Summer 11-24-2017

Public, Tax, and Health Policies and Institutional Performance

Antonios M. Koumpias
Georgia State University

Summer 11-24-2017

Public, Tax, and Health Policies and Institutional Performance

Antonios M. Koumpias

ABSTRACT

PUBLIC, TAX, AND HEALTH POLICIES AND INSTITUTIONAL PERFORMANCE

By

ANTONIOS MARIOS KOUMPIAS

AUGUST 2017

Committee Chair: Dr. Jorge L. Martinez-Vazquez

Major Department: Economics

This dissertation evaluates the effectiveness of public interventions in tax policy (such as a tax compliance campaign in Greece), the performance of public institutions that dictate land zoning (corruption of zoning officials in Greece and Spain) and public health (publicly-provided health insurance; namely, Medicaid). The common underlying theme of the dissertation is the public nature of the policies examines with an empirical emphasis. The ultimate goal of this research body is to provide credible policy solutions for the improvement of public administration.

PUBLIC, TAX, AND HEALTH POLICIES AND INSTITUTIONAL PERFORMANCE

BY

ANTONIOS MARIOS KOUMPIAS

A Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree

of

Doctor of Philosophy in Economics

in the

Andrew Young School of Policy Studies

of

Georgia State University

GEORGIA STATE UNIVERSITY

2017

ACCEPTANCE

This dissertation was prepared under the direction of Antonios M. Koumpias's Dissertation Committee. It has been approved and accepted by all members of that committee, and it has been accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Economics in the Andrew Young School of Policy Studies of Georgia State University.

Dissertation Chair:	Dr. Jorge L. Martinez-Vazquez
Committee:	Dr. Charles M. Becker Dr. Charles Courtemanche Dr. Thomas A. Mroz Dr. Frank A. Sloan

Electronic Version Approved:

Sally Wallace, Dean
Andrew Young School of Policy Studies
Georgia State University
August, 2017

DEDICATION PAGE

To my parents, Socrates Koumpias (DDS, Ph.D.) and Eugenia Koliniotou-Koumpia (DDS, Ph.D.).

ACKNOWLEDGMENTS

I would like to thank my dissertation committee members Prof. Jorge L. Martinez-Vazquez (chair), Prof. Charles M. Becker, Prof. Charles J. Courtemanche, Prof. Thomas A. Mroz, and Prof. Frank A. Sloan for their unwavering support.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

INTRODUCTION

My job market paper, "*The Effects of Compliance Reminders on Personal Income Tax Payments in Greece; Evidence from a Regression Discontinuity Design*", examines the effect of phone call compliance reminders by the tax authority on payments of overdue personal income tax (PIT) liabilities using administrative, individual-level information from 2013 tax records in Greece. The aim of the reminders is to increase awareness about the payment deadlines and warn about the potential tax levies in the case of non-compliance. I take advantage of the combined treatment assignment rule, which leads to a discontinuous relationship between tax delinquents with liability levels exceeding an arbitrary cutoff and the probability they receive a phone call reminder. I exploit this discontinuity to identify the local average treatment effect of the compliance reminders on payments at the cutoff using a fuzzy regression-discontinuity design (RDD). The findings indicate that information nudging may lead to two-way behavioral responses and even backfire under certain conditions. Recipients of phone call compliance reminders pay less. The result is driven by the particularly negative responses of those with past years' tax arrears. This behavior is compatible with a notion of Bayesian updating of taxpayer beliefs about tax debt collection enforcement. Personal income tax delinquents with prior experience with the tax authority eventually learn that the latter may not act on its sanction threats, which renders information nudging non-credible anymore.

In my essay on corruption in public administration, "*Housing Bubbles and Zoning Corruption; Evidence from Greece and Spain*", with Jorge Martinez-Vazquez and Eduardo Sanz-Arcega, we investigate whether the housing bubbles in the Eurozone periphery affected zoning corruption. We exploit variation in rent-seeking opportunities for zoning officials and profit opportunities for housing developers from improved credit conditions and the large rise in housing prices in Greece and Spain. Using aggregate information at the regional level and objective measures of zoning corruption, we estimate negative binomial models for each country separately that indicate a positive and significant association between housing prices and zoning corruption. Specifically, a monthly increase in housing prices (measured in euro per square meter) by 1% led to 3.79% more zoning corruption counts in Greece, on average. For Spain, we find comparable effects around 9% for increases of housing prices by 1% within a quarter, on average. We perform a series of robustness checks that include zero-inflated models, linear models, and dynamic panel analysis.

In my essay about public health "*The Short- and Long-term Effects of State Medicaid Expansions on Mortality: Evidence from a Synthetic Control Design*", with Charles Courtemanche and Daniela Zapata, we study the effect of state Medicaid expansions and one contraction on state-level mortality and morbidity outcomes. This is an important health policy question because expanding and contracting Medicaid is currently an active policy margin of the Affordable Care Act. We contribute a novel study of the dynamics of the effects of state Medicaid expansions to adults on all-cause adult mortality using a synthetic control design. Our analysis offers a comprehensive evaluation of all state Medicaid expansions in the 1990s and 2000s regardless of scope of coverage. Our results suggest that pre-existing estimates of mortality reductions due to Medicaid expansions are overstated. Mortality reductions due to Medicaid are more pronounced in the long-run and only where comprehensive insurance benefits are offered. However, no clear-cut relationship between the size of the Medicaid expansions in terms of total number of Medicaid enrollees and adult mortality can be deduced. It is possible that adult mortality is not driven by the extensive margin of health coverage but the intensive one; i.e., the quality and generosity of each state's Medicaid plan.

Essay 1: The Effects of Compliance Reminders on Personal Income Tax Payments in Greece; Evidence from a Regression Discontinuity Design

1 Introduction

Can behavioral strategies of information provision increase tax revenues in contexts of low voluntary compliance or is taxpayer behavior dictated by non-behavioral motives such as tax enforcement? In recent years, tax authorities throughout the world have enacted policies at an increasing rate that indicate their perception that non-pecuniary factors are central in tax compliance.[1] The aim of these behavioral strategies is improving "tax morale"; i.e., increasing voluntary compliance with the tax code to eventually establish a tax compliance norm (OECD, 2013). Mechanisms under which tax morale, and thus tax compliance, may operate include information imperfections among others (Luttmer and Singhal, 2014).[2]

Information nudging is only one behavioral intervention available to the tax agency among many others which can generally be classified as deterrent or non-deterrent. The former include the use of audits, third-party information reporting, changes in the remittance regime, penalties, shaming through public disclosure, and take-up of benefits. The latter comprise of neutral, information nudges or moral suasion nudges highlighting benefits from tax-funded projects, social

[1] Some of the earliest attempts to enhance tax compliance via nationwide campaigns took place in Spain, the Italian tax authorities put forth Television and print advertising campaigns in Italy to enhance tax compliance during the European debt crisis (Povoledo, 2011), and the Estonian Tax and Customs Board disseminated information in 2010 and 2011 to raise awareness among the population about the use of taxes by the state (Eurofound, 2013). Public disclosure of tax delinquent lists online have been used by tax administrators in Argentina, Bosnia and Herzegovina, Croatia, El Salvador, Greece, Mexico, Montenegro, Portugal, Serbia, Slovenia, Spain, the United Kingdom and U.S. states whereas names of tax evaders in Canada, Ireland, Italy and New Zealand have been listed in print.(Perez-Truglia and Troiano, 2015)

[2] Luttmer and Singhal (2014) define tax morale broadly as an "umbrella term capturing non-pecuniary motivations from tax compliance as well as factors that fall outside the standard, expected utility framework" operating through intrinsic motivation, reciprocity, peer and social effects, culture, information imperfections, and deviations from utility maximization.

norms, social comparisons, or emphasizing tax fairness or equity. A large body of empirical literature has emerged that examines behavioral strategies as a mechanism to enhance tax compliance using predominantly field experiments. The general takeaway is that both deterrent and non-deterrent strategies can substantially increase tax compliance in contexts of high tax morale. However, mainly deterrent strategies can effectively improve compliance in low tax morale environments. The existing evidence indicates that non-deterrent behavioral strategies generally have null effects on voluntary compliance. Slemrod (2015) provides an excellent review of the recent tax compliance and enforcement literature.

This paper evaluates the impact of a tax compliance information campaign on tax payments in Greece. In 2013, the Tax Compliance Unit (TCU) of the General Directorate of Public Revenues of the Ministry of Finance of Greece used e-mails and phone calls to remind personal income tax (PIT) delinquent individuals of upcoming payment deadlines and increase awareness about tax levies. All PIT delinquent individuals with a registered e-mail address received an e-mail reminder. Then, a fraction of this population received a phone call reminder, giving rise to a natural experiment of increased salience in information provision. Using individual-level information from de-identified tax records, I study the incremental effect of the phone call compliance reminder on payments of PIT liabilities. To disentangle the causal effect of the phone call reminder, I employ a fuzzy regression discontinuity design (RDD). I take advantage of TCU's treatment assignment rule, which leads to a discontinuous relationship between individuals with PIT liability levels exceeding €500 and the probability they received a phone call reminder. This discontinuity generates plausibly exogenous variation in treatment assignment which I exploit to identify the causal effect of the phone call reminder on payments of PIT liabilities. Specifically, there is a jump in the conditional mean of the treatment variable at the €500 cutoff that provides identifying variation. This allows me to estimate the effects of the reminder, within a neighborhood of the €500 cutoff, as if they were randomly assigned. Using a fuzzy RDD estimator, I identify the local average treatment effect (LATE) of the phone call reminder conditional on an e-mail reminder on payments of PIT liabilities.

This study's contribution is threefold. First, it provides a novel finding by showing that a behavioral tax compliance campaign may yield unintended consequences by decreasing tax revenues. Second, it is the first study of the effect of national-level behavioral nudging on tax payments in a OECD country with low tax morale context and weak tax debt collection enforcement. The majority of the applied behavioral tax compliance literature has examined either contexts of high voluntary tax compliance or cases of local taxes. Third, it extends our limited knowledge about the effectiveness of different delivery devices; namely, e-mails as opposed to phone calls. The focus of the pre-existing literature is almost exclusively on message content rather than delivery device. The findings of this paper also have important policy implications. They directly inform the revenue administration in Greece about the efficacy of non-pecuniary, behavioral strategies in increasing tax payments. The findings here are generalizable to contexts of low voluntary compliance where tax revenues are, also, more important for public finances.

Information nudging is expected to increase tax payments by minimizing informational asymmetries and taxpayer procrastination. I provide a simple theoretical model and empirical evidence that show under which conditions information nudging can backfire. Taxpayer uncertainty about the probability of debt collection enforcement may increase the present value of expected tax evasion and reduce expected tax payments. This is particularly true when taxpayers update their beliefs about the probability of tax debt collection enforcement in the future downwards following a period of weak enforcement. Therefore, the impact of compliance reminders emphasizing the likelihood of tax levies on tax compliance is not ex-ante clear and warrants empirical investigation. The estimated effects for tax delinquents with income tax liability levels around €500 indicate that the compliance reminders may actually reduce tax payments. On the extensive margin, the phone call reminders led to a 9.17%-9.14% decrease in the probability of any payment. On the intensive margin, the reminders caused a 21.4%-40.8% reduction in tax revenues, on average. When I split the sample by debt vintage, I find that the overall behavioral response is indeed driven by taxpayers' experience of tax debt collection enforcement. Newly-delinquent personal income taxpayers do not change their payment behavior in any significant way. On the contrary, those with tax arrears from previous years do react to the compliance reminders by paying substantially less. This behavior is compatible with a notion of Bayesian updating of taxpayer beliefs about tax debt collection enforcement. Personal income tax delinquents with prior experience with the tax authority eventually learn that the latter may not act on its sanction threats which renders information nudging non-credible anymore.

The rest of the paper is organized as follows. Section 2 reviews the literature of behavioral tax compliance with a focus on communication strategies from tax authorities. Section 3 presents a theoretical model of Bayesian tax delinquents with uncertainty about tax debt collection enforcement. Section 4 describe the empirical discusses the data and methodology and Section 5 presents the results. Section 6 provides robustness checks of the empirical analysis and Section 7 concludes.

2 Literature Review

The seminal theoretical work in tax compliance is due to Allingham and Sandmo (1972) who adapted the Becker (1968) economic model of crime to explain why taxpayers evade. Given that the actual level of tax compliance tends to be higher than what the expected utility model would predict, more recently, the focus shifted to what motivates taxpayers to pay their taxes (Andreoni et al., 1998). Often, tax declarations or tax deductions have been used as proxies of tax evasion (Alm, 2012). However, Slemrod and Weber (2012) question those findings due to the often insurmountable measurement problems these studies suffer from. Hallsworth et al. (2014) argue that focusing on tax payments rather than tax declarations would help mitigate such problems because the causal path between treatment and outcome (compliance or delinquency) is direct. An intuitive, and relevant to this study, way of distinguishing these findings is by the nature (deterrent or non-deterrent) and delivery device of the behavioral treatment.

2.1 Deterrence Behavioral Strategies

The first empirical study in behavioral tax compliance find that in Minnesota a higher perceived audit probability can generate substantial taxpayer responses (Slemrod et al., 2001). Similarly, Iyer et al. (2010) find that letters that increased taxpayers' perception of non-compliance detection risk and penalty awareness generated an improvement in tax compliance in the state of Washington. However, the dissemination of information that would increase the perceived audit probability will not necessarily improve tax compliance. Ariel (2012) reports null or negative deterrence effects of moral persuasion messaging on tax payments and tax deductions on behalf of Israeli firms. Often, the key to understanding what drives taxpayer behavior is the level of third-party information about the taxpayer's economic transactions that can be observed by the tax agency. Also, this appears to hold true over a number of different types of taxes. In a field experiment conducted in Denmark, Kleven et al. (2011) study the effects of tax enforcement to arrive at negligible tax evasion rates for income liable to third-party reporting, but significant for self-reported income. In the context of Chile, Pomeranz (2013) reports that an increase in the perceived audit probability of randomly chosen Chilean firms increases compliance to the VAT. Castro and Scartascini (2015) find that sanction-based messaging successfully generated more property tax revenues to the Argentine tax authority. Ortega and Sanguinetti (2013) show that enforcement messaging has the largest effect on compliance of a local business tax in Caracas, Venezuela. Del Carpio (2014) finds that a combination of a payment reminder and deterrence information enhances tax compliance of a local property tax in Peru because it increases the perception of enforcement. In Norway, Bott et al. (2014) show that an increase in the perceived audit probability can improve tax compliance on the extensive margin. Their results show that deterrence messaging induced a great proportion of the taxpayer population to file a foreign income tax return. Perez-Truglia and Troiano (2015) examine the effects of a shaming strategy on payments of state tax delinquencies in the US. Their results highlight that shaming can force more tax delinquents to settle their overdue liabilities but not in cases of large-sized debts.

2.2 Non-Deterrence Behavioral Strategies

Non-deterrence behavioral strategies involve neutral, non-deterrent nudges that disclose general information about payment deadlines and consequences of non-compliance. In addition, tax authorities also use non-neutral messages that attempt to increase the moral cost of delinquency through moral appeals to social norms, national identity, public goods, tax equity and fairness. The communication campaign by the Greek tax authorities that this paper examines belongs in this subset of tax compliance behavioral strategies.

The earliest findings from Minnesota suggest that normative appeals do not have an effect on reported income in Minnesota (Blumenthal et al., 2001). Evidence from firms and individuals in Australia by Wenzel (2005) and Wenzel and Taylor (2004) provide only weak support for the use of norm messaging. Furthermore, Torgler (2004, 2012) finds anemic effects of mail-based moral appeals on tax compliance in Switzerland. In the same fashion, Fellner et al. (2009) report that moral suasion and social information letters do not improve compliance of individuals who evaded TV and radio licensing fees in Austria. In fact, the authors show that appeals to morality or underlying social norms can backfire when letter recipients object to the moral values expressed or when local non-compliance is widespread. Appeals to tax fairness and tax equity do not effectively increase property tax revenues in Argentina according to Castro and Scartascini (2015). Experimental evidence from Venezuela by Ortega and Sanguinetti (2013) also suggests that moral suasion is ineffective in bringing about taxpayer responses. Doerrenberg and Schmitz (2015) examine the effect of informational provision on small accounting firms in a field experiment in Kranj, Slovenia but do not arrive at any statistically significant results (potentially due to a low sample size). On the contrary, Dwenger et al. (2014) report a positive degree of compliance with the local German Protestant church tax when actual and perceived enforcement is nonexistent. This results shows that, in the context of the German Protestant church tax, individuals comply with the law due to intrinsic motivations. Hallsworth et al. (2014) conduct two natural field experiments in the UK in 2011 and 2012 that challenged our knowledge surrounding the effectiveness of messaging onto payment rates. The authors use norm, public goods and fairness messages in a standard tax payment reminder letter to effectively increase tax payments. Bott et al. (2014) implement a natural field experiment on foreign income tax reporting behavior in Norway. They find that moral appeals can increase the amount reported by those who report any foreign income, generating more tax revenues on the intensive margin. In the US, Guyton et al. (2016) use reminders to examine the tax filing behavior of lower-income individuals.

2.3 Communication Delivery Devices

In earlier work, the experimental variation was almost always introduced on the dimension of message content rather than delivery device. Mail letters have served as the baseline communication device tax authorities employ. However, it is not clear whether the mode of communication can determine a campaign's success. In fact, the delivery device might play a crucial role in a campaign's success.[3] Indeed, Ortega and Scartascini (2015) report sizable differences across the effectiveness of a letter, an e-mail, and a personal visit by a tax inspector in a field experiment of the National Tax Agency in Colombia. Similarly, Doerrenberg and Schmitz (2015) show that in-home visits were more effective than letters to small accounting firms in Slovenia. These two studies provide evidence that taxpayer responses are stronger as the salience of the delivery device increases.

2.4 Empirical Designs

With the exception of Sanchez (2014), the recent wave of empirical evidence comes exclusively from field experiments. However, the latter are not necessarily immune to bias (Imai, 2005). More importantly, they are hard to replicate by the tax agency, let alone initiate by the average unaccustomed bureaucracy. From the taxman's perspective, policies which are straightforward to implement, exhibit ease of scalability, and are highly accessible may be preferable to finely-tuned field experiments that typically require a research unit. The only evidence in the recent tax compliance behavioral literature that is not obtained from a field experiment comes from Sanchez (2014) who examines a natural experiment in Ecuador. He finds that simple notifications of outstanding

[3] In the context of get-out-the-vote calls, Imai (2005) shows that phone calls are associated with more substantial behavioral responses than non-personalized methods such as flyers. In the same fashion, Stollwerk (2007) find e-mails to be ineffective in voter mobilization in terms of registration and turnout.

tax liability send by the national tax agency to firms with under-reported corporate income tax can dramatically improve tax payments.

In sum, this paper's main contributions to the literature comes from the examination of the effectiveness of behavioral tax compliance strategies at the national level in a context of low tax morale. In addition, this study extends our limited knowledge about the effectiveness of different nudging technologies (i.e., e-mail vs phone-call). Finally, it deviates from the latest trend in behavioral tax compliance of using field experiments and, instead, relies on a natural experiment that is easily reproducible by any tax agency.

3 Theoretical Framework

Tax debt collection enforcement is a key feature of the tax system. However, it is a relatively under-studied topic within the theoretical tax compliance literature. Perez-Truglia and Troiano (2015) present a model of taxpayer behavior that incorporates the degree of individuals' income garnishability. They examine the efficacy of pecuniary and non-pecuniary tax debt enforcement strategies for the tax authority, namely financial and shaming penalties. Paramonova (2016) provides a dynamic model of tax debt with taxpayer heterogeneity in income and ability to escape tax debt payment. She shows that alternative tax collection tools such as non-pecuniary tax sanctions in the form of a driver's license suspension can serve as complementary means of tax debt collection enforcement under certain conditions. Here, I consider how former tax debt collection enforcement experiences affect current taxpayer behavior. This section provides the theoretical foundation for the empirical investigation that follows.

First, I adapt the framework Snow and Warren (2007) offer by shifting the focus from taxpayer beliefs about the audit probability to tax debt collection enforcement. The result is a model of Bayesian tax delinquents with uncertainty about tax debt collection enforcement. With regards to the taxpayer's problem, the interest lies on how exposure to collection enforcement of tax debt in the past affect current tax compliance. Second, I consider optimal tax policy for a range of government preferences over tax revenues and private welfare following Perez-Truglia and Troiano (2015). With respect to the government's problem, I characterize the optimal mix of deterrent (pecuniary) and non-deterrent (behavioral, non-pecuniary) strategies. The solutions to the tax delinquent's and the government's problem are given in Appendix A.

3.1 A Model of Bayesian Income Tax Delinquents with Uncertainty about Tax Debt Collection Enforcement

Consider a two-period economy with a continuum of taxpayers indexed by subscript i, who each has a tax delinquency of measure one.[4] Tax delinquents prefer to defer paying in period two given liquidity constraints. In other words, the cost of complying in period one exceeds the level of the tax liability, $R_i > 1$. The severity of a tax delinquent's liquidity constraint is uniformly distributed between $\underline{R}$ and $\overline{R}$. The variation in R_i reflects individual access to credit through formal (retail banking) or informal (family, loanshark) lines. R_i is observable to the tax agency with a degree of error through individual repayment history, employment status and other factors. The tax agency also prefers revenues in the first period rather than in second one; i.e. it has a discount rate $R_g > 1$. (Perez-Truglia and Troiano, 2015).

Each period is characterized by a different state of the world regarding tax debt collection en-

[4] What follows in this subsection is an adaptation of Snow and Warren (2007)'s model of taxpayer behavior appended with a consideration of the government's problem.

forcement. Specifically, there are $k \in \{e,n\}$ states of the world where $k = e$ denotes perfect prior enforcement whereas $k = n$ denotes an imperfect one.[5] The state of tax debt collection enforcement is revealed by nature reflecting that is is not fully within the control of the tax authority. Imperfect enforcement may arise either from administrative capacity issues or represent government preferences.[6] Unpaid tax liabilities accumulate as a stock in the form of tax arrears.

The objective probability of perfect debt collection enforcement is denoted by p and is only observable by the tax agency. Tax delinquents form subjective beliefs q_i about the probability of perfect tax debt collection enforcement. Their uncertainty is reflected by the cumulative distribution function $F(q_i)$. The tax authority can influence subjective beliefs q_i through its communication strategy α such that effective subjective beliefs given by q_i^{α}. The parameter $\alpha \leq 1$ is an amplifier of the taxpayer's recollection of each tax debt collection enforcement experience. This variable is determined by the tax authority through its communication strategy. The value of α is decreasing with the salience of the nudging device used (e.g. e-mail, phone-call, letter, in-home visit) such that more salient devices are more influential (detrimental) in the formation of taxpayers' subjective beliefs about debt collection enforcement. If no communication strategy (no nudging) is employed by the tax agency then $\alpha = 1$. Essentially, parameter α reflects how invested the government is in shifting taxpayers' beliefs about debt collection enforcement. Since taxpayers maximize expected utility, the key parameter for their decision-making is π_i, the mean of $F(q_i)$, which is equal to $\pi_i = \int_0^1 q_i^{\alpha} dF(q_i)$. In both periods, the taxpayer's decision of whether to skip

[5] Imperfect enforcement of tax debt collection is defined as zero or partial enforcement of collection of outstanding delinquent debt by the tax agency.

[6] Such preferences provide a reasonable explanation for the tax repayment plans that tax agencies around the world are offering. A prominent example in the US is the IRS' Fresh Start, a 72-month installment repayment plan. Such repayment plans are policy relevant to cash-strapped governments as in the context I examine. As of 2016, the Greek tax agency offers a repayment plan of 100 installments. This is a pilot program that commenced in 2014 after the period of the subsequent empirical analysis. For my study population, tax liabilities that become delinquent in 2013, a shorter 12-month installment plan was available to individuals.

paying entirely $[x_i(R,\pi_i)=1]$ the predetermined, reported tax liability that just became delinquent or pay in full $[x_i(R,\pi_i)=0]$ is based on her prior mean expectation of tax debt collection enforcement π_i. However, if the tax agency warns about payment deadlines in a relatively salient way in period one but fails to enforce them, then beliefs of imperfect debt collection enforcement in period two become more prevalent.

Period One: Taxpayer Formation of Beliefs about Tax Debt Collection Enforcement

In period one, the taxpayer has no former experience as a tax delinquent in period one. A reasonable, uninformative prior a newly-delinquent taxpayer might form is the expectation that the outstanding debt of unity will be collected in full through tax liens, $q_i=1$. In a context of weaker underlying enforcement or overall institutional quality, she might alternatively expect imperfect debt collection, $q_i \in (0,1)$, though. The taxpayer's problem in period one narrows down to choosing the optimal levels of payment x_i^1 and saving under uncertainty about the probability of perfect debt collection enforcement to maximize period one expected utility given the level of the financial penalty, $\theta > 1$. From the additional income x_i^1 available due to imperfect debt collection enforcement in period one, an amount s_i is saved for period two.

$$\max_{x_i^1,s_i\in[0,1]} EU(x_i^1;s_i;\pi_i) = (1-\pi_i)\cdot[u(y_i+x_i^1-s_i)+R_i\cdot V(s_i,\pi_i^n)]+\pi_i\cdot[u(y_i-\theta x_i^1-s_i)+R_i\cdot V(s_i,\pi_i^e)] \tag{1}$$

where y_i is exogenously-determined, period one, post-tax income. The parameter $\theta > 1$ is a multiplier that accounts for the penalties incurred on first period's outstanding tax delinquencies.

The taxpayer's beliefs about perfect debt collection in the second period have only an indirect income effect (operationalized through the choice of savings) on the likelihood of not paying the outstanding tax liability in the first period and risk penalties in the case of enforcement. This is true when the taxpayer exhibits decreasing absolute risk aversion (DARA) which is an empirically relevant assumption. Thus, first-period net income increases as the current tax debt collection enforcement experience becomes more informative about the probability of perfect tax debt collection enforcement in the second period. However, there is no effect on tax evasion when the taxpayer exhibits constant absolute risk aversion (CARA) because the income change has no effect on the decision to not pay. The discrepancy arises due to the relatively more risk-loving nature of the DARA taxpayers. Under DARA, saving decreases due to the positive effect of increased π^e on second-period expected utility.

Period Two: Tax Collection under Informational Asymmetries

After the experience of the debt collection enforcement in the first period, the tax delinquent's posterior belief is updated following Bayes' rule to the conditional probability distribution $F(q_i|k)$. The expected conditional probability of perfect debt collection enforcement in the second period is equal to $\pi_i^k = \int_0^1 q_i^\alpha dF(q_i|k)$. Following Snow and Warren (2007), it is assumed that the posterior probability decreases relative to the prior if enforcement is imperfect ($\pi_i^n < \pi_i$) and vice-versa ($\pi_i^e > \pi_i$). Internally consistent taxpayer beliefs require the ex-ante expected value of posterior beliefs about the probability of perfect debt collection enforcement to be equal to the prior's value; i.e., $(1-\pi_i)\pi_i^n + \pi_i\pi_i^e = \pi_i$. The income available in period two from evading in period one is also taxable. The taxpayer's indirect expected utility maximization problem in period two is choosing the fraction x_i^2 of the second period tax arrears to skip paying. Now, the taxpayer decides how much of her tax arrears to pay based on her posterior subjective belief of being enforced to service her outstanding debt.

$$\max_{x_i^2 \in [0,1], \pi_i^k \in [0,1]} V(x_i^2; \pi_i^k) = (1-\pi_i^k) u[(1+r)s_i + x_i^2] + \pi_i^k u[(1+r)s_i - \theta \cdot x_i^2] \tag{2}$$

where s_i is the level of savings from period one available in period two at the prevailing interest rate r.

Assuming Decreasing Absolute Risk Aversion (DARA), it is not clear whether tax evasion will increase or decrease in the second-period. Second period tax evasion may in principle be lower but is assumed to be dominated by the first-period increase such that tax evasion increases with Bayesian updating over time. Due to lower savings in period one, the taxpayer is less inclined to bear risk in period two. However under Constant Absolute Risk Aversion (CARA), expected second-period income tax evasion increases as the taxpayer updates her beliefs about perfect debt collection enforcement based on previous experience. For the case of constant relative risk aversion (CRRA) equal to one, tax evasion is higher in period one, lower in period period two, and higher overall due to Bayesian updating (assuming rate of return equal to social discount factor).

3.2 Optimal Tax Policy

The benefit to the government when the tax delinquency is resolved in period 1 relative to period 2 is reflected by $R_g > 1$. Note that the government treats all tax revenues in period one equally, regardless of the liquidity constraints the tax delinquents face. Let x_i^* denote the best-response function of tax delinquent i. Similar to Perez-Truglia and Troiano (2015), the government maximizes social welfare which is a weighted average of the tax agency's tax revenues and tax delinquents' private welfare. It has two choice variables to its disposal: the level of the financial penalty θ and the intensity of the tax debt collection enforcement experience, $\alpha \leq 1$, through the choice of its communication strategy with tax delinquents. Essentially, the government is considering the socially optimal mix of incentives to avoid crowding-out prosocial behavior (Benabou and Tirole, 2006).

$$\max_{\theta>1,\alpha\leq 1} SW(\theta,\alpha) = \psi \cdot T(\theta,\alpha) + (1-\psi) \cdot PW(\theta,\alpha) \tag{3}$$

The parameter $\psi \in [0,1]$ denotes the government's preference in the second period for an extra dollar in its coffers relative to the taxpayers' pockets. It could be interpreted as the rate of return from government spending. For a non-benevolent government that enjoys a higher marginal return from government expenditures than private welfare $\psi \in (\frac{1}{2},1)$. This might be preferable because tax revenues can be used to provide the efficient level of public goods (Samuelson, 1954). For a benevolent government that faces a higher marginal return from taxpayers' private welfare than its own tax revenues $\psi \in (0,\frac{1}{2})$. For example, an inefficient government may simply prefer its constituents to pocket the tax revenues its tax collection system would have otherwise failed to capture. When the government is indifferent between tax revenues and private welfare, the first-best can be attained by eliminating all informational asymmetries about the probability of perfect debt collection enforcement. That is, the government should not allow the tax delinquents to form beliefs about debt collection enforcement but engage in a communication campaign if debt collection is perfectly enforced. Clearly, this is also true when the government places higher value on private welfare than tax revenues allowing tax delinquents to optimize their re-payment behavior. However, when the government puts more weight on tax revenues than private welfare and debt collection is imperfectly enforced, it is optimal to not correct all informational asymmetries in first period and surprise tax evaders in the second period by collecting both principal and interest on the tax liability.[7]

Bayesian updating implies an expected net loss of tax revenues, inclusive of penalties. The theoretical predictions about the incidence of these losses in tax revenues depend on the risk preferences of the taxpayer. Following CARA, tax evasion is unaffected in period one, saving is lower which leads to a narrower tax base for period two, and expected tax evasion in period two increases. Following CRRA of unity, both saving is decreased and expected evasion is increased such that expected tax revenues are lower. For the empirically-relevant case of risk preferences, DARA, there will be an increase in expected tax evasion in period one when allowing taxpayers to form beliefs about tax debt collection enforcement. In period two, expected tax evasion could decrease if tax debt collection in period one was strongly enforced but the net effect on tax revenues may still be negative due to a narrower tax base following decreased first-period savings. If tax debt was weakly enforced in period one, then the reduction in tax revenues due to a narrower tax base is compounded by increased non-compliance. Taxpayer Bayesian updating provides another window of opportunity for tax evasion. The policy implication with regards to tax revenues is that reduced uncertainty about the probability of perfect debt collection enforcement may enhance tax compliance. Thus, the compliance reminders can play a prominent role in enhancing tax compliance as devices that allow the tax authority to appropriately re-anchor taxpayer beliefs about tax debt collection enforcement.

[7] Since some taxes are paid by non-citizens this could plausibly be the case.

4 Empirical Analysis

This section discusses this study's empirical approach.

4.1 Institutional Background: Taxation in Greece

Greece is a country with a large shadow economy. Schneider et al. (2010) find that among the OECD countries Greece features the second largest shadow economy, representing 28.9% of GDP during the years leading up to the recent financial crisis. Tax withholding in Greece is handled as in a conventional Pay-As-You-Earn (PAYE) program where employers must withhold income tax from their employees' compensation. However, large-scale employee underreporting on behalf of the employer (to evade social contributions) implies that even the PAYE system captures only a fraction of income tax base. Indeed, Matsaganis and Flevotomou (2010) estimate income underreporting at 10%, precipitating a 26.1% shortfall in tax yield. Artavanis (2016) exploits VAT tax rate changes in 2011 and 2013 to document the disclosure of "hidden" sales as restaurant firms adjusted their reported VAT sales ratio (reported sales over reported inputs). Using confidential, micro-level bank credit information, Artavanis et al. (2015) estimate a lower bound of €28 billion of unreported income from self-employed people alone for Greece in 2009. Litina and Palivos (2016) provide empirical evidence and a theoretical framework that highlight the key role of political corruption in the exacerbation of tax evasion in Greece in recent years.

Revenue administration in Greece is facing a bleak outlook with regards to tax revenues. As of 2013, tax liabilities and arrears by individuals to the Greek government totaled US$64.67 billion. To put this in perspective, the general government debt owed by Greece to its international creditors was US$342.31 billion (174.9% of GDP). Undoubtedly, the severe economic crisis (on-going recession from 2009) has largely contributed to the rise in tax delinquencies. But, weak tax debt collection enforcement has been another important reason for such sluggish tax collections. For example, the general government in 2012 alone accrued \$3.48 billion of new tax debt (25% of total assessed taxes) of which only \$1.57 billion was collected. By international standards, the available instruments for tax debt collection enforcement are deemed inadequate IMF (2013). Figure 1 shows that in 2011 tax debt that went uncollected almost matched net tax revenue collections (89.5% relative to a 12.4% and 11.4% EU and OECD average, respectively). To make things worse, policy makers have not been able to collect from the "hard to tax" and, largely, rely on "easier" tax handles instead (Alm and Martinez-Vazquez, 2003). For a elaborate illustration of all the available pathways to tax evasion in Greece refer to Figure C.1 in Appendix C.

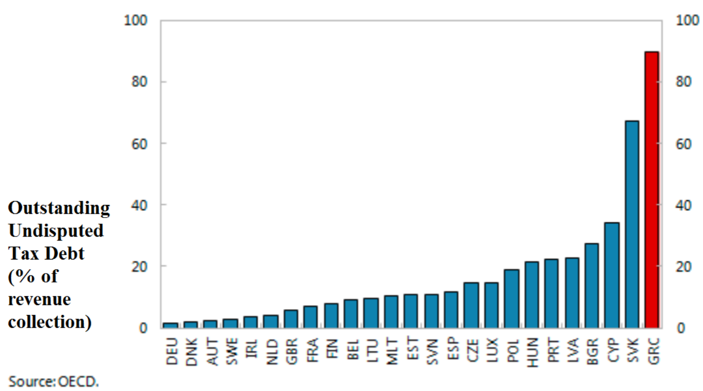

Figure 1: Comparison of Outstanding Undisputed Tax Debt among EU Member States

It is important to note that the tax burden borne by individuals and firms in Greece increased significantly and frequently, in recent years. Just in 2013, two legislative acts amended provisions of the country's income and VAT tax code that raised new rates, and lowered tax eligibility thresholds and tax reliefs.[8]

[8]In January 2013, the Tax Law No.4110 (No. 4110 / Government Gazette FEK No. 17 A'/23.01.2013 "Income Tax Provisions, Issues related to the authorities of the Ministry of Finance and Other Regulations") amended provisions of the Income Tax Code. Only six months later, The New Law No. 4111 (No. 4111 / Government Gazette FEK No. 18

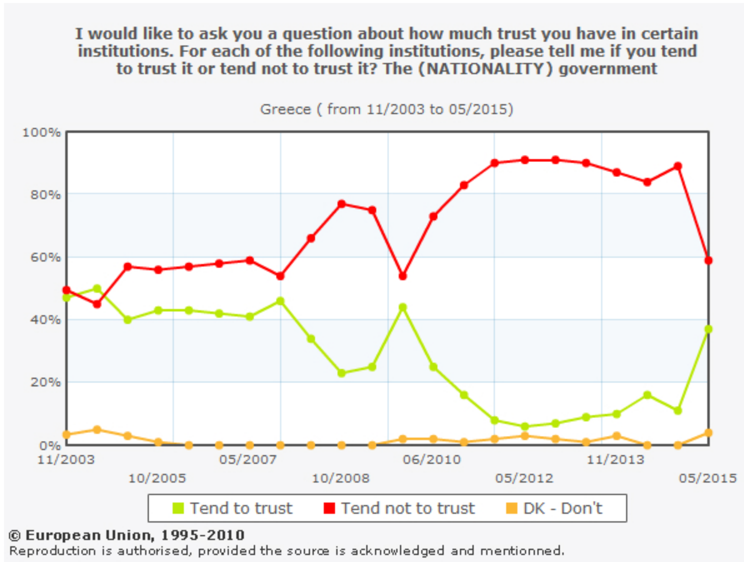

Figure 2: Eurobarometer: Trust in Government; Nov 2011 - May 2015

4.2 Description of the Natural Experiment

In an effort to increase much-needed tax revenues in 2013, the Tax Compliance Unit (TCU) of the Department of Public Revenues of the Ministry of Finance of Greece used phone-calls and e-mails to remind personal income tax delinquent individuals of upcoming payment deadlines and warn about potential tax liens. In particular, the reminder warned that measures will be taken in the case of non-compliance but did not explicitly mention financial (as interest accrual) or non-financial penalties (in the form of barriers to entering the labor market) for individuals. The nature of the intervention could be understood as a neutral nudge of information provision and, thus, operate within the fifth class of suggested tax morale mechanisms above.

The e-mail campaign was massive as e-mails were sent to virtually any tax delinquent, individual or firm, with an e-mail address registered to the tax agency. However, the TCU never perceived the e-mail campaign to be a particularly effective communication strategy. Given the

A'/25.01.2013 "New Law relating to various legislative provisions of urgent nature") was published in the Government Gazette replacing the Income Tax Code.

increasing complexity of the tax code in Greece that occurred during the crisis years, many individuals resorted to accountants for support in filing their taxes. In recognition of this phenomenon, the TCU was concerned that a large proportion of the e-mail reminders were sent to the tax delinquent's accountant e-mail address rather than himself. To increase the salience of the informational nudge, the TCU decided to enact a phone call campaign. An internal call center reached delinquent individuals to restate the information contained in the e-mail over the phone. Given the limited resources that the TCU could devote to a call center, the reach of the phone call campaign was much smaller relative to the one of the costless e-mails. To maximize expected payments, the tax delinquents who received a phone call reminder were selected following a tax repayment risk-analysis that highlighted those most likely to pay.

4.3 Data

Data Description

The confidential, individual-level tax information this paper uses was made available by the TCU in de-identified form. The dataset includes the 2013 cross-section of personal income tax delinquents in Greece as well as tax arrears from earlier years. Information includes the taxpayer's unique ID, e-mail or phone call treatment indicators, liability levels and end-of-year payments thereof. Tax arrears represent levels of outstanding debt from tax delinquencies of years prior to 2013. Demographic information includes the legal entity of the tax delinquent (whether it is an individual or a firm). For individuals, there is information regarding their sex, date of birth, possible date of death, zip code, whether the individual is currently employed, and occupation. Summary statistics by treatment status and tax debt vintage are given in Tables D1 and D2 in Appendix D.

To accommodate liquidity constraints, the tax agency collects income taxes in quarterly installments. This makes possible for the same individual to appear as a new tax delinquent multiple times in the dataset over the course of 2013 (e.g., a personal income tax delinquent not paying her monthly installment in quarter one, servicing both current and outstanding liabilities in the second quarter, failing to do so in the third). A time delinquency identifier allows me to create a panel of individual observations over installment periods. This permits testing whether the effect of the reminder is transitory or permanent, within the limited span of a calendar year.

Experimental Variation

The base treatment arm of the analysis is the combined compliance reminder; i.e., taxpayer overdue notice via phone call conditional on earlier e-mail notice. The control group is comprised of taxpayers who only received an e-mail notice.[9] In other words, I examine the incremental effect of a phone reminder relative to an e-mail one. Image E1 in Appendix E provides a translation of the message that was communicated. Clearly, the treatment should be understood as a neutral, non-deterrent behavioral nudge that corrects informational asymmetries about the payment deadline and sets high expectations about tax debt collection enforcement in a more salient manner (Thaler and Sunstein, 2012).

[9]The choice of the e-mail reminder recipients as opposed to a pure control group of tax delinquents who did not receive any notice at all is driven by three reasons. First, doing so the analysis is in line with the recent practice in behavioral tax compliance field experiments. For example, Hallsworth et al. (2014) uses taxpayers who received a letter compliance reminder as their control group, introducing variation through additional behavioral nudges. Second, e-mail compliance reminders appear to be the default option from the Greek tax authority's perspective.[10] Tax administrators are typically bound by law to provide notice to tax delinquents before applying penalties and levies. Thus, a setting where the control group does receive some type of notice is, generally, more policy relevant. Third, the fact that the tax authority failed to provide any notice to some individuals indicates limited third-party information on them or taxpayer negligence to register with the integrated tax system. To avoid using information from a potentially outlier population, I use e-mail compliance reminder recipients as the control group.

Exploratory Data Analysis of Phone Call Campaign

First, I present evidence of the separation in covariate balance between the income tax delinquents who received a phone call and those who only received an e-mail. Figure 3 illustrates histograms of the propensity score for the treated (phone call and e-mail reminder recipients) and untreated population (e-mail reminder recipients). The mass of untreated individuals have a very low probability of receiving a phone call. This highlights the extensive selection on these observables in the design of the phone call campaign. It also questions the validity of the matching design since the untreated population provides individuals who are only comparable with a sliver of the treated population; i.e., only those with a very low probability of receiving a phone call reminder.

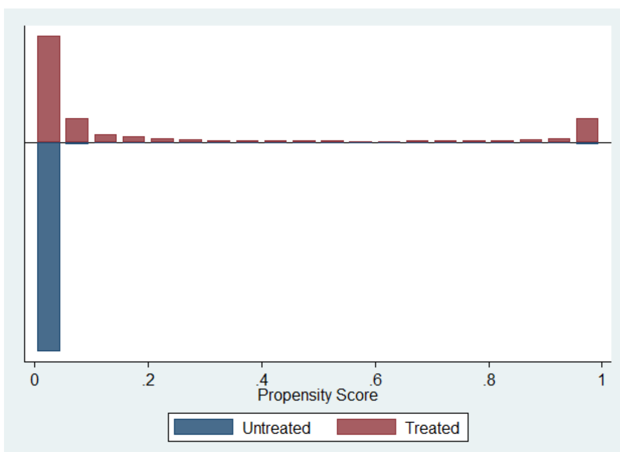

Figure 3: Histograms of Propensity Scores for Phone Call and non-Phone Call Reminder Recipients

Reported income tax delinquencies were naturally bounded from below at €50 by the TCU. Figure 4 below presents a histogram of personal income tax liabilities. For purposes of exposition, the figure is scaled to liability levels up to €2,000 to highlight the discontinuity in the treatment assignment decision rule at €500. Each bin represents a €10 range within the support. Evidently, there is very strong adherence to the treatment assignment rule. Only 0.004% with total income tax liabilities under €500 received a reminder phone call. This figure provides preliminary graphical evidence that the TCU called income tax delinquents on the basis of whether their liability exceed €500 or not. The illustrations of the raw data clearly indicate that the RDD is far superior to the matching design. This is part of the reason why the RDD results serve as the baseline of the paper.

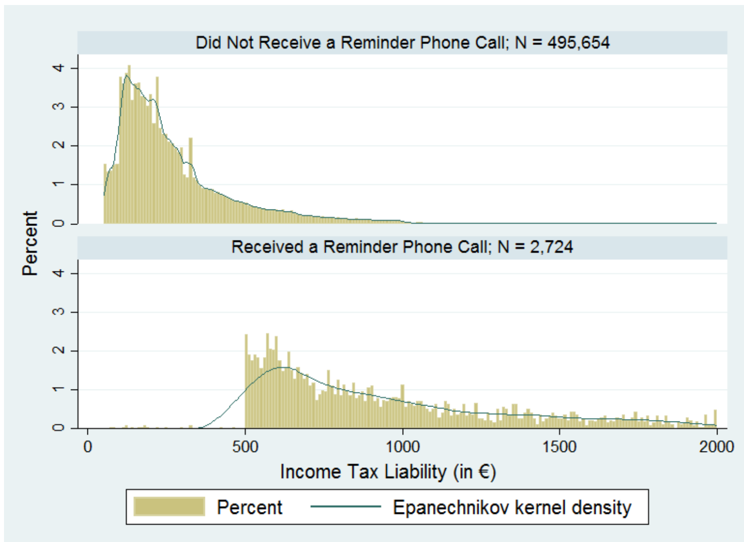

Figure 4: Histograms of Forcing Variable (Total Personal Income Tax Liabilities) for Phone Call and non-Phone Call Reminder Recipients

4.4 Identification

I begin with a simple but likely naive approach that compares all phone call reminder recipients to all non-recipients conditional and unconditional on observable characteristics. To disentangle the causal effect of the reminders I employ two quasi-experimental methods; matching methods and a regression discontinuity design (RDD). The TCU arbitrarily contacted only those with tax liabilities greater or equal to €500; but not in their entirety, which gives rise to a natural experiment. First, in the presence of crossovers, I use matching to recover the global effect over the entire distribution of tax liabilities. Then, I exploit the discontinuity in the policy rule that determines treatment assignment at the cut-off. Using a fuzzy regression discontinuity design, I identify the local average treatment effect (LATE) of the compliance reminders on payments of overdue tax liabilities. Section B in the Appendix provides more details on the each one of the identification strategies employed.

