Responses to Poverty: Essays in Public Economics and Labor

Andrew Johnston

University of Pennsylvania, acjohnston3@gmail.com

Follow this and additional works at: http://repository.upenn.edu/edissertations

Part of the Labor Economics Commons, and the Public Policy Commons

Recommended Citation


http://repository.upenn.edu/edissertations/1791

This paper is posted at ScholarlyCommons. http://repository.upenn.edu/edissertations/1791

For more information, please contact libraryrepository@pobox.upenn.edu.
Responses to Poverty: Essays in Public Economics and Labor

Abstract
Throughout my work, I seek to understand policy issues that have consequences for poverty. Data limitations and non-random assignment of policy treatment make meaningful analysis challenging in what I argue are important areas of policy; to overcome the empirical challenge I combine clean natural experiments with rich administrative data to make progress in each literature. In this collection of essays I look at policy factors for three important determinants of economic wellbeing: labor supply, human capital formation, and labor demand. (1) Regarding labor demand, my coauthor and I study the impact of unemployment insurance benefit extensions on the employment of displaced workers. Using a sharp policy change and rich administrative data I procured, we are able to make considerable progress. (2) Teacher quality is the most powerful school-input in human capital formation. To understand how teacher pay affects the quality of teachers, I leverage a federal policy that provides additional compensation to teachers serving sufficiently poor schools. Using a regression discontinuity design and rich education data, I am able to provide further light on this pressing policy issue. (3) Finally, I assess the consequences of a payroll tax that firms pay on the labor demand. Because the tax increases after recessions, workers may bear the tax when the labor market is already weak. I use a discontinuity in the tax schedule and administrative UI data to estimate the consequences of the tax as a deterrent and the effect of the tax once raised.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Economics

First Advisor
Robert Jensen

Subject Categories
Economics | Labor Economics | Public Policy

This dissertation is available at ScholarlyCommons: http://repository.upenn.edu/edissertations/1791
RESPONSES TO POVERTY: ESSAYS IN PUBLIC ECONOMICS AND LABOR

Andrew Cowley Johnston

A DISSERTATION

in

Applied Economics

For the Graduate Group in Managerial Science and Applied Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2016

Supervisor of Dissertation

__________________________________

Robert Jensen, David B. Ford Professor of Business Economics and Public Policy

Graduate Group Chairperson

__________________________________

Eric Bradlow, K.P. Chao Professor; Professor of Marketing, Statistics, and Education

Dissertation Committee

Robert Jensen, Professor of Business Economics and Public Policy

Mark Duggan, Professor of Economics

Alexander Mas, Professor of Economics

Olivia S. Mitchell, Professor of Business Economics and Public Policy

Todd Gormley, Assistant Professor of Finance
I am grateful to my adviser, Mark Duggan, whose broad vision and insight helped me enormously as a scholar to approach the questions that matter. I have nothing but thanks for Robert Jensen who is not only a terrific scholar, but a kind and gentle soul. My interactions with Alexandre Mas were perhaps the most influential and transformative in my graduate career; I owe him a great deal. Olivia S. Mitchell consistently guided me along, helping me to see the larger picture while aiding in every detail. Todd Gormley’s empirical insight improved my work tremendously. Their investment in me cannot be repaid, but it is appreciated, and remembered. Also owed are my fellow students; to Boris Vabson I owe the fortitude to withstand common nonsense for a better path; to Andy Wu I owe a debt for constant and unselfish friendship, never wavering; to Ana Gazmuri I owe a debt for the gift of honesty, which is seldom imparted among friends. And I owe my dear wife—the greatest person I know, the best thing in my life, and my favorite part of the future.
ABSTRACT

RESPONSES TO POVERTY: ESSAYS IN PUBLIC ECONOMICS AND LABOR

Andrew Cowley Johnston

Robert T. Jensen

Throughout my work, I seek to understand policy issues that have consequences for poverty. Data limitations and non-random assignment of policy treatment make meaningful analysis challenging in what I argue are important areas of policy; to overcome the empirical challenge I combine clean natural experiments with rich administrative data to make progress in each literature. In this collection of essays I look at policy factors for three important determinants of economic wellbeing: labor supply, human capital formation, and labor demand. (1) Regarding labor demand, my coauthor and I study the impact of unemployment insurance benefit extensions on the employment of displaced workers. Using a sharp policy change and rich administrative data I procured, we are able to make considerable progress. (2) Teacher quality is the most powerful school-input in human capital formation. To understand how teacher pay affects the quality of teachers, I leverage a federal policy that provides additional compensation to teachers serving sufficiently poor schools. Using a regression discontinuity design and rich education data, I am able to provide further light on this pressing policy issue. (3) Finally, I assess the consequences of a payroll tax that firms pay on the labor demand. Because the tax increases after recessions, workers may bear the tax when the labor market is already weak. I use a discontinuity in the tax schedule and administrative UI data to estimate the consequences of the tax as a deterrent and the effect of the tax once raised.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENT</td>
<td>III</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>IV</td>
</tr>
<tr>
<td>CHAPTER 1</td>
<td>1</td>
</tr>
<tr>
<td>CHAPTER 2</td>
<td>31</td>
</tr>
<tr>
<td>CHAPTER 3</td>
<td>60</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>97</td>
</tr>
</tbody>
</table>
Introduction

The debate over unemployment insurance duration reflects the tradeoff between consumption smoothing, on the one hand, and moral hazard on the other. In the recent recession, unemployment benefits extended to 99 weeks in some states, and during the recovery several states cut the duration of their regular benefits programs, leading policy makers to face important questions: How do recipients respond the permitted duration of UI benefits, and how do these responses affect the labor market as a whole? These questions are particularly relevant for understanding the performance of the US labor market in the Great Recession and its aftermath where evidence is thin. An additional consideration for evaluating how UI extensions affect the labor market is that the aggregate effects of these policies may differ from those implied by the micro response if there are general equilibrium effects or spillovers, as would be the case if UI recipients crowded out other jobseekers. With a few notable exceptions (Levine (1993), Lalive et al. (2013) and Valleta (2014)) we know relatively little about the relationship between the micro and macro responses to UI extensions.

---

1 Studies that have found this relationship include Card and Levine (2000), Katz and Meyer (1990), Moffitt (1985) and Solon (1979) in the United States and Card, Chetty, Weber (2007), Lalive (2008), Schmieder, Von Wachter, and Bender (2012), and Schmieder, Von Wachter, and Bender (2014) in Western Europe.

2 Levine (1993) estimates the relationship between state and year variation in UI replacement rates and unemployment durations for uninsured workers. Using data from the CPS and NLYS for 1979-1987 he finds evidence of displacement. Valleta (2014) uses linked CPS data to examine the relationship between potential benefit duration by state and exit to unemployment for workers who are likely UI ineligible. On average he finds no relationship, but he finds that ineligible workers in higher unemployment states have
In this paper we study the micro and macro effects of a large benefit duration cut that occurred in 2011 in Missouri using newly available administrative data and regression discontinuity and difference-in-difference designs. Following the 2007 recession, eight US states reduced regular UI durations, partly in response to diminished reserves in state UI trust funds as well as the changing political environment. While there is a precedent for cutting UI benefit levels, to our knowledge this was the first time states cut maximum UI benefit durations. Since then, eight states (Arkansas, Florida, Georgia, Kansas, North Carolina, Missouri, Michigan, and South Carolina) have cut the duration of UI benefits to below 26 weeks of maximum benefits, the standard level in place for over half a century.\(^3\)

We examine the effect of potential UI benefit duration on the duration of UI receipt, reemployment, wages, and the unemployment rate by examining the cut in UI benefit weeks implemented in Missouri in April 2011. This reduction, which occurred while Emergency Unemployment Compensation (EUC) was in effect, resulted in dislocated workers receiving up to 16 fewer weeks of UI eligibility than they would have had received if they had applied previously.\(^4\) The policy change was sudden and unanticipated; only five days passed between when the legislation was first proposed

\(^3\) In 2010 all states had a maximum duration of benefit eligibility of at least 26 weeks.

\(^4\) Specifically, the cuts were 16 weeks for UI recipients previously eligible for 26 weeks of regular state UI and eligible to participate in the EUC program.
and when the law applied to UI claimants. The timing was such that there was almost no opportunity for claimants to shift the timing of their claims.

We use rich unemployment insurance administrative data and wage records from Missouri and a regression discontinuity design (RDD) to estimate the effects of this policy, where the running variable is calendar time and the threshold of interest is the exact week the law was enacted. The novel administrative data we obtained not only allow us to see UI receipt but also re-entry in employment and wages which has not been possible in the vast majority of domestic papers investigating UI, none of which uses data from the recent decade (see Chetty (2010); Landais (2015); Krueger and Meyer (2002)). By observing employment we can rule out alternative explanations of UI exit including migration and labor-force exit, and demonstrate persuasively that early exit from UI flows into gainful employment.

Our findings indicate economically and statistically significant rates of exit from UI for claimants subject to the shorter benefit duration relative to claimants with the longer duration at the cutoff, resulting in an estimated sensitivity of unemployment duration to potential UI duration that is at the upper end of the literature. As found in Card, Chetty, and Weber (2007) in the case of Austria, and Schmieder, Von Wachter, and Bender (2012) in the case of Germany, we find evidence that some UI recipients are forward-looking. For example, UI recipients subject to the benefit cuts had 57 weeks of eligibility, but were 12 percentage points less likely to be receiving UI by week 20 of their spell, from a base of 46 percent. We estimate that a one-month reduction of UI duration reduces the duration of UI receipt by 15 days, on average, and that approximately 54

---

5 The legislation was a compromise aimed at breaking a Republican filibuster in the Missouri State Senate.
6 More precisely, this is an interrupted time-series design, but we use RDD methods and for convenience refer to the design as a RDD throughout.
percent of this change is through changes in exit rates occurring prior to benefit exhaustion.

Analysis of wage records for the universe of legally employed Missouri workers indicate that those exiting early from UI enter employment. The estimates imply that a one-month cut in potential duration resulted in a reduction of nonemployment duration of approximately 10 days. The findings suggest that the benefit cut increased job search intensity. However, we find limited effects of shorter benefit durations on the UI exit hazard rate after 20 weeks of UI, and for the long-term unemployed we find no evidence that lower potential duration leads to higher employment rates after exhaustion. As in Card, Chetty, and Weber (2007), we find no significant differences in the average quarterly earnings for the first job of recipients, conditional on employment relative to the comparison group, suggesting that those induced to exit unemployment earlier are not penalized with lower wages.

The effects of extended UI on other job seekers is theoretically ambiguous. If there is job rationing, which can arise in search models with diminishing returns to labor and wage stickiness (Michaillat 2012), increased search effort leads to negative externalities on other workers. However, there are no externalities in models with constant marginal returns to labor and perfectly elastic labor demand (Landais et al. 2010; Hall 2005). In models of Nash bargaining (such as Pissarides 2000) the macro elasticity of UI benefits is larger than the micro elasticity as a result of the “wage externality”. To assess spillovers, we calculate the predicted path of the unemployment rate from the benefit, using the survivor function from the RDD and the flow of initial UI claims. In the simulation we assume that jobseekers not affected by the UI cut are not displaced from employment by UI recipients who were exposed to the policy, or other spillovers. We compare this predicted path to the actual path of the unemployment rate
from a difference-in-difference (DiD) estimate of the cut. We find that the simulated and estimated paths of the macro effect closely match. The predicted and estimated paths are approximately the same in levels, and follow a similar U-shaped pattern peaking at approximately a 1 percentage point drop in the state unemployment rate. We find no evidence that the cut led to changes in the size of the labor force. The analysis suggests the labor market absorbed the jobseekers without displacement, even though the unemployment rate was high at the time of the cut at 8.6 percent. The findings are more consistent with a labor market characterized by a flat labor demand curve in the framework of Landais et al. (2010).

Our study also speaks to the question of the labor market effects of UI extensions during the Great Recession. During this period, UI benefits increased from the near-universal length of 26 weeks to up 99 weeks in some states. Subsequently, declining unemployment led to reductions in extended benefits, and benefit duration largely returned to pre-recession levels following the expiration of the EUC program in 2013. The labor market effects from these changes in benefit duration are a central question for labor market policy and have been the focus of a number of studies. Notably, recent papers studying this period in the United States have used state level variation in benefit lengths to estimate the effects of increases (Rothstein 2011; Farber and Valleta 2015) and declines (Farber, Rothstein, and Valleta 2015) in UI potential duration in the US over the 2007 recession period and its aftermath. These researchers found fairly small effects of extended benefit lengths on unemployment. Hagedorn et al. (2013) and Hagedorn et al. (2015) find small effects of changes in potential duration on jobseekers, but large macro effects on wages, job vacancies, labor force participation and employment. Ours is the first study to use a design-based approach with administrative micro data covering UI receipt, employment, and wages to study the labor
market effects of changes in maximum duration in this period. While we find no evidence
of moral hazard for the long-term unemployed, we identify a fairly large response to the
benefit cut for a subset of participants early in the spell.

II. Institutional Background

In the United States, UI is administered by state governments but is overseen
and regulated by the federal government. Before 2011, eligible laid-off workers received
up to 26 weeks of regular unemployment insurance benefits if they were not reemployed
before their benefits are exhausted. During periods of unusually high unemployment,
state and federal governments have extended potential benefit duration, to support the
long-term unemployed after regular benefits are exhausted. Two programs provide these
extended benefits: the Extended Benefit (EB) and the EUC programs.

EB is a permanent program that provides extended benefits in states with high
unemployment to unemployed workers who exhaust their regular state benefits. Until
recently, the federal government split the cost of EB with state governments. Through
the Recovery Act passed in February 2009, Congress temporarily suspended cost
sharing and the federal government bore all the cost of EB through December 2013. EB
extended benefits are triggered as a function of a state’s total and insured
unemployment rate, and triggering thresholds vary by state. In Missouri the 13-week EB extension can be triggered in two ways. First, EB is triggered if the unemployment rate among insured workers (IUR) is at least 5 percent over the previous 13 weeks and the IUR is 120 percent of the IUR for the same 13-week period in the previous two years. Second, EB can be triggered if the IUR for the previous 13 weeks is at least 6 percent, regardless of the IUR in previous years. If IUR crosses either of these thresholds, the state automatically enrolls unemployed workers in 13 additional weeks of benefits if they exhaust their regular benefits.
government took on all of the costs of EB, Missouri temporarily enacted legislation to implement an additional trigger that would increase EB duration from 13 to 20 weeks.\(^8\)

EUC has been enacted periodically through federal legislation when unemployment is high. During the Great Recession, the EUC program was active from June 2008 through December 2013. In its most recent version, federal benefits provided longer extensions for states with higher rates of insured unemployment.\(^9\)

The benefit cut in Missouri was the byproduct of a Republican filibuster, led by four lawmakers in the Missouri State Senate; they objected to legislation that would accept federal money to extend UI benefits under the EB program. The bill would have allowed for the continuation of 20 additional weeks of benefits to unemployed workers who exhausted their EUC and regular benefits at no cost to Missouri.\(^10\) The extension had already passed the Missouri State House by a margin of 123 to 14. The first news reports of the filibuster were on March 4, 2011 (Wing 2011). On April 6 a report indicated that the lawmakers had agreed to end their filibuster, though the article did not specify terms (Associated Press 2011). On April 8 the *St. Louis Post Dispatch* published the first article detailing the possible compromise. Under the compromise, regular benefits would be cut from 26 to 20 weeks in exchange for Missouri accepting federal dollars and maintaining EB benefits for the long-term unemployed (Young 2011). In effect, the agreement traded-off longer UI durations in the short run (for the long-term unemployed) in exchange for shorter UI durations in the long run. We found no press reports prior to

---

\(^8\) If the total unemployment rate (TUR) was at least 8 percent and 110 percent of the TUR for the same 3-month period in either of the two previous years, the duration of EB would increase from 13 to 20 weeks (http://www.cbpp.org/cms/index.cfm?fa=view&id=1466).

\(^9\) If a state had less than 6% unemployment, EUC provided 14 additional weeks after regular benefits were exhausted; 28 weeks if less than 7% (but greater than 6%); 37 weeks if less than 9 (but greater than 7%); and 47 weeks if greater than 9%.

\(^10\) The lawmakers leading the filibuster argued that accepting these funds would increase the federal deficit unnecessarily.
April 8 regarding the possibility of cutting the duration of regular benefits as a possible compromise for the filibuster. This legislation appears to have been unanticipated. On April 13 the Missouri House of Representatives passed the bill which Jay Nixon, the Democrat governor, signed into law on the same day (Selway 2011). All new claims submitted after that date were subject to the abbreviated benefits (Mannies 2011). Federal regulations calculate EUC weeks eligible in proportion to regular state UI benefits. Thus, the cut in regular state UI benefits triggered an additional 10-week reduction in EUC, and the maximum UI duration fell from 73 weeks for claimants approved by April 13, to 57 weeks for claimants approved afterwards resulting in a total change in potential duration of 16 weeks. EB was a non-factor for new claimants at this time (with or without the benefit cut) because EB phased out by the time they were eligible.

The change in potential UI duration was the only change in Missouri’s UI system in the legislation. We corresponded with Missouri UI program administrators who told us that there were no changes in the administration of the program, including search requirements or communications with UI recipients. For example, they did not send additional notices informing UI recipients affected by the policy change. For convenience, we label recipients applying for UI after the law the “treatment group” and recipients applying before the policy change the “control group.”

III. Data

Our analysis utilizes administrative data from the state of Missouri covering workers, firms, and UI recipients from 2003 to 2013. We use three data files for the analysis. The first is a worker-wage file detailing quarterly earnings for each worker with
unique (but de-identified) employee and employer IDs. The second is an unemployment claims file that contains the same worker and employer IDs as the wage file. For each claim, we observe the date the claim was filed, the weekly benefit amount, the maximum benefit amount over the entire claim, the dates weekly benefits were issued, the wage history used to calculate benefits and duration, and the benefit regime (i.e., regular benefits, EB, or EUC). For every claim, we link the records for regular benefits, EB, and EUC claims to construct a single continuous history associated with each claim. The third dataset reports a limited set of employer characteristics including detailed industry categories. The raw data contains 1,635,993 initial UI claims over 2003-2013 and 184,191 in 2011. We remove claims ineligible for UI, including unemployed workers who were fired for cause or quit voluntarily, observations with missing claim types (regular, EB, or EUC) or missing base-period earnings, and EB or EUC claims that could not be traced to an initial regular claim. To aid in interpreting the effects, we also limit the sample to those workers who, based on their earnings histories, would have been eligible for the full 26 weeks of regular UI benefits without the policy change. Specifically, the formula for maximum potential duration of regular benefits is:

\[
\text{Regular Potential Duration} = \min\left(X, \left(\frac{E}{3}\right)\left(\frac{1}{B}\right)\right)
\]

where \(E\) is a measure of total base period earnings, \(B\) is the average weekly benefit, and \(X\) is 26 weeks on or before April 13, 2011 as well as 20 weeks after this date. Because we want to focus on workers who are affected by the cut in maximum duration we select recipients for whom \(\frac{E}{3B} \geq 26\). This procedure does not induce any mechanical change in the characteristics of workers across the policy change threshold. These “full eligibility” claimants represent 72 percent of all claimants in 2011 and 67 percent of all
claimants for the entire 2003-2013 period. After these screens we have 1,064,652 claims over the 2003-2013 period and 127,710 claims in 2011.

