76	2.33	0.08	26.86	590
Hydrocodone	2.33	1.87	1.72	0.04	14.30	590
Morphine	1.04	0.89	0.68	0.04	8.58	590
2017						
All opioids	6.97	6.30	3.50	0.19	27.47	590
Oxycodone	3.75	3.42	2.25	0.11	15.34	590
Hydrocodone	1.86	1.63	1.17	0.04	10.57	590
Morphine	0.92	0.82	0.50	0.03	5.27	590

Notes: This table presents summary statistics for our measure of the prescription opioids supply and the distribution of oxycodone, hydrocodone, and morphine for the years 1997, 2007, and 2017. Data come from the ARCOS and are expressed in morphine-equivalent mg. This table is referenced in Section 3.3.1

Table A.2: Summary statistics: Pre-period and sample period.

	1989 - 1995		1999 - 2018	
	Mean	SD	Mean	SD
	(1)	(2)	(3)	(4)
Cancer Mortality per 1,000				
Cancer mortality rate 1994-1996	2.53	0.58	2.53	0.58
Cancer mortality rate	2.53	0.59	2.48	0.55
Opioid-related Mortality per 1,000				
Prescription opioids	0.01	0.01	0.04	0.05
Any opioids	0.01	0.02	0.07	0.07
Other Mortality Measures per 1,000				
All-cause mortality (+20 years old)	9.81	2.07	9.87	2.06
Deaths of despair	0.24	0.08	0.27	0.10
Deaths of despair - alcohol only	0.09	0.04	0.12	0.06
Deaths of despair - suicide only	0.13	0.05	0.15	0.06
Demand for Social Services				
Share SNAP	0.10	0.06	0.12	0.07
Infant and Fertility Outcomes				
Infant MR (per 1,000 births)	8.87	3.22	6.86	2.87
Birth weight	3416.31	80.77	3274.25	79.47
Share low birth weight	0.07	0.02	0.08	0.02
Share preterm	0.11	0.02	0.12	0.03
APGAR score - all infants	8.24	2.65	8.82	0.19
APGAR score - dead infants	6.14	2.15	5.62	2.28
Median gestation	39.12	0.32	38.95	0.24
Fertility rate	0.08	0.03	0.08	0.01
Fertility rate 25-29	0.12	0.04	0.13	0.02
Fertility rate - unmarried women	0.02	0.01	0.03	0.01

Notes: This table presents summary statistics for our main outcomes and cancer mortality incidence for the period before the launch of OxyContin (1989-1995) and the period of analysis (1999 - 2018). We leverage data from multiple sources. The last two columns reproduce columns (2) and (4) of Table 2.1. Data on opioid, cancer, birth, and fertility outcomes come from the NVSS. We use data from the Food and Nutrition Service of the Department of Agriculture and the SSA to construct demand for the Supplemental Nutrition Assistance Program (SNAP), Supplemental Security Income (SSI), and Social Security Disability Insurance (SSDI). This table is referenced in Section 3.

Table A.3: Determinants of Cancer Mortality Rate 94-96

Dependent variable: Cancer MR 94-96	(1)		(2)
Sh. of population over 66	11.13***	Adult MR excluding cancer	0.0439**
	[1.895]		[0.0179]
Sh. of population 18-65	-0.664	Income per capita	-0.00000857
	[1.361]		0.118
Sh. of population under 1	2.156	Share with some college	0.518*
	[9.066]		[0.274]
Share Black	0.127	Share with high school or less	0.124
	[0.241]		[0.191]
Share Hispanic	-1.215***	Share working in manufacturing	-0.199
	[0.303]		[0.133]
Share female	-1.48	Labor Force Participation	0.528
	[1.565]		[0.399]
Prescription Opioids MR	1.093	Employment rate	-1.984*
	[1.078]		[1.118]
Infant Mortality rate	-0.00288	Share SNAP	0.484
	[0.00337]		[0.383]
Fertility rate	0.311	Share SSDI	1.856
	[0.426]		[1.929]
Observations	590	R^2	0.847

Notes: This table presents estimates of the determinants of the 1994-1996 cancer mortality rate at the commuting zone level. This regression includes state fixed effects. Robust to heteroskedasticity standard errors are in brackets. MR stands for Mortality rate. * $p<0.10$, ** $p<0.05$, *** $p<0.01$. This table is referenced in Section 4.4.2

Table A.4: Baseline Results with Different Time Periods

Panel A: First Stage				
Dependent variable:	Prescription Opioids pc			
	(1)	(2)	(3)	(4)
Cancer MR 94-96	1.078***	0.916***	1.047***	1.474***
se	[0.264]	[0.258]	[0.277]	[0.330]
t-stat	4.08	3.55	3.78	4.46
Effective F-stat	16.63	12.62	14.25	19.90
Observations	11,800	7,080	8,850	5,310
Adjusted R^2	0.564	0.565	0.582	0.425
Sample	**All**	**1999-2010**	**1999-2013**	**2010-2018**

Panel B: Instrumental Variables				
Dependent variable:	Prescription Opioids Mortality Rate			
	(1)	(2)	(3)	(4)
Presc. Opioids pc	0.00679***	0.00785***	0.00769***	0.00533***
	[0.00200]	[0.00259]	[0.00230]	[0.00169]
Observations	11,800	7,080	8,850	5,310
Sample	**All**	**1999-2010**	**1999-2013**	**2010-2018**
Dependent variable:	Any Opioid Mortality Rate			
	(1)	(2)	(3)	(4)
Presc. Opioids pc	0.00646***	0.00677***	0.00672***	0.00562**
	[0.00231]	[0.00256]	[0.00232]	[0.00237]
Observations	11,800	7,080	8,850	5,310
Sample	**All**	**1999-2010**	**1999-2013**	**2010-2018**

Notes: Panel A presents results for the first-stage regression using alternative periods. Column (1) reproduces the main results for 1999-2018, column (2) presents estimates for the first wave of the opioid epidemic, column (3) presents estimates for the first and second waves pooled together, and column (4) presents estimates for the after-OxyContin reformulation period. Panel B presents results from a regression of the opioid mortality measure on all prescription opioids distribution per capita, instrumenting the latter by the cancer incidence in the commuting zone in 1994-1996; i.e., reproduces the results presented in Table 2.5 under alternative periods. All regressions include state times year fixed effects and a set of control variables: contemporaneous cancer mortality rate, share of population under 1 year old, share of population between 18 and 65, share of population over 66 years old, share of Black, White, and Hispanic population, and share of female population. Standard errors are clustered at the CZ level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$. This table is referenced in Section 5.5.1

Table A.5: Baseline Results with Different Time Periods. IV Estimates.

	SNAP			SSDI		
	(1)	(2)	(3)	(4)	(5)	(6)
Presc. Opioids pc	0.00455*	0.00487**	0.00680***	0.00584***	0.00605***	0.00718***
	[0.00250]	[0.00219]	[0.00205]	[0.00144]	[0.00141]	[0.00135]
Effective F	15.22	17.06	25.70	15.22	17.06	25.70
Sample	**1999-2010**	**1999-2013**	**2010-2018**	**1999-2010**	**1999-2013**	**2010-2018**

	SSI			IMR		
	(1)	(2)	(3)	(4)	(5)	(6)
Presc. Opioids pc	0.00226*	0.00248*	0.00320*	0.0458	0.0512	0.0846
	[0.00133]	[0.00141]	[0.00174]	[0.185]	[0.160]	[0.113]
Effective F	15.22	17.06	25.70	15.22	17.06	25.70
Sample	**1999-2010**	**1999-2013**	**2010-2018**	**1999-2010**	**1999-2013**	**2010-2018**

	Birth weight			Fertility		
	(1)	(2)	(3)	(4)	(5)	(6)
Presc. Opioids pc	-5.989**	-5.093**	-2.915*	0.00210***	0.00233***	0.00350***
	[2.811]	[2.316]	[1.623]	[0.000696]	[0.000674]	[0.000778]
Effective F	15.22	17.06	25.70	15.22	17.06	25.70
Sample	**1999-2010**	**1999-2013**	**2010-2018**	**1999-2010**	**1999-2013**	**2010-2018**

Notes: This table presents results from a regression of outcome y on prescription opioids distribution per capita, instrumenting the latter by the cancer incidence in the commuting zone in 1994-1996; i.e., reproduces the results presented in Tables 2.7 and 2.9 under alternative periods. Columns (1) and (4) present estimates for the first wave of the opioid epidemic, columns (2) and (5) present estimates for the first and second waves pooled together, and columns (3) and (6) present estimates for the after-OxyContin reformulation period. All regressions include state times year fixed effects and a set of control variables: contemporaneous cancer mortality rate, share of population under 1 year old, share of population between 18 and 65, share of population over 66 years old, share of Black, White, and Hispanic population, and share of female population. Standard errors are clustered at the CZ level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$. This table is referenced in Section 5.5.2

Table A.6: Effects on Fertility Rate by Marital Status and Age

Dependent variable: Fertility rate						
	(1)	(2)	(3)	(4)	(5)	(6)
Pres. Opioids pc	0.00166***	-0.000119	-0.00107	0.00327***	0.0000223	-0.00123**
	[0.000475]	[0.000517]	[0.00111]	[0.00115]	[0.000446]	[0.000497]
Sample	Non-marital births	Marital births	All 20-24	All 25-29	All 30-34	All 35-39
Observations	11,800	11,800	11,800	11,800	11,800	11,800
Clusters	590	590	590	590	590	590

Notes: This table presents results from a regression of measures of fertility rate on prescription opioids distribution per capita, instrumenting the latter by the cancer incidence in the commuting zone in 1994-1996. All regressions include state times year fixed effects and a set of control variables: contemporaneous cancer mortality rate, share of population under 1 year old, share of population between 18 and 65, share of population over 66 years old, share of Black, White, and Hispanic population, and share of female population. Standard errors are clustered at the CZ level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$. This table is referenced in Section 5.5.2

Table A.7: First Stage Results with Population Size Controls

Dependent variable: Prescription opioids per capita					
	(1)	(2)	(3)	(4)	(5)
Cancer MR 94-96	1.078***	1.635***	1.072***	1.046***	1.608***
se	[0.264]	[0.483]	[0.276]	[0.266]	[0.490]
t-stat	4.08	3.39	3.88	3.94	3.28
Effective F-stat	16.63	11.49	15.05	15.52	10.76
Share pop +65 yo	No	Yes	No	No	Yes
Total pop +65 yo	No	No	Yes	No	No
Total population	No	No	No	Yes	Yes
Observations	11,800	11,800	11,800	11,800	11,800
Clusters	590	590	590	590	590
Adj. R^2	0.56	0.57	0.56	0.57	0.57

Notes: All specifications include as control variables: contemporaneous cancer mortality rate, share of population under 1 year old, share of population between 18 and 65, share of population over 66 years, share of Black, White, and Hispanic population, and share of female population. Standard errors are clustered at the commuting-zone level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$. This table is referenced in Section 6.6.1

Table A.8: First Stage Robustness Check

Panel A. Alternative choices of instruments

Dependent variable:	Prescription Opioids pc				
	(1)	(2)	(3)	(4)	(5)
Cancer MR	0.868***	1.171***	0.930***	0.754***	1.417***
	[0.229]	[0.272]	[0.260]	[0.223]	[0.284]
Mean cancer MR	2.5168	2.5403	2.5477	2.5221	2.2582
Instrument version:	Age adjusted MR 94-96	1994	1995	1996	Weighted
Observations	11,800	11,800	11,800	11,800	11,800
Clusters	590	590	590	590	590
Adj. R^2	0.553	0.565	0.557	0.551	0.553

Panel B. Alternative choices of instruments

Dependent variable:	Prescription Opioids pc				
	(1)	(2)	(3)	(4)	(5)
Cancer MR	1.186***	0.402***	0.210**	0.127**	11.72***
	[0.315]	[0.149]	[0.0988]	[0.0563]	[4.317]
Mean cancer MR	0.6836	9.8072	13.1382	17.5892	0.1342
Instrument version:	Excluding lung cancer	55+	65+	75+	Sh. Pop 66+
Observations	11,800	11,800	11,800	11,800	11,800
Clusters	590	590	590	590	590
Adj. R^2	0.55	0.55	0.56	0.56	0.57

Panel C. Alternative measures of prescription opioids supply

Dependent variable:	Oxycodone pc	Non-oxycodone prescription opioids
	(1)	(2)
Cancer MR	0.605***	0.473***
	[0.186]	[0.107]
Mean cancer MR	2.5312	2.5312
Instrument version:	Baseline	Baseline
Observations	11,800	11,800
Clusters	590	590
Adj. R^2	0.526	0.594

Notes: All regressions include state times year fixed effects and a set of control variables: contemporaneous cancer mortality rate, share of population under 1 year old, share of population between 18 and 65, share of population over 66 years, share of Black, White, and Hispanic population, and share of female population. Standard errors are clustered at the commuting-zone level. * $p<0.10$, ** $p<0.05$, ***$p<0.01$. This table is referenced in Section 6.6.1

Table A.9: Baseline Results under a Shift-share Instrument

Dependent var:	Presc. Opioids pc	Presc. Opioids MR	Any Opioids MR	SNAP	SSDI
	(1)	(2)	(3)	(4)	(5)
Shift Share	0.00417***				
	[0.000997]				
Effective F	17.47				
Presc. Opioids pc		0.00644***	0.00635***	0.00927***	0.00553***
		[0.00188]	[0.00219]	[0.00277]	[0.00127]
Model	FS	IV	IV	IV	IV
Dependent var:		SSI	Infant Mortality Rate	Fertility rate	Birth weight
		(6)	(7)	(8)	(9)
Presc. Opioids pc		0.00319**	-0.0218	0.00149***	-4.344**
		[0.00158]	[0.120]	[0.000548]	[1.964]
Model		IV	IV	IV	IV

Notes: Column 1 reports the estimated coefficient for the first stage. Columns 2 to 9 present results from IV regressions using the shift-share instrument. Each regression is run over a sample of 11,800 observations with 590 clusters (commuting zones). All regressions include state times year fixed effects and a set of control variables: contemporaneous cancer mortality rate, share of population under 1 year old, share of population between 18 and 65, share of population over 66 years old, share of Black, White, and Hispanic population, and share of female population. Standard errors are clustered at the CZ level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$. This table is referenced in Section 6.6.1

Table A.10: Placebo Check - Alternative Instruments

Dependent variable: Prescription opioids per capita						
	(1)	(2)	(3)	(4)	(5)	(6)
CVD MR 94 96	0.372			-2.023**		
	[0.611]			[0.822]		
Accidental MR 94 96		1.067			-1.639	
		[1.411]			[1.406]	
Homicides MR 94 96			0.214			-0.474
			[3.379]			[3.173]
Cancer MR 94 96				1.381***	1.015***	0.923***
				[0.347]	[0.245]	[0.233]
Model	FS	FS	FS	FS	FS	FS
Observations	11,800	11,800	11,800	11,800	11,800	11,800
Clusters	590	590	590	590	590	590
Adjusted R^2	0.55	0.549	0.549	0.565	0.561	0.562

Notes: CVD stands for cerebrovascular diseases. Columns 1-3 report first-stage regression with alternative instrument. Columns 4-6 add our baseline instrument. All regressions include state times year fixed effects and a set of control variables: labor force participation, contemporaneous cancer mortality rate, share of population under 1 year old, share of population between 18 and 65, share of population over 66 years old, share of Black, White, and Hispanic population, and share of female population. Standard errors are clustered at the CZ level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$. This table is referenced in Section 6.6.2

Table A.11: Direct Effects. Alternative Measure of Opioid Mortality

Dependent var:	Drug Induced Mortality Rate		
	(1)	(2)	(3)
Prescription opioids pc	0.00505***		0.0112***
	[0.00152]		[0.00369]
tF 0.05 se			0.00518
t-stat using tF 0.05 se			2.16329
AR p-value			0.00010
Cancer MR 94-96		0.0121***	
		[0.00314]	
Effect size (%)	20.96		46.94
Model	OLS	RF	IV
Observations	11,800	11,800	11,800
Clusters	590	590	590
Adjusted R^2	0.4304	0.3908	
Effective F-stat			16.63
Cragg-Donald Wald F-stat			358.58

Notes: Control variables are contemporaneous cancer mortality rate, share of population under 1 year old, share of population between 18 and 65, share of population over 66 years old, share of Black, White, and Hispanic population, and share of female population. Effect size indicates the percent change in the dependent variable relative to its mean when doses of prescription opioids per capita increase from the 25th to the 75th percentile. Standard errors in square brackets are clustered at the CZ level. Using these standard errors, we report * $p<0.10$, ** $p<0.05$, *** $p<0.01$. tF 0.05 se, t-stat using tF0.05 se, and the *AR p-value* correspond to weak-instrument-robust inference procedures. This table is referenced in Section 6.6.3

Table A.12: Alternative Measure of Opioid Supply.

Dependent var:	Oxycodone pc	Presc. opioids MR	All opioids MR
	(1)	(2)	(3)
Cancer MR 94-96	0.605***		
	[0.186]		
Oxycodone pc		0.0121***	0.0115***
		[0.00412]	[0.00436]
tF 0.05 se		(0.00578)	(0.00612)
t-stat using tF 0.05 se		2.0932	1.8799
Effect size (%)	38.00	91.50	40.37
Model	FS	IV	IV
Observations	11,800	11,800	11,800
Clusters	590	590	590
Adjusted R^2	0.526		

Notes: All regressions include state times year fixed effects. Control variables are contemporaneous cancer mortality rate, share of population under 1 year old, share of population between 18 and 65, share of population over 66 years old, share of Black, White, and Hispanic population, and share of female population. This table reproduces the main analysis using Oxycodone shipments as the measure of opioid supply. Effect size in column (1) is computed as the predicted changes in doses of oxycodone and prescription opioids per capita from an increase in cancer mortality that would change a commuting zone in the 5th percentile of the cancer distribution to the 95th percentile. Effect sizes in columns (2) and (3) indicate the percent change in the dependent variable relative to its mean when doses of prescription opioids per capita increase from the 25th to the 75th percentile. Standard errors in square brackets are clustered at the CZ level; using these standard errors, we report * $p<0.10$, ** $p<0.05$, *** $p<0.01$. tF 0.05 se, and t-stat using tF0.05 se correspond to weak-instrument-robust inference procedures. This table is referenced in Section 6.6.3

Table A.13: Baseline Results under Alternative Sample Restrictions

Dependent var:	Presc. Opioids pc			Prescription Opioids MR		
	(1)	(2)	(3)	(4)	(5)	(6)
Cancer MR 94-96	1.191***	1.055***	1.018***			
	[0.249]	[0.297]	[0.288]			
Presc. Opioids pc				0.00355***	0.00684***	0.00826***
				[0.00134]	[0.00231]	[0.00268]
Sample	**15,000+**	**40,000+**	**55,000+**	**15,000+**	**40,000+**	**55,000+**
	Any Opioids MR			All-cause mortality over 20		
	(1)	(2)	(3)	(4)	(5)	(6)
Presc. Opioids pc	0.00152	0.00697**	0.00885***	0.0137	0.0515	0.102
	[0.00171]	[0.00273]	[0.00329]	[0.0361]	[0.0477]	[0.0668]
Sample	**15,000+**	**40,000+**	**55,000+**	**15,000+**	**40,000+**	**55,000+**
	SSDI			SSI		
	(1)	(2)	(3)	(4)	(5)	(6)
Presc. Opioids pc	0.00504***	0.00586***	0.00652***	0.00204**	0.00339**	0.00438*
	[0.00106]	[0.00155]	[0.00173]	[0.000851]	[0.00169]	[0.00239]
Sample	**15,000+**	**40,000+**	**55,000+**	**15,000+**	**40,000+**	**55,000+**
	SNAP			IMR		
	(1)	(2)	(3)	(4)	(5)	(6)
Presc. Opioids pc	0.00941***	0.00997***	0.00919***	0.175	-0.0297	0.0604
	[0.00248]	[0.00336]	[0.00307]	[0.130]	[0.142]	[0.150]
Sample	**15,000+**	**40,000+**	**55,000+**	**15,000+**	**40,000+**	**55,000+**
	Birth weight			Fertility		
	(1)	(2)	(3)	(4)	(5)	(6)
Presc. Opioids pc	-4.896***	-3.770*	-6.480**	0.00108***	0.00156**	0.00160**
	[1.852]	[2.240]	[2.624]	[0.000404]	[0.000632]	[0.000706]
Sample	**15,000+**	**40,000+**	**55,000+**	**15,000+**	**40,000+**	**55,000+**

Notes: This table presents results for the first-stage regression and IV results using alternative sample definitions. Our preferred specification restricts the sample to commuting zones with population higher than 25,000 residents. When the sample is restricted to population above 15,000, the sample size is 12,820 observations and 641 clusters. Analogously, when restricted to population above 40,000, sample size is 10,880 and 544 cluster, and 9,620 and 481 clusters when restriction is above 55,000. All regressions include state times year fixed effects and a set of control variables: contemporaneous cancer mortality rate, share of population under 1 year old, share of population between 18 and 65, share of population over 66 years old, share of Black, White, and Hispanic population, and share of female population. Standard errors are clustered at the CZ level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$. This table is referenced in Section 3.3.3 and in Section 6.6.4

Table A.14: Alternative Sample Results for SNAP

Dependent variable:	Share SNAP					
	(1)	(2)	(3)	(4)	(5)	(6)
Presc. Opioids pc	0.000144		0.00982***	0.000213		0.0106***
	[0.51]		[3.28]	[0.74]		[3.23]
Cancer 94 96		0.0106***			0.0116***	
		[4.67]			[5.53]	
Effective F-stat			16.63			13.70
Model	OLS	RF	IV	OLS	RF	IV
Sample	Baseline	Baseline	Baseline	Restricted	Restricted	Restricted
Observations	11,800	11,800	11,800	9,962	9,962	9,962
Clusters	590	590	590	533	533	533

Notes: Columns 1-3 report baseline results and columns 4-6 report results only for commuting zones where county-level data were available. All regressions include state times year fixed effects and a set of control variables: labor force participation, contemporaneous cancer mortality rate, share of population under 1 year old, share of population between 18 and 65, share of population over 66 years old, share of Black, White, and Hispanic population, and share of female population. Standard errors are clustered at the CZ level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$. This table is referenced in Section 6.6.4

Table A.15: Alternative Specifications

Dependent var:	Presc. Op Mortality	Any Op. Mortality	SSDI	SSI	SNAP	Fertility
Presc. Opioids pc	0.00684***	0.00643***	0.00579***	0.00322**	0.00922***	0.00145***
	[0.00204]	[0.00232]	[0.00136]	[0.00152]	[0.00270]	[0.000529]
Extra covariate	**Empl.**	**Empl.**	**Empl.**	**Empl.**	**Empl.**	**Empl.**
Observations	11,800	11,800	11,800	11,800	11,800	11,800
Clusters	590	590	590	590	590	590
Dependent var:	Presc. Op Mortality	Any Op. Mortality	SSDI	SSI	SNAP	Fertility
Presc. Opioids pc	0.00684***	0.00643***	0.00579***	0.00322**	0.00922***	0.00145***
	[0.00204]	[0.00232]	[0.00136]	[0.00152]	[0.00270]	[0.000529]
Extra covariate	**Unemp.**	**Unemp.**	**Unemp.**	**Unemp.**	**Unemp.**	**Unemp.**
Observations	11,800	11,800	11,800	11,800	11,800	11,800
Clusters	590	590	590	590	590	590

Notes: All regressions include state times year fixed effects and a set of control variables: contemporaneous cancer mortality rate, share of population under 1 year old, share of population between 18 and 65, share of population over 66 years old, share of Black, White, and Hispanic population, and share of female population. Standard errors are clustered at the CZ level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$. This table is referenced in Section 6.6.4

Table A.16: First Stage with Additional Control Variables: Recession, China Shock & Unemployment

Presc. Opioids pc	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Cancer MR 94 96	1.078***	1.137***	1.101***	1.104***	1.075***	1.074***	1.075***
	[0.266]	[0.272]	[0.268]	[0.268]	[0.264]	[0.264]	[0.263]
Extra control	**Recession**	**NTR Gap**	**ADH 1990**	**ADH 2000**	**Unemp. 94**	**Unemp. 95**	**Unemp. 96**
Observations	11,800	11,800	11,740	11,740	11,800	11,800	11,800
Adjusted R^2	0.57	0.57	0.57	0.57	0.56	0.56	0.56
Clusters	590	590	587	587	590	590	590

Notes: This table estimate the first stage including additional control variables to account for the 2001 Economic Recession and the China Shock. All regressions include state times year fixed effects and a set of control variables Standard errors are clustered at the CZ level. All regressions are run on panel at the CZ level with 11,800 observations and 590 clusters. * $p<0.10$, ** $p<0.05$, *** $p<0.01$. This table is referenced in Section 6.6.5

1.3 Alternative Sources of Variation in the Marketing of OxyContin

In this appendix, we compare the approach presented in this paper to the work of Alpert et al. (2022). We pay special attention to the ways in which we improve on their work, with focus on the differences in specifications, the statistical power of the two approaches, and we highlight concerns around the source of variation.

Both papers exploit geographic exposure to the initial marketing of OxyContin. Alpert et al. (2022) use a state-level binary variation, whereas we exploit continuous commuting-zone variation.[1] Alpert et al. (2022) show that five states with early versions of prescription drug monitoring programs, or triplicate prescriptions, received less marketing from Purdue Pharma. These early versions of PDMPs were often referred to as "triplicate" programs. We exploit a different dimension of the initial marketing strategy of OxyContin: the fact that prescription opioids were initially promoted to the cancer pain market, with the plan to expand from there to the much larger non-cancer pain market. This produces variation in the marketing of opioids that tracks the cancer pain market and that we measure as commuting-zone level cancer mortality between 1994 and 1996.