For each tax delinquent i, the observed outcome is $Y_i = T_i Y_i(1) + (1 - T_i) Y_i(0)$. However, only one of these potential outcomes can be observed as each respondent will receive either treatment or control, not both; "the fundamental problem of causal inference" (Holland, 1986). Using potential outcomes notation, the treatment effect for respondent i, $TE_i = Y_i(1) - Y_i(0)$, is unobserved.

Regression Adjustment

As a first, naive estimate of the phone call reminders' treatment effect, I calculate the raw (unadjusted and adjusted) difference in the level of payments between treatment and control group means. As mentioned in 4.2, the recipients of the phone call reminder were selected by the tax authority as the more likely to pay. Specifically, they were selected on the basis of observable characteristics most associated with tax compliance such as age, gender, employment status, and the level and vintage of the outstanding liability. The non-random assignment of the reminder introduces upwards bias to the estimated treatment effect. As a result, OLS estimates of the parameter of interest will be potentially overestimated. Despite its shortcomings, this calculation is still interesting from a policy perspective because it shows the implications of the phone call campaign on tax revenues. To assert its effectiveness, the tax authority would first like to know the net, aggregate effect of the phone call campaign on total tax payments. By partitioning the distribution of tax liabilities in increments, the tax authority can also examine heterogeneity in taxpayer responses by the size of their tax liability. This could help the design of future phone call campaigns by optimizing the targeting of tax delinquents by the size of their tax liability.

Coarsened Exact Matching

The fuzzy RD method has excellent internal validity but its identifying region is only local, around the cutoff. To recover the global effect of the reminders over the entire distribution of liabilities (not just within a neighborhood of the €500 cutoff), I rely on matching methods. The TCU reminded individuals with liabilities greater than €500 but not in their entirety. The existence of cross-overs in the study population - individuals with liabilities greater than €500 that were not treated - provides a very large number of untreated observations. This allows me to construct control groups that are comparable to the treated ones using matching. Matching is not an estimation technique but a pre-processing procedure. I use matching to mimic the TCU's assignment of treatment through the ex-post construction of a control group. The generated weights offer a post-matching sample that is less "model-dependent", has lower bias and more efficiency.

For identification, I am assuming selection on observables. That is, selection of tax delinquents into the treatment population was solely based on the observable demographic and socio-economic characteristics available used for pre-processing. This is a plausible assumption because the TCU's selection of phone call recipients was guided by a tax repayment risk analysis informed by the same observable information available to me. This allows me to give a causal interpretation of the estimated treatment effects for the entire distribution of liabilities as well, not just around the €500 cutoff.

I choose coarsened exact matching (CEM) as the baseline matching method of the analysis because it provides post-matching samples that are better balanced and estimates of the causal quantity of interest that have lower root mean square error than alternative matching techniques (Iacus et al., 2011). CEM assigns weights w_i to matched and unmatched observations that eliminate all differences (covariate imbalancess) across treatment and control groups up to the level of coarsening. It is important to note that CEM prunes both treated and control respondents. The computed weights w_i lead to a post-matching subsample that exhibits superior covariate balance (Tables D3 through D5 in Appendix D). This process changes the quantity of interest under study to the treatment effect in the post matching subsample.

Regression Discontinuity Design

To isolate causal effects, the RDD differentiates the non-linear and non-continuous indicator function $I(X_i \geq X_0)$ with $x_0 = 500$ from the smooth function $g(X_i)$ that determines treatment. I exploit the discontinuity at $X_i = X_0$ in the probability of treatment T_i conditional on liability level X_i to instrument treatment assignment D_i:

$$P[D_i = 1|X_i] = \begin{cases} g_1(X_i), & \text{if } x_i \geq x_0. \\ g_0(X_i), & \text{if } x_i < x_0. \end{cases} \tag{4}$$

where $g_0(X_i) \neq g_1(X_i)$. The relationship between the probability of treatment T_i and level of total liability X_i can be expressed as:

$$P[D_i = 1|X_i] = g_0(X_i) + [g_1(X_i) - g_0(X_i)] \cdot T_i \tag{5}$$

For identification I use the same assumptions as the binary instrumental variables design. Inference comes solely from the complier population. These are the observations for which treatment status changes as I change the value of x_i from just to the left of x_0 to just to the right of x_0.

4.5 Estimation

Regression Adjustment

I regress payments of total liabilities on a constant, the treatment indicator and controls using the full sample. The latter include age and its quadratic, an indicator for male gender, the level of the outstanding tax liability and its quadratic, an indicator of whether the tax delinquent has outstanding tax arrears("former debtor"), indicators of salaried, self-employed and unemployed status, and the number of delinquent installments in 2013. I also include quadratics of age and level of tax liability and interactions of treatment with the former debtor indicator and age with the male indicator. I estimate using OLS and clustering standard errors at the tax delinquent's zip code.

Coarsened Exact Matching

Applying the CEM weights, I regress payments of total liabilities on a constant and the treatment variable. If the covariate balance of the post-matching sample is sufficiently high then the coefficient of the treatment variable shows the causal effect of the compliance reminder on payments of overdue tax liabilities. CEM permits for a direct assessment of the covariate imbalance via the L_1 statistic. The estimates are based on OLS and standard errors are clustered by the tax delinquent's zip code.

Additionally, I exploit the quarterly and monthly nature of income tax payment structure to estimate the treatment effect on the likelihood of becoming delinquent again. I follow recipients of a phone call reminder in the first quarter of 2013 for the remainder of the year conditional on an e-mail reminder. Then, I compare their probability of re-appearing in the delinquency dataset relative to recipients of an e-mail only.

Regression Discontinuity Design

Following Imbens and Kalyanaraman (2009), I employ a non-parametric fuzzy regression discontinuity estimator that applies local linear regression on both sides of the €500 cutoff, using a triangular kernel. This is equivalent to a local linear IV model where I use T_i to instrument for D_i and include its interaction with the first order polynomial of X_i. The fuzzy RD estimand takes the form of a ratio of two sharp RD estimands: one for the main reduced-form outcome equation (that is, the regression of $g(Y_i)$ on X_i) and another for the first-stage equation (that is, the regression of T_i on X_i).

$$\rho = \frac{\lim_{c\to 500^-} E[Y|X=c] - \lim_{c\to 500^+} E[Y|X=c]}{\lim_{c\to 500^-} P[D=1|X=c] - \lim_{c\to 500^+} P[D=1|X=c]} \tag{6}$$

Also, I examine the robustness of my findings to the inclusion of a second-order polynomial (quadratic form). Following Gelman and Imbens (2014), I do not use any higher-order polynomials because of the unsatisfactorily large values the implicit weights may take, the sensitivity of results to the order of the polynomial approximation, and the misleading confidence intervals that can be obtained. The noisiness of the implicit weights can lead to global polynomial regressions that have both poor coverage and wide confidence intervals. Alternatively, in IV format, the first-stage is:

$$D_i = \gamma_{00} + \gamma_{01} x_i + \gamma_0^* T_i + \gamma_1^* x_i T_i + \xi_{1i} \tag{7}$$

The second-stage is:

$$Y_i = \alpha + \beta_{01}x_i + \rho D_i + \beta_1^* D_i x_i + \xi_{2i} \tag{8}$$

I re-center at €500 such that $z_i = x_i - 500$ and estimate via 2SLS. I re-weight observations within the bandwidth h_{IK}, putting more emphasis (larger weights pw_i) on those closer to the cutoff as denoted by the distance z_i. I transform the outcome variable using the IHS transformation $g(\cdot)$ because it leads to a substantial efficiency gains in estimation of the local linear regression model without the expense of bias. As the average value of Y_i is sufficiently larger than zero, the coefficient ρ could be interpreted in approximately per cent terms. The second stage becomes:

$$g(Y_i) = \alpha + \beta_{01}x_i + \rho D_i + \beta_1^* D_i x_i + \xi_{2i} \tag{9}$$

5 Results

This section discusses the paper's results.

5.1 Baseline Results

Regression Adjustment

Overall, the phone call campaign was successful in that those targeted did bring in more tax revenues to the tax authority. The raw, unadjusted difference in group means is €1199.81, on average, or approximately €6.279mil. in total. Adjusting with controls, the 5,234 income tax delinquents who received phone call reminders (following an e-mail reminder) paid an additional €310.68, on average, or €1.626mil., in total income tax revenues. Partitioning the sample into three classes by debt size, I find substantially heterogeneous treatment effects. Whereas the smaller- and medium-sized debt holders responded to the phone call reminders by paying more, no discernible difference in payment behavior by treatment can be reported for those with larger debts. First, I consider income tax delinquencies less than €10,000 which corresponds to approximately 99% of the sample or 477,240 income tax delinquents. A total of 4,316 tax delinquents received both an e-mail and phone call reminder as opposed to just an e-mail reminder. The unadjusted and adjusted average treatment effects of the phone call reminders are €588.72 and €178.18, respectively, which translate to €2.54mil. and €0.769mil. in additional total income tax revenues.

Figure 5 below illustrates the raw differences in payment levels of tax liabilities up to €10,000 per €250 tax liability increment. As expected, the recipients of the phone call reminders paid more income tax liabilities than those who just received the e-mail reminder. In fact, their payment levels maintain a robust difference of approximately €180 relative to e-mail recipients over the entire range of income tax liabilities considered. However, this merely may be evidence of the underlying selection of the phone call recipients on the basis of a higher likelihood of repayment. Next, the second class is comprised of 5,050 income tax delinquents with debt levels between

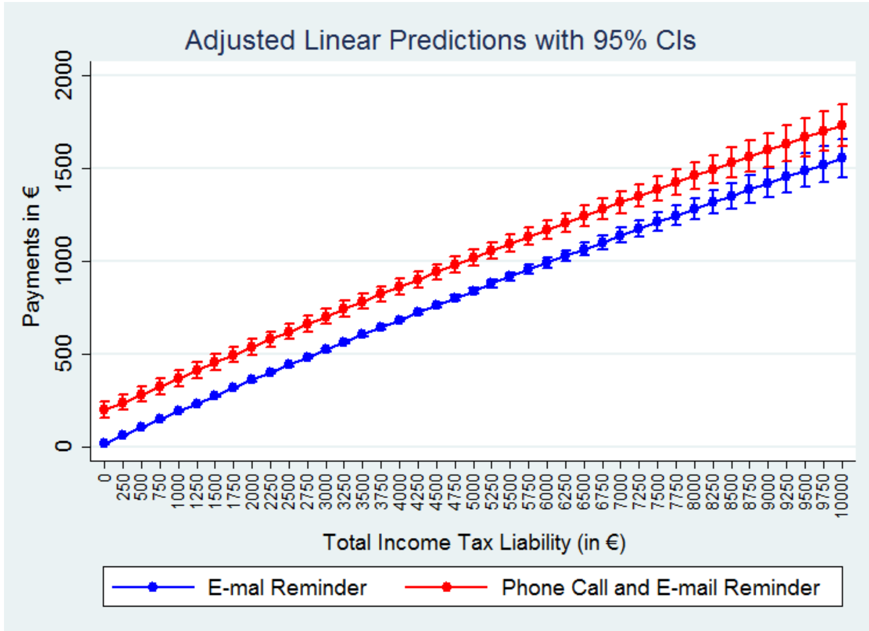

Figure 5: Naive Estimate of Phone Call Reminder Effect on Income Tax Payments

€10,000 and €100,000 of whom 880 received a phone call reminder. The unadjusted and adjusted average treatment effects of the phone call reminders are €1046.69 and €982.04, respectively, which translate to €0.921mil. and €0.864mil. in additional total income tax revenues. Finally, the largest debtor class ((€100000,€1186515]) includes just 97 income tax delinquents who received an e-mail reminder of whom 38 also received a phone call reminder. In this case, the raw average treatment effect of the phone call reminders is €1383.4 but is very imprecisely estimated. When I adjust the regression estimates with controls, the coefficient turns negative €4500.41 but remains imprecisely estimated. As a result, the raw gains in tax revenues are equal to €0.052mil but the adjusted estimate points towards tax losses of €0.171mil.

Coarsened Exact Matching

Pre-processing the data using CEM leads to small decreases in covariate imbalance. A detailed account of balance across treatment and control for all covariates specified is given in Tables D3 through D5 in Appendix D. A comparison of the L_1 statistic of the pre-matching sample in row 5 to the one of the post-matching sample in rows 6 of Table D6 in Appendix D transparently shows the large margins of improvement across all PIT liability ranges due to the CEM. Still, the post-matching covariate imbalance is not entirely eliminated as anticipated by Figure 3.

Table D6 in Appendix D reports the CEM-weighted linear regression estimates of the effect of incremental informational nudging on tax payments on personal income tax delinquents. A small, positive but very significant effect is recovered. This is in disagreement with the fuzzy RD estimate. Yet the statistical significance is smaller over the debt range that coincides with the bandwidth, somewhat resembling the earlier point estimates. Furthermore, the CEM estimates indicate that informational nudging has an increasingly weaker behavioral impact on medium-sized and large debtors.

Figure C3 in Appendix C shows the impact of the combined treatment arm on the probability of becoming income tax delinquent again. The prediction for the treated individuals is 7.2% lower and very significant. This test provides suggestive evidence that the effect of information provision was not ephemeral. An explanation for this result could be increased awareness about the payment deadlines, overall. It has been argued that taxpayers may become delinquent on their tax liabilities due to procrastination (Hallsworth et al., 2014). Thus, even if the compliance reminders may have not improved awareness about the financial penalty, it appears that personal income tax delinquents became more perceptive of the payment schedule.

Regression Discontinuity Design

Figure 6 illustrates the extensive margin of the estimated effect, focusing on whether any positive payment was received. There is a substantial (5.1%-7.7%) reduction in the repayment rate of total PIT delinquencies (including tax arrears). This finding is significant at conventional levels and robust to the polynomial order used.

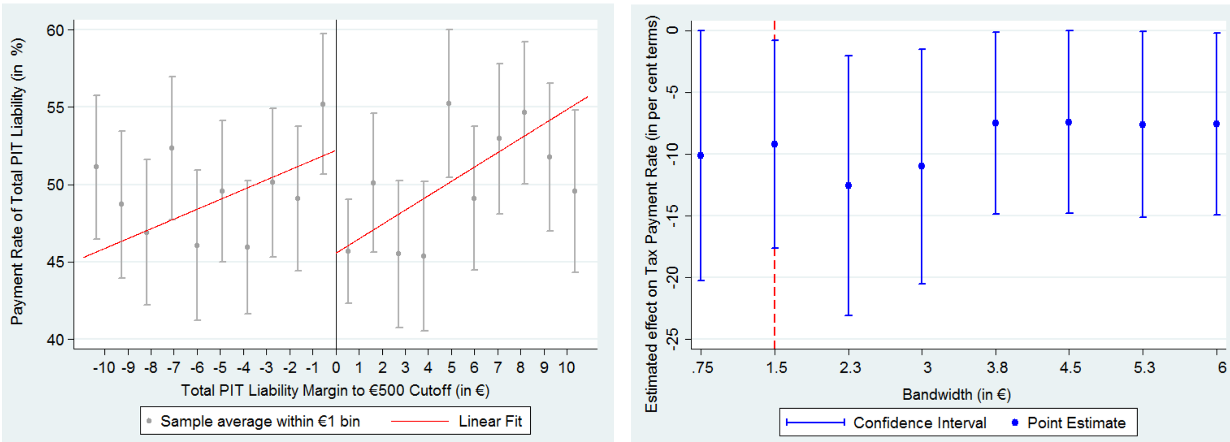

Figure 6: Incremental Effect of Phone Calls relative E-mails on Repayment Rate

Figure 7 illustrates the intensive margin of the incremental effect of a neutral information nudge via phone call on tax compliance. That is, the estimate of interest is the impact of an additional, phone call compliance reminder given an e-mail reminder on the level of tax payments. It directly compares the benefits from a treatment arm which is more salient but also more costly to the tax agency's baseline. The left panel of Figure 8 presents the LATE at the optimal bandwidth whereas the right panel shows the LATEs over a range of different bandwidths.

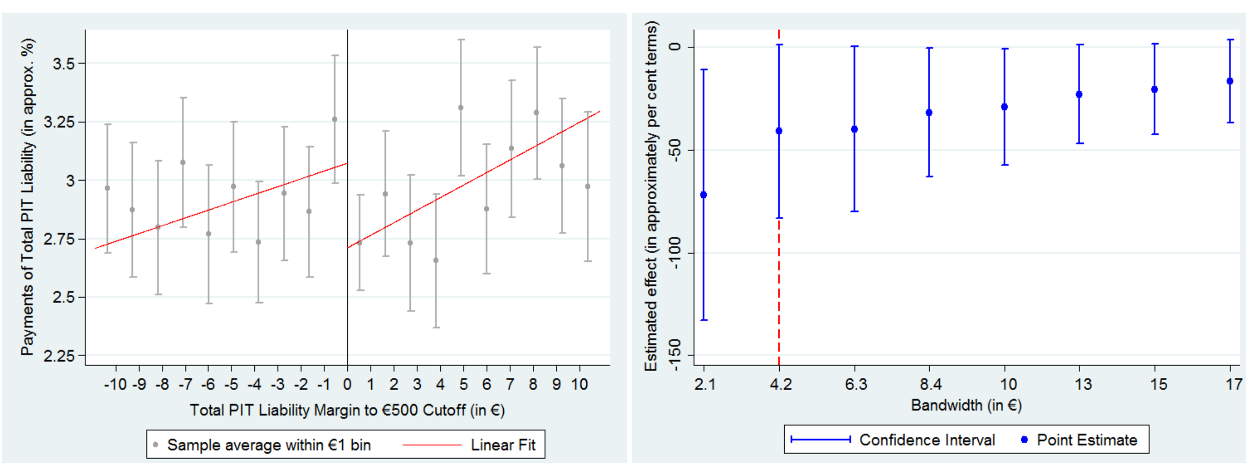

Figure 7: Incremental Effect of Phone Calls relative E-mails on Tax Payments

There are two main take-aways from Figure 8. First, one can reject with certainty a causal positive effect of compliance reminders on tax payments. Second, at the optimal bandwidth, there is a large and statistically significant decrease in tax payments among those who receive an additional phone call compliance reminder following an e-mail reminder. This finding is also given in tabular form by the 2SLS estimates in Table D7 in Appendix D. The evidence suggests that "soft" phone call reminders lead to lower tax compliance. Receiving a phone call reminder relative to just an e-mail reduces tax payments by $41.5\% - 39.8\%$ depending on whether a first or a second polynomial order term for the forcing variable is used. This finding is statistically significant at the 10% level for the outcome transformed by the IHS but not for payments in levels. The width of the confidence intervals should be attributed to the low power of the local regression analysis.

5.2 Heterogeneity Analysis by Tax Debt Vintage

In this subsection, I investigate why do the compliance reminders backfire in the sense that they causally reduce tax payments. A potential explanation is that PIT delinquents disregard the tax authority's nudging as non-credible. In this case, the tax delinquent's payment decision may be driven by her inferred probability of perfect tax debt collection enforcement. Taxpayers with former experience as tax delinquents may respond differently to a call for payments to those with no experiences for a number of reasons. For instance, they may treat the tax lien warning reminder as an inconsequential payment threat if they experienced weak tax debt collection enforcement in the previous fiscal year. In an attempt to go beyond speculation, I explore the heterogeneity of the estimated LATEs by debt vintage as a proxy of PIT delinquents' tax debt collection enforcement experience. That is, I examine whether a compliance reminder led to heterogeneous taxpayer responses between first-time tax delinquents ("new debtors") and those with tax arrears from previous years ("old debtors"). In this setting, the latter were likely exposed to a similar 2012 compliance reminder that sent via e-mail to tax delinquents (2012 treatment status is unobservable in the present analysis).[11] Even if only a fraction of the 2012 delinquent population was exposed to the earlier compliance reminder, all tax delinquents with tax arrears in 2012 have been exposed to the intensity of tax debt collection enforcement prior to 2013. If old debtors do not hold onto their former priors but exhibit adaptive expectations about the probability of perfect tax debt collection enforcement, then they should revise their expectations downwards given a mere 10% enforcement of tax liens in 2012 (IMF, 2013). I test whether the payment decision of old and new debtors is driven by tax debt collection enforcement rather than information nudging by splitting the sample and estimating the effect of the compliance reminders separately. To counteract the inevitable loss of power due to sample-splitting, I use substantially wider bandwidths and introduce covariates to capture the increased variability.

A tabular representation of the 2SLS estimates of the local linear effects of the compliance reminders by debt vintage as a proxy of tax debt collection enforcement is given in Table D8 in Appendix D. The behavioral response of new debtors is small as their payment behavior does not change in a significant way. On the contrary, old debtors do react to the compliance reminders by paying significantly less. A combined reading of the two implies that the baseline negative result across the entire population was driven by old debtors' payment response. The latter were likely recipients of the 2012 e-mail compliance reminders regarding their 2013 tax arrears and experienced only minimal collection enforcement. Thus, it is appears that the taxpayer responsive is driven by tax debt collection enforcement rather than information nudging. Generally, tax compliance

[11] The tax agency is obliged by law to notify tax delinquents during the initial, one-month grace period about the interest rate to be accrued prior to the final payment deadline.

heavily relies on exaggeration of the audit probability by tax authorities and the over-assessment of their perceived audit probability. This baseline is the likely information set under which new debtors decide whether to comply or not. However, as the tax authority did not take any further action (e.g. no levies issued), old debtors learned that, despite their tax delinquency, no further action would be likely. Thus, the tax authority's "soft" approach in tackling tax delinquency may have bred the perception that it is inconsequential. As a result, taxpayers who have acquired that knowledge from prior interaction with the tax authority are not affected by sanction threats as they have a corrected perception of the true ramifications of non-compliance. On the contrary, new debtors are more easily intimidated by the tax authority and, thus, increase payments because they still operate under an environment of exaggerated audit probabilities. In sum, when information nudging is framed as an imperative duty by an untrustworthy institution then there is a potential it may backfire.

5.3 Placebo Test of Tax Debt Collection Enforcement Experience

If taxpayer behavior is indeed driven by perceptions about tax debt collection enforcement rather than information nudging then, in the absence of the latter, one should expect the former to maintain its influence over the tax payment decision. I test this hypothesis using the pure control group of tax delinquents who do not receive any notice from the tax authority at all but vary by debt vintage. Specifically, I estimate the LATE of being an old debtor on tax payments conditional on observable characteristics such as tax debt size, age of the delinquent and their quadratics as well as employment status. Table D9 in Appendix D shows the results of the 2SLS estimation. As one would expect this tax compliance is influenced by tax debt collection enforcement, being an old debtor results to a reduction in tax payments. I find small reductions in payments for the subset of old debtors relative to new debtors. This finding is statistically significant only for the estimation in levels in column (1). The influence of former experiences of tax debt collection enforcement that was extremely weak (11% collection rate in 2012) on tax payments is negative. The evidence from the specification augmented by taxpayer characteristics does not hint towards any other potential explanation for this finding. Again, it appears that weak tax debt collection enforcement is instrumental for the tax payment decision.

6 Robustness: Validity of Regression-Discontinuity Design

This section includes a number of robustness checks surrounding the validity of the RDD. I follow the recommended "checklist" for implementation of RDDs by Lee and Lemieuxa (2010). The following subsections address every potential issue that relates to the analysis, presentation, and estimation of the RDD in succession.

6.1 Manipulation of Forcing Variable: McCrary's Density Test

The fundamental condition for identification in fuzzy RD is continuity of the conditional expectation of counterfactual outcomes in the forcing variable. McCrary (2008) notes that these continuity assumptions may be invalid if individuals can manipulate the reported value of the forcing variable; i.e., the level of PIT liability. To this end, I test the discontinuity in the density at the €500 cutoff (left panel of Figure 8). Graphically, I look for evidence of one-sided manipulation of total liabilities. A smooth density function would suggest that there was no manipulation of the forcing variable. However, the evidence suggests that this is not the case. There appears to be a jump immediately after the cutoff. An excess mass of PIT liabilities at €500 leads to the observed jump in the density estimates. At first blush, it seems that there is evidence of manipulation of the reported tax liability on behalf of the tax delinquents to receive a phone call compliance reminder. I argue that the jump in the density of the forcing variable at the cutoff is not evidence of manipulation but simply rounding. For example, consider the right panel of Figure 8: even over regions that are irrelevant to the treatment assignment rule, there is a sharp increase in the density estimates of just above €300 and €400 as well.

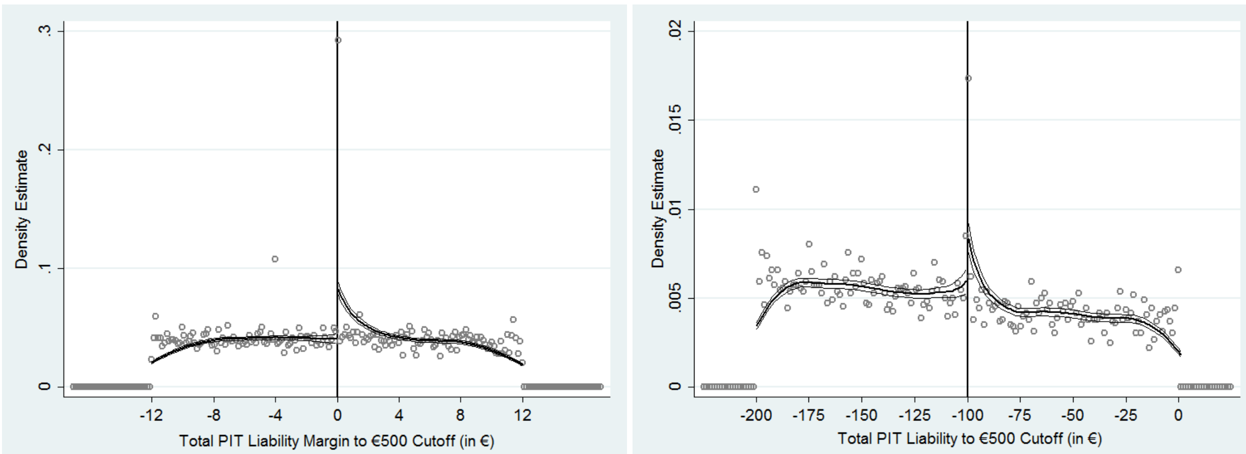

Figure 8: McCrary's Density Test of Manipulation of Forcing Variable - Full Sample

Kleven and Waseem (2013) similarly encounter the potential of rounding in income reporting in Pakistan and particularly for the case of self-employed taxpayers. Following their approach in addressing rounding, I split the sample by those whose tax liability is reported in hundred €increments ("rounders") and those who do not ("non-rounders").[12] Figure 9 presents the equivalent McCrary's density tests to Figure 8 ones' for the subsample of non-rounders. The resulting estimates of the density of PIT liabilities indicate a smooth underlying distribution in both panels. Now there is no evidence of bunching of PIT liabilities above the cut-off. If anything, there is minimal bunching at €25 increments now over the $[300, 500)$ region; again, a by-product of a natural

[12]Alternatively, Kleven and Waseem (2013) also specify fixed effects at multiples of income increments where rounding might occur. But, this alternative approach is more applicable in their context that focuses on the estimation of the global empirical distribution of the outcome variable rather than locally examining the smoothness of the forcing variable.

tendency towards rounding.

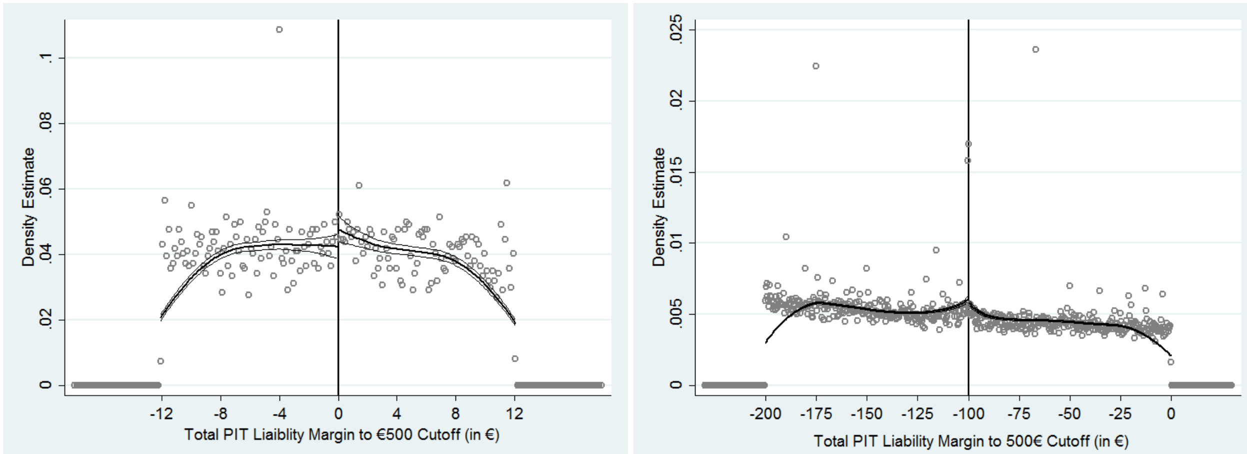

Figure 9: McCrary's Density Test of Manipulation of Forcing Variable - Non-rounders Sample

In sum, the results of the successive McCrary's density tests of manipulation of the forcing variable do not suggest that it would be a mistake to view the treatment assignment rule as quasi-random.

6.2 Inspection of Baseline Covariates: Parallel RDD Test

An indirect approach for testing the validity of the RDD is to assess to what extent the observed baseline covariates are "locally" balanced on either side of the threshold, such that treatment assignment can be considered as locally randomized. This test is first conducted graphically where the outcome variable is replaced with observed baseline covariates. Therefore, one should expect a smooth, continuous distribution of covariate averages across the cut-off. A discontinuity would violate the identifying assumption that predicts local random assignment. Figure 10 below presents combined evidence from the parallel RDD test. The x-axis presents the binned averages of the forcing variable in €1 increments over the optimal bandwidth of €10.90. The y-axis contains the pseudo-dependent variable. This exercise is a placebo test in that no substantial discontinuity to the left and to the right of the cut-off should be apparent in the binned averages of the pseudo-outcomes. The latter are comprised of covariates that were determined prior to the realization of the phone-call campaign such as sex, age, and employment status (indicators for salaried, self-employed, and unemployed PIT delinquent individuals). I also examine the continuity of old debtor status around the cut-off since tax arrears were formed before the phone-call campaign.

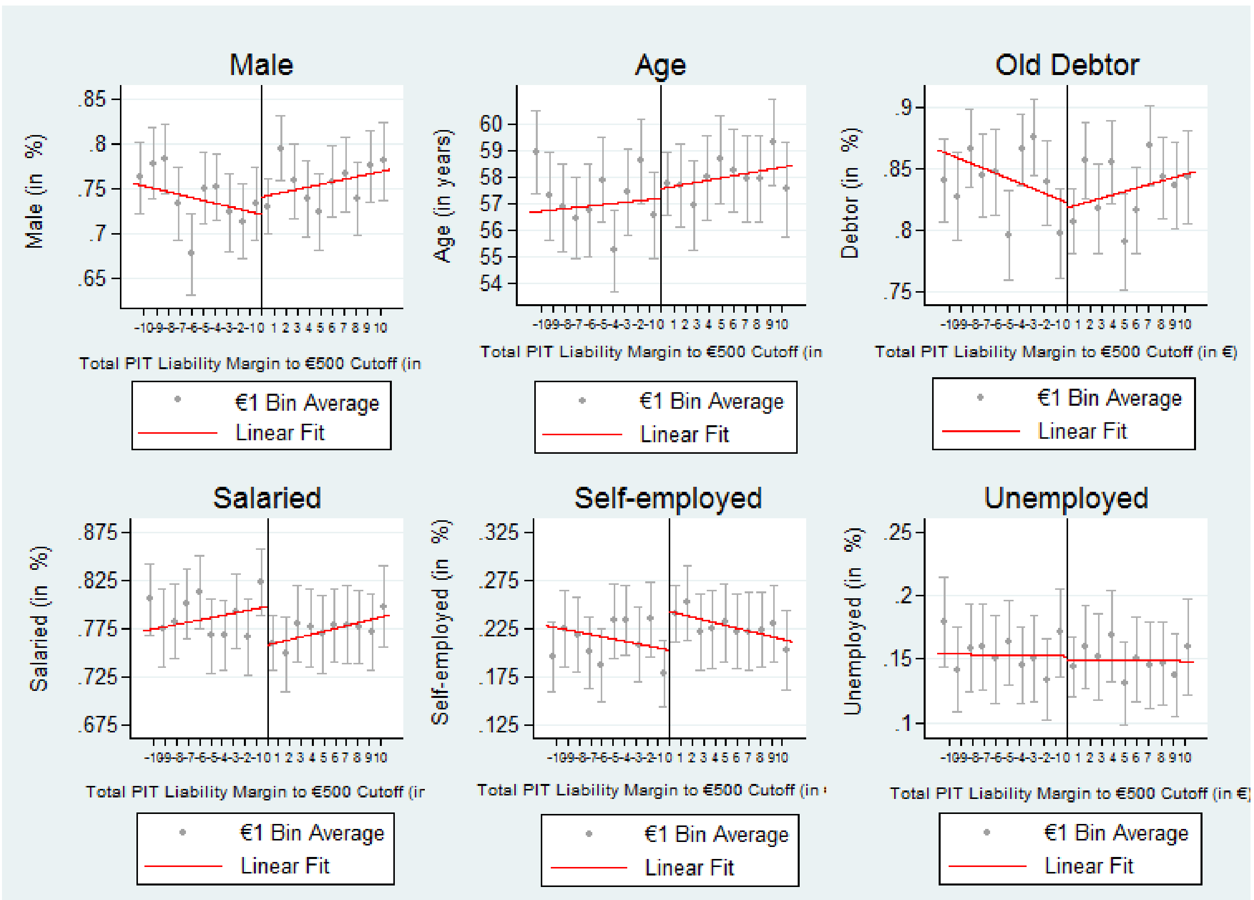

Figure 10: Inspection of Baseline Covariates: Parallel RDD Test

The distribution of the €1 binned averages over the optimal bandwidth is smooth for four out of a total of six baseline covariates. There appears to be a small, 5% difference in the averages of the fraction of self-employed and salaried individuals around the cut-off. However, the large confidence intervals around the point estimates at every €1 bin average imply that no statistically discernible differences below and above the cut-off are present.

The graphical evidence of no discontinuity in baseline covariate around the cut-off above is complemented by a formal computational test of covariate balance. Specifically, I am t-tests to assess whether average covariate values below and above the cut-off are statistically different within the optimal bandwidth. If the p-value's is less than 0.1, 0.05, or 0.01 then the mean difference between the baseline covariate's observations just below and just above the cut-off is statistically significantly different from zero. The results are shown in Table D10 in Appendix D. Around the cut-off, the sample of tax delinquents is well-balanced with regards to male, the old debtor, and the unemployed baseline covariates within the optimal bandwidth. However, there is a statistically discernible difference in the fraction of the tax delinquent population around the cut-off with regards in age, salaried and self-employed employment status. Nevertheless, these differences are essentially indistinguishable from a policy perspective. For instance, the largest discrepancy in means is for self-employed employment status from 21.1% for those below as opposed to 22.5% for those above the cutoff. Thus, the inspection of discontinuity in baseline covariates shows that the validity of the RDD based on the arbitrary €500 cut-off is valid. Locally, treatment assignment is as good as random.

6.3 Strength of €500 Cut-off as an Instrument

Another reason of caution about the fuzzy RD estimates could stem if the €500 cutoff was not a strong instrument of treatment assignment. However, the plots of the first stage for PIT delinquent individuals confirm that the cutoff is a very strong instrument of treatment assignment. This could have also been predicted by the high adherence to the treatment assignment rule as demonstrated by the histogram in Figure 4 in the exploratory data analysis. Figure C2 in Appendix C show an arguably flat line at zero probability for income tax delinquents with liabilities under €500 and a marginally sloped line for those with liabilities just above the cutoff. One could interpret the flatness of the probability estimate after €500 as evidence that the cutoff is not a strong instrument. However, the jump in the probability of receiving treatment after the cutoff is simply masked by the very large number of untreated observations. Even if the jump is not stark due to the existence of cross-overs, there is still very high adherence to the rule with virtually no phone call recipients with total tax liabilities less than €500.

7 Conclusions

This study examines the effect of information nudging on payments of overdue tax liabilities in Greece. Overall, the campaign targeting personal income tax delinquents who are more likely to pay did increase tax revenues. This findings echo the recent policy change by the TCU to adopt e-mail only informational nudging. However, I find that information nudges via phone calls may not causally enhance tax compliance in a context of low tax morale. In fact, unintended consequences are possible as the reduced tax payments for tax delinquents at the €500 cutoff.

The results in this paper are largely consistent with other studies in the literature showing that simple contact from the tax agency can have a substantial impact on taxpayer behavior, at any direction. Guyton et al. (2016) find that the mere receipt of a neutral information reminder can induce taxpayer responses, at least in the short run. It is important to note that this study similarly finds that even simple information provision to a taxpayer can induce behavioral responses that generate both gains and losses to the tax agency. Specifically, they find that the treatment individuals were more likely both to claim refunds and pay taxes owed. Hallsworth et al. (2014) uses a similar neutral, non-deterrent informational nudge as the control group letter to just inform personal income tax delinquents about the deadline and payment methods. They report a small but positive and significant effect on tax payments. Furthermore, Bott et al. (2014) used a base treatment of neutral information provision which closely resembles the one examined in this study. Specifically, their base treatment involved a mailed letter containing information about where and when to submit the tax returns and why foreign income is subject to income taxation. Following the receipt of the base treatment, they report substantial taxpayer responses in foreign income reporting. My findings are reversed and remind that "soft" nudging on behalf of the tax agency will not necessarily enhance compliance. In this regard, my results are in agreement with Fellner et al. (2009) and Ariel (2012) who report that potential individual evaders of TV license fees in Austria and firms in Israel responded to communication by the tax agency by registering to pay license fees and reporting corporate income tax less. On the contrary, empirical evidence suggests that increased compliance among taxpayers with substantial prior-year audit assessments (Erard, 1992). Regarding the effectiveness of different delivery devices, Ortega and Scartascini (2015) find phone calls to be more effective than e-mails in motivating individuals to pay overdue tax liabilities. On the contrary, I find that the salience of the nudging technology does not necessarily guarantee better outcomes for the tax agency. On the contrary, factors such as prior experience in tax delinquent status are more important in shaping individual beliefs about debt collection enforcement. Weak prior debt collection enforcement renders information nudging non-credible which, in turn, leads to lower tax revenues.

Finally, it should be noted that the study, at its current form, is subject to an important limitation. The empirical analysis does not account for spatial factors. However, taxpayer location may affect her payment behavior either through the horizontal tax reciprocity channel given the disproportionately higher public spending in Athens (due to events such as the 2004 Olympics) relative to the rest of Greece or because of the inherently weaker tax debt collection enforcement in distant location. Currently, the analysis is only allowing the unobservables of tax delinquents to be correlated by clustering standard errors at the zip-code level. Introducing location fixed effects will be a first step in assessing the extent of geographical heterogeneity. However, including additional covariates in the outcome equation of the 2SLS model in uncommon in RDDs. Moving forward with this project, I will bring-in GIS data on (centroid) distance between the tax delinquent's zip code and the national (Athens) and regional capitals as well as the nearest tax office. This new information will allow me to go beyond simply controlling for time-invariant characteristics that are specific to the region or the zip-code. Complemented with information on driving times, it will permit a rigorous examination of the horizontal tax reciprocity channel. More importantly, it will provide with a meaningful measure of variation in tax debt collection enforcement. Tax delinquents situated in harder-to-reach locations from the tax authority's perspective (e.g. island without a tax office) are likely subject to weaker tax debt collection enforcement. Since reminders are the same independent of tax delinquent location, I will exploit the influence of Greek geography as identifying variation in tax debt collection enforcement.

Essay 2: Housing Bubbles and Zoning Corruption: Evidence from Greece and Spain

1. Introduction

The creation of the European Monetary Union (EMU) in January 1, 1999 and the circulation of the euro in January 1, 2002 led to significant changes in the financing systems of many EMU member countries. Cheap credit and a surge in intra-EMU capital flows with the perception of a dramatic decrease in Southern-Europe financial risk transformed the economies of some of the peripheral countries such as Greece and Spain. The aim of this paper is to examine whether the housing bubbles experienced by Greece and Spain in the years preceding the Great Recession of 2008 exacerbated corruption of zoning officials. The mechanism under study worked as follows. Financial markets' perception of risk convergence among EMU member states significantly reduced the cost of mortgage lending in its periphery. The availability of cheap credit in Greece and Spain sparked an increase in economic activity excessively focused on housing construction. Gopinath et al. (2015) develop a theoretical framework that shows how the decline in the real interest rate in the southern EMU member states, often attributed to the euro convergence process, led to a decline in sectoral total factor productivity because capital inflows were misallocated toward investments that were not necessarily the most productive. The housing bubbles that ensued created significant new profit opportunities for housing developers while zoning officials' corruption penalties remained unchanged. As a result, the real estate boom in Greece and Spain altered housing developers' incentives to offer illegal bribes for land rezoning and officials' incentives to accept them. Other than discussions in the popular media of the potential link between the housing bubbles and zoning corruption, to date there has not been any formal testing of this relationship.[1]

[1] For an example of a discussion of the potential link between housing bubbles and corruption in the media for the case of China see the "China Daily" of April 22nd 2014: http://www.chinadaily.com.cn/opinion/2010-10/08/content_11382933.htm. Bujko et al.

To better understand the underlying process it is important to highlight the institutional designs that mortgage markets in Greece and Spain followed. Unlike the US experience, neither government acted as a significant mortgage insurer (like the Federal Housing Administration in the US) nor did they sponsor institutions that did so (like Fannie Mae or Freddie Mac also in the US) over the period of analysis. In fact, the Greek government guaranteed only 4 percent of standing mortgages and the Spanish government none at all. In addition, borrowers in Greece and Spain were fully exposed to the risks associated with servicing mortgage loans with draconian personal bankruptcy laws. Finally, Greek and Spanish households were able to continue to deduct primary-dwelling mortgage interest rate payments from individual income tax liability, as had been the case before the introduction of the euro. Thus, the variation in the mortgage interest rates that followed the introduction of the euro was the main –if not the only- mechanism affecting the decision of taking out a mortgage loan by Greek and Spanish households.

The story line we weave together in this paper starts with a dramatic drop in mortgage interest rates in Greece and Spain following EMU accession and their stabilization at historically low rates, thereafter. Figure 10 below shows the evolution of the benchmark mortgage interest rates on new housing loans in Greece and Spain from January 2000 through December 2005. This period corresponds to two years before and four after the adoption of the euro in 2002. We consider only the last two years leading to the EMU because the Spanish and, especially, the Greek economies were among the last to converge to EMU's inflationary criterion of 3 percent. For Greece, we present the trajectory of the interest rate on euro-denominated housing loans at a floating rate. It is a good benchmark of lending costs in Greece because the majority of euro-

(2015) uses corruption as an independent variable to explain "land grabbing" when here housing prices is the key independent variable that determines zoning corruption.

denominated housing loans from domestic credit institutions to households until December 2006 were arranged on a variable rate basis.[2] In Spain, mortgages typically featured a variable (or adjustable) rate determined by the lender. We depict the evolution of the most widely used benchmark in retail-banking in Spain; i.e., the one-year EURIBOR rate (Bank of Spain, 2013).