Descriptive statistics for the administrative data appear in Table 1. Column (1) reports summary statistics for the full 2003-2011 period and column (2) for 2011. The average weekly benefit in 2011 in the sample was $260. UI recipients eligible for the maximum benefit duration had an average of 14.5 quarters of tenure in their previous employer and their earnings in the last complete quarter of employment prior to collecting UI benefits was $8,259. Earnings in the first complete quarter of employment after the UI spell average $7,240. On average, recipients claiming benefits in 2011 received 29.3 weeks of unemployment benefits.

For the aggregate analysis we use data from the Local Area Unemployment Statistics (LAUS) program of the Bureau of Labor Statistics. For outcomes we use the state-by- calendar month unemployment rate, the natural log of number of unemployed, and the natural log of the size of the labor force. We deseasonalize these variables by regressing each outcome on state × month dummies over the 2001-2005 period and then deviating each outcome in 2005-2013 from the predicted value of this regression. We also use these variables derived from the Current Population Survey (CPS) to assess robustness.

IV. Empirical Design

To identify the causal effect of longer UI duration, we utilize the discrete change in the maximum UI duration resulting from a rapid and unexpected policy change: claimants who applied just before April 13, 2011 were eligible for 73 weeks of benefits and those who applied after were eligible for 57 weeks. We use this discontinuity to compare similar displaced workers entering the same labor market who experienced very different UI benefit durations.
The natural experiment’s contribution is significant, in part, because it implicitly controls for labor-market conditions that may be affected by the reform. In cross-state settings, treatment not only affects the search effort of a worker, but potentially the health of the labor market through simulative effects. More fundamentally, state-panel analyses may suffer because treatment affects significantly different populations, facing already dissimilar job-market dynamics, even if the policy treatment does not stimulate employment. Our data and setting allow us to compare extremely similar claimants, searching in the same labor market who only face different benefit durations, thus accounting for possible effects on employment that are channeled through the labor market rather than through changes in search. We model the outcome variable $Y_i$ as a continuous function of the running variable, the claim week, and estimate the outcome discontinuity that occurs at the threshold, the date of the policy change:

\[ Y_i = \beta T_i + f(x_i - x') + u_i, \]

where $x_i$ is the calendar week of the UI claim for person $i$, $x'$ is the week of the policy change, and $T_i$ equals one if worker $i$ applied after the policy change and zero if she applied before.\(^{11}\) Thus, $f(x_i - x')$ is a continuous function of the running variable which captures the continuous relationship between the application date and the outcome of interest. Because we control flexibly for the running variable, the model accommodates smooth seasonal changes, allowing for unbiased estimation of the discrete policy change. In practice we first collapse the data to the claim week level and weight the observations by the number of claims in the week, a process that yields identical point estimates to the micro data. As shown by Lee and Card (2008), heteroskedasticity-consistent inference with collapsed data is asymptotically equivalent to clustering on the

\(^{11}\) We use the claim week because the data can be sparse when using the claim application calendar date, and there are days with no claims, such as administrative holidays and weekends.
running variable. We estimate the model using local linear regression (Hahn, Todd and Van der Klaauw, 2001) with the Imbens and Kalyanaraman (2012) (IK) optimal bandwidth and a triangular kernel. We consider a range of alternative bandwidths to assess robustness, as well as estimation of a local quadratic using the Calonico, Cattaneo, and Titiunik (2014) (CCT) optimal bandwidth.

V. Results

Diagnostics

We begin by testing for manipulation of the running variable, which would occur if claimants could strategically time their applications around the policy change. Figure 1 plots the frequency distribution of the number of UI claims by week, over the 2009–2012 period. The solid vertical line denotes the time of the policy change, and the dashed vertical lines denote the same date in the previous years. It is evident in Figure 1 that there is a great deal of seasonality in claims, with a large spike in claims around the new year. The policy change occurred after the large seasonal increase, in April, and by this time claims were at moderate levels. There is no visual evidence of an abnormal spike in claims before the policy change, as would be the case if claimants could time their applications apply for longer-lasting UI benefits. Column 1 of Table 2 formally tests for a discontinuity in claims (McCrary 2008). Estimating a local quadratic model to fit the curvature in the distribution, we find no significant discontinuity in the relative frequency of claims.12

Inspection of the frequency distribution does reveal a moderate jump in claims two weeks after the change in policy. As will be seen, this applicant cohort looks different in a number of dimensions than recipients who applied before or after this group, and in particular they appear to have characteristics correlated with being lower duration

12 Appendix Figure 1 displays the fitted quadratic in the frequency distribution.
claimants. This outlier might be random noise, or it might reflect a failed attempt to time claims to obtain UI before the cut. To keep the analysis as transparent as possible, we keep this group in the main sample. However, we have also estimated all models excluding this cohort. Estimates presented in the Online Appendix show precise but somewhat smaller estimates on UI receipt and nonemployment when this cohort is excluded.

As a second examination of design validity, we test for discontinuities in predetermined covariates of UI applicants around the policy change. Because there are numerous predetermined variables from which we can select, we construct an index of predicted log initial UI duration using all covariates available in the data set following the same procedure as Card et al. (2015). To construct the index, we regress log UI duration on a fourth-order polynomial of earnings in the quarter preceding job loss, indicators for four-digit industry, and previous job tenure quintiles. Figure 2 plots the mean values of the covariate index over 2009–2012 by claim week. The continuity in the index around the threshold is borne out visually, and the RDD estimate of this predicted value at the cutoff is small and statistically insignificant (column (2) of Table 2). The lack of evidence of sorting and differences in predetermined characteristics around the threshold reinforces the claim that the policy change was unanticipated and difficult or impossible to game.¹³

UI Receipt

Figure 3 exhibits the mean duration of realized UI spells by application week. There is a clear drop in the number of weeks claimed as a function of the claim week. Column (1) of Panel A in Table 3 shows that the benefit reduction of 16 weeks is

¹³ As previously discussed, in Figure 2 we see that the cohort receiving claims two weeks after the duration cut has substantially lower predicted durations. This pattern will be seen in all subsequent analyses.
associated with 8.7 fewer weeks of UI benefits claimed (s.e. = 1.4), on average.

Estimating the same model but setting the threshold for the same week one year prior to the cut shows an insignificant difference in the duration of UI receipt between treatment and control (Table 3/Panel B/column 1). Appendix Figure 4 shows the point-estimate estimated over a range of alternative bandwidths. The estimate is stable for a wide range of bandwidths, including bandwidths smaller than the IK bandwidth and up to twice as large as the IK bandwidth. Appendix Table 1 reports the estimate excluding the negative outlier cohort two weeks after the policy change. The estimate is somewhat smaller but remains highly significant. Appendix Table 2 reports the estimate using a local quadratic model with the CCT optimal bandwidth. The estimated effect is somewhat larger than the local linear case, and statistically significant.

To examine the timing of UI receipt, we estimate the probability that an individual remains on UI through each of the first 73 weeks of the spell. Figure 4 presents binned scatterplots of the probability that claimants remained on UI in weeks 20, 40, 55, and 60 weeks of UI benefits as a function of their initial claim week. The figure shows that there is a response to the cut in maximum duration fairly early in the spell. In weeks 20, 40, and 55, before the treatment group exhausted benefits, it can be seen visually that the duration cut is associated with a lower probability of receipt. By week 60, the probability of remaining in UI for the treated group falls to close to zero, consistent with all remaining claimants in the treatment group exhausting their benefits, while 26 percent of the comparison group was still receiving UI at that point. In none of these series do we see a similar break one year prior to the policy change (denoted by the dashed vertical line).

Table 3 columns (2)–(5) report the point estimates for the probability that the UI spell lasted until weeks 20, 40, 55, and 60. The RDD estimate for UI receipt is -12.3
percentage points in week 20, -11.8 percentage points in week 40, -10.1 percentage points in week 55, and -23.6 percentage points in week 60. All estimates are highly significant. These shifts are not seen in the corresponding placebo estimates in Panel B. Placebo estimates are indistinguishable from 0 in all cases except for the probability of receiving benefits in week 20, which is positive.\textsuperscript{14} As with unemployment duration, the estimates are somewhat smaller excluding the outlier two weeks after the policy change (Appendix Table 1), and somewhat larger when estimating a local quadratic with a CCT bandwidth (Appendix Table 2) but significant in both cases.

To estimate the timing of the effects over the whole period, we fit variants of equation (1) where, in each specification, $Y_i$ is the probability that the claimant received at least $T$ weeks of benefits, where $T$ spans 1 to 73. These estimates give the relative survival probabilities between the two groups by week. Figure 5 plots each of these RDD estimates with the associated confidence intervals. Figure 5 shows that the survival function diverges between the two groups starting early on in the UI spells, until around week 20 of the UI spell, and then levels out. This pattern implies that the hazard rate for UI exit is larger for the treatment group than for the control group in the first five months of the UI spell and then stabilizes.

Note that there is a sharp drop in the survivor rate for the treatment group in week 20 and a similar drop for the comparison group in week 26. These drops represent individuals who did not receive benefits beyond the regular state benefits, either because they were ineligible since the federal government automatically enrolls the eligible. Because of these drops in the survivor rate at regular benefit exhaustion date, we do not interpret the 20–26 week span because any differences over this term

\textsuperscript{14} Because Easter was on April 24, 2011, we also estimated a placebo specification setting the policy change just prior to Easter 2010. We found no significant effects for this placebo as well suggesting that our estimates are not being driven by this holiday.
reflect a combination of eligibility and behavioral effects. Nevertheless, this pattern
demonstrates the RDD estimates detect expected changes in UI receipt behavior,
confirming our intuition.

Excluding this 20–26 week period, the treatment-control differences in the survivor rate
is stable from week 20 of the UI spell through week 57, at which point there is a
significant drop in the relative survivor rates as the treatment group exhausts EUC
benefits while the control group continues to receive EUC benefits until week 73. The
error bands in Figure 5 show that the first significant difference between the two groups
occurs in week 14, and the differences remain significant for all subsequent weeks.
These estimates indicate claimants respond in a forward-looking way to UI exhaustion,
and most of the response to the duration cut occurs fairly early in the spell, within the
first three months. This time pattern of exit is robust to alternative bandwidths. Appendix
Figure 3 shows the same plots with double the IK bandwidth and the pattern persists.
One way to assess the consistency of the duration and survival RDD estimates is to note
that mean UI duration is the integral of the survival function. Using the discrete analog to
this relationship, summing the estimated survival probabilities through 73 weeks yields
an expected duration of 31.1 weeks in the control group and 21.5 weeks in the treatment
group. The 9.7 week difference in the expected duration implied by the survival
probabilities is close to the 8.7 week RDD estimate. The consistency of the estimates
further reinforces the validity of the design.

The reduction in weeks of UI receipt is a possible combination of “mechanical”
effect of earlier exhaustion for the treatment group and pre-exhaustion UI exit. We can
use the estimated survival probabilities to decompose the overall change in weeks of UI
receipt into two parts: the part due to changes in behavior prior to exhaustion and the
part due to pre-exhaustion exit. The estimated survival function implies that the expected
duration conditional on duration being less than 58 weeks is 27.6 weeks in the control and 21.2 weeks in the treatment. Because $E[\text{Duration}] = E[\text{Duration}|\text{Duration} < 58] \times \text{Pr}(\text{Duration}<58) + E[\text{Duration}|\text{Duration} \geq 58] \times \text{Pr}(\text{Duration} \geq 58)$, and $\text{Pr}(\text{Duration}<58) \approx 0.74$ in the control group, approximately 54 percent of the change in the overall duration of UI receipt comes from changes in the response to the cut before exhaustion.

**Employment**

Using the quarterly wage files we can measure the employment rate for the treatment and control groups following the policy change. Figure 6 plots the employment rate by UI application week for four quarters after the benefit cut. Consistent with the pattern seen for UI exits, in 2011 Q3—the first full quarter after the cut—there is a noticeable jump in the employment rate for applicants claiming after the duration cut. The elevated employment rate for the treated group can also be seen in 2011 Q4, 2012 Q1 and 2012 Q2.

Figure 7 presents the RDD estimates and associated 95 percent confidence intervals for employment rates by quarter, starting in the quarter the policy went into effect in the second quarter of 2011 through the second quarter of 2013. The RDD estimate for employment is insignificant in 2011 Q2, the quarter of the policy change. In 2011 Q3—the first complete quarter after the duration cut—the treated group has a 11.9 percentage point higher employment rate than the comparison group. The difference in employment rates is similar to the 10-12 percentage point difference in the probability of receipt in the early part of the UI spells over the relevant range, suggesting that those individuals who leave UI before exhaustion tend to enter employment. The employment effect fades out by 2012 Q4 at which point both treatment and control have exhausted
their benefits. The point estimates and standard errors for the employment RDD are presented in Table 4.\textsuperscript{15}

Conveniently, the 16-week period when the treated group had exhausted benefits and the control group was still eligible for benefits covers the entire third quarter of 2012 (as well as part of the second quarter of 2012). Therefore, to assess the effects of benefit exhaustion for the long-term unemployed in the treatment group, relative to the control who still received benefits, we can look at the change in the relative employment rate between the two groups in 2012 Q3 relative to earlier quarters. If exhausting benefits results in people scrambling and successfully finding employment, we would expect to see an increase in the RDD estimate for employment relative to the estimate in the previous quarter and the subsequent quarter. This is not what we find, rather, the relative employment rates in the treatment and control groups fell over the period. This suggests that, for the long-term unemployed who did not respond to the policy early, exhausting UI benefits did not hasten reemployment relative to the control. Instead, the positive employment effects we observe come from the group of UI recipients who responded to the changing weeks of eligibility well before exhaustion.\textsuperscript{16} A caveat to this conclusion is that at the time the treatment group exhausts the composition of the two groups differs since there were more exits from UI in the treated group among the “forward-looking” subset of participants. It is possible that an increase in the exit rate from this group in the control masks any positive effect of exhaustion on employment in the treatment group.

\textsuperscript{15}Appendix Table 3 reports local linear estimates excluding the outlier cohort two weeks after the policy change. Appendix Table 4 reports local quadratic estimates with the CCT optimal bandwidth. We continue to see significant employment effects in both cases.

\textsuperscript{16}Appendix Figure 4 shows the same charts using twice the IK bandwidth.
Figure 8 shows the “placebo” estimate for the employment effect of the benefit cut. Specifically, we estimate the same model with quarterly employment outcomes for quarters starting one year prior to the duration cut, setting the placebo duration cut to April 2010. There are no significant employment estimates over this period.

We can use the estimates corresponding to the relative nonemployment probabilities by quarter (shown in Figure 7) to calculate the expected difference in the duration of mean nonemployment between the two groups. If we assume that the relative employment probabilities between the two groups are the same after the third quarter of 2012, after which all recipients have exhausted their benefits, summing the estimates in Figure 7 from the quarter of the policy change through 2012 Q3 implies that a one-month reduction in unemployment duration reduces the number of days of nonemployment by an average of 10.4 days, with a 95 percent confidence interval of (6.7,14.1).\(^\text{17}\)

**Reemployment Earnings**

A class of job search models predict that longer provision of unemployment benefits allow workers to increase their reservation wage and find a more desirable job match. Longer UI duration could also depreciate human capital resulting in lower wages. The literature has mixed findings on the relationship between UI benefit duration and reemployment wages. Card, Chetty, and Weber (2007) found no significant effect of delay while Schmieder, Von Wachter, and Bender (2013) find that workers with longer potential UI spells have lower wages. We find that post-employment earnings do not change significantly following the cut in duration. Figure 9 shows mean log reemployment earnings for the first complete quarter after the individual has been

\(^\text{17}\) The confidence interval, which is constructed from the standard errors for each quarterly estimate, assumes no covariance term between the RDD estimates of employment by quarter.
reemployed, by application week.\textsuperscript{18} There is no evidence of a break at the threshold, a finding that is confirmed by the positive and insignificant estimate on the log reemployment wage outcome in column (5) of Table 4.

VI. Reconciling the Individual and Market-Level Effect of the Policy

We have documented fairly large responses of the duration of UI receipt and nonemployment to changes in potential duration. In this section we ask how the cut affected the aggregate unemployment rate and, further, what the relative magnitude of the change in the unemployment rate and the change implied by the RDD estimates implies about possible spillovers, particularly displacement effects from the treated group crowding out other jobseekers. To this end, we estimate DiD models comparing the unemployment rate in Missouri to a comparison group of states. We then compare the estimated change in the Missouri unemployment rate over the period to the change in the unemployment rate predicted by the estimated change in the survivor function from the RDD models, assuming no market-level spillovers. A comparison of the two series is informative about the degree of spillovers.

In Figure 10 we plot the raw difference between the deseasonalized unemployment rates in Missouri and the average of all other states by month. The figure shows what appears to be a decline in the unemployment rate in Missouri coinciding with the duration cut as we see a relative reduction in the Missouri unemployment rate, peaking at just over 1 percentage point, following the April 2011 cut.

In Figure 11 we compare Missouri to a synthetic control using the method of Abadie and Gardeazabal (2003) and Abadie, Diamond, and Hainmueller (2010) which assigns weights to states as to minimize the mean squared prediction error between the treatment and control states in the pre-intervention period for a set of outcomes. To

\textsuperscript{18} Our data contains information on quarterly earnings.
construct weights for the comparison group, we use as predictors the unemployment rate for each month from January 2009 – March 2011 and 1-digit NAICS industries.\textsuperscript{19} The figure plots the Missouri unemployment rate against the weighted unemployment rate for the synthetic control. The figure shows a similar drop as when we use the unweighted comparison group of states, with the relative unemployment rate declining, peaking at almost a one-percentage point decline, and then gradually reverting back to the control.

Next we compare these relative changes in the state unemployment rate to the changes in the unemployment rate predicted by the RDD estimates assuming no spillovers. For every week $\tau$ relative to the week of the benefit cut ($\tau=0$), we compute the predicted change in the number of unemployed ($\Delta \hat{n}_\tau$) due to the policy as:

$$
\Delta \hat{n}_\tau = \sum_{t=0}^{57} (\hat{p}_t^T - \hat{p}_t^C) \ast c_{\tau-t} + \sum_{t=58}^{73} (-0.05) \ast c_{\tau-t},
$$

where $c_{\tau-t}$ is the number of initial UI claims in week $\tau - t$ if $\tau - t \geq 0$, $c_{\tau-t} = 0$ if $\tau - t < 0$, and $\hat{p}_t^T$ and $\hat{p}_t^C$ are the estimated probabilities that UI recipients are receiving benefits $t$ weeks into the spell for the treatment and control groups respectively. An underlying assumption, which the analysis above supports, is that pre-exhaustion exits out of UI represent moves out of unemployment and into employment. For UI recipients who first received benefits 58–73 weeks prior to the week of April 13, we make the assumption that the relative difference in the relative exit rate out of unemployment between treatment and control is the RDD estimate for the employment probability outcome in 2012 Q3. We assume that after 73 weeks, beyond the duration of the program in the control period, there are no differences in relative unemployment exit rates, an assumption that is consistent with the insignificant employment probabilities between the

\textsuperscript{19}The procedure assigns weights of 4.3\% to Arizona, 37.2\% to Georgia, 7.7\% to Idaho, 2.1\% to Indiana, 7.6\% to Massachusetts, 23.2\% to Minnesota, 12.1\% to Oklahoma, 5.9\% to Pennsylvania, and 0 to all other states.
two groups after they both exhaust. We then compute the predicted change in the unemployment rate in each week after April 13, 2011 as $\Delta \hat{n}_t/l_t$, where $l_t$ is labor force participation.

