From the holdings of the Department of Special Collections , Princeton University Library, Princeton, NJ

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyrightlaw. We do not charge any permission or usage fees for the publication of images of material inour collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff.For more information regarding copyright, credit, and citation please review the information onthe Copyright, Credit and Citation Guidelines page of our website.
https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. https://library.princeton.edu/special-collections/ask-us

---

Doctoral Dissertations, 2011-2026 Economics

2013

Essays in Empirical Microeconomics

Park, Yoon Soo

Princeton, NJ : Princeton University

---

http://arks.princeton.edu/ark:/88435/dsp01j098zb23z

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. https://www.copyright.gov/title17/title17.pdf

ESSAYS IN EMPIRICAL MICROECONOMICS

YOON SOO PARK

A DISSERTATION

PRESENTED TO THE FACULTY

OF PRINCETON UNIVERSITY

IN CANDIDACY FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE

BY THE DEPARTMENT OF ECONOMICS

ADVISER: HENRY S. FARBER

SEPTEMBER 2013

© Copyright by Yoon Soo Park, 2013. All rights reserved.

Abstract

This collection of essays performs causal inference to questions of employment, health, and education using empirical strategies developed in modern empirical economics.

Chapter 1 studies why the incidence of long-term employment relationships has been declining in the U.S. labor market. This chapter asserts that an employment protection, which is likely to be more favorable to workers who have spent longer period of time with their employers, can incentivize employers to refrain from employing such workers in order to avoid the regulation and hence can be a reason for the decline in long-term employment relationships. Using cross-state and over-time variation in the timing of the law adoption, this chapter finds that the employment protection accounts for about 40 percent of the observed decline in long-term employment relationships.

Chapter 2 empirically tests whether religious participation can improve health conditions, which has been widely discussed in various fields of social science. Using repeals of regulations on secular activities on Sundays by state governments as an exogenous source of variation in religious participation, the chapter concludes that strong correlations between religious participation and health outcomes reported in previous studies are likely to have been driven by selectivity bias rather than reflecting a causal relationship.

Chapter 3, which is collaborated with Changhui Kang, examines a causal effect of receiving private tutoring on academic performance of students. This chapter finds that the mean effect of private tutoring on test scores is at most modest, while the distributional effects are positive for students whose test scores are located at upper percentiles of test score distribu-

tion but statistically insignificant from zero for those at lower percentiles of the distribution. These results suggest that private tutoring is not an effective remedial educational measure for students left behind, although it seems to have a modest effect of facilitating learning processes of students in good standing.

Acknowledgements

I would like to express my heartfelt gratitude to my principal adviser, Henry Farber. His patience and encouragement has supported me throughout my doctoral study in Princeton, not to mention that this dissertation is greatly benefited from his insight and curiosity. It is fortunate for me to have such a wonderful person as my academic adviser. I am also deeply indebted to Orley Ashenfelter, Will Dobbie, Kirill Evdokimov, Bo Honore, Changhui Kang, and Harvey Rosen for their willingness to help at various stages of this dissertation and my graduate study. Without their helps, I would not be able to finish my doctoral program. Stephane Bonhomme, Angus Deaton, Daniel Hungerman, Satoshi Shimizutani, and Izumi Yokoyama deserve special thanks for their helps that substantially improved this dissertation.

I am very grateful to Linda Belfield, Bobray Bordelon, Kathleen DeGennaro, Laura Hedden, Kathleen Hurley, Eugenia Moore, Linda Oppenheim, Matthew Parker, and Patricia Tracey for their excellent technical and administrative support. These people taught me that helping others cheerfully really brightens up my life. I also thank to my friends including, but not exclusively, Sehyoun Ahn, Qingqing Cao, Dongbeom Choi, Eleanor J. Choi, Hyunsoo Choi, Cheng Chen, Josephine Duh, Qi Ge, Nikolaj Harmon, Tatiana Homonoff, Stephanie M. Hwang, Hanjun Kim, John Kim, Penka Kovacheva, Yan Lau, Nicholas Lawson, Alice H. Lee, Jihyung Lee, Yi Li, Xiaotong Niu, Woongyong Park, Yeje Park, Doron Ravid, Minkee Song, Neel Sukhatme, Phillys Sun, Jade Vichyanond, Jungeun Yoon, and Meixia Zhang. Memories with them are "something else" of my life in Princeton than academic research. In

addition, I deeply appreciate Princeton University and the Korea Foundation for Advanced Studies for generous financial support during my graduate study.