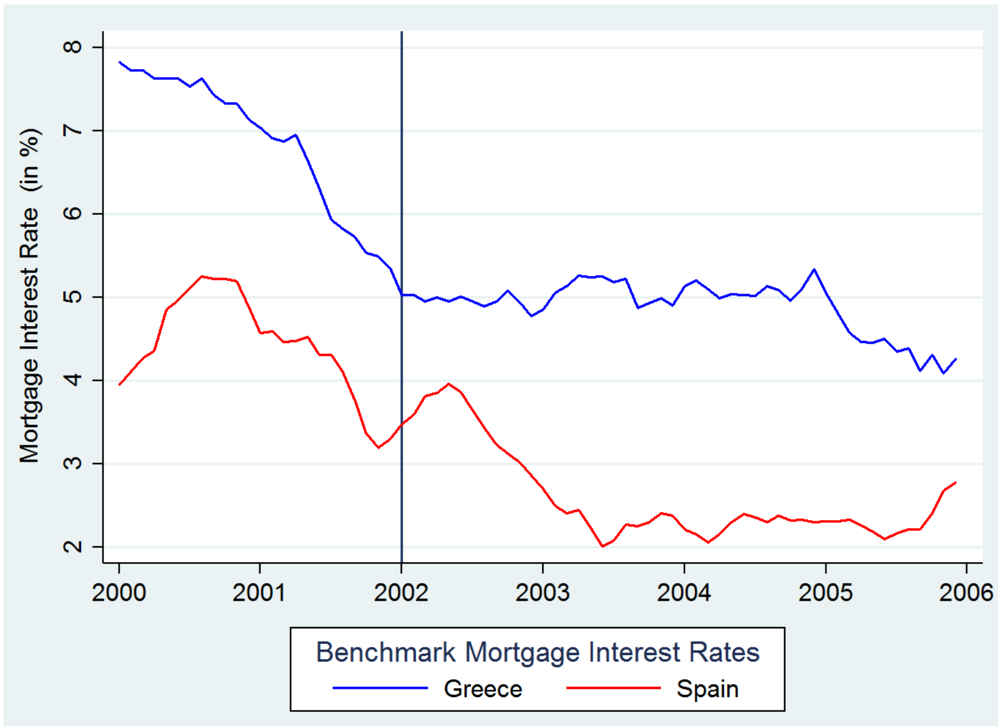

Figure 11: Mortgage Lending Benchmark Interest Rates, January 2000 – December 2005

Source: Bank of Greece, Bank of Spain

The average monthly rate in Greece and Spain from January 2000 through December 2001 were 6.87 percent and 4.43 percent, respectively. However, over the next two years these figures markedly declined to a 5.01 percent (a 27 percent decrease) and a 2.91 percent (a 34 percent decrease) average, respectively. These are large swings, especially considering that they represent two-year averages. Indeed, the Wilcoxon rank-sum test suggests that mortgage lending in Greece and Spain was significantly cheaper two years after accession in the EMU than before

[2] Housing loans in non-euro currencies represented 2.5 percent of the total and their rates are not presented.

with very high precision. Therefore, we take these as evidence that the housing markets in Greece and Spain experienced substantial decreases in mortgage lending that we exploit as identifying variation.

These substantial reductions in the cost of borrowing prompted many more households in Greece and Spain to enter the mortgage market. Between 1999 and 2007, the average annual growth rate of loans for house purchase in Greece was 30.3 percent (the highest among Eurozone-founding member states), and in Spain 19.8 percent (third highest) whereas the volume of mortgages in the EMU as a whole grew only by 10.4 percent (ECB Structural Issues Report, 2009).

The significant growth in housing finance and the associated strong increase in housing demand resulted in a rapid increase in construction. Housing starts per 100 dwellings in Greece grew from 1.6 in 1999 to 2.1 in 2004 and in Spain from 1.5 in 1999 to 1.6 in 2007. In contrast, average housing starts in the EMU remained stable at 1.1 dwellings during the same period (ECB Structural Issues Report, 2009).

There are many potential determinants of housing prices and causality between interest rates and housing prices cannot be established from observational inference. Low interest rates do not necessarily lead to housing price increases as the recent housing markets in the US, Greece, and Spain have shown but a link between monetary policy and housing prices has been firmly documented (Sutton, 2002; Tsatsaronis and Zhu, 2004; Holt, 2009). Several studies have identified a positive relationship between housing prices and the availability of credit (mortgage lending) in Greece (Himoniti-Terroviti, 2005; Brissimis and Vlassopoulos, 2009) and Spain (Gimeno y Martínez-Carrascal, 2006; Gentier, 2012). The unprecedented supply of cheap credit due to the adoption of the euro added impetus to the pre-existing inflationary trend of nominal housing prices in Greece and Spain that followed the general inflation. This paper uses the transition to the euro and the associated windfall in mortgage lending in Greece and Spain to exploit the ensuing housing bubble as an exogenous shock that increased zoning corruption.[3] This interpretation of the role played by the euro on housing prices in both countries is supported by the data. Figures 12 and 13 below plot the average national housing price per square meter in Greece and Spain from the adoption of the euro (January 2002) till the end of the period of our analysis (December 2008), respectively. We use the innovational outlier (IO) unit root test statistic by Clemente, Montanes, and Reyes (1998) to identify the existence of a structural break in the housing price time series.

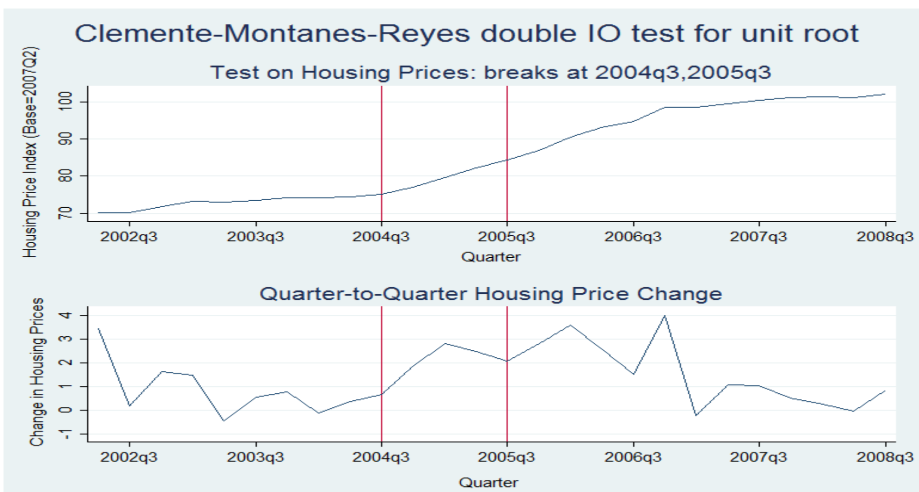

Figure 12: Housing Prices in Greece and EMU creation (2002), 3rd Quarter of 2002 – 3rd Quarter of 2008

Source: Bank of Greece, Real Estate Market Analysis

[3] We cannot exclude the possibility that zoning corruption itself may have had an effect on housing prices (reverse causality). But note that zoning corruption works to increase the supply of housing—shifting the supply curve rightward as more buildable land becomes available—and therefore putting downward pressure on housing prices. As Glaeser, Gyourko, and Saiz (2008) note, fewer and shorter bubbles with smaller price swings should be expected in regions with more elastic housing supply. Sowell (2009) finds that the largest housing price increases occurred in housing markets where local governments imposed land use restrictions which reduced the supply of available land for housing. Even if a causal loop between housing prices and zoning corruption is at play, the direction of the bias suppresses the magnitudes of our estimates and, thus, making our results more conservative.

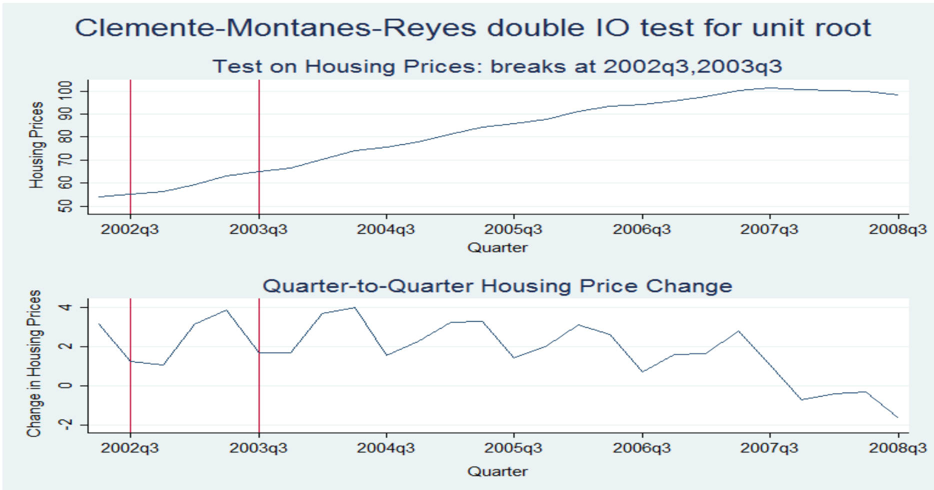

Figure 13: Housing Prices in Spain and EMU creation (2002), 3rd Quarter of 2002 – 3rd Quarter of 2008

Source: Spanish Ministry of Public Works

In both countries, a structural break in the time series of housing prices occurred within two years after the the adoption of the euro.[4] This provides formal evidence that housing markets in Greece and Spain underwent substantial changes following the adoption of the euro.

Following Rose-Ackerman (1975), in this paper we define zoning corruption as "special treatments" of housing developers by zoning officials for personal gain. These "special treatments" usually took the form of granting illegal building permits in Greece and the rezoning of land tracts for residential and commercial use in Spain. The causal mechanisms that determine the supply and demand for "special treatment" may work either through the incentives individuals face or through the opportunities to actually engage in zoning corruption.

[4] This statistic is preferred to the Perron-Vongels and and Zivot-Andrews statistics because it can capture more than one structural breaks in the time series while retaining its ability to identify only one (Baum, 2001).

First, we discuss how increased financial incentives favored more zoning corruption. Surging housing prices in combination with fixed, inflation-adjusted construction costs (Brissimis et al., 2007) and decreasing lending costs meant greater profit opportunities for housing developers. As new housing construction projects became more profitable, this provided housing developers with larger incentives to bribe. As Rose-Ackerman (1988, p.278) puts it, “if bribes are offered there must be some prospective excess profits out of which to pay them”. The size of the bribe may be related to the structure of uncertainty about the expected costs a corrupt zoning official faces (Bliss and Di Tella, 1997). In essence, the surge in housing prices increased the expected benefits to developers of extending a bribe to a zoning official (Becker, 1968).

It is important to also highlight the opportunities for zoning officials to engage in corruption. The institutional framework in Greece and Spain hugely empowered zoning officials on urban planning matters. Decision making on zoning issues was at the local level and the process was overseen by municipal authorities only. As Tanzi (1998, p.14) has noted: “decisions as to the particular use of private land (zoning laws) that determine whether the same piece of land can be used only for agriculture, and thus have low market value, or for high rise buildings, and thus can be very expensive… are prone to bribery”. The natural variation in the supply of housing and zoning laws made land a commodity of differential value across regions. The opportunity for zoning corruption remained available since the zoning officials' audit probability remained unchanged. The delay of necessary institutional reforms could be attributed to the bonanza years of expanding incomes and rising tax revenues in the EU periphery (Fernández-Villaverde et al., 2013). Therefore, the expected costs to violators remained stable leaving the willingness to accept a bribe and the *supply of special treatment* unchanged.[5] Bribing opportunities persisted as long as government-generated shortages in the form of zoning laws and regulations were in place (Lee, 1994). In a way, the institutional setup allowed zoning officials to function as "discriminating monopolists fixing market clearing rates for the services being offered" (Jagannathan, 1986).

For our baseline results, and based on the availability and nature of the data, we estimate separate negative binomial models for each country. We model corruption counts using monthly regional data from 2003 through 2008 in Greece and quarterly regional data from 2006 through 2008 in Spain. We find a positive and significant association between housing prices and zoning corruption. The non-constant marginal effects indicate that zoning corruption increased at a faster rate than housing prices. Our baseline findings are robust to a series of robustness checks. We estimate zero-inflated variants of the unconditional negative binomial model to test whether our baseline model is mis-specified. We investigate different specifications to address any omitted variable bias concerns. We estimate linear models to assess the validity of the "memoryless" property embedded in the count data analysis, and the importance of dynamic panel bias, particularly for the case of Spain.

The rest of the paper is organized as follows. Section 2 presents a review of the empirical literature on the determinants of corruption. Section 3 discusses the data and methodology and Section 4 presents the baseline results. Section 5 discusses econometric issues pertaining to our analysis, and addresses them in a succession of robustness checks. Section 6 concludes.

---

[5] Of course, housing bubbles make acceptance of a bribe relatively more appealing to zoning officials when it is not accompanied by a tightening of the penalties associated with misconduct (Aidt, 2003).

2. Review of the Literature on the Determinants of Corruption

Much of the empirical literature on corruption has been hampered by the use of subjective measures of corruption and the lack of a proper identification strategy. The first wave of empirical studies of the determinants of corruption, often cross-national analyses, made use of subjective measures of corruption from survey data from Transparency International Global Corruption Barometer, the United Nations World Value Survey, and the World Bank Worldwide Governance Indicators (WGI).[6] The main weakness of this literature is some well-known shortcomings with the use of subjective measures of corruption. Subjective measures of corruption rely on perceptions or on personal experiences of bribery. Surveys use responses of businesses or households to questionnaires that inquire whether or not they have ever given an illegal payment to a public official, or their opinion about how corrupt the government is. However, individuals or businesses may want to mask their actual behavior in survey responses to avoid potential penalties or further actions. As a result, subjective measures of corruption may be plagued by imprecision (Kaufman et al., 2010). In addition, systematic biases may undermine empirical analysis when different classes of respondents differ by design, or when the ideological orientation of the institution matters. In addition, subjective assessments might be driven by "halo effects"; i.e., negative evaluations when conditions are worsening and vice versa. Finally, when different data providers use each other's evaluations, endogeneity is introduced since perception errors are correlated. In this paper we bypass the aforementioned challenges by

[6] There has been a host of studies based on the survey's data. Just to mention some of the findings in this literature on the determinants of corruption, Serra (2006) conducts a cross-country Global Sensitivity meta-analysis and finds that income, history of democratic institutions, Protestant religion, colonial heritage, and political instability are all robust determinants of corruption. Mauro (1997), Beets (2005), Lederman et al. (2005), and Cheung and Chan (2008) report a negative association between education level and corruption; however, Frechette (2006) arrives at the opposite conclusion. Beets (2005) finds that higher levels of unemployment are associated with higher corruption and Emerson (2006) finds a negative relation between market competition and subjective measures of corruption. Lambsdorff (2006), Svensson (2005), Beets (2005), Treisman (2000), Husted (1999), Mauro (1997), and Chang (2010) find a negative association between income level and corruption. On the contrary, Braun and Di Tella (2004) and Frechette (2006) find GDP per capita to be positively correlated with corruption using panel data.

employing objective counts of corruption. We measure actual counts of zoning corruption coming from public records in Greece counts and media reports in Spain. The data are described extensively in section 3.1 below.

The literature on corruption has also struggled with the lack of proper identification strategies. Due to the practical difficulties in conducting natural field experiments on corruption, only very limited evidence from the field exists. And by design, these studies can only reveal effective approaches to curbing corruption rather than identifying its causal determinants. Olken (2007) did carry out a field experiment in over 600 Indonesian village road projects to find that government audits reduced corruption significantly whereas increased grassroots participation in monitoring had negligible effects. The absence of clear-cut comparative case study settings has prevented the use of standard quasi-experimental techniques to study corruption. As a result, the methodology employed has been typically restricted to either instrumental variable (IV) and dynamic panel techniques or has exploited instances of exogenous, natural variation in the determinants of corruption.

Some of these studies include Ferraz and Finan (2009) who use audits of incumbent municipal politicians to construct enhanced measures of corruption that control for unobserved characteristics of the locale. Based on reduced-form as well as IV techniques, Brollo et al. (2013) report that an increase in federal transfers in Brazil by 10 percent amplifies the incidence of severe corruption by about 16 percent. Following the same designs, Dong and Torgler (2013) report that Chinese provinces with resource abundance exhibit greater corruption whereas higher educational attainment and fiscal decentralization leads to significantly less corruption. Batzilis (2014) uses a IV approach to show that electoral competition at the municipality level reduced public spending corruption in Greece. Also, he finds a higher incidence of corrupt spending in less populated, more rural municipalities. Del Monte and Papagni (2007) apply an autoregressive distributed lag model to regional data from 1963 through 2001 in Italy to investigate the determinants of corruption, defined as crimes against the public administration reported to the police. They find that per capita GDP, public expenditure on consumption goods and services, and institutional and judiciary changes are the most important contributors to public corruption in Italy.

A different strand of the empirical literature has relied on instances of natural resource windfalls as sources of exogenous variation to examine corruption. Maldonado (2010) exploits exogenous variation in economic conditions in Peru from mineral price shocks due to the relative abundance of mineral resources across regions. His results suggest that the increases in transfer funds due to positive shocks in international mineral prices affected corrupt practices in citizens' interactions with public officials, and that these corrupt practices differed according to the size of the shock. Caselli and Michaels (2013) use variation in oil output among Brazilian municipalities that was exogenously dictated by world oil prices to find evidence of embezzlement in oil-rich municipalities. It is in this latter strain of the empirical literature where our paper fits. Similarly to those papers, we exploit variation in corruption incentives (rents reflected by housing prices) that naturally occurred following the adoption of the euro in Greece and Spain. We focus on a period of housing bubbles as documented by structural breaks in the residential property price time series. The steep acceleration in housing prices allows us to identify their role in zoning officials' corruption motives separately from other confounding factors of corruption.

3. Empirical Analysis

We conduct separate empirical analyses for Greece and Spain. We refrain from pooling the observations from the two countries because that would be responsible for a large fraction of the overall variation. Even though Greece and Spain are not dramatically different in many respects, there are significant institutional and economic differences.[7] Furthermore, the observations of zoning corruption are coming from distinct data generating processes for each country. This is why we estimate parameters using a panel of a single country at a time, and, thus, exploit greater cross-sectional variation due to the regional disaggregation.

Given the within-country nature of our study, we assume institutional homogeneity; i.e., relative wages of zoning officials and anti-corruption efforts by auditors are equal across regions, within country. This may be quite appropriate given the nationwide salary base and trends that applied in each country. It is also an empirical necessity because controls of institutional quality such as "Rule of Law" and "Quality of the Bureaucracy" are only available at the national level for the period of analysis.[8] And even though the low mortgage interest rates that triggered the housing bubbles in Greece and Spain were uniformly available across regions, it could be argued that there may still be differences in the quality of public administration. This would imply substantial regional heterogeneity in the level of corrupt activities but does not seem to be the case as discussed below. Our analysis captures regional institutional differences through the inclusion of education attainment as a control variable. Education is often regarded as a proxy of overall development and has been found to correlate with enhanced government performance

---

[7] For example, Golden (2015) argues that "there is little likelihood that Spain, for instance, will tumble to Greece's…levels of corruption".

[8] An exception is the European Quality of Government Index (EQI), which is the result of novel survey data on corruption and governance at the regional level within the EU, conducted first in 2010 and then again in 2013. However, these survey data only cover years later than the scope of our analysis

and reduced levels of evasion.[9] Lastly, any time-invariant, region-specific characteristics we failed to control are captured by regional dummies. Arguably, these should capture the bulk of institutional differences among regions given the slowly-adjusting nature of institutional quality.

3.1 Data

Our main dependent variable is an objective (non-perception, non-survey-based) metric of corruption: legal indictments of zoning officials from prosecution records in Greece, and zoning corruption scandals reported in the media in Spain. In both cases, zoning corruption is measured at the regional level. The observations of the zoning corruption are non-negative and discrete; thus, they qualify as count data (Bujko et al., 2015).

In the case of Greece, information comes from monthly prosecution records of zoning officials by the Inspectors-Controllers Body for Public Administration of Greece from January 2003 through December 2008. This is an internal audit service of the public administration which independently monitors all public officials in Greece.[10] There are six characterizations of public administration malpractices, each of which reflects a different degree of corruption intensity. We only use the most extreme case of zoning corruption intensity to make results for Greece directly comparable to those for Spain. More importantly, it allows us to rule out the possibility that increased counts of zoning corruption are due to the reporting of multiple prosecutions of the same defendant. The use of corruption audits data is a relatively recent development in the literature, but it is becoming more prevalent. A number of studies has relied on corruption audits data as the research community is becoming increasingly more engaged and

[9] See, for example, Mauro (1997), Beets (2005), Lederman et al. (2005), and Cheung and Chan (2008).

[10] The data were obtained during our personal visit to the Inspectors unit in Athens in August 2014. Inspections are delineated into 16 categories that fully describe the nature of the violation. The six corruption-related categories incorporated in our measure of the dependent variable are: (1) Corruption - Other Legal Indictments; (2) Illegal Action; (3) Organizational Problems; (4) Violations of Code of Conduct; (5) Violation of Transparency; (6) Omission of Designated Action.

works directly work with public officials for the improvement of the quality of the public administration (Olken, 2007; Ferraz and Finan, 2009; Caselli and Michaels, 2013; Brollo et al., 2013).

The data on zoning corruption in Spain were obtained from the GISAS Research Group at the University of La Laguna in Spain. Although the original dataset comprised corruption scandals from 2000 to 2008, just 3 per cent of the counts were recorded for the period 2000-2005. This is the case because information for these two periods was collected independently by different sources. To prevent measurement error from combining information by two different data generating processes, we discard any observations prior to 2006. The dependent variable for Spain represents the quarterly number of media reports of zoning corruption scandals from the first quarter of 2006 to the fourth quarter of 2008. These scandals involve real estate crimes committed by zoning officials at the municipal level, and are aggregated up to the regional level. The calculus of corruption counts is based on web-scrapping of media reports. This approach of corruption enumeration is not new. It has been used in previous studies of public corruption in Spain (Fundación Alternativas, 2007; Costas-Pérez et al., 2012), financial scandals and election law violations in Japan (Nyblade and Reed, 2008) and oil revenue embezzlement in Brazil (Caselli and Michaels, 2013). As a matter of fact, the same dataset has already been used in publications that examine the impact of zoning corruption on political outcomes (Martín-Martín et al., 2010; Jerez-Darias et al., 2012).

Housing prices is the key explanatory variable used in the analysis and reflects profit incentives for bribing. They are measured in euros per square meter. We use this normalization to account for any effects the size of property may have on its price. Considering merely contemporaneous prices we cannot explore whether a rise in housing prices triggered zoning corruption in a future time period. To account for the time-consuming process a corruption count takes to mature we use 18-month housing price lags.[11] This corresponds to the average duration for a dwelling to be constructed in Greece and accounts for the time required for the gradual exposure and media coverage of a zoning scandal indictment in Spain.

For Greece, housing prices are constructed based on the Housing Price Index compiled by the Real Estate Market Analysis Section of the Bank of Greece. Four time series are available of which only two perfectly match our regional disaggregation for Greece. We interact these time series with actual median residential property prices from Mitrakos et al. (2014) to introduce further cross-sectional housing price variation. The latter study provides information on 28 cities across 12 regions of Greece. For the one region which is not covered in Mitrakos et al. (2014), we use its neighboring region's housing price information. Our results are robust to the exclusion of the region for which we lack a local housing indicator. Housing prices are expressed as housing prices per square meter in a region of Greece.

For Spain, housing prices were extracted from the website of the Spanish Ministry of Public Works and Transport and match regions perfectly. They represent property values of new housing units per square meter from 2000 through 2008 in quarterly frequency. We use values earlier than 2006 to create the necessary housing price lags without sacrificing any zoning corruption observations when differencing.

[11] The use of lags of independent variables to study corruption is not novel. In a study of the regional determinants of corruption in Italy, Del Monte and Papagni (2007) also resort to lagged terms because "time is needed for bargaining" and inaccuracies may arise between committing and reporting the corrupt activity.

Next, we use the number of new building permits to approximate the window of opportunity of zoning corruption. Figures 14 and 15 below plot zoning corruption and building permit averages for each region in Greece and Spain, respectively, without adjusting for regional population. However, confounding will arise if housing prices and zoning corruption have common causes that we fail to condition on.

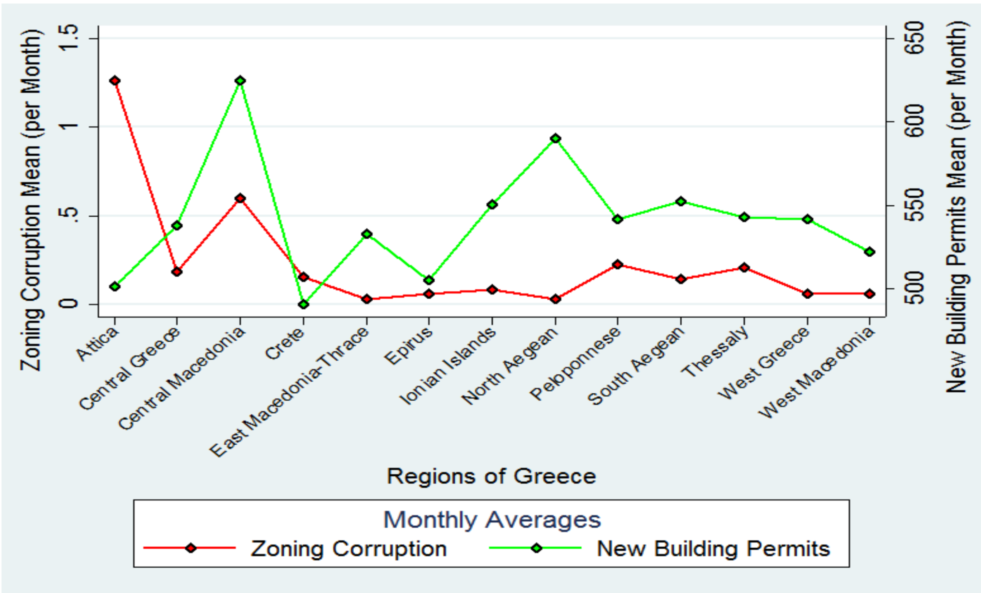

Figure 14: Average Zoning Corruption and New Building Permits per Region in Greece

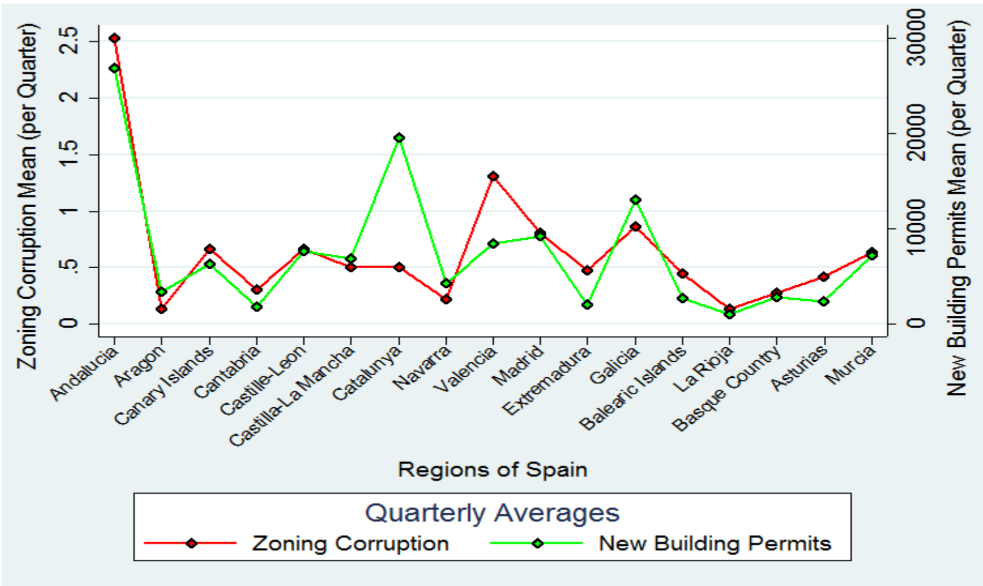

Figure 15: Average Zoning Corruption and New Building Permits per Region in Spain

Note that the ratio of the zoning corruption mean to the building permit mean is fairly constant across regions within Greece and within Spain. First, this is encouraging for the validity of our design because it illustrates that the opportunity to engage in zoning corruption followed a common pattern, across regions, as the institutional and judicial uniformity we assumed, within country, would suggest. Second, the predictive power of building permits demonstrates its appropriateness as the exposure variable in our modeling. In Greece, the notable deviation from the corruption-permit country ratio is for Attica, the region of Athens. This could be explained by its extremely high population density (highest among all 30 EMU regions we consider). If population density is operationalized through higher-rise buildings then the number of new building permits need not be as large as in other regions. In addition, the fact that the homeownership rate in Greece was high even before the adoption of the euro implies smaller new building needs at the largest population center. As a matter of fact, many new residential properties in Greece were constructed as second-home properties in regions where tourism is prevalent such as the South and North Aegean and the Ionian Islands for residents from the large cities. The fact that Central Macedonia, a region of Thessaloniki, the second largest city in Greece, exhibits a large number of new building permits could be attributed to the presence of a major tourist destination in the region (Chalkidiki) where second-home housing construction grew substantially post-EMU. In Spain, Catalonia stands out with a much smaller zoning corruption to building permit ratio compared to other regions in Spain. Its capital, Barcelona, is considered a top haven for expatriate retirees with an estimated population of over 35,000.[12] The associated increased demand for housing is one of the contributing factors for such outlier behavior. And, similarly to the experience we document for the case of Greece, demand from individuals of landlocked regions that

[12] 9 Top Havens for Expat Retirees: http://money.usnews.com/money/blogs/on-retirement/2014/08/28/9-top-havens-for-expat-retirees

experienced steep growth such as in Madrid, Leon, Aragon, and Extremadura who are seeking a coastal second-home could also explain part of the excess construction as these individuals are entering the housing market seeking new coastal establishments.

Besides housing prices and building permits, we allow for other factors to influence the proclivity of public officials towards corruption on the basis of findings in the within-country literature on the determinants of corruption. For our baseline estimation we adjust for the unemployment rate, population, population density, and the enrollment rate in tertiary education. All control variables are obtained from Eurostat's regional statistics portal. They are available at the regional level (EU classification: NUTS level 2); the latter three in annual frequency and unemployment in quarterly frequency.

The measurement of education attainment via enrollment rates in tertiary (higher and continuing) education may be subject to bias because large metropolitan areas feature more tertiary educational institutions, skewing enrollment rates higher in more populous regions. However, this concern is mitigated by the fact that public tertiary education is uniformly provided across regions of Greece and Spain, a feature that is quite common across public continental European education systems. Literacy rates are an alternative measure commonly found in past empirical research but would have yielded negligible variation across regions given the EMU context.

Population measures the number of inhabitants in a region (all ages, both sexes). A combined reading of Figures F.1 through F.4 in Appendix F reveal that construction activity is positively related to increases in population. The highest counts of zoning corruption are recorded in the most populous regions in both Greece and Spain; Attica and Andalucía, respectively. This mechanically widens the window of opportunity for corruption in more populous regions. Thus, we include population to control for any population size effects (level effects) on the counts of zoning corruption. Clearly, more populous regions will be mechanically associated with higher demand for housing.

Population density measures the number of inhabitants in a region (all ages, both sexes) per square kilometer. We believe that population density is an important determinant of zoning corruption because it affects its monitoring intensity. In densely populated regions, more people compete for land. This produces more checks on the performance of the zoning officials and, subsequently, less corruption. As a matter of fact, in Greece most zoning corruption is exposed after adjacent residential property owners file a complaint to the public administration auditors. This, in turn, triggers investigation of the case by the auditors. If any misconduct on behalf of the zoning official is revealed, arraignment ensues. Our zoning corruption observations for Greece come from the subset of arraigned officials that were, eventually, legally indicted. In essence, more "eyes" provide more checks of corrupt zoning activity.

We use the unemployment rates of individuals aged 15 years or over for both sexes as a predictor of housing demand. GDP could have alternatively been used to proxy housing demand. However, since GDP is more strongly correlated with housing prices this would have attenuated our coefficient of interest.[13]

[13] Indeed, housing prices are more collinear with GDP (ρ= 0.29 for Greece; ρ= 0.68 for Spain) than unemployment (ρ=-0.28 for Greece; ρ= -0.14 for Spain)

Tables G.1 and G.2 in the Appendix G provide summary statistics for Greece and Spain, respectively. Table G.3 contains a short description of the variables used as well as their source. Tables G.4 and G.5 report correlation coefficients between all of the study variables for Greece and Spain, respectively.

3.2 Estimation approach

We model zoning corruption (Y_{it}) as a function of housing prices (HP_{it}) and adjust for educational attainment, unemployment, population, population density, and number of new building permits $\boldsymbol{X}_{it}$ issued in region i at time t. We also use housing price lags to account for the inherently slowly process of construction. Time-invariant, region-specific variation is captured by the regional dummy variables α_i.

Because the count of zoning corruption includes many zero values, our modeling strategy needs to take this into account. The coexistence of a high proportion of zeros along with large values of counts makes it difficult for a linear specification to appropriately model corruption. Assumptions such as normally distributed residuals and constant variance required for identification in an ordinary least squares regression model rarely fit count data. The Poisson and the negative binomial are the primary specifications used in the literature under these conditions (Cameron and Trivedi, 2005).

A Poisson data generating process assumes the rate parameter $\alpha_i \exp\left(\boldsymbol{X}_{it}'\boldsymbol{\beta}\right)$ is known when, in fact, it is a random variable itself. Further, Poisson models require that the conditional mean is equal to the conditional variance. In our case, the latter exceed the former for both Greece and Spain. In the presence of overdispersion (λ_i), negative binomial models are preferred to Poisson ones due to their flexibility in modeling the moments of the observed distribution of corruption. Namely, negative binomial models permit the conditional mean to differ from the conditional variance. The superiority of the negative binomial to the Poisson models in fitting identified corruption in Greece and Spain can be seen, respectively, in Figures 16 and 17 below.

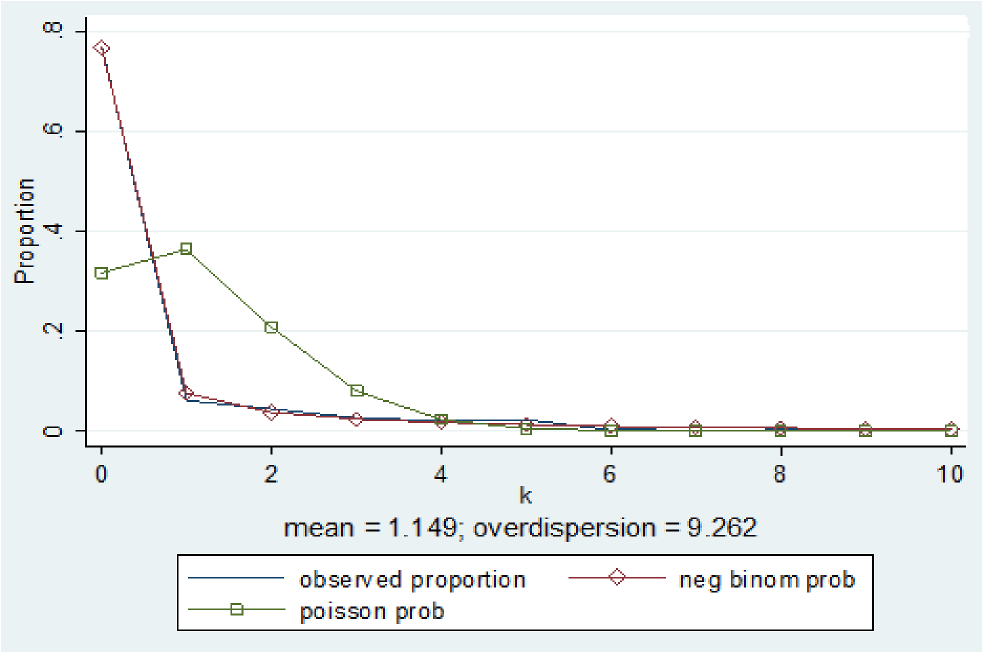

Figure 16: Poisson and Negative Binomial Fits of Zoning Corruption, Greece

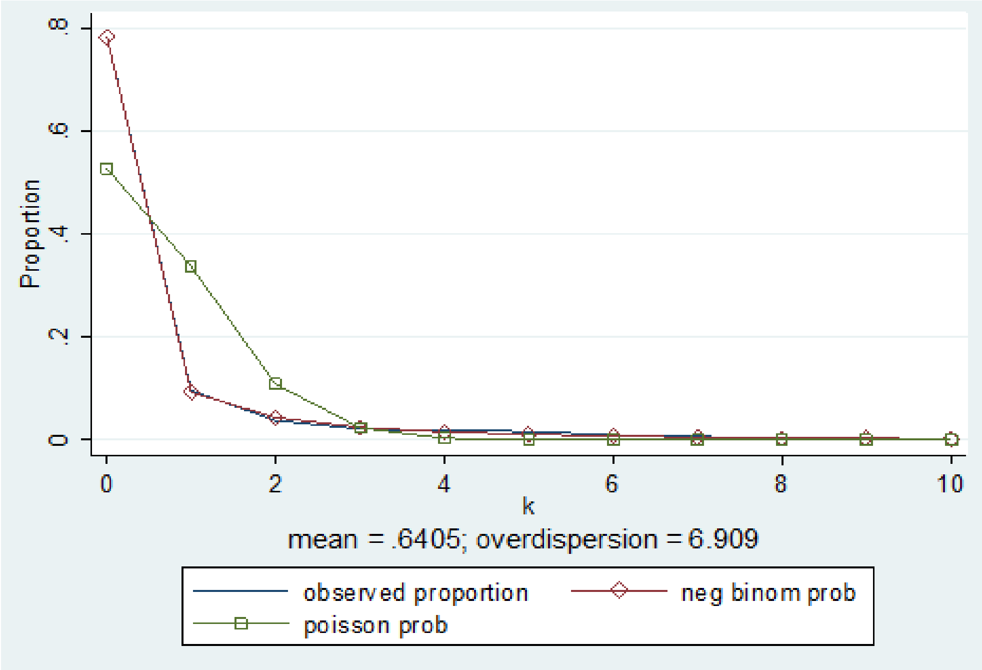

Figure 17: Poisson and Negative Binomial Fits of Zoning Corruption, Spain

Likelihood-ratio tests formally confirm the validity of our choice of negative binomial as opposed to Poisson models, too.

Hausman, Hall and Griliches (1984) use the Andersen's conditional log likelihood to estimate fixed-effects negative binomial models. Allison and Waterman (2002) suggest that Hausman et al. (1984) did not formulate a true fixed effects model in the mean of the random (dependent) variable because they did not control for all time invariant covariates. Their model layers the fixed effect into the heterogeneity portion of the model and not the conditional mean. This portion is then conditioned out of the distribution to produce the model Hausman et al. (1984) estimate. However, the parameters that are conditioned out of the likelihood function do not correspond to different intercepts in the log-linear decomposition of the overdispersion parameter. Allison and Waterman (2002) propose an unconditional negative binomial model that uses dummy variables to represent fixed effects. We follow their approach and specify 13 and 17 regional dummy variables for Greece and Spain, respectively. The probability mass function of a single count y_{it}, its mean $E(y_{it})$ and variance $Var(y_{it})$ is obtained from

$$f(y_{it}|\mu_{it},\lambda_i) = \frac{\Gamma(\lambda_i + y_{it})}{\Gamma(\lambda_i)\Gamma(1+y_{it})}\left(\frac{\mu_{it}}{\mu_{it}+\lambda_i}\right)^{y_{it}}\left(\frac{\lambda_i}{\mu_{it}+\lambda_i}\right)^{\lambda_i} \quad (1)$$

$$E(y_{it}) = \mu_{it} \quad (2)$$

$$var(y_{it}) = \mu_{it}\left(1+\frac{\mu_{it}}{\lambda_i}\right) \quad (3)$$

where the mean μ_{it} fluctuates with time but the overdispersion parameter λ_i is time-invariant.

By design, the dependent variable could be interpreted as a *rate* of zoning corruption; zoning corruption cases per number of new building permits issued in a month in Greece or a quarter in Spain.[14] Therefore, our estimates should be adjusted for the amount of opportunity for corruption events. To do so, we introduce an exposure variable which reflects the amount of exposure over which the outcome events were observed for each region-time observation. Construction permits are chosen because they captured the window of bribing opportunities to a zoning official. In the polar case, if zero building permits were issued, no zoning corruption could have been detected.

We, implicitly, assume that the likelihood of zoning corruption events did not change over time. That is, the denominator of the zoning corruption rate did not affect the numerator beyond opportunity. If, for example, it took auditors and prosecutors different times to clear a zoning corruption case then arrival time of each zoning corruption event was not just a matter of exposure. But, there is no reason to believe that this was the case. In Greece, we use corruption events of only the same intensity, thus, clearing time of each prosecution should have been the same. For Spain the assumption would be violated if reporting of each case differed by the time it took to be either judicially processed or released in the news. We discard these possibilities because of the uniformity of the judicial system in Spain and the fact that the arrival time of zoning corruption scandals to all media outlets is concurrent by briefs from the Associated Press in Spain. Indeed, media coverage of zoning corruption across regions of Spain is not a concern as 97 percent of the zoning corruption scandals locally reported also appeared in national media.

[14] To transform the dependent variable from a rate back to a count one can multiply both sides of the estimated equation by the exposure variable. Logging yields a right side of the equation that includes the natural logarithm of the exposure variable. This is called the offset variable and is required to have a coefficient of 1. Moving the offset variable back to the left side of the equation turns the count back into a rate.

Our baseline specification is the unconditional negative binomial dummy variable model below:

$$\ln(\mu_{it}) = \beta_1 HPI_{it} + \boldsymbol{\beta_2 X_{it}} + \boldsymbol{I}(Region_i) + \beta_3 \ln(Permits_{it}) \quad (4)$$

where $\beta_3 = 1$. It is also referred to as the NB2 model (Cameron and Trivedi, 1998).

4. Results

Recall that the Poisson model is a special case of the negative binomial one – it restricts the over-dispersion parameter λ_i to be equal to zero. To test whether $\lambda_i = 0$, we perform a likelihood-ratio test. Standard errors are clustered by region to allow for intra-regional correlation.

4.1 The Case of Greece

The likelihood-ratio test of $\lambda_i = 0$ has an associated χ^2 value of 132.93 which suggests the probability that these observations were generated under the assumption that $\lambda_i = 0$ is zero. This finding favors the use of negative binomial models in expense of the Poisson ones.

Parameter estimates are therefore based on negative binomial regression models. The coefficient of interest is β_1 and shows the impact of the housing prices during the bubble on zoning corruption. Given the model's non-linearity, its interpretation is not straightforward. Therefore, the focus in Table 1 is only on the signs of the coefficients. Table 2 presents marginal effects that allow for a more natural interpretation of the coefficients' magnitudes.

Table 1: Negative Binomial Regression Model – Greece, 2003-2008

Variables	(1) Contemporaneous Housing Prices	(2) 18-month Housing Price Lag
Housing Prices(€/m^2)	0.000182	0.00194**
	(0.000399)	(0.000879)
Unemployment	0.0811	0.140
	(0.179)	(0.243)
Population Density	−0.112***	−0.0973***
	(0.0226)	(0.0300)
Population	0.00002***	−0.000058
	(0.000002)	(0.000062)
Education	0.386***	−0.0152
	(0.107)	(0.190)
Central Greece	−40.51**	−300.7
	(19.17)	(224.5)
Central Macedonia	−62.62***	−214.9
	(18.83)	(140.5)
Crete	−39.80**	−294.2
	(18.86)	(220.7)
East Macedonia-Thrace	−43.94**	−298.7
	(19.13)	(221.1)
Epirus	−39.64 **	−314.8
	(19.45)	(237.8)
Ionian Islands	−27.71	−316.9
	(17.96)	(245.8)
North Aegean	−34.69*	−322.4
	(18.98)	(246.4)
Peloponnese	−41.31**	−298.1
	(19.27)	(222.3)
South Aegean	−32.59*	−312.9
	(18.56)	(239.0)
Thessaly	−44.67**	−287.2
	(19.26)	(212.1)
West Greece	−43.20**	−290.6
	(18.86)	(215.3)
West Macedonia	−38.38**	−318.3
	(19.37)	(241.4)
Constant	19.14	324.9
	(18.98)	(258.5)
Observations	936	702

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Standard errors clustered at the regional level (in parentheses)

Columns (1) and (2) of Table 6 above present the results of specifications using contemporaneous and lagged housing prices, respectively. Column (2) suggests that for a one euro per square meter increase in monthly housing prices, the difference in the logs of expected counts of zoning corruption is expected to increase by 0.00194 unit, while holding the other controls constant. This relationship is statistically significant at the 5 percent level.

Unemployment is positively but not significantly correlated with zoning corruption. More densely populated regions have significantly lower zoning corruption. This is in line with the practice of the audit process in Greece which is initiated and augmented by adjacent property owners' reports. This result highlights the importance of an active –or at least self-interested- citizenry. The findings on the effects of educational attainment and level of population on zoning corruption are mixed. Regional dummies are estimated in reference to Attica, the region of Athens which is the omitted region. This is the most populous region of Greece by far; approximately half of the country's population resides in Attica. This is reflected on zoning corruption counts as well, with the majority of corruption incidences occurring there. As a result, all other region dummies compare negatively to Attica. We interpret this as a sheer size effect.

A tabular representation of the housing prices marginal effects is given in Table 7 below.

Table 2: Housing Price Marginal Effects – Greece, 2003-2008

	(1)	(2)
Housing Prices(€/m^2)	Contemporaneous	18-month Lag
1250	0.385**	0.119*
	(0.161)	(0.061)
1500	0.403***	0.193***
	(0.129)	(0.058)
1750	0.421***	0.314***
	(0.093)	(0.028)
2000	0.441***	0.510***
	(0.054)	(0.078)
2250	0.461***	0.828***
	(0.015)	(0.305)
2500	0.483***	1.346*
	(0.040)	(0.789)
2750	0.505***	2.188
	(0.091)	(1.762)
Observations	936	702

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Standard errors clustered at the regional level (in parentheses)

Variation comes in increments of €250 per square meter, starting from €1500 up to €3000 per square meter. We base our results on the lagged housing price specification. Column (2) indicates a wide range of effects. The average predicted monthly corruption count is 0.193 when 18-month lagged housing prices equaled €1500 per square meter and the other controls are at their mean values. We estimate as many as 1.3 more cases of zoning corruption per month when housing prices double to €3000 per square meter. This increasing relationship to housing prices is as expected; zoning corruption increased when profit opportunities became larger. This can also be seen in the graphical representation of the 18-month lagged housing price marginal effects in Figure 18 below.