Figure 12 plots the predicted change in the state unemployment rate by week against the DiD estimates (by month) of the change in the Missouri unemployment rate expressed relative to the value in March 2011, the month before the cut. The DiD estimates not only line up closely to the predicted change, but the series exhibit the same U-shaped pattern with the unemployment decline, peaking at close to 1 percentage point and kinking up at the same time as the predicted change. It appears that the assumption of no spillovers used to form the predicted response is appropriate as the increased exit rate of the UI applicants translated into a lower unemployment rate.

Table 5 reports the estimates for the DiD models fit over the 2009–2013 period and with the intervention period defined as April 2011 through December 2013. The unit of observation is at the month-by-state level, and we estimate all models with state fixed effects, calendar month dummies, and with and without a Missouri-specific trend. Computing standard errors is complicated in cases where there is only one intervention unit. The primary concern when using grouped data in a DiD analysis is how to account for possible serial correlation (Bertrand, Duflo, and Mullainathan 2004). Though we use data from all 50 states and the District of Columbia, we cannot cluster on state because the relevant degrees of freedom is the number of intervention units (Imbens and Kolesar 2012), which in this case is a single state. As an alternative, we employ a number of different approaches for inference. For the unweighted DiD estimates we report OLS standard errors, panel-corrected standard errors, confidence intervals from a wild

---

20 We have also estimated models with state-specific trends, which yield almost the same point estimates. However, these models are not well suited for bootstrapping so we opted for the more parsimonious model.
bootstrap using the empirical t-distribution (Cameron, Gelbach, and Miller 2008), and the percentile rank of the coefficient from a permutation exercise where we estimate a placebo effect of the cut for every state for the post-April 2011 period. For the synthetic control estimates, we report the percentile rank from the permutation exercise. Specifically, for every state we form its state-specific synthetic control and compute the mean difference in the outcome between the state and the state-specific control as if the state were treated. Table 5 also includes the average post-intervention predicted change in the unemployment rate from the RDD estimates, which can be compared to the DiD estimates as assess the degree of spillovers.

The DiD estimate using the unweighted control is -0.94 percentage points (column 1), and -0.82 percentage points with a Missouri-specific trend. These estimates are interpretable as the difference in the Missouri unemployment rate in the period April 2011-December 2013 relative to January 2009-March 2011 and relative to the average change in all other states. The estimates are statistically significant from 0 as well as from the predicted change in the unemployment rate, in both models using OLS standard errors, panel corrected standard errors, and the wild bootstrap confidence intervals. The percentile ranks are 9.8% (column 1) and 0.0% (column 2) meaning that in specification 1, 9.8 percent of states have more negative estimated effects while in specification 2 no states have more negative estimated effects. Column (3) presents the synthetic control estimates. The DiD point-estimate is -0.78, which has an associated percentile rank of 3.9 percent. The estimate is also close but somewhat larger than the predicted change.

Next we separately look at the numerator and denominator of the unemployment rate. In Table 5 columns (4)–(6) we estimate the same models using the log of the number of unemployed as the dependent variable. Across specifications, we see large
and significant declines in the number of unemployed, in the range of 10–12 percent depending on the specification. These estimates are close to the predicted change in the number of unemployment from the RDD estimates of 8.7 percent. Columns (7)–(9) report the estimates for log size of the labor force. The estimates range from a −0.5 percent decline in the unweighted control with a Missouri-specific trend (column 7) to a −0.9 percent decline in the synthetic control (column 9). While the upper wild bootstrap confidence interval is below 0 in column (7), the permutation percentile ranks of the estimates are large at 25.5 percent and 19.6 percent suggesting that there is little evidence of a statistically significant change in the size of the labor force. The change in the unemployment rate appears to be driven instead by a change in the number of unemployed rather than a change in the size of the labor force.

In Appendix Table 2 we reproduce this analysis using these measures derived from the Current Population Survey. The magnitudes are close to those from LAUS, and while noisier they are still reasonably precise. This analysis shows that our findings are not driven by how the LAUS data are constructed.

We have also computed p-values for the difference-in-differences estimate of the effect of the policy change on the unemployment rate based on the approach of Ibragimov and Müller (2014). To implement this test we limit the sample to 28 months on each side of the policy change, and collapse the monthly difference between the Missouri and the average of the comparison group unemployment rates (denoted for convenience $U_{MO-CO,t}$) into blocks of months of varying sizes (28, 14, 7, 4, 3, and 2 blocks in each of the pre and post periods). We then conduct a two-sample t-test of equality of $U_{MO-CO}$ in the pre and post periods using the collapsed data and N-2 degrees of freedom. In these tests the sampling variances are estimated from variation in $U_{MO-CO}$ across blocks of months, and in doing so we assume independence of $U_{MO-CO}$ across blocks of
months, but allow for arbitrary correlation within blocks. Under the conventional assumption of weak dependence in time series data, observations that are far apart will be less correlated to each other than those close together, and we would therefore expect less auto-correlation when grouping more months together into larger blocks than smaller blocks. By comparing p-values across block groups we can assess the degree to which the inference is serially robust. Looking across the columns of Table 6, this indeed appears to be the case. For the unweighted and synthetic controls we can reject equality of the pre and post period values of $U_{MO-CO}$ for all block groupings, even when we collapse the sample to just two blocks on either side of the cut-off, where auto-correlation should be minimal. Appendix Table 3 shows the same test for the CPS derived sample.

Our conclusion from this analysis is that there is reasonably strong evidence that the increase in exit rates translated into a lower unemployment rate. Moreover, while an important caveat is that in a single unit intervention it is not straightforward to compute correct standard errors, the point-estimates suggest that there were limited displacement effects due to the higher employment rates from the treated group. This analysis also supports another assumption: that the behavioral response is not local to the time of the policy change. If the effect were transitory, we would not expect to see a pronounced and growing change in the state unemployment rate.

VII. Discussion

The UI receipt estimates imply that a one-month reduction in potential UI duration leads to a 15-day reduction in UI spells (marginal effect = 0.5) and 10 fewer days of nonemployment (marginal effect = 0.3). These estimates are larger than what has been typically found in the literature using data from earlier decades in the US and in Western
Europe. Katz and Meyer (1990) estimate marginal effects of changes in potential UI duration on UI spells in the range of 0.13–0.2, and Card and Levine (2000) who find a marginal effect on UI spells of 0.065. The estimates on nonemployment duration are larger than those in Card, Chetty, and Weber (2007), Schmieder, Von Wachter, and Bender (2012), and Lalive (2008) who find marginal effects on nonemployment in the range of 0.09–0.13. Our estimates are closer to Le Barbanchon (2012) (marginal effect = 0.3) and Centeno and Novo (2009) (marginal effect = 0.22). As pointed out in Schmieder, Von Wachter and Bender (2012), who draw from Baily (1978) and Chetty (2008), a summary measure of the disincentive effects of changes in UI duration is the ratio of the effects of changes in potential UI duration on nonemployment and UI receipt. In our setting this ratio (≈ 0.67) is approximately twice as large as the ratio in Schmieder, Von Wachter and Bender (2012) (≈ 0.3–0.4).

We find that the increased hazard rate out of unemployment insurance occurs in the first twenty weeks of the UI spell and then stabilizes. While there is previous evidence of this kind of anticipatory effects (Schmieder, Von Wachter, and Bender 2012; Card, Chetty, and Weber 2007), it is perhaps surprising that the hazard rate of exit spikes early and then stabilizes. It is possible that the media attention following the policy made the duration cut more salient in the minds of some UI recipients, resulting in increased search intensity. However, this explanation would imply that the change in behavior is mainly local to the time of the cut, and less pronounced for subsequent cohorts of UI recipients. As discussed, since the path of the unemployment rate tracks the predicted path, which is based on the assumption that the change in the survivor function is permanent, this explanation is less compelling.

Another explanation is that recipients were confused by the policy change, believing that the cut would give them only 20 weeks of benefits and not the federal benefits which
were an additional 37 weeks. It is possible that recipients interpreted the law in this way, but our review of media reports and Missouri communications to UI recipients provide no evidence that the information disseminated would lead to this kind of confusion. The media coverage at the time emphasized that the reduction was a compromise to preserve extended benefits (e.g. Young 2011). The initial packet sent to claimants before and after the law change was identical and did not explicitly state the number of weeks of eligibility for regular UI. Rather, the notice states the maximum benefit and the weekly benefit. The number of weeks of eligibility would be derived from the ratio of these two numbers (see Appendix Figure 5 for an example of this document). No other wording was changed and no information about extended benefits was provided in the initial packet for either the treatment or control group. Instead, the claimant were informed when they logged into Missouri UI website (MODES) whether extended benefits were in effect and they also received a call informing them that extended benefits are available. When the claimant exhausted their benefits they were reminded in correspondence that EUC was available and eligible claimants were automatically enrolled. These procedures did not change with the law. Because the policy change was so clearly described even in the headlines and the information regarding regular and extended benefits were continuous at the time of the policy change, we find it difficult to sustain an argument that policy understanding was affected discontinuously at the threshold. Nevertheless, if there was confusion, it is interesting that some exiting recipients responded well before the 20 week mark and were largely able to find employment.

The findings suggest that there is a forward-looking group of recipients who respond early to changes in potential duration. However, the long-term unemployed who exhausted benefits did not have higher rates of reemployment relative to the group that
remained on UI. This can be seen most clearly in the comparison of employment rates during the period that the treated group had no benefits remaining while the comparison group remained eligible. There is no evidence that the employment rate rose for the group exhausting benefits during this period (with the caveat that the control group at this point has a different composition near exhaustion as it contain a subset of the “forward-looking” types). This finding suggests that the benefit cut increased reemployment rates for a subset of individuals who responded early in the spell, but for the remaining recipients UI continued to serve an insurance function with limited moral hazard response.

The estimated macro effect of the cut is also larger than what other papers have found. Marinescu (2014) estimates that a 10 percent increase in benefits corresponds to a 0.7 percent decline in the unemployment rate and Hagedorn et al.’s (2015) estimates imply that a 10 percent decrease in maximum benefit durations led to a 1.7 percent decrease in the unemployment rate as a result of the decrease in the number of unemployed. In our case, a 10 percent decrease in benefits was associated with approximately a 5 percent decrease in the unemployment rate. Unlike Hagedorn et al. (2015), however, we find no evidence that the benefit cut increased the labor force participation rate.

Finally, we provide direct evidence on the relative magnitudes of the micro and macro elasticities with respect to potential UI duration. Unlike Lalive et al. (2013), we find that the macro elasticity is at least as large as the micro elasticity. Within the framework of Landais et al. (2010), this finding is consistent with a horizontal aggregate labor demand curve and suggests that the assumptions of the Baily-Chetty model of optimal UI, which assume no spillovers, are appropriate in this setting. We note that while the seasonally-adjusted Missouri unemployment rate was high at the time of the benefit cut, at 8.6 percent, the labor market nationally was mending, and the finding that the market
largely absorbed the larger number of workers exiting UI without displacement may not hold when the unemployment rate is even higher or on an upward trajectory.
CHAPTER 2

Teacher Compensation, Quality, Retention, and Student Achievement:
Evidence from a Regression Discontinuity

Introduction

The United States spends over $304 billion per year in public-school teacher compensation each year ($6,100 per public school student).\(^{21}\) Despite this extensive public spending, we know very little about how increased compensation would influence key education outcomes including student achievement. Because effective teaching improves a student's future income and citizenship, improving teacher quality promises a productive response to poverty (Heckman, 2000; Chetty et al., 2011). Unfortunately, teacher quality is measurably worse at low-income schools (Mansfield, 2013). Perhaps compensation can be increased at disadvantaged schools to attract and retain capable teachers, improving equal opportunity. This paper exploits a policy that attempts to do just that.

Efforts to estimate the relationship between compensation and outcomes including teacher quality, teacher retention, and student achievement have yielded mixed results. Several authors have emphasized the difficulty in identifying the causal relationship between teacher salaries and these variables (Aaronson, Barrow, and Sander 2003; Hanushek 1997; Murnane 1975; Murnane and Phillips 1981, Hanushek et al., 2005). Historically, two challenges exist. First, most variation in teacher compensation is confounded with local priorities, resources, compensating differentials, and student ability and motivation. Second, there is much disagreement on how to measure teacher quality (Hanushek et al., 2005).

\(^{21}\) This does not account for unfunded pension liabilities or the 11 percent of students in private schools
In this paper, I implement a regression discontinuity design to isolate exogenous variation in teacher compensation to evaluate how early-career compensation influences teacher quality, retention, and student performance. Since 1997, the federal government has offered teachers in low-income schools loan forgiveness for most of their federal education loans. A teacher is automatically eligible if at least 30 percent of students at her school are considered "low-income". Specifically, at least 30 percent of the students enrolled must meet a measure of poverty under 1113(a)(f) of the Elementary and Secondary Education Act (ESEA); in practice this is the proportion of students that qualify for reduced-price lunches since these data are easily available to school personnel.

Over the years, teacher salary schedules have become increasingly back-loaded as real starting salaries have fallen and real salaries for experienced teachers have risen. A question of current interest is whether, in principle, compensation should be more "front-loaded" (Lankford and Wyckoff, 1997). Answering this question is difficult due to multicollinearity between portions of the salary schedule, the gradual onset of back-loading, and omitted variables describing selection into this practice. The loan forgiveness program is a valuable natural experiment because it only raises compensation in the first five years of teaching and does not affect benefits, salary, or pensions.

I exploit this loan forgiveness policy to estimate the effect of a substantial increase in teacher compensation on teacher quality, which provides two distinct contributions to the literature. First, this study is the first to use a quasi-experiment to

---

22 There exists loan forgiveness for all Perkins, Stafford, Direct, and Federal Family Education loans if the teacher remains five years; only PLUS loans are excluded which are federal loans parents can take out for their children's education expenses and make up a minority of federal loan volume.
overcome unobservables and endogenous choice in measuring the effect of teacher compensation. Second, I isolate the impact of raising starting salaries, which my theoretical model suggests is the strongest compensation determinant of teacher quality and retention.

Moreover, this intervention does not come from local school funding so treatment does not impact schools by affecting other expenditures through the school’s budget constraint. Finally, I link student test scores to teachers to estimate an objective measure of teacher quality rarely available to previous researchers studying the effects of teacher compensation. I apply my estimators to a rich North Carolina data set that merges student, teacher, and school data from 1995 to 2010 including annual measures of student test scores, teacher assignment, teacher characteristics, and school characteristics.

Treatment does not appear to affect traditional measures of teacher quality; schools are no more likely to hire applicants with a master’s degree, a math or science major, or come from more selective undergraduate institutions. However, they are significantly more likely to hire an education major, which positively predicts student gains in my data. Treatment also appears to have significant positive effects on teacher retention, increasing the probability a teacher remains at the school for five years by about 20 percent. Loan forgiveness does not appear to increase teacher effort, as proxied by a teacher’s non-sick absences. Higher compensation through loan forgiveness appears, however, to improve student achievement significantly (conservative estimates are 0.1 standard deviations in math and 0.07 standard deviations in reading).

The fundamental causal link brought forth by my theoretical model is that compensation can affect hiring quality through two channels: One, compensation can
improve the quality of applicants and thus improve the quality of new hires even if the hiring process is non-discriminating; two, compensation can improve the quality of hires by expanding the applicant pool, but only if schools can discriminate quality. It is likely that the loan forgiveness only works through the second channel (expanding applicant pools); thus the results suggest that schools successfully discriminate on characteristics that are not observed by researchers. This finding suggests that raising compensation can raise the quality of new hires, even if the compensation doesn't improve applicant quality.

Because this treatment can be interpreted as an exogenous supply shifter, I can estimate the demand for various applicant characteristics. These parameters represent how schools value various applicant characteristics when hiring. The results indicate schools prefer applicants who majored in education and dislike applicants who majored in social science. Somewhat surprisingly schools have no preference for applicants with master's degrees, applicants from more selective colleges, or applicants with rigorous degrees in math and science. Schools appear to have no pronounced preference for teacher race or sex.

Ballou (1996) suggests that schools do not prioritize the applicant qualities that best predict teaching success. My results suggest that, in North Carolina, teacher selection is consistent with a model of achievement maximization subject to a budget constraint. That is, schools appear to prioritize applicant characteristics in approximate proportion to the characteristic's predicted student-achievement effect.

There is no evidence that schools strategically sort into the program or that treatment schools differ systematically. Moreover, results are robust to a variety of alternative specifications and controls. Taken as a whole, results suggest that loan forgiveness enlarges the applicant pool at treatment schools and schools appear to
select applicants “correctly” by hiring teachers with characteristics that predict effective teaching.

Although some of the literature reports a positive correlation between teacher compensation and student or teacher outcomes (student retention and hiring from selective schools), my results suggest that unobserved factors may be biasing these previous estimates. For instance, Figlio (1997) reports a significant correlation between geographically varied starting salary and whether a hire had a hard science major. By contrast, I find no correlation when using quasi-experimental variation. Instead, teacher salaries may be correlated with unobserved factors including school amenities, student characteristics, and principal networks.

This paper fits into a small set of papers that estimate the impact of teacher compensation on student outcomes or teacher characteristics. Loeb and Page (2000) use changes in local average wages as an instrument on the local relative wages of teachers; this design implies that a 10 percent increase in relative wages reduced the student dropout rate by three to four percent. These results are hard to interpret since the instrument fails the exclusion restriction: local wages directly affect the opportunity cost of a high school student remaining in school. Hanushek, Kain, and Rivkin (2005) used changes in first-year teacher salaries and found no correlation with teacher effectiveness (as measured by value-add) or certification scores. Clotfelter, Glennie, Ladd, and Vigdor (2008) used a difference-in-difference-in-difference approach to determine the effect of a retention bonus on turnover; the bonus was announced after hiring decisions were finalized so the bonus could not affect teacher quality, but the incentive reduced the turnover rate by 17 percent. In general, this literature has been haunted by endogeneity and omitted variables bias, and has been able to only estimate
effects on a few outcomes at a time, accounting for meager progress on this critical issue.