Finally, I sincerely thank to my family with all my heart for their love and prayers.

Contents

Abstract iii

Acknowledgements v

1 Employment Protection and Long-Term Employment Relationships 1

1.1 Introduction 1

1.2 Institutional background 4

1.3 Previous literature 7

1.4 Expected effects of the implied-contract exception on job tenure 9

1.5 Data 11

1.5.1 Data on current employer tenure 11

1.5.2 Data on the timing of the legal reforms 12

1.5.3 Sample 18

1.6 Empirical analysis 25

1.6.1 Difference-in-differences 25

1.6.2 Distributional difference-in-differences 41

1.7 Conclusion 50

2 Does Religious Participation Cause Better Health Outcomes? 54

2.1 Introduction 54

2.2 Previous literature 58

2.3 Data 59

2.4 Empirical analysis 63

2.4.1 Correlation between religious participation and health condition . . . 63

2.4.2 Effects of the blue laws 78

2.5 Instrumental variable strategy 91

2.6 Summary and conclusion 97

3 The Impact of Private Tutoring on Academic Performance of Students 99

3.1 Introduction 99

3.2 Previous literature 103

3.3 Data 108

3.4 Empirical analysis 114

3.4.1 Mean effects of private tutoring 114

3.4.2 Distributional effects of private tutoring 122

3.5 Conclusion 130

A Appendix for Chapter 1 134

A.1 Details on reweighting procedure in figure 1-4 134

A.2 Details on deriving equation 1.9 137

A.3 Details on choosing the truncation parameter in equation 1.10 140

A.4 Results based on the alternative legal classification 141

B Appendix for Chapter 3 **147**

B.1 Details on deriving equation 3.10 147

B.2 Details on deriving equation 3.13 150

B.3 Details on chooing the truncation parameter in equation 3.15. 153

Chapter 1

Employment Protection and Long-Term Employment Relationships

1.1 Introduction

Long-term employment relationships, or sometimes called "life-time jobs", have been an important feature of the U.S. labor market (Hall, 1982[54]; Farber, 2008[44]). However, the incidence of long-term employment relationships has been declining particularly among male workers in the private sector[1] (Swinnerton and Wial, 1995[102]; Rose, 1995[92]; Jaeger and Stevens, 1999[64]; Valletta, 1999[106]; Stewart, 2002[101]; Farber, 2006[43], 2008[44]). In spite of the continuing decline in stable employment relationships, little[2] has been known

[1] For example, Farber (2008)[44] analyzed data from 21 supplements to the Current Population Survey (CPS) over the 1973-2006 period and found that mean tenure and incidence of long-term employment relationships, defined as jobs longer than 10 or 20 years of tenure, have been declining by about 20 percent and 10 percentage points, respectively, among private-sector male workers over the period.

[2] Some studies suggested relative wage movements against unskilled workers (Levenson, 2000[75]), growth in alternative employment relationships (Farber, 1999[42]), and firm downsizing (Cappelli, 2000[26]) as

about the source of the decline (Neumark, 2000[85]).

Given the lack of explanations for the observed decline in long-term employment relationships, I try to contribute to the literature by providing evidence that an employment protection may be a reason for the decline. The employment protection that I consider is the so-called *implied-contract exception* to the conventional employment-at-will doctrine in the U.S. common law system. A unique feature of the implied-contract exception is that it is more favorable to workers who have spent longer time with their employers (Autor et al., 2006[12]). To the extent that the implied-contract exception provides disproportionately more job security for workers with longer tenure, the common-law protection could incentivize employers to substitute long-tenured employees with short-tenured ones in order to minimize potential firing costs. This potential feedback of cost-minimizing employers on the regulation may be a reason why long-term employment relationships have become less common in the U.S.