Specification

Exploiting different dimensions of the initial marketing strategy of OxyContin translates into different empirical strategies. The event study specification proposed in each paper is given by:

$$\text{AELP:}\;\; y_{st} = \sum_{\tau=1983}^{2017} \beta_\tau \times \mathbf{1}(\text{ Nontriplicate })_s \times \mathbf{1}(\text{t} = \tau) + \tilde{\alpha}_s + \tilde{\gamma}_t + \zeta X_{st} + \varepsilon_{st} \ ,$$

$$\text{AB:}\;\; \Delta\, y_{sct} = \sum_{\tau=1989}^{2018} \phi_\tau \times \text{CancerMR}_{sct_0} \times \mathbf{1}(Year = \tau) + \gamma_{st} + \alpha\, \Delta\, X_{ct} + \upsilon_{ct} \ .$$

[1] Using state-level data implies 50 potential units of analysis: 49 continental US states and the District of Columbia. Using commuting-zone level data implies 740 potential units of analysis.

where s indexes state, t indexes time, c indexes commuting zone, t_0 defines the pre-OxiContin period (1994-1996), and τ indexes event time. AELP reproduces Equation (1) in Alpert et al. (2022), augmented to include the baseline controls (X_{st}) considered by the authors: the fraction of the population that is white non-Hispanic, Black non-Hispanic, Hispanic, the fraction ages 25–44, 45–64, 65+, the fraction with a college degree, and log population. This specification includes state (α_s) and year (γ_s) fixed effects. The indicator variable $\mathbf{1}(\text{Nontriplicate})_s$ is based on the initial triplicate status of the state in 1996.

AB reproduces Equation 2.3 of this paper and its terms have been defined before. Different from Alpert et al. (2022), this specification allows the inclusion of state times year fixed effects (γ_{st}) to account for important confounders at the state and year level. During the period of analysis there was relevant state-level variation in response to the opioid epidemic, such as the implementation of Prescription Drug Monitoring Programs (PDMP), the regulation of "pill mill" clinics, and the availability of naloxone. These policy changes were quite common, for example, between 2007 and 2013, 17 states implemented some version of a PDMP (Buchmueller and Carey, 2018*a*). Between 2001 and 2017, every US state passed a law that facilitates the widespread distribution and use of naloxone (Doleac and Mukherjee, 2019). These changes are likely related to both the levels of opioid in the population and downstream outcomes of the epidemic, biasing regression estimates.

Statistical precision

Next, we study how the two strategies compare in their statistical power to measure the impacts of the opioid epidemic on the outcomes of interest. To assess the statistical power of each model, we compute the *minimum detectable effect* and compare its distribution. We define this statistic as follows:

$$\text{Minimum detectable effect } = t_{critical} \times \frac{se_{\min}}{\text{SD}_y} \ , \tag{4.1}$$

where $t_{critical}$ corresponds to the xth percentile of the $t-$distribution and SD_y is the standard deviation of the outcome variable, e.g., opioid-related mortality or change in the share of the population on the SNAP. In the previous expression, $se_{\min}$ is given by:

$$se_{\min} = \min\{se_\tau\} = \min\{se_{\tau_0}, se_{\tau_1}, se_{\tau_2}, \ldots, se_{\tau_T}\} \ ,$$

where se_τ corresponds to the standard error of the $\hat{\beta}_\tau$ and $\hat{\phi}_\tau$ coefficients respectively, e.g., $se_\tau = se(\hat{\beta}_\tau)$. Thus, the minimum detectable effect of the vector of parameters $\boldsymbol{\beta}_\tau$ is the smallest effect size on the outcome variable y for which the researcher can reject the null hypothesis that the β_τ with the smallest variance equals zero. Analogously, the minimum detectable effect of the vector of parameters $\boldsymbol{\phi}_\tau$ requires the computation of $se_\tau = se(\hat{\phi}_\tau)$.

We construct the distribution of this statistic for each model $\{AELP, AB\}$. To do so, we perform S simulations of each model, and for each iteration s, we compute the minimum detectable effect as defined in Equation 4.1. In doing this exercise, we need to take a stand on the data-generating process for the outcome variable of interest (y and Δy, respectively). We consider alternative distributions of the outcome variable and parameter values. Table A.17 summarizes the results of this exercise and Figure A.22 presents selected distributions. We run 500 simulations for each proposed distribution of the outcome variable and work with a $t_{critical} = 1.96$. For example, the series labeled Beta(1,3) in Panel (a) corresponds to the distribution of the *minimum detectable effect* of the vector of parameters $\boldsymbol{\beta}_\tau$ when the outcome y follows a Beta (1,3) distribution. The Log-normal(8.85,7.15) closely captures the distribution of opioid-related mortality as defined by Alpert et al. (2022). Similarly, the Log-normal(0.02, 0.04) closely captures the distribution of the change in prescription opioid

mortality as defined in this paper. Our results in Figure A.22 and Table A.17 show that across distributions, the variation and specification presented in this paper has substantially higher statistical power. Specifically, at the median, we can identify effects that are 25% of the size that the model proposed by Alpert et al. (2022).

Definition of triplicate status

From our review of Purdue Pharma and other pharmaceutical companies' internal documents, we believe that when Purdue referred to "Triplicate States" it meant a group of nine states and not five as stated in Alpert et al. (2022). We base this statement on the fact that at least on two separate occasions, Purdue explicitly referred to triplicates as the *"nine states"* (Figure A.23), and to our knowledge, *never* mentioned only five. A footnote in Alpert et al. (2022) comments on *one* of these references to the nine states and deemed it an incorrect reference by Purdue. Academic documents that explain the prescription drug monitoring programs that were in effect at the time also refer to a group of nine states. These documents are more precise in their language and refer to these programs as multiple-copy prescription programs (Joranson, D et al., 2020 and Fishman, S et al., 2004). Similar to today's PDMPs, different states had different versions of the program, but the informal industry name for these programs was "triplicate programs". In an internal email between Mallinckrodt sales specialists, also disclosed as part of the opioid litigation, one sales specialist lists and explains to the other the history of the triplicate programs and lists the original nine states (Figure A.24). These are California, Hawaii, Idaho, Indiana, Michigan, Illinois, New York, Rhode Island, and Texas.[2]

In light of these alternative definitions of the group of states with triplicate programs, we inspect the time trends of overdose mortality in triplicate states and replicate the main

[2]Mallinckrodt is a pharmaceutical company that is also part of the opioid litigation for their role in the opioid epidemic. More precisely, *"Collectively, Purdue, Actavis, Cephalon, Janssen, Endo, Insys, and Mallinckrodt are referred to as "Marketing Defendants"* Case No. 17-md-2804. United States District Court for the Northern District of Ohio Eastern Division.

results in Alpert et al. (2022).[3] First, in Figure A.25 we inspect patterns in the raw data. Panels (a) and (b) show the evolution of overdose mortality in five triplicate states and nine triplicate states, respectively, compared to the evolution in the rest of the country. Using the alternative definition of triplicates provides less clear evidence that "triplicate states" fare better regarding overdose mortality.

Event studies models in Figure A.26 suggest a similar story. While the main results are not overturned, they are attenuated and are often equal to zero statistically, suggesting a smaller effect of the triplicate status on overdose mortality. We estimate the event studies with and without population weights. The unweighted version is more sensitive to the definition of triplicate status, which is natural since even though the sample of "treated" states is increasing by 80%, the treated population is only changing by 21%. Finally, Table A.18 replicates the main estimate in Alpert et al. (2022). Consistent with the event study estimates, results are attenuated when using the alternative definition and are more sensitive when regressions are not weighted by population.

Finally, we test whether our instrument has predictive power in the group of triplicates states as defined by Alpert et al. (2022). We find that the positive relationship between mid-nineties cancer mortality and the supply of opioids is present in both triplicate and non-triplicate states (Figure A.20). In fact, the first stage is stronger in the five triplicate states. This result is consistent with a story in which pharmaceutical companies need to be more strategic in promoting opioids in places where they face additional barriers and do not avoid the largest markets in the country (For example, California, Texas, New York, and Illinois).[4] Anecdotal evidence is consistent with this fact: All five "triplicate" states filed lawsuits against Purdue Pharma and other pharmaceutical companies for their direct

[3]We define overdose deaths as deaths involving underlying cause of death ICD-9 codes E850-E858, E950.0-E950.5, E962.0, or E980.0-E980.5 and ICD-10 codes X40-X44, X60-64, X85, or Y10-Y14.

[4]The first stage coefficient is 1.542 in the five triplicate states defined in Alpert et al. (2022), and 0.917 in the other states. This difference is not statistically different, and the p-value of such test is 0.203.

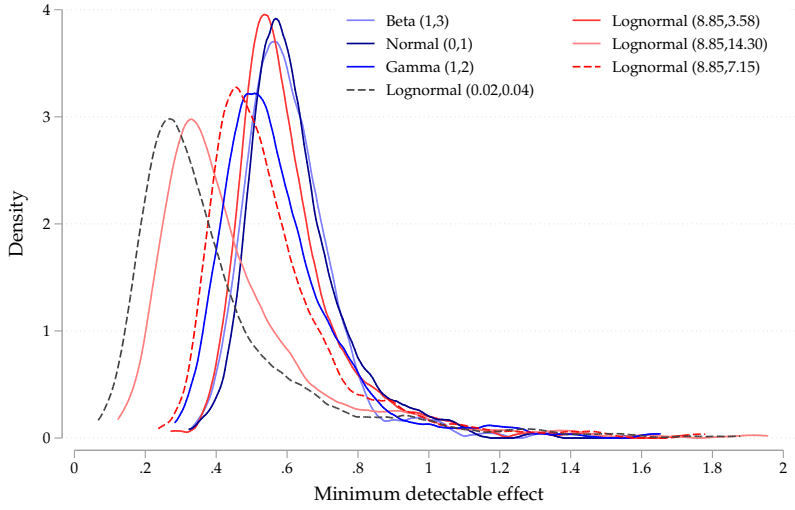
responsibility in the unfolding of the opioid epidemic.[5] We corroborate this, in the reduced form estimates, our results are indistinguishable, both qualitatively and statistically, in the five states triplicate states and the rest of the states, as panel b of Figure A.20 shows.[6]

[5]For example, "New York State is in the throes of an opioid epidemic that has ravaged the lives of its residents and drained its public coffers for more than two decades. This statewide catastrophe happened because the Defendants in this case deliberately betrayed those duties through a persistent course of fraudulent and illegal misconduct, in order to profiteer from the plague they knew would be unleashed"—the People of New York, -against-Purdue Pharma et al.; March 28, 2019.

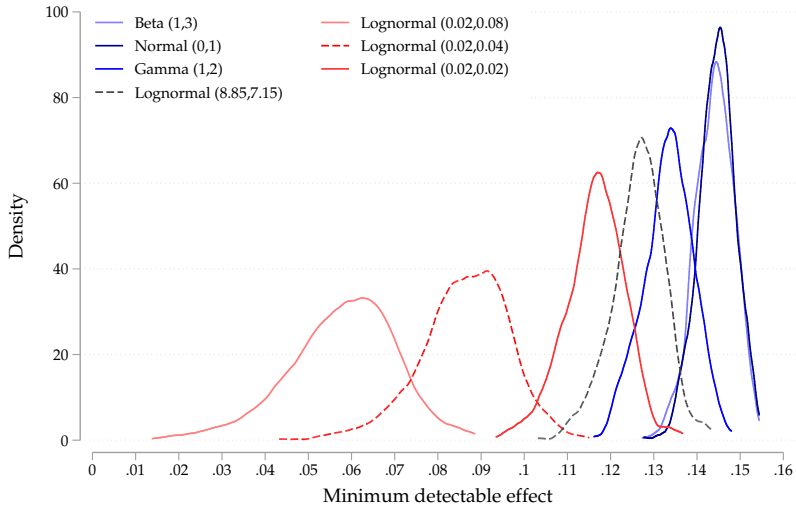
[6]The reduced form coefficient is 0.0072 in the five triplicate states defined in Alpert et al. (2022), and 0.0073 in the other states. This difference is not statistically different, and the p-value of such test is 0.97.

Figure A.22: Distribution of the Minimum Detectable Effect: Alternative Specifications

(a) State-level Variation: Triplicate programs

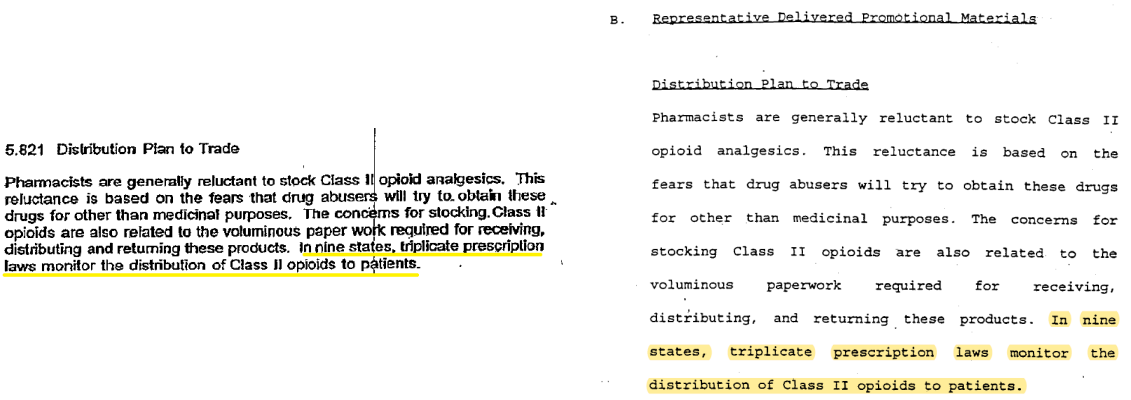


(b) Commuting-zone level variation: Cancer market



Notes: Panel (a) shows the distribution of the minimum detectable effect for the AELP model under alternative distributions of the outcome variable. For example, the series labeled Beta(1,3) corresponds to the distribution of the *minimum detectable effect* of the vector of parameters $\boldsymbol{\beta}_\tau$ when the outcome y follows a Beta(1,3) distribution. The Log-normal(8.85,7.15) closely captures the distribution of opioid-related mortality as defined by Alpert et al. (2022). The additional Log-normal distributions have the same mean but change the variance. Similarly, panel (b) shows the distribution of the minimum detectable effect of the vector of parameters $\boldsymbol{\phi}_\tau$, i.e., it corresponds to this paper's model under alternative distributions of the outcome variable. The Log-normal(0.02, 0.04) closely captures the distribution of the change in prescription opioid mortality as defined in this paper.

Figure A.23: Reference to Nine Triplicate States in OxyContin Launch Plan



5.821 Distribution Plan to Trade

Pharmacists are generally reluctant to stock Class II opioid analgesics. This reluctance is based on the fears that drug abusers will try to obtain these drugs for other than medicinal purposes. The concerns for stocking Class II opioids are also related to the voluminous paper work required for receiving, distributing and returning these products. In nine states, triplicate prescription laws monitor the distribution of Class II opioids to patients.

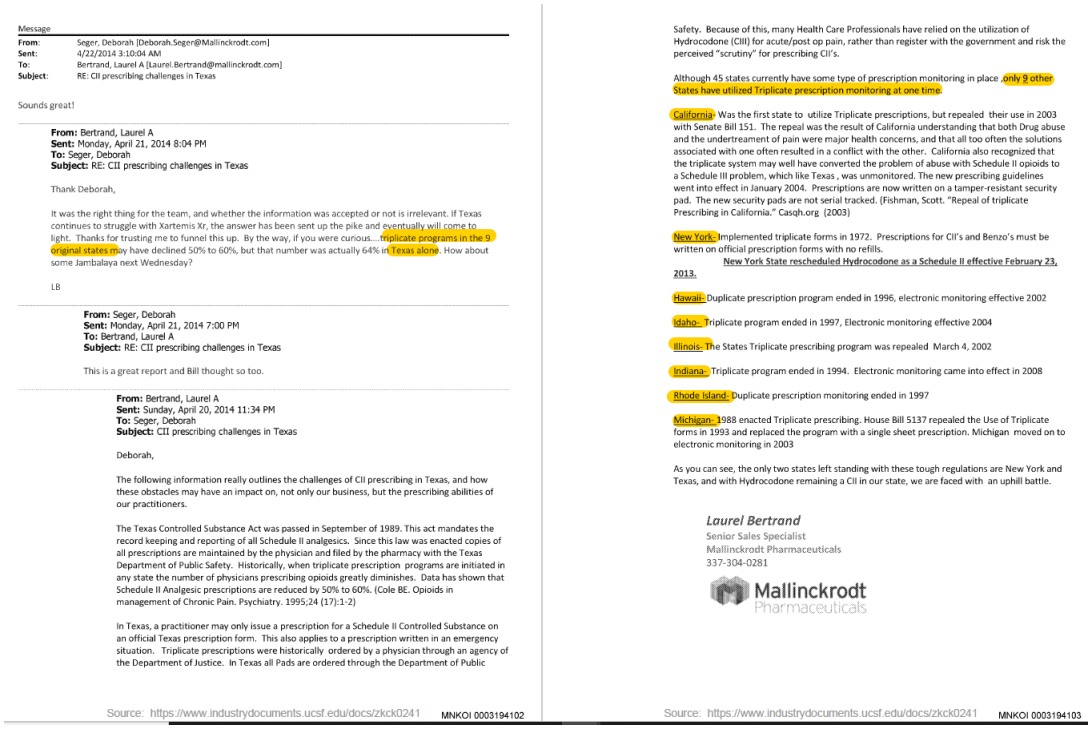
B. Representative Delivered Promotional Materials

Distribution Plan to Trade

Pharmacists are generally reluctant to stock Class II opioid analgesics. This reluctance is based on the fears that drug abusers will try to obtain these drugs for other than medicinal purposes. The concerns for stocking Class II opioids are also related to the voluminous paperwork required for receiving, distributing, and returning these products. In nine states, triplicate prescription laws monitor the distribution of Class II opioids to patients.

Notes: This figure shows extracts of OxyContin Launch plans. The left panel reproduces a segment of the OxyContin Launch Plan, page 27 September 27th 1995. The right panel is an extract from OxyContin Budget Plan 1996, page 29. This figure is referenced in Appendix 1.3.

Figure A.24: Reference to Nine Triplicate States in Internal Communications



Message

From: Seger, Deborah [Deborah.Seger@Mallinckrodt.com]
Sent: 4/22/2014 3:10:04 AM
To: Bertrand, Laurel A [Laurel.Bertrand@mallinckrodt.com]
Subject: RE: CII prescribing challenges in Texas

Sounds great!

From: Bertrand, Laurel A
Sent: Monday, April 21, 2014 8:04 PM
To: Seger, Deborah
Subject: RE: CII prescribing challenges in Texas

Thank Deborah,

It was the right thing for the team, and whether the information was accepted or not is irrelevant. If Texas continues to struggle with Xartemis Xr, the answer has been sent up the pike and eventually will come to light. Thanks for trusting me to funnel this up. By the way, if you were curious…triplicate programs in the 9 original states may have declined 50% to 60%, but that number was actually 64% in Texas alone. How about some Jambalaya next Wednesday?

LB

From: Seger, Deborah
Sent: Monday, April 21, 2014 7:00 PM
To: Bertrand, Laurel A
Subject: RE: CII prescribing challenges in Texas

This is a great report and Bill thought so too.

From: Bertrand, Laurel A
Sent: Sunday, April 20, 2014 11:34 PM
To: Seger, Deborah
Subject: CII prescribing challenges in Texas

Deborah,

The following information really outlines the challenges of CII prescribing in Texas, and how these obstacles may have an impact on, not only our business, but the prescribing abilities of our practitioners.

The Texas Controlled Substance Act was passed in September of 1989. This act mandates the record keeping and reporting of all Schedule II analgesics. Since this law was enacted copies of all prescriptions are maintained by the physician and filed by the pharmacy with the Texas Department of Public Safety. Historically, when triplicate prescription programs are initiated in any state the number of physicians prescribing opioids greatly diminishes. Data has shown that Schedule II Analgesic prescriptions are reduced by 50% to 60%. (Cole BE. Opioids in management of Chronic Pain. Psychiatry. 1995;24 (17):1-2)

In Texas, a practitioner may only issue a prescription for a Schedule II Controlled Substance on an official Texas prescription form. This also applies to a prescription written in an emergency situation. Triplicate prescriptions were historically ordered by a physician through an agency of the Department of Justice. In Texas all Pads are ordered through the Department of Public

Source: https://www.industrydocuments.ucsf.edu/docs/zkck0241 MNKOI 0003194102

Safety. Because of this, many Health Care Professionals have relied on the utilization of Hydrocodone (CIII) for acute/post op pain, rather than register with the government and risk the perceived "scrutiny" for prescribing CII's.

Although 45 states currently have some type of prescription monitoring in place ,only 9 other States have utilized Triplicate prescription monitoring at one time.

California- Was the first state to utilize Triplicate prescriptions, but repealed their use in 2003 with Senate Bill 151. The repeal was the result of California understanding that both Drug abuse and the undertreament of pain were major health concerns, and that all too often the solutions associated with one often resulted in a conflict with the other. California also recognized that the triplicate system may well have converted the problem of abuse with Schedule II opioids to a Schedule III problem, which like Texas , was unmonitored. The new prescribing guidelines went into effect in January 2004. Prescriptions are now written on a tamper-resistant security pad. The new security pads are not serial tracked. (Fishman, Scott. "Repeal of triplicate Prescribing in California." Casqh.org (2003)

New York- Implemented triplicate forms in 1972. Prescriptions for CII's and Benzo's must be written on official prescription forms with no refills.
New York State rescheduled Hydrocodone as a Schedule II effective February 23, 2013.

Hawaii- Duplicate prescription program ended in 1996, electronic monitoring effective 2002

Idaho- Triplicate program ended in 1997, Electronic monitoring effective 2004

Illinois- The States Triplicate prescribing program was repealed March 4, 2002

Indiana- Triplicate program ended in 1994. Electronic monitoring came into effect in 2008

Rhode Island- Duplicate prescription monitoring ended in 1997

Michigan- 1988 enacted Triplicate prescribing. House Bill 5137 repealed the Use of Triplicate forms in 1993 and replaced the program with a single sheet prescription. Michigan moved on to electronic monitoring in 2003

As you can see, the only two states left standing with these tough regulations are New York and Texas, and with Hydrocodone remaining a CII in our state, we are faced with an uphill battle.

Laurel Bertrand
Senior Sales Specialist
Mallinckrodt Pharmaceuticals
337-304-0281

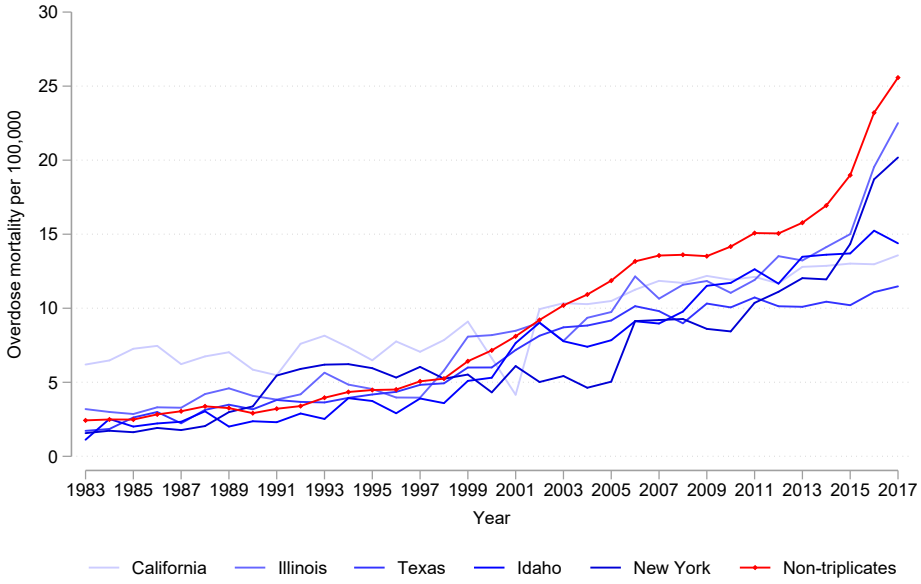
Mallinckrodt Pharmaceuticals

Source: https://www.industrydocuments.ucsf.edu/docs/zkck0241 MNKOI 0003194103

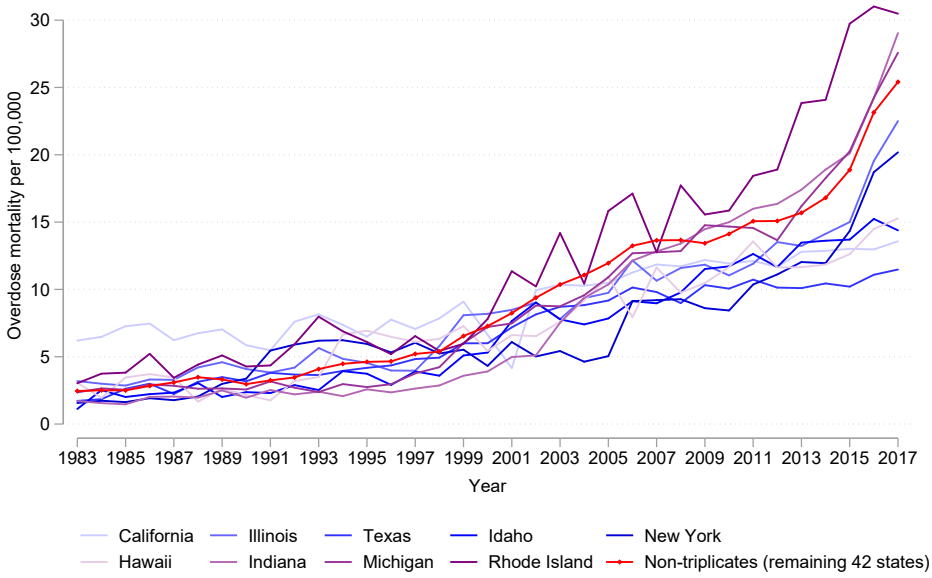
Notes: This figure shows extracts of the internal email from the opioid litigation with details on the list of triplicate states. This figure is referenced in Appendix 1.3.