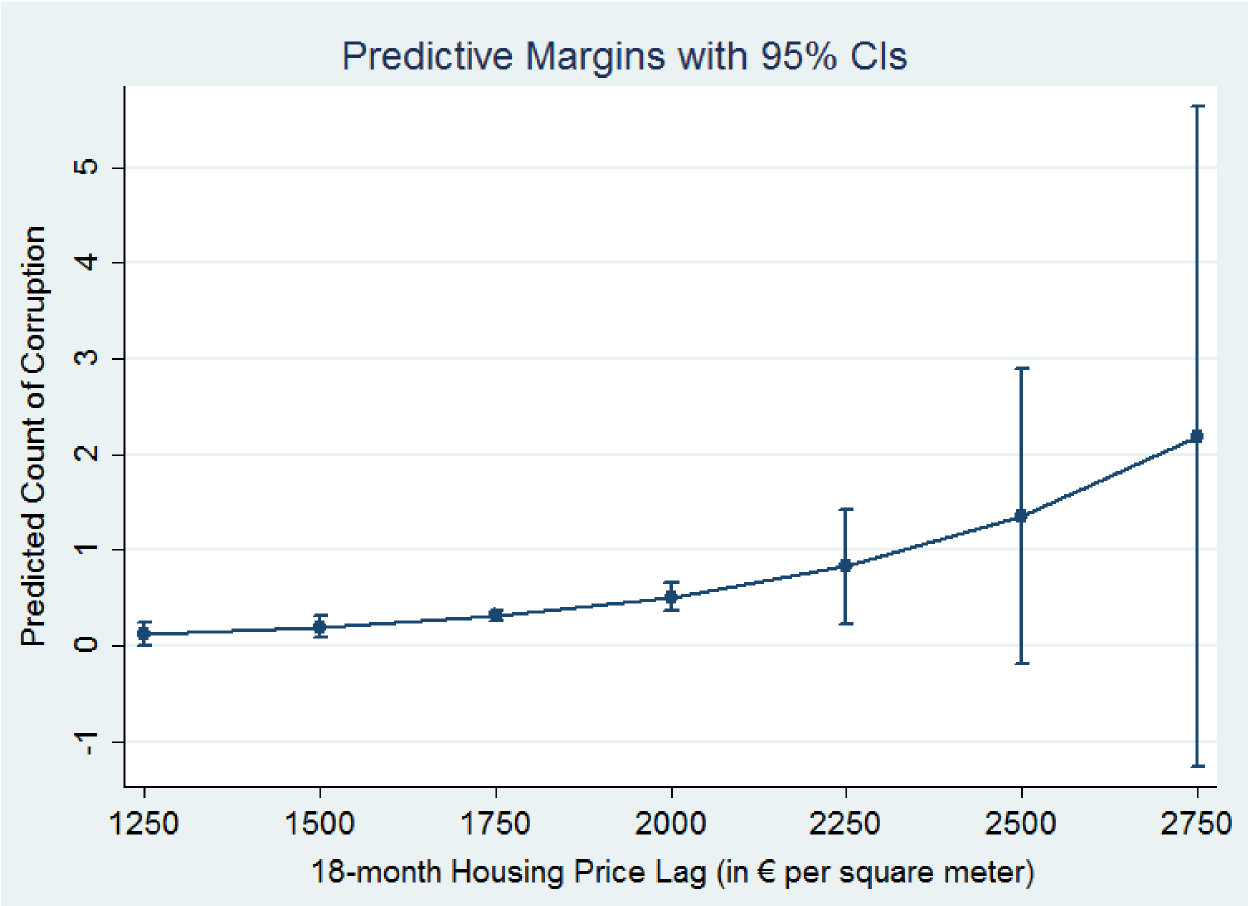

Figure 18: Lagged Housing Price Marginal Effects, Greece

The graph covers a range of housing price values starting from €1250 through €2750 per square meter in increments of €100 per square meter. It is evident from Figure 8 that as housing prices increased, zoning corruption was intensified, as conjectured.

4.2 The Case of Spain

First, we discuss the results from the likelihood-ratio test of $\lambda_i = 0$. The χ^2 value of 46.09 suggests that the probability that these observations were generated conditional on $\lambda_i = 0$ is zero. This implies that the negative binomial model is also far more appropriate than the Poisson model in modeling zoning corruption in Spain.

Table 3 presents parameter estimates of negative binomial regression models for Spain. Again, the focus should be primarily on the signs of the coefficients. Table 4 presents marginal effects that allow us to interpret the magnitudes of the coefficients directly.

Table 3: Negative Binomial Regression Model – Spain, 2006-2008

Variables	(1) Contemporaneous Housing Prices	(2) 6-quarter Housing Price Lag
Housing Prices(€/m^2)	0.00807***	0.00382***
	(0.00167)	(0.000777)
Unemployment	0.0171	0.00179
	(0.149)	(0.164)
Population Density	−0.0107	−0.0386*
	(0.0185)	(0.0200)
Population	−0.00000006	0.00000005
	(0.00000002)	(0.00000142)
Education	0.0696	0.0892
	(0.120)	(0.116)
Aragon	−3.471	−4.252
	(13.60)	(10.01)
Canary Islands	1.829	7.449
	(12.99)	(9.886)
Cantabria	−1.356	1.087
	(14.66)	(10.51)
Castille-Leon	−0.411	−3.477
	(11.87)	(9.243)
Castille-La Mancha	2.426	−0.216
	(9.554)	(6.864)
Catalonia	−5.658	1.074
	(4.667)	(3.685)
Navarra	−2.476	−3.349
	(15.42)	(11.54)
Valencia	3.736	6.615
	(7.936)	(6.296)
Madrid	−4.246	19.43
	(16.26)	(15.20)
Extremadura	6.270	1.907
	(11.20)	(8.293)
Galicia	0.358	−0.311
	(10.73)	(8.078)
Balearic Islands	−1.948	5.172
	(14.19)	(10.10)
La Rioja	0.908	−0.119
	(14.46)	(10.58)
Basque Country	−8.868	2.465
	(16.43)	(12.15)
Asturias	0.0175	0.533
	(14.51)	(10.81)
Murcia	1.320	2.068
	(12.90)	(9.511)
Constant	−24.64**	−15.66**
	(10.11)	(7.781)
Observations	204	204

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Standard errors clustered at the regional level (in parentheses)

Columns (1) and (2) of Table 3 present the results of specifications using contemporaneous and lagged housing prices, respectively. As for Greece, for Spain we base our findings on the lagged housing prices due to the time-consuming zoning corruption generating process. For a one euro increase in quarterly housing prices, the difference in the logs of expected counts of zoning corruption is expected to increase by 0.00382 units. This relationship is statistically significant at the 1 percent level.

Population density is negatively associated with zoning corruption in Spain. This finding again underscores the importance of civic engagement in curbing zoning corruption. More civilian surveillance can substantially narrow the window of opportunity for zoning corruption.

The estimates of the regional dummies are obtained in reference to Andalucia, which is the most populous and also one of the regions with the highest incidence of zoning corruption scandals in Spain. However, when we adjust for housing prices and other control variables, Andalucia becomes the median region in terms of zoning corruption levels.

Table 4 below presents marginal effects of housing prices on zoning corruption in Spain.

Table 4: Housing Price Marginal Effects – Spain, 2006-2008

Housing Prices(€/m^2)	(1) Contemporaneous	(2) 18-month Lag
750	0.012^{*}	0.038^{***}
	(0.007)	(0.009)
900	0.035^{***}	0.096^{***}
	(0.012)	(0.016)
1050	0.102^{***}	0.241^{***}
	(0.013)	(0.038)
1200	0.303^{***}	0.606^{***}
	(0.062)	(0.122)
1350	0.899^{***}	1.523^{***}
	(0.395)	(0.422)
1500	2.662	3.834^{***}
	(1.822)	(1.401)
1650	7.888	9.647^{**}
	(7.352)	(4.431)
Observations	204	204

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Standard errors clustered at the regional level (in parentheses)

To capture the mass of observations, we consider values from €750 to €1650 per square meter. The marginal effects of housing prices on zoning corruption are estimated in increments of €150 per square meter. We base our results on the lagged housing price specification. Again, the marginal effects exhibit a large degree of variation. The average predicted quarterly corruption count is merely 0.038 when lagged housing prices stood at €750 per square meter but grows exponentially at each housing price increment. The significance of housing prices as a predictor of zoning corruption declines at higher values. This is a consequence of low power at the highest housing price levels. For instance, only four regions ever recorded housing prices greater than €2000 per square meter (Madrid, Catalonia, the Balearic Islands, and the Basque Country). But even so and similarly to Greece, zoning corruption in Spain was amplified when profit opportunities were heightened. This is also evident in the graphical representation of the housing price marginal effects in Figure 19 below.

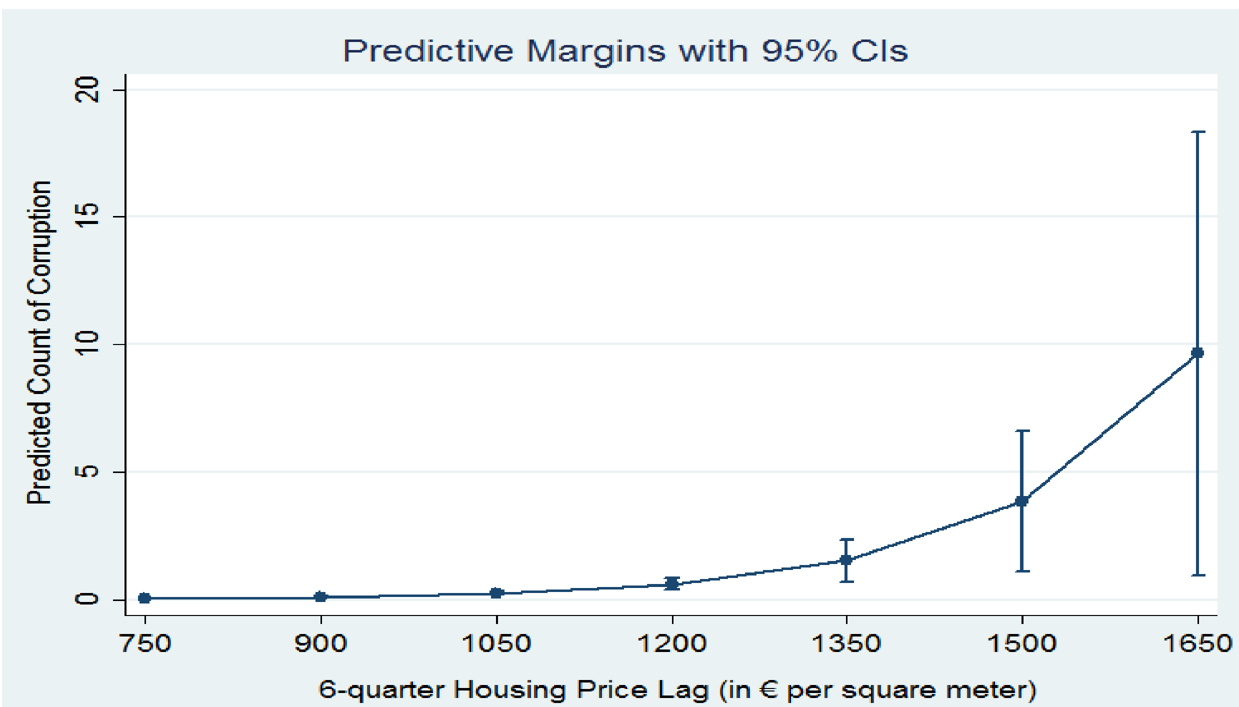

Figure 19: Negative Binomial Regression Housing Price Marginal Effects, Spain

We take a closer look at housing prices and consider values from €750 per square meter to €1650 per square meter in increments of €150 per square meter. We choose this range of housing prices because they are the most frequently documented.

5. Sensitivity Analysis

A potential limitation of the analysis is the implementation of inference from a non-random sample. Given the observational nature of our study, selection bias could be present if, for example, zoning corruption is more likely in some regions than others due to unobservable factors. Such constraints that mask the presence of underlying heterogeneity could lead us to erroneous estimates.

The first potential source of bias could come from the presence of omitted variables. Housing prices may have increased faster in regions with higher demand for housing coming from factors that our baseline analysis does not include. For example, coastal regions where tourism is more prominent could have experienced relatively greater price swings possibly because land there yields higher rents.[15] One would expect zoning authorities in these regions to be more susceptible to receiving bribes for unlawful residential construction. To test whether tourism-intensive regions are affected differently by the housing bubbles, we augment the negative binomial model relevant controls. In particular, in order to capture the influence of tourism in the housing stock of a region we adjust our baseline estimates with the number of nights spent at tourist accommodation establishments and the number of bed-places (total number of beds available to tourists) in a region.

[15] In addition, the majority of new residential properties in Greece and Spain were second homes that were typically erected in coastal areas (summer housing). Also, the percentages of total dwellings occupied by its owner were in both countries among the highest in the EMU. By 2007, the owner-occupancy rates in Greece and Spain were 79.6 percent (4th highest) and 86.3 percent (highest), respectively, when the EMU-average was a mere 62.3 percent (ECB Structural Issues Report, 2009).

The results of the augmented negative binomial regression model are shown in Table 5.

Table 5: Augmented Negative Binomial Regression – Greece, 2003-2008; Spain, 2006-2008

Variables	(1) Greece, Lagged Housing Prices	(2) Greece, Lagged Housing Prices	(3) Spain, Lagged Housing Prices	(4) Spain, Lagged Housing Prices
Housing Prices (€/m^2)	0.002***	0.007***	0.006***	0.005***
	(0.002)	(0.008)	(0.007)	(0.007)
Unemployment	0.114	0.151	−0.107	−0.042
	(0.254)	(0.249)	(0.097)	(0.093)
Population Density	−0.098***	−0.081***	−0.045	−0.048
	(0.031)	(0.021)	(0.032)	(0.032)
Population	−0.000104	−0.000044	−0.000004	−0.000004
	(0.000092)	(0.000046)	(0.000003)	(0.000003)
Education	0.00844	−0.0289	0.153	0.132
	(0.189)	(0.175)	(0.119)	(0.110)
Tourist Bed-Places	0.000085		0.000007	
	(0.000104)		(0.000008)	
Tourist Nights Spent		−0.000022		0.0000003
		(0.000029)		(0.0000001)
Regional Dummies	YES	YES	YES	YES
Constant	245.9	253.5	−23.475***	−33.01***
	(371.1)	(179.7)	(5.675)	(10.39)
Observations	702	702	204	204

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Standard errors clustered at the regional level (in parentheses)

The positive and statistically significant relationship between housing prices and zoning corruption is robust to the inclusion of the tourism intensity controls. The variables 'nights spent at tourist accommodation establishments' and the 'number of bed-places in a region' are not significant predictors of zoning corruption.

The second robustness check tests whether the baseline modeling of zero values is mis-specified. The zero-inflated negative binomial (ZINB) model may be more appropriate when modeling count data with a significant number of zeros. As mentioned above, our time series for zoning corruption events in Greece and Spain contain many zero values. Lambert (1992) and Greene (1994) suggest that excess zeros may be generated by a separate process from the count values and models them independently using a logit model for predicting excess zeros. Here we use issued construction permits as the predictor of the number of zero zoning corruption cases in a region. Intuitively, the window of opportunity to engage in corrupt acts should be greater when more construction activity is present.

To formally test for model misspecification, we conduct the Vuong test to determine whether a ZINB versus a negative binomial model should be specified (Vuong, 1989). In essence, we check whether the presence of a zero was actually true or is simply due to misspecification. This test statistic has a standard normal distribution with large positive values favoring the ZINB model, large negative values favoring the negative binomial one, and with values close to zero favoring neither model (Long, 1997). In the case of Greece, the Vuong test's z-values when modeling contemporaneous and lagged housing prices are 0.67 and 5.85, respectively. For the model of contemporaneous and lagged housing prices in Spain the test's z-value are equal to .25 and 0.23, respectively. As a result, we cannot discard the earlier results in favor of the ZINB models but still provide results from the latter as a supplement.

Columns (1) and (2) of Table 6 below present estimates of the ZINB models for Greece and Spain, respectively. We only present results for our baseline specifications that use lagged housing prices.

Table 6: Zero-Inflated Negative Binomial Regression – Greece, 2003-2008; Spain, 2006-2008

Variables	(1) Greece, Lagged Housing Prices	(2) Spain, Lagged Housing Prices
Housing Prices (€/m^2)	0.00122**	0.00623***
	(0.000510)	(0.00065)
Unemployment	0.133	−0.102
	(0.167)	(0.082)
Population Density	−0.0713***	−0.0533**
	(0.0151)	(0.0246)
Inflate(Population)	−0.000038	−0.0000017**
	(0.000037)	(0.0000007)
Education	−0.0433	0.1062
	(0.134)	(0.1156)
Regional Dummies	YES	YES
Constant	501.6	−6.997
	(371.1)	(5.096)
Observations	702	204

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Standard errors clustered at the regional level (in parentheses). Population specified as the inflated variable.

The positive and statistically significant relationship between housing prices and zoning corruption is robust to the specification of a zero-inflated model as well. The results of this robustness check suggest that the baseline modeling of the zero-dominated corruption generating process is adequate. With respect to housing prices, the ZINB model for Greece yields very similar estimates to the baseline model whereas for Spain it indicates our baseline estimates are, if anything, conservative.

In our third robustness check, we ponder a necessary assumption of our baseline model. The foundation for the negative binomial model is the Poisson distribution. It comes from the family of exponential distributions and exhibits a “memoryless property”. It depends on key assumptions regarding independence, stationarity, and homogeneity (Cameron and Trivedi, 1998, pp.5-8). However, when modeling individual-level, sequential behavioral choices these assumptions may be potentially violated. In our case, it is unlikely that zoning corruption events took place independently of former levels of corruption. In general, corruption is not "memoryless". Bribes are offered because of the prevailing culture of bribery-prone zoning activity (for example, regarded as the most corrupt branch of the Greek bureaucracy). Such cultural stylized facts are certainly not formed overnight. If bribing a zoning official was a phenomenon alien to Greek and Spanish societies, then the increase in profit opportunities might not have prompted the corrupt behavioral responses we find. Also, Andvig and Moene (1990) argue that "corruption may corrupt" since the profitability of corruption may be related to its frequency. This would violate the independence assumption, a fundamental Poisson postulate, and put into question our baseline estimates from the negative binomial model. This motivates the robustness check involving linear OLS regression (Mroz, 2012). More specifically, to test whether our results are robust to relaxing the "memoryless" property in the manner zoning corruption events arise, we estimate linear regression models. We re-parameterize zoning corruption to a continuous variable because count data fit linear models poorly. The dependent variable now expresses zoning corruption per 100,000 individuals. The results are shown in Table 7 below.

Table 7: Linear Regression – Greece, 2003-2008; Spain, 2006-2008

Variables	(1) Greece, Lagged Housing Prices	(2) Spain, Lagged Housing Prices
Housing Prices (€/m^2)	0.000380*	0.00265*
	(0.000203)	(0.001475)
Unemployment	0.0275	−0.1431
	(0.0190)	(0.2830)
Population Density	−0.0682***	−0.1149*
	(0.00686)	(0.0614)
Population	−0.0000065	0.000014*
	(0.0000082)	(0.0000007)
Education	−0.0135	0.1402*
	(0.0228)	(0.07603)
Regional Dummies	YES	YES
Constant	98.42**	−101.84*
	(37.06)	(52.26)
Observations	702	204
R-squared	0.625	0.401

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Standard errors clustered at the regional level (in parentheses).

Columns (1) and (2) present estimates from the lagged housing price specification for Greece and Spain, respectively. Our baseline results are robust to the linear specification. We report a positive and significant association between housing prices and zoning corruption. Figures 20 and 21 use binned scatterplots ("binscatters") to provide a graphical representation of the OLS estimates for Greece and Spain, respectively.

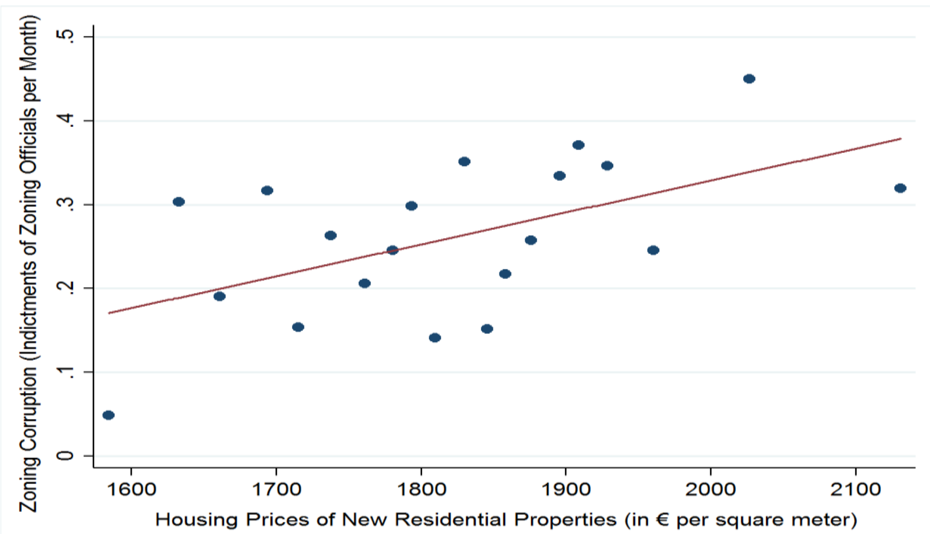

Figure 20: Graphical Representation of OLS Estimates for Greece using a Binscatter

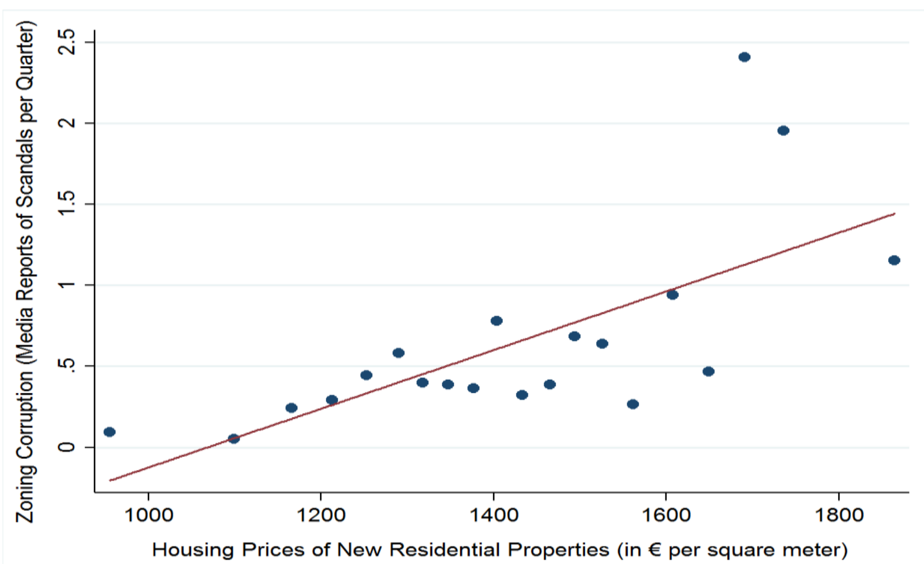

Figure 21: Graphical Representation of OLS Estimates for Spain using a Binscatter

Graphically, it is clear that as housing prices increased, zoning corruption was amplified. The "binscatters" also indicate that this relationship is more significant in the case of Spain, in a statistical manner. Our baseline results are robust even when we assume away the "memoryless property" in the incidence of zoning corruption.

In our fourth robustness check, we employ Generalized Method of Moments (GMM) estimators to explore the importance of "dynamic panel bias". As previously mentioned, this is a plausible assumption because zoning corruption may not be independent over time. If current values are related to their own past realizations, lagged values of zoning corruption should be used as explanatory variables. Then, OLS estimates are subject to "dynamic panel bias" because lagged values of zoning corruption are endogenous to the fixed effect in the error term (Roodman, 2006). However, if T is large, dynamic panel bias becomes insignificant, and the OLS estimates are reliable. Also, GMM estimators permit independent variables that are not strictly exogenous, meaning correlated with past and possibly current realizations of the error. We can test the robustness of our results under the assumption that housing prices are endogenous to zoning corruption. Given the observational nature of our study, this allows us to assess the causal interpretation of the housing price coefficient of our baseline results.

Since GMM estimators assume a linear functional relationship, we use zoning corruption per 100,000 individuals, a continuous dependent variable. We take first differences to eliminate unobserved region-specific fixed effects and use lagged instruments to correct for simultaneity in the first-differenced equations (Arellano and Bond, 1991). This yields the difference GMM estimator. However, if the dependent variable follows approximately a random walk, then this estimator will perform poorly. Specifically, past levels of zoning corruption will not be informant of future changes and, therefore, untransformed lags will be weak instruments for transformed variables (Roodman, 2006). To increase efficiency, we difference the instruments to make them exogenous to the fixed effects (Arellano and Bover, 1995; Blundell and Bond, 1998). This system GMM estimator will be valid assuming that differences in any instrumenting variables are uncorrelated with the fixed effects. The system GMM estimator is superior to the difference GMM in terms of root mean squared error. However, the latter is important primarily for prediction, in which the negative binomial model does already an excellent job as Figure 16 and 17 indicate. Bun and Windmeijer (2010) show that the 2SLS biases can be greater than the OLS biases in finite samples due to a weak instruments problem. Therefore, we employ both the two-step difference GMM and the two-step system GMM estimators with robust standard errors and small-sample adjustments (Windmeijer, 2000).

GMM estimators are designed for panels with few time periods (T) and a large number of cross-sectional units (N). Generally, the instrument count should not exceed the number of regions because as T increases the number of instrumenting variables tends to explode and the Arellano-Bond autocorrelation test becomes imprecise (Roodman, 2006). To conform with this rule of thumb, we collapse the panel of Greece to annual frequency (7 time periods) and the panel of Spain in semi-annual frequency (6 time periods) and do not make extensive use of lags as instruments. Autocorrelation is not a concern in any of the specifications according to the Arelland-Bond test for an AR(1) and AR(2) in the first differences of the lagged instruments. Table 8 below shows the results of the two-step system GMM estimations for Greece and Spain, collectively.

Table 8: GMM Estimation of Dynamic Panel Model – Greece, 2003-2008, annual; Spain, 2006-2008, semi-annual

Variables	(1) Greece, Baseline	(2) Greece, System GMM	(3) Spain, Difference GMM	(4) Spain, System GMM
Corruption Rate per 100k, 1-Year Lag	-0.2744518	-0.215224	-0.2110873	0.0984775^{*}
	(0.1085221)	(0.2141529)	(0.2257008)	(0.0506335)
Housing Prices	0.0001923	0.0006047	0.0021948^{*}	0.0001186^{***}
	(0.0003862)	(0.0003569)	(0.0010582)	(0.0000289)
Unemployment			-0.2563092	
			(0.1831954)	
Density	-0.0266624		-0.0135126	-0.0003502^{**}
	(0.0246729)		(0.0157812)	(0.0001322)
Education	-0.0078916		0.0517698	
	(0.0509265)		(0.073579)	
Constant		-0.5610423		
		(0.6461527)		
Number of Regions	13	13	17	17
Number of Instruments	10	13	10	17
AR(1) in First Differences	$Pr > z = 0.285$	$Pr > z = 0.272$	$Pr > z = 0.088$	$Pr > z = 0.018$
AR(2) in First Differences	$Pr > z = 0.611$	$Pr > z = 0.515$	$Pr > z = 0.593$	$Pr > z = 0.095$
Hansen Test of Overidentification Restrictions	$Pr > \chi^2 = 0.455$	$Pr > \chi^2 = 0.328$	$Pr > \chi^2 = 0.142$	$Pr > \chi^2 = 0.272$

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Two-step GMM Estimation with robust standard errors and small-sample adjustments (Windmeijer correction).

Column (1) shows the estimates of a specification for Greece where zoning corruption is endogenous in time, housing prices are endogenous, and augmented by density and education as IV-style instruments. The coefficient of housing prices in column (1) is positive but not statistically significant. The rule of thumb is preserved and we do not overidentify using few instruments. Still, the latter are not particularly strong as the Hansen test indicates. Column (2) shows the estimates of a specification for Greece where zoning corruption is endogenous in time, housing prices are not strictly exogenous, and only year dummies serve as additional instruments. The coefficient of housing prices is positive and almost significant at the 10 percent level. The estimates in column (2) are more reliable since the instruments are stronger and the instrument count does not exceed the number of regions in Greece. When we introduce more controls (e.g. population density or educational level) the relationship of interest does turn significant. However, this comes at the expense of instrument power as the instrument count then exceeds the number of regions in Greece. We refrain from violating this minimal rule of thumb of GMM estimation but note that the positive and significant relationship between zoning corruption and housing prices virtually holds.

Column (3) presents the difference GMM estimates for Spain. We specify zoning corruption to be endogenous in time, and housing prices as an endogenous regressor. We use density and education as instruments from the baseline controls. Unemployment is considered predetermined but not strictly exogenous because its inclusion in the instrument pool severely reduces the instruments' strength. The instrument count in this specification is well below the number of regions and, as a result, this set of instruments is the strongest among all estimations. Accounting for any dynamic panel bias, we retrieve a positive and statistically significant relationship between housing prices and zoning corruption. Column (4) shows the estimates of the system GMM estimator for Spain where zoning corruption is endogenous in time, housing prices are treated as an endogenous regressor, and only population density is added as an instrument. The coefficient of housing prices is positive and very significant. The instrument count in this estimation is just equal to the number of regions and, as a result, density is not a weak instrument[16]. The Arellano-Bond tests show that autocorrelation is only minor.

The GMM estimates suggest that our baseline results are fairly robust to dynamic panel bias, even when the panels are collapsed to fewer time units. The first takeaway is that the positive association between zoning corruption and housing prices is preserved. Second, even if housing prices are assumed to be endogenous, results do not change substantially. The statistical significance in the case of Greece dissipates when parameters are estimated using the annual

[16] Note that we drop the constant from the system GMM estimation to achieve equal instrument counts to regions.

panel. However, any dynamic panel bias should be negligible in the longer, monthly panel that served as the basis for estimation of the negative binomial and earlier linear models. Thus, we cannot interpret the insignificant coefficient as evidence against our baseline results for Greece. The positive and statistically significant relationship between zoning corruption and housing prices in Spain is robust to the GMM estimations. This is encouraging for the validity of our baseline findings because the "small T, large N" panel for Spain might have leaked dynamic panel bias in those estimates. Clearly, GMM estimators are more enlightening in the case of Spain (T=12, N=17) whereas the fixed effects estimator is more fitting for the baseline panel of Greece (T=72, N=13).[17]

Finally, we conduct two country-specific robustness checks that address potential measurement error.

In the case of Greece, we explore whether our results are subject to measurement error of the housing price information. Recall that the housing price observations are the result of our interaction of national time series with local cross-sections of housing prices. Here, we consider a sub-sample of Greek regions for which the official housing price information matches our geographical disaggregation perfectly.[18] Our baseline findings are robust. To preserve space, estimates are not shown. Even in this restricted sample, the positive relationship between zoning corruption and housing prices persists.

[17] The unconditional negative binomial model is susceptible to a potential parameters problem which there is no way to explore. When using the dummy variable approach, estimates of negative binomial models in short panels might be inconsistent due to the incidental parameters problem (Cameron and Trivedi, 1998, p.282). These perform well for sample size of at least 200 observations, a sample size requirement met only marginally in the panel for Spain. In light of potentially inconsistent baseline estimates for Spain, the system GMM estimates for Spain become even more important.

[18] These are the regions of Attica and Central Macedonia that account for more than 50 percent of the country's population and include the mass of corruption counts.

In the case of Spain, we conduct a sensitivity analysis of our findings across all methods employed by introducing deeper geographical disaggregation. We employ information at the provincial level (Eurostat NUTS-level 3), a more local level of governance relative to the regional level (Eurostat NUTS-level 2). Specifically, the number of cross-sectional units almost triples from 17 to 50. As a result, we are able to exploit substantially more cross-sectional variation in the identification of model parameters. Moreover, this geographical disaggregation makes the causal relationship between housing prices and zoning corruption more tenable. By design, the use of smaller geographical units reduces the measurement error in the association of an instance of zoning corruption with a particular housing price level. The natural mapping of zoning corruption to housing becomes less crude by almost 300%.

Data limitations prevent us from carrying out a similar geographical robustness check for Greece as zoning corruption information is available only at the regional level.

Table 9 below revisits our findings from count data models using provincial-level information. Specifically, columns (1) and (2) present results for the negative binomial and zero-inflated model regressions for housing price in levels. Column (3) replicates our baseline negative binomial model of lagged housing prices, whereas in column (4) we implement its zero-inflated variant. Note that the set of controls varied slightly relative to the regional analysis. The Eurostat does not provide a wealth of provincial information (NUTS level 3) which severely restricts our choice of control variables. As a result, we use GDP level per capita in the place of education and unemployment rate to proxy the level of development in the province. We maintain population density as a control variable and use population as the exposure variable.

Table 9: Count Data Models – Spanish Provinces

Variables	(1) Neg Bin	(2) Lag Neg Bin	(3) Zero-inflated	(4) Lag Zero-inflated
Housing Prices	0.00636***	0.00640***		
	(0.00161)	(0.00142)		
6-quarter Housing Price Lag			0.00165*	0.00175*
			(0.00090)	(0.00090)
Density	−0.00677	−0.00674	−0.00626	−0.00598
	(0.00568)	(0.00910)	(0.00831)	(0.00944)
GDP per capita	0.00036**	0.00038**	0.00050***	0.00050***
	(0.00014)	(0.00016)	(0.00016)	(0.00019)
Constant	−30.00345***	−16.51667***	−25.59613***	−11.86371***
	(2.68045)	(2.67890)	(3.09566)	(3.28132)
Number of Regions	50	50	50	50

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Standard errors clustered at the provincial level in columns (1), (2)

The positive and statistically significant relationship between zoning corruption and housing prices persists at the provincial panel. We do not report any substantial differences between the negative binomial models and their zero-inflated counterparts. This is due to the fact that the two specifications are not different at all as the Vuong test statistic indicates. Specifically, with a value of 0.06 and 0 for the contemporaneous and lagged specifications, the two models are virtually the same.

Next, using the provincial panel we estimate linear models where the dependent variable takes the form of a rate: zoning corruption per 100,000 individuals. It should be noted that the increased cross-sectional variation naturally leads to a substantial improvement in our estimates obtained via panel methods. This is particularly true for the GMM models that are ideal for "small T, large N" panels. Thus, the findings of the following fixed effects and dynamic panel models in Table 10 are preferable to the ones for Spain in Table 8 and 9.

Table 10: Linear Models – Spanish Provinces

Variables	(1) Contemporaneous FE	(2) Lagged FE	(3) Difference GMM	(4) System GMM
Corruption Rate per 100k, 1-Year Lag			0.00557	0.00557**
			(0.09217)	(0.03273)
Housing Prices	0.00044***	0.00018*	0.00345	0.00022*
	(0.00016)	(0.00010)	(0.00332)	(0.00012)
GDP per capita	0.000028**	0.000029**	0.000104	−0.000014***
	(0.000012)	(0.000011)	(0.000100)	(0.000005)
Density	−0.00083**	−0.00045*	−0.00168	−0.00044*
	(0.00032)	(0.00025)	(0.00131)	(0.00022)
Population	0.00000004	−0.00000016	−0.00000029	
	(0.00000022)	(0.00000028)	(0.00000094)	
Constant	−1.19204186***	−0.62168329**		
	(0.21568277)	(0.24577716)		
Number of Regions	50	50	50	50
Number of Instruments	0	0	23	43
AR(1) in First Differences			$Pr > z = 0.003$	$Pr > z = 0.001$
AR(2) in First Differences			$Pr > z = 0.386$	$Pr > z = 0.332$
Hansen Test of Overidentification Restrictions			$Pr > \chi^2 = 0.545$	$Pr > \chi^2 = 0.579$

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Two-step GMM Estimation with robust standard errors and small-sample adjustments (Windmeijer correction).

Again, our results from the linear models are insensitive to the geographical disaggregation of the zoning corruption and housing price information. We recover a positive and statistically significant result between the two in columns (1), (2) and (4). Controlling for unobservable characteristics that are unique to a province, we show that both contemporaneous and lagged housing prices substantially increase zoning corruption. The dynamic panel methods in columns (3) and (4) address serial correlation. The coefficient of housing prices becomes insignificant for the difference GMM estimator that relies primarily on corruption lags. However, when we improve GMM estimation by also exploiting controls as instruments we are able to replicate the earlier system GMM results for Spain.

Overall, the robustness of our baseline results to the use of provincial information indicates they are not subject to significant error in space in the measurement and attachment of zoning corruption to local housing prices.

6. Concluding remarks

The understanding of the determinants of corruption and its effects on the economy is a fast-growing field of study in public and urban economics. We follow an emerging literature of the empirical studies of corruption that exploits shocks to corruption incentives brought about by nature or by policy. We use the adoption of the euro and its associated windfall of cheap mortgage credit in Greece and Spain to examine the effect of housing price bubbles on zoning corruption. The increased supply of loanable funds at cheaper rates helped expand economic activity for a decade but also postponed important structural reforms. The latter had significant side effects that became imminent years later.

The empirical analysis relies on objective measures of zoning corruption in Greece and Spain, measured at the regional level. We estimate negative binomial models and test the robustness of our findings using zero-inflated and linear regression models. We document a positive and significant relationship between housing prices and zoning corruption. The positive association is robust to a series of checks that include zero-inflated variants of the negative binomial model, and linear models that address model misspecification, omitted variable and dynamic panel bias.

Essay 3: The Short- and Long-term Effects of State Medicaid Expansions on Mortality: Evidence from a Synthetic Control Design

1. Introduction

Does public provision of free or low-cost health insurance to the uninsured improve their health outcomes? This is an important health policy question because expanding Medicaid is currently the only active policy margin of the Affordable Care Act. Theoretically, Medicaid eligibility is expected to reduce mortality by improving access to medical care which, in turn, may lead to improved physical and mental health. However, very little evidence exists to date as to whether this is the case for adults.

Finkelstein et al. (2013) find no evidence of an effect of a randomized Medicaid expansion in Oregon on adult mortality, but given the relatively small number of deaths in their sample they may have lacked sufficient statistical power to detect reasonably sized effects. Sommers et al. (2012) conduct a difference-in-difference analysis comparing three states that expanded Medicaid in the 2000s to three adjacent control states, finding a significant reduction of 6.1 per cent in adjusted all-cause mortality in the expansion states. However, Kaestner (2012) questions the validity of this study's difference-in-differences design because pre-intervention mortality trends in the treatment and control states exhibited statistically significant differences. Sommers (2017) uses propensity-score matching to find that all-cause and healthcare-amenable mortality at the county level declined by 6 and 6.7 per cent, respectively. As results are driven by New York, it is unclear to what extent mortality reductions can be causally attributed to Medicaid, the divergence in HIV-related mortality to pre-expansion trends or other idiosyncratic factors to mortality in New York. Furthermore, disaggregation to the county level artificially inflates the sample size as the expansions were implemented at the state level, overstating levels of significance. Given the small number of clusters (7), the issue is not addressed by clustering standard errors by state and other techniques that introduce randomization inference are preferable (Bertrand, Duflo, and Mullainathan, 2004).

The purpose of this paper is to examine the long- and short-term effects of all major state Medicaid expansions to adults in the 1990s and 2000s on all-cause, state-level adult mortality. First, we implement a difference-in-differences (DID) design to estimate the effect of all state Medicaid expansions on state-level, all-cause mortality. To identify the causal effect of state Medicaid expansions on mortality, we employ a synthetic control design (SCD) that creates a control group as a weighted average of non-expansion states that most closely resembles the pre-intervention mortality levels of expansion states (Abadie, Diamond, and Hainmueller, 2010). Finally, to shed light on what drives the observed heterogeneity in Medicaid's efficacy in reducing mortality levels across states, we assess to what extent the expansions of eligibility increased actual coverage. Specifically, we examine how many more individuals enrolled in Medicaid in each state following expansion. This allows us to gain a deeper understanding of the reasons behind the underlying heterogeneity across states.

The contribution of this work is threefold. First, we contribute a novel study of the dynamics of the effects of state Medicaid expansions to adults on all-cause adult mortality. Second, our evaluation is comprehensive since it examines all state Medicaid expansions in the 1990s and 2000s regardless of scope of coverage. Each state-specific finding is complemented with corresponding estimates of the size of each Medicaid expansion in terms of the levels of total number of enrollees (total coverage). Finally, we provide a sensitivity analysis of the inferential assumptions of pre-existing, quasi-experimental estimates for Arizona, New York, and Maine.

Our results suggest that pre-existing estimates of mortality reductions due to Medicaid expansions are largely overstated. Mortality reductions due to Medicaid, if any, were more pronounced in the long-run. We find that mortality reductions were possible only in states that recorded substantial enlargement of their total Medicaid population and when comprehensive insurance benefits are offered.

The rest of the paper is organized as follows. Section 2 reviews the literature discussing the effectiveness of public health provision on curbing mortality. Section 3 discusses the data and the methods employed. Section 4 presents the baseline results and, further, explores heterogeneous effects of Medicaid across sub-groups of the treated population. Section 5 concludes.

2. Literature on Public Health Care Coverage Effects on Health Outcomes

The experiment that could ideally be used to estimate this treatment effect would involve a randomized control trial of a Medicaid expansion in a state, similar to the Oregon Medicaid experiment (Finkelstein et al., 2013). But, by and large, the empirical findings are somewhat mixed, possibly misguided, and outdated as increasingly more states have expanded Medicaid eligibility for adults in the past decade at varying levels. Seminal work has found mortality reductions following infant and children Medicaid expansions in the 1980s (Currie & Gruber, 1996a; Currie & Gruber, 1996b). However, others have found negligible effects (Epstein & Newhouse, 1998; Howell, 2001; Levy & Meltzer, 2004).

Compared to the access and utilization literature, the study of Medicaid's effect on health is far less conclusive (Howell, 2001; Howell & Kenney, 2012). This is largely because health is a difficult construct to capture and it is affected by a large number of factors that are typically not measured, but are correlated with enrollment in Medicaid, and there is a clear recursive relationship between health and enrollment that can easily bias standard OLS regression estimates (Levy & Meltzer, 2008). Nonetheless, a handful of researchers have attempted to isolate Medicaid's independent effect on health. Health in these studies has been operationalized as birth weight, self-reported health status, disability, and mortality.

Lykens and Jargowsky (2002) find that Medicaid expansions during the late 80s and early 90s appeared to decrease the number of acute health conditions and functional limitations among low-income White children under the age of 15, but not among other racial groups. In many studies, results for minority racial groups are hampered by small sample sizes. A priori it is reasonable to expect greater gains by persons from minority racial groups due to the correlation of poverty and race. Within the context of the Oregon experiment, Baicker et al. (2013) consider a subset of health outcomes potentially affected by Medicaid coverage that are important contributors to morbidity and mortality; namely, hypertension, high cholesterol levels, diabetes, and depression. They report detectable improvement only in the case of depression (but not self-reported happiness); however, their investigation was underpowered by the relatively small number of patients with these conditions. Black et al. (2017) compare the mortality hazard rates of near-elderly individuals that received insurance (public or private) in 1992 to those uninsured over a 20-year period. Based on waves of the Health and Retirement Study, they find that the mortality hazard rates of the two groups are not substantially different. They interpret their findings as evidence that "prior studies have greatly overestimated the health and mortality benefits of health insurance for the uninsured". Many other studies have examined the connection between Medicaid and health, but due to weak study designs that fail to account for unobserved confounding and selection bias, little confidence can be placed in their results.

Hadley (2003) provides a summary the literature on the impact of health insurance on mortality that compares an insured to an uninsured study population. The pattern that emerges from existing studies that may suffer from methodological flaws is that Medicaid does indeed improve health; however, the effect is often modest, but notably stronger among younger and more disadvantaged populations.

3. Empirical Analysis

This section first provides a brief discussion of the data we use. Then, we offer a description of the two identification strategies we undertake. Finally, we proceed to specify the empirical models we estimate our coarsened exact matching procedure in detail.