Background

Not only are teachers the most significant public input in the education production function, but also high value-add teachers improve the life outcomes of their students (Mansfield, 2013; Chetty et al., 2011a, 2011b). Thus the distribution and allocation of teacher quality affects income mobility and economic growth in the coming decades. Unfortunately, since 1960, teacher quality measures that correspond to teacher value-add have been steadily falling in the United States (Murnane et al., 1991, Bacolod, 2007). For instance, while 20 percent of top graduates entered teaching in 1964, fewer than four percent did in 1992 (Corcoran et al., 2002). At the same time, teacher experience, the most predictive characteristic of teacher value-add, has quickly fallen due to rising turnover and reduced class-size (Ingersoll, 2003).

Numerous explanations exist for these falling quality and retention patterns, including greater opportunities for women and wage compression demanded by teacher's unions (Hoxby and Leigh, 2004; Bacolod, 2007). I argue that the evolution of teacher salary schedules has also reduced teacher quality and retention. Policymakers have allowed the real value of starting teacher salaries to fall even as real salaries for experienced teachers rose (Lankford and Wyckoff, 1997). Thus a starting teacher earned less in 1994 than she did in 1970, while a teacher with 20 years of experience earned 20 percent more. In the North Carolina data, this pattern of shrinking starting pay and growing pay for veteran teachers was sustained throughout the 1990s and into the 2000s.

Of course, all compensation is likely to affect teacher labor supply, but starting wages may be especially important in attracting and retaining quality teachers. Both
theory and survey data suggest the relative importance of early compensation levels. In Appendix A, I present a sequence of simple models demonstrating that starting pay is especially influential when individuals choosing a career are shortsighted, have limited information about salary progression in various careers, or are uncertain about whether they will enjoy their initial career choice. From survey data we learn that 79 percent of teachers leaving the profession (most of whom are new teachers) cite low salary as their primary reason for departing; this is by far the most common rational teachers give for departing and it accounts for a larger proportion of responses than the next three largest reasons combined (Ingersoll, 2003). My analysis is the first to test the impact of increased starting pay on teacher quality, effort, and retention.

A federal program introduced in 1997, the Teacher Cancellation Low Income Program, offered student-loan forgiveness to new teachers serving in low-income schools across the country. Loan forgiveness covered most federal education loans (subsidized and unsubsidized Stafford and Perkins loans). To promote eligible low-income schools states submit a list of the schools where at least 30 percent of students qualify as low-income to the federal government each year. These lists are made available on the internet, so that prospective teachers can easily identify qualifying schools.

Perkins and Stafford loans undergo different forgiveness schedules. Perkins forgiveness is pro-rated: 15 percent of the principal is forgiven after the first year of teaching at a qualified school; an additional 15 percent is forgiven after the second year, 20 percent after the third, another 20 percent after the fourth, and 30 percent after the fifth year, totaling 100 percent at the end of five years. The current maximum cumulative

---

23 Schools and Staffing Survey (SASS) provided by the National Center for Education Statistics
24 The only federal loans forgiveness does not cover are PLUS loans which accounted for a small proportion of federal loan volume.
Perkins loan is $27,500 for undergraduates and a lifetime maximum of $60,000 for graduate students. By contrast, $5,000 of Stafford debt is forgiven only at the end of the fifth year teaching. Up to $17,500 of Stafford loans can be forgiven for teachers in low-supply positions like mathematics and science. In the value-add portion of the analysis below, I focus on elementary school teachers where I can reliably link students to their teachers; for this subset, the larger Stafford forgiveness was rarely available.\textsuperscript{25}

The loan-forgiveness eligibility rule is central to my identification strategy. Eligible schools are those which had at least 30 percent of students designated as low-income. A teacher was eligible if her school was eligible, and the teacher retains her eligibility even when the school does not. For instance, consider a teacher hired at a qualifying low-income school. After her first year, the school falls below the eligibility requirement of 30 percent. Hence the school’s new teachers would not be eligible, but a teacher who taught in an eligible year still qualifies. If a teacher initially chose an ineligible school and the school then became eligible, she would qualify for loan forgiveness with the first year of school eligibility being the first to count towards her loan forgiveness. Thus teachers could have been covered by the program intentionally or unintentionally, allowing me to differentiate the effects of compensation on retention from the retention associated with teacher selection into treatment.

This is an attractive environment to study the effects of teacher compensation for several reasons. First, the loan forgiveness program acts more like a compensation package than an add-on incentive scheme such as an explicit retention incentive, which could exert effects independent of compensation. Second, the environment allows me to test the impact of the most visible form of compensation (upfront compensation) on a

\begin{footnote}
\textsuperscript{25} The full $17,500 Stafford forgiveness would be available to elementary school teachers providing special education.
\end{footnote}
sensitive margin of teacher supply (new hires). Third, teacher compensation is locally randomized at the threshold which changes from year to year, providing not only exogenous variation in teacher compensation within and across schools.

The compensation increase for a given teacher depends on how much qualifying debt she has. The data does not include information on a teacher's debt load or whether she receives loan forgiveness. Moreover, it's unclear whether teachers perceive loan-forgiveness and salary increases of the same amount to be equally valuable. Because there is no equivalent quasi-randomization in salary, I am unable to re-scale the loan-forgiveness effects into salary effects. Thus I rely on the fact that a dollar less in loan payments is equivalent to a dollar increase in salary, for low levels of taxation.

In 2009, the average recent U.S. graduate had $15,705 in debt (in 2009 dollars), a tally that includes 38 percent of students who graduated without debt. The average student with debt had $24,822 in loans (Avery and Turner, 2012). 81 percent of the $115 billion in loans disbursed that year were public, most of which would be eligible for loan forgiveness if the graduate became a teacher in a low-income school. Like students in other states, most students in North Carolina take out Stafford loans, half taking unsubsidized and half taking subsidized Stafford loans each year. Thus, most graduates in North Carolina would have received around $5,000 in additional compensation for teaching in a low-income elementary school. A relatively small proportion of students (3 percent of students per year) receive a Perkins loan with an average yearly amount of $2,125. I could not obtain individual loan data for teachers in North Carolina, but my estimate is that among those who had debt, they would have qualified for an average of $5,319 in loan forgiveness if they taught in a low-income elementary school for five years. This represents the equivalent of a three to four percent salary increase over the
Theoretical Model

I motivate the empirical work to follow with a simple model. The fundamental result of the model is that increased compensation can raise teacher effectiveness through three clear channels, two of which raise the quality of new hires. First, it can increase the average quality of applicants. Second, it can increase the size of the applicant pool. Third, it can retain teachers which improves their effectiveness but not their baseline quality (Wiswall, 2013). Understanding whether compensation does improve the quality of the applicant pool and whether schools can discern quality from a larger applicant pool is critical for policy recommendations. If compensation does not improve the quality of applicants and schools cannot discern applicant quality, increased compensation is unlikely to yield a better workforce.

In the first stage of the model, the agent chooses a career based on the average expected utility of two choices, teaching and her best alternative. In the second stage, she chooses which firms (including schools) to apply to given that career. Choice functions are monotonic and binary descriptions of the worker's expected utility. Teacher utility is a function of job-match quality ($\alpha$), total compensation ($s$), working conditions ($k$), and entry costs ($c$) of individual $i$ in firm $f$ in career $J$:

$$U_{if}^J = u(\alpha_i^J, S_f^J, k_f^J, c^J)$$

The match-quality parameter, $\alpha_i^J$, for individual $i$ in career $J$ is drawn from a distribution $F$ on $(-\infty, \infty)$. Agents do not directly observe their match-quality parameters, but they receive a noisy signal of match quality $\theta_i^J$ for each career $J$ before making their initial career choices when entering the labor market. The $\theta_i^J$ signals are assumed to be...
normally distributed with mean $\alpha_i^J$ and variance $\sigma_i^2$. Once agents are in the labor market they receive an additional signal, $\theta_i^J$, each period they work in career $J$; these signals are independent and identically distributed. Thus, through experience in career $J$, the agent acquires a more precise estimate of the match quality, approaching $\alpha_i^J$.

The agent chooses either to teach or to pursue her alternative career every period. At time 0, the agent picks the career for which her expected utility is higher, with utility of the form:

$$U_i^T = \alpha_i^T - c^T + \bar{k}^T + \sum_{t=1}^{Y} \beta^{(t-1)} S_t^T$$

$$U_i^A = \alpha_i^A - c^A + \bar{k}^A + \sum_{t=1}^{Y} \beta^{(t-1)} S_t^A(\psi_i)$$

As before, $\alpha_i^J$ describes the match quality between individual $i$ and career $J$; $c^J$ is the one-time cost of entering career $J$; and $\bar{k}^J$ is the average working conditions in career $J$. $\beta$ is the discount factor; $S_t^J$ is the average salary sequence for career $J$ and $\psi_i$ is the ability of the worker. In words, agents make a first-stage career choice based on match, entry cost, average working conditions, and expected compensation. The alternative career may pay more to a worker with high $\psi$. By contrast, teacher compensation is determined by a schedule and is not influenced by employee quality. People choose to teach if the expected value of teaching exceeds that of the alternative career. Thus an individual becomes a teacher if and only if:
Accordingly, the probability of being a teacher depends on a monotonic increasing function of teaching-career match quality and a monotonic decreasing function of the match quality of the alternative, *ceteris paribus*. With higher entry costs of teaching, the individual is more likely to choose the alternative. Finally, as the relative compensation of teaching declines (including through higher $\psi_i$), the individual becomes less likely to enter teaching.

The first channel describing how increased teacher compensation could affect teacher quality is now visible. Namely, higher salary sequences in teaching will induce higher $\psi$ individuals to apply for teaching positions.

The expected difference in match quality is uncertain, and agents update every period as new information shocks arrive. Let

$$v = (c^A - c^T) + (\bar{k}^T - \bar{k}^A) + \sum_{t=1}^{T} \beta^{t-1} (S_t^T - S_t^A(\psi_i))$$

which is the net value of teaching conditional on match quality. Therefore individuals enter teaching if and only if:

$$E[\alpha_i^T - \alpha_i^A] + v > 0$$

$$E[\alpha_i^T] - E[\alpha_i^A] > -v,$$

Because the distribution of shocks $\theta_i^J$ has mean $\alpha_i^J$, the expected value is the average of shocks received:

$$E[\alpha_i^J] = \frac{1}{n} \sum_{i=0}^{n} \theta_i^J .$$

We call this expectation $\overline{\theta}^J_i$. In period 0 the agent chooses teaching if: $\theta_{(i,0)}^T - \theta_{(i,0)}^A > -v.$
As information shocks arrive, the difference between $\bar{\theta}_{iT}$ and $\bar{\theta}_{iA}$ may fall below $-v$ but the teacher will not necessarily leave teaching because she has already paid $c^T$ but has not paid $c^A$.

People who initially choose teaching may be divided into three categories, defined as $A'$, $B$, and $C$ for switchers. These are defined as follows:

$$A' = \{ i \mid \bar{\theta}_{iT} - \bar{\theta}_{iA} > -v \}$$

$$B = \{ i \mid v - c^A < \bar{\theta}_{iT} - \bar{\theta}_{iA} < -v \}$$

$$C = \{ i \mid \bar{\theta}_{iT} - \bar{\theta}_{iA} < -v - c^A \}$$

In any given period, the $A'$ teachers are those who chose teaching because in expectation the career delivers higher utility, and their current stock of information shocks confirms their expectations. The $B$ teachers are those who now believe that they made a mistake in choosing teaching, but the differential does not justify the upfront cost of moving to their alternative career. Finally, those in set $C$ are those for whom the differential is sufficient to justify incurring the cost of switching careers, including $c^A$ and potentially starting at the beginning of the alternative career's salary progression. In other words, the expected difference in the match quality must also surmount the cost of entering the alternative career. Notice also, that as $c^A \to 0$, the set $B$ shrinks and the proportion of switchers grows. It seems that in many careers, teachers might consider alternatives, the cost of entering are quite low (e.g. secretarial work).

Here we can see the second insight of the model---that raising teacher salaries could increase teacher retention. Since experience increases a teacher's effectiveness (Wiswall, 2012), though not her latent quality, increasing salary may increase student
achievement, without even affecting applicant quality or, as we will see, principal quality predictions.

Next I add another step, the application process, which incorporates the hiring decision, heterogeneity in the salary sequence, and working conditions of various firms in a career. Once individuals choose a career, they apply to employers that offer sufficient utility. Since the entry costs are already paid, they apply to firms that provide more utility than switching careers. For instance, if an individual chooses teaching, she applies to schools satisfying:

$$U^T_{i(f)} > \overline{U^A} - c^T,$$

If $$z = \alpha^A - \alpha^T - c^A$$, she will apply to those schools for which:

$$k^T_f(\rho) - \overline{k^A} + \sum_{t=1}^{T} \beta^{(t-1)} (S^T_t(\rho) - \overline{S^A}_t(\psi_i)) > z$$

In words, each agent applies to the firms that deliver higher utility than the utility of switching careers. Here, the salary and working conditions in teaching depend on the percent of students that are low income. Low-income areas tend to pay less, but teachers receive a loan-forgiveness subsidy if a sufficient proportion of the students at the school are low-income ($$\rho \geq 30\%$$). Working conditions are also highly correlated with the student poverty. Teachers leaving teaching commonly cite student discipline and motivation, both of which are highly correlated with student poverty ($$\rho$$) (Ingersoll, 2001).

School principals select the vector $$h$$ of hires from the applicant pool according to a selection function based on teacher expressed match quality ($$\theta^T_{i(0)}$$), observed by the principal as enthusiasm for teaching; the principal also observes a vector $$\overline{x}_i$$ which researchers can observe, which is correlated with $$\psi_i$$. $$\overline{x}_i$$ includes variables commonly observed at the time of hire that represent applicant training, work ethic, and natural ability, including undergraduate selectivity, certification scores, grade point average,
additional certificates and degrees. Principals also observe \( \vec{i}_t \) which is a vector of interview qualities including personality which a researcher cannot observe but may predict teacher quality. Accordingly:

\[
h = h(\theta^T_{t,0}; \vec{x}_t, \vec{i}_t)
\]

That is, the Boolean vector function \( h \) represents the hiring decision for each applicant as a function of the applicants' research-observable characteristics (\( \vec{x}_t \)), interview-observable characteristics \( \vec{i}_t \), and the match quality (\( \theta^T_{t,0} \)). While little is known about the teacher hiring function, it appears that hiring largely ignores academic success and other observable characteristics that appear to predict effective teaching (Ballou, 1996; Harris et al., 2010).

In a competitive market, principals would value student achievement intrinsically or be compensated for student achievement. In such a case, principals would know through experience, training or research, the relative importance of the elements of \( \vec{i}_t \) and the value of \( \theta^T_{t,-1} \). The appropriate approach then would maximize student achievement, a non-trivial problem since those with high academic performance may have better outside options and thus be more likely to leave. Teacher turnover, replacement costs, and the expected quality of the replacement are all factors. The teacher hiring function \( h(\cdot) \) would indicate 1 for the applicants that satisfy:

\[
max_{i \in B} V_i = \sum_{t=1}^{Y} p_{(t, \vec{i}_t)} q_t (\vec{\psi}_t(\vec{x}), t) + (1 - p_{(t, \vec{i}_t)})(V_r - k)
\]

where \( V_r = E[max_{i \in B'} V_i] \).
In other words, the principal will hire applicant $i$ from applicant pool $B$ to maximize total student achievement, balanced with hiring costs.

- $p_{t,\bar{\psi}_i}$ is the probability that in year $t$ a teacher of expected quality $\bar{\psi}_i$ will remain at the school, a dynamic separating probability;
- $q_t(\bar{\psi}_i(\bar{x}), \tau)$ is the expected achievement gains of a teacher of expected quality $\bar{\psi}_i$ in her $\tau$th year of teaching at the school;
- $V_r$ is the expected value of hiring a replacement;
- and $k$ is the cost of replacement.

This reveals the third channel described by the model; if principals can discriminate quality from $\bar{x}_i$ and $\bar{i}_i$ then a larger pool of applicants can generate higher quality hires, even if applicants are no better on average.

Ballou (1996) shows that, given reasonable ranges for retention, quality, and replacement-cost parameters, current hiring practices in education are far from optimal. It appears that principals discriminate against more successful graduates, graduates with degrees in mathematics and science, and degrees from more prestigious institutions. My results indicate that schools do not discriminate but are indifferent.

The properties of the function $h = h(\theta^T_{i, -1}; \bar{x}, \bar{i})$ are crucial for determining whether an increase in the sum of $\beta^T s^T$ will improve the average quality of teachers if applicant quality or the number of applicants change. If principals do not infer quality accurately from $\bar{x}$ and $\bar{i}$ or have hiring priorities orthogonal to teacher quality, raising compensation will not have a positive impact on teacher quality unless it affects the quality of the
average applicant. The compensation increase I use is not likely to bring agents into teaching since it is not part of the compensation package someone out of education would likely be aware of. Moreover, the quality of teachers interested in debt forgiveness is likely no greater than teachers who do need debt forgiveness. Thus, my treatment is able to estimate the impact of compensation through this third channel and reveal whether principals are able to accurately discern quality through unobservable $i$.

Empirical Research Design

In this section I describe my regression discontinuity design (RDD) which approximates a randomized experiment. Teacher eligibility for federal loan forgiveness constitutes "treatment." Whether a teacher is treated depends on whether the school has at least 30 percent of students are designated as low-income. The idea of the design is to compare teachers at schools that were barely above and barely below the threshold. Conditional on the low-income share, the covariates—including unobservables—are balanced across the threshold, yielding an unbiased estimate of the effect of loan forgiveness on any outcome variable. In effect, some schools randomly have an additional component of compensation.

Suppose that school $j$ has $\rho_j$ percent of students eligible for reduced-price lunches (relative to the required threshold $\rho^*$). Let $b_{ijt} = 1(\rho_j \geq \rho^*)$ be an indicator for school $j$'s loan-forgiveness eligibility. We can write some outcome $y_{ijt}$ (e.g. the retention or quality of newly hired teacher $i$ at school $j$ in year $t$, for instance) as follows:

$$y_{ijt} = \kappa + b_{ijt} \theta + u_{ijt}$$

where $\theta$ is the causal effect of loan-forgiveness compensation and $u_{ijt}$ represents all other determinants of the outcome (with $E[u_{ijt}] = 0$).
In general, school eligibility may be correlated with other school characteristics that influence teacher quality and retention, so $E[u_{i(j)} b_j] \neq 0$. If so, a simple regression of $y_{i(j)}$ on $b_j$ will yield a biased estimate of $\theta$. However, as long as the outcome variable is smooth function of other factors at the cut-off, eligibility around the threshold approximates a randomized experiment. In other words, the correlation between school eligibility for teacher loan forgiveness and unobserved district characteristics can be kept arbitrarily small, by focusing on schools sufficiently close to the threshold.

One complication to the empirical design is that schools use different measures of low-income status. I compile the historical data which declared the eligible schools each year. I merge this data into the NCERDC data and use this to create the treatment indicator. Because I merged using school name, the matching is imperfect; for schools that should definitely be treated (e.g. the reduced-price lunch proportion is above 50 percent) about 85 percent of schools have the correct treatment status. I implement a fuzzy RD design, exploiting the large discontinuity in treatment status that occurs when the percent of students on reduced-price lunch crosses the 30 percent threshold.