Figure A.25: All Drug Overdose Mortality by Triplicate Status.

(a) Triplicate definition as in Alpert et al. (2022).



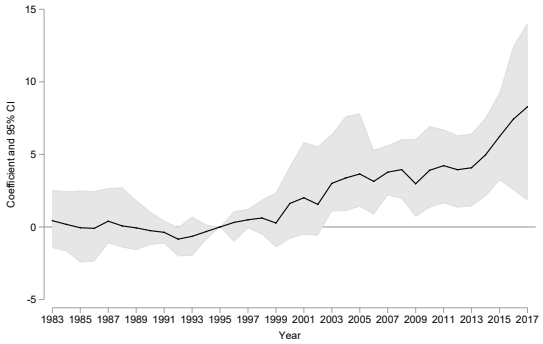
(b) Alternative triplicate definition



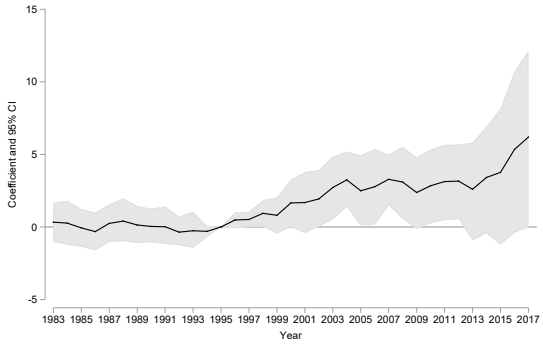
Notes: Time series for all drug overdose mortality. Panel (a) defines triplicates as California, Idaho, Illinois, New York, and Texas. Panel (b) adds Hawaii, Indiana, Michigan, and Rhode Island for a total of 9 triplicate states. This figure is referenced in Appendix 1.3.

Figure A.26: All Drug Overdose Mortality By Triplicate Status - Unweighted analysis.

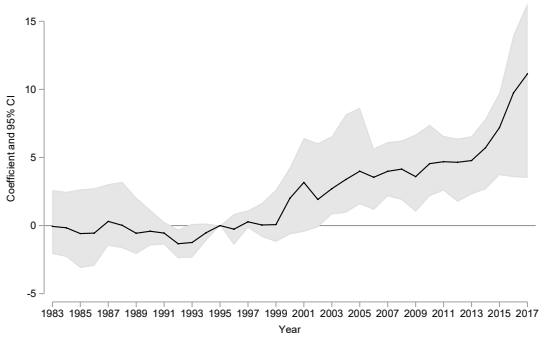
(a) 5 Triplicates unweighted



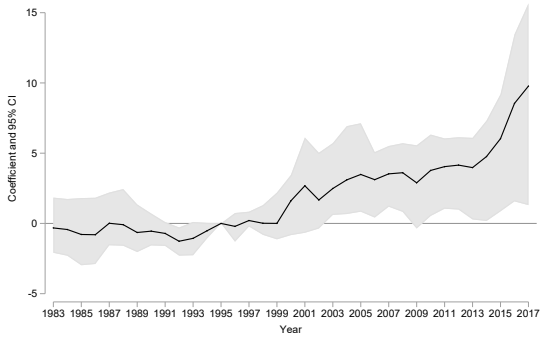
(b) 9 Triplicates unweighted



(c) 5 Triplicates weighted



(d) 9 Triplicates weighted



Notes: Figures in panels (a) and (c) reproduce Figure 4 in Alpert et al. (2022). Panels (b) and (d) present the analysis using the alternative definition of triplicate states: we add Hawaii, Indiana, Michigan, and Rhode Island for a total of 9 triplicate states. Event study models include state and year fixed effects. 95% confidence intervals are generated using a clustered (at state) wild bootstrap. Estimates are normalized to zero in 1995. This figure is referenced in Appendix 1.3.

Table A.17: Distribution of the Minimum Detectable Effect for Alternative Specifications and Distributions of the Outcome Variables

	Outcome variable		Minimun effect size							
	Mean	SD	Min	1^{st} pctile	5^{th} pctile	Mean	Median	95^{th} pctile	99^{th} pctile	Max
Distribution	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Alpert et al. (2022) model										
Beta (2, 2)	0.50	0.22	0.36	0.41	0.46	0.62	0.60	0.85	0.96	1.24
Beta (1, 3)	0.25	0.19	0.33	0.38	0.44	0.61	0.59	0.82	1.05	1.34
Gamma (1, 2)	2.00	1.99	0.28	0.33	0.38	0.60	0.54	0.95	1.53	3.03
Gamma (2, 2)	4.00	2.82	0.31	0.36	0.40	0.61	0.58	0.90	1.31	1.85
Normal (0, 1)	0.00	1.00	0.32	0.39	0.46	0.63	0.60	0.91	1.08	1.58
Log-normal (2.5, 0.5)	2.50	0.50	0.30	0.39	0.46	0.62	0.59	0.92	1.12	1.59
Log-normal (7, 5)	7.01	5.02	0.24	0.31	0.39	0.59	0.53	1.00	1.64	3.87
Log-normal (8.85, 3.58)	8.80	3.56	0.27	0.38	0.44	0.61	0.57	0.94	1.32	2.26
Log-normal (8.85, 7.15)	8.84	7.14	0.24	0.29	0.36	0.58	0.51	1.01	1.77	4.40
Log-normal (8.85, 14.3)	8.86	14.28	0.12	0.17	0.22	0.51	0.39	1.09	2.51	8.92
Log-normal (0.02, 0.04)	0.02	0.04	0.07	0.11	0.17	0.48	0.33	1.14	2.85	11.16
Panel B: Arteaga and Barone (2023) model										
Beta (2, 2)	0.50	0.22	0.13	0.14	0.14	0.15	0.15	0.15	0.16	0.16
Beta (1, 3)	0.25	0.19	0.13	0.13	0.14	0.14	0.14	0.15	0.15	0.15
Gamma (1, 2)	2.00	2.00	0.12	0.12	0.12	0.13	0.13	0.14	0.15	0.15
Gamma (2, 2)	4.00	2.83	0.12	0.13	0.13	0.14	0.14	0.15	0.15	0.15
Normal (0, 1)	0.02	0.04	0.13	0.13	0.14	0.14	0.14	0.15	0.15	0.15
Log-normal (2.5, 0.5)	2.50	0.50	0.13	0.13	0.14	0.14	0.14	0.15	0.15	0.15
Log-normal (7, 5)	7.00	5.00	0.11	0.11	0.12	0.13	0.13	0.14	0.14	0.15
Log-normal (0.02, 0.02)	0.02	0.02	0.09	0.10	0.10	0.12	0.12	0.13	0.13	0.14
Log-normal (0.02, 0.04)	0.02	0.04	0.04	0.06	0.07	0.09	0.09	0.10	0.11	0.12
Log-normal (0.02, 0.08)	0.02	0.08	0.01	0.02	0.04	0.06	0.06	0.08	0.09	0.09
Log-normal (8.85, 7.15)	8.83	7.11	0.10	0.11	0.12	0.13	0.13	0.14	0.14	0.14

Notes: This table presents summary statistics for the distribution of the minimum detectable effect. Panel A considers the model proposed by Alpert et al. (2022) and panel B considers the model proposed in this paper. Columns (1) and (2) present the mean and standard deviation of the simulated outcome of interest (y and Δy respectively). Columns (3) to (10) present moments of the distribution of the minimum detectable effect, pctile stands for percentile. This table is referenced in Appendix 1.3.

Table A.18: Replication of Table 1 in Alpert et al. (2022)

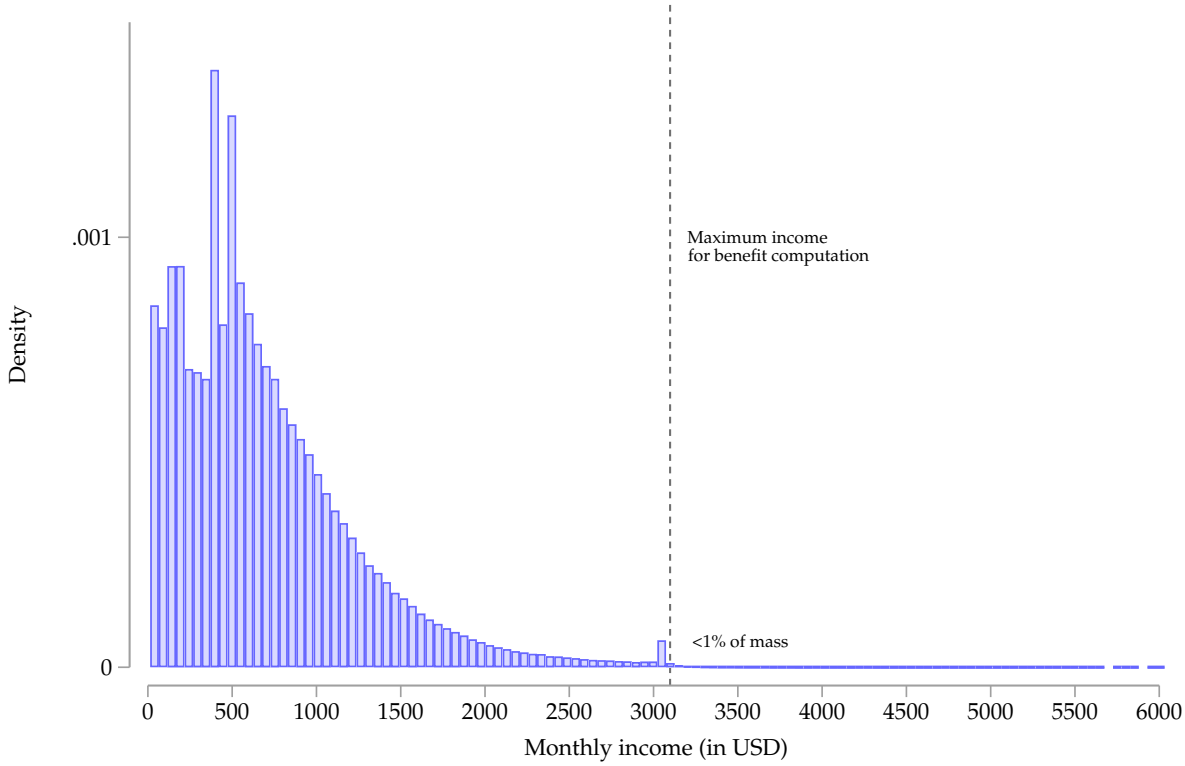
Triplicate state group (n)	Nine	Five	Nine	Five
Nontriplicate ×	(1)	(2)	(3)	(4)
1996–2000	0.998***	1.173	0.711	1.229**
SE, CI	[0.356]	[0.390, 2.374]	[0.538]	[0.017, 2.483]
Coeff. change		14.9%		42.1%
2001–2010	2.257**	3.667**	1.998**	3.232**
SE, CI	[0.913]	[1.521, 6.210]	[0.994]	[1.011, 5.318]
Coeff. change		38.5%		38.2%
2011–2017	2.793	6.061**	3.203**	4.714***
SE, CI	[1.891]	[2.812, 9.371]	[1.337]	[1.811, 7.253]
Coeff. change		53.9%		32.1%
Weighted	No	No	Yes	Yes
Covariates	No	No	Yes	Yes
Region-time dummies	No	No	Yes	Yes
Observations	1,377	1,377	1,377	1,377

Notes: Columns (2) and (4) of this table reproduce columns (1) and (4) of Table 1 in Alpert et al. (2022) respectively. Columns (1) and (3) present the analysis using an alternative definition of triplicate status. This table is referenced in Appendix 1.3.

2. Appendix to "On the Design of Paid Sick Leave: A Structural Approach"

2.1 Additional Figures

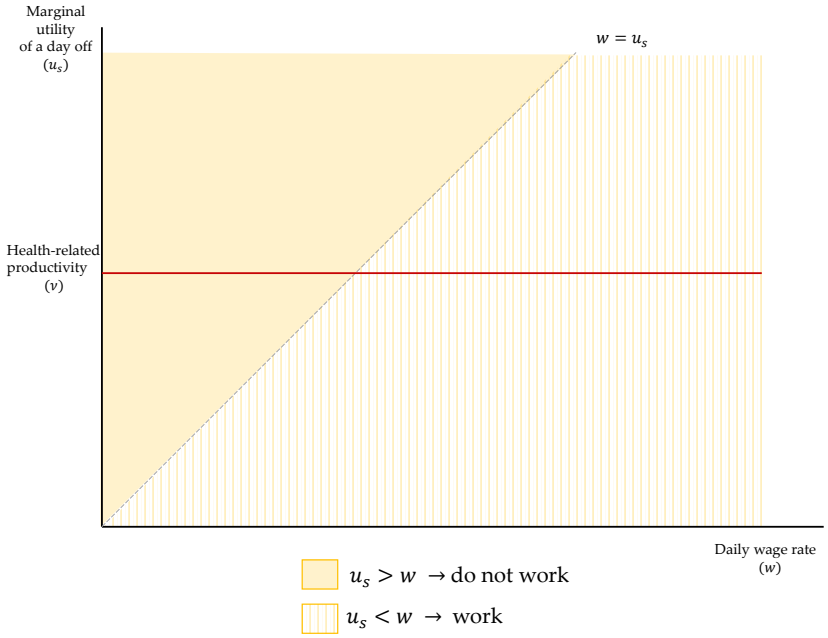
Figure A.27: Distribution of monthly income (in USD). Workers eligible to file a sick leave claim



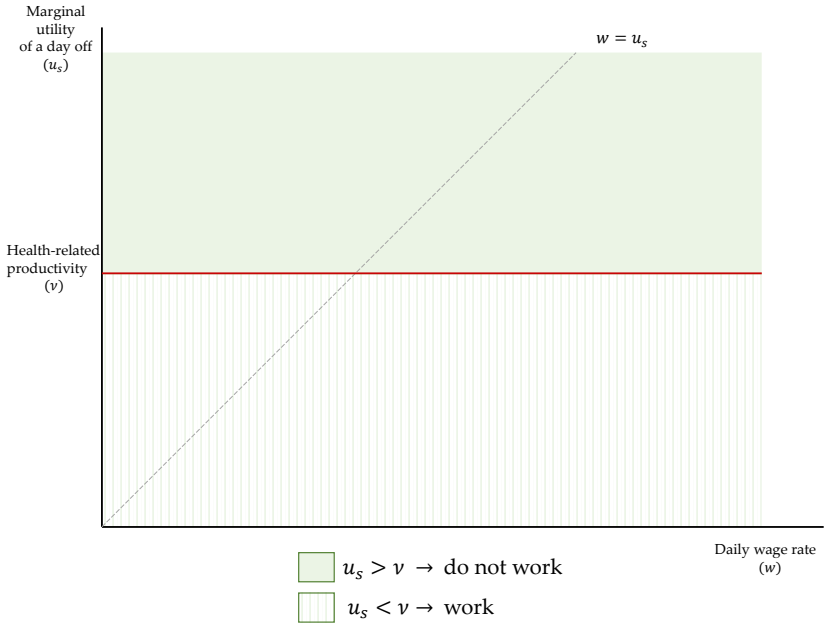
Notes: This figure shows the distribution of monthly income (in USD) for workers eligible to file a sick leave claim in 2017. The vertical line indicates the income level associated with the maximum benefit threshold. This figure is referenced in Section 2.2.

Figure A.28: Worker vs Social Planner trade-offs: externalities and no insurance provision

(a) Worker trade-off



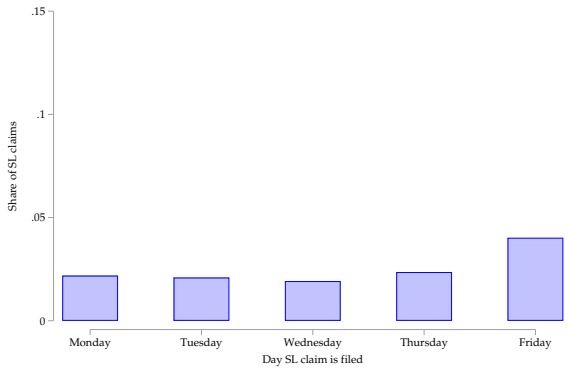
(b) Social Planner trade-off: optimal employment



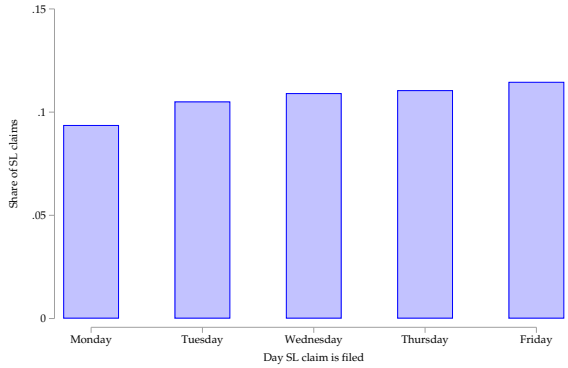
Notes: Panel (a) summarizes workers' choices. Absent of sick pay, worker i takes a day off if $u_s^i > w^i$, this corresponds to the solid fill area. Note that some workers optimally choose to work even if their value of a day off is above their productivity (ν). This is the pattern fill area located above the horizontal line. This is a consequence of the fact that wages do not longer reflect productivity. Panel (b) shows the optimal employment decision. This trade-off compares the productivity of working with the value of a day off. Absent of sick pay, it would be efficient that worker i takes a day off if her valuation is above her marginal product when sick, i.e., when $u_s^i > \nu$. This corresponds to the solid fill area located above the horizontal line. Note that some workers would find optimal to not work regardless: those with $u_s^i > w^i$. This figure is referenced in Section 3.3.

Figure A.29: Distribution of sick leave claims by duration and day of the week.

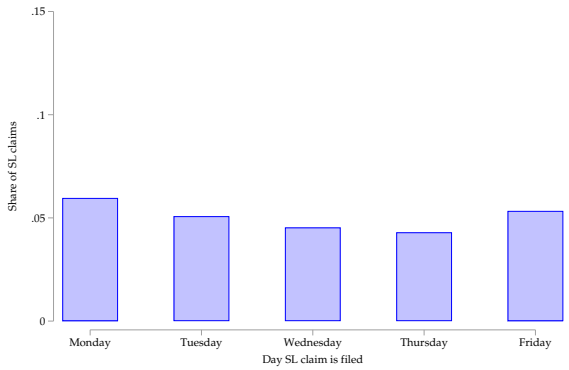
(a) Eight and nine days long



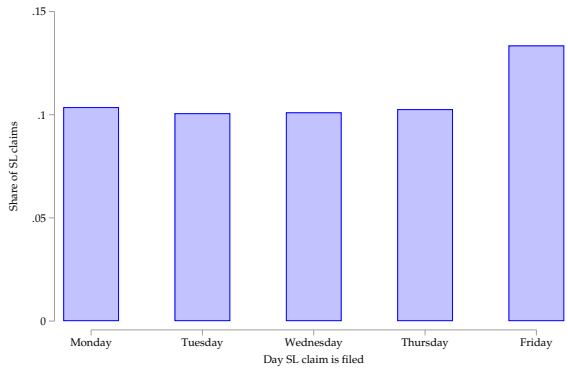
(b) Ten and eleven days long



(c) Twelve and thirteen days long



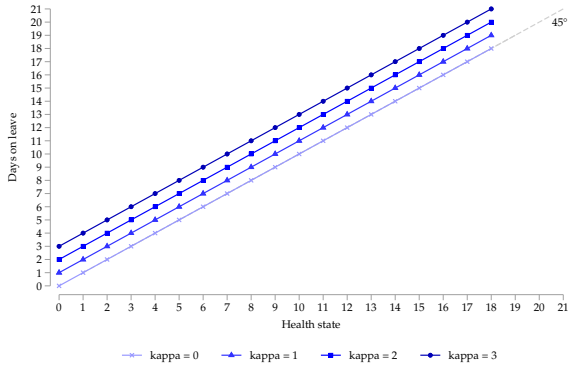
(d) Fourteen and fifteen days long



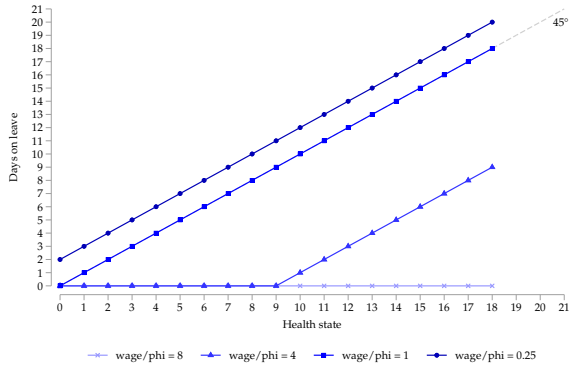
Notes: This figure shows the share of sick leave claims of duration s filed on each day of the week. Each panel aggregates sick leave claims with consecutive duration as stated in the title. This figure is referenced in Section 4.2.

Figure A.30: Sick Pay Utilization with Linear Contract

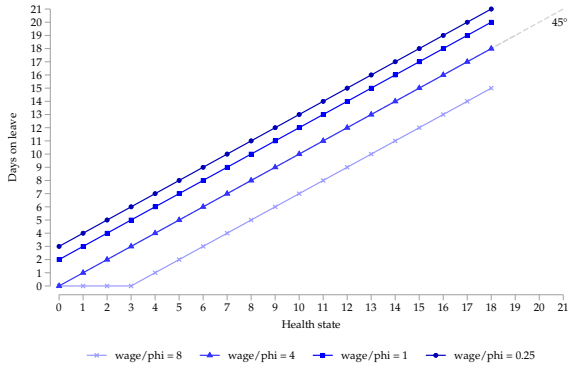
(a) Full coverage ($b = 1$)



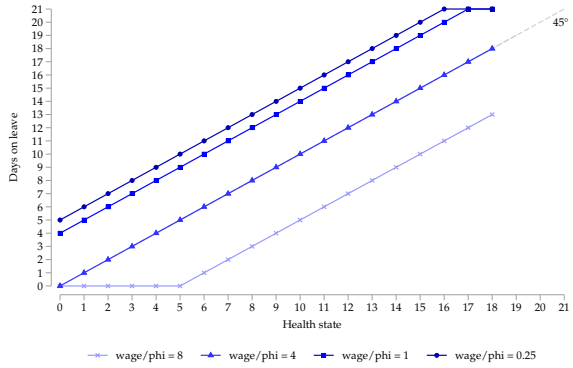
(b) Unpaid sick leave ($b = 0$) and $\kappa = 3$



(c) Partial coverage ($b = 0.75$) and $\kappa = 3$



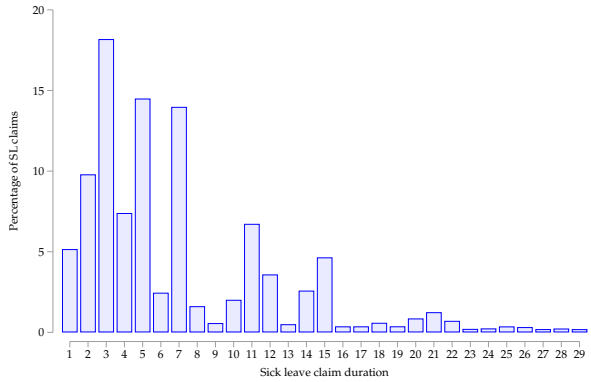
(d) Partial coverage ($b = 0.75$) and $\kappa = 5$



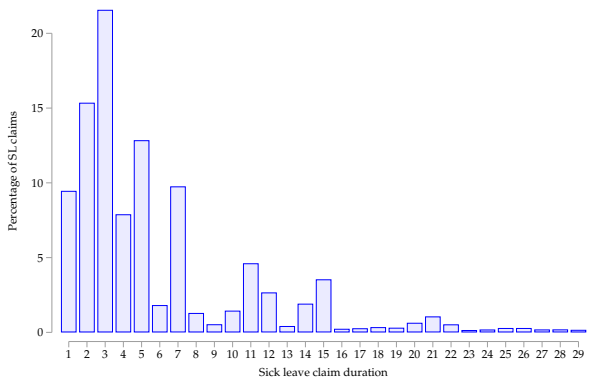
Notes: This figure shows the optimal demand of days on leave $s^*(\theta)$ as a function of worker's health status (θ) under the assumption of linear contracts with different levels of coverage. This figure is referenced in Section 3..