3.1 Data

Mortality data are obtained from the Compressed Mortality File of the Centers for Disease Control and Prevention (CDC) and are available from 1986 to 2010. Socio-economic characteristics are obtained from the Bureau of Labor Statistics whereas demographic information comes from the Census. We employ state-level information such as average unemployment rate, median income, population, per cent of the population in poverty, married, male, 25-years-of-age and above with a Bachelor's degree, in 20-34, 35-44, 45-54, 55-64 age groups, white, Hispanic or other race. Information on the total number of Medicaid beneficiaries is coming from the University of Kentucky's Center for Poverty Research (UKCPR) National Welfare Data that combines resources from the Centers for Medicare & Medicaid Services, the Medicaid Budget and Expenditure System (MBES) and Quarterly Medicaid Enrollment and Expenditure Reports. It should be noted that this figure includes all Medicaid eligible subpopulations such as children, not just adults, in June. We net out the (rounded 12-month fiscal year average) number of participants in each state's Women, Infants and Children (WIC) program to approximate the number of adult Medicaid beneficiaries, even though that still does not exclude teenage children.

3.2 Identification

Difference-in-Differences Design

We begin with a simple but likely naïve difference-in-differences design (DID). For control states, we first use all non-expanding states and then choose region-specific ones. This shows the effect of Medicaid on mortality in all states that expanded over the span of two decades relative to the national and regional averages in non-expanding states. Implicitly, this approach uses a linear combination of the untreated units with coefficients that sum to one. In other words, the regression estimator could also be considered as a weighting estimator with weights that sum up to one. However, there is no restriction on the values these weights may take, enabling extrapolation outside the support of the data (Abadie, Diamond, and Hainmueller, 2014).

The identifying assumption for a causal interpretation of the parameter of interest in the DID design is that, conditional on the other covariates, changes in the mortality pre-intervention would have been the same in expansion and non-expansion states in the absence of the intervention. The validity of the DID is confirmed in Figures I.1 through I.3 in Appendix I that clearly show that the assumption of common trends does hold.

Synthetic Control Design

Next, we turn to the SCD, a data-driven procedure introduced by Abadie et al. (2010). In this case, controls are weighted averages of states that did not expand Medicaid that most closely resemble pre-intervention mortality predictor values in states that did expand. The intuition behind the approach is that a combination of units often provides a better comparison for the unit exposed to the intervention than any single unit alone. This offers reasonable counterfactuals of mortality rates in treated states if the expansion had not occurred. The procedure reweights the control group such that the synthetic control expanding states match observable characteristics and pre-intervention mortality values. Unlike the regression approach, the SCD estimator assigns weights ranging between zero and one to shield off extrapolation bias. This prevents estimation of "extreme counterfactuals" by forcing to show the proximity between treated and control.

For identification, a SCD makes two implicit assumptions. First, there exists no interference between units that comprise the donor pool (Rosenbaum, 2007). Given the state-specific nature of the expansion and our state-year panel, this holds by design. More importantly, unbiasedness requires that there exists a synthetic control with a vector of weights

$$\boldsymbol{W}^{*} = \left\{ w_2^*, \dots, w_{J+1}^* \middle| w_j^* \geq 0 \; for \; j = 2, \dots, J+1 \, , w_2^* + \cdots w_{J+1}^* = 1 \right\} \;(1)$$

a vector $\boldsymbol{Z}_i$ of observed covariates (unaffected by the intervention), and a vector $\boldsymbol{\mu}_i$ of unknown factor loadings such that

$$\sum_{j=2}^{J+1} w_j^* \boldsymbol{Z}_j = \boldsymbol{Z}_1, \sum_{j=2}^{J+1} w_j^* \boldsymbol{\mu}_j = \boldsymbol{\mu}_1 \;(2)$$

Even if equation (3) holds only approximately for a synthetic control that fits well but not perfectly the treated state, the bias of the estimator will tend to zero as the number of pre-intervention treatment periods increases. In our sample, the post-intervention period commences the first full year after the Medicaid expansions since we drop the transitional year of the expansion.[1] It extends for 11 years while the pre-intervention period is defined as the six years leading up to the expansion. This is the maximum number of pre- and post-intervention years that can be accommodated for each state. By using the same number of pre- and post-treatment periods for each state, we can easily transition into the national-level estimation where we pool all treated and all control units together.

Inference is based on placebo tests in time and space. The placebo test in space estimates the probability of finding mortality reductions of the magnitude of the observed reduction in treated states under a random permutation of the expansion to generate the p-values of our estimated effects. In addition, leave-one-out sensitivity checks show that our findings are robust to the exclusion of any particular control state from the donor pool.

3.3 Estimation

Difference-in-Differences Implementation

We estimate the following specification in the DID design:

$$Mortality\ Rate_{it} = \beta_0 + \beta_1 MedicaidExpansion_{lt} + \boldsymbol{\beta_1 X_{ijk}} + \boldsymbol{\beta_2 State-level\ Factors_{lt}} + \beta_3 ImputationDummy_{ijklt} + \mu_t + \Omega_l + \varepsilon_{ijklst}\ (3)$$

[1] In doing so, we lose information from the months before and after the Medicaid expansion pertaining to the transitional years.

where i indexes age, j race, k gender, l state, s state-treatment-control pairing, and t year. $MedicaidExpansion_{lt}$ indicates whether state s expanded Medicaid in year t. $\boldsymbol{X_{ijk}}$ is a vector of demographics (percentage of the population that is white, Hispanic, percentage of the population that is female, and age indicators for groups 20-34, 35-44, 45-54, and 55-64 years old). $\boldsymbol{State - level\ Factors_{lt}}$ include state-year-specific poverty rate, median income (measured in constant 2012 $), population size, unemployment rate, and percentage of the population with a high school diploma and percentage of the population that is married. The $ImputationDummy_{ijklt}$ is a binary variable indicating whether the dependent variable was censored by the CDC for a death count between 1 and 5. Finally, we include year and state fixed effects μ_t and Ω_l, respectively. We adjust our estimates by state population weights. The adjustment gives a greater weight to mortality rates from states with relatively larger populations. However, panel data methods require that the weight remains constant across all years for a particular state (panel unit). To this end, we rely on constant weights that reflect each state's population as measured by the 2001 Census.[2] The parameter of interest is β_1 that denotes the effect of the state Medicaid expansion on adult mortality.

Synthetic Control Design Implementation

To implement the SCD numerically, we solve a nested optimization problem. First, we minimize the multivariate distance between values of mortality predictors $\boldsymbol{X_{it}}$ of states expanding Medicaid to their corresponding synthetic controls subject to weight constraints $\boldsymbol{W}$ ranging [0,1]. Then, we introduce matrix $\boldsymbol{V}$ that applies different weights to $\boldsymbol{X_{it}}$ depending on the

[2] Alternatively, we assume that each state-year combination is a separate observation and cluster standard errors at the state level, so that states may have a larger weight in 2007 than in 1992. Estimating via indicator variables for state and years, gives rise to only very minor increases in the standard errors that does not change statistical significance of the point estimates.

covariate's predictive power of mortality. Our choice of $\boldsymbol{V}$ which is the solution to the inner minimization problem assigns weights that minimize the mean square error of the synthetic control estimator.

Based on weights calculated by the SCD, our main results come from estimation of the factor model in equation (4) below:

$$Mortality\ Rate_{it} = \delta_t + \sum_{g=-j}^{0} \mu_g MortalityRate_{it+g} + \boldsymbol{\beta_1 X_{it}} + \lambda_t \mu_i + \varepsilon_{it} \quad (4)$$

where $j = 1, 2, \dots, 6$ (but in our preferred specification $j = 1$), i indexes state and t year. X_{it} now contains a set of observed state-year demographic and socio-economic information; namely, percentage of the population that is female, aged 20 to 34, aged 35 to 44, aged 45 to 54 and aged 55 to 64, black, Hispanic or white, under the Federal poverty line, unemployed, aged over 25 with a high-school diploma, that is married and finally the level of median income (measured in constant 2012 $). The term $\lambda_t \mu_i$ captures heterogeneous responses to multiple observed factors. The parameter δ_t is an unobserved, common, time-dependent factor. Finally, $\sum_{g=-j}^{0} MortalityRate_{it+g}$ is a collection of lagged dependent variables over the pre-expansion years that start from the previous full calendar year preceding a Medicaid expansion ($MortalityRate_{it-1}$) up to the earliest pre-treatment period year ($MortalityRate_{it-6}$). The model selection is data-driven so that the final specification is resulting into the lowest Mean Squared Prediction Error (MSPE).. Following Kaul et. al (2016), our baseline specification includes only the last value of the pre-treatment outcome lags in addition to the demographic and socio-economic controls above.[3] We also experiment with a specification that includes either only pre-treatment outcome lags or the average across all pre-treatment outcome lags and

[3] Kaul et. al (2016) show that including all pre-treatment year lags as predictors does result optimal pre-treatment fit but also renders all other covariates irrelevant that may lead to substantial bias.

controls. Our point estimate of interest is defined as the average difference between the state-level adult mortality in a state with a Medicaid expansion during the years after it took effect to its synthetic control representation and measured in levels; i.e., difference in the level of the mortality rate per 100,000 adults.

Finally, we estimate an equivalent factor model to the one specified in equation (4) where we substitute the outcome variable to be the number of total Medicaid enrollees in a state i in year t:

$$Log(TotalMedicaidEnrollees)_{it} = \delta_t + \sum_{g=-j}^{0} \mu_g Log(TotalMedicaidEnrolles)_{it+g} + \boldsymbol{\beta_1 X_{it}} + \lambda_t \mu_i + \varepsilon_{it} \quad (5)$$

The notation denotes the same terms as in equation (4). Similarly, the final model is a combination of the pre-treatment log-level of total Medicaid enrollees in a state preceding its Medicaid expansion ($g = -1$) and the demographic and socio-economic controls. Given the log-levels specification, the point estimate of interest is now interpreted in per cent terms; i.e., per cent change in mortality rate per 100,000 adults

4. Results

Difference-in-Differences Estimates

For a graphical representation of the state-specific DID results, refer to Figures I.4.A through I.14.A in Appendix I. Table 11 presents population-weighted estimates of the effect of state Medicaid expansions on the average US mortality per 100,000 individuals. Due to the large number of adopting states and their geographical dispersion this is an interesting finding from a policy perspective. It represents a nationwide estimate of the projected impact of expanding Medicaid on all-cause mortality, overall.

Table 11: Aggregate DID Estimates of Effect of State Medicaid Expansions on Mortality			
	(1)	(2)	(3)
	Full Sample	Contiguous U.S.	Hand-matched Sample
VARIABLES			
Medicaid Expansion	-22.532**	-23.078**	-22.071**
	(9.432)	(9.457)	(8.543)
% Female	-18.891**	-19.569**	-22.401*
	(8.182)	(9.441)	(11.219)
% Hispanic	-5.149***	-5.037***	-4.789***
	(0.942)	(1.056)	(1.135)
% White	2.303	2.362	5.147*
	(2.395)	(2.447)	(2.871)
% Aged 20-34	-4.531*	-4.479	-5.001
	(2.646)	(2.672)	(2.973)
% Aged 35-44	-11.278***	-10.983***	-9.630**
	(3.447)	(3.561)	(4.127)
% Aged 45-54	1.715	1.411	3.045
	(5.649)	(5.622)	(6.360)
% Aged 55-64	11.854***	12.635***	14.231***
	(3.257)	(3.362)	(4.051)
Unemployment	0.888	0.811	1.214
	(0.733)	(0.731)	(0.970)
Poverty	0.193	0.236	-0.636
	(0.478)	(0.477)	(0.545)
Log Average Income	44.325**	45.809**	30.293
	(18.979)	(19.405)	(19.261)
% Married	0.565*	0.524	0.378
	(0.330)	(0.342)	(0.326)
% H-S Diploma	-13.667***	-13.606***	-14.155***
	(1.936)	(1.936)	(2.885)
Log State Population	96.343**	93.294**	95.395**
	(45.311)	(45.417)	(43.514)
Constant	-422.414	-372.766	-345.099
	(1,168.047)	(1,214.235)	(1,371.818)
Observations	1,200	1,176	840
R-squared	0.777	0.782	0.829
Number of state	50	49	35

Notes: Estimates in column (1) are based on the DID design using the full donor pool of 39 control states; namely, excluding Massachusetts. In column (2) we further discard Hawaii and focus on Medicaid expansions in states in the contiguous US. Column (3) provides results from a hand-matched donor pool of neighboring/regional controls. Time and state fixed effects included in all specifications. Robust standard errors in parentheses, clustered at the state level; *** p<0.01, ** p<0.05, * p<0.1

Using the sample of the entire US excluding Massachusetts, we find that Medicaid had a statistically significant and negative effect on average, all-cause, US mortality rates. Excluding Hawaii or using hand-matched control states based on geographical proximity to the expanding state, does not have a substantial impact on the magnitude of our point estimates or our standard errors. The previously documented negative and significant relationship between Medicaid expansion and mortality is very robust to a different set of samples.

Next, we narrow-in on each state that expanded Medicaid. Table 12 presents population-weighted estimates of the effect of the Medicaid expansion on all-cause mortality, on a state-by-state basis.

Table 12: State-Specific DID Estimates of Effect of State Medicaid Expansions on Mortality

	(1)	(2)	(3)
VARIABLES	AZ	HI	ME
Medicaid Expansion	6.890	9.628	8.241*
	(3.775)	(25.409)	(3.122)
Observations	56	42	56
R-squared	0.9629	0.9866	0.8844
# Control States	4	3	4
	(4)	(5)	(6)
	MN	NM	NY
Medicaid Expansion	0.213	-0.755	-18.6188***
	(3.422)	(7.833)	(2.522)
Observations	70	56	70
R-squared	0.9313	0.9595	0.9385
# Control States	5	4	5
	(7)	(8)	(9)
	OR	TN	UT
Medicaid Expansion	-11.943	1.913	9.101***
	(4.968)	(1.754)	(1.239)
Observations	42	140	70
R-squared	0.9982	0.9189	0.8281
# Control States	3	10	5
	(10)	(11)	

	VT	WA
Medicaid Expansion	-21.640**	-18.21257**
	(6.372)	(2.804)
Observations	56	42
R-squared	0.8307	0.9956
# Control States	4	3

Notes: Estimates are based on the following hand-matched donor pool based on geographical proximity, AZ: CA, NV, CO; OR & WA controls: CA, ID; MN controls: WI, IA, ND, SD; UT controls: NV, WY, ID, CO; NM controls: TX, CO, OK; ME & VT controls: NH, RI, CT; NY: PA, NJ, CT, RI; TN: KY, AK, MO, IL, MS, AL, GA, NC, VA. Robust standard errors in parentheses, clustered at the state level; *** p<0.01, ** p<0.05, * p<0.1

The state-specific estimates reveal that the large, negative, and statistically significant findings from the US aggregate analysis above was primarily driven by New York and to a certain extent by Washington. The state-specific analysis for New York confirms that the state experienced large reductions in mortality following Medicaid expansion. But, due to the fact that New York is so populous relative to other Medicaid-expanding states, it is assigned such a large weight that its influence dominates any competing effects from other states. For instance, mortality rate increases followed the Medicaid expansions in Maine and Utah. However, their population sizes dwarf the ones of New York and Washington which, subsequently, depress their influence on the estimated average treatment effect of Medicaid on U.S. overall average mortality rate. Combined with the fact that the only other large and statistically significant findings are mortality reductions in Vermont and Washington, it is not surprising that the effect of Medicaid expansions on average U.S. mortality per 100,000 individuals is negative and similar in magnitude to ones of these states. Oregon also experienced mortality reductions but not at conventional levels of statistical significance. Finally, mortality rates were not affected by Medicaid expansions in Hawaii, Minnesota, New Mexico and Tennessee.

Synthetic Control Design Estimates

A graphical illustration of the SCD findings is given in Figures I.4.B through I.14.B in Appendix I. Table 13 below presents the estimated average treatment of each Medicaid expansion over time, state-by-state. Table J.2 in the Appendix illustrates the dynamic treatment effect of each Medicaid expansion, year-by-year, over the span of seven post-treatment years.

Table 13: State-Specific SCD Estimates of Effect of State Medicaid Expansions on Mortality

VARIABLES	(1) AZ	(2) HI	(3) ME
Medicaid Expansion	3.009	10.632	5.611
# Chosen Donor Pool States	4	3	4
	(4) MN	(5) NM	(6) NY
Medicaid Expansion	-19.877	11.880**	-19.065**
# Chosen Donor Pool States	5	4	5
	(7) OR	(8) TN	(9) UT
Medicaid Expansion	14.401	20.421	-31.678
# Chosen Donor Pool States	3	10	5
	(10) VT	(11) WA	
Medicaid Expansion	-7.873*	-4.704	
# Chosen Donor Pool States	4	3	

Notes: Chosen donor pool states for (1)AZ: CA, NV, CO; (2)HI: AK, CO, ND (7)OR & (11)WA: CA, ID; (4)MN: WI, IA, ND, SD; (9)UT: NV, WY, ID, CO; (5)NM: TX, CO, OK; (3)ME & (10)VT: NH, RI, CT; (6)NY: PA, NJ, CT, RI; (8)TN:KY, AK, MO, IL, MS, AL, GA, NC, VA.
Pseudo p-values from permutation tests: *** p<0.01, ** p<0.05, * p<0.1

No clear and consistent pattern across states emerges from Table 13. Mortality reductions are concentrated to a handful of states. We recover a large and statistically significant negative effect on state-level mortality in New York and Vermont. Mortality reductions were also achieved in Washington following Medicaid expansion. Interestingly, we show that Medicaid may even be associated with an increase in state-level mortality in the case of New Mexico. A potential explanation for this counter-intuitive result is either a confounding effect of the opioid epidemic in New Mexico that was commencing or a direct influence of Medicaid's access to medical care and opioids. However, this is a conjecture that needs to be investigated more thoroughly using more information. For instance, state levels of opioid usage could be included as a regressor. Alternatively, we could obtain diagnosis-specific mortality data and exclude mortality counts with opioid usage as the primary cause of death to assess the robustness of our results to the influence of the opioid epidemic in New Mexico. An increase in mortality was also recorded for Hawaii that might simply due to the poor fit of the synthetic control. Similarly, we recover mortality reductions in Utah and Minnesota but discard those results because of the failure of the SCD to construct reasonable synthetic representations of Utah and Minnesota, pre-treatment. Specifically, the mortality gap between the observed state mortality levels and the synthetic ones remains large and negative before and after the respective Medicaid expansion.

A graphical illustration of the SCD findings with respect to coverage and access to care is given in Figures I.4.C through I.14.C in Appendix I. Table 14 presents the SCD estimates of the effect of state Medicaid expansions on the total number of Medicaid beneficiaries in each state in analogous fashion to Table 13.

Table 14: State-Specific SCD Estimates of Effect of State Medicaid Expansions on Total Medicaid Beneficiaries

	(1)	(2)	(3)
VARIABLES	AZ	HI	ME
Medicaid Expansion (in %)	43.948	-19.891	27.200
	(4)	(5)	(6)
	MN	NM	NY

Medicaid Expansion (in %)	-12.319	3.184	11.056

	(7) OR	(8) TN	(9) UT
Medicaid Expansion (in %)	33.102	47.400	15.102

	(10) VT	(11) WA	
Medicaid Expansion (in %)	30.731	10.789	

Our synthetic control design estimates of the effect of Medicaid expansions on the total number of beneficiaries enrolled in Medicaid do not indicate a monotonic relationship between the size of each Medicaid expansion and its impact on adult mortality. As one would expect, the significant reductions in adult mortality were recorded in the states of New York and Vermont were associated with substantial increases in the per cent of total Medicaid beneficiaries by 11.1 per cent and 30.7 per cent, respectively. Similarly, the mortality increases in Hawaii occurred over a period when fewer beneficiaries were enrolled in the state's Medicaid plan. However, states like Arizona, Oregon and Tennessee saw an even larger expansion in the number of total Medicaid enrollees but adult mortality, if anything, increased. In the same fashion, the significant increase in mortality in New Mexico did not coincide with a contraction in the number of total Medicaid enrollees. Still, one should interpret these associations with caution due to measurement error. Specifically, adult mortality could have been affected by changes in the mortality levels of adult pregnant women which were covered by WIC. Also, the number of total Medicaid enrollees might have changed due to increases of coverage to teenagers who do not contribute to the average adult mortality levels.

5. Concluding Remarks

Using a synthetic control design with randomization inference, there is no evidence that Medicaid expansions reduce adult mortality in most states. Our state-specific findings are robust to the research design we employ. Specifically, the estimated state-specific treatment effects using a SCD in Table 13 are more conservative relative to the DID ones in Table 12 above. The estimated effect sizes are fairly comparable and the direction of the sign differs only for states for which statistically insignificant results are found. Still, our DID findings suggest that, at the national level, Medicaid did reduce mortality, a result primarily driven by the large influence of populous New York on the national average.

Our results from suggest that pre-existing estimates of mortality reductions due to Medicaid expansions are overstated. Our findings about New York are in line with Sommers et al. (2012) and Sommers (2014). However, our results for the case of Arizona and Maine point toward a null effect of Medicaid expansions on state-level adult mortality. Moreover, our findings suggest that these interventions may even lead to an increase in mortality when other behavioral factors are at play such as the opioid epidemic. It should be noted that the size of these Medicaid expansions differed across state. However, no clear-cut relationship between the size of the Medicaid expansions in terms of total number of Medicaid enrollees and adult mortality can be deduced. It is possible that adult mortality is not driven by the extensive margin of health coverage but the intensive one; i.e., the quality and generosity of each state's Medicaid plan.

Appendix A: Solving the Model

From the taxpayer's perspective, the critical action of tax administration is whether collection of the tax liabilities will be enforced. It is important to clarify that tax evasion in this model occurs either directly through taxpayer non-compliance (potentially due to beliefs of weak tax debt collection enforcement) or indirectly through actual weak debt collection enforcement on behalf of the tax administration.

The first-order condition with respect to the level of the tax delinquent debt to evade is:

$$\frac{\partial EU}{\partial x_i^1} = (1-\pi_i) \cdot \frac{\partial u(c_i^n)}{\partial x_i^1} - \theta \cdot \pi_i \cdot \frac{\partial u(c_i^e)}{\partial x_i^1} = 0 \tag{10}$$

where $c_i^e = y_i - \theta \cdot x_i - s_i$ is period-one consumption if debt collection is perfectly enforced and $c_i^n = y_i + x_i - s_i$ if it is not enforced. At the optimal amount of non-compliance of the tax delinquency, the expected utility from consumption augmented by income from the unenforced collections of the outstanding tax debt is equal to the expected costs from reduced consumption due to the enforced tax lien, $(1-\pi_i) \cdot \frac{\partial u(c_i^n)}{\partial x_i^1} = \theta \cdot \pi_i \cdot \frac{\partial u(c_i^e)}{\partial x_i^1}$.

The first-order condition with respect to the level of savings of the portion of the debt to be evaded is:

$$\frac{\partial EU}{\partial s_i} = -(1-\pi_i) \cdot [\frac{\partial u(c_i^n)}{\partial s_i} - \delta \frac{\partial V(s,\pi^n)}{\partial x_i^1}] - \pi_i \cdot [\frac{\partial u(c_i^e)}{\partial s_i} - \delta \frac{\partial V(s,\pi^e)}{\partial s_i}] = 0 \tag{11}$$

The taxpayer adjusts savings to equalize aggregate marginal utility of the two periods across potential states of the tax debt collection enforcement.

The first-order condition for $k \in \{e,n\}$ state of the world of tax debt collection enforcement is :

$$\frac{\partial V}{\partial x_i^2} = (1-\pi^k)\cdot\frac{\partial u(c_i^{kn})}{\partial x_i^2} - \pi^k\cdot\theta\cdot\frac{\partial u(c_i^{ke})}{\partial x_i^2} = 0 \tag{12}$$

which implies a marginal rate of substitution across tax debt collection enforcement states of the world:

$$MRS^k = \frac{\frac{\partial u(c_i^{kn})}{\partial x_i^2}}{\frac{\partial u(c_i^{ke})}{\partial x_i^2}} = \theta\cdot\frac{(\pi^k)}{(1-\pi^k)} \tag{13}$$

The optimality condition in period two resembles the one in period one only that it is conditional on first-period tax debt collection enforcement. At the optimal level of x_i^2, marginal utility derived from additional consumption due to income from unpaid tax debt whose collection is not enforced is equal to the marginal cost from the enforcement of the tax lien and the application of the financial penalty due to perfect tax debt collection enforcement. Consumption in the second period then depends on whether the second-period debt collection is perfectly enforced conditional on first-period tax debt collection enforcement; i.e., either $c_i^{ke} = (1+r)s_i - \theta\cdot x_i^2$ under perfect or $c_i^{kn} = (1+r)s_i + x_i^2$ under imperfect tax debt collection enforcement.

Comparative statics show that the effect on second period tax evasion of updating beliefs about debt collection enforcement following taxpayer information gains in the first-period.

$$\frac{dx_i^B}{d\pi^e} = \frac{\frac{dMRS^k}{d\pi^e}}{\frac{dMRS^k}{dx_i^B}} \tag{14}$$

The optimal amount of tax evasion is equal to

$$(y-s)\cdot\frac{1-\pi(1+\theta)}{\theta}+\rho\cdot[(1-\pi)\frac{1-\pi^{n}(1+\theta)}{\theta}+\pi\cdot\frac{1-\pi^{e}(1+\theta)}{\theta}]\cdot(1+r)\cdot s \tag{15}$$

From the government's perspective, the present value of tax revenues is given by T below:

$$T(\theta,\alpha)=\int(1-p)[1-x_i^*]\cdot R_g+p\cdot\theta\cdot x_i^*(R_i,q_i^{\alpha})dR_i \tag{16}$$

The total present value of private welfare is given by:

$$PW(\theta,\alpha)=-\int\{(1-p)[1-x_i^*(R_i,q_i^{\alpha})]\cdot R_i+p\cdot\theta\cdot x_i^*(R_i,q_i^{\alpha})\}dR_i \tag{17}$$

Clearly, the government is better off by not employing information nudging when it only values tax revenues, $\psi=1$. When the government is is indifferent between private welfare and tax revenues, $\psi=\frac{1}{2}$, then it is strictly better off by employing information nudging than relying solely on the financial penalty. By the mean value theorem, there exists a unique $\psi^*\in(\frac{1}{2},1)$ for which the government is better off employing information nudging ($\alpha<1$) if $\psi<\psi^*$, is indifferent between the financial penalty and information nudging if $\psi=\psi^*$ and is better off relying solely on the financial penalty ($\alpha=1$) if ($\hat{q}^{\alpha}>q$) if $\psi>\psi^*$.

Appendix B: Identification and Estimation

Coarsened Exact Matching

There exists a plethora of matching techniques in addition to the ideal but improbable exact matching. The main distinction of matching techniques is their approach in measuring the distance between observations of different treatment status. Propensity score matching and Mahalanobis distance matching are characterized as Equal Percent Bias Reducing matching methods. CEM is a generalized class of Monotonic Imbalance Bounding matching methods that dominates Equal Percent Bias Reducing matching methods (Iacus et al., 2011). It provides post-matching samples that are better balanced and estimates of the causal quantity of interest that have lower root mean square error than matching based on propensity scores, Mahalanobis distance, nearest neighbors, and optimal matching. These satisfy weaker properties and only in expectation.

Intuitively, CEM constructs a set of strata $s \in S$ each with identical coarsened values for covariates. Then, the observations coming from strata that do not retain both a treated and a control observation are removed. Let the treated and control observations in stratum s be T^s, C^s, respectively, whereas the number of matches is given by $m_T^s = T^s$, $m_C^s = C^s$. Then, the total number of matched and controls over all strata is $m_T = \bigcup_{s \in S} m_T^s$ and $m_C = \bigcup_{s \in S} m_C^s$, respectively.

CEM makes an explicit statement about the overall covariate imbalance between treated respondents and untreated ones. Specifically, the statistic L_1 expresses the global difference between the multidimensional histogram of all pre-treatment covariates of T^s and C^s. This measure includes imbalance with respect to the full joint distribution, including all interactions, of the covariates. Thus, L_1 works for imbalance as R^2 works for model fit: the absolute values mean less than comparisons between matching solutions. A $L_1 = 0$ statistic indicates perfect global balance up to the level of coarsening. Larger values imply larger imbalance, with a maximum of $L_1 = 1$, which indicates complete separation. CEM assigns the following weights w_i to matched and unmatched observations:

$$w_i = \begin{cases} 1, & \text{if } i \in T^s \\ \frac{m_C}{m_T} \cdot \frac{m_T^s}{m_C^s} & \text{if } i \in C^s \\ 0, & \text{if } i \text{ is unmatched} \end{cases} \tag{18}$$

Regression Discontinuity Design

The choice of the bandwidth h_n may reflect a trade-off between bias and variance or come from cross-validation. It is widely preferred to use a data-driven bandwidth that asymptotically minimizes Mean Squared Error, the sum of squared bias, and variance (Imbens and Kalyanaraman, 2009). In this treatment, a local linear estimator is used with a correction-term that prevents small denominators in moderate-sized samples as mine. However, the preliminary bandwidths b_n used for the construction of the optimal one may not be optimally chosen. A second-generation plug-in estimator proposed by Calonico et al. (2014) is based on b_n's which are consistent estimators of the corresponding population h_n. This is achieved by using a fixed-matches, nearest-neighbor-based "estimate" of the residuals as in Abadie and Imbens (2006). This makes confidence intervals more robust to the bandwidth choice ("small" or "large"), as they are not only valid when the usual bandwidth conditions are satisfied (being asymptotically equivalent to the conventional confidence intervals), but offer correct coverage rates in large samples even when the conventional confidence intervals do not. Lastly, quantile-specific bandwidths have been proposed that use cross-validation as in Ludwig and Miller (2007). Within the h_{IK} bandwidth proposed by Imbens and Kalyanaraman (2009), observations are weighted according to the following formula that penalizes those further away from the cut-off.

$$pw_i = \max\{0, \frac{h_{IK} - |z|)}{h_{IK}}\} \tag{19}$$

The left panel of Figure B1 provides a residual plot of the outcome equation. The distance of each point from the red line of zero indicates how poor the prediction is for that point. Clearly, the residuals are highly heteroskedastic. They exhibit increasing variance as they take on increasingly larger values along the x-axis. To obtain homoskedastic residuals, I transform Y_i using the inverse hyperbolic sine transformation (IHS), $g(Y_i) = log[Y_i + (Y_i^2 + 1)^{\frac{1}{2}}] = sinh^{-1}(Y_i)$. This is a convenient transformation for the outcome I am interested due to two appealing properties: (a) it is well-defined over zero but (b) retains logarithmic properties. Its first attribute prevents the empiricist from discarding observations that take zero values, unlike the log transformation. This allows for a logarithmic interpretation except for very small values of Y_i (Burbidge et al., 1988).

Figure B1: Heteroskedastic and Homoskedastic Residuals

The right panel of Figure B1 shows the residual plot of transformed values of the outcome. The cloud of residuals indicates that the residuals of observations with positive payments are normal. Second, there is a downwards trending line of residuals that envelops the cloud from below and tends increasingly farther for higher values of the outcome. These plotted residuals denote with zero payments. It is worth noting that this downwards trending line is the main difference to the residual plot of the log-transformed outcome variable. Thus, a logarithmic transformation will also effectively address heteroskedasticity. However, by discarding observations with zero tax payments and relying only on those with positive ones it will introduce sample selection bias.

Appendix C: Figures

This subsection presents a graphical representation of the RDD findings for personal income tax delinquents.

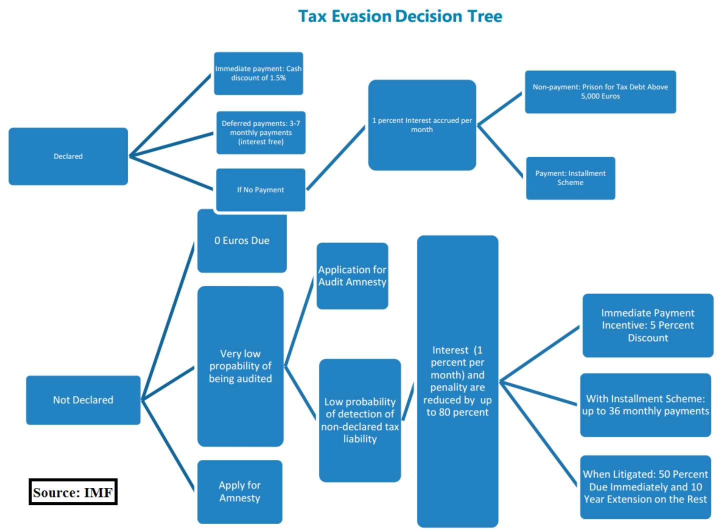

Figure C1: Universe of Tax Evasion Windows of Opportunity

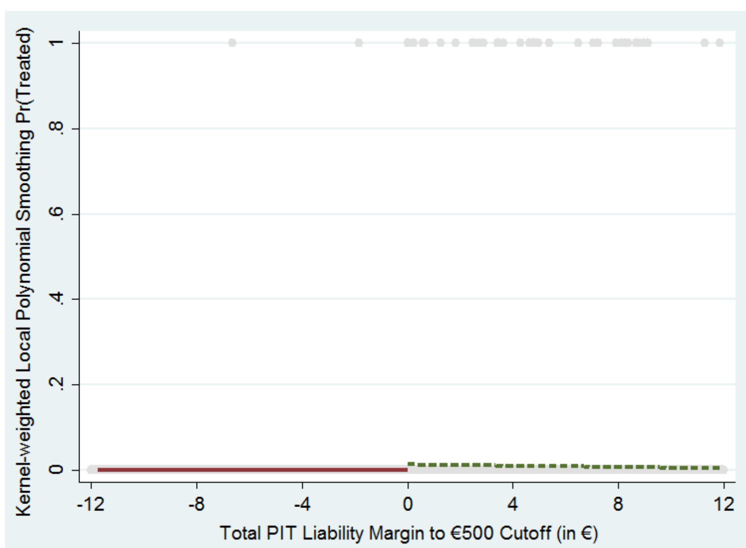

Figure C2: First Stage - €500 Cutoff as an Instrument

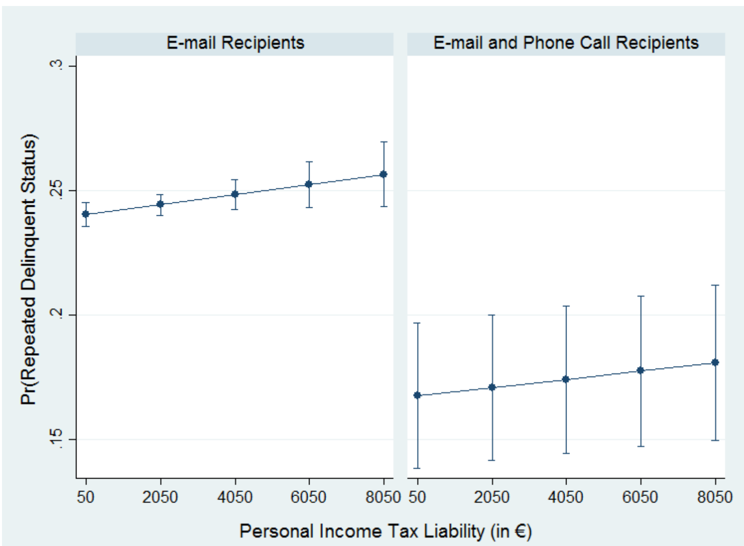

Figure C3: Effect of Reminder on Probability of Becoming Personal Income Tax Delinquent Again

Appendix D: Tables

Table D1: Summary Statistics - New Tax Debtors

Variables	Mean	Std Dev	Min	Max
E-mail Reminder Recipients - N = 174,613				
Total PIT Liabilities (in €)	292.44	448.11	50	10000
Payment of Total PIT Liabilities (in €)	99.11	277.52	0	9961.93
Payment of New PIT Liabilities (in €)	292.44	448.11	50	10000
New PIT Liabilities (in €)	99.06	277.45	0 9961.93	
Age	53	17	18	111
Male	0.73	0.44	0	1
Salaried	0.80	0.39	0	1
Self-employed	0.19	0.39	0	1
Unemployed	0.15	0.35	0	1
Phone Call and E-mail Reminder Recipients - N = 2,762				
Total PIT Liabilities (in €)	2235.446	1946.417	146.6	10000
Payment of Total PIT Liabilities (in €)	468.5465	1171.45	0	9770.91
New PIT Liabilities (in €)	2235.44	1946.41	146.6	10000
Payment of New PIT Liabilities (in €)	468.45	1171.30	0	9770.91
Age	55	16	18	99
Male	0.78	0.41	0	1
Salaried	0.76	0.42	0	1
Self-employed	0.23	0.42	0	1
Unemployed	0.18	0.38	0	1

Notes: Universe of "new tax debtors", taxpayers with newly-delinquent PIT ranging [€0,€10,000] with and zero tax arrears. N = 177,375.

Table D2: Summary Statistics - Old Tax Debtors

Variables	Mean	Std Dev	Min	Max
E-mail Reminder Recipients - N = 297,519				
Total PIT Liabilities (in €)	1094.92	1217.76	51.71	9999.73
Payment of Total PIT Liabilities (in €)	187.07	445.33	0	9972.47
Payment of New PIT Liabilities (in €)	441.93	536.51	50	9926.27
New PIT Liabilities (in €)	140.32	311.05	0	8771.91
Age	56	17	18	111
Male	0.75	0.43	0	1
Salaried	0.73	0.44	0	1
Self-employed	0.26	0.44	0	1
Unemployed	0.15	0.36	0	1
Phone Call and E-mail Reminder Recipients - N = 1,550				
Total PIT Liabilities (in €)	5246.85	2323.66	252.22	10000
Payment of Total PIT Liabilities (in €)	1239.91	1842.48	0	9565.65
New PIT Liabilities (in €)	2836.71	1612.04	77.61	9333.13
Payment of New PIT Liabilities (in €)	1040.31	1533.34	0	9167.54
Age	52.81	13.74	20	90
Male	0.83	0.37	0	1
Salaried	0.37	0.48	0	1
Self-employed	0.62	0.48	0	1
Unemployed	0.10	0.31	0	1

Notes: Universe of "old tax debtors", taxpayers with overdue PIT ranging [€0,€10,000] and positive tax arrears. N = 299,069.

Table D3: CEM-weighted Post-Matching Sample. Covariate Balance of Treated VAT Delinquents with Combined (E-Mail and Phone Call) Compliance Reminders and with only an E-mail Compliance Reminder

Matched Controls	N = 241	Total Liability Range	$[400, 600]$	
Variable	**Mean**	**Standard Deviation**	**Min**	**Max**
Total Liabilities	543.64	30.04	500	599.97
Male	0.502	0.501	0	1
Age	42.8	8.1	23	67
Self-employed	1	0	1	1
Salaried	0	0	1	1
Formerly Self-employed	0	0	0	0
Debtor	0	0	0	0
Combined Compliance Reminder Recipients	N = 70			
Variable	**Mean**	**Standard Deviation**	**Min**	**Max**
Total Liabilities	543.73	29.51	503.69	599.49
Male	0.528	0.502	0	1
Age	44.9	8.9	23	66
Self-employed	1	0	1	1
Salaried	0	0	0	0
Formerly Self-employed	0	0	0	0
Debtor	0	0	0	0

Table D4: CEM-weighted Post-Matching Sample. Covariate Balance of Treated PIT Delinquents with Phone Call Compliance Reminders and Controls

Matched Controls	N = 545	Total Liability Range	$[600, 1000]$	
Variable	**Mean**	**Standard Deviation**	**Min**	**Max**
Total Liabilities	748.14	108.94	600	997.69
Male	0.776	0.417	0	1
Age	59.6	18.9	19	95
Self-employed	0.034	0.183	0	1
Salaried	0.965	0.183	0	1
Formerly Self-employed	0.058	0.235	0	1
Debtor	0.071	0.257	0	1
Combined Compliance Reminder Recipients	N = 309			
Variable	**Mean**	**Standard Deviation**	**Min**	**Max**
Total Liabilities	761.78	112.96	600	995.2
Male	0.663	0.473	0	1
Age	56.2	20.5	22	93
Self-employed	0.051	0.221	0	1
Salaried	0.948	0.221	0	1
Formerly Self-employed	0.055	0.228	0	1
Debtor	0.009	0.098	0	1

Table D5: CEM-weighted Post-Matching Sample. Covariate Balance of Treated PIT Delinquents with Phone Call Compliance Reminders and Controls

Matched Controls	N = 363	Total Liability Range	$[1000, 2000]$	
Variable	**Mean**	**Standard Deviation**	**Min**	**Max**
Total Liabilities	1349.04	281.34	1000	1995.81
Male	0.801	0.399	0	1
Age	0.077	0.267	0	1
Self-employed	0.034	0.183	0	1
Salaried	0.922	0.267	0	1
Formerly Self-employed	0.085	0.279	0	1
Debtor	0.187	0.390	0	1
Combined Compliance Reminder Recipients	N = 220			
Variable	**Mean**	**Standard Deviation**	**Min**	**Max**
Total Liabilities	1341.45	269.8	1000	1994.58
Male	0.718	0.45	0	1
Age	63.6	18	24	96
Self-employed	0.05	0.218442	0	1
Salaried	0.95	0.218	0	1
Formerly Self-employed	0.086	0.281	0	1
Debtor	0.04	0.198	0	1

Table D6: CEM Estimates of Effect of Combined Compliance Reminder on Payments of PIT

Variables	(1)	(2)	(3)	(4)	(5)
Combined Compliance Reminders	1.22**	1.85***	1.20***	0.85***	0.11
	(0.47)	(0.31)	(0.16)	(0.17)	(0.26)
Constant	3.239**	2.805***	2.42***	2.79***	3.58***
	(0.273)	(0.187)	(0.14)	(0.18)	(0.41)
R-Squared	0.019	0.051	0.05	0.04	0.01
Pre-Matching Covariate Imbalance L_1	0.9593	0.9499	0.8024	0.8076	0.8248
Post-Matching Covariate Imbalance L_1	0.2137	0.4783	0.4956	0.5861	0.5899
Total Liability Range	$[400, 600]$	$[600, 1000]$	$[1000, 2000]$	$[2000, 5000]$	$[5000, 10000]$
# of Matched Strata	45	98	346	333	157
# of Matched Treated Observations	70	140	900	646	389
# of Matched Control Observations	241	538	1441	955	332
Observations	311	678	2,341	1,601	721

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Standard errors in parentheses, clustered at the zip code level.

Table D7: 2SLS Estimates of Compliance Reminders' LATEs on Payments of Total PIT Liabilities

Variables	(1) Payments in Levels	(2) Payments in Levels	(3) Payments in %	(4) Payments in %	(5) Payment Rate	(6) Payment Rate
Phone Call	-728.69	-1033.079	-40.80^{*}	-46.67^{**}	-9.17^{**}	-9.14^{*}
	(658.25)	(916.50)	(21.51)	(21.58)	(4.28)	(4.87)
Distance to Cutoff	1.14	6.94	0.16^{***}	0.42	0.08	0.06
	(1.10)	(4.35)	(0.07)	(0.28)	(0.06)	(0.22)
Distance to Cutoff Squared		0.30		0.07		-0.02
		(0.24)		(0.07)		(0.16)
Above Cutoff $*$ Distance to Cutoff	-10.57	-1.11	-0.22	-0.12	-0.03	-0.03
	(1.59)	(6.37)	(0.14)	(0.51)	(0.11)	(0.40)
Above Cutoff $*$ Distance to Cutoff Squared		-0.12		-0.01		-0.04
		(0.50)		(0.09)		(0.26)
Constant	132.60^{***}	128.48^{***}	3.33^{***}	3.46^{***}	0.60^{***}	0.59^{***}
	(9.54)	(8.07)	(0.17)	(0.22)	(0.04)	(0.06)
Polynomial Order	1	2	1	2	1	2
Bandwidth (in €)	10.90	10.90	4.18	4.18	3.01	3.01
Observations	8,376	8,376	3,547	3,547	2,606	2,606

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. 2SLS estimation. Standard errors in parentheses, clustered at the zip code level.

Table D8: 2SLS Estimates of Compliance Reminders' LATEs on Payments in Per Cent Terms of Total PIT Liabilities by Debt Vintage

	(1)	(2)
Variables	New Debtors	Old Debtors
Compliance Reminder(s)	0.97	-28.05^{*}
	(4.39)	(16.23)
Distance to Cutoff	0.01	0.01^{***}
	(0.04)	(0.00)
Above Cutoff-Distance	0.04	-0.01^{***}
	(0.06)	(0.00)
Constant	3.64^{***}	3.38^{***}
	(0.24)	(0.03)
Polynomial Order	1	1
Bandwidth (in €)	11.67	250
# Treated Observations	41	117
Observations	1,415	121,037

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Standard errors in parentheses, clustered at the zip code level.

Table D9: 2SLS Estimates of Compliance Reminders' LATEs on Payments of Total PIT Liabilities

Variables	(1) Payment, Levels	(2) Payment, %	(3) Payment Rate
Old Debtor	-49.62^{***}	-0.13	-0.01
	(15.58)	(0.24)	(0.03)
Unemployed	2.34	0.03	0.01
	(17.41)	(0.30)	(0.03)
Self-Employed	10.69	0.17	0.02
	(17.16)	(0.31)	(0.04)
Distance to Cutoff	0.87^{*}	0.01	0.01
	(0.49)	(0.01)	(0.01)
Distance to Cutoff Squared	-0.05	-0.01	-0.01
	(0.04)	(0.01)	(0.01)
Age	1.85	0.02	0.01
	(2.13)	(0.03)	(0.01)
Age Squared	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)
Male	18.32	0.34	0.04
	(11.97)	(0.22)	(0.02)
Bandwidth (in €)	10.90	21.38	35.12
Observations	703	1,277	2,106

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Standard errors in parentheses, clustered at the zip code level.