The first-stage regression is thus:

$$b_{(jt)} = \mu + \delta 1(\rho_{(jt)} > 30\%) + P_{g}(\rho_{(jt)} - 30\%, \gamma_a) + P_{g}^{e}(\rho_{(jt)} - 30\%, \gamma_a) + \zeta_{(jt)}$$

Here we are modeling how the probability of treatment changes with the percent of reduced-price students. $\mu$ will represent the average probability of being treated $\epsilon$ below the threshold, because the running variable is now distance to the threshold. $\delta$ represents the discontinuous jump in the probability of treatment when the reduced-price lunch proportion crosses the eligibility threshold. The results of these regressions with various bandwidths are found in Appendix B. The results suggest that about 39 percent
of schools are eligible just before reduced-price-lunch share crosses the threshold causing the treatment probability to jump by about 27 percentage points.

My RD strategy uses two flexible polynomials of the running variable, so as to use more of the available information. I present a series of results that start with estimates that depend on a fifteen-point bandwidth around the threshold and narrow the focus successively to ten-, five-, and two-points within the threshold. For all bandwidths, I use flexible controls for low-income share which absorbs variation coming from schools far away from the threshold, but reduce the polynomial order as the bandwidth narrows.

Assuming that $E[u_{ij} | \rho_j]$ (the conditional expectation of the unobserved determinants of $y$ given the realized low-income share) is continuous, I can approximate it by a polynomial of order $g$ with coefficients $\gamma_u, P_g(\rho_j, \gamma_u)$, and the approximation can be made arbitrarily accurate as $g \to \infty$. I estimate a polynomial of this form on either side of the treatment threshold. Under this assumption we can rewrite (1) as

$$y_{ijt} = \kappa + \tilde{b}_{jt} \theta + P_g(\rho_j, \gamma_u) + P_g^\theta(\rho_j, \gamma_u) + \omega_t + u'_{ijt}$$

where $u'_{ijt} \equiv u_{ij} - P_g(\rho_j, \gamma_u) = (u_{ij} - E[u_{ij} | v_j] - P_g(\rho_j, \gamma_u))$ is asymptotically uncorrelated with $\rho_j$ (and therefore with $b_j$). A regression of hire characteristics on the qualifying indicator controlling flexibly on low-income share thus consistently estimates $\theta$. I include time fixed effects in all specifications ($\omega_t$) and run additional specifications using a control vector $X$ which consists of other variables that might influence the attractiveness of teaching at a given school including student demographic make-up and school type; the results are identical with and without controls but I lose precision.

I run a series of regressions to infer the effect of teacher characteristics and treatment on student achievement. These regressions are called value-added models (VAM) in the education literature. In their most flexible form, which I adopt, one
regresses the student’s test score this year on factors (teacher characteristics or treatment) and a flexible polynomial of the student’s score last year. The coefficients on the factors is meant to capture the marginal effect of the characteristic or treatment on student achievement. The practice potentially biases estimates because teachers are not randomly assigned to students. I follow Rothstein (2010) and use student math and reading scores as the dependent variable, with the student's last score as an element of X. Rothstein finds evidence that value-added measures are biased when the lagged test score is linear. I experiment with a flexible polynomial of the lagged score which dramatically reduces Rothstein’s measure of bias, the predictive power of fifth grade teacher assignment on fourth grade achievement gains. Intuitively the polynomial accommodates nonlinear relationships generated by performance mean reversion and measurement error.

First I estimate an achievement production function, associating achievement gains with various teacher characteristics while controlling for student characteristics where possible. The regression takes the form:

$$A_{(kijg)} = \alpha + \lambda_t + m_{kijg} + \beta \psi_{(ijg)} + \gamma X_{(kijg)} + f_5(\rho_j) + g_5(A_{(kijg-1)}) + v_{(kijg)}$$

$A_{(kijg)}$ is student k’s test score in teacher i’s class at school j in grade g. $\lambda_t$ represents time fixed effects, $m_{(kijg)}$ represents a number of fixed effects for missing data in the student characteristic vector, $\psi_{(ijg)}$ represents a vector of teacher characteristics, $X_{(kijg)}$ represents student characteristics with zeros filled in for those data that are not available, $f_5(\rho_j)$ is a flexible fifth-order polynomial function of low-income share, and $g_5(A_{(kijg-1)})$ is a fifth-order polynomial of the student’s last test score. This estimates the student gains associated with various teacher characteristics including college selectivity, master’s degree, sex, race, and major.
I then modify the specification to test the effect of treatment. I replace the teacher characteristics with the treatment indicator, instrumented as in previous regressions. I add a polynomial on the treatment side and remove the indicators for missing data and student characteristics.

As discussed, these regressions may not reflect the average quality of treatment teachers relative to control since students are not randomly assigned. I make student characteristics the outcome variable in the regression discontinuity. Here the coefficient on treatment represents how much more likely a treatment teacher is to have students with a given student characteristic. This assesses the extent of student sorting and gauges whether treatment teachers are allocated to the intended low-income students.

Following the initial test associating students growth with the treatment status of their teacher, the subsequent regressions address the concern that students are not randomly sorted to teachers and thus the value-added scores associated with the previous regressions may reflect student sorting in addition to treatment. I modify the RD specifications to pool treatment to the school-grade level. Thus unless treatment teachers are sorted to whole grades who for some reason have different growth, student sorting cannot drive the effect. The specification takes the form:

\[ A_{ijtg} = \alpha + \beta T_{ijtg} + \delta \nu_{ijtg} + f(r_{tj} - .3) + f'(r_{tj} - .3) + g_5(A_{ij(t-1)g}) + \epsilon_{ijtg} \]

\( A_{ijtg} \) is the achievement score of student \( i \) in teacher \( j \)'s class in year \( t \) in grade \( g; T_{ijtg} \) is the share of teachers in \( i \)'s grade who are treated; \( \nu_{ijtg} \) reflects the share of teacher in \( i \)'s grade who are new teachers; \( f \) is a flexible function of the running variable, the share of the school's students who are designated as low-income; \( f' \) is another flexible polynomial on the treatment side of the threshold. \( g_5 \) is a flexible fifth-order
polynomial of the student i's score last year. The instrument on $T_{(jtg)}$ is an indicator for $r_{(tg)} > 30\%$. In all RD regressions, I cluster the standard errors at the school level.

Data

To estimate this model, I combine a number of datasets. Information on teachers and students (including student test scores) comes from a the North Carolina Education Research Data Center which created the data from administrative records from 1995 to 2010 in North Carolina. This data source allows me to link students to their teachers to estimate value-added models. Moreover it includes information on teachers including their race, sex, absences, and education history; information on students includes their race, sex, and age and, in some years, information on the student's home habits, parent education, gifted status, low-income status, and student absences. The data allows me to calculate the share of students on reduced-price lunch which constitutes my running variable.

A common measure of teacher quality is the Barron's Selectivity Index; the National Center for Education Statistics provided this dataset and I merged the university selectivity score to each applicant's undergraduate institution. Finally, to create the treatment variable, I compile historical records of what schools were eligible in each year; unfortunately the school names in the eligibility website are often misspelled, use abbreviations inconsistently, and have no unique identification. Thus I employ an iterative fuzzy match which standardizes common abbreviations, transforms capital letters to lower-case, and removes spaces. This fuzzy match is estimated to achieve 85 percent accuracy.

There are two main data compilations I use in the analysis. The first is a merging of all data above where each observation is a teacher in a year. These data are used for
analyses of teacher characteristics and includes all teachers in elementary, middle, and high schools. The second merges in yearly student test scores for grades three through six, where I can reliably identify a student's teacher. Thus the value-add analyses only applies to elementary school teachers. In this dataset, the observation level is a student in a year.

Importantly, I cannot identify whether a given teacher had eligible loans or whether she applied for loan forgiveness. Therefore my estimates reflect intent to treat (ITT) effects, which are the average effects of eligibility, not the average effect of treatment on recipients (TT).

The data includes approximately 100,000 unique teachers each year which grows from 88,900 in 1995 to a peak of 112,800 in 2007. Analysis is restricted to teachers hired in schools within plus and minus fifteen points of the eligibility threshold and the analysis focuses on narrowing bandwidths of plus and minus ten, five, and two. The largest teacher sample (that for fifteen points within the threshold) allows a sample of 38,751 new hires in 2,101 schools over the 14 years the program was in effect. The smallest bandwidth restricts the sample to 5,381 new hires. In some regressions the dependent variable was more sparsely reported reducing the usable sample size. Eligibility only requires that a school has 30 percent of students designated as low-income, so a large fraction of schools in the sample qualify. 81 percent of schools qualified in 1997, 88 percent qualified in 2000, and 94 percent qualified in 2005.

When the student data is merged in, I focus on those students in classes with new teachers so that I can ignore the complicated effects of teacher tenure and turnover (Wiswall, 2012) and focus on relevant teacher quality. In this sample we observe 325,561 students in classrooms with new teachers over the thirteen years of the
program. The largest bandwidth, fifteen points, includes 143,941 students and the smallest, two points, includes 20,573.

Evaluating the Quasi-Experiment

The regression discontinuity strategy relies on using schools slightly above and below the eligibility threshold, to approximate an intentional experiment. To satisfy the exclusion restriction, this requires that eligibility be as good as random and that other determinants of the outcome variable be continuous across the threshold (Angrist and Pischke, 2009). Here I test these assumptions.

Three diagnostics test the validity of the RD quasi-experiment: pre-program differences in outcome variables and characteristics across schools and hires, the distribution of the forcing variable, and whether observables are balanced across the threshold. Each strongly suggests that the exclusion restriction is satisfied.

In order for the RD estimate to be valid, we must assume that, absent the policy, \(E[\varepsilon_i | \rho_s = c] = 0\). This assumption would be violated if other factors affecting teacher supply such as school amenities, working conditions, or other compensation changed discontinuously when schools reached the 30 percent threshold. This might be observed if, for instance, treatment "effects" existed prior to the program, suggesting some other determinant was changing discontinuously. The results of such regressions are uniformly insignificant; this test is enabled by the panel structure of the data supports the approximate randomness of the design (Cellini, Ferreira, and Rothstein, 2010). Appendix B contains the regression tables and further discussion from these placebo tests. These results support the continuity of \(E[\varepsilon_i | \rho_s = c] = 0\) at the threshold.

A discontinuity in the density at the threshold would suggest that schools endogenously sorted into treatment status, violating the RD requirement (McCrary, 2008). There is no evidence of endogenous sorting.
Finally I also test whether observable school characteristics are balanced across the threshold. Rejecting the null hypothesis would imply that schools either sort endogenously.

In the analysis presented in Appendix B, each coefficient describes the "impact" of treatment on the share of students who are respectively Hispanic, Black, or White. The fact that coefficient estimates are null provides further substantiation that schools did not endogenously sort over the threshold.

Results

The results require some thought to interpret. First, the coefficients on treatment sometimes differ with bandwidth and statistical significance is seldom constant. Thus, I interpret an effect if the coefficients are relatively stable and at least one is statistically significant at the ten percent level. I also place a higher weight on the narrower bandwidths which are more likely to approximate a randomized experiment.

First I analyze the effect of treatment on the characteristics of new hires; recall that these estimates reflect demand parameters since treatment induces an exogenous supply shock. Schools appear to have no preference regarding the sex or race of the applicant; the coefficients here are uniformly small and statistically insignificant. Next I estimate demand parameters for characteristics that have been used as measures of teacher quality: whether the teacher has a master's degree, a teacher's major, and undergraduate institution selectivity. What may surprise some readers is that schools appear to have no significant preference for applicant's with a master's degree or applicants from more selective undergraduate institutions, but schools do express preferences for certain majors. Treatment schools at the cutoff are eight percentage points more likely to hire education majors, nine percentage points less likely to hire a social science major and 4 percentage points less likely to hire a language major,
relative to applicants with unspecified majors. Interestingly, schools have no distinct preference for science, math, or business majors. Because treatment is quasi-random these magnitudes represent school preferences for these characteristics. This is consistent with Ballou (1996) who finds that schools appear to prefer teacher applicants with education majors compared to those with math and science training.

I next look at the behaviors of treated teachers, principally in terms of retention and effort. Treated teachers are no more likely to stay one or two years (conditional on staying one) as virtually all stay during the first two years. However, treated teachers are four percent more likely to be retained three years (conditional on being retained two years), 17 percent more likely to be retained four year, 42 percent more likely to be retained five years, and 33 percent more likely to stay six and seven years, even after loan forgiveness. This reflects the selection effect (the fact that teachers sort into treatment) and the effect of the incentive itself. In order to separate out the pure effect of the incentive on retention, I look at teachers who were unintentionally treated, being those who initially chose untreated schools but then became eligible because their school crossed the threshold. In these regressions, treatment increases three-year retention by eleven percent, four-year retention by 19 percent, five-year retention by 21 percent, six-year retention by 20 percent, and seven-year retention by 20 percent. I cannot directly observe effort, but the number of non-sick absences a teacher has during a year could represent effort's extensive margin. Treated teachers appear to exert no more effort, having the same number of non-sick absences as untreated teachers.

I estimate a student-achievement effect of each teacher characteristic. The university selectivity predicts larger gains in both reading and mathematics, although the effect is small (a one standard deviation increase in college selectivity is associated with a .007 standard deviation (\( \sigma \)) increase in math and .003 standard deviation in reading). A
master's degree is associated with a $.01\sigma$ increase in both reading and math. Males appear to produce $.01\sigma$ more in reading but $.01\sigma$ less in math. There appear to be only statistically insignificant gains associated with teacher race. The most significant gains are those associated with having an education major, yielding a $.06\sigma$ increase in math and $.01\sigma$ increase in reading. Math and science majors ironically appear to predict small gains in reading and small losses in math. Social Science majors appear to predict small losses $.01\text{-}.02\sigma$ in both reading and math. Business majors appear to promote increases in math ($.04\sigma$) and reading ($.02\sigma$). To summarize, the best predictor of teacher effectiveness is having an education major, second is having a business major, third is having a master's degree. Having a math, science, or language major appears to predict very little either way. Having a social science major predicts lower achievement gains. I also include the teacher's certification score, the PRAXIS, and find that it predicts a precise zero effect in math and reading growth (a one $\sigma$ increase in PRAXIS score is associated with a $.002$ increase in math achievement and $.006$ decrease in reading). Why policy makers would exclude any teacher based on this measure is a puzzle.

I estimate the same type of model using these same teacher characteristics to predict retention, as experience is perhaps the most successful explanatory variable in teacher effectiveness. Here we learn that students from more selective schools are significantly more likely to leave teaching; a teacher from a $1\sigma$ more selective school is three percentage points less likely to be retained (on a base of 69 percent). Interestingly, males are significantly more likely to leave at every stage of their career, being five to seven percent more likely to leave each year. Whites appear modestly more likely to be retained in the long-run (two to four percentage points each year) while Hispanics are
significantly more likely to quit; for instance, they are 13 percentage points less likely to be retained five years, conditional on staying for four years. Having an education major predicts seven percentage points higher retention. Math and science majors predict six percentage points lower retention. Master's degree, social science, language, and business major predict very little about retention.

What is most interesting about these results is that the preference parameter is highly correlated with the value-add parameter for a given teacher characteristic, but is no correlation between the preference parameter and the retention probability of the characteristic, as seen in Figures 3 and 4. However, teacher retention is a critical factor in teacher value-add and the preference parameters appear completely unrelated to the teacher's likelihood of remaining in teaching. Schools appear able to predict inherent quality fairly well, but do not respond to significant differentials in retention probabilities which may be more important for student achievement.

Next I see if treatment is associated with larger student gains. Using students with new teachers near the cutoff, treated teachers induce .1 σ larger growth in mathematics and .07 σ growth in reading. One concern is that students are not randomly assigned to teachers and thus these correlations could reflect differences in students rather than differences in the underlying quality of treated teachers. I find, for instance that treatment teachers at the cutoff are more likely to be sorted students who are black, younger, and whose parents' highest degree was a high school diploma. Conversely, they are less likely to be sorted students who are white or who have parents that graduated from a four-year college. Interestingly, they are no more or less likely to be sorted students who are gifted, learning disabled, or poor. This is evidence that students are non-randomly sorted based on a small number of observables and suggests that the program is benefiting disadvantaged students if treated teachers are better.
In order to address non-random sorting of students to teachers, I pool treatment at the grade-school level following Chetty et al. (2011). That is, I estimate the model as if a teacher had an effect on the whole grade she is assigned to, not just her class. I modify Chetty et al.'s approach by making the treatment variable the share of teacher's in a student's grade who are treated. Using this approach I can estimate the impact of a treatment teacher on student achievement with much weaker, more realistic assumptions on student sorting; estimates from these regressions are even larger. A treated teacher causes .19 \sigma more achievement in math and .25 \sigma in reading, which are imprecisely estimated. These seem implausible large, especially since the reading effect is larger than the math effect. But these results confirm that the positive effects estimated in the class-level regressions were not driven by student sorting.

One question is why, given that a master's degree predicts positive growth, why don't school's prefer applicant's with a master's degree. First, there is mixed evidence regarding the effect of a master's degree. Hanushek summarizes the literature, indicating that having a master's degree "bear[s] no consistent relationship with student achievement," and new teachers with masters degrees cost schools an additional $3,200 on average in the first year alone.

Conclusions

Public school teacher compensation makes up the majority of public education spending and constitutes a key personnel policy (Lankford and Wyckoff, 1997). This paper uses a regression discontinuity approach to estimate the effect of additional teacher compensation on the quality of new hires and student outcomes. I identify the effects of this loan forgiveness program by comparing schools that barely met loan-forgiveness eligibility with those that barely missed. Since sorting is not an issue, the regression discontinuity overcomes several concerns typical of past research.
I find that the compensation boost caused schools to hire teachers that were intrinsically more effective, even in the first year of teaching. The program attracts individuals interested in teaching long-term and exerts a significant effect on retention even independent of selection. Because treatment acts as an exogenous supply shifter, we learn about school preferences in hiring. Schools appear to prefer characteristics that predict intrinsic quality but fail to incorporate into their preferences characteristics that predict retention. They prefer education majors over hard science, humanities, or social sciences, and have only insignificant preferences for applicant's with master's degrees or applicants from more prestigious undergraduate institutions. In response to the subsidy, schools captured some of the surplus by reducing teacher salary by $204.

The results provide evidence that additional compensation does increase teacher supply quality. My theoretical model helps suggest that this quality improvement happens solely through a larger applicant pool, not by better applicants. This suggests, forcefully, I think, that schools can discern teacher quality even beyond characteristics observable to the researcher. For instance, by a simple calculation of the quality improvement brought about by the differential hiring of treatment schools, only a small fraction of the value-add differences of treatment teachers is attributable to these observable variables, suggesting that hiring agents discern other strong predictors of intrinsic quality in the application process. Together the results indicate that higher starting compensation for teachers would increase new-teacher quality under relatively weak assumptions.
CHAPTER 3

Unemployment Insurance Taxes and Labor-Market Recovery:
Evidence from Florida and Missouri

Introduction

During the Great Recession and its aftermath from 2008 to 2012, US unemployment insurance programs paid $520 billion in benefits to some 53 million claimants (Unemployment Insurance Data Summary). These benefits provide significant insurance and simulative benefits, and many economists argue they impose only modest distortionary costs (Gruber, 1997; Card, Chetty, and Weber, 2007; Farber, Rothstein, and Valletta, 2015; Di Maggio and Kermani, 2015).