Applying this idea to the Current Population Survey (CPS) data on current employer tenure between 1978 and 1999, I examine how the implied-contract exception affected job tenure in the U.S. labor market. Exploiting the cross-state and over-time variation in adopting the implied-contract exception by state courts, I find evidence that the law reduced mean tenure by about four percent and the probability of being in long-term employment relationships, defined as jobs with 10 years or longer tenure, by about three percentage points. These findings account for about 20 percent of the observed decline in average tenure and

potential sources of the decline. However, Neumark (2000[85]) pointed out that these studies do not suggest an exogenous source impacting employment relationship but discuss how various changes in employment relationship may reinforce each other.

about 37.5 percent of the observed decline in the incidence of the long-term employment relationships over my sample period. The observed labor market shift away from long-term employment relationships seems to have been driven by cost-minimizing employers' reaction to the legal intervention by gradually substituting the share of long-tenured jobs with that of short-tenured ones in their workforce. I find evidence that, as soon as the law had been introduced, short-tenured jobs became more and more common over time, while long-tenured ones became less and less common.

In order to see a more complete picture of the impact of the implied-contract exception, I also examine how the law affected the overall tenure distribution. Recently, Bonhomme and Sauder (2011)[20] showed that the logic of standard difference-in-differences model can be extended to the entire distribution. Applying their distributional difference-in-differences model to the CPS data on job tenure, I find evidence that the legal intervention decreased the duration of long-tenured jobs, while increasing that of short-tenured ones, which is what is predicted to happen if employers substitute long-tenured jobs with short-tenured ones. Based on these results, I conclude that the law changed the tenure distribution from one state where long-term employment relationships are more common to another state where such relationships are less common and employers rely on short jobs more than they used to.

Section 1.2 discusses the institutional background. The traditional employment-at-will doctrine and the three major exceptions in the U.S. common law system are introduced. Section 1.3 summarizes previous research on the impact of the common-law employment protections on U.S. labor market outcomes. Previous works commonly observed that the

impact of the implied-contract exception is significant enough to change firms' utilization of their workforce. Section 1.4 discusses the expected effects of the implied-contract exception on job tenure. Section 1.5 reviews the data on current employer tenure and the timing of the doctrine recognitions. The differences in legal classifications across different sources will be discussed in detail. Section 1.6 explains the empirical framework and summarizes the results, and section 1.7 concludes.

1.2 Institutional background

Traditionally, all the employment relationships in the U.S. without a formal contract for duration of the employment were "at will", meaning that they were terminable by either the employee or employer for any reason whatsoever (Muhl, 2001[82]). This legal presumption, often called the *employment-at-will doctrine*, is probably best summarized in the judge's opinion in *Payne v. Western & Atlantic Railroad Co.* (1884)[3] where the Supreme Court of Tennessee held "either party may terminate the service, for any cause, good or bad, or without cause and the other cannot complain in law."

The conventional employment-at-will doctrine, however, has been steadily eroding since the 1930s (Muhl, 2001[82]). A series of employment protection laws were introduced and provided protections for union members (National Labor Relations Act, 1935), minorities in terms of race, color, religion, sex or national origin (Title VII of the Civil Rights Act, 1964), the elderly (Age discrimination in Employment Act, 1967), and physically impaired workers

[3] 81 Tenn. 507

(Americans with Disabilities Act, 1990). Following these federal legislative protections, most state courts in the U.S. also limited discretion of employers to discharge their workers by recognizing exceptions to the conventional employment-at-will doctrine, particularly in the 1980s (Autor et al., 2006[12]). These exceptions are often categorized into public policy, good faith, and implied-contract exceptions.