Table D10: Inspection of Baseline Covariate Balance at Optimal Bandwidth around Cut-off-T-tests

Statistic Relative to Cutoff	Mean Below	St.Dev. Below	Mean Above	St.Dev. Above	$Pr(\|T\| > \|t\|)$
Male	0.738	0.440	0.756	0.430	0.054^{*}
Age	57.425	17.295	58.094	17.151	0.076^{*}
Old Debtor	0.851	0.356	0.844	0.363	0.3869
Salaried	0.782	0.413	0.764	0.425	0.047^{**}
Self-employed	0.218	0.404	0.236	0.420	0.049^{**}
Unemployed	0.158	0.365	0.152	0.359	0.4525
Observations	4,451	4,451	4,643	4,643	

Notes: optimal bandwidth is [€489.1,€510.9]. Columns (1) and (2) contain the mean and standard deviation of the baseline covariate for observations strictly below the €500 cut-off. Columns (3) and (4) contain the mean and standard deviation of the baseline covariate for observations above the €500 cut-off. Column (5) shows the two-tailed p-value computed using the t distribution which represents the probability of observing a greater absolute value of t under the null hypothesis. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

Appendix E: Images

Hellenic Republic
Ministry of Finance
General Directorate of Tax and Customs Affairs
General Directorate of Tax Audits and Public Revenues Collection
Tax Compliance Unit
1 Thessalonikis & Chandri
Moschato

Recipient: XXXXXX Username: XXXXXXX

Dear Sir/Madame,

Following an inspection of the electronic records of the General Directorate of Information Systems of the Ministry of Finance on 10/26/2012, due, and potentially overdue, tax delinquencies were determined. Especially, for the cases of overdue tax delinquencies we urge you to settle your outstanding debt.

If you have a counterclaim against the State on this outstanding debt, you are required to submit the necessary application at your local Tax Clearing House.

Warning: With regards to the tax delinquencies that have become overdue, the supervisor of your local Tax Clearing House reserves the right to enforce collection using all provisions in the legislation including non-negotiable administrative and financial measures to protect the State's interest.

Tax Compliance Unit Supervisor

Hara Mavridou

Note: Every electronic message we send you from the General Directorate of Information Systems, insofar it is not a response to one of yours, includes at the upper part your name and username. If you received a message suggesting it was sent from the General Directorate of Information Systems without the inclusion of this information and is not a reply to previous message of yours should be deleted since it is counterfeit and of possibly malicious content.

Figure E1: 2013 Compliance Reminder E-mail (translated from Greek by the author)

Appendix F: Figures

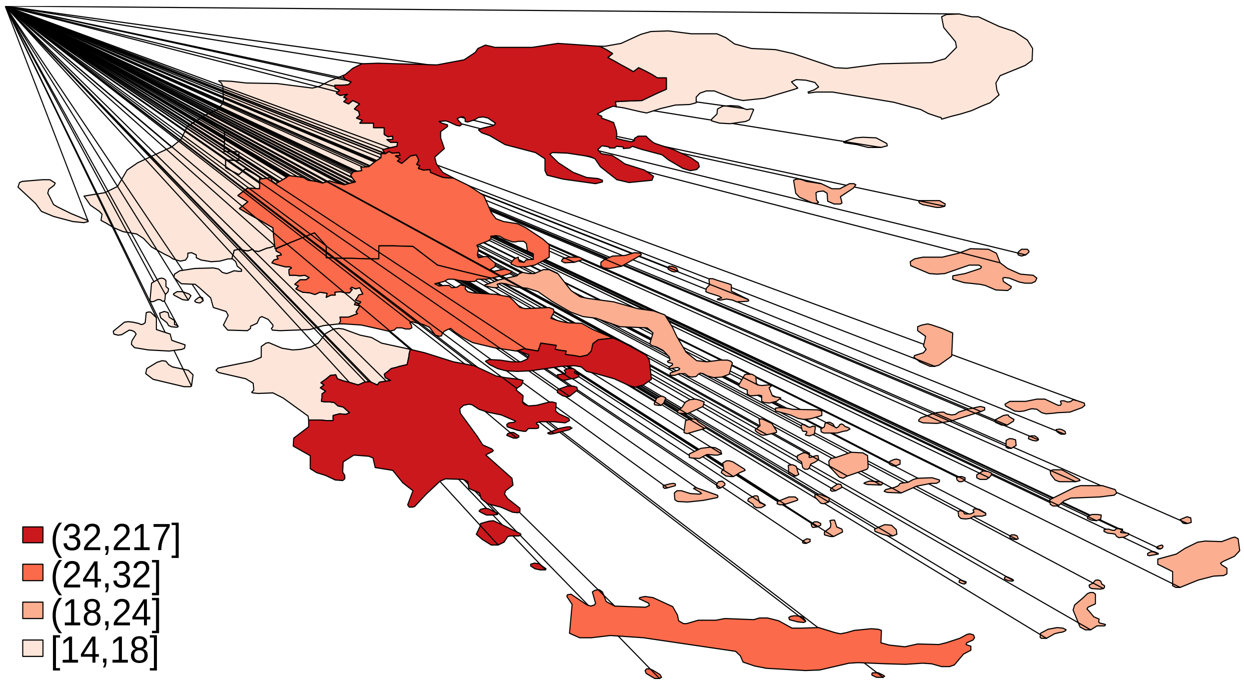

Figure F.1: Aggregate Corruption Counts from 2003 to 2008 – Map of Regions of Greece and Number of Zoning Corruption Counts

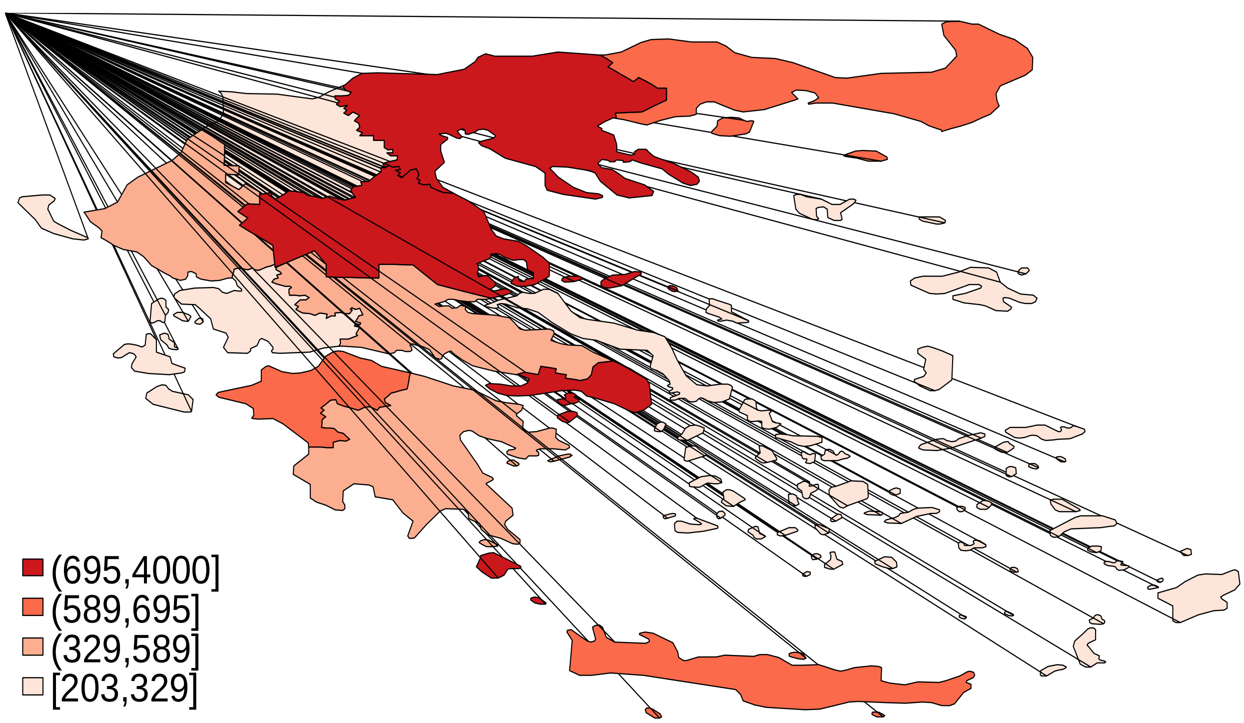

Figure F.2: Average Regional Population from 2000 through 2008 – Map of Regions of Greece and Number of Thousand Inhabitants

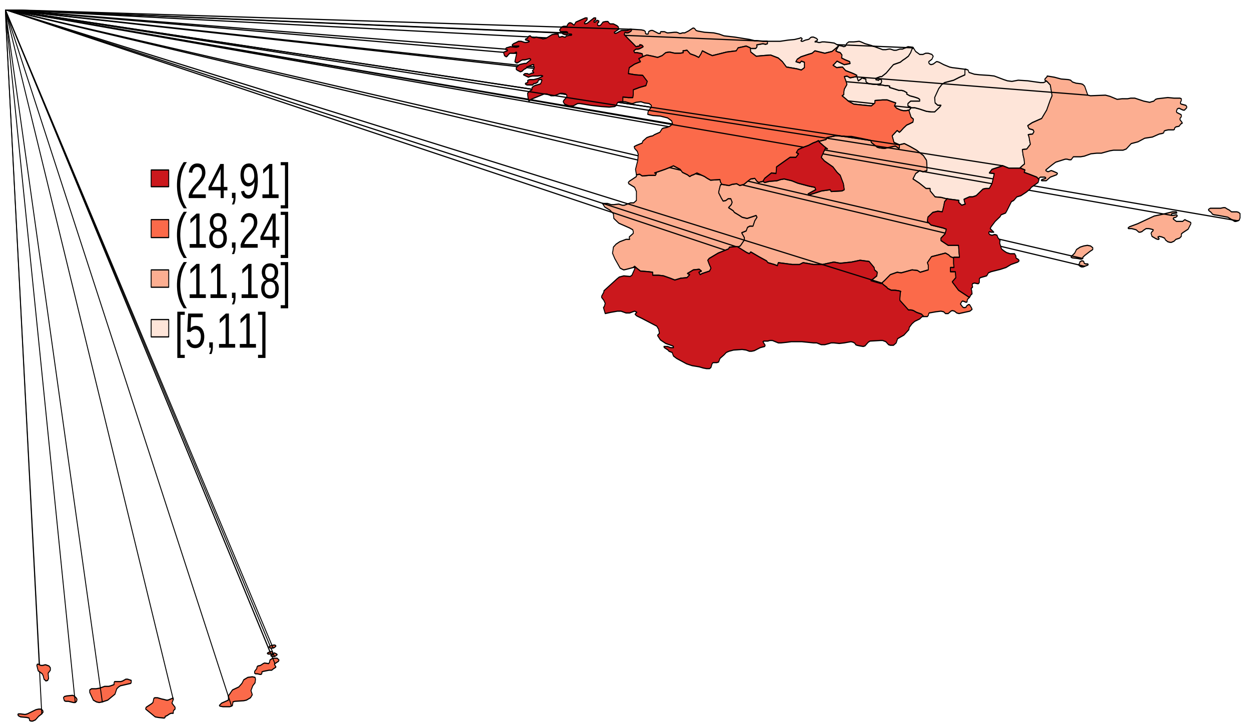

Figure F.3: Aggregate Corruption Counts from 2006 through 2008 – Map of Regions of Spain and Number of Zoning Corruption Counts

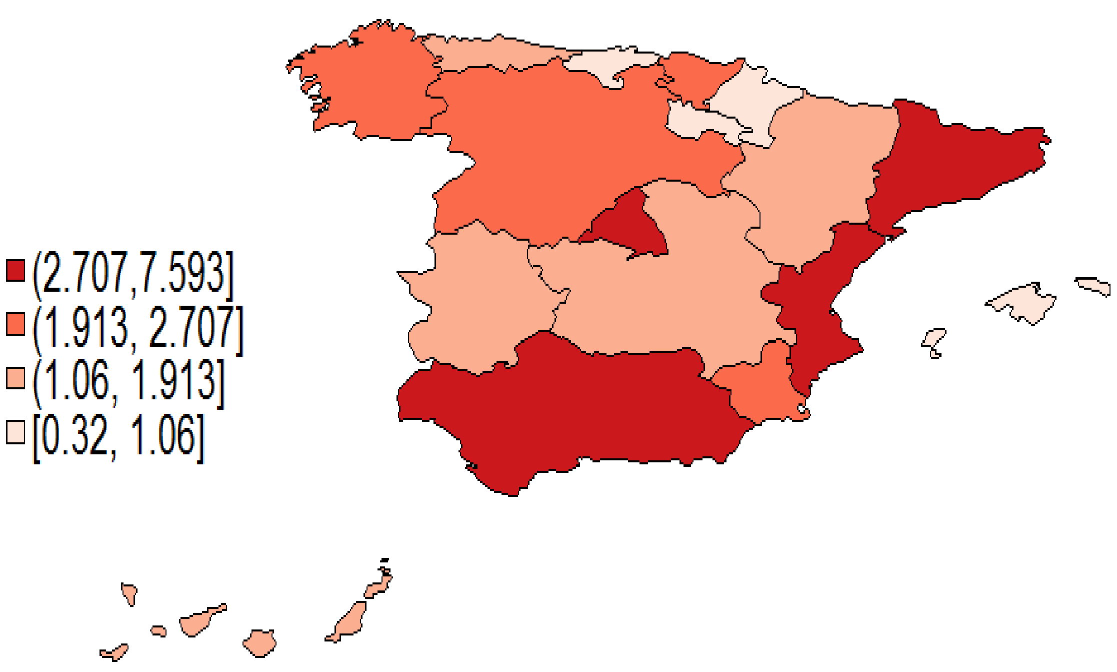

Figure F.4: Average Regional Population from 2000 through 2008 – Map of Regions of Spain and Number of Thousand Inhabitants

Appendix G: Tables

Table G.1: Summary Statistics – Greece, N = 936

Variables	Mean	Std Dev	Min	Max
Corruption Counts per Month	0.23	0.61	0	6
Housing Prices (€/m^2)	1,922.25	405.29	1,228	3,071
Unemployment Rate	9.6	2.3	4.4	18.54
Population Density	133	265	31	1,058
Population	851,855	996,541	201,796	4,000,137
College Attainment Rate	17.2	4.1	10.5	28.1
New Building Permits per Month	540	998	6	12,114

Sources: Inspectors-Controllers Body for Public Administration, Eurostat, Hellenic Statistical Authority

Table G.2: Summary Statistics – Spain, N = 204

Variables	Mean	Std Dev	Min	Max
Corruption Counts per Month	1.86	2.90	0	23
Housing Prices (€/m^2)	1847.49	509.59	914.9	3,035.8
Unemployment Rate	7.8	2.4	4.2	15.6
Population Density	159	175	24	789
Population	2,611,837	2,307,037	300,821	8,046,131
College Attainment Rate	53.4	12.9	26.5	77.6
New Building Permits per Quarter	8,525	8,251	1,045	34,600

Sources: GISAS Research Group, University of La Laguna, Eurostat, Spanish Ministry of Public Works

Table G.3: Data Notes

Description	Source	Frequency
Legal indictments of zoning officials	Inspectors-Controllers Body	Monthly
Zoning corruption scandals	University of La Laguna, Spain	Quarterly
Housing Prices in Greece (€/m^2)	Bank of Greece	Monthly
Housing Prices in Spain (€/m^2)	Ministry of Public Works and Transport	Quarterly
Seasonally adjusted unemployment rate	Eurostat	Quarterly
Inhabitants per square kilometer	Eurostat	Annual
Inhabitants	Eurostat	Annual
College-degree holders aged 30-34 years	Eurostat	Annual

Greek housing prices are taken from the Housing Price Index compiled by the Real Estate Market Analysis Section of the Bank of Greece. Four regional indices are available taking values from 24.18 to 102.6, overall, 2006 being the base year. There is a index for the cities of Athens and Thessaloniki that precisely maps onto the Attica and the Central Macedonia regions of analysis, respectively. The third index is computed as an average of housing prices of all prefecture (smaller administrative unit than regions) capitals - excluding Athens and Thessaloniki - and otherwise cities with populations larger than the capital of each prefecture. The fourth index is computed as the average of "other areas" which are not capitals of the prefecture they are located in. Next, we determine whether a region is rural or urban according to the 2010 OECD typology classification. Regions with a greater percentage of their population living in urban areas are assigned the third housing price ("Rest Large Cities") while more rural regions are assigned the fourth housing price index ("Smaller Urban Areas"). Out of the remaining 11 regions of Greece, excluding Attica and Central Macedonia, 2 rural regions; namely, Epirus and Thessaly, are assigned the fourth housing price index. Then, we bring information on the median price in 2013 for each prefecture from Bank of Greece's Real Estate Market Analysis Section, aggregate it to the regional level to derive the population-weighted average residential property value in each region. Lastly, we interact these latter values with the four housing price indices to introduce cross-sectional variation.

Table G.4: Correlation Coefficients – Greece, N = 936

Corruption Counts	1						
Housing Prices	0.333**	1					
Housing Price 18-month Lag	0.335**	0.941**	1				
Population	0.490**	0.565**	0.597**	1			
Population Density	0.528**	0.509**	0.524**	0.922**	1		
Unemployment Rate	−0.118**	−0.538**	−0.576**	−0.194**	−0.183**	1	
College Attainment Rate	0.408**	0.524**	0.437**	0.635**	0.751**	−0.243**	1

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Bonferroni adjustment;

Table G.5: Correlation Coefficients – Spain, N = 204

Corruption Counts	1						
Housing Prices	0.261**	1					
Housing Price 18-month Lag	0.244**	0.971**	1				
Population	0.238**	0.210**	0.187**	1			
Population Density	0.048**	0.583**	0.592**	0.922**	1		
Unemployment Rate	0.081**	−0.142**	−0.388**	0.261**	−0.112**	1	
College Attainment Rate	0.063**	0.327**	0.335**	0.286**	0.364**	−0.135**	1

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Bonferroni adjustment;

Appendix I: Figures

Aggregate Analysis

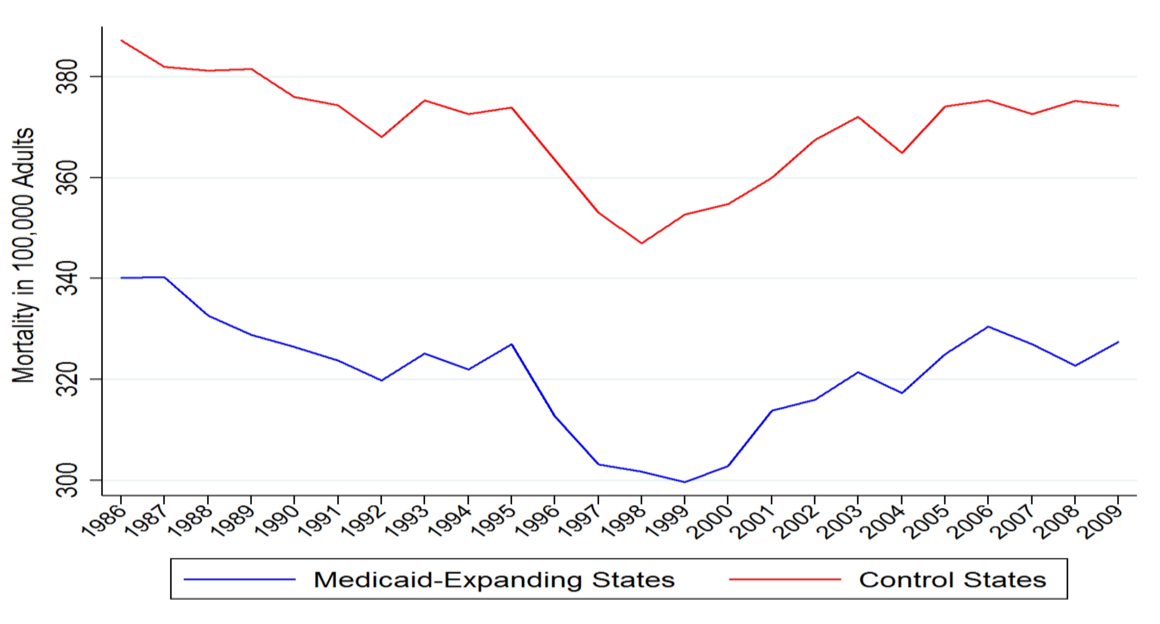

Figure I.1: Mortality Rate Trends in Adopting and non-Adopting Medicaid States (Full Sample)

Notes: Medicaid expansions occurred in 1992, 1993, 1994, 1995, 2001, and 2002

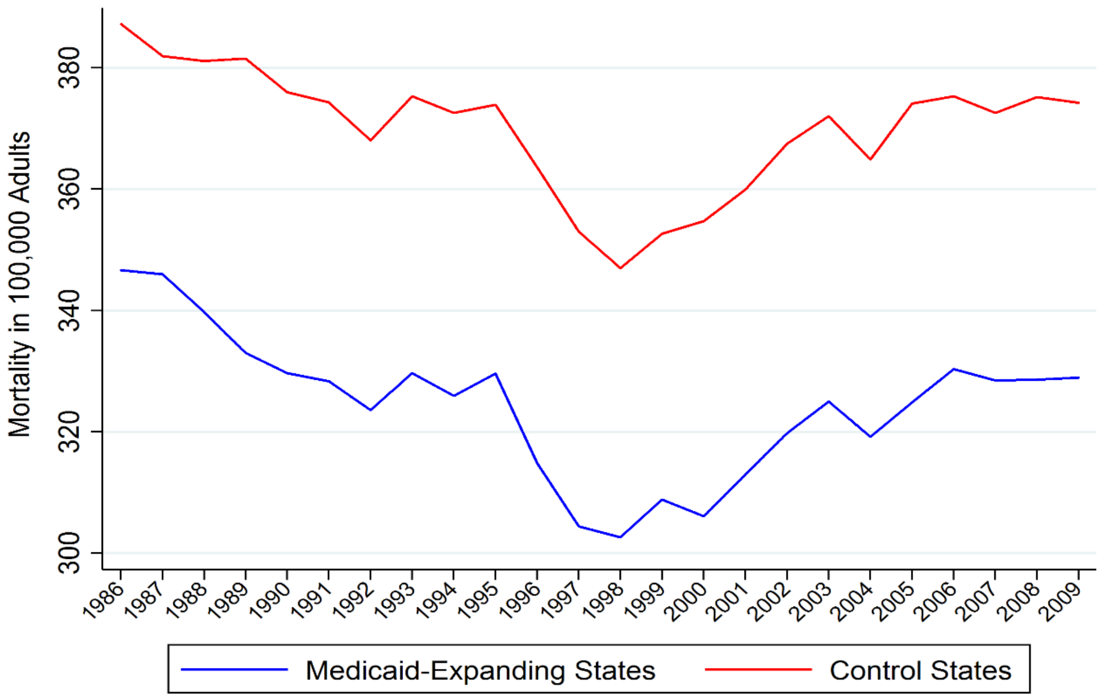

Figure I.2: Mortality Rate Trends in Adopting and non-Adopting Medicaid States (contiguous U.S. sample excluding Hawaii)

Notes: Medicaid expansions occurred in 1992, 1993, 1994, 1995, 2001, and 2002

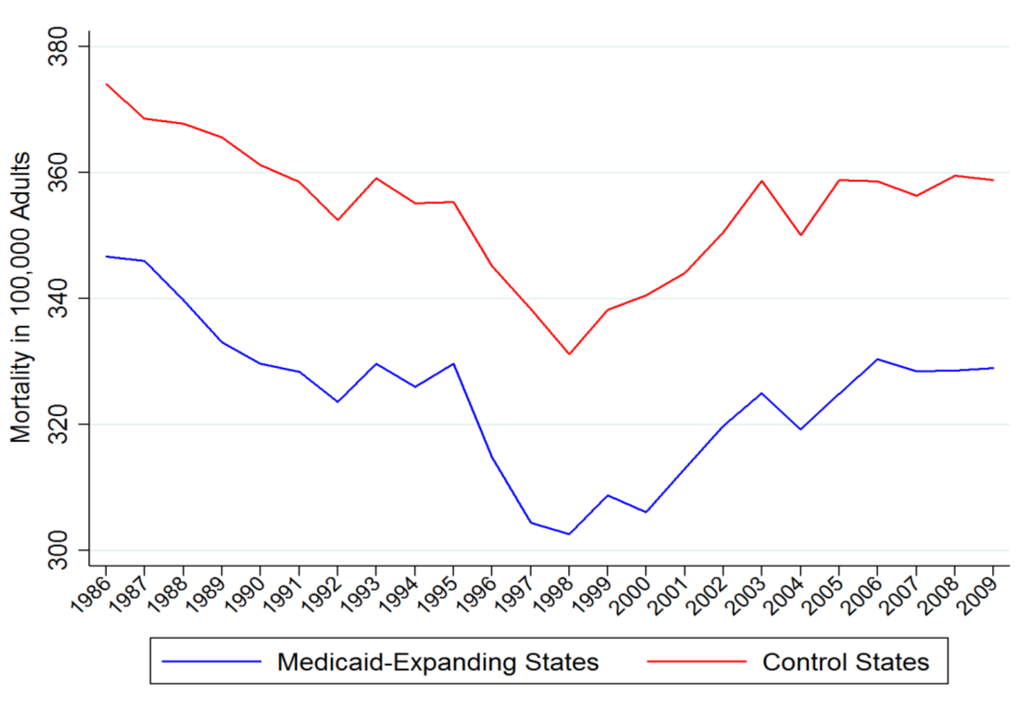

Figure I.3: Mortality Rate Trends in Adopting and non-Adopting Medicaid States (contiguous U.S. sample excluding Hawaii and hand-matched control states)

Notes: Medicaid expansions occurred in 1992, 1993, 1994, 1995, 2001, and 2002

State-Specific Analysis

1992 Medicaid Expansion in Minnesota

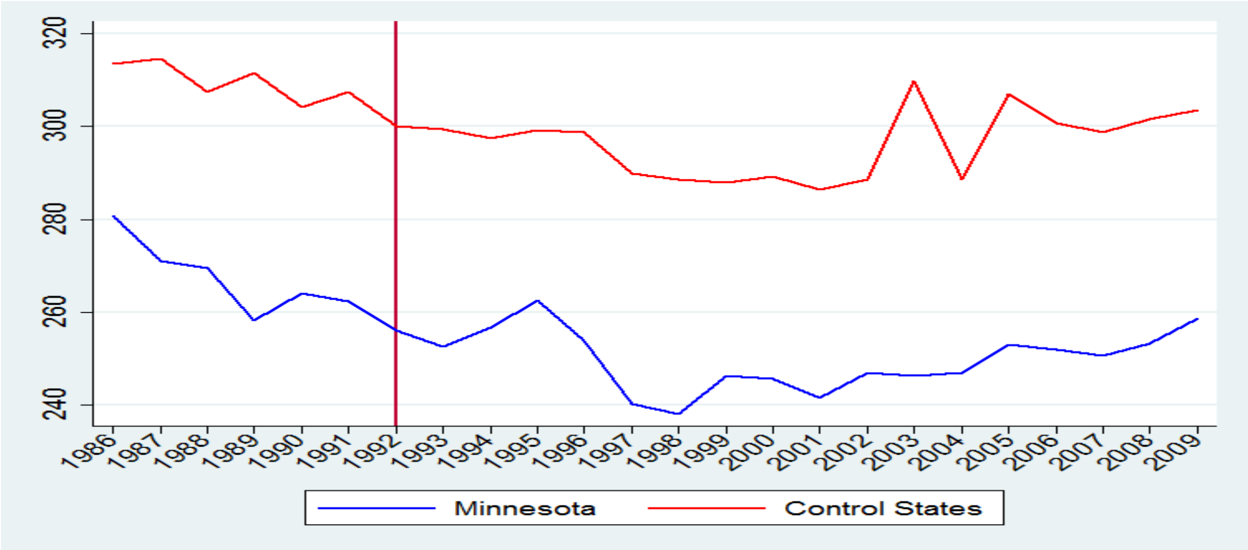

Figure I.4.A: Mortality Rate Trends in Minnesota

Notes: Medicaid expansion occurred in 1992 and is denoted by the red vertical line

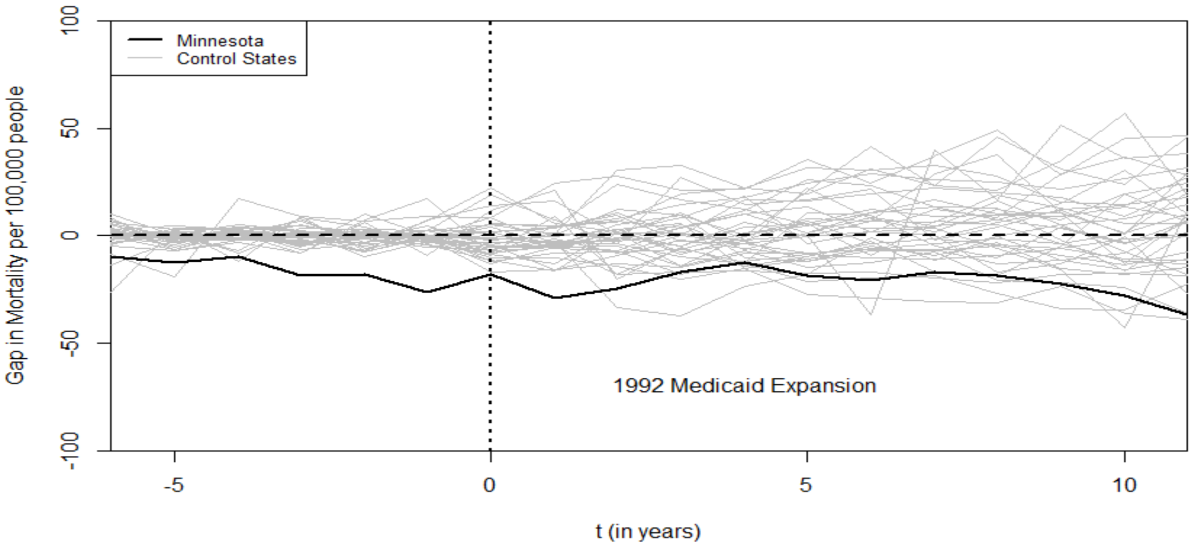

Figure I.4.B: Gap in Mortality per 100,000 people between Synthetic Control and Minnesota

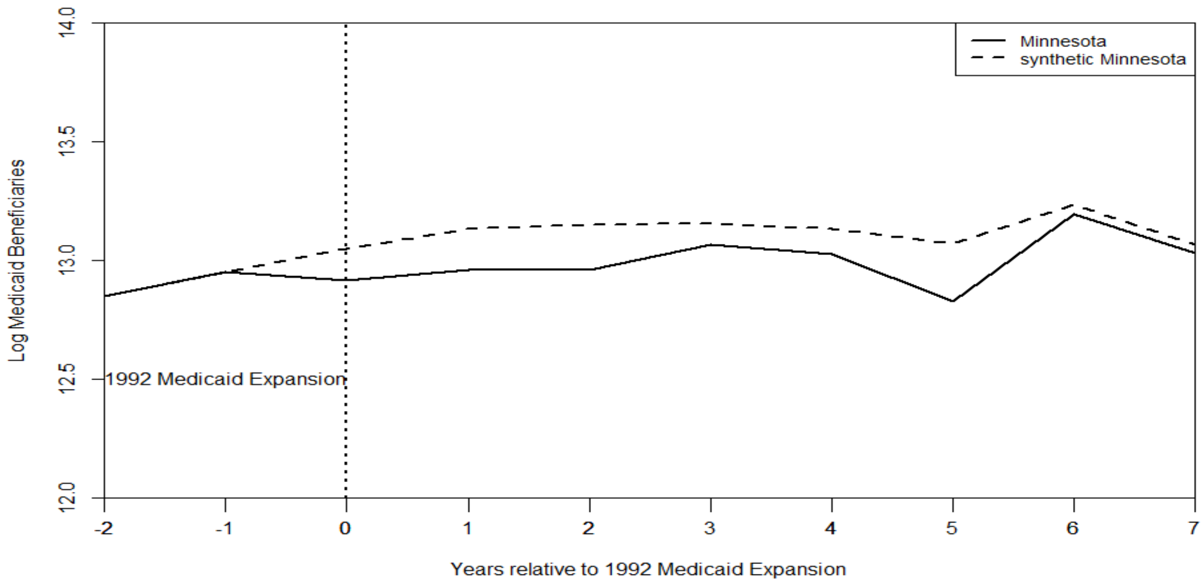

Figure I.4.C: Total Number of Medicaid Enrollees Trends in Minnesota and Synthetic Control

1993 Medicaid Expansion in Hawaii

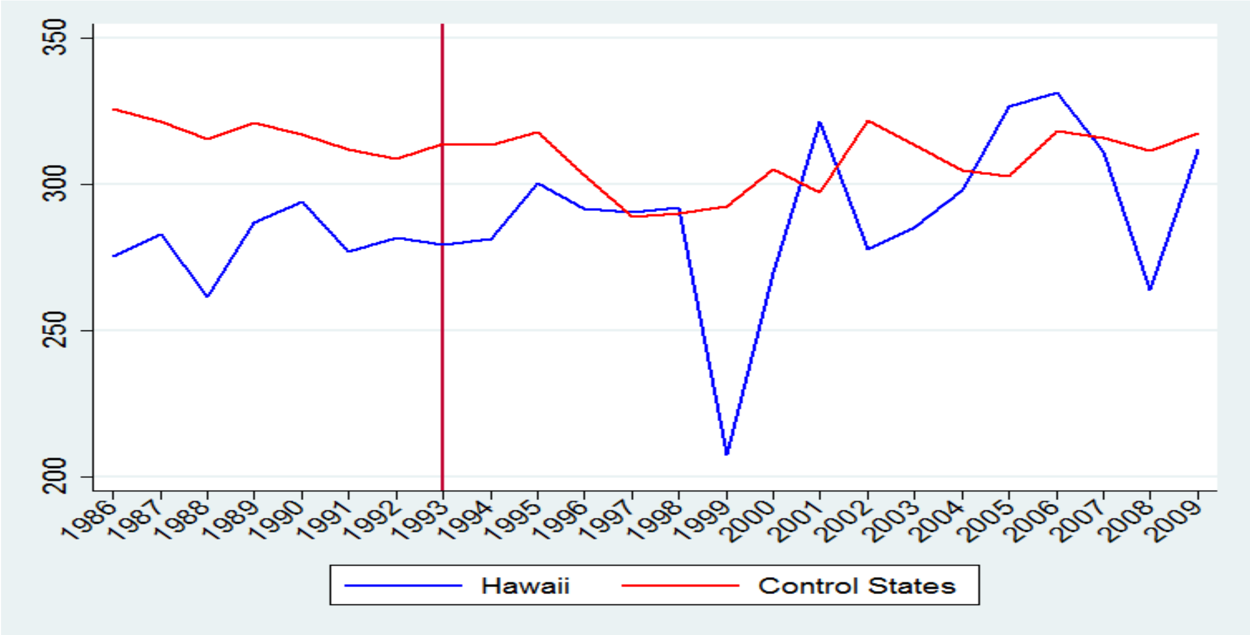

Figure I.5.A: Mortality Rate Trends in Hawaii

Notes: Medicaid expansion occurred in 1993 and is denoted by the red vertical line

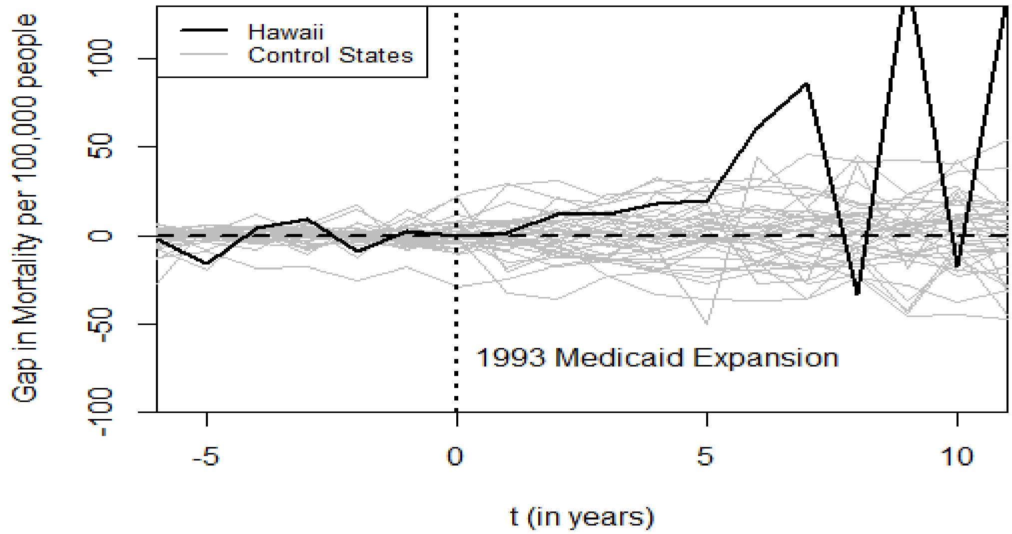

Figure I.5.B: Gap in Mortality per 100,000 people between Synthetic Control and Hawaii

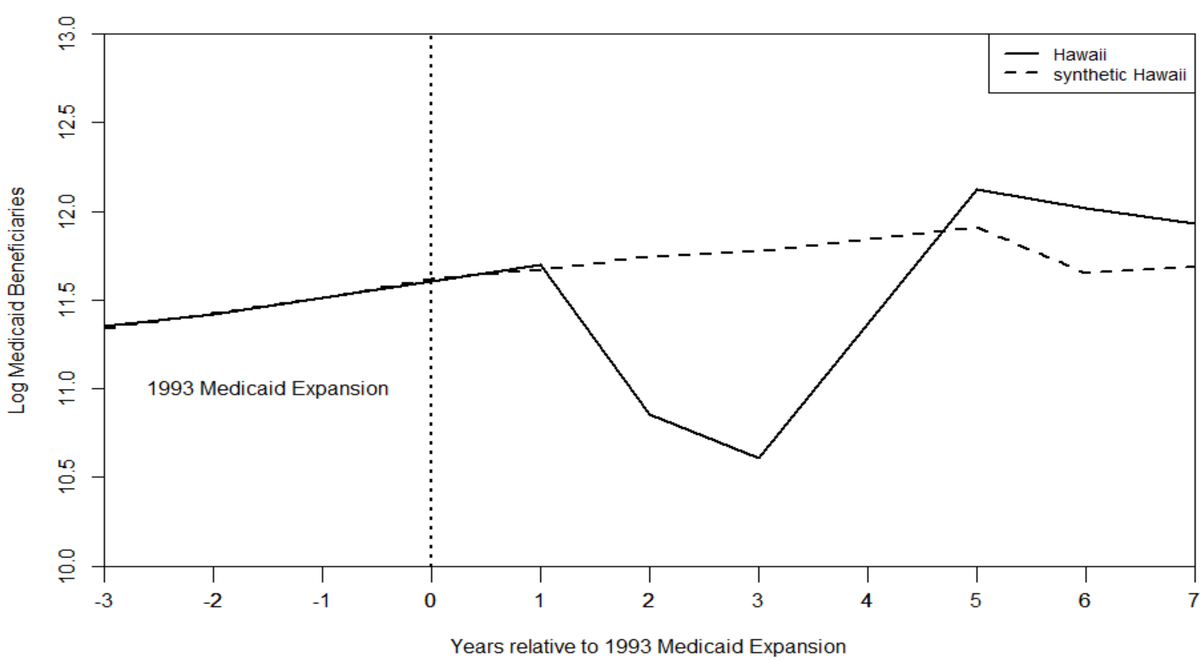

Figure I.5.C: Total Number of Medicaid Enrollees Trends in Hawaii and Synthetic Control

1993 Medicaid Expansion in Washington

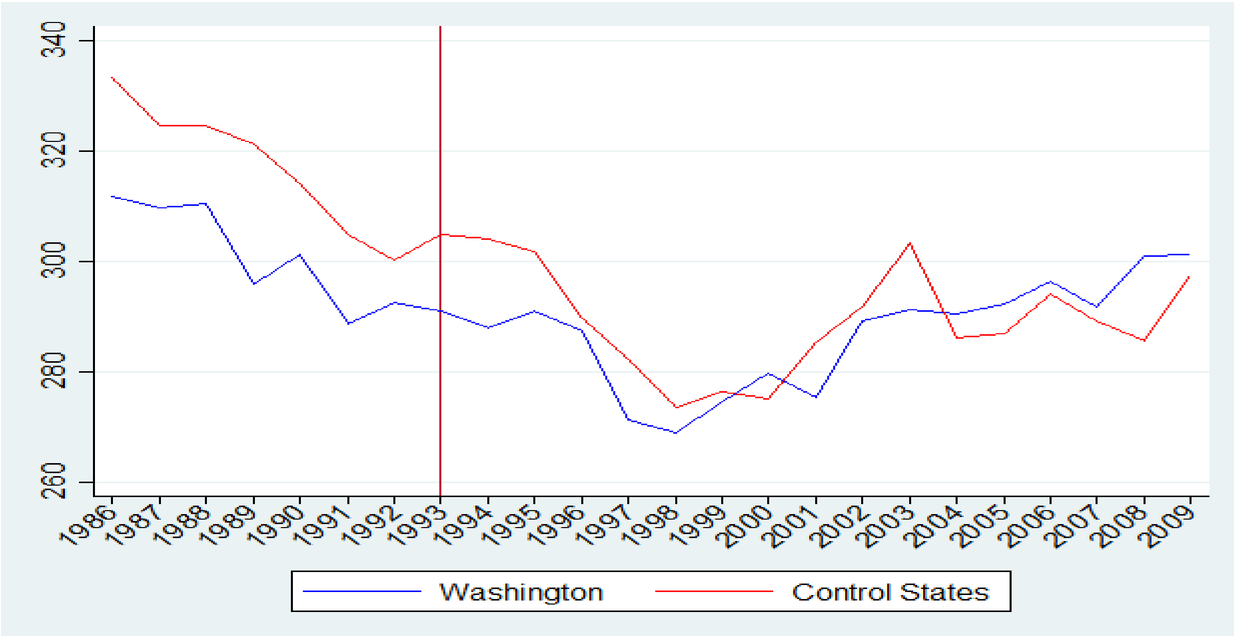

Figure I.6.A: Mortality Rate Trends in Washington

Notes: Medicaid expansion occurred in 1993 and is denoted by the red vertical line

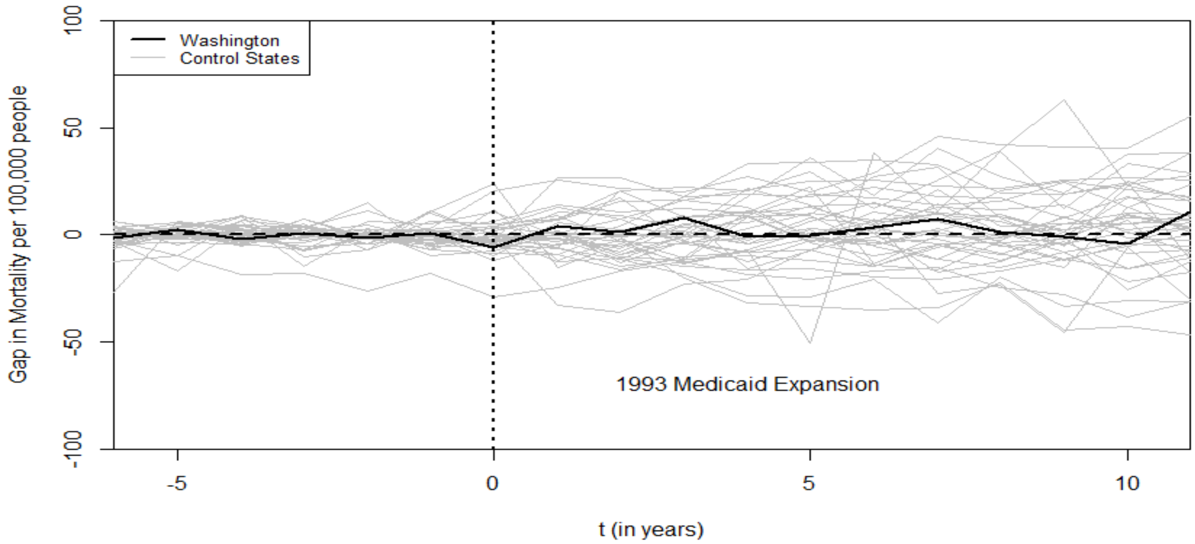

Figure I.6.B: Gap in Mortality per 100,000 people between Synthetic Control and Washington