Figure A.31: Histogram of days on leave by worker characteristics

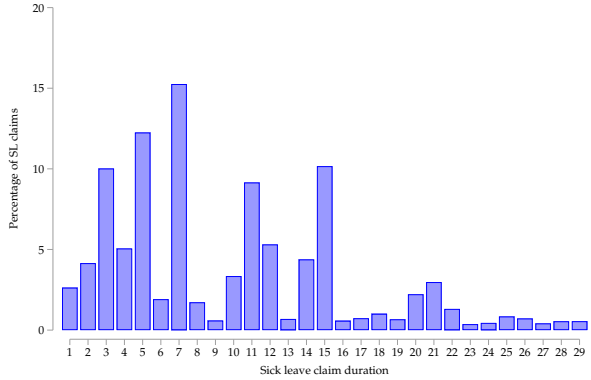
(a) 25–34 years old, blue-collar



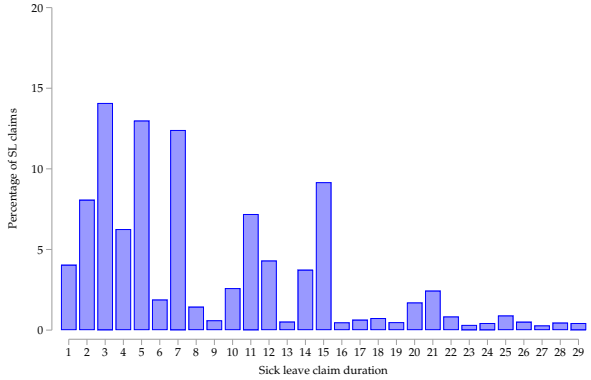
(b) 25–34 years old, white-collar



(c) 55–64 years old, blue-collar



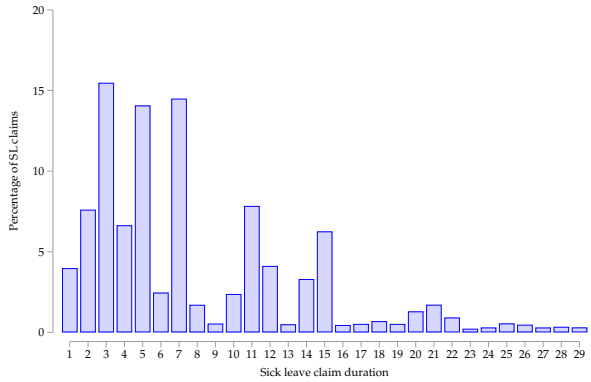
(d) 55–64 years old, white-collar



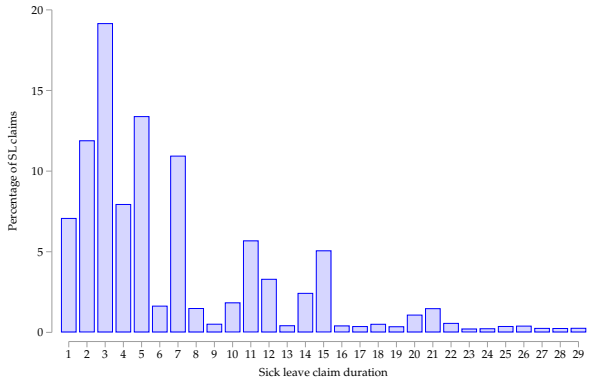
Notes: This figure shows the distribution of days on leave by worker age and occupation for the youngest and oldest group of workers. The sample includes male private-sector employees. Blue-collar workers refers to workers who engage in hard manual labor, typically agriculture, manufacturing, construction, mining, or maintenance. White-collar workers refers to workers whose daily work activities do not involve manual labor—e.g., teachers or administrative staff. Additional groups are presented in Appendix Figure A.32. This figure is referenced in Section 2.4 and in Section 4.1.

Figure A.32: Histogram of days on leave by workers characteristics

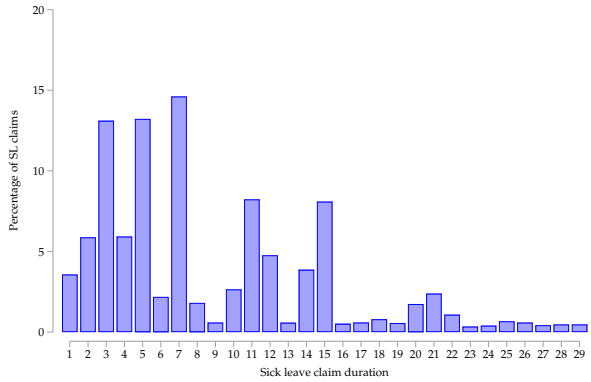
(a) 35-44 years old. Blue-collar



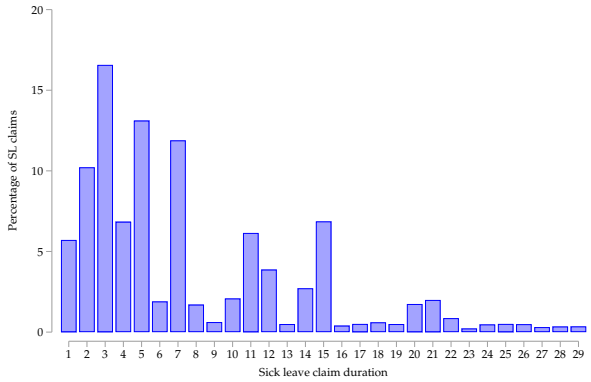
(b) 35-44 years old. White collar



(c) 45-54 years old. Blue-collar



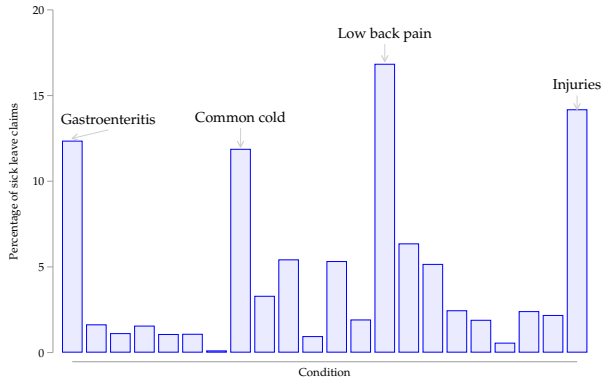
(d) 45-54 years old. White collar



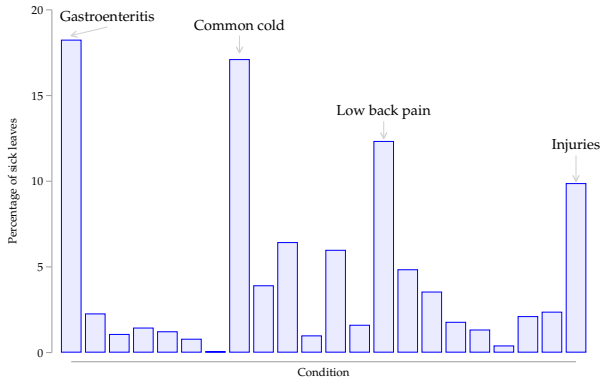
Notes: This figure shows the distribution of days on leave by workers' age and occupation for the youngest and oldest group of workers. Sample includes male private-sector employees. Blue-collar worker refers to workers who engage in hard manual labor, typically agriculture, manufacturing, construction, mining, or maintenance. White-collar worker refers to workers whose daily work activities do not involve manual labor—e.g., teachers or administrative staff. Additional groups are presented in Appendix Figure A.31. This figure is referenced in Section 2.3 and in Section 4.1.

Figure A.33: Histogram of days on leave by workers characteristics

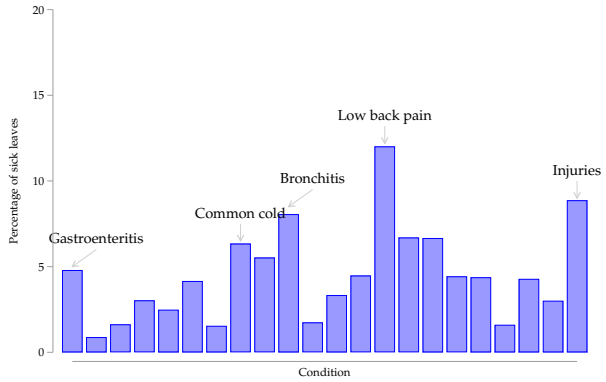
(a) 25-34 years old. Blue-collar



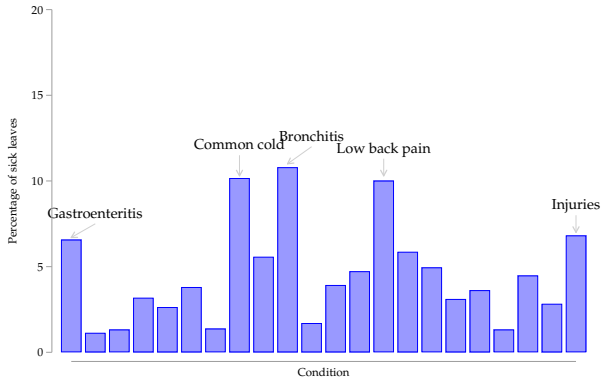
(b) 25-34 years old. White collar



(c) 55-64 years old. Blue-collar



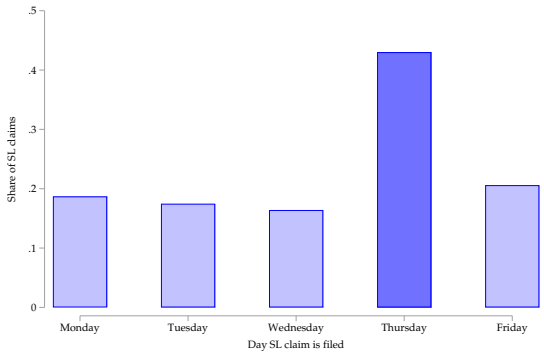
(d) 55-64 years old. White collar



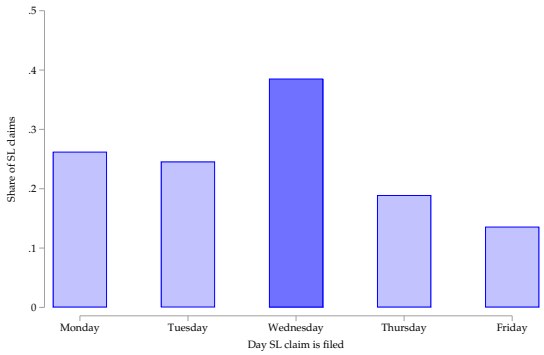
Notes: This figure shows the probability that a worker contracts disease d by by workers' age and occupation for the youngest and oldest group of workers. Diseases are ordered as presented in Table A.28. Sample includes male private-sector employees. Blue-collar worker refers to workers who engage in hard manual labor, typically agriculture, manufacturing, construction, mining, or maintenance. White-collar worker refers to workers whose daily work activities do not involve manual labor—e.g., teachers or administrative staff. This figure is referenced in Section 4.2.

Figure A.34: Identification of compliance costs parameter: Sick leave claims by duration and day of the week.
Health shock (θ) equals 2-days-long.

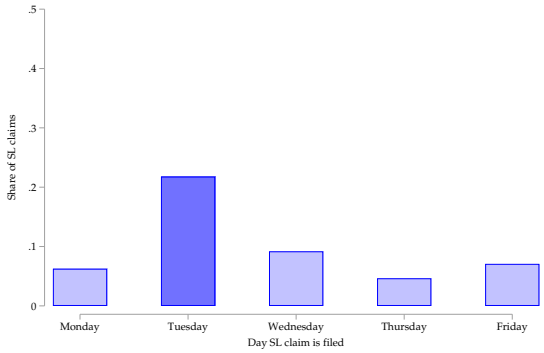
(a) Deviation = 0, duration = 2



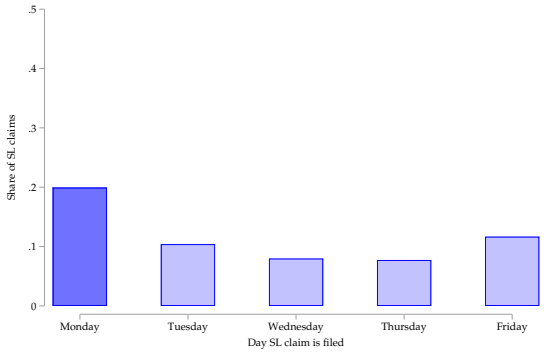
(b) Deviation = 1, duration = 3



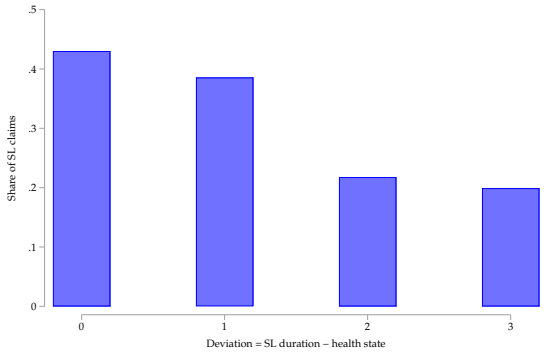
(c) Deviation = 2, duration = 4



(d) Deviation = 3 , duration = 5



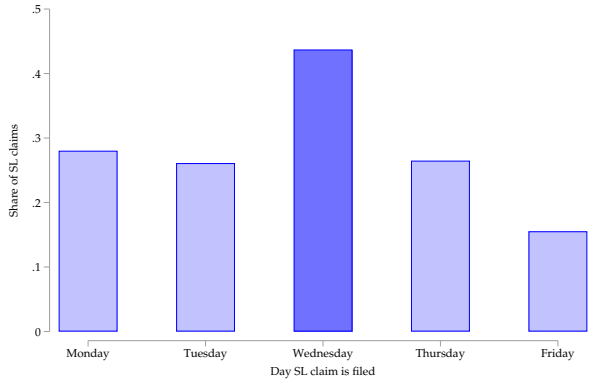
(e) Aggregate across days of the week



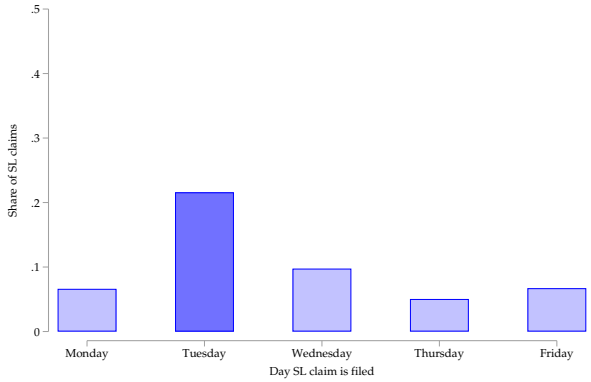
Notes: Panels (a) to (d) show the share of sick leave claims with duration s for workers whose main diagnose would implied a health state of 2 days on leave. Panel (e) aggregates the share of sick leave claims across days of the week, including only the weekend-streak combinations, e.g., from panel (a) I only consider the share for Thursday. This figure is referenced in Section 4.2.

Figure A.35: Identification of compliance costs parameter: Sick leave claims by duration and day of the week.
Health shock (θ) equals 3-days-long.

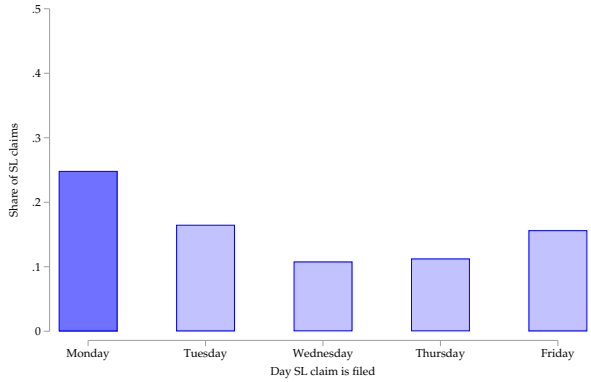
(a) Deviation = 0, duration = 3



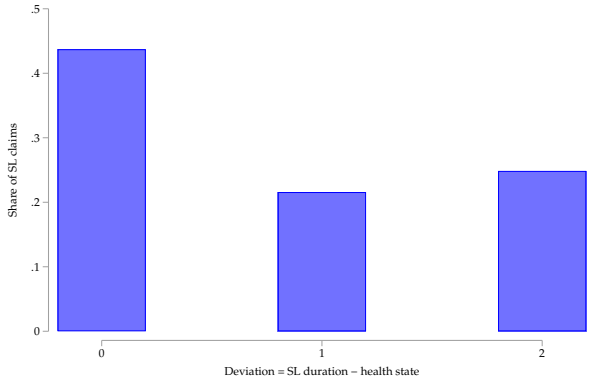
(b) Deviation = 1, duration = 4



(c) Deviation = 2, duration = 5

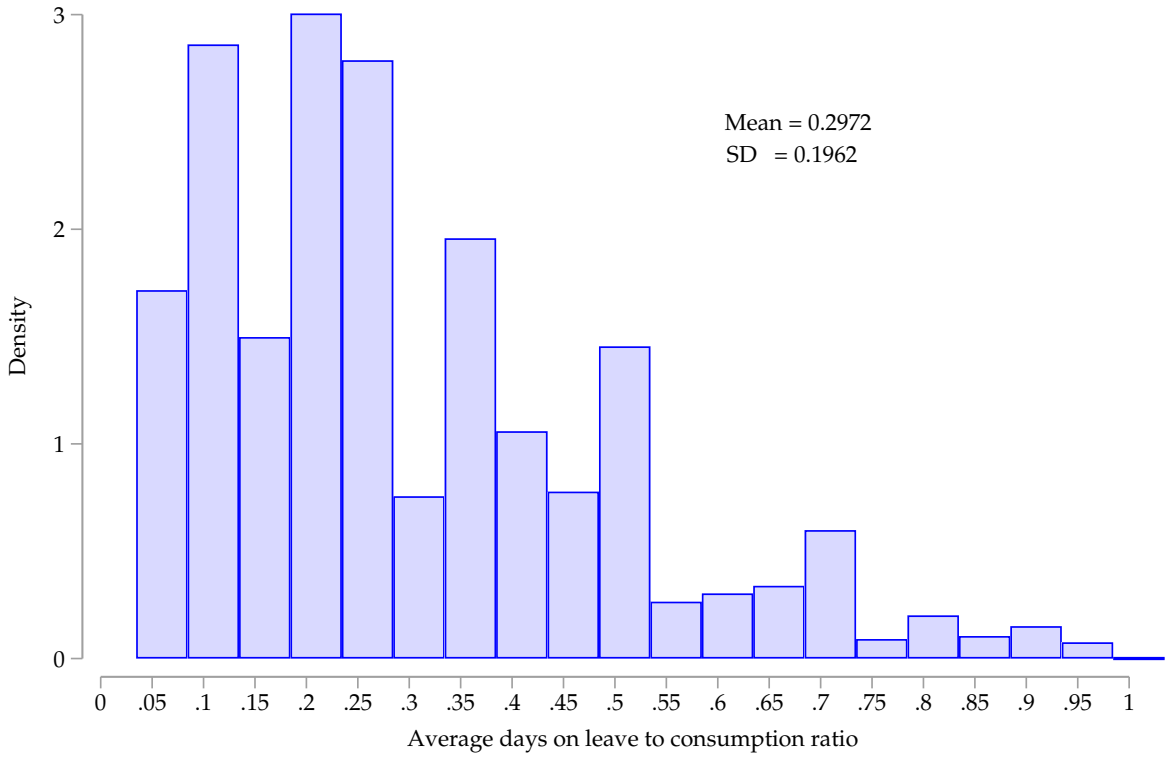


(d) Aggregate across days of the week



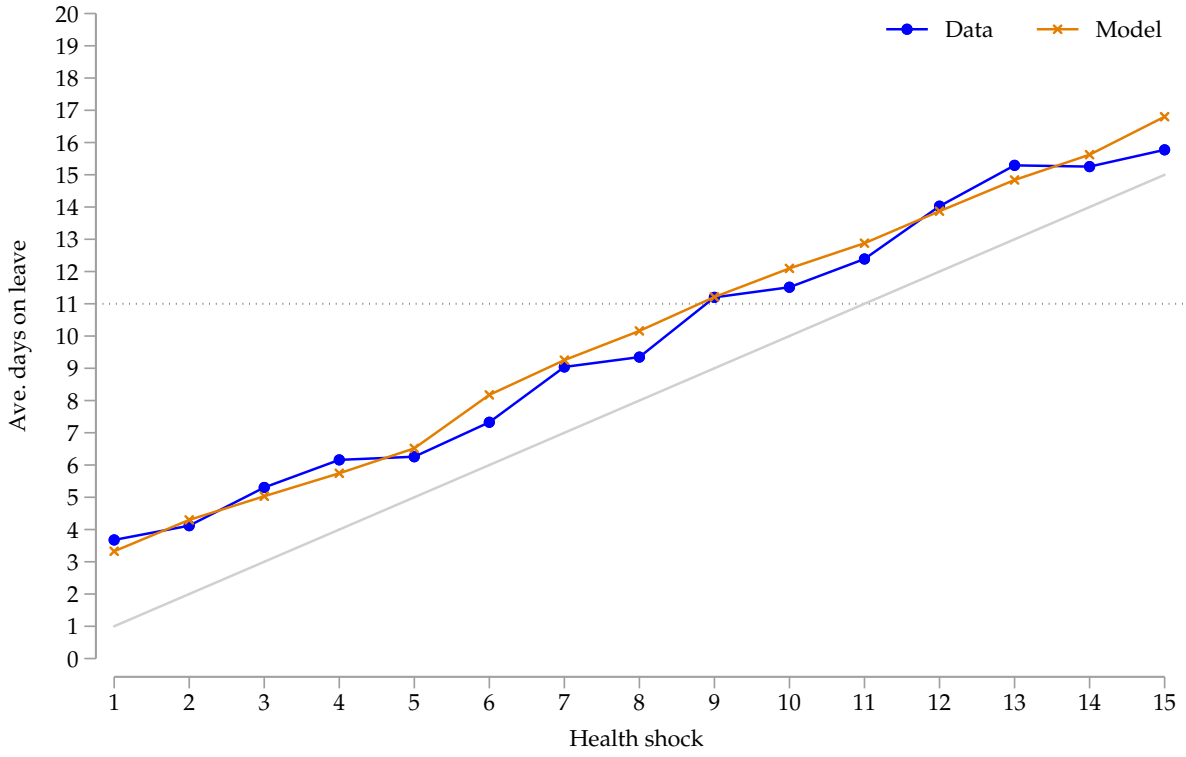
Notes: Panels (a) to (c) show the share of sick leave claims with duration s for workers whose main diagnose would implied a health state of 3 days on leave. Panel (d) aggregates the share of sick leave claims across days of the week, including only the weekend-streak combinations, e.g., from panel (a) I only consider the share for Wednesday. This figure is referenced in Section 4.2.

Figure A.36: Distribution of leisure to consumption ratio from raw data



Notes: This figure shows the distribution of the leisure to consumption ratio LC_i. This figure is referenced in Section 4.2.

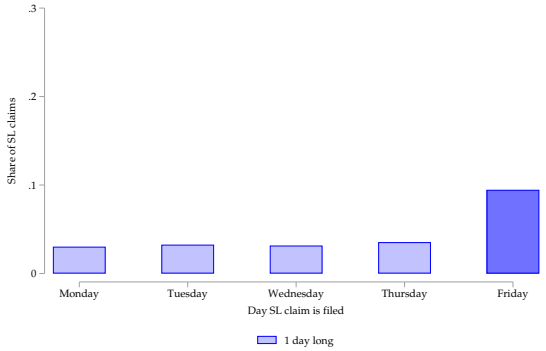
Figure A.37: Demand for days on leave as a function of health shock



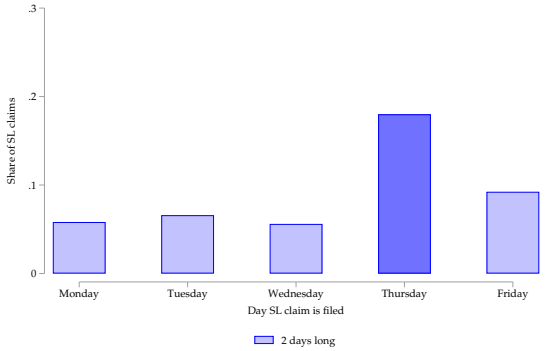
Notes: This figure shows the demand for days on leave as a function of the duration of the health shock from the data and a model-simulated sample. For each duration, I compute how many days; on average, workers request to be on leave. The 45 degrees line can be interpreted as the demand for days on leave when workers report their health. The horizontal line at 11 days indicates the position of the discontinuity in the sick pay scheme. This figure is referenced in Section 5.3.

Figure A.38: Days of the week and sick leave claim duration. Conditional to first quarter of the year (Summer quarter).

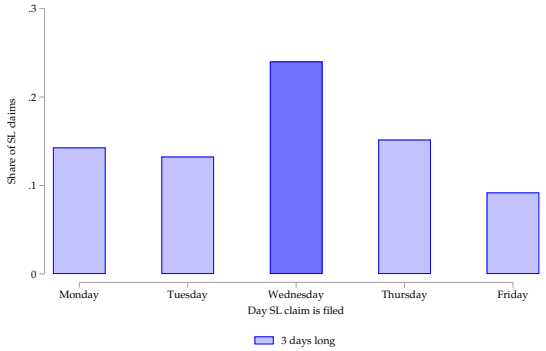
(a) One day long



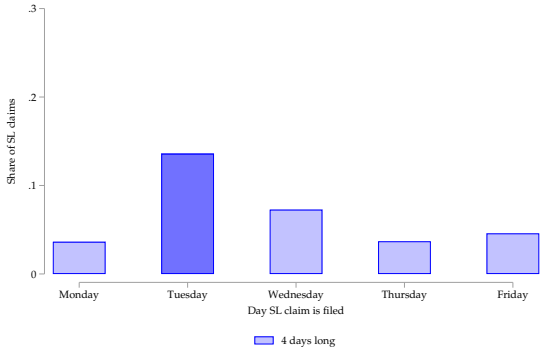
(b) Two days long



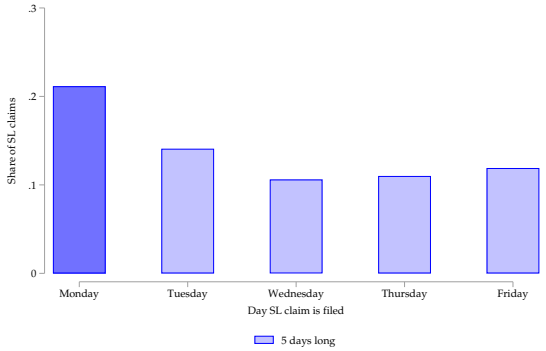
(c) Three days long



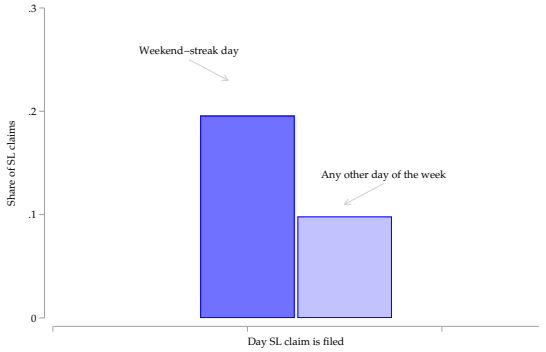
(d) Four days long



(e) Five days long



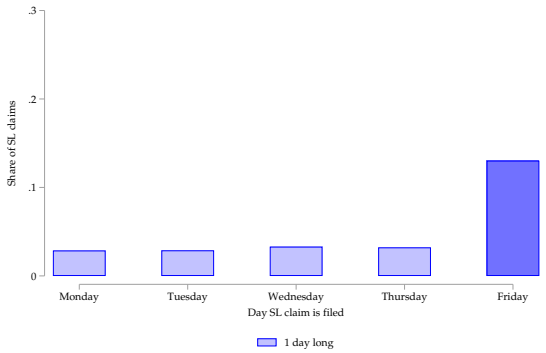
(f) Aggregate across days of the week



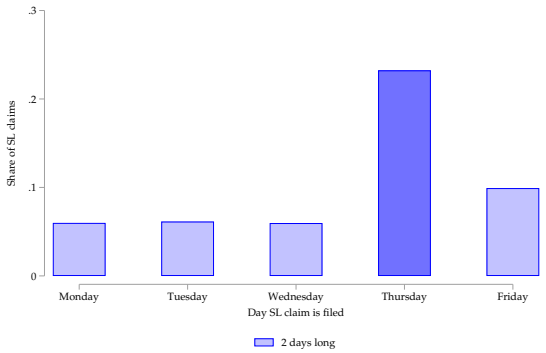
Notes: Panels (a) to (e) show the share of sick leave claims with duration s and the share of seven-days-long sick leave claims filed on each day of the week. Panel (f) aggregates across duration and days of the week: the first bar—labeled "weekend streak"—averages the share of one-to-five-days-long sick leave claims that end of a Friday and are filed any day of the week. For example, one-day-long on Friday, two-days-long on a Thursday, and so on. The second bar—labeled "non-weekend streak"—averages the share one-to-five-days-long sick leave claims filed any other day of the week. For example, two-days-long claims file on Friday. This figure is restricted to sick leave claims filed during the first quarter of the year (Summer quarter in Chile).