To finance UI benefits, employers pay a dynamic payroll tax: each year, states calculate a tax rate for each firm so as to reflect the cost of UI benefits incurred by firm layoffs. States vary significantly in UI tax policy, reflecting uncertainty about the consequences of UI taxation. The size of the UI tax penalty represents an important tradeoff for policymakers: larger tax penalties likely discourage layoffs (a “deterrent” effect), but the resulting tax increases likely discourage employment (an “overhang” effect from the previous negative shock) (Hopenhayn and Rogerson, 1993; Anderson, 1993).

Although an enormous literature has considered the influence of UI benefits on labor supply (Barro, 2010; Landais, Michaillat, and Saez, 2010, Krueger and Mueller, 2010; Schmieder and von Wachter, 2012; Kroft et al., 2012; Farber, Rothstein, and Valletta, 2015, Card et al., 2015, Johnston and Mas, 2015), few authors have assessed the consequences of UI financing on labor demand. I contribute to the UI literature
measuring these countervailing forces by bringing to bear detailed micro data and quasi-experimental variation. Previous research focused on the effect of experience rating on temporary layoffs (Feldstein, 1976; Topel, 1983; Anderson, 1993; Card and Levine, 1995).

To estimate the “overhang” effect of higher UI taxes, I exploit a discontinuity in the UI tax schedule in Missouri that permits me to isolate exogenous variation in the tax rate. I show that increases in UI tax rates lead to significant increases in firm exit and declines in hiring and employment, with no effect on the rate of separation or average wages. Firms respond to their new tax as soon as they are informed the November before it is implemented in January of the next year. I present additional evidence of the UI tax effect on employment by leveraging a change in the UI tax formula in Florida. Again, I find that tax hikes significantly increase firm exit and reduce firm employment. On average, a 1-percentage-point rise in the UI tax rate increased firm exit by about 1 percentage point and reduced firm-level employment by 1 percent, consistent with an own-wage elasticity of 3.5. The average labor-demand elasticity estimate is 0.45 (Lichter, Peichl, and Siegloch, 2014; Hamermesh, 1996). I explore a number of explanations for the large effect sizes and demonstrate that these effects are consistent with a model of cash-constrained firms.

To study the “deterrent” effect of tax penalties on layoffs, I compare how firms adjust to industry employment shocks when they face different UI tax penalties arising from the firm’s placement on the tax schedule. Firms at or near the maximum rate face virtually no penalty for layoffs because their tax rate cannot increase, but firms with tax rates below the maximum face a tax penalty roughly proportional to their distance from the maximum tax rate. To account for possible firm differences, I supplement this approach by deploying a regression kink design which exploits the kink in the tax penalty
a firm faces as it approaches the maximum tax rate. Using parametric and non-parametric tests, I find no evidence that tax penalties discourage firms from shrinking in response to negative industry employment shocks. This response is consistent with a model of cash-constrained firms since the layoff allows the firm to survive the period and future tax increases are not deterrent.

I test whether these results hold in national data. An important prediction of the results is that, contrary to the intention of UI taxation, tax penalties for layoffs could increase labor-market volatility over the business cycle because penalties do not discourage layoffs but do increase labor costs during recoveries (Anderson, 1993). I test this hypothesis in a state-level analysis, where I find that experience rating is negatively related to employment and positively related to employment volatility, conditional on state and year fixed effects, and time trends. I show graphically that states with more experience rating recover more haltingly after recessions. Taken together, these results suggest that UI tax penalties do not discourage firms from reducing employment in response to negative employment shocks, but resulting tax increases do significantly decrease employment.

The results are important for a broader literature. First, I show UI taxes, like benefits, track the business cycle, diminishing the overall stabilizing influence of the UI program. Second, the results speak to the optimal UI literature which seeks to balance the costs and benefits of UI provision; while the work-horse model assumes workers pay UI taxes out of their wages in the high state of the world, I show tax hikes do not affect worker wages but reduce the probability that the worker is employed. Third, UI tax

26 Because the UI tax penalty follows layoffs, UI taxes are highest during labor-market recoveries and lowest when labor markets are tight. When the Great Recession began in 2007, the average US firm paid $280 per worker in UI taxes (in 2014 dollars). By 2012, the average firm paid $470, a 70 percent increase. This is likely an understatment of the tax increase since businesses with larger tax increases were more likely exit, thus disappearing from the data.
penalties for layoffs—which disproportionately fall on recovering firms—have increased by 65 percent over the past 25 years, plausibly slowing labor-market recovery in the subsequent years, contributing to the recent rise of jobless recovery in the United States (Berger, 2012).  

The remainder of the article proceeds as follows. Section II describes the US unemployment insurance program and, in particular, the institutional features I leverage for empirical identification. Section III describes the data sources used in my analysis. Section IV presents a conceptual framework of UI taxation. Section V describes the research design. Section VI presents the effects of UI taxation on labor demand and section VII discusses the results.

Unemployment Insurance in the U.S.

In the United States, workers who are laid off qualify for unemployment insurance benefits. The weekly benefit amount is typically a function of the worker’s prior year’s earnings and the worker’s highest quarterly earnings in that year. State UI programs paid $80 billion in benefits in 2009 when unemployment peaked. By 2013, annual benefits paid had fallen to $39 billion, closer to the $30 billion pre-recession expenditure in 2006. From 2008 to 2012, state and federal UI programs paid out over $520 billion in UI benefits.

Benefits paid by state UI programs are financed by an employer payroll tax where the tax rate is calculated for each firm every year. Although the federal government sets requirements on the structure of state programs, states vary in how they determine the tax rate for employers. Forty-seven states subscribe to one of two systems for determining firm tax rates: a benefit-ratio model, or a reserve-ratio model.

---

27 A jobless recovery is one in which labor productivity rebounds but employment does not. The three most recent recessions have been followed by so-called jobless recoveries (Figure 2).
The benefit-ratio model ties a firm's tax rate to the ratio of the benefits drawn by the firm's employees over the past \( n \) years divided by the firm's taxable payroll during those same years. If \( B_t \) represent the cost of benefits drawn by former employees in year \( t \), a firm's benefit ratio is the benefit cost (over the past \( n \) years) divided by the taxable wages over those same years:

\[
BR_t = \frac{\sum_{j=1}^{n} B_{t-j}}{\sum_{j=1}^{n} W_{t-j}}
\]

where \( W_t \) is the firm's taxable wage base in year \( t \). The UI tax rate is a function of the benefit-ratio, subject to a maximum and minimum. Specifically, between the maximum and minimum rates \( \tau_{\text{max}} \) and \( \tau_{\text{min}} \), the firm is assigned a tax rate \( \tau_t = \psi + \lambda \times BR_t \), where the parameters \( \psi \) and \( \lambda \) are chosen by the state.

The principal alternative to a benefit-ratio (BR) model is a reserve-ratio (RR) model, which is broadly similar. In reserve-ratio states, the state Department of Labor keeps an account for each firm; the firm's UI tax payments are credits to the account, and benefits drawn by former employees are debited the firm's account. States calculate a reserve ratio, \( RR_t \), for each firm expressing the account’s reserves divided by the \( n \)-year rolling average of the firm’s taxable wages:

\[
RR_t = \frac{RES_t}{\frac{1}{n} \sum_{i=0}^{n-1} W_{t-i}}
\]

A firm's tax rate declines as its reserve ratio increases, subject to a minimum and maximum rate like the benefit-ratio system. Unlike benefit-ratio formulas which calculate a tax rate from a simple formula, reserve-ratio states use a schedule or step-function of the reserve ratio to determine each firm's tax rate.

Unlike other tax instruments used in the U.S., taxes for unemployment insurance may exaggerate the business cycle. Because of UI tax penalties, taxes rise in response
to economic downturns. Figure 1 plots the real (2014$) average per-employee UI tax bill over time. The continuous line represents the BLS measure of US unemployment (U3). In general, the tax bill closely follows the unemployment rate which rises sharply during recessions, represented by shaded periods. One simple way of measuring whether a tax is stabilizing or pro-cyclic is to test the correlation between average tax rates and concurrent unemployment. Here, unemployment has a strong positive correlation with the tax. When the labor market is robust, taxes fall, increasing labor demand. After the recessions in 1990 and 2001, the average per-worker tax bill increased to approximately $340. In the recent recession, the per-worker tax bill has increased by 70 percent to $470 per worker in inflation-adjusted dollars.

**Experience Rating over Time**

A series of policy changes over the recent decades has increased the tax penalty associated with layoffs, making UI taxes more cyclical due to larger tax swings when unemployment is high. Prior authors use the term *experience rating* to refer to the fraction of UI costs borne by the claimant’s employer. A firm also pay’s a larger penalty for layoffs when benefits are more generous or the unemployed claim benefits for longer periods. This presents two general factors that increase the average tax penalty firms face for laying off workers. First, the penalty is limited by the extent to which a firm’s tax bill can rise. Firms with layoffs will pay for less of their benefit charges if the maximum per employee tax fee is low since it limits the ability of the state to raise revenue from a firm. Second, the generosity of benefits and how long recipients receive unemployment insurance affect the tax bill. When a state provides a more generous weekly benefit to the unemployed, for instance, the firm faces a proportionally larger tax increase for an

---

28 Under a progressive income tax, for instance, the average income tax rate automatically declines during a downturn and increases when wages are high.
otherwise identical layoff. All four of these elements (wage base, maximum rate, average benefit duration, and weekly benefit amount) have changed since 1985 in ways that increase the penalties firms acquire for layoffs.

The Tax Equity and Fiscal Responsibility Act (TEFRA) of 1982 increased the minimum wage base in 1983 and required states to raise their maximum tax rates to 5.4 percent or higher by 1985. Hamermesh (1993) calculates that TEFRA boosted experience rating by 15 percentage points or 30 percent. The 1991 Emergency Unemployment Compensation Act increased firms’ tax liability during recessions by increasing the duration of extended benefits from 13 to 20 weeks, partly financed by state UI programs. In addition, over the past 30 years, the unemployed have drawn more benefits through longer unemployment spells and increased benefit generosity.

Figure 3 shows that in the mid-1980s the average maximum rate rose from 4.2 percent to 6.9 percent. This dramatically reduced the share of firms at the maximum rate, increased experience rating, and raised the potential tax cost of laying off workers. States cannot fully finance benefits for firms already at the maximum rate, so in practice states often raise the minimum rate to balance their UI trust funds, exploiting the broad tax base below the maximum rate. The dramatic reduction in the average minimum rate supports the interpretation that the increase in maximum rates increased experience rating by reducing the fraction of firms at the maximum. From 1990 to 2015, the real weekly benefit amount increased by 19 percent and the average duration of UI receipt increased by 36 percent (Figures 5 and 6),29 resulting in the average layoff costing 65 percent more since 1990 (Figure 8). Because these costs are charged back to the firm,

29 Reliable estimates of experience rating over time are not readily available as they require a precise knowledge of the slope of the “rated” portion of the tax schedule and the percent of firms who are at or near the maximum tax rate in each state in each year. Instead I show that key indicators suggest a marked increase in experience rating.
$300 billion in state benefits also represent $300 billion in labor tax increases principally to firms in distress.

Analysts in the 1970s and 1980s estimated that firms pay about 50 percent of the benefit costs they originate due to a low maximum tax rate. This meant a significant fraction of firms effectively could not be charged for marginal benefits (Topel, 1983). Using administrative records I estimate that firms tend to pay a larger share of benefits charged than they did three decades ago. I estimate that an average firm in Florida pays 87 percent of benefits originated by the firm, and as high as 98 percent for firms who are not new and thus can be charged a variable rate. In Missouri, firms pay an estimated 86 percent of benefits originated by the firm.

Cross-State Comparison of Experience Rating

In what follows, I use an industry case study to demonstrate some of the features of experience rating by comparing how a hard-hit industry was affected in Missouri and Florida. Florida has relatively low experience rating. Its taxable wage base over this period was $7,000 and the maximum tax rate was 5.4 percent, thus the highest possible UI yearly tax fee per worker was $378. By contrast, the taxable wage base in Missouri was $13,000 and the maximum tax rate 7.8 percent; therefore the highest UI yearly tax fee is $1,014—nearly three times greater in Florida.

In Figure 9, I plot the average per-employee tax bill of building contractors hard hit during the recessions in 2001 and 2007. The red line represents the per-employee tax bill of the average Missouri firm in this industry and the blue line represents that of the average Florida firm in this industry. Firms in Missouri faced a significantly larger increase in their tax bill, in part, because the maximum rate and the wage base there were significantly higher.

Conceptual Framework
The framework is designed to study the consequence of unemployment insurance taxes during an initial negative shock and during the recovery. The partial-equilibrium framework is described by a three-period model. Period 0 represents the firm's behavior in a period of normalcy in which product demand is perceived as stable. In period 1 the firm receives a negative demand shock, represented by a reduced price per output, where the timing and depth of the shock is unknown prior to the shock’s arrival. In the final period, price remains low and the UI tax bill from layoffs arrives. This classic profit-maximizing model does not rationalize the behavior of firms I observe in the data, so I point to sensible additions to the model that better describe the firm’s response each period.

Each firm takes its production function as given and uses labor as its only variable input; a firm employing \( N_t \) workers in period \( t \) who receives an average price \( p_t \) earns profits represented by:

\[
\Pi_t = p_t f(N_t) - N_t c - \tau L_{t-1} \frac{N_t}{N_{t-1}}
\]

A few items require some explanation. First, the wage rate has been normalized to one so the price of goods is relative to the wage rate. Second, capital is fixed in the short-run and thus output becomes a function of the variable input, labor. Furthermore, \( c \) represents fixed operating costs including those for the fixed level of capital and \( \tau \) represents the cost of a layoff from the last period divided by the employment level from the previous period, so the firm pays a tax per employee this period. It is assumed that function \( f(N_t) \) has a positive first derivative and negative second derivative so that production increases with employment, but at a declining rate.

Firm employment shrinks by steady attrition at a rate \( 1 - \delta \); the firm can increase its employment by hiring at a rate higher than \( N_t (1 - \delta) \), or the firm can decrease
employment gradually by not hiring. Firms can reduce their employment level more quickly by engaging in layoffs which increase future taxes:

\[ N_t = N_{t-1} \delta - L_t + H_t \]

The firm begins at some initial employment \( N_0 \) in the pre period, which can be regarded as the optimal employment for current price \( p \).

In period 1, the firm receives a negative shock, represented by an unexpected decline in price from period 0. Firms with a sufficiently low new price will engage in layoffs, shedding workers to increase the marginal productivity of labor. At the same time, firms are discouraged from laying off workers because they may be required to pay the future cost of a UI benefit spell for the former employee. Thus a firm’s decision in period one depends also on the expected profits of period two:

\[ \Pi_1 = p'f(N_1) - N_1 - c \]
\[ \Pi_2 = p'f(N_2) - N_2 - c - \frac{\tau L_1}{N_1}N_2 \]

Firms have no incentive to alter their employment levels in period zero since period zero is assumed to be a long-run best response to perceived stable prices. Accordingly, firms do not engage in layoffs and hire at a rate \( N_0(1 - \delta) \) in period zero to maintain their employment level.

The firm maximizes profits by increasing layoffs \( L_1 \) if \( p' \) is sufficiently low. Likewise, hires \( H_1 \) reduces to zero if \( p' \) is sufficiently low. This will occur even for smaller negative shocks than those inducing layoffs because of the tax penalties associated with layoffs. If the tax penalty were simply a bill for benefits discharged the previous year, then \( H_2 \) and \( L_2 \) would not be influenced by the tax since the optimal level of employment
is not affected by the level decrease in profits.\textsuperscript{30} In practice, the firm can reduce its tax by further reducing employment and so the tax penalty also reduces hiring in period two.

Thus the basic model predicts firms reduce their layoffs in response to expected tax penalties and the arrival of tax penalties do not affect hiring in the recovery period.

These predictions alter significantly when cash-constraints are introduced to the model; this feature is likely to be a significant feature of firms suffering negative shocks and performing layoffs. A cash-constraint is a requirement that a firm ceases to exist if it suffers negative profits in a period. In a model without cash-constraints, a firm continues to employ workers so long as positive employment maximizes profits, even if those profits are negative. If firms are cash-constrained, as they likely would be during and after negative shocks large enough to induce layoffs, then the consequences of UI taxation are somewhat unexpected.

If firms are cash-constrained they must make layoffs to survive the recession period, and the firm cannot be deterred from layoffs by costs incurred in the next period. Moreover, if firms are cash-constrained, the increased tax bill in the second period pushes firms against their budget constraint, forcing them to reduce their variable costs, driving the firm to reduce employment in the period the tax bill arrives. In a model without cash constraints, there is no firm exit. With cash constraints, some firms will exit during the recessionary period while others will exit when the tax bill arrives because no available choices deliver positive profits if the price fall is significantly large.

Finally we could consider the tax penalty associated with layoffs as a random variable, \( \sim B(1, p)N(\mu, \sigma^2) \). The firm does not know the benefit cost of layoffs from last

\textsuperscript{30} A firm can avoid some of the tax penalty by shrinking their employment. To simplify the exposition, I force the firm to confront the cost in period 2 rather than complicating the issue with additional periods and notation. This model demonstrates the influence of cash constraints on firm layoff and hiring decisions.
period, \( \tau \), which can be represented as the product distribution of a Bernoulli and Normal distribution. There is some probability \( p \) that a given layoff will not claim benefits. It is estimated that only about half of eligible unemployed persons claim benefits. These workers may not claim because they immediately find another job or have other sources of supplemental income (Anderson and Meyer, 1997b). Conditional on claiming, I assume for simplicity beneficiary costs are approximately normally distributed with a cap on the maximum arising from benefit exhaustion. Therefore, a firm may lay off workers and not know the size of the tax penalty or whether they will receive a tax penalty at all. Because of this feature, firms may be less dissuaded by tax penalties and may also respond by reducing employment once the tax bill is revealed in the recovery period. The tax penalty-as-random-variable and the cash-constraint feature of the model can explain the large effects of tax increases on a firm’s hiring.

UI tax penalties function as a linear adjustment cost to layoffs (Anderson, 1993). As the penalty increases, layoffs become more costly and in expectation, the cost of hiring is greater, reducing the volatility of a firm’s employment. Firms with a zero-cost of layoffs at the maximum tax rate should be more responsive to employment shocks than firms with larger potential tax penalties well below the maximum rate.