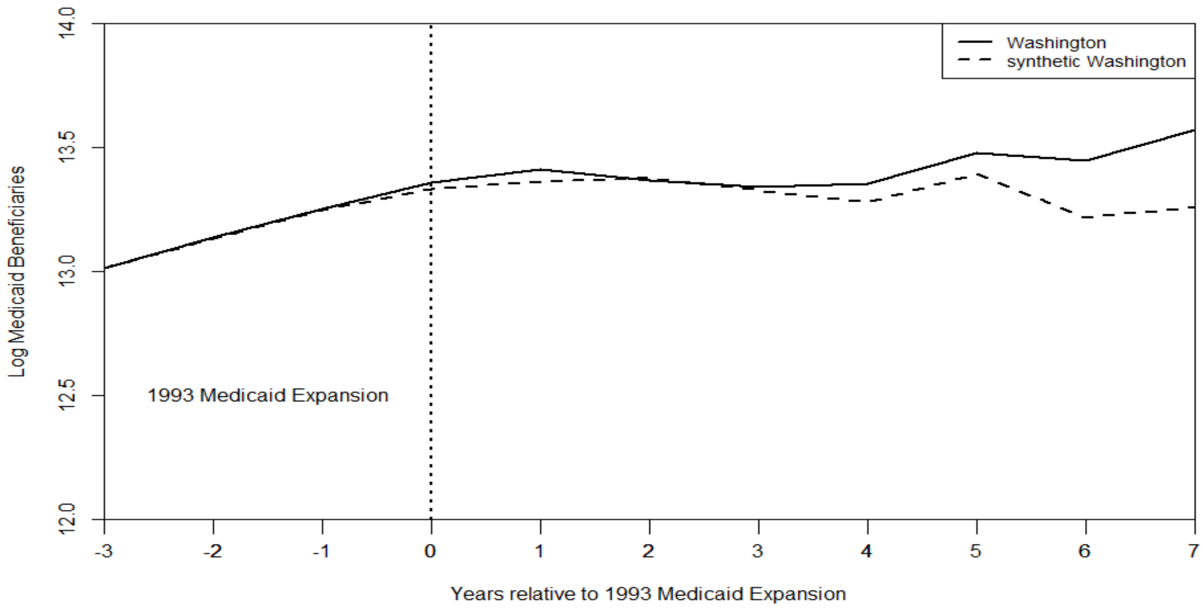

Figure I.6.C: Total Number of Medicaid Enrollees Trends in Washington and Synthetic Control

1994 Medicaid Expansion in Oregon

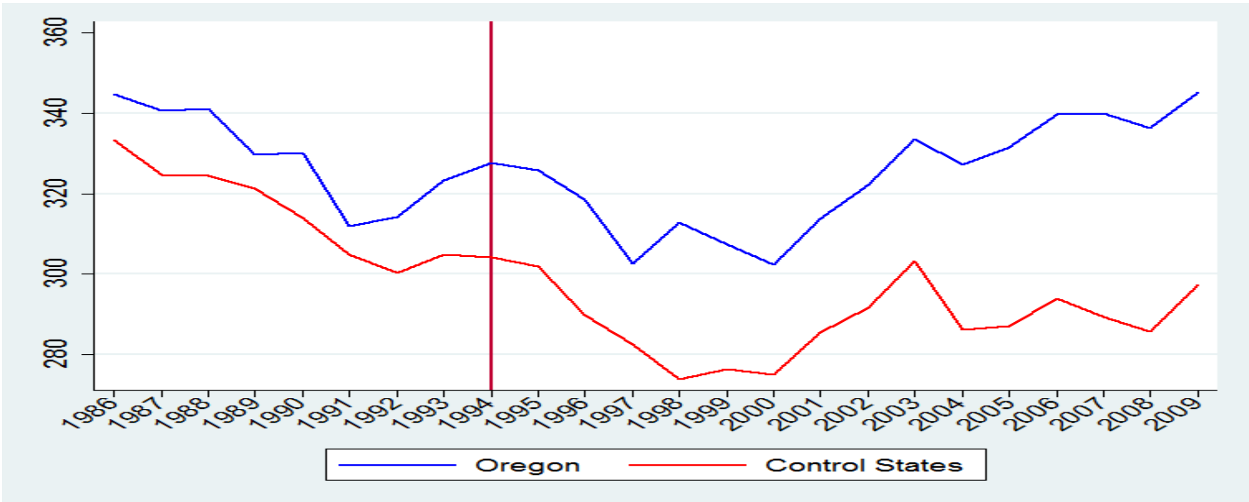

Figure I.7.A: Mortality Rate Trends in Oregon

Notes: Medicaid expansion occurred in 1994 and is denoted by the red vertical line

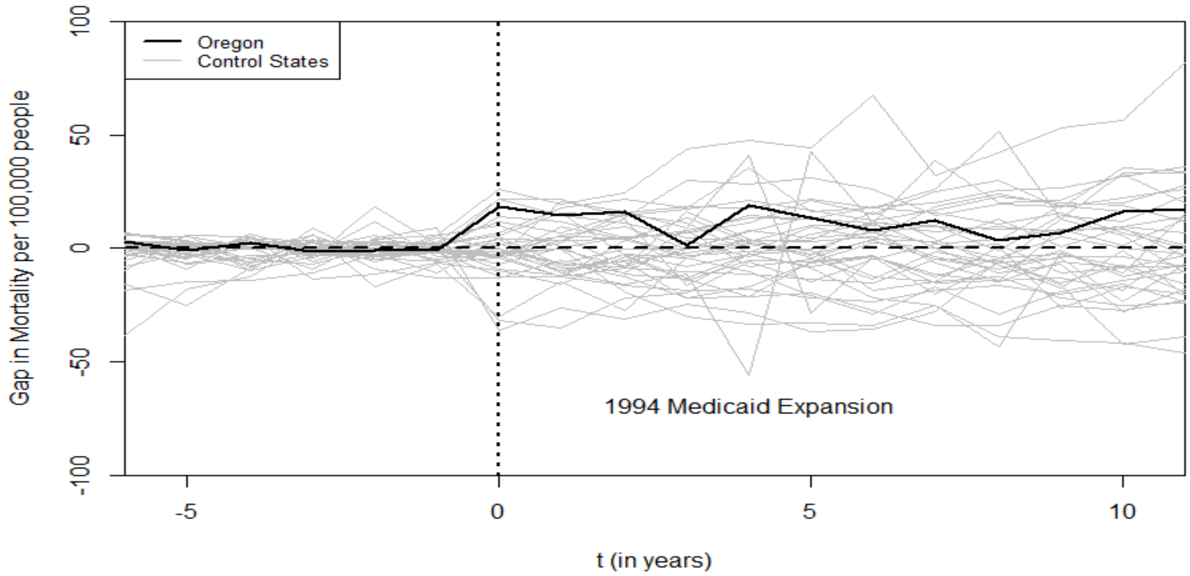

Figure I.7.B: Gap in Mortality per 100,000 people between Synthetic Control and Oregon

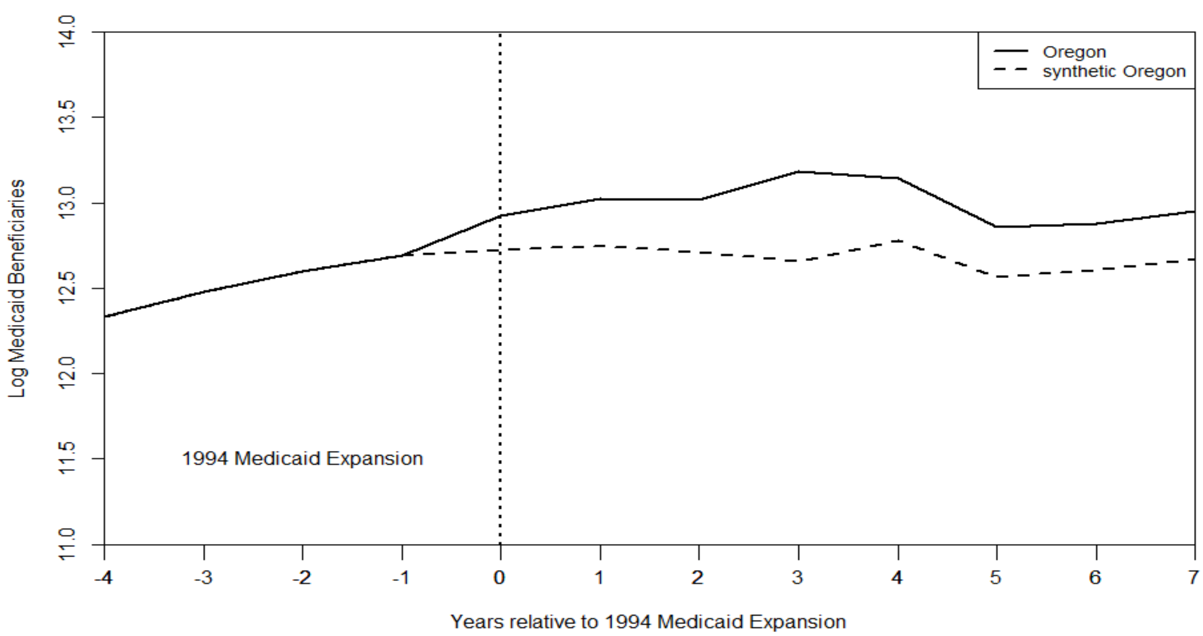

Figure I.7.C: Total Number of Medicaid Enrollees Trends in Oregon and Synthetic Control

1994 Medicaid Expansion in Tennessee

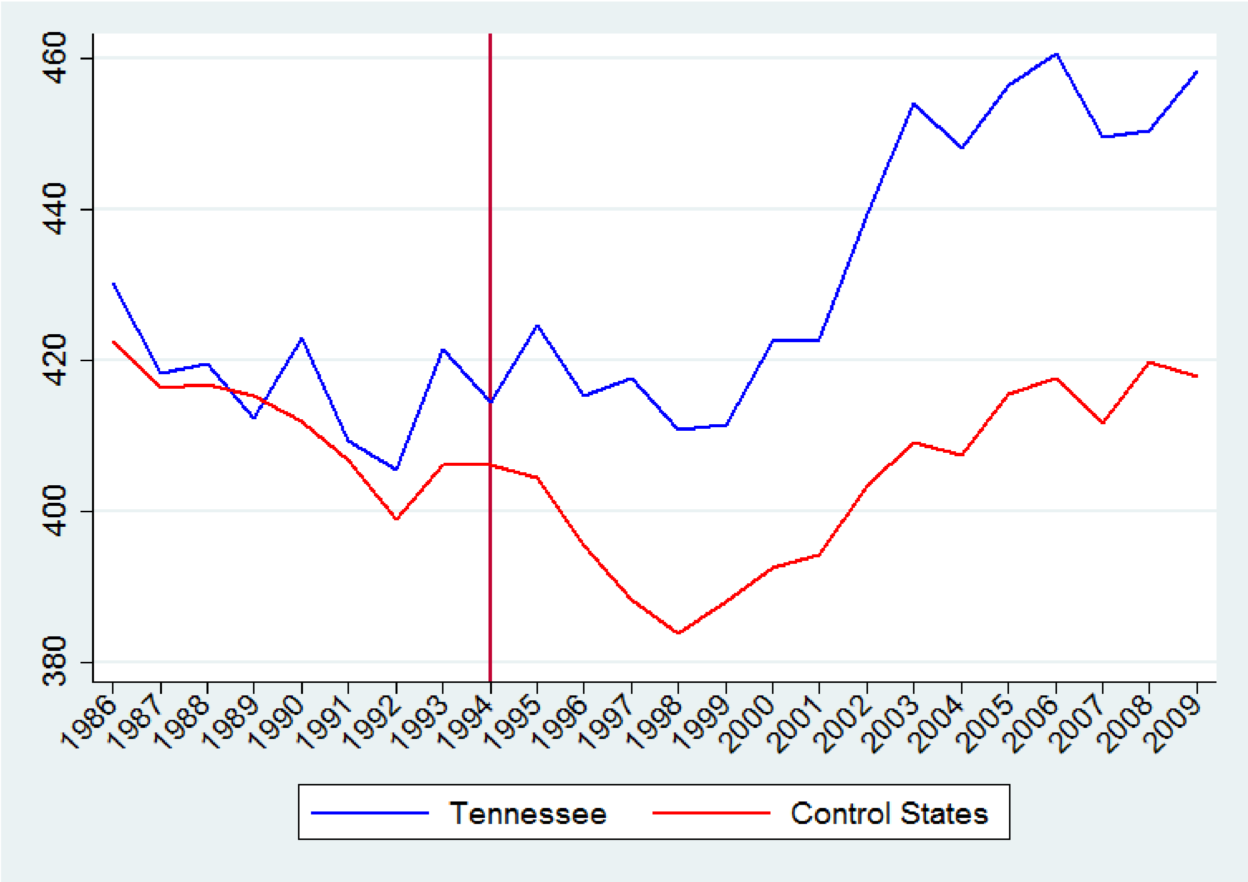

Figure I.8.A: Mortality Rate Trends in Tennessee

Notes: Medicaid expansion occurred in 1994 and is denoted by the red vertical line

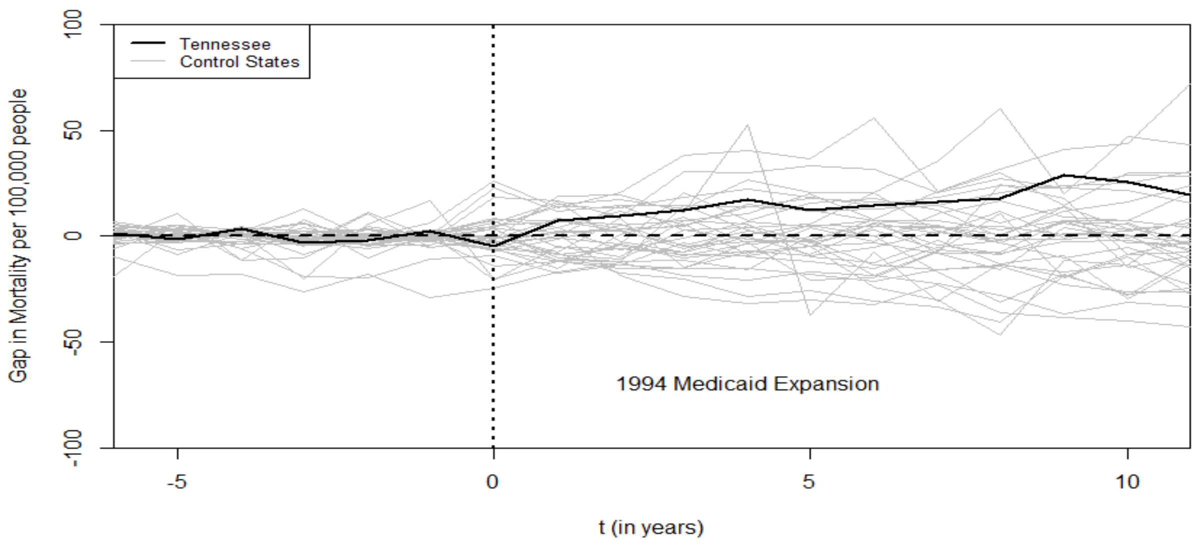

Figure I.8.B: Gap in Mortality per 100,000 people between Synthetic Control and Tennessee

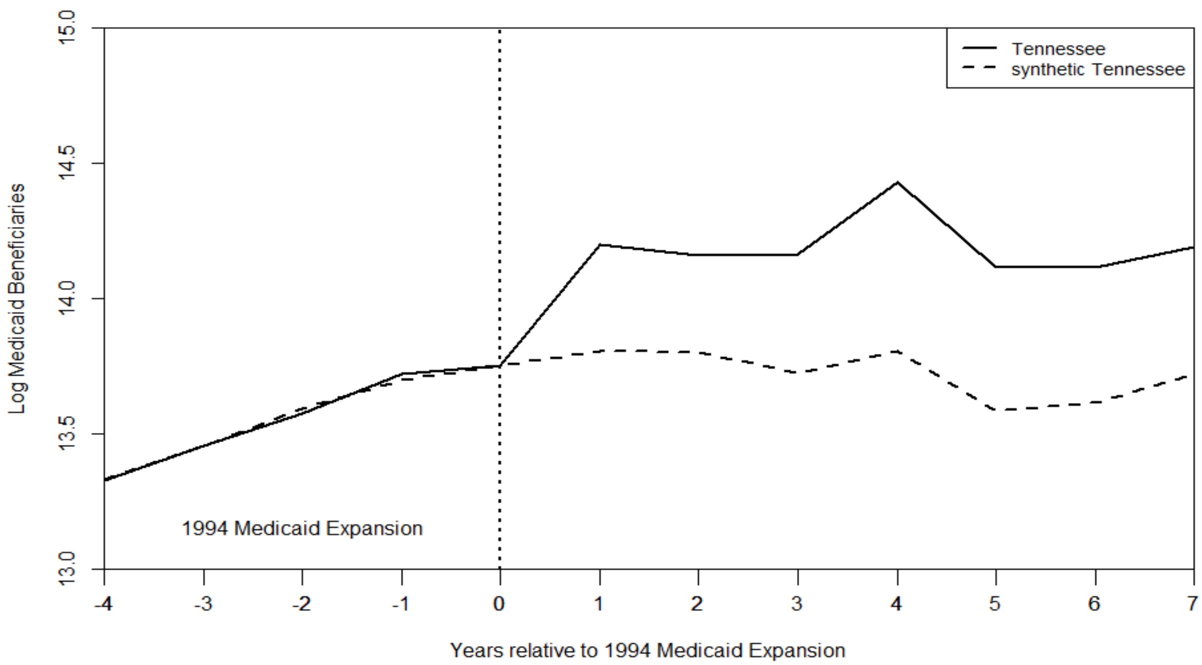

Figure I.8.C: Total Number of Medicaid Enrollees Trends in Tennessee and Synthetic Control

1995 Medicaid Expansion in Vermont

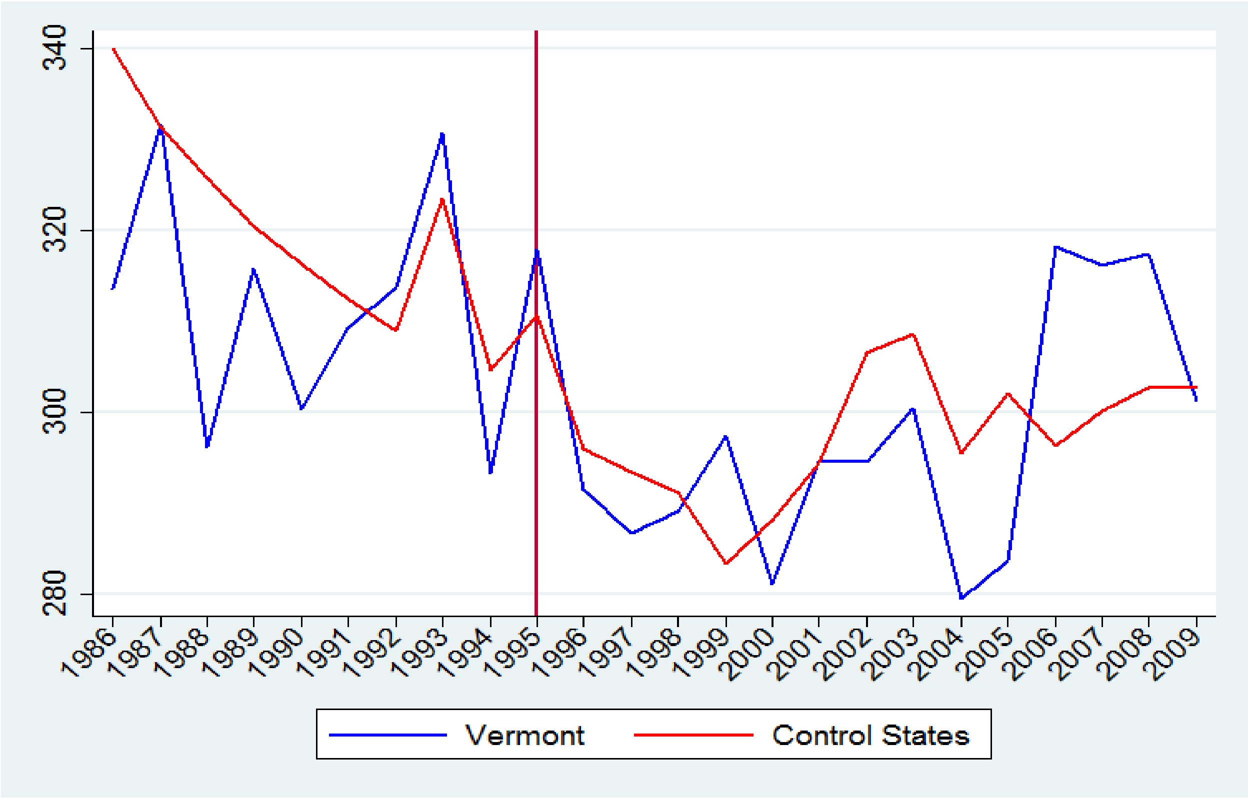

Figure I.9.A: Mortality Rate Trends in Vermont

Notes: Medicaid expansion occurred in 1995 and is denoted by the red vertical line

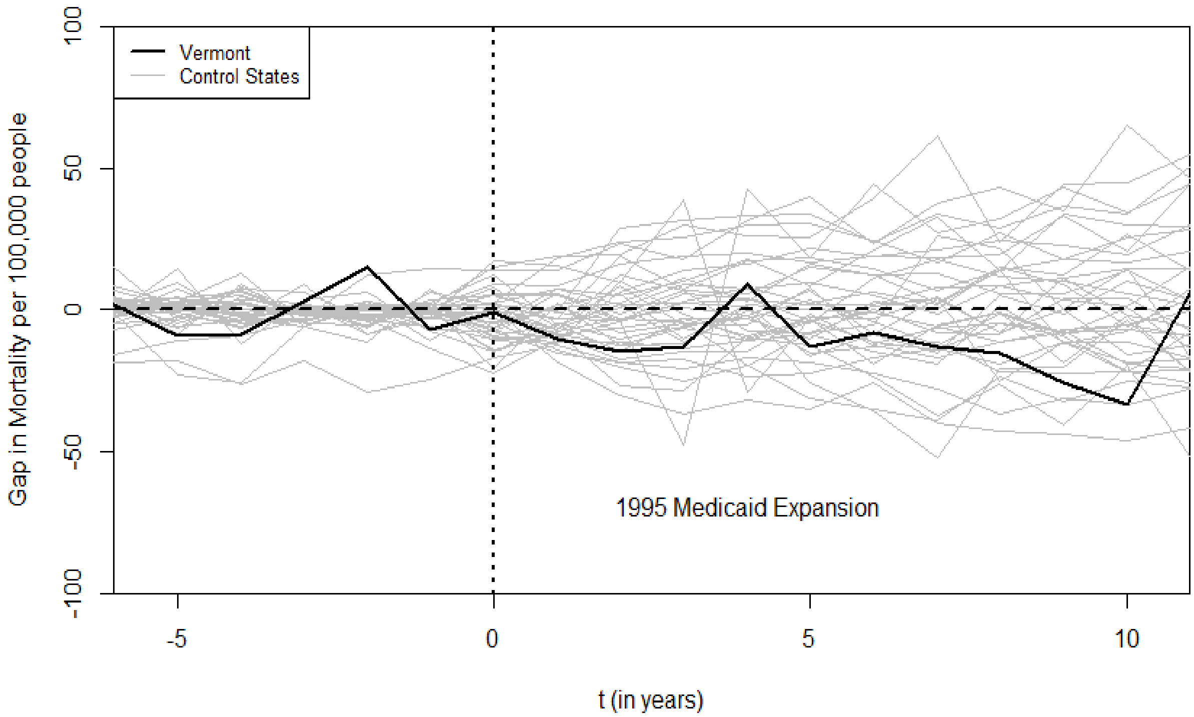

Figure I.9.B: Gap in Mortality per 100,000 people between Synthetic Control and Vermont

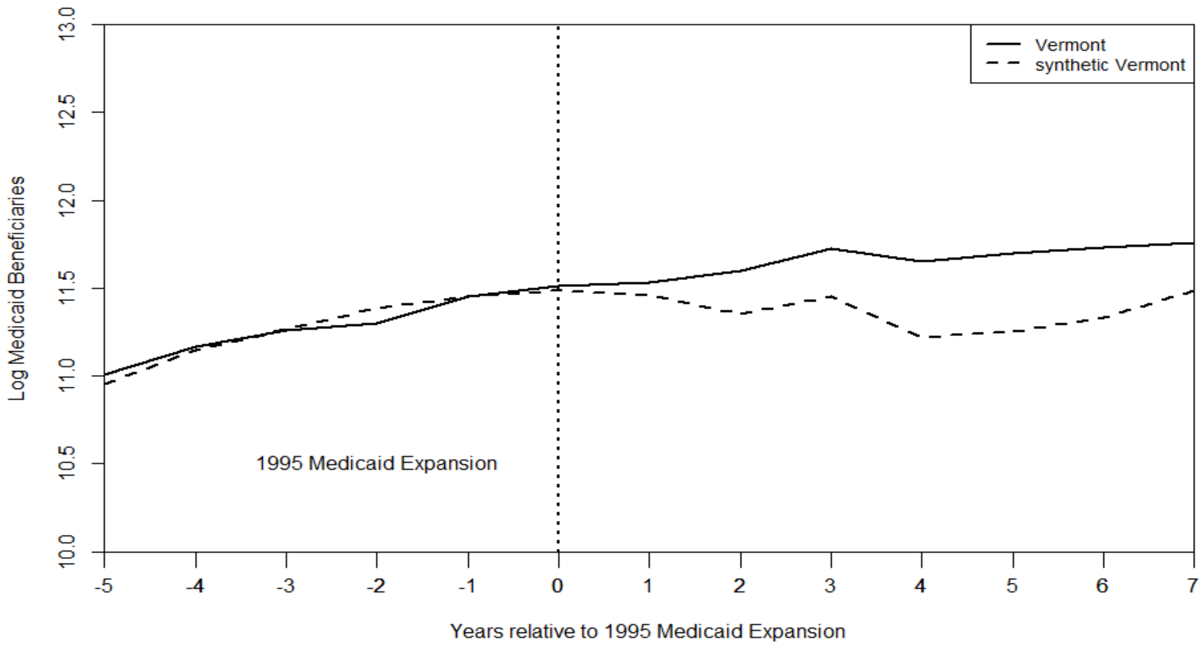

Figure I.9.C: Total Number of Medicaid Enrollees Trends in Vermont and Synthetic Control

2001 Medicaid Expansion in Arizona

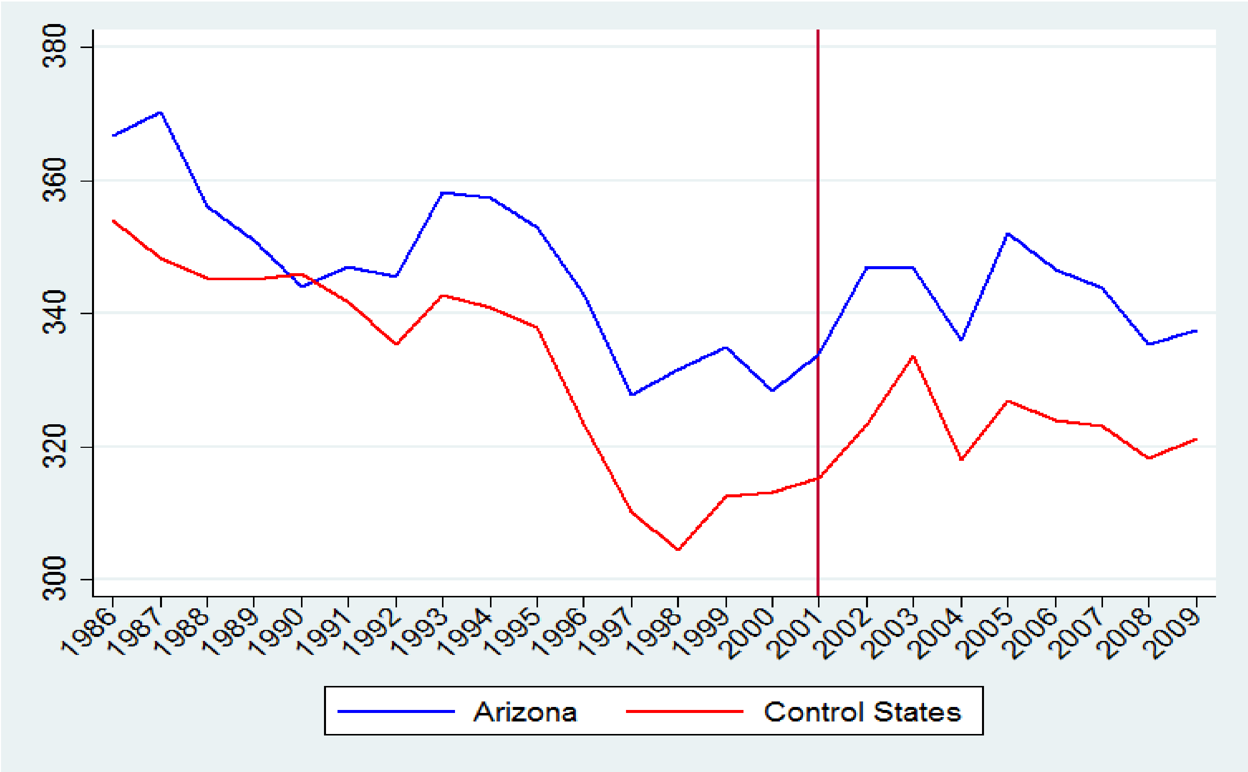

Figure I.10.A: Mortality Rate Trends in Arizona

Notes: Medicaid expansion occurred in 2001 and is denoted by the red vertical line

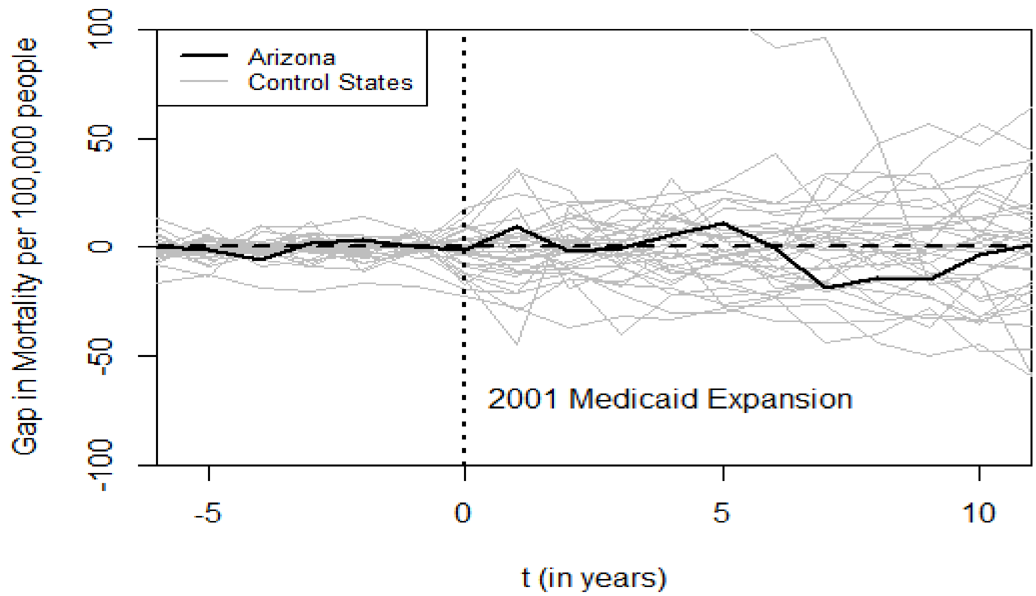

Figure I.10.B: Gap in Mortality per 100,000 people between Synthetic Control and Arizona

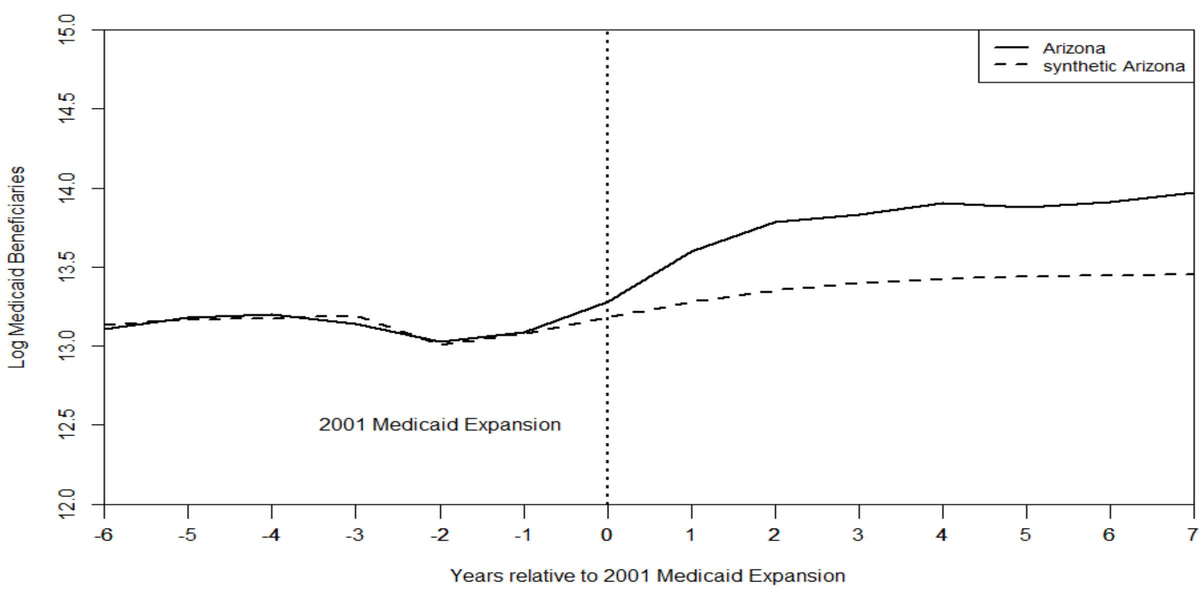

Figure I.10.C: Total Number of Medicaid Enrollees Trends in Arizona and Synthetic Control

2001 Medicaid Expansion in New York

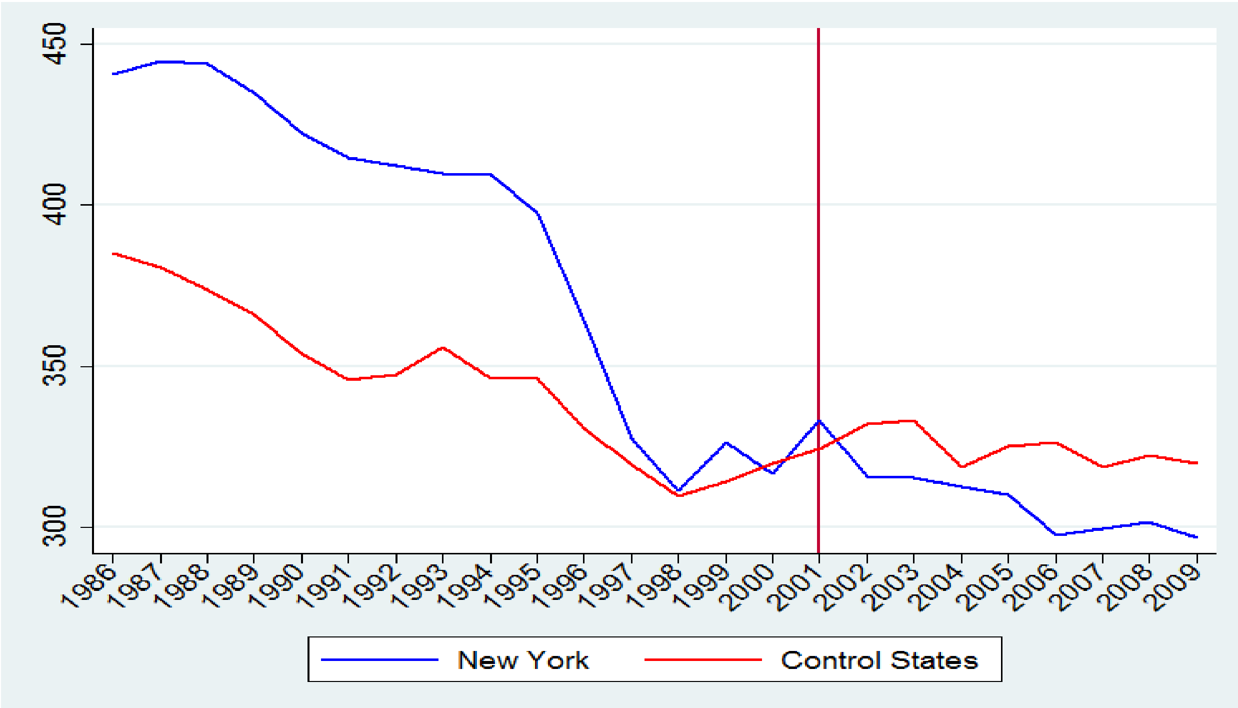

Figure I.11.A: Mortality Rate Trends in New York

Notes: Medicaid expansion occurred in 2001 and is denoted by the red vertical line

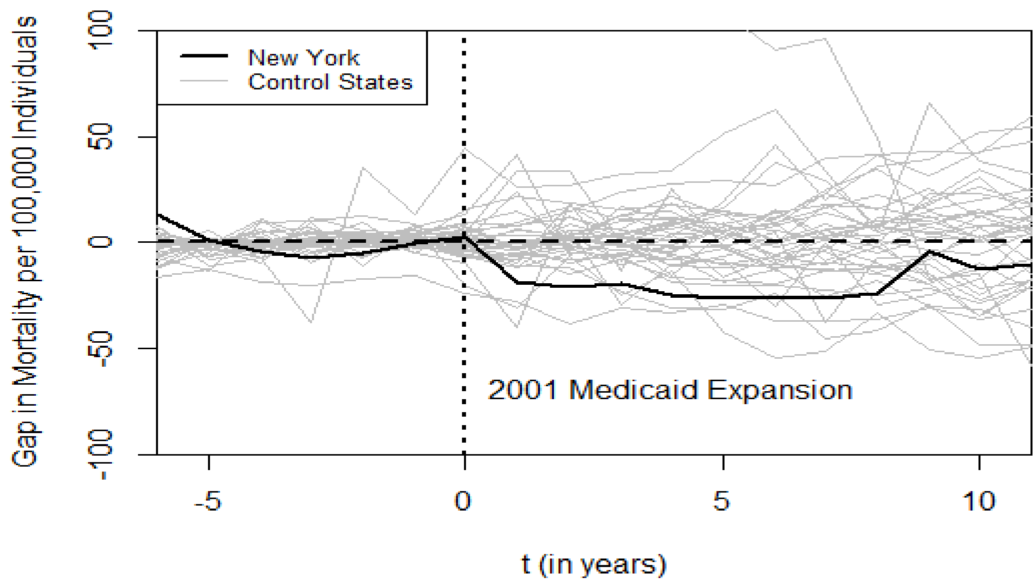

Figure I.11.B: Gap in Mortality per 100,000 people between Synthetic Control and New York

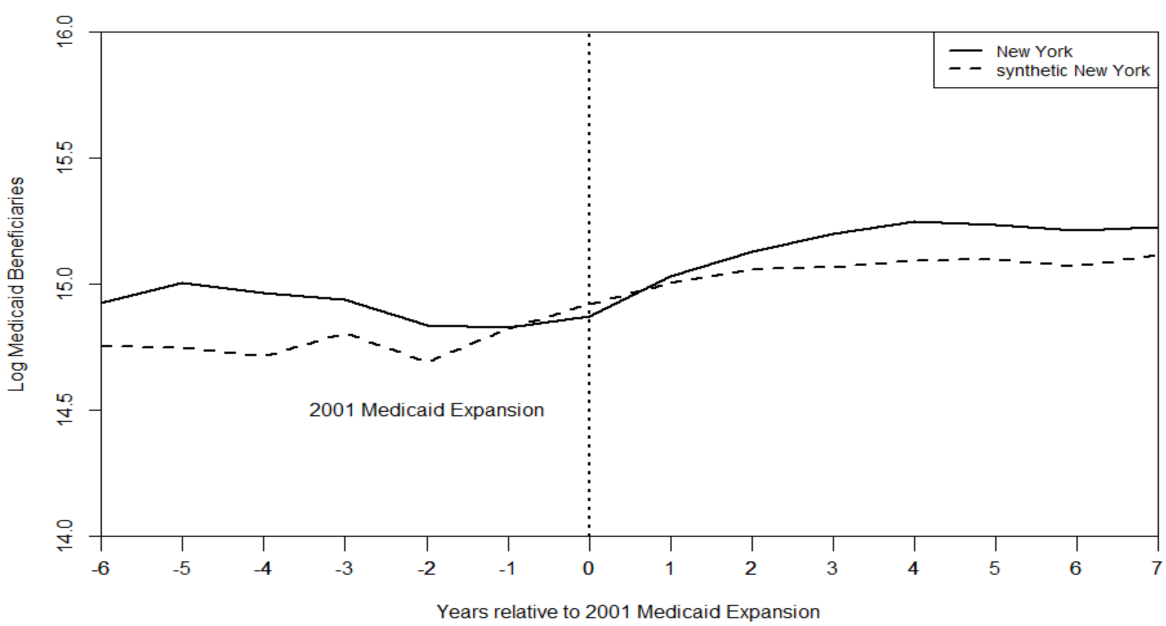

Figure I.11.C: Total Number of Medicaid Enrollees Trends in Synthetic Control and New York

2002 Medicaid Expansion in Maine

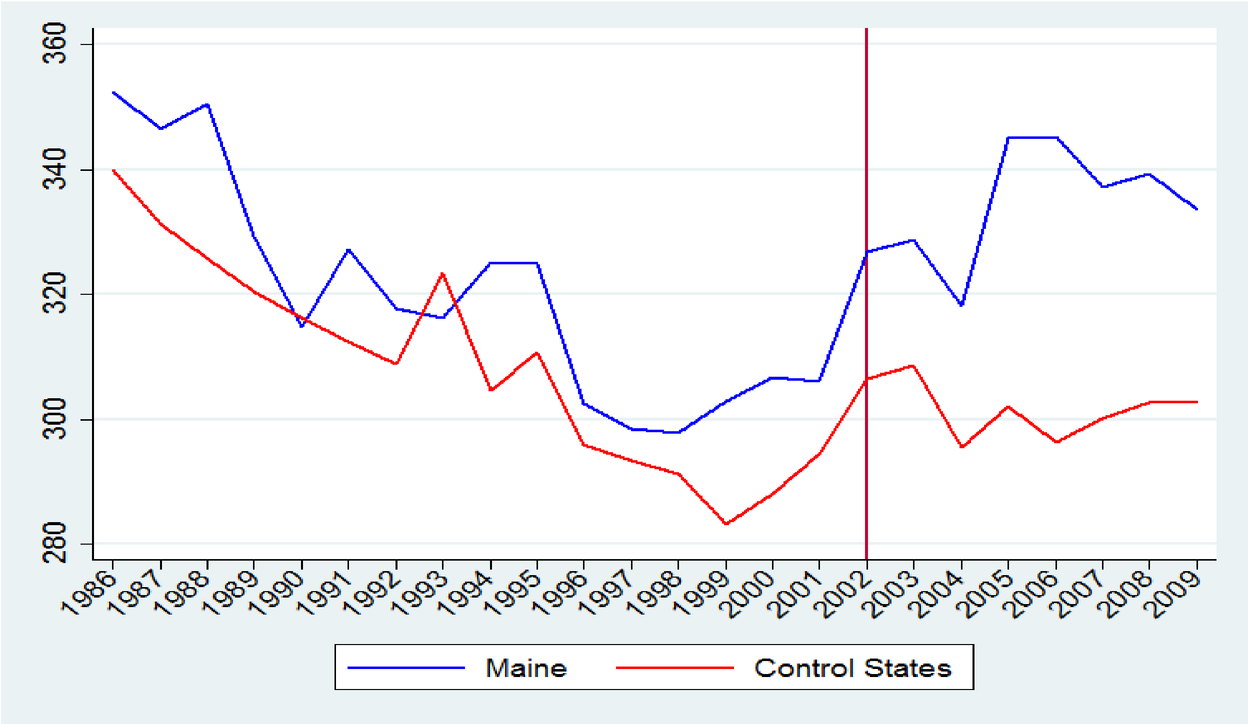

Figure I.12.A: Mortality Rate Trends in Maine

Notes: Medicaid expansion occurred in 2002 and is denoted by the red vertical line