Figure A.39: Days of the week and sick leave claim duration. Conditional to the second quarter of the year (Fall quarter).

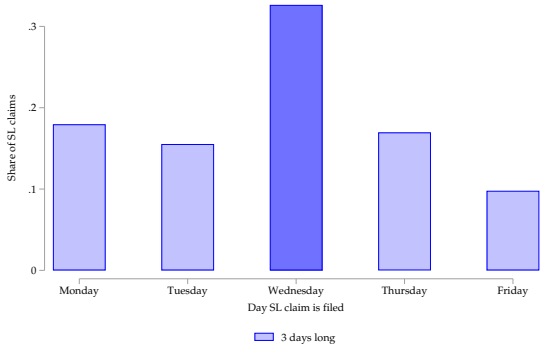
(a) One day long



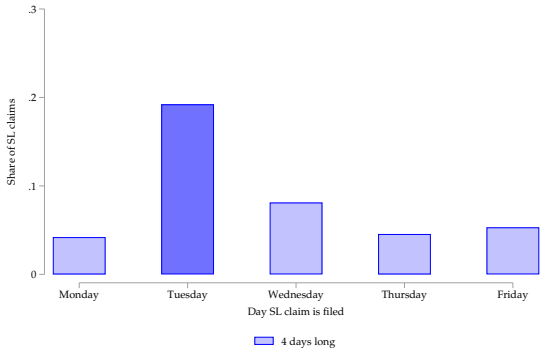
(b) Two days long



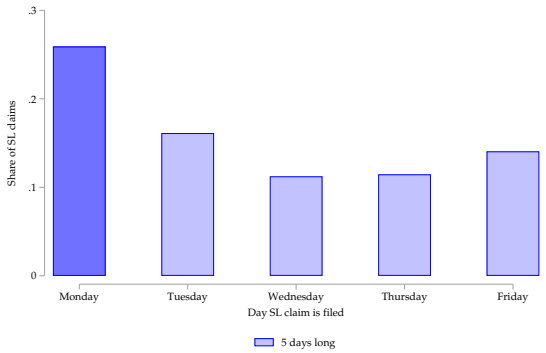
(c) Three days long



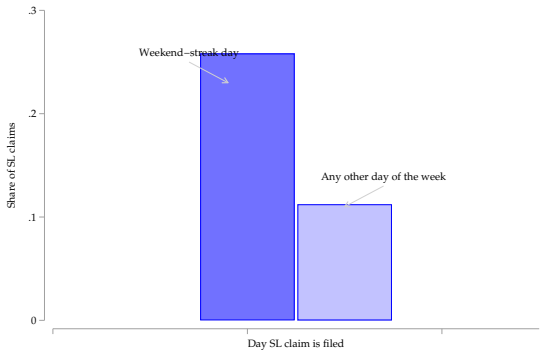
(d) Four days long



(e) Five days long



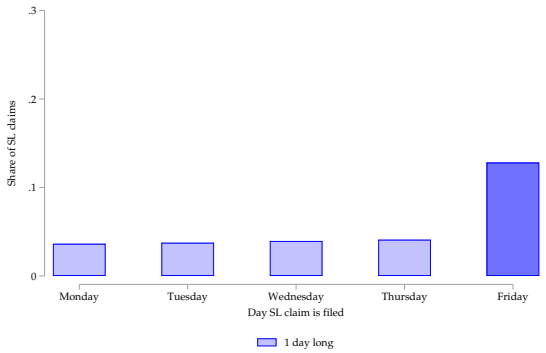
(f) Aggregate across days of the week



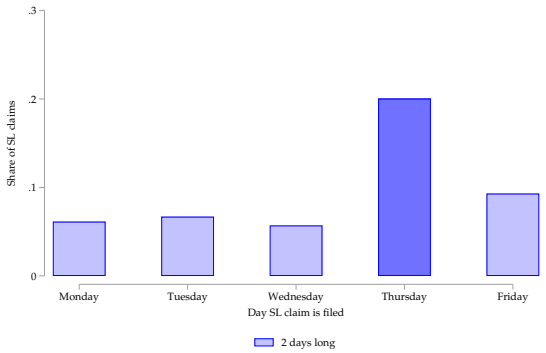
Notes: Panels (a) to (e) show the share of sick leave claims with duration s and the share of seven-days-long sick leave claims filed on each day of the week. Panel (f) aggregates across duration and days of the week: the first bar—labeled "weekend streak"—averages the share of one-to-five-days-long sick leave claims that end of a Friday and are filed any day of the week. For example, one-day-long on Friday, two-days-long on a Thursday, and so on. The second bar—labeled "non-weekend streak"—averages the share one-to-five-days-long sick leave claims filed any other day of the week. For example, two-days-long claims file on Friday. This figure is restricted to sick leave claims filed during the second quarter of the year (Fall quarter in Chile).

Figure A.40: Days of the week and sick leave claim duration. Conditional to the third quarter of the year (Winter quarter).

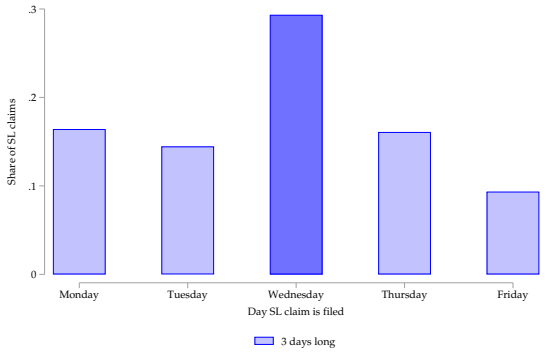
(a) One day long



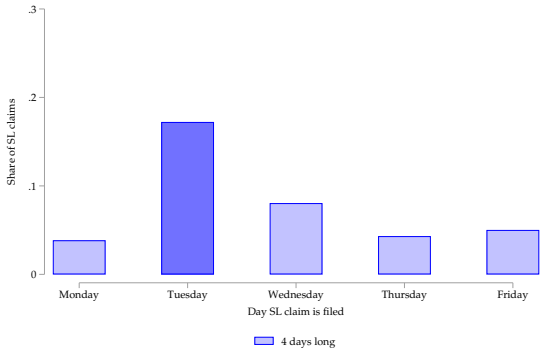
(b) Two days long



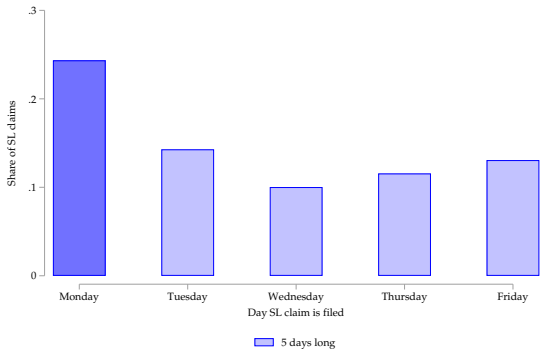
(c) Three days long



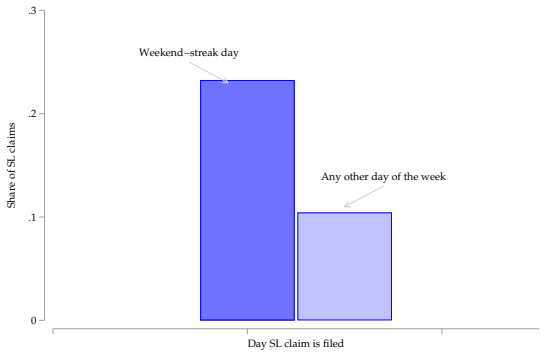
(d) Four days long



(e) Five days long



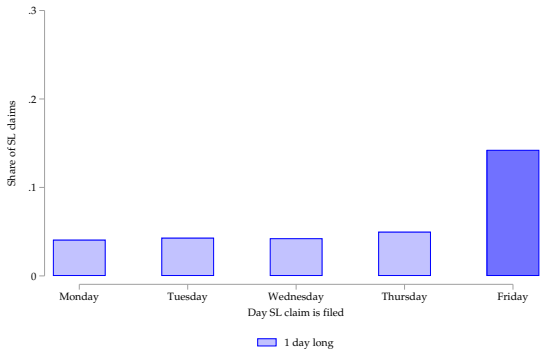
(f) Aggregate across days of the week



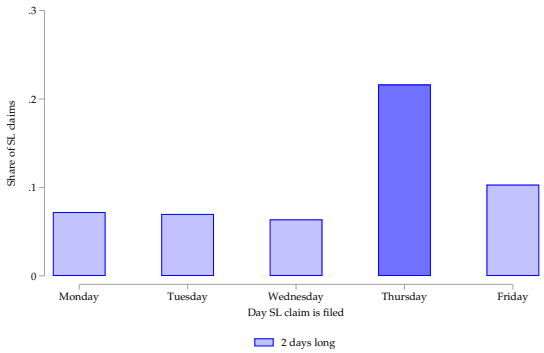
Notes: Panels (a) to (e) show the share of sick leave claims with duration s and the share of seven-days-long sick leave claims filed on each day of the week. Panel (f) aggregates across duration and days of the week: the first bar—labeled "weekend streak"—averages the share of one-to-five-days-long sick leave claims that end of a Friday and are filed any day of the week. For example, one-day-long on Friday, two-days-long on a Thursday, and so on. The second bar—labeled "non-weekend streak"—averages the share one-to-five-days-long sick leave claims filed any other day of the week. For example, two-days-long claims file on Friday. This figure is restricted to sick leave claims filed during the third quarter of the year (Winter quarter in Chile).

Figure A.41: Days of the week and sick leave claim duration. Conditional to the third quarter of the year (Spring quarter).

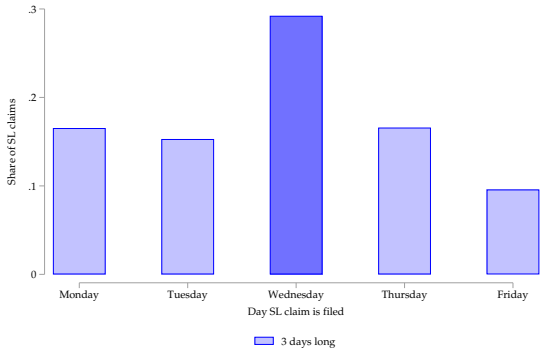
(a) One day long



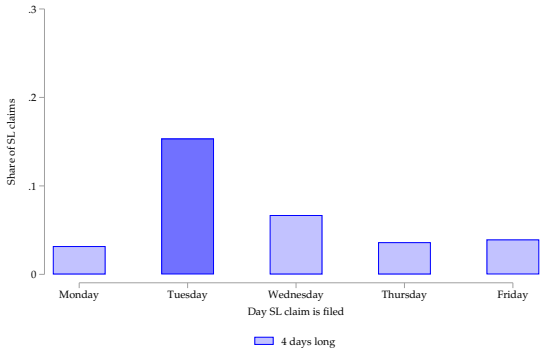
(b) Two days long



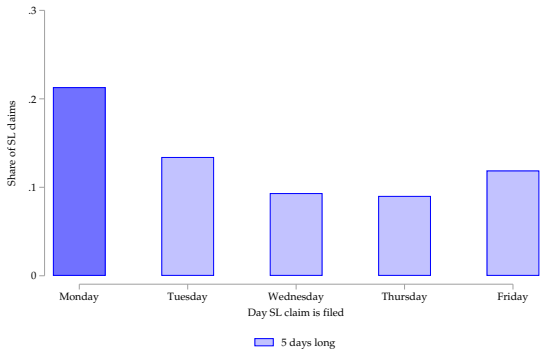
(c) Three days long



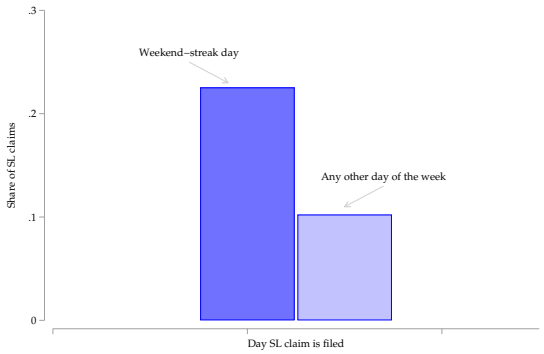
(d) Four days long



(e) Five days long



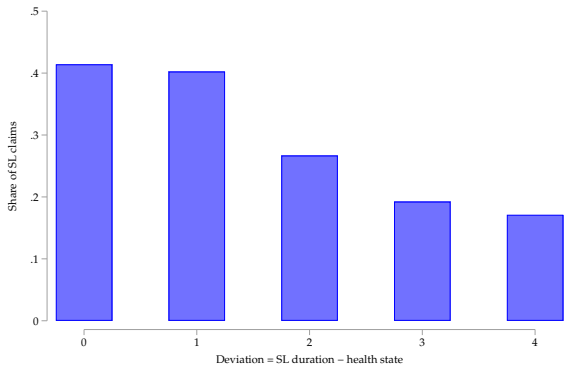
(f) Aggregate across days of the week



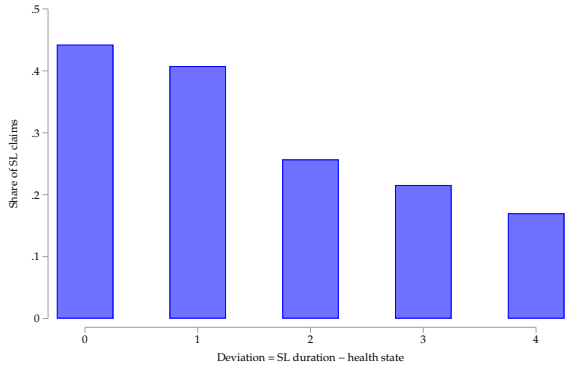
Notes: Panels (a) to (e) show the share of sick leave claims with duration s and the share of seven-days-long sick leave claims filed on each day of the week. Panel (f) aggregates across duration and days of the week: the first bar—labeled "weekend streak"—averages the share of one-to-five-days-long sick leave claims that end of a Friday and are filed any day of the week. For example, one-day-long on Friday, two-days-long on a Thursday, and so on. The second bar—labeled "non-weekend streak"—averages the share one-to-five-days-long sick leave claims filed any other day of the week. This figure is restricted to sick leave claims filed during the fourth quarter of the year (Spring quarter in Chile).

Figure A.42: Compliance cost function by quarter.

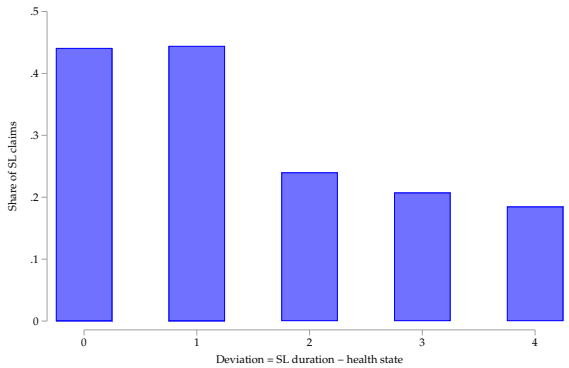
(a) Quarter 1 (Summer)



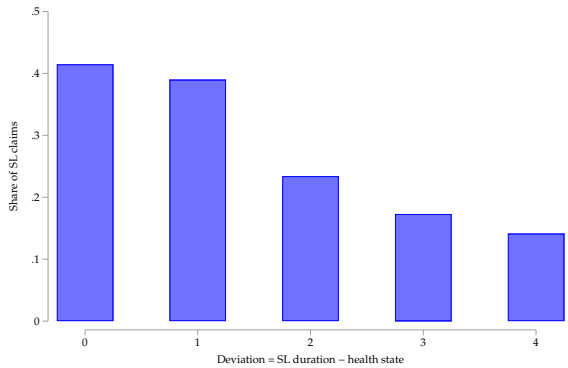
(b) Quarter 2 (Fall)



(c) Quarter 3 (Winter)



(d) Quarter 4 (Spring)



Notes: Panels (a) to (d) show the average share of sick leave claims with deviations between 0 and 4 days for each quarter. Summer quarter goes from Jan to March. The average is computed over sick leave claims with primary diagnosis requiring 1, 2 or 3 days of rest filed on a weekend streak days.

Table A.19: Summary statistics by healthcare insurance provider

	Government-run insurance (1)	Private insurance (2)	Year(s) (3)
Panel A. Enrollees Characteristics			
Share of enrollees aged			
25 - 34	25.06	31.76	2015-2019*
35 - 44	21.51	28.84	2015-2019*
45 - 54	21.11	20.12	2015-2019*
55 - 64	14.64	11.47	2015-2019*
Share female enrollees	0.44	0.35	2015-2019*
Wages (in USD monthly)			
Average	761.27	1,824.94	2015-2019*
Enrollees w/ wage above median (%)	34.44	86.69	2015-2019*
Metropolitan region (%)	38.04	60.01	2015-2019*
N of enrollees	4,503,474	1,689,240	2015-2019*
Share (%)	72.72	27.28	2015-2019*
Panel B. Sick Leave Claims			
Ratio SL claims to enrollees (%)			
2015	77.66	86.53	2015
2019	98.42	90.66	2019
Approved SL claims (%)	91.94	74.54	2015-2019
Rejected SL claims (%)	5.31	14.76	2015-2019
Ratio days on leave to SL claim	13.09	10.24	2015-2019
Annual cost per enrollee (in USD)	240.69	463.61	2015-2019
Ratio of total annual cost	24.91	58.90	2015-2019
to paid days on leave			
Annual cost			
percentage of mandatory contribution	2.6	2.1	2015-2019
as percentage of GDP	0.51	0.37	2015-2019
N of sick leave claims	3,910,482	1,473,540	2015-2019
Share (%)	72.63	27.37	2015-2019

Notes: Panel A presents summary statistics of individuals enrolled in plans offered by each healthcare insurance provider. Only individuals eligible to file a sick leave claim are included in the computations. Panel B shows characteristics of the sick leave claims handled by each insurer. Data come from the Annual Statistics of the Sick Leave System published by the Social Security Administration (SUSESO, 2020; 2019; 2018; 2017; 2016). Statistics in this table correspond to averages for 2015 - 2019, * indicates that data for 2018 are not available. The median monthly wage is computed from the 2017 CASEN survey, and using this figure I compute the share of workers with monthly salary above the median. GDP data comes from the World Bank national accounts data. SL stands for sick leave.

Table A.20: Paid sick leave systems across countries

Country	**Design**	**Benefits' computation**
	(1)	(2)
Bulgaria	bracket system	- Days 1 to 3: rep. rate = 0.7
		- Day 4 onward: rep. rate = 0.8
Chile	bracket system	- Days 1 to 3: rep. rate = 0
		- Days 4 to 10: rep. rate = 1
		- Days 11 onward: rep. rate = 1
		and the waiting period is reimbursed
Estonia	bracket system	- Days 1 to 3: rep. rate = 0
		- Day 4 and onward: rep. rate = 0.7
Finland	bracket system	- Days 1 to 9: rep. rate = 0
		- Day 10 and onward: rep. rate = 0.7
France	bracket system	- Days 1 to 3: rep. rate = 0
		- Day 4 and onward: rep. rate = 0.5
Greece	bracket system	- Days 1 to 3: rep. rate = 0.5
		- Days 4 to 30 are paid with a rep. rate of 1
Ireland	bracket system	-Days 1 to 3: 6 working days
		- Rates vary by earnings
Hungary	bracket system	- Days 1 to 15: rep rate = 0.7
		- Days 16 onward: rep. rate = 0.5
Italy	bracket system	- Days 1 to 3: rep rate = 0
		- Days 4 to 20: rep. rate = 0.5
		- Days 21 onward: rep. rate = 0.66
Portugal	bracket system	- Days 1 to 3: rep rate = 0
		- Days 4 onward: rep rate between 0.65 and 0.75
Spain	bracket system	- Days 1 to 5: rep rate = 0
		- Days 6 to 20: rep. rate = 0.60
		- Days 21 onward: rep. rate = 0.75
United Kingdom	bracket system (hybrid)	- Waiting period = 3 days , rep. rate = 0
		- 99.35 pounds per week
Denmark	linear system	Replacement rate = 0.9
Netherlands	linear system	Replacement rate = 0.7
Norway	linear system	Replacement rate = 1
Poland	linear system	Replacement rate = 0.8
Switzerland	linear system	Replacement rate = 1
Germany	linear system*	- Days 1 to 42 : rep. rate = 1
		- Week 7 onward: rep. rate = 0.7
Australia	credit account	10 sick days per year
Austria	credit account	Six weeks full paid sick leave
Belgium	credit account	30 sick days per year
United States	credit account	Average private-sector: $<$ 10 days per year

Source: several sources, author's own collection and illustration. Notes: This table summarizes sick paid systems for 22 countries. Bracket system refers vary the replacement rate based on the duration of a leave. Linear system feature a constant replacement rate. Credit account refers to the case where paid leave is earned over time and unused leave accumulates, producing an employee-specific "leave balance." The UK system is hybrid because the second bracket proposes a lump sum transfer. The German system a hybrid because it features a change in the replacement rate that kicks in after a long period. This table is referenced in Section 2.2.

Table A.21: Average recovery times - Examples from Peruvian Handbook

Workers' characteristics and diagnoses	Correction factor and recovery time
Example 1	
Lumbago with sciatica (M544)	14
43 years old	1.05
Operator/manual worker (blue collar)	1.5
Optimal time	22.05
Example 2	
Common cold (J00)	3
25 years old	0.87
Teacher (white collar)	0.75
Optimal time	2
Example 3	
Infectious gastroenteritis (A09)	2
57 years old	1.3
Office manager (white collar)	0.75
Optimal time	2

Notes: This table presents examples on how to construct the average recovery time based on workers' characteristics and sick leave diagnoses. This table is referenced in Section 2.3 and Section 4.2.

Table A.22: Sick leave claims and sick leave spells definitions

	(1)	(2)	(3)
Number of sick leave claims	1,483,103	657,125	551,647
Number of sick leave spells	1,030,613	437,418	365,127
N of SL claims in a spell (% of claims)			
One claim	55.43	51.71	51.22
Two claims	16.85	17.19	17.30
Three claims	7.30	7.93	8.00
Four claims	4.49	5.06	5.12
Five claims	3.19	3.63	3.67
Six or more claims	12.75	14.48	14.70
Among sick leave spells with more than 1 claim* (% of claims)			
Diagnoses change within spell			
Yes — 4 digits disease code	30.83	30.01	29.97
Yes — 3 digits disease code	28.67	27.86	27.82
Physician change within spell	31.21	30.87	30.83
Sample: Private sector workers			
Gender	All	Male	Male
Age	18-70	18-70	25-64

Notes: This table presents counts and summary statistics of sick leave claims and sick leave spells. A spell is a group of consecutive claims—these are considered one claim for the computation of sick leave benefits. The first row counts each sick leave claim as one observation and the second row considers the number of sick leave spells. The subsequent rows explore the composition of a spell in terms of number of claims and whether diagnoses and physicians changed withing a spell. * indicates that proportions are computed for spells composed by two to five sick leave claims. This table is referenced in Section 2.2 and in Section 2.3.

Table A.23: Conditions included in the analysis by ICD-10 group

ICD Group	Description	Included	Sick leave claims	
		(=1 if yes)	Number	Share (%)
		(1)	(2)	(3)
A00-B99	Certain infectious and parasitic diseases	1	31,244	8.56
C00-D49	Neoplasms	0	6,515	1.78
D50-D89	Blood and blood-forming organs	0	478	0.13
E00-E89	Nutritional and metabolic diseases	0	3,842	1.05
G00-G99	Nervous system	1	8,758	2.40
H00-H59	Eye and adnexa	1	6,141	1.68
H60-H95	Ear and mastoid process	1	6,246	1.71
I00-I99	Circulatory system	1	15,139	4.15
J00-J99	Respiratory system	1	64,823	17.75
K00-K95	Digestive system	1	25,854	7.08
L00-L99	Skin and subcutaneous tissue	0	8,762	2.40
M00-M99	Musculoskeletal system	1	108,908	29.83
N00-N99	Genitourinary system	1	11,605	3.18
O00-O9A	Pregnancy and childbirth	0	<50	0.01
P00-P96	Certain conditions of the perinatal	0	149	0.04
Q00-Q99	Congenital malformations	0	331	0.09
R00-R99	Abnormal clinical and laboratory findings	1	9,840	2.69
S00-S99	Injuries	1	44,922	12.30
T00-T88	Poisoning and external causes	0	4,385	1.20
U00-U85	Codes for special purposes	0	<50	0.00
V00-Y99	External causes of morbidity	0	3,578	0.98
Z00-Z99	Contact with health services	0	3,577	0.98
Total included			329,312	90.19
Total			365,127	

Notes: This table reports the health conditions included in the analysis, the number of sick leave claims filed in 2017, and what share these represent of the universe of claims. The criteria for excluding the selected health conditions is discussed in detailed in Section 2.3 This table is referenced in Section 2.3.