The public policy exception prohibits discharges against a government policy. The leading case recognizing this exception is *Petermann v. International Brotherhood of Teamsters* (1959)[45] where an employer, Teamsters, solicited an employee, Petermann, to commit perjury to a legislative committee investigating corruption of the employer. Petermann, however, testified truthfully and was discharged the day after his truthful testimony. The Californian Courts of Appeal ruled that the termination was a wrongful discharge even though Petermann was an at-will employee because an employer's right to discharge their workers may be limited by considerations of public policy. The public policy exception is the most widely accepted exception among the three major exceptions (Muhl, 2001[82]). Although the number of adopting states varies across different classifications, which will be further discussed below, about 43 state courts currently recognize the public policy exception[6]. In spite of its widespread recognitions, however, the actual application of the exception tends to be limited to the cases where a discharge clearly violates a well-established public policy such as performing jury duty or filing a worker's compensation claim (Autor et al., 2006[12]; Perritt, 2011[90]).

[4] 174 Cal.App.2d 184

[5] I refer to the Westlaw database for the summary of court decisions in this section.

[6] The list of states that recognizes each exception as of April 2008 can be found in the webpage of National Conference of State Legislatures (http://www.ncsl.org/default.aspx?tabid=13339)

The good faith exception prevents discharges with deceitful motivation. A landmark case is *Fortune v. National Cash Register Co.* (1977)[7] where a salesman, Fortune, who was supposed to receive a sales bonus proportional to his sale, was fired on the next business day after a company, National Cash Register Co., received a $5,000,000 order from the territory of Fortune. The Supreme Judicial Court of Massachusetts found that, even though Fortune was an at-will employee, the discharge was wrongful because it was motivated by the employer's desire to deprive the employee of the promised bonus. Since the good-faith exception basically prohibits any discharge motivated by a "bad cause", it could be interpreted in a way that the court only allows termination for a "good cause" and thus taken as potentially very restrictive protection against employers. However, most state courts tend to narrowly apply the doctrine to "timing" cases as in *Fortune v. National Cash Register Co.* (1977)[8] (Autor et al., 2006[12]). Currently, 20 state courts recognized the good faith exception[9].

Finally, the implied-contract exception prohibits any discharges against an "implied contract" between an employer and an employee. Typical examples of the "implied contract" are oral or written promises made by an employer for job security or termination procedures (Muhl, 2001[82]). A leading case is *Toussaint v. Blue Cross & Blue Shield of Michigan* (1980)[10] where Toussaint, an at-will employee who was discharged by Blue Cross, claimed that the termination was wrongful because it violated an employment handbook provision stating that the company would discharge its workers only for a just cause. The Supreme Court of Michigan ruled that the employer's handbook provision, although it was not a

[7] *364 N.E.2d 1251*
[8] *364 N.E.2d 1251*
[9] Source: National Conference of State Legislatures (http://www.ncsl.org/default.aspx?tabid=13339)
[10] 292 N.W.2d 880

part of formal employment contract, was an "implied contract" and hence legally enforceable because it engendered the employee's *legitimate expectations* that he would not be fired without a valid reason. In another landmark decision, *Pugh v. See's Candies, Inc.* (1981)[11], the California Court of Appeal expanded the factors creating an "implied contract" from oral or written promises made by an employer to the employment context itself such as an employee's longevity of service (i.e. job tenure), a history of promotion or salary increases of an employee, general personnel policies of a company, or typical practices in an industry (Autor et al., 2006[12]). Currently, 43 state courts adopted the implied-contract exception[12].