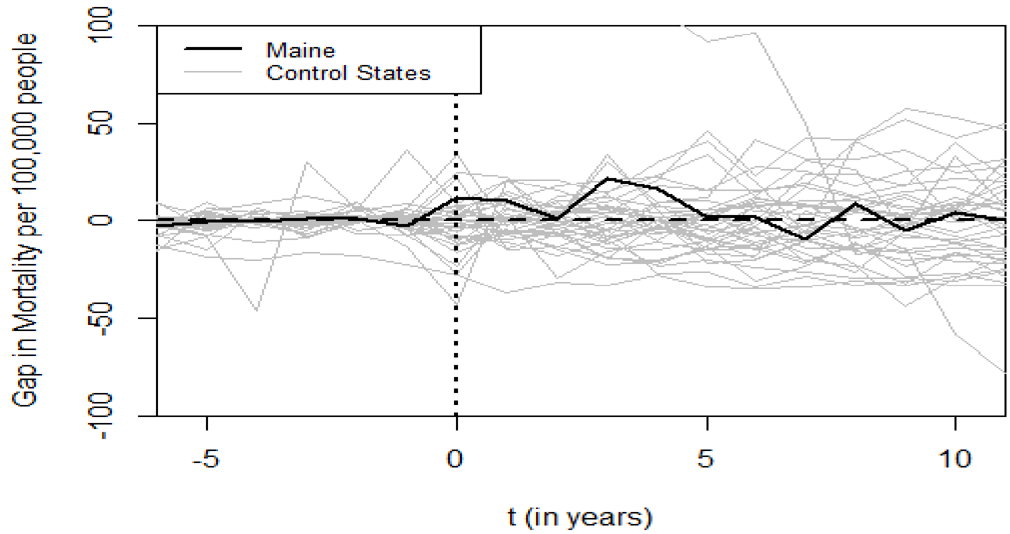

Figure I.12.B: Gap in Mortality per 100,000 people between Synthetic Control and Maine

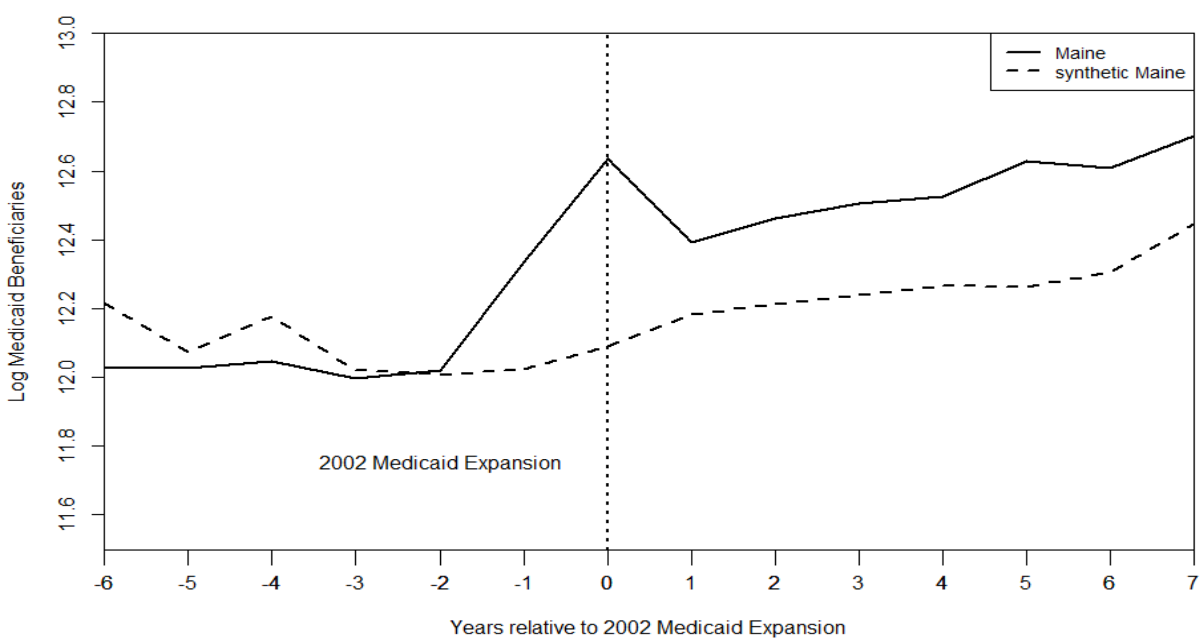

Figure I.12.C: Total Number of Medicaid Enrollees Trends in Maine and Synthetic Control:

2002 Medicaid Expansion in New Mexico

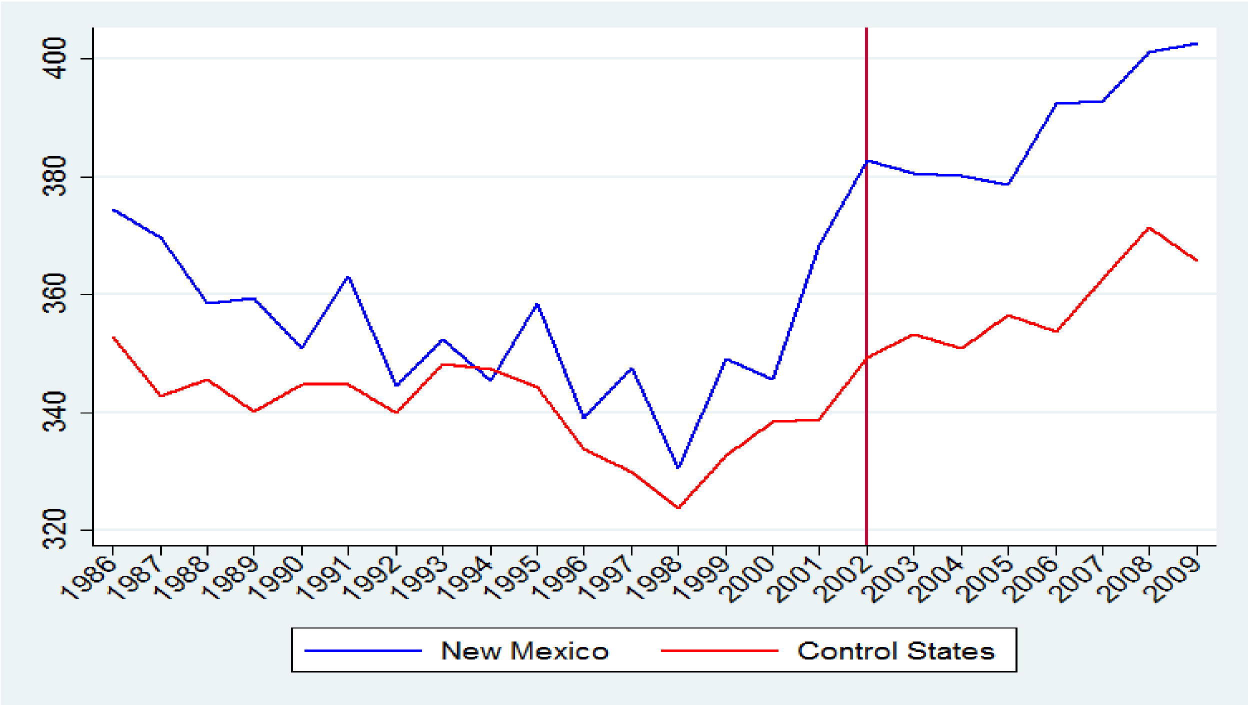

Figure I.13.A: Mortality Rate Trends in New Mexico

Notes: Medicaid expansion occurred in 2002 and is denoted by the red vertical line

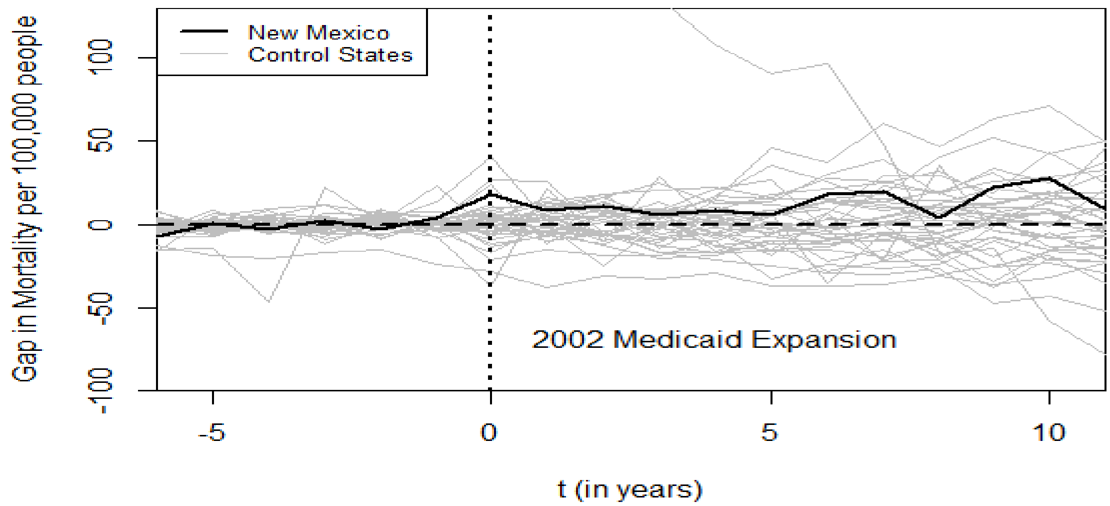

Figure I.13.B: Gap in Mortality per 100,000 people between Synthetic Control and New Mexico

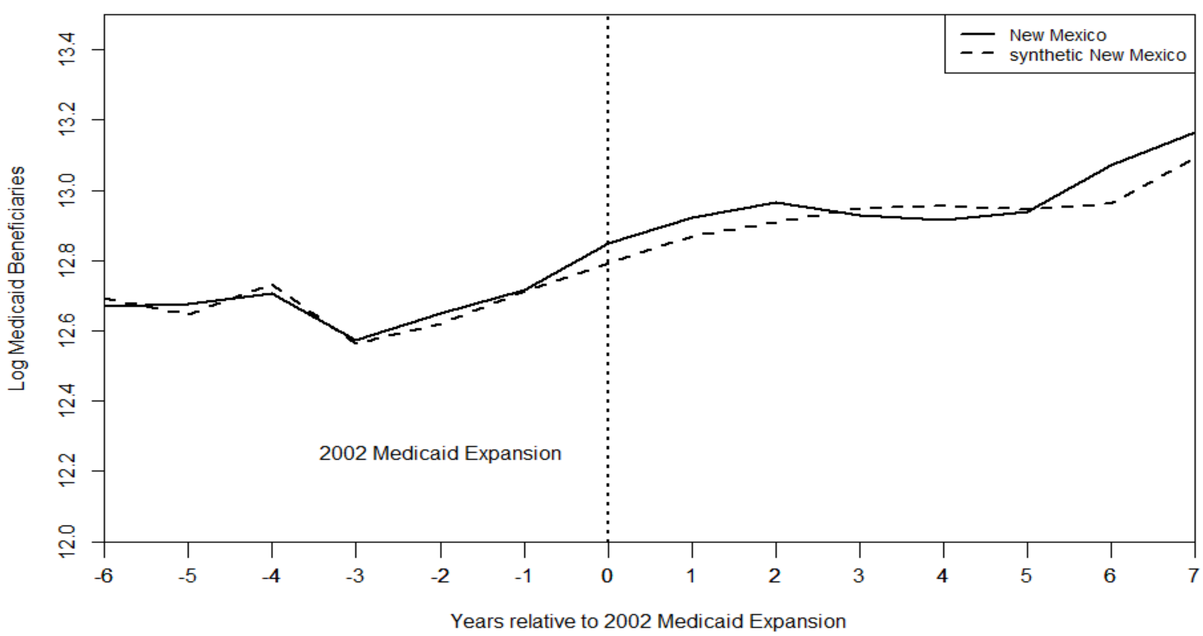

Figure I.13.C: Total Number of Medicaid Enrollees Trends in New Mexico and Synthetic Control

2002 Medicaid Expansion in Utah

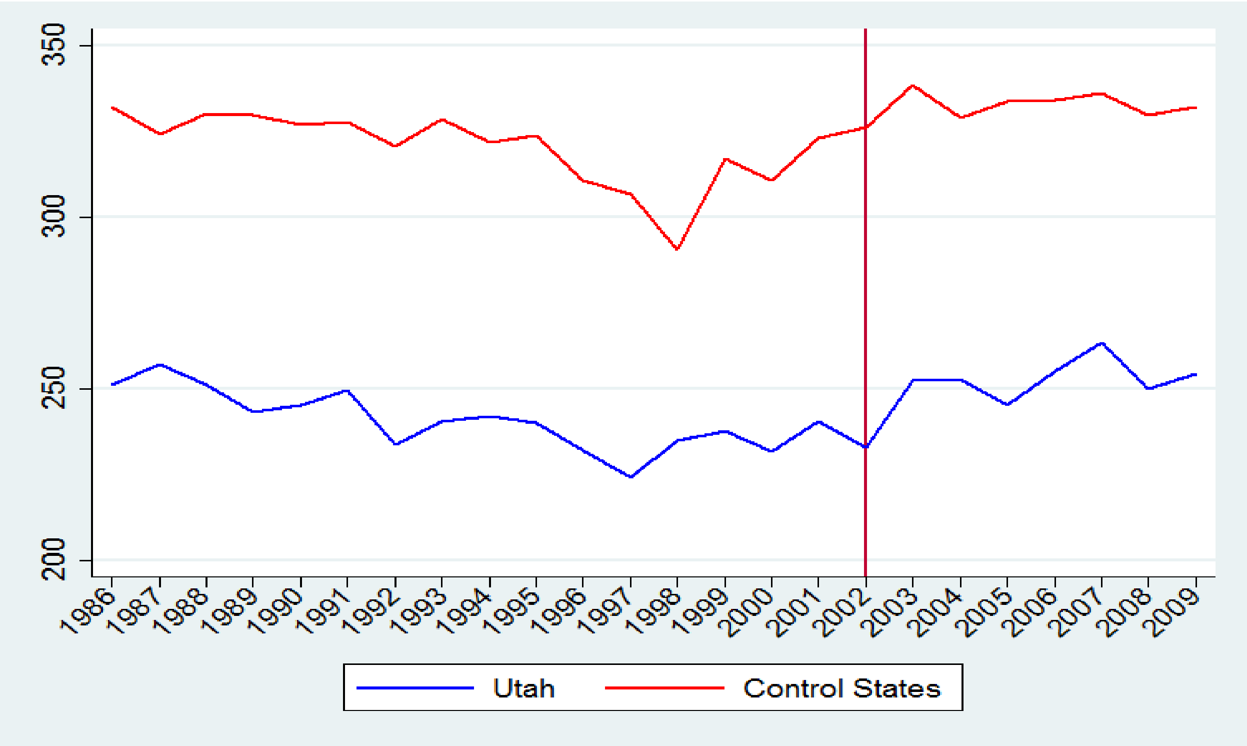

Figure I.14.A: Mortality Rate Trends in Utah

Notes: Medicaid expansion occurred in 2002 and is denoted by the red vertical line

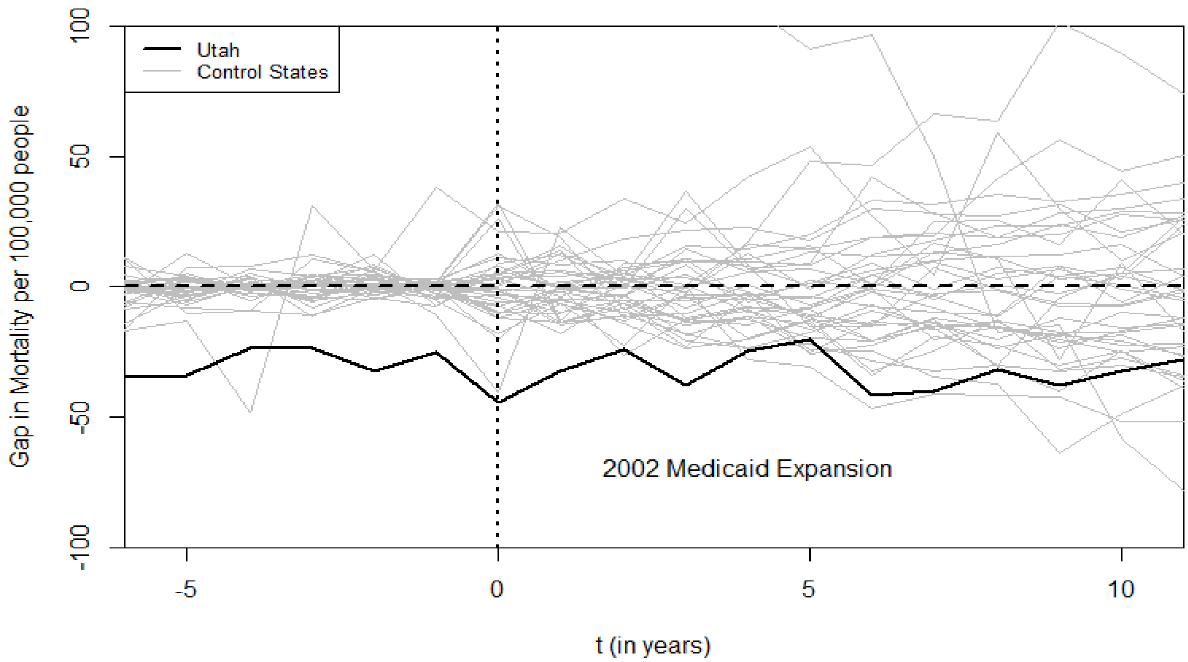

Figure I.14.B: Gap in Mortality per 100,000 people between Synthetic Control and Utah

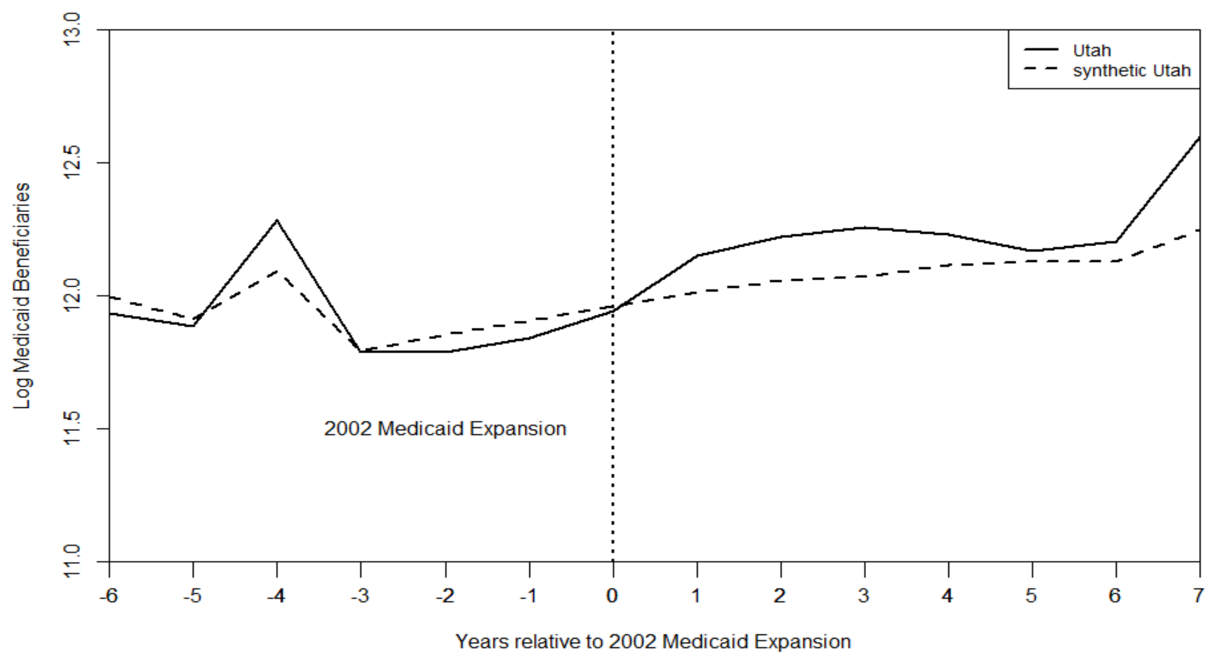

Figure I.14.C: Total Number of Medicaid Enrollees Trends in Utah and Synthetic Control

Appendix J: Tables

Table J.1: Summary Statistics

Medicaid Expanding States: AZ, HI, ME, MN, NM, NY, OR, TN, UT, VT, WA; N = 308

VARIABLES	Mean	Std. Dev.	Min	Max
Mortality Rate	323.036	56.656	224.517	474.264
% female	0.506	0.006	0.488	0.522
% Hispanic	0.109	0.121	0.005	0.473
% white	0.836	0.195	0.240	0.989
% aged 20-34	0.218	0.024	0.168	0.281
% aged 35-44	0.147	0.015	0.115	0.173
% aged 45-54	0.126	0.022	0.074	0.166
% aged 55-64	0.096	0.020	0.060	0.153
% married	9.550	3.563	4	22.6
% high school diploma	24.884	4.675	13.835	36.834
Average income	53132.78	7322.88	35988	72335
Unemployment	5.630	1.681	2.4	11.1
Poverty	12.799	3.679	5.7	25.5

Medicaid Non-Expanding States: Rest of the US excluding Massachusetts; N = 1,120

VARIABLES	Mean	Std. Dev.	Min	Max
Mortality Rate	369.3958	79.31537	258.5546	863.1047
% female	0.509	0.009	0.472	0.535
% Hispanic	0.069	0.075	0.004	0.383
% white	0.827	0.121	0.307	0.985
% aged 20-34	0.218	0.026	0.172	0.318
% aged 35-44	0.147	0.015	0.109	0.201
% aged 45-54	0.126	0.022	0.074	0.166
% aged 55-64	0.097	0.017	0.052	0.146
% married	9.783	11.252	4	114
% high school diploma	23.440	6.612	10.993	60.129
Average income	52,098.79	8,714.83	31,819	77.506
Unemployment	5.772	1.982	2.3	13.8
Poverty	12.918	3.810	2.9	27.2

Sources: CDC, Census, Bureau of Labor Statistics, National Cancer Institute

Table J.2: Dynamic SCD Estimates of Medicaid's Impact on Mortality

	(1)	(2)	(3)
Post-treatment Year	AZ	HI	ME
1	9.01	-2.57	3.02
2	-0.47	15.61	1.35
3	2.22	18.43*	18.40
4	10.83	23.02	17.05
5	15.90	25.61*	1.16

6	1.11	2.14	2.44
7	-17.54	14.96	-4.14

	(4) MN	(5) NM	(6) NY
1	-29.12	11.88	-22.48
2	-24.82	13.44*	-22.12**
3	-16.65	9.73*	-17.09*
4	-12.72	22.86**	-25.87**
5	-18.42	26.20**	-32.06**
6	-20.65	33.10**	-15.65
7	-16.76	36.31**	-17.60*

	(7) OR	(8) TN	(9) UT
1	12.94	13.02	-32.56
2	17.84*	14.45	-24.34
3	5.14	17.02	-38.05
4	23.75	22.34	-24.87
5	17.59	22.78	-20.21
6	12.03	28.98	-41.51
7	14.96	24.36	-40.20

	(10) VT	(11) WA
1	1.39	-6.76
2	-7.70	-1.80
3	22.25	0.37
4	-17.94	-5.19
5	-19.09	-8.47
6	-33.79**	-8.48
7	-26.18	-2.60

Notes: Donor pool states for AZ: CA, NV, CO; OR & WA controls: CA, ID; MN controls: WI, IA, ND, SD; UT controls: NV, WY, ID, CO; NM controls: TX, CO, OK; ME & VT controls: NH, RI, CT; NY: PA, NJ, CT, RI; TN: KY, AK, MO, IL, MS, AL, GA, NC, VA. Pseudo p-values from permutation tests; *** p<0.01, ** p<0.05, * p<0.1.

References

Abadie, A. and Imbens, G. (2006). Large sample properties of matching estimators for average treatment effects. *Econometrica*, 74(1):235–267.

Allison, P., and Waterman, R. (2002). Fixed-Effects Negative Binomial Regression Models. *Sociological Methodology*, 247-265.

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

Abadie, A., Diamond, A., & Hainmueller, J. (2014). Comparative Politics and the Synthetic Control Method. *American Journal of Political Science*, 1-16.

Aidt, T. (2003). Economic Analysis of Corruption. *The Economic Journal*, 632-652.

Allingham, M. G. and Sandmo, A. (1972). Income tax evasion: a theoretical analysis. *Journal of Public Economics*, 1(3-4):323–338.

Alm, J. (2012). Measuring, explaining, and controlling tax evasion: lessons from theory, experiments, and field studies. *International Tax and Public Finance*, 19(1):54–77.

Alm, J. and Martinez-Vazquez, J. (2003). "sizing" the problem of the hard-to-tax. International center for public policy working paper series, at aysps, gsu, International Center for Public Policy, Andrew Young School of Policy Studies, Georgia State University.

Alternativas, F. (2007). *Urbanismo y democracia. Alternativas para evitar la corrupcion.* Madrid: Fundacion Alternativas.

Andreoni, J., Erard, B., and Feinstein, J. (1998). Tax compliance. *Journal of Economic Literature*, 36(2):818–860.

Andvig, J., and Moene, K. O. (1990). How Corruption May Corrupt. *Journal of Economic Behavior and Organization*, 63-76.

Arellano, M., and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies*, 277-297.

Ariel, B. (2012). Deterrence and moral persuasion effects on corporate tax compliance: Findings from a randomized controlled trial. *Criminology*, pages 27–69.

Arellano, M., and Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 29-51.

Artavanis, N. (2016). Vat rates and tax evasion: Evidence from the restaurant industry in Greece.

Artavanis, N., Morse, A., and Tsoutsoura, M. (2015). Tax evasion across industries: Soft credit evidence from Greece. NBER Working Papers 21552, National Bureau of Economic Research, Inc.

Bank of Spain. (2013). *MORTGAGE LOAN ACCESS GUIDE.* Madrid: Bank of Spain.

Batzilis, D. (2014). *Does Electoral Competition Reduce Corruption? Evidence from Municipality Audits in Greece.* Chicago: University of Chicago PhD Dissertation.

Baum, C. (2001). Stata: The language of choice for time series analysis? *The Stata Journal*, 1-16.

Baicker, K., Taubman, S. L., Allen, H. L., Bernstein, M., Gruber, J. H., Newhouse, J. P., ... & Finkelstein, A. N. (2013). The Oregon experiment—effects of Medicaid on clinical outcomes. *New England Journal of Medicine*, 368(18), 1713-1722.

Becker, G. S. (1968). Crime and Punishment: An Economic Approach. *Journal of Political Economy*, 169-217.

Beets, D. (2005). Understanding the Demand-Side Issues of International Corruption. *Journal of Business Ethics*, 65-81.

Benabou, R. and Tirole, J. (2006). Incentives and prosocial behavior. *American Economic Review*, 96(5):1652–1678.

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How Much Should we Trust Differences-in-Differences Estimates. *Quarterly Journal of Economics*, 249-275.

Black, B., Espín-Sánchez, J. A., French, E., & Litvak, K. (2017). The Long-Term Effect of Health Insurance on Near-Elderly Health and Mortality. *American Journal of Health Economics*.

Bott, K., Cappelen, A. W., Sorensen, E. O., and Tungodden, B. (2014). You've got mail: A randomised field experiment on tax evasion. Discussion Paper 26, Norwegian School of Economics, Department of Economics.

Bliss, C., and Di Tella, R. (1997). Does Competition Kill Corruption? *Journal of Political Economy*, 1001-1023.

Blundell, R., and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. *Journal of Econometrics*, 115-143.

Blumenthal, M., Christian, C. W., and Slemrod, J. (2001). Do normative appeals affect tax compliance? evidence from a controlled experiment in minnesota. *National Tax Journal*, 54(n. 1):125–38.

Braun, M., and Di Tella, R. (2004). Inflation, Inflation Variability, and Corruption. *Economics and Politics*, 77-100.

Brissimis, S., and Vlassopoulos, T. (2007). The Interaction Between Mortgage Financing and Housing Prices in Greece. *Bank of Greece Working Paper*.

Brollo, F., Nannicini, T., Perotti, R., and Tabellini, G. (2013). The Political Resource Curse. *American Economic Review*, 1759-1796.

Burbidge, J. B., Magee, L., and Robb, A. L. (1988). Alternative transformations to handle extreme values of the dependent variable. *Journal of the American Statistical Association*, 83(401):123– 127.

Bujko, M., Fischer, C., Krieger, T., and Meierrieks, D. (2015). *How Institutions Shape Land Deals: The Role of Corruption.* CESIFO WORKING PAPER NO. 5178.

Bun, M. J., and Windmeijer, F. (2010). The weak instrument problem of the system GMM estimator in dynamic panel data models. *The Econometrics Journal*, 95–126.

Calonico, S., Cattaneo, M., and Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82:2295–2326.

Cameron, C., and Trivedi, P. (2005). *Microeconometrics: Methods and Applications.* Cambridge: Cambridge University Press.

Caselli, F., and Michaels, G. (2013). Do Oil Windfalls Improve Living Standards? Evidence from Brazil. *American Economic Journal: Applied Economics*, 208-38.

Castro, L. and Scartascini, C. (2015). Tax compliance and enforcement in the pampas evidence from a field experiment. *Journal of Economic Behavior and Organization*, 116(C):65–82.

Chang, H.-J. (2010). Institutions and Economic Development: Theory, Policy and History. *Journal of Institutional Economics*, 1-25.

Cheung, H. Y., and Chan, A. (2008). Corruption Across Countries; Impacts from Education and Cultural Dimensions. *The Social Science Journal*, 223-239.

Clemente, J., Montanes, A., and Reyes, M. (1998). Testing for a unit root in variables with a double change in the mean. . *Economic Letters*, 175-182.

Costas-Pérez, E., Solé-Ollé, A., and Sorribas-Navarro, P. (2013). Corruption scandals, voter information, and accountability. *European Journal of Political Economy*, 469-484.

Del Carpio, L. (2014). Are the neighbors cheating? evidence from a social norm experiment on property taxes in peru.

Del Monte, A., and Papagni, E. (2007). The determinants of corruption in Italy: Regional panel data analysis. *European Journal of Political Economy*, 379-396.

Doerrenberg, P. and Schmitz, J. (2015). Tax compliance and information provision: A field experiment with small firms. ZEW Discussion Papers 15-028, ZEW - Zentrum fÃijr Europaische Wirtschaftsforschung / Center for European Economic Research.

Dong, B., and Torgler, B. (2013). Causes of Corruption: Evidence from China. *China Economic Review*, 152-169.

Dwenger, N., Kleven, H., Rasul, I., and Rincke, J. (2014). Extrinsic vs intrinsic motivations for tax compliance. evidence from a randomized field experiment in germany. Technical report.

ECB Structural Issues Report. (2009). *Housing Finance in the Euro Area.* Frankfurt: European Central Bank.

El Bahnasawy, N., and Revier, D. (2012). The determinants of corruption: Cross-Country-Panel-Data Analysis. *The Developing Economies*, pp. 311-333.

Emerson, P. (2006). Corruption, Competition and Democracy. *Journal of Development Economics*, 193-212.

Erard, B. (1992). *The influence of tax audits on spending behavior.*

Eurofound (2013). Information campaign on tax compliance, estonia. Technical report, Eurofound.

Fellner, G., Sausgruber, R., and Traxler, C. (2009). Testing enforcement strategies in the field: Legal threat, moral appeal and social information. Working Paper Series of the Max Planck Institute for Research on Collective Goods 2009 31, Max Planck Institute for Research on Collective Goods.

Fernández-Villaverde, J., Garicano, L., and Santos, T. (2013). Political Credit Cycles: The Case of the Eurozone. *Journal of Economic Perspectives*, 145-166.

Ferraz, C., and Finan, F. (2009). Electoral Accountability and Corruption: Evidence from the Audits of Local Governments. *NBER Working Paper Series*.

Finkelstein, A., Taubman, S., Wright, B., Bernstein, M., Gruber, J., Newhouse, J. P., ... & Baicker, K. (2012). The Oregon health insurance experiment: evidence from the first year. *The Quarterly Journal of Economics*, 127(3), 1057-1106.

Frechette, G. R. (2001). A Panel Data Analyis of the Time-Varying Determinants of Corruption. *EPCS Conference Proceedings*.

Gelman, A. and Imbens, G. (2014). Why high-order polynomials should not be used in regression discontinuity designs. NBER Working Papers 20405, National Bureau of Economic Research, Inc.

Gentier, A. (2012). Spanish Banks and the Housing Crisis: Worse than the Subprime Crisis? *International Journal of Business*, 343-351.

Gimeno, R., and Martínez-Carrascal, C. (2006). L ainteracción entre el precio de la vivienda y el crédito a hogares destinado a su adquisición. *Boletín Económico*.

Glaeser, E. L., Gyourko, J., and Saiz, A. (2008). Housing Supply and Housing Bubbles. *Journal of Urban Economics*, 198-217.

Golden, M. (2015). *A Global View of Political Corruption.* Barcelona: Institut d' Economia de Barcelona.

Guyton, J., Manoli, D. S., Schafer, B., and Sebastiani, M. (2016). Reminders and recidivism: Evidence from tax filing and eitc participation among low-income nonfilers. NBER Working Papers 21904, National Bureau of Economic Research, Inc.

Hadley, J. (2003). Sicker and poorer—The consequences of being uninsured: A review of the research on the relationship between health insurance, medical care use, health, work, and income. *Medical Care Research and Review*, 60(2_suppl), 3S-75S.

Hallsworth, M., List, J., Metcalfe, R., and Vlaev, I. (2014). The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. NBER Working Papers 20007, National Bureau of Economic Research, Inc.

Himoniti-Terroviti, S. (2005). *Recent developments in the Greek housing market.* Athens: Centre of Planning and Economic Research Report No. 43.

Holland, P. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, 81:945–960.

Holt, J. (2009). A Summary of the Primary Causes of the Housing Bubble and the Resulting Credit Crisis: A Non-Technical Paper. *The Journal of Business Inquiry*, 120-129.

Husted, B. W. (1999). Wealth, Culture, and Corruption. *Journal of International Business*, 339-59.

Jagannathan, V. (1986). Corruption, delivery systems, and property rights. *World Development*, 127–132.

Iacus, S., King, G., and Porro, G. (2011). Multivariate matching methods that are monotonic imbalance bounding. *Journal of the American Statistical Association*, 106(493):345–361.

Imai, K. (2005). Do get-out-the-vote calls reduce turnout? the importance of statistical methods for field experiments. Natural field experiments, The Field Experiments Website.

Imbens, G. and Kalyanaraman, K. (2009). Optimal bandwidth choice for the regression discontinuity estimator. NBER Working Papers 14726, National Bureau of Economic Research, Inc.

IMF (2013). Greece; selected issues. IMF Staff Country Reports 13/155, International Monetary Fund.

Iyer, G. S., Reckers, P. M., and Sanders, D. L. (2010). Increasing tax compliance in washington state: A field experiment. *National Tax Journal*, 63(1):7–32.

Jerez Darias, L. M., Martin-Martin, V. O., and Perez Gonzalez, R. (2012). Approximacion a una geografia de la corrupcion urbanistica en Espana. *Eria*, 5-18.

Kaestner, B. (2012). Mortality and Access to Care after Medicaid Expansions. *New England Journal of Medicine*, *367*(25), 2453-2454.

Kaufman, D., Kraay, A., and Mastruzzi, M. (2010). The Worldwide Governance Indicators: Methodology and Analytical Issues. *World Bank Policy Research Working Paper 5430*.

Kaul, A., Kloessner, S., Pfeifer, G., & Schieler, M. (2015). Synthetic control methods: Never use all pre-intervention outcomes as economic predictors. *Unpublished. URI: http://www.oekonometrie.uni-saarland.de/papers/SCM_Predictors.pdf*

Kleven, H. J., Knudsen, M. B., Kreiner, C., Pedersen, S., and Saez, E. (2011). Unwilling or unable to cheat? evidence from a tax audit experiment in denmark. *Econometrica*, 79(3):651–692.

Kleven, H. J. and Waseem, M. (2013). Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from pakistan. *The Quarterly Journal of Economics*, 128(2):669–723.

Lambsdorff, J. (2006). Causes and Consequences of Corruption: What Do we Know from a Cross-Section of Countries? In Rose-Ackerman, *International Handbook on the Economics of Corruption* (pp. 3-51). Edward Elgar.

Lederman, D., Loyaza, N., and Soares, R. (2005). Accountability and Corruption: Political Institutions Matter. *Economic and Politics*, 1-35.

Lee, M. (1994). THE COST OF THE SECOND BEST: THE PRICE OF LAND REGULATION IN ASIA . *DPU Working Paper No. 62*.

Lee, D. S. and Lemieuxa, T. (2010). Regression discontinuity designs in economics. *Journal of economic literature*, 48(2):281–355.

Litina, A. and Palivos, T. (2016). Corruption, tax evasion and social values. *Journal of Economic Behavior & Organization*, 124:164–177.

Long, J. S. (1997). *Regression Models for Categorical and Limited Dependent Variables.* Thousand Oaks, CA: Sage Publications.

Ludwig, J. and Miller, D. L. (2007). Does head start improve children's life chances? evidence from a regression discontinuity design. *The Quarterly Journal of Economics*, 122(1):159–208.

Luttmer, E. and Singhal, M. (2014). Tax morale. IZA Discussion Papers 8448, Institute for the Study of Labor (IZA).

Lykens, K. A., & Jargowsky, P. A. (2002). Medicaid matters: children's health and Medicaid eligibility expansions. *Journal of Policy Analysis and Management*, *21*(2), 219-238.

Maldonado, S. (2010). Resource Windfall and Corruption: Evidence from a Natural Experiment in Peru. *Working Paper*.

Martín-Martín, V. O. (2010). Georeferenciando la corrupcion urbanistica en Espana. In J. O. al., *La informacion geografica al servicio de los ciudadanos: de lo global a lo local* (pp. 917-929). Seville: University of Seville Press.

Matsaganis, M. and Flevotomou, M. (2010). Distributional implications of tax evasion in Greece. Greese - hellenic observatory papers on greece and southeast europe, Hellenic Observatory, LSE.

Mauro, P. (1997). The Effects of Corruption on Growth, Investment, and Government Expenditure: A Cross-Country Analysis. In K. Elliot, *Corruption and the Global Economy* (pp. 83-107). Washington, DC: Institute of International Economics.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics 142*, page 714.

Mitrakos, T., Akantziliotou, C., Vlachostergiou, V., and Tsolacos, S. (2014). Exploring the determinants of residential property values in a crisis: evidence from Greece. *European Real Estate Society 21st Annual Conference.* Bucharest, Romania: Bucharest University of Economic Studies.

Mroz, T. A. (2012). A Simple, Flexible Estimator for Count and Other Ordered Discrete Data. *Journal of Applied Econometrics*, 646-665.

Nyblade, B., and Reed, S. (2008). Who Cheats? Who Loots? Political Competition and Corruption in Japan, 1947-1993. *American Journal of Political Science*, 926-941.

OECD (2013). Tax and development: What drives tax morale? Oecd working paper series, Organization for Economic Cooperation and Development.

Olken, B. (2007). Monitoring Corruption: Evidence from a Field Experiment in Indonesia. *Journal of Political Economy*, 200-249.

Ortega, D. and Sanguinetti, P. (2013). Deterrence and reciprocity effects on tax compliance: Experimental evidence from Venezuela. Caf working paper no. 8, CAF.

Ortega, D. and Scartascini, C. (2015). Don’t blame the messenger: A field experiment on delivery methods for increasing tax compliance. Working Paper Series 627, Inter-American Development Bank.

Paramonova, Y. A. (2016). Tax debt collection enforcement: When does suspension of a driver’s license help? Department of economics working paper, Higher School of Economics in St.Petersburg.

Perez-Truglia, R. and Troiano, U. (2015). Shaming tax delinquents: Theory and evidence from a field experiment in the united states. NBER Working Papers 21264, National Bureau of Economic Research, Inc.

Pomeranz, D. (2013). No taxation without information: Deterrence and self-enforcement in the value added tax. NBER Working Papers 19199, National Bureau of Economic Research, Inc.

Povoledo, E. (2011). Italy tries to get tax cheats to pay up. Technical report.

Roodman, D. (2006). *How to Do xtabond2: An Introduction to "Difference" and "System" GMM in Stata* . Washington, DC: Center for Global Development.

Rose-Ackerman, S. (1975). The economics of corruption. *Journal of Public Economics*, 187-203.

Rose-Ackerman, S. (1988). Bribery. In J. Eatwell, M. Milgate, and P. Newman, *New Palgrave Dictionary of Economics.* London: MacMillan.

Rosenbaum, P. (2007). Interference Between Units in Randomized Experiments. *Journal of the American Statistical Association*, 191-200.

Samuelson, P. A. (1954). The pure theory of public expenditures. *Review of Economics and Statistics*, 36(4):387–389.

Sanchez, G. (2014). The impact of low-cost intervention on tax compliance: Regression discontinuity evidence.

Schneider, F., Buehn, A., and Montenegro, C. (2010). New estimates for the shadow economies all over the world. *International Economic Journal*, 24(4):443–461.

Serra, D. (2006). Empirical Determinants of Corruption: A Sensitivity Analysis. *Public Choice*, 225-256.

Slemrod, J. (2015). Tax compliance and enforcement: New research and its policy implications. Working paper, University of Michigan.

Slemrod, J., Blumenthal, M., and Christian, C. W. (2001). Taxpayer response to an increased probability of audit: Evidence from a controlled experiment in minnesota. *Journal of Public Economics*, 79(3):455–483.

Slemrod, J. and Weber, C. (2012). Evidence of the invisible: toward a credibility revolution in the empirical analysis of tax evasion and the informal economy. *International Tax and Public Finance*, 19(1):25–53.

Snow, A. and Warren, R. (2007). Audit uncertainty, bayesian updating, and tax evasion. *Public Finance Review*, 35(5):555–571.

Sommers, B. D., Baicker, K., & Epstein, A. M. (2012). Mortality and access to care among adults after state Medicaid expansions. *New England Journal of Medicine*, *367*(25), 1025-1034.

Sommers, B. D. (2017). State Medicaid Expansions and Mortality, Revisited: A Cost-Benefit Analysis. *American Journal of Health Economics.*

Sowell, T. (2009). *The Housing Boom and Bust.* New York: Basic Books.

Stollwerk, A. E. (2007). Does e-mail boost turnout? *Quarterly Journal of Political Science*, 2:369–379.

Sutton, G. (2002). Explaining Changes in Housing Prices. *BIS Quarterly Review*, 46-55.

Svensson, J. (2005). Eight Questions About Corruption. *Journal of Economic Perspectives*, 19-42.

Tanzi, V. (1998). Corruption Around the World: Causes, Consequences, Scope, and Cures. *IMF Staff Papers*, 559-594.

Thaler, R. H. and Sunstein, C. R. (2012). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press, 7th edition.

Torgler, B. (2004). Moral suasion: An alternative tax policy strategy? evidence from a controlled field experiment in switzerland. Crema working paper series, Center for Research in Economics, Management and the Arts (CREMA).

Torgler, B. (2012). A field experiment on moral suasion and tax compliance focusing on underdeclaration and over-deduction. School of Economics and Finance Discussion Papers and Working Papers Series 285, School of Economics and Finance, Queensland University of Technology.

Treisman, D. (2000). The Causes of Corruption: A Cross-National Study. *Journal of Public Economics,* 399-457.

Tsatsaronis, K., and Zhu, H. (2004). What Drives Housing Price Dynamics: Cross-Country Evidence. *BIS Quarterly Review*, 65-78.

Wenzel, M. (2005). Misperceptions of social norms about tax compliance: From theory to intervention. *Journal of Economic Psychology*, 26(6):862–883.

Wenzel, M. and Taylor, N. (2004). An experimental evaluation of tax-reporting schedules: a case of evidence-based tax administration. *Journal of Public Economics*, 88(12):2785–2799.

Windmeijer, F. (2000). *A finite sample correction for the variance of linear two-step GMM estimators.* London: IFS Working Papers, No. W00/19.

VITA

Antonios M. Koumpias is a Ph.D. candidate in Economics and a research associate at the International Center for Public Policy in the Andrew Young School of Policy Studies of Georgia State University. His primary research interests lie in the fields of Public Finance and Health Economics. Specifically, he is fascinated by the challenges to revenue administration posed by tax evasion due to low voluntary compliance or informality. His job market paper estimates the effect of compliance reminders on payments of overdue income taxes in Greece using a regression-discontinuity design. His second dissertation essay examines the relationship between housing prices and zoning corruption in Greece and Spain with Jorge L. Martinez-Vazquez of Georgia State University and Eduardo Sanz-Arcega of the Institute of Fiscal Studies in Spain. In his third dissertation essay, he examines the impact of increased eligibility for publicly-provided health insurance (Medicaid) to adults in U.S. states in the 1990s and early 2000s on state-level adult mortality with Charles J. Courtemanche of Georgia State University and Daniela Zapata of IMPAQ International employing a synthetic control design. In other works, he documents the extent of Value-Added tax evasion due product misclassification in Greece with Nikolaos Artavanis of the University of Massachusetts, Amherst. Also, he investigates the effect of income tax adoption by school districts in Ohio on total local tax revenue growth and variability with Joshua C. Hall of West Virginia University. In addition, he studies the effect of mass media campaigns on tax morale in Pakistan with Jorge L. Martinez-Vazquez and Musharraf R. Cyan of Georgia State University. His co-authored work on the determinants of tax morale in Pakistan with Jorge L. Martinez-Vazquez and Musharraf R. Cyan was published in 2016 by the Journal of Asian Economics. Before his studies at the Andrew Young School of Policy Studies of Georgia State University, Antonios obtained his B.Sc. in Economics from the University of Macedonia in Thessaloniki, Greece, fulfilled his compulsory military duty as a Greek national and earned his M.A. in Economics from Duke University. Outside of academia, Antonios enjoys outdoor sports, in particular basketball and soccer.