Table A.24: Average recovery time by workers characteristics

ICD group	Main diagnoses	25 - 34 years old		35 - 44 years old		45 - 54 years old		55 - 64 years old	
		Blue c.	White c.	Blue c.	White c.	Blue c.	White c.	Blue c.	White c.
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A00-A99	Infectious gastroenteritis	2	1	3	2	3	2	3	2
G00-G99	Migraine and headaches	3	2	4	2	4	2	5	3
G00-G99	Carpal tunnel syndrome	13	8	17	10	17	10	21	13
H00-H59	Conjunctivitis	5	3	7	4	7	4	9	5
H60-H95	Vertigo	4	2	5	3	5	3	6	4
I00-I99	Hypertension	4	3	6	3	6	3	7	4
I00-I99	Myocardial infarction	16	10	21	13	21	13	26	16
J00-J06	Common cold	3	2	4	2	4	2	5	3
J09-J18	Influenza and pneumonia	4	3	5	3	5	3	6	4
J20-J22	Bronchitis	5	3	7	4	7	4	8	5
J23-J99	Other respiratory diseases	8	5	9	6	9	6	11	7
K00-K95	Noninfective gastroenteritis	2	1	2	1	2	1	3	2
K00-K95	Inguinal hernia	6	4	9	5	9	5	11	7
M50-M54	Chronic low back pain	10	6	12	7	12	7	14	8
M50-M54	Lumbago with sciatica	10	6	12	7	12	7	14	8
M60-M79	Tendinitis	8	5	9	6	9	6	10	6
M60-M79	Shoulder lesions	8	5	9	6	9	6	10	6
Other M	Arthritis	9	5	10	6	11	6	12	7
Other M	Knee injuries	12	7	14	8	14	8	16	10
N00-N99	Renal colic	4	3	5	3	5	3	6	4
R00-R99*	Abdominal and pelvic pain	2	1	3	2	3	2	3	2
S00-S99	Injuries (e.g., sprain ankle)	14	8	16	9	16	9	18	11

Notes: This table shows the average recovery time by workers' age and occupation type for 22 disease groups. Blue c. stands for blue collar and white c. stands for white collar. Table A.25 indicates what occupations and industries are classified as blue and white collar. This table is referenced in Section 2.3 and Section 4.2.

Table A.25: Workers' occupation, industry and manual work classification

Occupation	Industry	Blue collar (=1 if yes)
		(1)
Executive, managers	Any	0
Professor, lecturer, teacher	Any	0
Other professional	Any	0
Sales representative	Any	0
Admin staff	Any	0
Factory worker	Any	1
Trabajador de casa particular	Any	1
Technician	Any	1
Unknown	Agriculture	1
Unknown	Natural Resources and mining	1
Unknown	Manufacturing	1
Unknown	Construction	1
Unknown	Utilities	1
Unknown	Retail trade	0
Unknown	Transportation, warehousing and telecommunications	1
Unknown	Service-Providing Industries	0
Unknown	Public administration	0
Unknown	Not specified	n.a.

Notes: This table reports workers' occupation, industry, and whether its combination implies the worker is considered a blue-collar (or manual) worker or not. If information is available on occupation and industry, I use worker's occupation to classified the worker as a blue-collar. If occupation is not available, I use workers' industry information. When neither occupation or industry is available, I drop observations for this worker. "n.a." stands for not applicable. This table is referenced in notes to Table A.24 and in Section 2.3.

Table A.26: Sample construction

	2017
Panel A. Sick leave claims	
Single claims - from clean dataset	2,698,993
Observations without demographic information	22,757
Workers' age not in the interval [18,70]	50,369
Worker is not Chilean	61,740
Worker not enrolled in a public insurance plan	33,560
Observations without income information	8,920
Observations	2,521,647
Panel B. Sick leave spells	
Single spells	1,825,904
Condition on private sector workers	1,030,613
Condition on male workers	437,418
Condition on ages 25-64	365,127
Condition on diagnoses included in analysis	329,312

Notes: Panel A of this table shows the counts of sick leave claims drop due to each sample selection criterion. Panel B shows the counts of sick leave spells—consecutive claims with continuous start and end dates—for each sample selection criterion. A complete list of diagnoses included in the analysis is provided in Table A.23. This table is referenced in Section 2.3.

Table A.27: Summary statistics: Private sector workers who have used SL benefits

	Any	Included conditions	
		All	Up to 30 days
	(1)	(2)	(3)
Age			
Mean	43.39	43.25	42.00
Share of workers aged (%)			
25 - 34 years old	29.21	29.51	33.07
35 - 44 years old	23.71	23.98	25.11
45 - 54 years old	24.70	24.63	23.50
55 - 64 years old	22.38	21.89	18.32
Income (monthly USD)			
Mean	857.97	862.72	870.30
Standard deviation	367.59	368.79	367.30
25th percentile	555.28	559.42	569.20
Median	782.85	788.21	797.88
75th percentile	1,089.33	1,095.32	1,103.05
90th percentile	1,408.30	1,414.46	1,418.81
Region (%)			
Central	46.78	47.32	48.35
Mining intensive	6.77	6.67	6.01
Health - chronic conditions (%)			
Hypertension	16.89	16.71	14.34
Diabetes	8.24	7.79	6.22
Share of workers (%)	100.00	92.59	72.16
Observations	246,017	227,797	177,531

Notes: This table presents summary statistics for all male workers who had used sick leave benefits in the past year based on the conditions and duration of sick leave claims. The sample includes private sector employees age 25 to 64 years old. Income statistics are based on the winsorized distribution where the lowest and highest 5% of the income values are excluded. Sick leave claims of up to 30 days account for 95% of all claims filed in a year. This table is referenced in Section 2.3.

Table A.28: Probability of filing a SLC for each disease group by workers' characteristics

ICD group	Main diagnoses	25 - 34 years old		35 - 44 years old		45 - 54 years old		55 - 64 years old	
		Blue c.	White c.	Blue c.	White c.	Blue c.	White c.	Blue c.	White c.
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A00-A99	Infectious gastroenteritis	12.36	18.26	8.91	13.06	6.51	9.36	4.80	6.57
G00-G99	Migraine and headaches	1.64	2.28	1.57	1.86	1.19	1.64	0.89	1.14
G00-G99	Carpal tunnel syndrome	1.13	1.08	1.31	1.30	1.48	1.31	1.63	1.33
H00-H59	Conjunctivitis	1.57	1.46	1.92	2.40	2.40	2.40	3.03	3.19
H60-H95	Vertigo	1.08	1.24	1.32	1.23	1.85	1.68	2.49	2.64
I00-I99	Hypertension	1.09	0.80	1.88	1.70	2.95	2.85	4.16	3.81
I00-I99	Myocardial infarction	0.12	0.08	0.31	0.34	0.77	0.71	1.55	1.38
J00-J06	Common cold	11.89	17.12	10.64	14.87	8.62	12.79	6.34	10.17
J09-J18	Influenza and pneumonia	3.31	3.92	3.78	4.66	4.68	4.88	5.53	5.58
J20-J22	Bronchitis	5.44	6.44	5.91	7.33	7.18	8.35	8.06	10.80
J23-J99	Other respiratory diseases	0.96	1.00	1.11	1.13	1.30	1.37	1.74	1.71
K00-K95	Noninfective gastroenteritis	5.34	5.99	4.19	4.99	3.95	4.57	3.34	3.93
K00-K95	Inguinal hernia	1.93	1.62	3.02	2.87	3.82	4.02	4.48	4.73
M50-M54	Chronic low back pain	16.85	12.35	16.73	12.78	14.17	11.43	12.02	10.02
M50-M54	Lumbago with sciatica	6.37	4.86	8.11	6.33	7.54	6.60	6.70	5.87
M60-M79	Tendinitis	5.17	3.56	5.99	4.37	6.55	4.98	6.67	4.95
M60-M79	Shoulder lesions	2.47	1.79	3.78	2.49	4.62	3.61	4.43	3.11
Other M	Arthritis	1.91	1.34	2.25	1.79	3.11	2.63	4.38	3.63
Other M	Knee injuries	0.57	0.41	0.70	0.66	1.16	0.99	1.61	1.33
N00-N99	Renal colic	2.42	2.12	3.06	3.23	3.63	3.47	4.29	4.48
R00-R99*	Abdominal and pelvic pain	2.19	2.38	1.99	2.37	2.38	2.72	3.01	2.83
S00-S99	Injuries (e.g., sprain ankle)	14.19	9.89	11.53	8.25	10.13	7.64	8.87	6.82

Notes: This table shows the probability of filling a sick leave claim for each disease (d) by workers' group (b). Each of these probabilities is computed as the ratio of sick leave claims with diagnosis d and all claims from group b, thus columns add up to 100. Main diagnoses indicates the most common condition for a disease group. Blue c. and white c. stand for blue-collar and white-collar occupations respectively. These probabilities are plotted in Figure A.33. This table is referenced in Section 2.4.

Table A.29: Number of business days on leave (s_l)

	Number of days on leave (s_c)							
Day of the week (*dow*)	1	2	3	4	5	6	7	8
Monday	1	2	3	4	5	5	5	6
Tuesday	1	2	3	4	4	4	5	6
Wednesday	1	2	3	3	3	4	5	6
Thursday	1	2	2	2	3	4	5	6
Friday	1	1	1	2	3	4	5	6

Notes: This table shows the number of business days on leave (s_l) as a function of (total) days on leave (s_c) and day of the week (*dow*) a sick leave claim is filed. This table is referenced in Section 3.1.

Table A.30: Identification of weekend-streak utility parameter (q): estimates from raw data

Duration	Day of the week		
	Weekend streak	Non-weekend streak	Difference
	(1)	(2)	(3)
1 day long	0.1219	0.0355	0.0864
2 days long	0.2062	0.0672	0.1391
3 days long	0.2872	0.1482	0.1390
4 days long	0.1640	0.0489	0.1151
5 days long	0.2330	0.1216	0.1114
Simple average	0.2025	0.0843	0.1182
Weighted average	0.2274	0.1041	0.1233

Notes: This table presents the distribution of sick leave claims by duration and day of the week. Weekend streak refers to the day of the week a sick leave claim should start to finish on Friday. For example, when duration is one day, weekend streak refers to Friday, when duration is two days, it refers to Thursday. The non-weekend streak category groups all the other days of the week. The share of sick leave claims of duration s filed on day dow is computed as the ratio between the number of claims with duration s filed on dow and the number of claims of filed on dow with duration between one and fifteen days. Figure 1.5 presents this table graphically. This table is referenced in Section 4.2.

2.3 Distribution of health states

The Peruvian Handbook of Recovery Times specifies an average recovery time for 2,763 unique disease codes at the fourth digit level of the 10th revision of the ICD. This paper focuses on non-mental health conditions; excluding these diagnoses reduces the number of unique diseases to 2,690.[7] Estimating the model with such level of granularity is unfeasible: it would require estimating a probability for each disease and group of observable characteristics (2,690 diagnoses $\times$ 4 age group $\times$ 2 occupation groups = 21,520). For this reason, I group diagnoses in more aggregated categories.

I these categories considering (i) the type of diseases they represent and (ii) how frequently these diagnoses are used in the claims data. Table A.23 lists the conditions included in the analysis by their ICD 10th revision group and the share of sick leave claim data with these diagnoses. To compute these shares, I used the sample constructed for the quantitative analysis of this paper. Table A.26 provides details on sample construction.

The first criteria for excluding a group of conditions are those not listed in the Peruvian handbook. These are conditions originating in the perinatal period (codes in groups P00-P96) and congenital malformations, deformations, and chromosomal abnormalities (codes in groups Q00-Q99). In fact, 0.15% of the sick leave claim data is reported under these diagnoses. I dropped such observations.

The second criteria for excluding a group of conditions is the nature of the diagnosis which makes very challenging to assign a benchmark recovery time. I exclude poisonings and burns (codes in group T00-T98)—these diagnoses accumulate 1.22% of the sick leave claim data. Additionally, I exclude diseases coded under "special purposes codes" (codes U00-U85), external causes of morbidity (codes V00-Y99), and factor influencing health status and contact with health services (codes Z00-Z99). All these together represent 3.16% of the sick

[7] I use codes F01-F99 to define mental health conditions, these are grouped under the chapter "Mental, Behavioral and Neurodevelopmental disorders".

leave claims. These conditions associated with longer recovery times or impairments where full recovery might not be foreseeable, for example, leg amputations and organs transplants. Finally, I exclude conditions with diagnoses C00-D49; these codes are used for neoplasms, which in most cases, are chronic conditions or diseases that would require a longer recovery time. In terms of claims data, these represent 1.78% of the claims. The final sample includes about 86% of all sick leave claims filed by private-sector male workers.

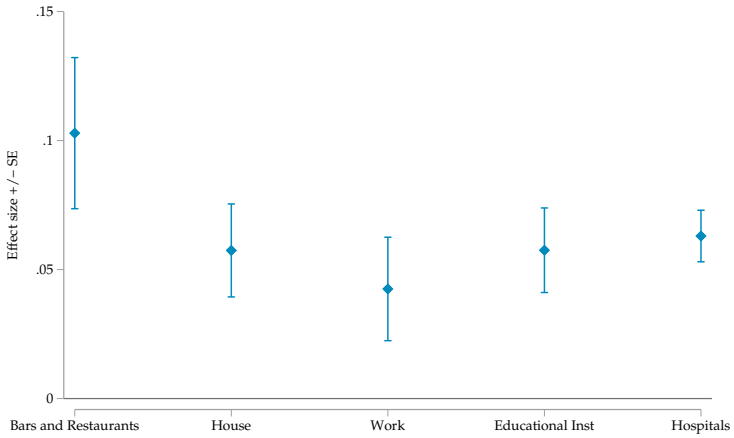
3. Appendix to "Tobacco Consumption Habits in Argentina: Causal Evidence from a New Regulation"

3.1 Additional Figures

Figure A.43: Examples of law-regulated packaging.

Notes: This figure provides an example of the Tobacco Graphic Warnings printed on cigarette packages. Big font messages: (i) smoking reduces years of life, (ii) smoking causes cancer, (iii) smoking might cause leg amputation, (iv) pregnant women who smoke harm her child (v) smoking causes death by suffocation. Small font messages: (i) tobacco drives half of smoker's deaths, (ii) every cigarette poisons you, (iii) smoking causes gangrene, (iv) every cigarette damages your respiratory capacity.

Figure A.44: Exposure to Environmental Tobacco Smoke



Notes: This figure presents estimates of the change in the share of non-smokers exposed to environmental tobacco smoke after the implementation of the national law. Each coefficient corresponds to a separate regression of the share of individuals reporting to not have noticed someone smoking inside a given venue. All coefficients are expressed as effect sizes to ease the comparison of the effects across venues.

Table A.31: Evolution of Cigarette Sales and Consumption

	Sales data		Survey data	
Year	Consumption	Percent change	Consumption	Percent change
2008	181.06		106.75	
2009$^{(a)}$	177.64	-1.89%	103.42	-3.11%
2010	174.86			
2011$^{(b)}$	182.65	1.40%	105.84	1.16%
2012	178.36			
2013$^{(b)}$	174.07	-2.38%	99.39	-3.09%

Notes: Consumption measures are expressed in millions of 20 cigarettes packages a month. Sales data is collected by the Ministry of Agriculture, time series are expressed in 20 cigarettes packages in a year. Consumption from survey data is constructed as the average consumption per day multiplied by 30. Summary statistics indicates that 62% of smokers smoke every day, this is the modal frequency. Percent changes are: (a) relative to the previous calendar year; (b) relative to the $n-2$ year.

Table A.32: Determinants of Strong Regulation prior to 2011

Dependent variable:	Indicator for Enacting a Strong Regulation	
	Coefficient	Marginal Effect
Peronist party (=1)	0.390	0.133
	(0.266)	(0.090)
State is tobacco producer (=1)	0.598	0.204
	(0.372)	(0.128)
Ln employment	-0.172	-0.058
	(0.516)	(0.177)
Unemployment rate	4.769	1.626
	(5.197)	(1.783)
Proportion of smokers	-0.327	-0.112
	(3.060)	(1.044)
Prevalence of COPD	1.156***	0.394***
	(0.378)	(0.130)
Share population 0 - 14 years old	0.066	0.022
	(16.370)	(5.584)
Share population 15 - 24 years old	48.93***	16.687***
	(17.580)	(5.832)
Share population 25 - 44 years old	20.400	6.958
	(16.510)	(5.682)
Share population 45 - 64 years old	44.210	15.077
	(30.110)	(10.360)
Ln population	0.542	0.185
	(0.529)	(0.183)
Observations	144	144
Pseudo R^2	0.2196	

Notes: This table presents the results of a probit model estimating the determinants of enacting a strong regulation before 2011. The second column presents estimated coefficients from the probit model and the last column presents the corresponding marginal effects computed at the means. Standard errors are robust to heteroskedasticity. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table A.33: Effects on Extensive Margin Outcomes

Probability of	Never smokers	Current smokers
	(1)	(2)
2008	-0.0054	-0.0056
	(0.0147)	(0.0104)
2009	-0.0155	-0.0076
	(0.0162)	(0.0087)
2013	0.0434**	-0.0617***
	(0.0175)	(0.0103)
Mean dep. var. in 2011	0.4608	0.2825
Observations	153,093	153,093
R-squared	0.0329	0.0259
Correctly predicted	0.5536	0.6881
Individual controls	Yes	Yes
State × time controls	Yes	Yes
State FE	Yes	Yes
Time FE	Yes	Yes

Notes: The omitted category corresponds to 2011, the year the federal law was passed. Treatment is a dichotomous variable, $Treat_s$, that equals one if the legislation index for state s in moment t is less than or equal to 3 before 2011. Individual-level controls include age, gender, educational attainment, employment status and income category of the household. State × time controls include total private employment and total population. Standard errors are block-bootstrapped at the state-level with 200 replications.

Table A.34: Probability of being a current smoker by age group

Age:	18-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60	60-65
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
2008	0.0009	-0.0252	-0.0363	-0.0212	0.0125	-0.0253	0.0074	0.0177	0.0365
	(0.0293)	(0.0267)	(0.0235)	(0.0319)	(0.0303)	(0.0267)	(0.0257)	(0.0179)	(0.0201)
2009	0.0038	0.0089	-0.0209	-0.0363	-0.0216	0.0163	-0.0227	-0.0038	-0.0088
	(0.0157)	(0.0225)	(0.0212)	(0.0224)	(0.0166)	(0.0249)	(0.0302)	(0.0201)	(0.0158)
2013	-0.1014***	-0.0379	-0.0148	-0.0434	-0.0983***	-0.1207***	-0.1191***	0.0152	-0.0156
	(0.0307)	(0.0251)	(0.0200)	(0.0217)	(0.0163)	(0.0250)	(0.0275)	(0.212)	(0.0187)
Mean dep. var. in 2011	0.27	0.31	0.29	0.29	0.26	0.30	0.32	0.27	0.23
Observations	24,822	19,409	20,329	18,804	16,121	14,205	13,576	11,992	13,835
R-squared	0.0497	0.0278	0.0377	0.0341	0.0348	0.0247	0.0253	0.0232	0.0240
Correctly predicted	0.6899	0.6604	0.6564	0.7059	0.5085	0.6758	0.6591	0.7331	0.8005
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State × time controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The omitted category corresponds to 2011, the year the federal law was passed. Treatment is defined as a dichotomous variable, $Treat_s$, that equals one if the legislation index for state s in moment t is less than or equal to 3 before 2011. Individual-level controls include gender, educational attainment, employment status and income category of the household. State × time controls include total private employment and total population. Standard errors are block-bootstrapped at the state-level with 200 replications. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table A.35: Probability of being a current smoker by age group and gender

Age:	18-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60	60-65
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
2008									
Female	-0.0448	-0.0507	-0.0252	-0.0123	-0.0247	-0.0055	-0.0322	-0.0174	0.0612
	(0.0338)	(0.0296)	(0.0212)	(0.0455)	(0.0375)	(0.0341)	(0.0433)	(0.0359)	(0.0356)
Male	0.0485	0.0146	-0.0442	-0.0204	0.0560	-0.0446	0.0423	0.0510	0.0035
	(0.0359)	(0.0316)	(0.0406)	(0.0341)	(0.04145)	(0.0499)	(0.0671)	(0.0342)	(0.0379)
2009									
Female	0.0142	-0.0257	-0.0111	-0.0047	-0.0068	-0.0330	-0.0276	-0.0362	0.0153
	(0.0257)	(0.03212)	(0.0367)	(0.225)	(0.0218)	(0.0359)	(0.0414)	(0.0294)	(0.0266)
Male	-0.0056	0.0422	-0.0312	-0.0669	-0.0338	0.0680	-0.0238	0.0220	-0.0391
	(0.0247)	(0.0329)	(0.0315)	(0.0363)	(0.0344)	(0.0306)	(0.0327)	(0.0274)	(0.0272)
2013									
Female	-0.1150	-0.0253	-0.0036	-0.0549	-0.1125	-0.1115	-0.1044	0.0413	0.0096
	(0.0252)***	(0.0381)	(0.0386)	(0.0266)	(0.0259)***	(0.0308)***	(0.0401)***	(0.0343)	(0.0283)
Male	-0.0891	-0.0599	-0.0271	-0.0261	-0.0859	-0.1292	-0.1318	-0.0215	-0.0508
	(0.0510)***	(0.0384)	(0.0342)	(0.0282)	(0.0324)***	(0.0473)***	(0.0304)***	(0.0437)	(0.0348)
$H_0 : \delta_{FEMALE} = \delta_{MALE}$	.432	.1235	.6888	.9247	.5955	.9301	.9443	.8385	.3941
Mean dep. var. in 2011									
Female	0.27	0.31	0.29	0.29	0.26	0.30	0.32	0.27	0.23
Male	0.33	0.36	0.36	0.35	0.32	0.35	0.37	0.30	0.25
Observations									
Female	13,107	10,700	11,360	10,451	8,859	7,666	7,464	6,697	8,193
Male	11,715	8,709	8,969	8,353	7,262	6,539	6,112	5,295	5,642
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State × time controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The omitted category corresponds to 2011, the year the federal law was passed. Treatment is defined as a dichotomous variable, $Treat_s$, that equals one if the legislation index for state s in moment t is less than or equal to 3 before 2011. Individual-level controls include educational attainment, employment status and income category of the household. State × time controls include total private employment and total population. Standard errors are block-bootstrapped at the state-level with 200 replications. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table A.36: Probability of being a current smoker by educational attainment

Educational level:	Elementary d/o	Elementary	HS d/o	HS	College d/o	College +
	(1)	(2)	(3)	(4)	(5)	(6)
2008	-0.0312	-0.0310	0.0030	0.0336	0.0036	-0.0093
	(0.4700)	(0.0175)	(0.0272)	(0.0212)	(0.0260)	(0.0238)
2009	-0.0630***	0.0093	-0.0177	-0.0141	0.0049	0.0167
	(0.0210)	(0.0170)	(0.0139)	(0.0179)	(0.0220)	(0.0223)
2013	-0.0080	-0.0437*	-0.0599***	-0.0448***	-0.0989***	-0.0986***
	(0.0361)	(0.0185)	(0.0216)	(0.0303)	(0.0296)	(0.0193)
Mean dep. var. in 2011	0.31	0.30	0.35	0.28	0.24	0.228
Observations	12,261	30,388	27,877	35,538	21,195	23,811
R-squared	0.0713	0.0434	0.0258	0.0217	0.0199	0.0134
Correctly predicted	0.5073	0.6906	0.6370	0.6173	0.7181	0.7607
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
State × time controls	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The omitted category corresponds to 2011, the year the federal law was passed. Treatment is defined as a dichotomous variable, $Treat_s$, that equals one if the legislation index for state s in moment t is less than or equal to 3 before 2011. Individual-level controls include age, gender, educational attainment, employment status and income category of the household. State × time controls include total private employment and total population. HS stands for High School and d/o stands for drop-out. Standard errors are block-bootstrapped at the state-level with 200 replications. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table A.37: Probability of being a current smoker by income

Income category	1st quintile	2nd quintile	3rd quintile	4th quintile	5th quintile
	(1)	(2)	(3)	(4)	(5)
2008	-0.0239	-0.0288	-0.0080	0.0282	-0.0028
	(0.0246)	(0.0241)	(0.0201)	(0.0195)	(0.0318)
2009	-0.0433	-0.0513**	0.0297	0.0093	-0.0213
	(0.0211)	(0.0216)	(0.0200)	(0.0146)	(0.0404)
2013	-0.0561***	-0.0313*	-0.0534***	-0.0919***	-0.0892***
	(0.0135)	(0.0187)	(0.0137)	(0.0183)	(0.0207)
Mean dep. var. in 2011	0.29	0.29	0.31	0.28	0.25
Observations	23,427	33,418	32,291	37,006	16,507
R-squared	0.0530	0.0291	0.0248	0.0231	0.0217
Correctly predicted	0.6846	0.7049	0.7032	0.7057	0.7023
Individual controls	Yes	Yes	Yes	Yes	Yes
State × time controls	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes

Notes: The 1st quintile corresponds to the lowest 20 percent households in the income distribution. Analogously, the 5th quitile corresponds to the highest 20 percent households in the income distribution. The omitted category corresponds to 2011, the year the federal law was passed. Treatment is defined as a dichotomous variable, $Treat_s$, that equals one if the legislation index for state s in moment t is less than or equal to 3 before 2011. Individual-level controls include age, gender, educational attainment, and employment status. State × time controls include total private employment and total population. Standard errors are block-bootstrapped at the state-level with 200 replications. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table A.38: Probability of risky alcohol consumption - Estimates from the Probit model

Dependent variable	Alcohol consumption	Beer abuse	Wine abuse	Spirits	Binge drinking
	(1)	(2)	(3)	(4)	(5)
2008	-0.0845	-0.0085	0.0132	-0.2020	0.0534
	(0.0591)	(0.1297)	(0.1108)	(0.1645)	(0.1109)
2009	0.0091	0.0342	-0.0807	-0.1013	-0.0277
	(0.0741)	(0.0985)	(0.0901)	(0.1329)	(0.0816)
2013	0.0288	-0.2256**	-0.1527**	0.6968***	-0.1526**
	(0.0355)	(0.1079)	(0.0556)	(0.1553)	(0.0730)
Marginal Effects (at means)					
2008	-0.0268	-0.0011	0.0021	-0.0171	0.0179
	(0.0187)	(0.0162)	(0.0174)	(0.0139)	(0.0373)
2009	0.0029	0.0043	-0.0127	-0.0086	-0.0093
	(0.0235)	(0.0123)	(0.0142)	(0.0111)	(0.0274)
2013	0.0091	-0.0282**	-0.0241**	0.0589***	-0.0512**
	(0.0113)	(0.0133)	(0.0089)	(0.0117)	(0.0246)
Mean dep. var. in 2011	0.7232	0.1095	0.1291	0.0318	0.3605
Observations	29,391	21,561	21,561	21,561	21,561
Pseudo R-squared	0.0534	0.1768	0.1035	0.1613	0.1054
Correctly predicted	0.9816	0.8605	0.8663	0.9258	0.6909
Individual controls	Yes	Yes	Yes	Yes	Yes
State × time controls	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes

Notes: The omitted category corresponds to 2011, the year the federal law was passed. Treatment is defined as a dichotomous variable, $Treat_s$, that equals one if the legislation index for state s in moment t is less than or equal to 3 before 2011. The sample is restricted to current smokers who have reported drinking alcohol in the last month. Individual controls include age, gender, educational attainment, employment status and income category of the household. State × time controls include total private employment and total population. Standard errors in parentheses are clustered at the state-level.