1.3 Previous literature

The impacts of the three exceptions on labor market performance have been discussed by many scholars. Although the impacts on other variables such as worker flows (Kugler and Saint-Paul, 2004[73]) and firm productivity (Autor et al., 2007[13]) were also examined, the discussions were mainly focused on the effects on employment level. Dertouzos and Karoly (1992)[34] reported large negative impacts of the protections on the state employment level over the 1980-1987 period. On the contrary, Miles (2000)[79] found no significant disemployment effect for any of the three exceptions from 1965 to 1994. However, Miles (2000)[79] observed that temporary employment was increased by about 15 percent following the adoption of the implied-contract exception. This finding was confirmed by Autor (2003)[10] which found the adoption of the implied-contract exception accounts for about 20 percent of the

[11] 171 Cal.Rptr. 917

[12] Source: National Conference of State Legislatures (http://www.ncsl.org/default.aspx?tabid=13339)

growth in temporary help industry between 1973 and 1995.

Given the contrasting results between Dertouzos and Karoly (1992)[34] and Miles (2000)[79], Autor et al. (2006)[12] re-examined the effect of the three exceptions on the employment level. Using the data between 1978 and 1999, they found that recognition of the implied-contract exception decreased a state's employment-population ratio by about 0.8 to 1.6 percent. These estimates are smaller in magnitude than what Dertouzos and Karoly (1992)[34] reported, but clearly larger than the statistically insignificant estimates by Miles (2000)[79].

In a separate paper, Autor et al. (2004)[11] tried to reconcile the contrasting results between what they found in Autor et al. (2006)[12] and what Dertouzos and Karoly (1992)[34] and Miles (2000)[79] reported. Autor et al. (2004)[11] pointed out that the large negative effects on employment estimated by Dertouzos and Karoly (1992)[34] largely results from a problematic choice of instrumental variables, while the zero effect on employment estimated by Miles (2000)[79] is presumably because legal classification used in Miles (2000)[79] is too lagged to identify any economic effect.

In sum, contrasting views exist on whether the three major exceptions reduced employment. However, the previous literature commonly suggests that the implied-contract exception, among the three exceptions, exerted the most significant effects on firms' utilization of workforce. Dertouzos and Karoly (1992)[34] reported large negative effects of the implied-contract exception as well as the other two exceptions on employment. Miles (2000)[79] and Autor (2003)[10] observed that the implied-contract exception increased employment of temporary help industry. Autor et al. (2006)[12] found that only the implied-contract exception has small but robust negative effect on employment, whereas the other two exceptions do

not have such impact at all.

Given these findings from the previous works, I examine in the following sections whether the implied-contract exception accounts for the observed decline in mean tenure and the incidence of long-term employment relationships in the U.S. economy.

1.4 Expected effects of the implied-contract exception on job tenure

There are at least two channels through which the implied-contract exception could affect job tenure. First, the implied-contract exception could increase job tenure by raising firing costs. Theoretically, it is well known that employment protection in general reduces both hiring and firing (Bentolila and Bertola, 1990[17]; Hopenhayn and Rogerson, 1993[59]). The implied-contract exception could discourage employers from firing an incumbent worker in order to minimize potential risks of wrongful discharge litigations. On the other hand, the implied-contract exception could also make employers reluctant to hire a new worker because, once hired, those workers would be more costly to be discharged in the future. The decrease in both firing and hiring would contribute to an increase in job tenure because current employments are more likely to be maintained, while new jobs are less likely to be created. In sum, the implied-contract doctrine could increase job tenure through increasing firing costs of current and potential employees.

However, the implied-contract exception could also decrease job tenure. Autor et al. (2006)[12] indicated that the implied-contract exception is likely to provide a higher level of