Once a firm lays off a worker, their UI payroll tax increases. In equilibrium, payroll tax increases reduce wages with little impact on employment to the extent labor supply is inelastic (Hamermesh, 1996; Gruber, 1997; Chetty et al., 2011). Unlike other payroll taxes, UI taxes can vary annually. Because nominal wages are rigid and UI tax changes may be largely unexpected (Kaur, 2014), UI tax increases do not result in lower wages but lower firm-level employment. Tax increases represent real-wage increases in the short run. The exogenous increase in input prices affects firm profits, inducing some firms to miss their target profit and exit (Hamermesh, 1993). If firms have access to
credit markets, temporary increases in input price should have a relatively small effect on firm exit. If firms are cash-constrained and lack access to credit, however, the effects of unexpected cost increases can lead to more firm exit (Hamermesh, 1996).

Data

I use detailed administrative unemployment insurance and firm UI tax records from the Florida Department of Economic Opportunity and the Missouri Department of Labor and Industrial Relations which administer the unemployment insurance programs in Florida and Missouri, respectively.

In Florida, the data cover the universe of firms participating in UI from 2003 to 2012, with quarterly records for 903,000 unique firms. This information includes each firm’s industry, employment, wages paid, county, entity type, tax parameters, tax rate, and the benefit ratio used to calculate the tax rate for each firm. The average firm in the data has 18 employees and faces a tax rate of 1.7 percent. Firms in the data pay average yearly earnings of $38,800 per employee in 2014 dollars. The Florida data also include a UI claim file which contains an observation for each claim, a unique employee ID, employer ID, employee wages, hire date, separation date, and reason for separating from 1990 to 2013. The employer ID in this file and the firm file are not the same and so I am unable to merge the two data sets.

The Missouri data include an observation for each firm in each quarter including the firm’s account balance, taxable payroll, reserve ratio, tax rate, and six-digit NAICS industry code. Missouri also provided an employee file which indicates the wages and the employers of each worker in each quarter. Using these data, I calculate the employment of each firm and count the number of new hires and separations for each firm in each quarter.
Nationally, an employer must enroll in UI taxes, if it has a payroll of $1,500 or more in a calendar year or has at least one employee working at least a portion of one day during any 20 weeks of a calendar. The coverage includes businesses, nonprofit organizations, state or local government employers, and Indian tribal units (Florida, 2012). In practice this means all lawful employers are represented in the dataset.

I also use a number of datasets to study unemployment insurance taxation across states. The Bureau of Labor Statistics provides the Unemployment Insurance Data Summary (UIDS) which includes an observation for each state in each quarter describing the state’s benefits paid, number of UI claims, average duration of UI benefits, exhaustion rate, average weekly benefit, average tax rate, taxable wages, taxable wage base, fund balance, total loans, and unemployment rate from 1987 through 2013. I supplement this with information from the Commerce Clearinghouse UI Data which includes records of the maximum tax rates and taxable wage base by state from 1976-2004. I also use local unemployment rates data (U3) from the Bureau of Labor Statistics for each state. Finally, I use County Business Patterns Data which has county-level employment figures for each industry, to construct measures of industry employment shocks using all states but Missouri and Florida, similar to Bartik (1991).

Variables

A number of variables I impute from the wage and employer file require some discussion. The variable *firm size* simply sums the number of wage earners reported to the state Department of Labor each quarter for a particular firm. I infer a *new hire* if a worker starts working for a firm he had not previously worked for and remains there for two or more consecutive quarters. A large number of workers are employed at a given firm for only one quarter which are not counted as new hires. Similarly, a *separation* is inferred when a worker no longer works for an employer who employed him for two or
more consecutive quarters. One reason for not counting new hires and separations for those who are employed at a given firm for only one quarter is that these would count as both a new hire and a separation, dulling the measure’s meaning. Firm exit is inferred as the last quarter the firm is in the dataset, unless that quarter is also the last quarter of the dataset.

I calculate a number of variables to describe wages paid by each firm. The most basic is the average wage which represents the mean earnings paid to a firm’s workers each year. The average wage is affected by two factors—the skill of labor and the firm’s wage premium (Abowd and Kramarz, 1999). That is, a firm may pay more for two primary reasons: the firm employs higher skilled workers (worker type), or the firm pays more for a given level of skill (firm premium). I estimate each firm’s worker type and firm premium using a high-dimensional fixed-effect regression. I regress real wages on worker fixed effects and year-by-firm fixed effects, using a high-dimensional fixed-effect package in Stata. The year-by-firm fixed effect measures what a given firm pays above what other firms would pay the same worker which can only be estimated if turnover is sufficiently high at a given firm to separately identify the firm premium from the worker type. The yearly firm average of captured worker fixed effects reflects the average worker type at a given firm. As a second measure of measure worker type, I also capture each employee’s wage at his former employer.

Empirical Approach

The empirical aim of the paper is to estimate two primary effects of UI taxes. The second is the deterrent effect of tax penalties in discouraging firm downsizing—the consequence of UI firing costs. The first is the overhang effect of higher taxes on employment and firm exit once tax penalties are in place—in essence the impact of a sudden increase in the payroll tax rate.
To identify the “overhang” effect of increased UI taxes on labor demand, I leverage a discontinuity in Missouri’s tax schedule to implement a sharp regression discontinuity design (RDD). I also exploit a change in Florida’s minimum tax rate deploying a first-differences design (FD).

Second I investigate the role of tax penalties in deterring firms from laying off workers. I exploit the fact that penalties vary based on the firm distribution to the maximum to estimate whether firms more exposed to UI tax penalties react differentially to industry shocks by using parametric and non-parametric approaches including a regression kink design (RKD). I test whether firms exposed to larger tax penalties are less responsive to negative shocks as predicted by past literature (Anderson, 1993).

I complement these strategies with ancillary identification strategies that prove to be less robust but support the results of the primary identification strategies. I implement a regression kink design leveraging the kinks in the tax formulae for causal identification of the overhang effect. Moreover, I exploit the tax changes that occur when new firms become experience rated for the first time, which allows me to precisely estimate the overhang effect throughout the business cycle. Regardless of the identification strategy, the results are remarkably similar.

**Deterrence Effect: The Effect of Tax Penalties in Discouraging Layoffs**

The policy intention of UI tax penalties is to align firm incentives with the social cost of unemployment benefits, discouraging layoffs and encouraging lower firm employment volatility. UI tax penalties are a form of experience rating, the insurance practice of calculating premiums to reflect cost. A number of studies have found associations between experience rating and lower employment fluctuations, usually focusing on temporary layoffs (Topel, 1983; Anderson, 1993; Card and Levine, 1994; Ratner, 2013). I implement a parametric and a non-parametric test of this prediction,
which is the first time detailed administrative records have been combined with quasi-experimental variation, exploiting the kink in the tax rate that arises from the maximum allowable rate.

I estimate a triple difference (DDD) model in which a firm's employment change is a function of an industry-wide employment shock and an interaction of the shock with a measure of the firm's marginal tax penalty for laying off workers. The maximum tax rate shields firms at the maximum from tax penalties, while firms just below the maximum face limited penalties, and firms well below the maximum rate face large penalties for a given layoff:

\[
\ln(E_{fit}) - \ln(E_{fit-1}) = \beta y_{it} + \delta(y_{it} \times MTC_{fit}) + \alpha_t + \phi_f + \epsilon_{fit}
\]

Here, \(E_{fit}\) represents the employment of firm \(f\) in industry \(i\) at time \(t\), calculated from administrative Florida records. To represent industry shocks, I compute \(y_{it}\) from the County Business Patterns data by calculating the employment in each industry in all other states and calculating the log-employment change from the previous year (Bartik, 1991). Therefore, \(\beta\) represents the percent change in a firm's employment resulting from an exogenous industry shock, a unit representing a 1 percent decline in industry employment. The interaction between this shock and the tax penalty captures how firms more exposed to penalties respond to industry shocks differentially, captured by \(\delta\). I demonstrate robustness by including firm fixed effects which represent firm-specific trends in the FD framework. I also demonstrate that the treatment of interest is not correlated with differential firm trends using a distributed lag model.

I measure the firm's marginal tax cost \((MTC_{fit})\) following prior work. This measure indicates the present value of the taxes the firm expects to pay per dollar of benefits received by former employees (Topel, 1983; Anderson, 1993) who calculate MTC as:
\[ \Delta PV \text{ tax} = \frac{(\rho \gamma)^2 \eta}{i + (\rho \gamma)^2 \eta} \]

In this calculation, \( \rho \) represents 1 plus the rate of growth in the firm’s employment, and \( \gamma \) represents 1 plus the rate of growth in average (per employee) taxable wages at the firm; \( i \) indicates the interest rate, and \( \eta \) represents the slope of the tax schedule. In Florida, the tax schedule sometimes changes so I experiment with several computations of \( \eta \) including using the contemporaneous slope, the average slope over the past two years, or the average slope over the decade following Ratner (2013). The results are robust to any measure used.

One feature of experience rating that has been ignored in previous research due to data limitations is how experience rating partially depends on the distance a firm is from the maximum rate. A firm \( \varepsilon \) below the maximum rate does not suffer a significant penalty from marginal layoffs, so the slope of the tax rate locally is not a good measure of experience rating as the firm approaches the maximum rate. At the other end, a firm with a minimum rating faces a large potential tax penalty. To capture this variation in experience rating, I convert Topel’s measure of experience rating into the per-employee penalty of a 1 percent layoff, a measure developed by the Bureau of Labor Statistics (BLS) to describe the tax penalty. I test to see if my results come from this computation of the marginal tax cost. The results are consistent with and without the adjustment, described as follows:

\[ \Delta PV \text{ taxes} = \left( 1 + \frac{(\rho \gamma)^2 \eta}{i + (\rho \gamma)^2 \eta} \right) \frac{c}{300} \]

\(^{31}\) C is the cost of the average layoff, divided by three because cost is divided over three years. E is the employment at the firm. It’s divided by 100 so that the variable measures the tax increase associated with laying off 1% of the firm, a standard measure used by DOL.
I also implement a non-parametric test of experience-rating’s effect on adjustment by separating the firms near and below the maximum rate into bins based on their reserve ratio. Within each bin, I estimate the relationship between changes in log firm employment and negative industry shocks. In theory, if the firms are not cash constrained, the effect of industry shocks should be smaller for firms with benefit ratios below the maximum than above because they face a tax penalty for additional layoffs. I plot the coefficient on the industry shocks against the marginal tax cost to visualize the relationship between the tax incentive and the response to industry shocks.

I document that there are not differential trends using a distribution lag model and I control for firm trends. But a more convincing avenue to exploit the kink I in the penalties firms face using a modified regression kink design. The regression kink approach exploits the fact that the potential tax penalty is kinked due to the maximum tax rate. As the firm approaches the maximum tax rate, the cost of a layoff decreases since the firm’s penalty is limited by the maximum rate. Firm characteristics are smooth across the kink so the expectation would be the firm’s response to industry shocks evolves smoothly as a firm approaches the maximum rate. If a firm’s response to shocks is kinked at the maximum tax rate, I infer that the tax penalty affected the firm’s response to industry shocks. I perform this test parametrically and non-parametrically. The parametric specification takes the form:

\[
\frac{\log(E_{fit})-\log(E_{fit-1})}{K^T} = \sum_{p=1}^{n} \{ \beta_p (y_{it} \times (w - k)^p) + \delta_p (y_{it} \times K \times (w - k)^p) \} + \beta y_{it} + \alpha_t + \epsilon_{it},
\]

where \(|w - k| < h\)

The outcome variable is the change in log employment, scaled by the size of the tax penalty kink, \(K^T\). Here, \(w\) is the assignment variable which is the firm’s distance from the benefit ratio at which the maximum tax rate becomes binding; \(K = 1(w > k)\) is an
indicator for being to the right of the kink point, $h$ is the bandwidth, and the slope change is captured by the parameter $\delta_1$. I provide estimates for various bandwidths and include quadratic terms for wider bandwidths. A typical regression kink estimation lacks the $\gamma_{it}$ variable and $\gamma_{it}$ interactions. The key is that normally the treatment kink directly affects the outcome variable. In this analysis, the kink of interest is the firm’s response to shocks. This approach can be considered a regression-kink difference-in-difference where a kinked variable may moderate the effect of a shock.

I present this evidence non-parametrically by estimating the relationship between firm adjustment and industry shocks within bins around the kink and plot the estimated firm responses to show visually what the regression kink estimates.

**Overhang Effect: Results of Tax Penalty Incentive**

The results of these analyses appear in Tables 1–3 and Figures 11–12 which consistently fail to demonstrate that firms are deterred from layoffs and tax penalties.

Recall from the conceptual framework that UI tax penalties represent an adjustment cost of reducing a firm’s workforce. Because UI taxes impose this cost on layoff adjustment, theory predicts that firms should retain more workers when they face negative shocks and hire fewer workers when they face positive shocks. Based on this, we expect that $\beta$ would be positive, representing the positive effect of a shock in industry employment on the firm’s employment, precisely, the effect of a 1 percent industry shock. The coefficient $\delta$ represents the effect of the tax penalty in attenuating the effect of the industry employment shock, and theory predicts $\delta < 0$. We expect the effect of the shock to be significant and positive, and we expect the tax penalty to dampen this effect. Represent by a negative coefficient on the tax-penalty-shock interaction. To complement this, I estimate distributed lag models to explore the common trends...
assumption. I find “treated” firms have common trends before the recovery but not during the recovery in Florida. Therefore, I limit the analysis to those years where the common trends assumption holds.

In Table 1 we see that a 1 percent negative industry shock reduces a firm’s employment by about 0.9 percent which is robust to a number of controls and highly significant. The coefficient on the tax-penalty-shock interaction is small and not negative, contrary to the prediction. This point estimate is remarkably consistent, regardless of how the tax penalty is calculated or what controls are included. Column 1 represents the calculated tax penalty using the concurrent tax slope while in column 2 I use the average tax slope over the entire data period. In both calculations, the results are nearly identical. Column 3 includes a control for the firm’s benefit ratio, and in column 4 I add a measure of the firm’s age. A 1 percent shock in industry wide employment is associated with a 0.7 percent change at an individual firm, and the tax penalty has a dampening effect on the impact of industry shocks.

This result might be an artifact of unobservable firm differences along the benefit ratio, making firms that have lower tax penalties less responsive to firm shocks for some other reason. To explore this possibility, I perform the same regressions but include firm fixed effects which capture individual firm trends. Each estimated firm fixed effect represents the average trend of that firm, accommodating possible non-parallel trends between firms. While the effect of an industry shock is attenuated, presumably due to serial correlation in industry shocks and firm adjustment, the coefficient on the tax-penalty-shock interaction is positive and statistically insignificant.

\[ ^{32} \] In addition, I experiment with a number of moving average calculations of the tax slope and the results remain remarkably consistent.
To perform a non-parametric version of this test, I create bins along the benefit ratio and estimate the relationship between industry shocks and firm employment adjustment within each bin. I plot these coefficients against the marginal tax penalties in Figure 11. The expectation is that firms are more responsive to industry shocks near or above the maximum tax rate where the tax penalties for marginal layoffs are small or non-existent. Instead, firm response does not systematically vary with the tax penalty.

It is possible that unobserved differences could bias this estimation if those unobserved differences vary within firm over time. To address this issue, I exploit the fact that a firm’s potential tax penalty is kinked at the maximum tax rate. That is, the tax penalty declines linearly as a firm approaches the maximum tax rate until it reaches the maximum point and the tax penalty stops decreasing. This generates a kink in the tax penalty that allows for careful quasi-experimental examination. If firm responses to industry shocks are kinked at that point, it provides strong evidence that the tax penalty influences the firm’s decision to reduce employment in response to industry shocks. Hence the change in the slope of the response measures the firm’s sensitivity to the tax rate.

The regression kink estimates are fairly noisy, but the point estimates are economically small and statistically insignificant at a range of bandwidths. To show this non-parametrically, I estimate the relationship between firm employment adjustment and industry shocks within bins along the running variable around the kink point. I plot these coefficients around the threshold as shown in Figure 12. The figure confirms the parametric estimation; namely, there is no discernable kink at the threshold though the standard errors are too large to rule out a meaningful effect.

Thus I am unable to provide any evidence that UI tax penalties discourage firms from downsizing in response to negative shocks. As we explored in the conceptual
framework, one potential explanation of this finding is firms considering layoffs are cash-constrained, implying future tax penalties are not deterring since these employers layoff workers in order to survive.

The Effect of the Tax

Regression Discontinuity Design

To identify the causal effect of UI payroll taxes, I leverage a relatively large discontinuity in the firm’s UI tax rate based on its reserve ratio. Recall that Missouri generates each firm’s reserve ratio based on the firm’s UI account balance and the state uses a tax schedule to determine each firm’s tax rate every year. The Missouri tax schedule includes a relatively large, 1.2-percentage-point discontinuity in the tax rate, which increases the per-employee tax by $152 annually. I use this discontinuity to compare firms with similar UI histories who experienced different UI tax rates, overcoming the omitted variables problem which challenged previous work.

I model the outcome variable $Y_{it}$ as a continuous function of the running variable, the firm’s reserve ratio, and estimate the outcome discontinuity that occurs at the threshold:

$Y_{it} = \beta T_{it} + f_l(x_{it} - x') + f_r(x_{it} - x') + \alpha_i + u_{it}$,

where $x_{it}$ is the reserve ratio of firm $i$ in year $t$, $x'$ is the value of the running variable at the tax discontinuity, and $T_{it}$ equals one if firm $i$ is on the left of the discontinuity in year $t$ causing a higher tax rate. Here, $f_l(x_{it} - x')$ is a continuous function of the running variable to the left of the threshold which captures the continuous relationship between the firm reserve ratio and the outcome of interest. Likewise, I allow for a different relationship between the outcome and running variable to the right of the threshold by including polynomial $f_r(x_{it} - x')$. When the residuals of this approach are serially
correlated, I take it as a sign of misspecification and add additional polynomial terms. Regardless the estimates themselves are broadly robust to lower-order polynomials. I include firm fixed effects which allows me to reduce the residual variation considerably and isolate within-firm variation in the tax rate (Greene, 1987). An additional virtue of the firm fixed effects in this context is the resulting estimates reveal the immediate effect of tax hikes rather than the effect of a stable payroll tax. One concern is whether firm fixed effects offers the same experimental interpretation as a traditional RDD because the fixed effects force the regression to use within-firm variation. While it is standard practice that controlling for exogenous covariates does not affect the causal interpretation of a traditional experiment, this issue requires more exploration.

To demonstrate that the fixed effects and RDD design work together to produce the parameter of interest, I use a Monte Carlo simulation. Recall that the long-run effects of a payroll tax increase are significantly different than those in the short-run; wages are rigid in the short run and so a payroll tax increase reduces employment rather than wages. In the longer-run a firm will adjust to a higher payroll tax by reducing wages such that the tax has little to no effect on employment (Gruber, 1997). The data-generating process I study creates heterogeneous firms that move randomly around the threshold. If they fall to the right of the threshold, their employment falls by 1, the true beta I intend to estimate. If the firm remains treated, the effect on employment attenuates and the running variable has a positive effect on employment. In Table 4b, the results of this simulation are demonstrated. A simple regression without fixed-effects or RDD controls estimates the wrong sign because of the differences of firms along the benefit ratio. Firm fixed effects get closer to the true parameter because they use within-firm variation, but they do not control for within firm variation in the benefit ratio that also affects employment. The RDD alone also systematically underestimates the effect of the
threshold because it estimates the average effect at the threshold, not the immediate effect of interest. When these two control designs are combined—the regression discontinuity with firm fixed effects—the effect is estimated accurately. The running-variable controls account for within firm differences in time and force the regression to estimate the effect at the threshold while the firm fixed effects allow the estimation of the true short-run effect by exploiting within firm variation.