Table A.39: Demographic Characteristics of Smokers by Alcohol Beverages

	Abusive consumption			Binge drinking
	Beer	Wine	Spirits	
Female	0.14	0.13	0.30	0.19
	(0.01)	(0.01)	(0.04)	(0.01)
Young (<25 years old)	0.45	0.29	0.61	0.35
	(0.02)	(0.02)	(0.04)	(0.01)
Less than high school	0.53	0.61	0.44	0.53
	(0.02)	(0.02)	(0.04)	(0.01)
Single	0.54	0.36	0.74	0.48
	(0.02)	(0.02)	(0.03)	(0.01)
High income hh	0.06	0.03	0.09	0.07
	(0.01)	(0.01)	(0.02)	(0.01)
Population share	0.12	0.13	0.03	0.38

*Notes:*This table presents demographic characteristics of smokers who consumed alcohol in the last month. Abusive consumption thresholds are specific to the alcohol beverage, see the main text for details. Binge drinking is defined as consuming 5 or more drinks during a single occasion in the last 30 days, either during the weekend or during a week day.

Table A.40: Health Outcomes

Diagnosis	COPD	Lung cancer
	(1)	(2)
2008	-0.00085	0.00317
	(0.01393)	(0.00658)
2009	-0.00492	-0.00031
	(0.01240)	(0.00518)
2010	-0.01130	0.00309
	(0.01043)	(0.00641)
2012	-0.01670	-0.01140
	(0.01955)	(0.01019)
2013	-0.031**	-0.0166
	(0.01810)	(0.01387)
2014	-0.0361*	-0.0255
	(0.02235)	(0.01704)
Mean dep. var.	0.1061	0.0644
Observations	1,512	1,512
R-squared	0.447	0.539
Time varying controls	Yes	Yes
State FE	Yes	Yes
Linear trend	Yes	Yes

Notes: This table presents estimated effects on health outcomes. The omitted category corresponds to 2011, the year the national law was passed. Standard errors are block-bootstrapped at the state-level with 200 replications. *** $p<0.01$, ** $p<0.05$, * $p<0.10$

Table A.41: Mechanisms

Policy	National regulation (δ)		Place-based bans (δ^{bans})		Graphic Warnings (δ^{tgw})	
Probability of	Never smokers	Current smokers	Never smokers	Current smokers	Never smokers	Current smokers
	(1)	(2)	(3)	(4)	(5)	(6)
2008	-0.0054	-0.0056	-0.00162	-0.0135	-0.00461	0.0012
	(0.0147)	(0.0104)	(0.0149)	(0.0115)	(0.0103)	(0.0083)
2009	-0.0155	-0.0076	0.00879	-0.02318*	-0.01988*	0.0040
	(0.0162)	(0.0087)	(0.0168)	(0.0087)	(0.0164)	(0.0092)
2013	0.0434**	-0.0617***	0.00546	-0.04635***	0.04067*	-0.038525***
	(0.0175)	(0.0103)	(0.0020)	(0.0080)	(0.0207)	(0.0102)
Observations	153,093	153,093	153,093	153,093	153,093	153,093
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
State x time controls	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table presents point estimates of the effects of the federal law, the effects of place-based bans, and the incorporation of graphic tobacco warnings following the decomposition proposed in equation (3.4). The first column reproduces results presented in Figure 3.2. Standard errors are block-bootstrapped at the state level with 200 replications. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

3.3 Construction of Legislation Index

Argentina has 23 states and one federal district.[8] Each state has the autonomy to dictate laws in their territory as long as these laws are not contrary to the federal laws. Broadly speaking, taxes on tobacco-related products are implemented at the national level but non-price policies have been implemented at the regional level before the 26,687 law was passed.

I coded a total of 47 regional laws and two national laws. The three main categories that composed the index are tobacco advertising, sales, and consumption. Each category is further divided into sub-categories, to allow for a better understanding of the regulation and to be able to assess the importance of each particular aspect of the law. The subcategories were defined following González-Rozada (2006) report on the status of tobacco legislation in Argentina.

The advertising category includes the sub-categories: publicity, advertising in radio, in television, to a certain audience, regulation on the content of advertising, events sponsoring, brand stretching, inclusion and size of tobacco graphic warnings, and inclusion of contact information about anti-smoking public services. Sales category—defined as bans on sales—includes the next sub-categories: sales to under 18 years old individuals, elementary school, high school, education institutions in general, hospitals or health institutions, government buildings, public transportation means, and sales by the unit. Finally, the consumption category includes the sub-categories: government buildings, workplaces, health institutions, elementary schools, high schools, universities, public transportation means, restaurants, bars, entertainment centers.

To construct the legislation index I summarize the previous categories with a dummy variable. The advertising category is summarized by a dummy that takes the value of 1 if the state has passed some regulation regarding advertising. The sales category is restricted

[8] I refer to this administrative unit as another state since the distinction between state and federal district is not relevant for this paper.

to bans in sales to underage individuals. Consumption sub-categories are grouped regarding similarities of the environments: public means of transportation, educational institutions, health institutions, and restaurants, bars, and other entertainment places together. Thus, the maximum value the index can take before the implementation of the national law is 6. I define strict states as those with an index strictly greater than 3.

3.4 Robustness Checks

Alternative Definition of Comparison and Treated States

My identification strategy relies on state-level differences in the strength of regulations on tobacco products before the national law was passed. I define the treatment as a dichotomous variable, $Treat_s$, that equals 1 if the legislation index for state s is less than or equal to 3 before 2011. In this section, I present the main results of my analysis using an alternative definition. I define the treatment as a dichotomous variable that equals 1 if the legislation index for state s is strictly less than 3 before 2011.

Table A.42 presents balancing test results. As in the main analysis, estimated differences indicate that individuals are statistically indistinguishable in terms of gender, age, marital status, employment status, and educational attainment. I find evidence that on average individuals in comparison states are poorer than individuals in treated states.

I reproduce results for the extensive margin in Table A.43. The estimates of the coefficients for 2008 and 2009 are not statistically distinguishable from zero, meaning that the comparison and treated states were on similar trends regarding cigarette consumption before the national law was passed. I find that results on the probability of being a never smoker and the probability of being a current smoker are smaller in absolute value and still significantly different from zero under the more strict definition of treatment.

I present results for the intensive margin in Table A.44. The estimated coefficients are

the percentage change in the probability of observing average daily consumption in bin b. I find evidence that the distribution of daily cigarettes smoked shifts to the right, i.e. it is more likely that smokers consume more than 15 cigarettes per day and less likely that they consume fewer than 15 cigarettes per day. The proportion of smokers who consume between zero and five cigarettes per day decreases by 23 percentage points. This estimate is slightly greater than the one found with the original treatment definition.

Table A.42: Pre Policy Balance Individual Characteristics

	Control	Treatment	Difference
Average age	37.66	37.71	-0.05
	(13.47)	(13.31)	(0.42)
Male	0.49	0.49	0.01
	(0.50)	(0.50)	(0.02)
Young (< 25 years old)	0.24	0.23	0.01
	(0.43)	(0.42)	(0.01)
Married or cohabitant	0.57	0.59	-0.01
	(0.49)	(0.49)	(0.02)
Employed	0.70	0.70	0.00
	(0.46)	(0.46)	(0.02)
Educational level			
Elementary school drop out	0.08	0.08	0.01
	(0.27)	(0.26)	(0.01)
Elementary school	0.20	0.19	0.01
	(0.40)	(0.39)	(0.01)
High school drop out	0.19	0.20	-0.01
	(0.39)	(0.40)	(0.01)
High school	0.23	0.24	-0.01
	(0.42)	(0.43)	(0.02)
College drop out	0.14	0.13	0.00
	(0.34)	(0.34)	(0.01)
College	0.14	0.14	0
	(0.35)	(0.34)	(0.01)
Income category			
First quintile	0.15	0.13	0.02***
	(0.34)	(0.32)	(0.01)
Second quintile	0.35	0.23	0.12***
	(0.47)	(0.41)	(0.01)
Third quintile	0.20	0.23	-0.03***
	(0.40)	(0.42)	(0.01)
Forth quintile	0.17	0.27	-0.09***
	(0.38)	(0.43)	(0.01)
Fifth quintile	0.13	0.14	-0.01
	(0.32)	(0.32)	(0.01)
Observations	27,348	33,101	60,449
Number of states	15	9	24

Notes: This table presents the mean and standard deviation of individual characteristics for comparison and treated states in columns 1 and 2. Column 3 presents estimated coefficients and standard errors from an OLS regression of the mean difference. Estimates use pooled data from national surveys on tobacco use for 2008 and 2009.

Table A.43: Effects on Extensive Margin Outcomes

Probability of	Never smokers	Current smokers
	(1)	(2)
2008	-0.0205	-0.0055
	(0.0149)	(0.0120)
2009	-0.0238	-0.0138
	(0.0165)	(0.0099)
2013	0.0402**	-0.0594***
	(0.0187)	(0.0100)
Mean dep. var. in 2011	0.4608	0.2825
Observations	153,093	153,093
R-squared	0.0328	0.0259
Correctly predicted	0.5536	0.6881
Individual controls	Yes	Yes
State × time controls	Yes	Yes
State FE	Yes	Yes
Time FE	Yes	Yes

Notes: The omitted category corresponds to 2011, the year the federal law was passed. Treatment as a dichotomous variable, $Treat_s$, that equals one if the legislation index for state s in moment t is strictly less than 3 before 2011. After 2011, all states *are treated* thus $Treat_s$ equals one for all states s after 2011. Individual controls include age, gender, educational attainment, employment status and income category of the household. State × time controls include total private employment and total population. Standard errors are block-bootstrapped at the state-level with 200 replications.

Table A.44: Effects on Intensive Margin Outcomes

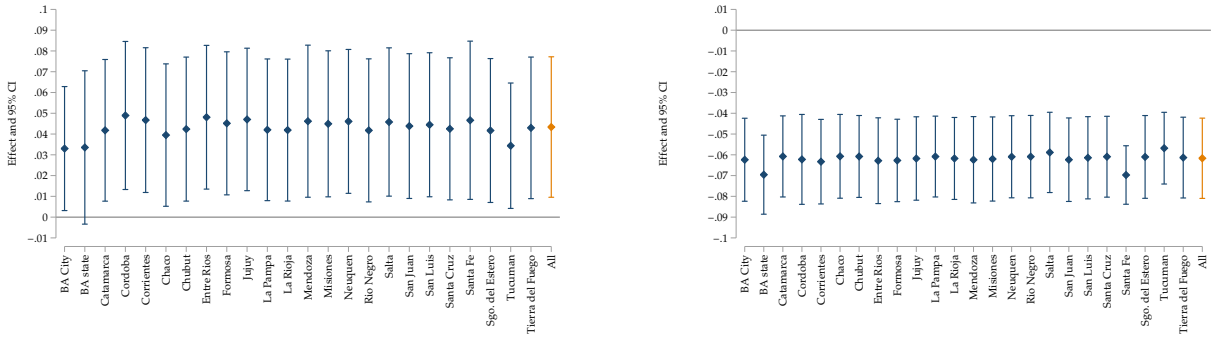
Cigarettes smoked	(0,5]	(5, 10]	(10, 15]	(15, 20]	(20, 30]	(30 , 40]	(40 , 50]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
2008	0.0017	0.0237	-0.0033	0.0084	0.0051	-0.0100	0.0010
	(0.0317)	(0.0185)	(0.0104)	(0.0123)	(0.0067)	(0.0062)	(0.0011)
2009	-0.0839**	0.0394*	0.0020	-0.0182	-0.0111***	-0.0097*	0.0008
	(0.0370)	(0.0201)	(0.0056)	(0.0150)	(0.0035)	(0.0054)	(0.0014)
2013	-0.2335***	-0.0224	-0.0147**	0.0110	0.0073	-0.0073**	-0.0003
	(0.0233)	(0.0159)	(0.0069)	(0.0144)	(0.0045)	(0.0034)	(0.0007)
Mean dep. var. in 2011	0.3397	0.2628	0.0937	0.2080	0.0374	0.0265	0.0024
Observations	45,585	45,585	45,585	45,585	45,585	45,585	45,585
R-squared	0.0797	0.0114	0.0069	0.0257	0.0108	0.0170	0.0014
Correctly predicted	0.8013	0.7775	0.7333	0.7634	0.7211	0.7182	0.7123
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State × time controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*Notes:*The omitted category corresponds to 2011, the year the federal law was passed. Treatment is a dichotomous variable, $Treat_s$, that equals one if the legislation index for state s in moment t is strictly less than 3 before 2011. After 2011, all states *are treated* thus $Treat_s$ equals one for all states s after 2011 Individual controls include age, gender, educational attainment, employment status and income category of the household. State × time controls include total private employment and total population. Standard errors are block-bootstrapped at the state-level with 200 replications.

Removing One State at a Time

Are the effects of the policy driven by one particular state? I address this question by performing a simple exercise: I estimate the effect on extensive margin outcomes in a sub-sample of states where I exclude one state at a time. I present the results of this exercise in Figure A.45, the category "All" replicates the effects discussed in section **??**. This figure presents estimates of the marginal effect and confidence intervals of the 2011 national law on the probability of never smokers (panel a) and current smokers (panel b). I find evidence that point estimates are robust to the exclusion of one state from the sample.

Figure A.45: Extensive Margin Outcomes



(a) Never Smokers

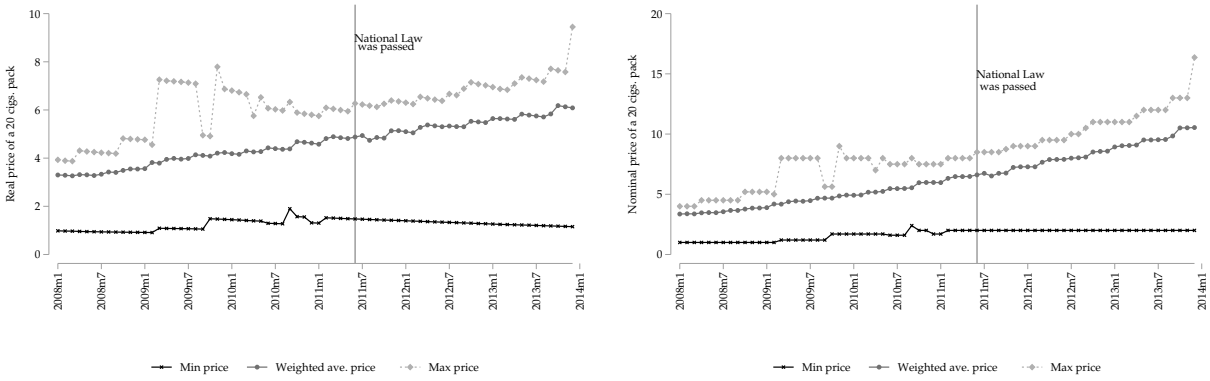
(b) Current Smokers

*Notes:*This figure presents estimates of the marginal effect and confidence intervals of the 2011 national law on the probability never smokers (panel a) and never smokers (panel b). Treatment is a dichotomous variable, $Treat_s$, that equals one if the legislation index for state s in moment t is less than or equal to 3 before 2011. All regressions include individual-level controls: age, gender, educational attainment, employment status, and income category of the household. The category "All" refers to the sample that includes all the states, see table **??**. Standard errors are clustered at the state-level.

Prices and Industry

Did the 2011 policy change prices of cigarettes? I use data on prices to address this question. The Ministry of Agriculture follows the sales of cigarettes at the national level and provides information on sales by the price paid by the consumer. I focus on three price ranges: the minimum price paid in the cigarette market, the average price paid and the highest price paid. Data are at the price-range month level so I can closely follow the evolution around the dates of the change in the regulation. Figure A.46 presents time series of real and nominal prices. I use the price index constructed by Cavallo (2013) to construct the real price time series.[9] There is no evidence of sharp changes in the price of the cigarette package. I pay closer attention to the minimum price because smokers could *easily* substitute among brands to avoid price changes in the more expensive brands.

Figure A.46: Cigarette Prices - Time



Notes: This figure presents time series of real and nominal prices paid by consumers split into three price ranges: the minimum price, the average price, and the highest price. I use the price index constructed by Cavallo, 2013) to deflate prices. Data is at the price-monthly level and comes from price series constructed by the Ministry of Agriculture.

[9]During the period under study, the Argentinean economy experienced high inflation. The lack of reliability on official estimates of the Consumer Price Index challenges the use of the official Consumer Price Index series, for a discussion see, Cavallo (2013).

3.5 Derivations: Bans or Tobacco Graphic Warnings?

The goal of this section is to discuss a framework that allows the identification of the effect of each policy branch: tobacco graphic warnings and place-based bans. The exogenous variation induced by the 2011's anti-tobacco law alone does not allow me to separately identify the effect of each policy branch. Thus, to disentangle the effects, I exploit an additional source of variation and put some structure on how each policy interacts.

Let y_{0i} be the smoking status of individual i in state s in absence of the federal policy and let y_{1i} be the smoking status of individual i in state s if the federal policy is implemented. I assume that the effect of the policy is constant across states (s) and time (t), so the conditional mean function $E[y_{1i}|s,t]$ can be written as: [10]

$$E[y_{1i}|s,t] = E[y_{0i}|s,t] + \delta \ , \tag{3..1}$$

where the parameter δ is the effect of the national regulation, i.e., the effects of the bundle policy. I argue that this effect is a linear combination of the effects of each branch of the policy: tobacco graphic warnings (δ^{tgw}) and place-based bans (δ^{bans}). While tobacco graphic warnings were implemented nationally in 2011, place-based bans were adopted by some states before 2011. Define S_1 as the subset of states that have implemented place-based bans before the national implementation—bans early adopters—and S_2 as the subset of states that have not implemented such bans until the national law—bans late adopters. Under the assumption that the effects of each policy branch are additive and allowing for the presence of spillover effects among the clean-indoor-air policy and the tobacco graphic

[10] In the rest of the derivation I implicitly condition on a vector of covariates $\boldsymbol{X}$ that includes age, educational achievement, among other observable characteristics. I omit this conditioning to ease notation.

warnings, the outcomes if the implementation happens are:

$$
\begin{aligned}
S_1 \quad \text{states:} \quad & E\left[y_{1i}|s \in S_1, t\right] = E\left[y_{0i}|s \in S_1, t\right] + \delta^{tgw} \ , \\
S_2 \quad \text{states:} \quad & E\left[y_{1i}|s \in S_2, t\right] = E\left[y_{0i}|s \in S_2, t\right] + \delta^{bans} + \delta^{tgw} + f(\delta^{bans}, \delta^{tgw}) \ , \qquad (3..2)
\end{aligned}
$$

where $f(\delta^{TGW}, \delta^{bans})$ is the spillover effect of implementing the policies jointly. Thus, from 3..1 it follows that the effect of the bundle policy can be written as:

$$
\begin{aligned}
\delta =& E\left[y_{1i}|s, t\right] - E\left[y_{0i}|s, t\right] \ , \\
\delta =& E\left[y_{1i}|s \in S_1, t\right] P(s \in S_1) + E\left[y_{1i}|s \in S_2, t\right] P(s \in S_2) - \\
& E\left[y_{0i}|s \in S_1, t\right] P(s \in S_1) - E\left[y_{0i}|s \in S_2, t\right] P(s \in S_2) \ .
\end{aligned}
$$

Let $P(s \in S_1) = \omega_1$ and $P(s \in S_2) = \omega_2$ and using the expressions in 3..2:

$$
\begin{aligned}
\delta =& \omega_1 \left(E\left[y_{1i}|s \in S_1, t\right] - E\left[y_{0i}|s \in S_1, t\right]\right) + \omega_2 \left(E\left[y_{1i}|s \in S_2, t\right] - E\left[y_{0i}|s \in S_2, t\right]\right) \ , \\
\delta =& \delta^{tgw} + \omega_2 \delta^{bans} + \omega_2 f(\delta^{bans}, \delta^{tgw}) \ . \qquad (3..3)
\end{aligned}
$$

Where ω_2 can be estimated as the proportion of states that have not implemented a place-based ban before the national regulation, and ω_1 can be estimated as the proportion of states that have implemented a place-based ban before the national regulation. Expression 3..3 shows that under the assumptions of (i) homogeneity of policy effects, and (ii) additive policy effects; the effect of the policy bundle (δ) is a linear combination of the effects of each branch of the policy. With this expression at hand, the next step is to ask whether each of these parameters are identified.

The effect of place-based bans is identified by the comparison of early adopters of place-based bans with late adopters. But, it can not be disentangled from the spillover effects.

Put another way, we can identified and estimate the parameter $\tilde{\delta}^{bans} = \delta^{bans} + f(\delta^{bans}, \delta^{tgw})$. The effect of the national policy is identified by the comparison between lenient states and strict states as discussed before. Thus, I can re-write 3..3 as follows:

$$\delta = \delta^{tgw} + \omega_2 \tilde{\delta}^{bans} \ . \tag{3..4}$$

Estimates of ω_2, δ and $\tilde{\delta}^{bans}$ can be used to construct estimates of the effect of tobacco graphic warnings. The running equation to estimate δ is (3.1), replicated here omitting individual and state-time varying control variables to ease exposition:

$$y_i = \sum_{\tau=-2}^{1} \delta_\tau \left[\text{Treat}_s \cdot (\text{Years after treat} = \tau)\right] + \alpha_s + \alpha_t + \varepsilon_i \ . \tag{3..5}$$

Next, the proportion of states that have not implemented a place-based ban before the national regulation is observed in the data: $\omega_2 = P(s \in S_2)$ and can be estimated as the number of states that are late-adopters of clean-indoor-air bans relative to the total number of states. Finally, the running equation to estimate $\delta^{bans} + f(\delta^{bans}, \delta^{tgw})$ is:

$$y_i = \gamma_0 + \gamma_1 \times 1\{Year = 2013\} + \gamma_2 \times 1\{s \in S_2\} + \tilde{\delta}^{bans} 1\{Year = 2013 \ , s \in S_2\} + u_i \ . \tag{3..6}$$

This regression can be augmented by leads and lags of relative treatment time in an event-study framework:

$$y_i = \sum_{\tau=-2}^{1} \tilde{\delta}_\tau^{bans} \left[\text{Bans}_s \cdot (\text{Years after treat} = \tau)\right] + \alpha_s + \alpha_t + u_i \ , \tag{3..7}$$

where $Bans$ is a dummy variable that equals 1 if the state s is a late-adopter of place-based bans.

Bibliography

Abadie, Alberto. 2003. "Semiparametric Instrumental Variable Estimation of Treatment Response Models." *Journal of Econometrics*, 113(2): 231–263.

Abascal, Winston, Elba Esteves, Beatriz Goja, Franco González Mora, Ana Lorenzo, Amanda Sica, Patricia Triunfo, and Jeffrey E Harris. 2012. "Tobacco Control Campaign in Uruguay: a Population-based Trend Analysis." *The Lancet*, 380(9853): 1575–1582.

Adda, Jérôme, and Francesca Cornaglia. 2010. "The Effect of Bans and Taxes on Passive Smoking." *American Economic Journal: Applied Economics*, 2(1): 1–32.

Aguayo Ormeño, Irina. 2019. "Estadísticas de Jornada Laboral de Trabajadores del Sector Privado." Comisión Trabajo y Previsión Social de la Cámara de Diputados.

Ailes, Elizabeth C., April L. Dawson, Jennifer N. Lind, Suzanne M. Gilboa, Meghan T. Frey, Cheryl S. Broussard, and Margaret A. Honein. 2015. "Opioid Prescription Claims Among Women of Reproductive Age - United States, 2008-2012." *MMWR. Morbidity and Mortality Weekly Report*, 64(2): 37.