job protection for workers with longer tenure for the following two reasons. First, workers who have spent more time with their employers are likely to have more past events that could be used to justify that they have developed a *legitimate expectation* for job security. Hence, long-tenured workers can more easily make a prima facie case that there is an enforceable implied contract in their employment relationships (Autor et al., 2006[12]). Some state courts even explicitly indicated that they would consider job tenure when deciding if there is an enforceable implied contract. For example, in *Pugh v. See's Candies, Inc.* (1981)[13], the Californian Court of Appeal acknowledged an employee's longevity of service as a potential factor creating an implied contract between an employer and an employee. Second, the amounts of compensatory damages in implied-contract litigations tend to be proportional to workers' prior earnings (Autor et al., 2006[12]), thus workers with longer tenure, who are also likely to receive higher earnings, tend to receive a larger amount of damage awards if they win the litigations. Since the implied-contract exception provides disproportionately more protection for workers with longer tenure in these ways, the doctrine could also discourage employers from employing long-tenured workers because those workers are potentially more costly to be discharged in the future. To the extent that the implied-contract exception increases the relative employment cost of long-tenured workers to short-tenured ones, cost-minimizing employers would react to the legal intervention by substituting more costly long-tenured employees with less costly short-tenured ones and thus job tenure would decline.

Since the implied-contract doctrine can have both positive and negative effects on job tenure, the net effect is ambiguous and should be measured empirically. Given that each

[13] 171 Cal.Rptr. 917

state adopted the implied-contract exception at different points in time, one may exploit this cross-state and over-time variation to identify the causal effect of the implied-contract exception on job tenure. The rest of the chapter describes my empirical analysis of estimating such impact.

1.5 Data

1.5.1 Data on current employer tenure

The data for this study are drawn from the Current Population Survey (CPS) over the period between 1978[14] and 1999[15]. Particularly, I use the mobility supplements in January 1978, 1981, 1983, 1987, 1991, and February 1996, 1998; the pension and benefit supplements in May 1979, 1983, 1988 and April 1993; and the contingent and alternative employment arrangement supplements in February 1995, 1997, 1999[16]. Most of these CPS supplements collect data on number of years that respondents have been working for their current employers in integer form. Following Farber (2008)[44], I interpret the integer report of job tenure as the nearest integer to the underlying true job tenure. For example, I assume that reported job tenure of 10 years indicate that true job tenure would be between 9.5 years and 10.5 years. Based on this interpretation, I recover true job tenure in the following way. If an integer

[14] Although the CPS data on job tenure are available in machine-readable form as early as 1973, the information on respondents' state of residence is only available since 1978. In the 1973 CPS, some small neighboring states are grouped together and thus cannot be distinguished from one another.

[15] Given that most states adopted the implied-contract exception by 1990, I restrict my sample to the observations before 1999 because a pre-post comparison beyond a decade after the adoption of the law is not useful.

[16] There are some differences in wording of questions on job tenure across different surveys. See Farber (2008)[45] for the issue of comparability of these surveys.