Another important advantage of including firm fixed effects is that it can account for considerable noise arising from firm heterogeneity. To demonstrate this, I run the same Monte Carlo simulation but now I introduce larger firm heterogeneity in the firm's initial size. With considerable firm heterogeneity, the RDD without firm fixed effects estimates the effect imprecisely and maintains the bias described in the first exercise (Table 4c). When firm fixed effects are included, the true beta is precisely estimated. In short, the Monte Carlo simulations demonstrate the purpose of the firm fixed effects: they provide precise estimates and identify the short-term effect of tax increases.

I focus on firms in Missouri whose average per-worker wages were less than $50,000 over the data period in real terms ($2014). This is intended to focus on firms for which the tax represents a meaningful fraction of employment costs. I also consider a range of alternative bandwidths to assess robustness. The standard errors are clustered at the firm level and I have collapsed the data at the firm-year level to use yearly data rather than quarterly in most analyses.

Regression Discontinuity Design Results

Two primary threats would undermine the validity of the RD design. The first is if firms can precisely manipulate their position on the UI tax schedule, creating selection bias. The second is if other determinants are also discontinuous at the tax threshold. I begin by testing for manipulation of the running variable, which would occur if firms could
strategically manipulate their reserve ratio around the tax threshold. If strategic manipulation had occurred, we would see an excess density of firms on the favorable side of the threshold and a deficit density on the less favorable side; this intuition is formalized by a McCrary test (2008). Figure 14 is a histogram representing the distribution of the running variable (the reserve ratio) around the threshold. The eye suggests and the McCrary test confirms that there is no statistically discernable manipulation around the threshold (Figure 15).

A second threat to identification is if some other determinant of the outcome is discontinuous at the threshold. There are numerous predetermined variables with which I construct an index of predicted outcomes using all fixed covariates in the data, similar to Card et al. (2015). To construct the index, I regress firm size, hiring, and firm exit respectively on two-digit industry indicators, a measure of firm age, and year. Figure 16 plots the mean values of the covariate indices over the running variable. Using my RDD specification, the predicted-value discontinuity at the cutoff is small and statistically insignificant. The lack of evidence of sorting and differences in predetermined characteristics around the threshold supports the assertion observable factors are balanced around the threshold.

I conducted informal interviews with employees of the state department of labor to determine if any other policies turn on at the cutoff of interest. Each of five employees indicated there were no other policies that were affected in any way by the reserve ratio and other state departments had no access to the reserve ratio measure preventing them from applying policy dependent on the reserve ratio.

I implement a number of placebo regressions to probe the validity of the design. First I estimate the model using outcome variables that preceded treatment and find no significant effects (Figure 17). I also estimate the model for firms who were excluded
from the sample for paying more in average wages. As expected, the estimated effects are smaller in magnitude and statistically insignificant.

Finally, I test whether the firm fixed effects may produce bias. Because controlling for firm fixed effects requires within-firm changes in tax rate, it may be that the RDD compares dissimilar firms at the threshold. In order to probe this concern, I implement a robustness test similar to Chang, Hong, and Liskovich (2014), limiting the analysis to firms who originated on one side of the tax discontinuity. Using this method, the RDD estimator compares only firms who originated on one side and continuously migrated toward the threshold with some firms quasi-randomly crossing the cutoff. These intuitive estimates match the baseline regressions in magnitudes and significance (Table 4). This supports the assertion that the RDD successfully compares like firms at the cutoff.

The regression discontinuity estimates the effect of a $152 per-employee tax increase on firm outcomes. The tax appears to have an economically significant effect on firm size, shrinking the average firm by 0.5 employees where the average firm in the data consists of 21.5 employees, or about 2 percent, although these estimates are insignificant at conventional levels (Table 5). The tax increase does not appear to affect the rate of worker separation, but does reduce the rate of quarterly hiring by 0.3 hires per quarter, highly significant at conventional levels (Table 5). The average firm hires 1.8 employees each quarter. The RDD estimates also suggest that the higher tax rate increases firm exit by 1 percentage point, up from an average rate of 10 percent annual exit (Table 5), about equivalent to the effect of a 4-6 point negative industry shock. These estimates are larger than predicted by average labor demand elasticities.

To complement this analysis, I consider a natural experiment in Florida where UI taxes increased for one group of firms, but left another group unaffected. I vary the
bandwidth and show that the results are robust to a variety of bandwidths in Figures 17e and 17f. In figure 17d I demonstrate the effect of the tax increase in event time. In the quarter before the tax increase is announced to the firm, the firm does not alter its hiring decision. Firms receive their tax letter in November. In the quarter in which the firm knows its new tax but does not yet face this new tax, it reduces hiring. When the tax is implemented, the firm’s hiring falls significantly. A significant literature finds minimal effects of payroll taxes. Since firms begin to respond when informed this harmonizes in part how I can document significant consequences of payroll taxes while well anticipated payroll tax changes show little effect when the tax is implemented.

*First-Differences Design*

In the aftermath of the 2008 recession, Florida’s UI trust fund was depleted. The state fund represented about 98 percent of wages 2007 falling to 6 percent by 2009. In 2007, Florida’s reserves were average among the states but fell to 49th by 2012. This dramatic decline took place, in part, because Florida has a small wage base and a low maximum tax rate, so a large fraction of firms was not charged for marginal layoffs. To shore up the UI trust fund, Florida raised the minimum tax rate substantially from 0.1% in 2009 to 1.5% by 2012, to leverage the large tax base of firms below the maximum rate with no change in the maximum rate. I deploy a first-differences (FD) design comparing firms consistently at the minimum rate to those who were consistently at the maximum rate. Figure 18 shows the tax changes in the minimum tax rate over time in Florida.

I identify firms who are consistently at the maximum or minimum rate and use them to estimate the model:

\[
\Delta Y_{it} = \beta \Delta \tau_{it} + \delta_t + \alpha_i + \mu_i
\]

33Florida also covered its UI operating expenses by borrowing from the federal government $2.2 billion or $330 for each employee.
Where $Y_{it}$ represents the outcome variable (e.g. firm employment and firm exit); $\tau_{it}$ represents the tax per employee in $100 (2014$); $\delta_t$ represent year fixed effects which captures the average change in $Y_i$ each year and $\alpha_i$ represents firm fixed effects which capture each firm’s linear trend. These firm trends, if unaccounted for, could bias estimates if trends correlate with the tax changes. The standard errors are clustered at the firm level.

One potential concern is that firms at the minimum and maximum may be very different and thus subject to different shocks which may be confounded with tax changes. As a response to this concern, I use the simulated tax increase from the formula change as an instrument on the firm’s tax change near the maximum to compare similar firms, some of whom receive an exogenous tax shock because of policy changes.

First-Difference Results

Florida dramatically raised its minimum UI tax rate after the recession to shore up the state’s UI trust fund. This policy change affected firms at the minimum rate, but not those at the maximum rate. Implementing a first-differences design comparing firms at the minimum who underwent significant tax increases to those at the maximum, I find that a $100 increase in per-employee taxes reduce employment by 0.24. The tax increase of $100 is associated with a 0.9 percentage point (9 percent) increase in the firm exit rate, on a base of 4.9 percent. The identification assumption is that firms undergoing a tax change would have trended parallel to the firms that did not experience a tax increase. To evaluate this assumption, I include firm-specific time trends, and all results remain robust to this inclusion, consistent with the identification assumption. The employment effect increases slightly to a 0.28 employee reduction. The firm exit
increase is 0.6 percent with firm trends. Another robustness test implements a placebo where I regress current outcomes on future tax rates. The coefficients in this regression are small and insignificant, demonstrating the effects are not driven by selection since firm size is not significantly correlated with future taxes. There can be no effect on firm exit rates because the exit in years before a firm exits must be zero.

*Regression Kink Design*

The maximum tax rate in Florida creates a kink in the tax rate as a function of the benefit ratio. The regression kink design (RKD) relates a kink in the outcome variable with a kink in a policy variable, the tax rate. Unbiased identification relies on two assumptions. First, the assignment variable must have a smooth marginal effect on the outcome of interest. Second, the density of the unobserved determinants of the outcome variable must evolve smoothly with the assignment variable at the kink point. If these conditions are not met, the kink in the outcome variable is confounded with other factors and the causal estimation is statistically biased. Although firms can know the placement of the tax kink, it is virtually impossible for firms to precisely manipulate their tax rate because it depends on the value of benefits drawn by laid-off workers. An employer would find it impossible to precisely control the duration of each former worker’s benefit receipt, precluding precise manipulation around the kink. Moreover, I have been unable to document any evidence that firms act strategically to game the unemployment insurance formula.

The RKD estimate measures the slope change in the outcome variable at the treatment kink and scales the slope change by the slope change in the treatment

---

34 Because exit is an absorbing condition, firm fixed effects will always attenuate the estimate towards zero. Since the trend will be positive for exiting firms and zero for _____ surviving firms.
variable. The numerator can be estimated by implementing a parametric polynomial model:

\[ Y_{it} = \alpha_0 + \sum_{p=1}^{n} \beta_p (w - k)^p + \delta_p (w - k)^p \times K + \mu_{it} + \varepsilon_{it}, \text{ where } |w - k| < h \]

Here, \( w \) is the assignment variable, \( K = 1(w > k) \) is an indicator for being above the kink point, \( h \) is the bandwidth, and the slope change is captured by the parameter \( \delta_1 \).

Estimates should be interpreted as the average treatment effect for firms near the kink. For estimation, I divide the outcome variable by the policy kink change in $100s of dollars in 2014 dollars so that the estimates reflect the average effect of a $100 increase in per-employee taxes. I implement a separate regression for each year with varying bandwidth and \( p = 1 \) and \( p = 2 \). All regressions are estimated with standard errors clustered at the individual firm level.

The estimates are only precise for employment, but yield similar point estimates. Like the RDD and the FD, the RKD estimand implies that a $100 increase in the tax rate decreases employment by 0.20 employees or about 1 percent (Figures 9 and 10).

**Experience Rate Introduction (ERI) Estimation**

In Florida, new firms become experience-rated the January after a new firm’s first 10 quarters, creating variation in tax rates for approximately 70,000 firms each year. This variation can be used for identification similar to Anderson and Meyer (2000) who use the introduction of experience rating in Washington State for estimation. In this strategy, I exploit the tax change that occurs when firms become experience rated to estimate the influence of the tax. This estimation is imperfect, in part because firms can respond to the tax rate in expectation so the strategy may provide estimates biased toward zero. The underlying identification assumption is that the tax change induced by experience rating’s introduction is not correlated with the firm’s trend in employment. The principal
value of this strategy is that it allows for the estimation of the effect of the tax over the business cycle.

I estimate the employment effect over the business cycle using the ERI estimator, associating changes in employment with tax changes arising from the introduction of experience rating each year for new firms. Like the RDD, FD, and RKD estimates, this estimator implies that a $100 increase in per-employee UI fees reduces employment by 0.30 employees. When estimated by year, I find that the effect of the tax is significantly larger in 2008 and 2009, reaching 0.5 in 2009. Afterward, the effect declines back to the pre-recession effect level but then declines again (Figure 17). This suggests that firms may be especially sensitive to the tax during the worst of a recession.

Discussion

In this paper I measure two counterailing influences of UI taxation: the intended consequence of deterrence and the unintended consequence of tax overhang. The intended effect of the UI tax in discouraging layoffs is not apparent in the data. The UI tax increase resulting from the layoffs appears to reduce hiring and increase firm exit significantly.

The results imply that a 1 point increase in tax rates decrease a firm’s employment by about 1 percent. Taking the average wage as the base, the implied elasticity is 3.5. Comparing this to the results of meta-analysis regressing own-price labor demand elasticity on study characteristics (whether it used administrative data, the time period, whether wages were instrumented, whether panel FE were employed, etc.), I find the average labor demand elasticity in contexts similar to mine is 0.9 (Lichter, Peichl and Siegloch, 2014). This is the average short-run elasticity estimate from reduced-form papers focusing on low-skill labor demand, using instrumented wages,
administrative data, and firm fixed effects using data from the 2000s. For instance, measure elasticity for more recent periods have been higher, presumably because of falling costs associated with mechanization and international corporations.

To explain why this paper finds very little role for UI tax penalties in discouraging layoffs, while the previous literature finds a rather large effect (Topel, 1983; Anderson, 1993; Card and Levine, 1994), I offer several thoughts. First, the small estimated response is consistent with a plausible model of cash constrained firms. I use new administrative data and within state quasi-experimental methods while previous studies have used cross-state variation in tax penalties for identification. One concern with cross-state variation is the possibility intentional location selection could drive firms with more unemployment risk to states with limited tax penalties, plausibly generating the correlations previous analysts have observed. It is also possible that my results and previous work are consistent. Perhaps firms are responsive to state-specific variation because they are more able to understand the tax penalties of their state, than the penalties they face from their location on a schedule.

The estimates presented here regarding the effect of UI payroll tax increase are large compared to estimates from other papers studying payroll taxes (Gruber, 1997; Kugler and Kugler, 2002; Egebark and Kaunitz, 2014). Among papers that identify a negative effect of payroll taxes on employment, prior estimates tend to imply a demand elasticity smaller than one. Two papers that specifically identify the effect of UI taxes specifically estimate larger effects: Anderson and Meyer (1997) find that a 1-point increase in UI taxes are associated with a 1.4 percent reduction in firm employment implying an elasticity of greater than one. Anderson and Meyer (2000) report enormous effects, though over a time frame when the tax would reduce wages rather than employment (1-point tax increase reduced wages by up to 4 percent). It is possible that
the unique features of UI taxation cause larger unemployment effects. Normally, Payroll
taxes increases are announced well in advance, allowing companies to gradually adjust
employment and wages in expectation of the tax change. The authors of this literature
tend to tax the tax change as the event of interest rather than the announcement plus
the implementation of the tax change, possibly leading to underestimation of the payroll
tax effect. In contrast, UI taxes are announced a month before they are implemented,
allowing little time for firms to adjust employment and wages, allowing for a careful
evaluation of the short-run payroll tax change.

In Missouri, a firm’s new reserve ratio is calculated in July, at the beginning of the
third quarter, but firms are not notified of their next year’s tax rate until November and
the tax does not rise until January. Firms are completely unresponsive to next year’s tax
rate in Q3 when their new reserve ratio is calculated but they are unaware of their new
tax. In Q4 firms learn their new tax rate and tax-hiked firms at the discontinuity reduce
their hiring by 0.18 hires that quarter, significant at conventional levels. The following
quarter when the tax rate increases, tax-hiked firms reduce hiring by 0.60 hires that
quarter. Hiring is lower for the rest of the year, but the effect moderates significantly.

Furthermore, UI taxes represent, in effect, a head tax because the taxable wage
base is smaller than most employees’ yearly wages. Unlike other taxes, that means the
tax discourages the quantity employed more so than the quantity of wages paid. Finally,
UI taxes are not predictable. About half of layoffs do not claim benefits and workers can
vary widely in how long they remain unemployed, receiving benefits; the large variation
in benefits drawn makes predicting the tax difficult for firms. Because of this, firms may
interpret unexpected tax increases as indicative of the future, and thus overreact to this
period’s tax change.
Another important feature discussed in detail in the conceptual framework section of UI tax is that they may broadly apply to firms who have experienced a negative shock and may be cash-constrained. As discussed, this explains why firms respond strongly to the tax bill but are not significantly deterred by tax penalties.

In order to survive the period, a cash constrained firm must reduce its variable costs so that no matter what the penalty, the firm is unlikely to be discouraged by future tax increases. Because the firm’s cash constraint may be serially correlated, it is likely to be cash constrained the taxes rise, causing the firms to further reduce variable costs by decreasing hiring and surviving with fewer workers. This explanation is helpful because it illuminates the unexpected pattern of firm responses, based simply on the firm’s need to meet its financial obligations to survive. Similarly, cash constraints explain why tax increases can induce some firms to exit.

The timing of these effects as well as the effect of the tax on firm exit is consistent with the explanation that firms adjust as information becomes available. The estimated effect of the tax increase on exit is unexpected. I cannot locate any literature that has exploited quasi-experimental variation in regulation or taxes and reported an effect on firm exit, which could be for a myriad of reasons. To assess the size and cause of the exit effect, I estimate the effect of industry shocks on firm exit. The effect of a 1 percent UI tax increase is about equal to a 4-6 percent negative industry shock. To compare my effects to the effects of other taxes I use changes in state corporate and individual income taxes to assess their effect on bank exit using credit union and non-affected corporations as a control group. This exercise shows that firm behavior is highly

---

35 Using data from Compustat, I verify that industries undergoing negative shocks appear to be also be from somewhat more cash-constrained industries.
36 Analysts may not find effects on firm exit, may not be interested in firm exit, or systematically report null effects on firm exit to preserve the interpretation of other results.
responsive to income taxation, and firms are about 0.2 percent more likely to exit when
tax rates rise by one point. Because income taxes only reduce surpluses and do not
introduce new costs, this makes sense in light of the leading explanation of the data,
namely, UI taxes increase input costs which put a firm’s viability at stake.

Conclusion

UI tax penalties are intended to discourage layoffs and recoup the cost of
unemployment benefits. In this paper, I measure two consequences of this tax program:
the effect of UI penalties in discouraging layoffs (the “deterrent” effect); and the effect of
UI tax increases on employment and exit (the “overhang” effect). Through a variety of
quasi-experimental designs, I report that a 1 point increase in UI tax rates results in an
employment reduction of about 1 percent and an increase in the firm exit rate of about 1
percentage point, though I find no evidence that tax penalties deter layoffs. A model of
profit-maximizing firms with cash constraints explains why firms react strongly to tax
increases but are not deterred by future tax increases.

Economists tend to think of the UI program as an automatic stabilizer because
benefits discharge at times when unemployment is high. I show that UI taxes also track
the business cycle, diminishing the overall stabilizing influence of the program.
Moreover, the results measure the cost of UI financing and thus have implications for the
optimal UI literature. The work-horse model of optimal UI assumes workers pay the tax
from their wages in the high state of the world, but my results demonstrate that, at least
in the short-run, tax increases do not reduce a worker’s wages, but instead affect the
probability the worker is in the low state of the world. This understanding of the
unintended costs of UI benefits may alter our perception of the optimal generosity and
duration of UI benefits.
Finally, tax penalties increased dramatically in the mid-1980s as the federal government induced states to increase their maximum tax rate and their taxable wage base. Since then, the cost of unemployment to firms increased as workers became eligible for more generous benefits and chose to receive unemployment insurance for longer periods of time. My evidence is consistent with the claim that these increases contribute to the rise of jobless recovery since the in1980s and I demonstrate that employment recovers more haltingly in states with higher tax penalties.