Alcaraz, Andrea, Joaquín Caporale, Ariel Bardach, Federico Augustovski, and Andrés Pichon-Riviere. 2016. "Carga de Enfermedad Atribuible al Uso de Tabaco en Argentina y Potencial Impacto del Aumento de Precio a Través de Impuestos." *Revista Panamericana de Salud Pública*, 40: 204–212.

Almond, Douglas, Hilary W. Hoynes, and Diane Whitmore Schanzenbach. 2011. "Inside the War on Poverty: The Impact of Food Stamps on Birth Outcomes." *The Review of Economics and Statistics*, 93(2): 387–403.

Alpert, Abby, David Powell, and Rosalie Liccardo Pacula. 2018. "Supply-Side Drug Policy in the Presence of Substitutes: Evidence From the Introduction of Abuse-Deterrent Opioids." *American Economic Journal: Economic Policy*, 10(4): 1–35.

Alpert, Abby, William N Evans, Ethan M J Lieber, and David Powell. 2022. "Origins of the Opioid Crisis and its Enduring Impacts." *The Quarterly Journal of Economics*. qjab043.

Anderson, Theodore W, Herman Rubin, et al. 1949. "Estimation of The Parameters of a Single Equation in a Complete System of Stochastic Equations." *Annals of Mathematical statistics*, 20(1): 46–63.

Andrews, Isaiah, James H Stock, and Liyang Sun. 2019. "Weak Instruments in Instrumental Variables Regression: Theory and Practice." *Annual Review of Economics*, 11: 727–753.

Autor, David. H., and David Dorn. 2013. "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market." *American Economic Review*, 103(5): 1553–97.

Azagba, Sunday, and Mesbah F. Sharaf. 2012. "The Effect of Graphic Cigarette Warning Labels on Smoking Behavior: Evidence from the Canadian Experience." *Nicotine and Tobacco Research*, 15(3): 708–717.

Behrman, Jere R, and Mark R Rosenzweig. 2004. "Returns to Birthweight." *Review of Economics and Statistics*, 86(2): 586–601.

Bitler, Marianne P, Christopher S Carpenter, and Madeline Zavodny. 2010. "Effects of Venue-specific State Clean Indoor Air Laws on Smoking-related Outcomes." *Health Economics*, 19(12): 1425–1440.

Blecher, Evan. 2008. "The Impact of Tobacco Advertising Bans on Consumption in Developing Countries." *Journal of Health Economics*, 27(4): 930–942.

Böckerman, Petri, Ohto Kanninen, and Ilpo Suoniemi. 2018. "A kink that makes you sick: the effect of sick pay on absence." *Journal of Applied Econometrics*, 33(4): 568–579.

Buchmueller, Thomas C, and Colleen Carey. 2018*a*. "The Effect of Prescription Drug Monitoring Programs on Opioid Utilization in Medicare." *American Economic Journal: Economic Policy*, 10(1): 77–112.

Buchmueller, Thomas C, and Colleen Carey. 2018*b*. "The Effect of Prescription Drug Monitoring Programs on Opioid Utilization in Medicare." *American Economic Journal: Economic Policy*, 10(1): 77–112.

Buckles, Kasey, William N Evans, and Ethan MJ Lieber. 2022. "The drug crisis and the living arrangements of children." *Journal of Health Economics*, 102723.

Buonanno, Paolo, and Marco Ranzani. 2013. "Thank You for Not Smoking: Evidence from the Italian Smoking Ban." *Health Policy*, 109(2): 192–199.

Burton, Anne. 2020. "The Impact of Smoking Bans in Bars and Restaurants on Alcohol Consumption, Smoking, and Alcohol-Related Externalities." *Smoking, and Alcohol-Related Externalities (November 13, 2020)*.

Cabral, Marika, and Marcus Dillender. 2020. "The Impact of Benefit Generosity on Workers' Compensation Claims: Evidence and Implications." NBER Working Paper No. 26976.

Cameron, A Colin, Jonah B Gelbach, and Douglas L Miller. 2008. "Bootstrap-based Improvements for Inference with Clustered Errors." *The Review of Economics and Statistics*, 90(3): 414–427.

Campolieti, Michele, and Douglas E Hyatt. 2006. "Further evidence on the "Monday Effect" in workers' compensation." *ILR Review*, 59(3): 438–450.

Card, David, and Brian P McCall. 1996. "Is workers' compensation covering uninsured medical costs? Evidence from the "monday effect"." *ILR Review*, 49(4): 690–706.

Carpenter, Christopher, Sabina Postolek, and Casey Warman. 2011. "Public-place Smoking Laws and Exposure to Environmental Tobacco Smoke (ETS)." *American Economic Journal: Economic Policy*, 3(3): 35–61.

Carpenter, Christopher S., Chandler B. McClellan, and Daniel I. Rees. 2017. "Economic Conditions, Illicit Drug Use, and Substance Use Disorders in the United States." *Journal of Health Economics*, 52: 63–73.

Case, Anne, and Angus Deaton. 2015. "Rising Morbidity and Mortality in Midlife Among White Non-Hispanic Americans in the 21st Century." *Proceedings of the National Academy of Sciences*, 112(49): 15078–15083.

Case, Anne, and Angus Deaton. 2017. "Mortality and Morbidity in the 21st Century." *Brookings Papers on Economic Activity*, 2017: 397.

Caudillo, Mónica L, and Andrés Villarreal. 2021. "The Opioid Epidemic and Nonmarital Childbearing in the United States, 2000-2016." *Demography*, 58(1): 345–378.

Cavallo, Alberto. 2013. "Online and Official Price Indexes: Measuring Argentina's Inflation." *Journal of Monetary Economics*, 60(2): 152–165.

Cawley, John, and Christopher J Ruhm. 2011. "The Economics of Risky Health Behaviors." 2: 95–199.

CDC. 2021. "Prescription Opioid Epidemic: Know the Facts."

Chang, Fong-ching, Chi-hui Chung, Po-tswen Yu, and Kun-yu Chao. 2010. "The Impact of Graphic Cigarette Warning Labels and Smoke-free Law on Health Awareness and Thoughts of Quitting in Taiwan." *Health Education Research*, 26(2): 179–191.

Chetty, Raj. 2008. "Moral Hazard Versus Liquidity and Optimal Unemployment Insurance." *Journal of Political Economy*, 116(2): 173–234.

Chiang, Chin Long. 1991. "Competing risks in mortality analysis." *Annual review of public health*, 12(1): 281–307.

Cid, Camilo P. 2006. "Evaluación del Subsidio por Incapacidad Laboral por Enfermedad Común en Chile: 1990-2000." *Superintendencia de Salud. Documento de Trabajo*.

Crawfurd, Lee, and Alexis Le Nestour. 2019. "Who Smokes in Developing Countries? Implications for a Tobacco Tax." *Center for Global Development*.

Cronin, Christopher J, Matthew C Harris, and Nicolas R Ziebarth. 2022. "The Anatomy of US Sick Leave Schemes: Evidence from Public School Teachers." NBER Working Paper No. 29956.

Currie, Janet, and Hannes Schwandt. 2021. "The opioid epidemic was not caused by economic distress but by factors that could be more rapidly addressed." *The ANNALS of the American Academy of Political and Social Science*, 695(1): 276–291.

Cutler, David M, and Richard J Zeckhauser. 2000. "The Anatomy of Health Insurance." In *Handbook of Health Economics*. Vol. 1, 563–643. Elsevier.

DEA. 2020. "ARCOS Drug Retail Summary Reports." U.S. Department of Justice, Drug Enforcement Administration, Office of Diversion Control. https://www.deadiversion.usdoj.gov/arcos/.

DeCicca, Philip, Donald S Kenkel, and Michael F Lovenheim. 2022. "The Economics of Tobacco Regulation: A Comprehensive Review." *Journal of Economic Literature. Forthcoming.*

Decker, Sandra L, and Amy Ellen Schwartz. 2000. "Cigarettes and Alcohol: Substitutes or Complements?" National bureau of economic research.

De Paola, Maria, Vincenzo Scoppa, and Valeria Pupo. 2014. "Absenteeism in the Italian public sector: The effects of changes in sick leave policy." *Journal of Labor Economics*, 32(2): 337–360.

Desai, Rishi J., Sonia Hernandez-Diaz, Brian T. Bateman, and Krista F. Huybrechts. 2014. "Increase in Prescription Opioid Use During Pregnancy Among Medicaid-Enrolled Women." *Obstetrics and Gynecology*, 123(5): 997.

Doleac, Jennifer L, and Anita Mukherjee. 2019. "The moral hazard of lifesaving innovations: naloxone access, opioid abuse, and crime." *Opioid Abuse, and Crime (March 31, 2019).*

Eichmeyer, Sarah, and Christina Kent. 2021. "Parenthood in Poverty." Working Paper.

Eichmeyer, Sarah, and Jonathan Zhang. 2020. "Can a Single Opioid Prescription Make a Difference? Evidence From Physician Prescribing Variation in Emergency Departments." Working Paper.

Einav, Liran, Amy Finkelstein, and Mark R Cullen. 2010. "Estimating Welfare in Insurance Markets Using Variation in Prices." *The Quarterly Journal of Economics*, 125(3): 877–921.

Einav, Liran, Amy Finkelstein, Stephen P Ryan, Paul Schrimpf, and Mark R Cullen. 2013. "Selection on Moral Hazard in Health Insurance." *American Economic Review*, 103(1): 178–219.

Einav, Liran, and Amy Finkelstein. 2018. "Moral Hazard in Health Insurance: What We Know and How We Know It." *Journal of the European Economic Association*, 16(4): 957–982.

Eklund, Britt-Marie, Siv Nilsson, Linnea Hedman, and Inger Lindberg. 2012. "Why Do Smokers Diagnosed with COPD Not Quit Smoking? A Qualitative Study." *Tobacco Induced Diseases*, 10(1): 1–7.

EsSalud. 2014. "Normas y Procedimientos Para la Emision, Registro Y Control De Las Certificaciones Médicas Por Incapacidad y Maternidad." Seguro Social de Salud.

Evans, Mary F, Matthew Harris, and Lawrence Kessler. 2020. "The hazards of unwinding the prescription opioid epidemic: implications for child abuse and neglect." *Claremont McKenna College Robert Day School of Economics and Finance Research Paper*, , (3582060).

Evans, William N., Ethan M.J. Lieber, and Patrick Power. 2019. "How the Reformulation of OxyContin Ignited the Heroin Epidemic." *Review of Economics and Statistics*, 101(1): 1–15.

Fernandez, Fernando, and Dijana Zejcirovic. 2018. "The Role of Pharmaceutical Promotion to Physicians in the Opioid Epidemic."

Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams. 2018. "What Drives Prescription Opioid Abuse? Evidence From Migration." *Stanford Institute for Economic Policy Research Working Paper*, 18–028.

Fuchs, Alan, Maria Fernanda Gonzalez Icaza, and Daniela Paula Paz. 2019. "Distributional Effects of Tobacco Taxation: a Comparative Analysis." *World Bank Policy Research Working Paper*, , (8805).

Gabler, Nachum, and Diane Katz. 2010. *Contraband Tobacco in Canada: Tax Policies and Black Market Incentives.* Fraser Institute.

Ganong, Peter, and Jeffrey B Liebman. 2018. "The Decline, Rebound, and Further Rise in SNAP Enrollment: Disentangling Business Cycle Fluctuations and Policy Changes." *American Economic Journal: Economic Policy*, 10(4): 153–76.

Gihleb, Rania, Osea Giuntella, and Ning Zhang. 2022. "The Effect of Mandatory-Access Prescription Drug Monitoring Programs on Foster Care Admissions." *Journal of Human Resources*, 57(1): 217–240.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift. 2020. "Bartik Instruments: What, When, Why, and How." *American Economic Review*, 110(8): 2586–2624.

González-Rozada, Martín. 2006. "Economía del Control del Tabaco en los Países del Mercosur y Estados Asociados: Argentina: 1996-2004."

Goodchild, Mark, Nigar Nargis, and Edouard Tursan d'Espaignet. 2018. "Global Economic Cost of Smoking-Attributable Diseases." *Tobacco Control*, 27(1): 58–64.

Gruber, Jonathan. 2000. "Disability Insurance Benefits and Labor Supply." *Journal of Political Economy*, 108(6): 1162–1183.

Hah, Jennifer M., Brian T. Bateman, John Ratliff, Catherine Curtin, and Eric Sun. 2017. "Chronic Opioid Use After surgery: Implications for Perioperative Management in the Face of the Opioid Epidemic." *Anesthesia and Analgesia*, 125(5): 1733.

Hammond, David. 2011. "Health Warning Messages on Tobacco Products: a Review." *Tobacco Control*, 20(5): 327–337.

Handel, Ben, Igal Hendel, and Michael D Whinston. 2015. "Equilibria in health exchanges: Adverse selection versus reclassification risk." *Econometrica*, 83(4): 1261–1313.

Heil, Sarah H, Hendree E Jones, Amelia Arria, Karol Kaltenbach, Mara Coyle, Gabriele Fischer, Susan Stine, Peter Selby, and Peter R Martin. 2011. "Unintended pregnancy in opioid-abusing women." *Journal of substance abuse treatment*, 40(2): 199–202.

Hendren, Nathaniel. 2017. "Knowledge of Future Job Loss and Implications for Unemployment Insurance." *American Economic Review*, 107(7): 1778–1823.

Ho, Kate, and Robin S Lee. 2020. "Health insurance menu design for large employers." NBER Working Paper No. 27868.

Honoré, Bo E, and Adriana Lleras-Muney. 2006. "Bounds in Competing Risks models and The War on Cancer." *Econometrica*, 74(6): 1675–1698.

Hopenhayn, Hugo A, and Juan Pablo Nicolini. 1997. "Optimal Unemployment Insurance." *Journal of Political Economy*, 105(2): 412–438.

Hoynes, Hilary, Doug Miller, and David Simon. 2015. "Income, the Earned Income Tax Credit, and Infant Health." *American Economic Journal: Economic Policy*, 7(1): 172–211.

Hungerman, Daniel, Tyler Giles, and Tamar Oostrom. 2022. "Opiates of the Masses? Deaths of Despair and the Decline of American Religion."

Johansson, Per, and Mårten Palme. 2005. "Moral Hazard and Sickness Insurance." *Journal of Public economics*, 89(9-10): 1879–1890.

Jones, Mark R., Omar Viswanath, Jacquelin Peck, Alan D. Kaye, Jatinder S. Gill, and Thomas T. Simopoulos. 2018. "A Brief History of the Opioid Epidemic and Strategies for Pain Medicine." *Pain and Therapy*, 7(1): 13–21.

Kleven, Henrik J. 2021. "Sufficient Statistics Revisited." *Annual Review of Economics*, 13: 515–538.

Krauss, Melissa J, Patricia A Cavazos-Rehg, Andrew D Plunk, Laura J Bierut, and Richard A Grucza. 2014. "Effects of State Cigarette Excise Taxes and Smoke-free Air Policies on State Per Capita Alcohol Consumption in the United States, 1980 to 2009." *Alcoholism: Clinical and Experimental Research*, 38(10): 2630–2638.

Kuehnle, Daniel. 2019. "How Effective Are Pictorial Warnings on Tobacco Products? New Evidence on Smoking Behaviour using Australian Panel Data." *Journal of Health Economics*, 67: 102215.

Lee, David L., Justin McCrary, Marcelo J. Moreira, and Jack Porter. 2020. "Valid t-ratio Inference for IV." *arXiv preprint arXiv:2010.05058*.

Low, Hamish, and Luigi Pistaferri. 2015. "Disability Insurance and the Dynamics of the Incentive Insurance Trade-Off." *American Economic Review*, 105(10): 2986–3029.

Lynch, Sean, Laura Sherman, Susan M. Snyder, and Margaret Mattson. 2018. "Trends in Infants Reported to Child Welfare with Neonatal Abstinence Syndrome (NAS)." *Children and Youth Services Review*, 86: 135–141.

Maclean, Catherine, Justine Mallatt, Christopher J. Ruhm, and Kosali Ilayperuma Simon. 2020. "Review of Economic Studies on the Opioid Crisis." NBER Working Paper No. 28067.

Maclean, Johanna Catherine, Stefan Pichler, and Nicolas R Ziebarth. 2020. "Mandated Sick Pay: Coverage, Utilization, and Welfare Effects." NBER Working Paper No. 26832.

Maestas, Nicole A, Kathleen J Mullen, and Stephanie Rennane. 2021. "Absenteeism and Presenteeism Among American Workers." *Journal of Disability Policy Studies*, 32(1): 13–23.

Malbran, Pia. 2007. "What's a Pill Mill?" *CBS News, May*, 31.

Marie, Olivier, and Judit Vall Castello. 2022. "Sick Leave Cuts and (Unhealthy) Returns to Work." *accepted at Journal of Labor Economics*.

Marone, Victoria R, and Adrienne Sabety. 2022. "When Should There Be Vertical Choice in Health Insurance Markets?" *American Economic Review*, 112(1): 304–42.

Meara, Ellen, Jill R Horwitz, Wilson Powell, Lynn McClelland, Weiping Zhou, A James O'malley, and Nancy E Morden. 2016. "State legal restrictions and prescription-opioid use among disabled adults." *New England Journal of Medicine*, 375(1): 44–53.

Meinhofer, Angélica, and Yohanis Angleró-Díaz. 2019. "Trends in Foster Care Entry Among Children Removed From Their Homes Because of Parental Drug Use, 2000 to 2017." *JAMA Pediatrics*, 173(9): 881–883.

Melzack, Ronald. 1990. "The tragedy of Needless Pain." *Scientific American*, 262(2): 27–33.

Miloucheva, Boriana. 2021. "Pharmaceutical Promotion, Physician Response, and Opioid Abuse: Identifying the Role of Physicians in the Opioid Crisis." Working Paper.

Montiel Olea, José Luis, and Carolin Pflueger. 2013. "A Robust Test for Weak Instruments." *Journal of Business & Economic Statistics*, 31(3): 358–369.

Office, United States. General Accounting. 2003. *Prescription Drugs OxyContin Abuse and Diversion and Efforts to Address The Problem: Report to Congressional Requesters.* DIANE Publishing.

of Health, US Department. 2014. "The Health Consequences of Smoking—50 years of Progress: a Report of the Surgeon General."

Oshlack, Benjamin, Mark Chasin, John J Minogue, and Robert Kaiko. 1996. "Controlled Release Oxycodone Compositions."

OxyContin Launch Plan. September 1995. Purdue Pharma.

OxyContin Team Meeting. April 1994. Purdue Pharma.

OxyContin Team Meeting. March 1995. Purdue Pharma.

Park, Sujeong, and David Powell. 2021. "Is the Rise in Illicit Opioids Affecting Labor Supply and Disability Claiming Rates?" *Journal of Health Economics*, 76: 102430.

Patrick, Stephen W, Matthew M Davis, CU Lehmann, and William O Cooper. 2015. "Increasing incidence and geographic distribution of neonatal abstinence syndrome: United States 2009 to 2012." *Journal of Perinatology*, 35(8): 650–655.

Pauly, Mark V. 1968. "The Economics of Moral Hazard: Comment." *The American Economic Review*, 58(3): 531–537.

Pichler, Stefan, and Nicolas R Ziebarth. 2017. "The pros and cons of sick pay schemes: Testing for contagious presenteeism and noncontagious absenteeism behavior." *Journal of Public Economics*, 156: 14–33.

Pierce, Justin R, and Peter K Schott. 2020. "Trade Liberalization and Mortality: Evidence from US Counties." *American Economic Review: Insights*, 2(1): 47–64.

Pieroni, Luca, Manuela Chiavarini, Liliana Minelli, and Luca Salmasi. 2013. "The Role of Anti-Smoking Legislation on Cigarette and Alcohol Consumption Habits in Italy." *Health Policy*, 111(2): 116–126.

Pollak, Catherine. 2017. "The Impact of A Sick Pay Waiting Period on Sick Leave Patterns." *The European Journal of Health Economics*, 18(1): 13–31.

Powell, David, and Seth Seabury. 2018. "Medical Care Spending and Labor Market Outcomes: Evidence from Workers' Compensation Reforms." *American Economic Review*, 108(10): 2995–3027.

Powell, David, Rosalie Liccardo Pacula, and Erin Taylor. 2020. "How Increasing Medical Access to Opioids Contributes to the Opioid Epidemic: Evidence From Medicare Part D." *Journal of Health Economics*, 71: 102286.

Purdue Pharma. 1995. "Purdue Frederick Company Focus Groups Research and Findings: OxyContin for Non-Cancer Pain Management."

Rinehart, Deborah J, Melanie Stowell, Adriana Collings, M Joshua Durfee, Tara Thomas-Gale, Hendrée E Jones, and Ingrid Binswanger. 2021. "Increasing access to family planning services among women receiving medications for opioid use disorder: A pilot randomized trial examining a peer-led navigation intervention." *Journal of substance abuse treatment*, 126: 108318.

Roberts, John M, and Devon D Brewer. 2001. "Measures and tests of heaping in discrete quantitative distributions." *Journal of Applied Statistics*, 28(7): 887–896.

Ruhm, Christopher J. 2019. "Drivers of the Fatal Drug Epidemic." *Journal of Health Economics*, 64: 25–42.

Ruhm, Christopher J. 2021. "Living and Dying in America: An Essay on Deaths of Despair and the Future of Capitalism."

Sales Force Memorandum. 2001. Purdue Pharma.

SAMHSA. 2020. "Key Substance Use and Mental Health Indicators in the United States: Results from the 2019 National Survey on Drug Use and Health."

Savych, Bogdan, David Neumark, and Randall Lea. 2019. "Do Opioids Help Injured Workers Recover and Get Back to Work? The Impact of Opioid Prescriptions on Duration of Temporary Disability." *Industrial Relations: A Journal of Economy and Society*, 58(4): 549–590.

Schnell, Molly, and Janet Currie. 2018. "Addressing the opioid epidemic: is there a role for physician education?" *American journal of health economics*, 4(3): 383–410.

Sharma, Bikash, Ann Bruner, Gabriela Barnett, and Marc Fishman. 2016. "Opioid Use Disorders." *Child and Adolescent Psychiatric Clinics*, 25(3): 473–487.

Shrestha, Vinish. 2018. "Do Young Adults Substitute Cigarettes for Alcohol? Learning from the Master Settlement Agreement." *Review of Economics of the Household*, 16(2): 297–321.

Stone, Rebecca H, Brooke Griffin, Rachel A Fusco, Kathy Vest, Tran Tran, and Savannah Gross. 2020. "Factors affecting contraception access and use in patients with opioid use disorder." *The Journal of Clinical Pharmacology*, 60: S63–S73.

Strayer, Mark, Joshua Leftin, Esa Eslami, et al. 2012. "Characteristics of Supplemental Nutrition Assistance Program Households: Fiscal Year 2011 (Executive Summary)." Mathematica Policy Research.

Sureda, Xisca, Jose M Martínez-Sánchez, Marcela Fu, Raúl Pérez-Ortuño, Cristina Martínez, Esther Carabasa, María J López, Esteve Saltó, José A Pascual, and Esteve Fernández. 2014. "Impact of the Spanish Smoke-free Legislation on Adult, Non-smoker Exposure to Secondhand Smoke: Cross-sectional Surveys Before (2004) and After (2012) Legislation." *PLoS One*, 9(2): e89430.

Tauchmann, Harald, Silja Lenz, Till Requate, and Christoph M Schmidt. 2013. "Tobacco and Alcohol: Complements or Substitutes?" *Empirical Economics*, 45(1): 539–566.

Terplan, Mishka, Dennis J Hand, Melissa Hutchinson, Elizabeth Salisbury-Afshar, and Sarah H Heil. 2015. "Contraceptive Use and Method Choice Among

Women with Opioid and Other Substance Use Disorders: A Systematic Review." *Preventive medicine*, 80: 23–31.

Thoursie, Peter Skogman. 2004. "Reporting Sick: Are Sporting Events Contagious?" *Journal of Applied Econometrics*, 19(6): 809–823.

Thrasher, James F, Rosaura Pérez-Hernández, Kamala Swayampakala, Edna Arillo-Santillán, and Matteo Bottai. 2010. "Policy Support, Norms, and Secondhand Smoke Exposure Before and After Implementation of a Comprehensive Smoke-Free Law in Mexico City." *American Journal of Public Health*, 100(9): 1789–1798.

Tolbert, Charles M, and Molly Sizer. 1996. "US Commuting Zones and Labor Market Areas: A 1990 Update."

Van Zee, Art. 2009. "The Promotion and Marketing of OxyContin: Commercial Triumph, Public Health Tragedy." *American Journal of Public Health*, 99(2): 221–227.

Warner, Margaret, Li Hui Chen, Diane M. Makuc, Robert N. Anderson, and Arialdi M. Miniño. 2011. "Drug Poisoning Deaths in the United States, 1980-2008." *NCHS Data Brief*, 81(81): 1–8.

WHO. 2009. *Global Health Risks: Mortality and Burden of Disease Attributable to Selected Major Risks.* Geneva: World Health Organization.

Ziebarth, Nicolas R. 2013. "Long-term Absenteeism and Moral Hazard - Evidence From A Natural Experiment." *Labour Economics*, 24: 277–292.

Ziebarth, Nicolas R, and Martin Karlsson. 2010. "A natural Experiment on Sick Pay Cuts, Sickness Absence, and Labor Costs." *Journal of Public Economics*, 94(11-12): 1108–1122.

Ziebarth, Nicolas R, and Martin Karlsson. 2014. "The Effects of Expanding the Generosity of the Statutory Sickness Insurance System." *Journal of Applied Econometrics*, 29(2): 208–230.

Ziedan, Engy, and Robert Kaestner. 2020. "Effect of Prescription Opioids and Prescription Opioid Control Policies on Infant Health." NBER Working Paper No. 26749.