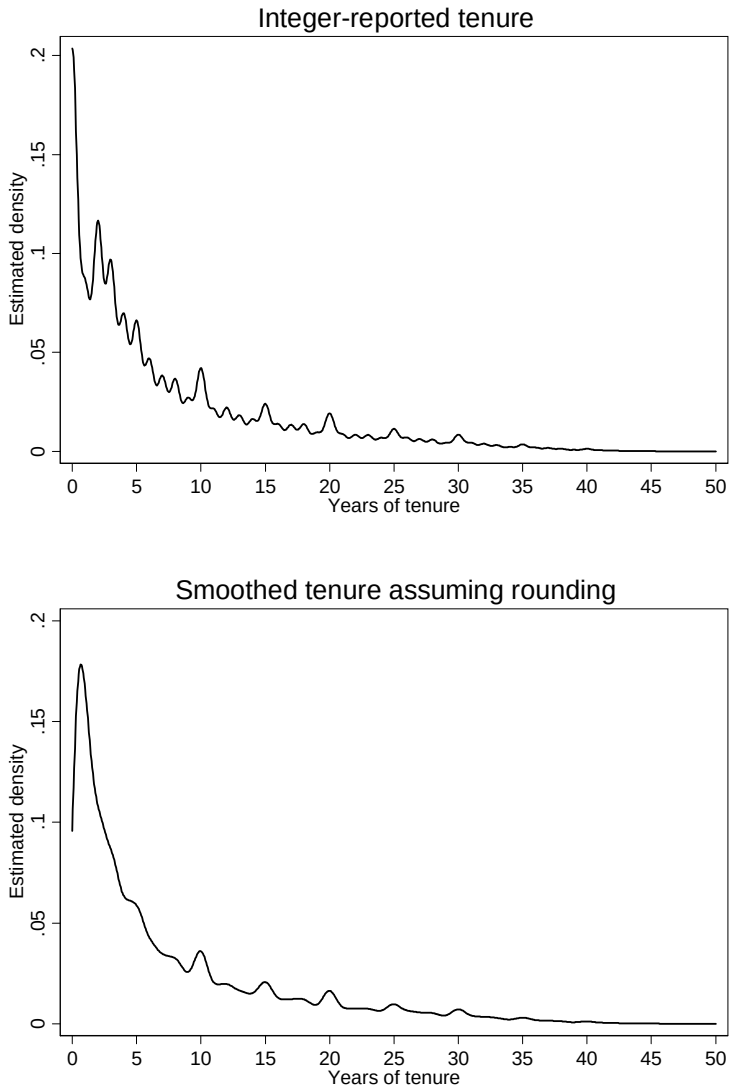
report of job tenure of T is greater than one, I replace T with $\widetilde{T} = T + U$ where U is a uniform random variable distributed on the unit interval $[-.5, .5)$, or $\widetilde{T} \sim Uniform\,[T - .5, T + .5)$. If a reported tenure of T is equal to one, I replace T with $\widetilde{T} \sim Uniform\,[1, 1.5)$. When T is equal to zero, I replace T with $\widetilde{T} \sim Uniform\,(0, 1)$.

Figure 1-1 compares kernel density estimates of the reported tenure with the smoothed tenure. The overall shapes of the two distributions are similar. However, the distribution of the integer-reported tenure reveals spikes at five-year or ten-year intervals, while this tendency is less clear for the smoothed tenure distribution. Although all the results in my analysis are not affected by the choice of tenure variables, the smoothing procedure has two major advantages for my analysis. First, since the smoothing procedure converts a reported tenure of $T = 0$ to a nonzero $\widetilde{T} \sim Uniform\,(0, 1)$, I can define the logarithm of years of tenure without loss of any data. Second, the smoothed tenure is continuously distributed and hence easier to handle when estimating distributional effects of the implied-contract exception on overall tenure distribution, which will be discussed in section 1.6.2.

1.5.2 Data on the timing of the legal reforms

There are two legal classifications covering my sample period of 1978-1999. One is classified by Walsh and Schwarz (1996; hereafter WS)[108] which has been updated to 2000 by Muhl (2001)[82] and the other is by Autor, Donohue, and Schwab (2006; hereafter ADS)[12]. Both classifications code whether or not each state adopted the implied-contract exception and, if adopted, the timing of the law adoptions. I compare the two classifications and find that

Figure 1-1: Kernel Density Estimates of Integer-Reported Tenure and Smoothed Tenure



Note. I use Gaussian kernel for the kernel density estimation. Given that tenure distribution is skewed to the right, I choose to use bandwidth smaller than (as large as two thirds of) the rule of sum bandwidth suggested by Silverman (1986).

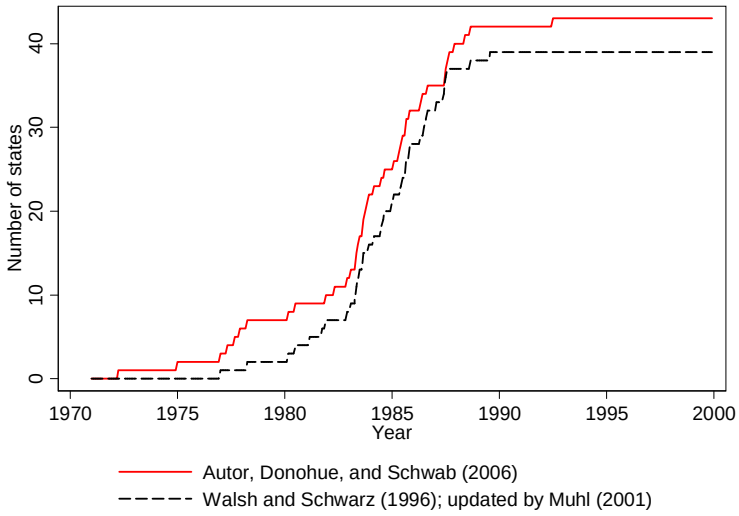
they are more or less similar but also substantially differ in some states[17]. Among the 50 states excluding the District of Columbia, 33 states are coded equivalently, while 17 states are coded differently. Among the 17 states in which the two classifications differ, 13 states have differences only in the timing of the law adoptions, while the other four states are coded inconsistently on whether or not the implied-contract doctrine was recognized.

A question that might naturally follow is to what extent these differences are significant. The results from the previous literature suggest that the differences in the two legal classifications are potentially important. Miles (2000)[79] examined the effect of the three exceptions on employment using the WS classification and concluded that none of the three exceptions has any significant effect. On the contrary, Autor et al. (2006)[12] examined the same question using their own classification and found that only the implied-contract exception, but not the other two exceptions, has a significant negative effect on employment. In a separate paper (Autor et al., 2004[11]), they also showed that the contrasting results between the two studies are largely driven by the differences in the two legal classifications that they relied on. Using the data and empirical models of Miles (2000)[79] but replacing the WS legal classification used in Miles (2000)[79] with the ADS classification used in Autor et al. (2006)[12], they found a significant disemployment effect of the implied-contract exception as reported in Autor et al. (2006)[12].

Autor et al. (2004)[11] also provided an explanation as to why the two legal classifications are inconsistent for some states. They argued that the ADS scheme tried to choose the earliest decisions that are likely to trigger management-side employment lawyers to write

[17]The timing of the law adoption classified by Walsh and Schwarz (1996) is provided in the appendix.

Figure 1-2: Number of States with the Implied-Contract Exception



awareness letters to their clients, while the WS scheme attempted to locate the cases that best articulate courts' rationale for recognizing the implied-contract exception. They further asserted that this difference in the way of choosing the relevant cases can explain why the WS scheme tends to pick up a later case than the ADS. Among the 13 states where the two classifications differ in the timing of the law adoption, the ADS cases precede the WS ones in 11 states. Figure 1-2 shows the tendency that the ADS classification chooses an earlier case than the WS scheme.

In order to confirm the validity of this argument, I compare the WS and the ADS classifications for each of the 50 states. For example, the ADS classifies that California adopted the implied-contract exception in 1972 referring to *Drzewiecki v. H&R Block* (1972)[18]. In this case[19], the Court of Appeal in California explicitly mentioned that it still maintained

[18] 101 Cal. Rptr. 169

[19] I refer to the Westlaw database for the summary of court decisions in this section.

the employment-at-will as a default doctrine[20]. However, the general decision was for employee who claimed that his discharge was wrongful because an employer had made a written promise for job security. While the ADS interpreted this case as the first case recognizing the implied-contract exception, the WS chose *Pugh v. See's Candies, Inc.* (1981)[21] in which the Court of Appeal in California clearly declared that any informal promise made by an employer is potentially actionable. This difference can be also seen in the states where the WS cases precede the ADS ones. Taking an example of Mississippi, the WS chose *Robinson v. Board of Trustees of East Cent. Jr. College* (1985)[22] where the Supreme Court of Mississippi clearly stated that provisions in an employment handbook could be a part of an employment contract. However, the decision was against the employee and the ADS might take this as suggesting that the court did not recognize the implied-contract exception "in practice" and therefore the case is not likely to trigger any management-side reaction. Based on these observations, I conclude that the explanation by Autor et al. (2004)[11] is convincing.

A potentially more important question that needs to be answered is why the two classifications are opposite for the four states: Indiana, Massachusetts, Texas, and Virginia. The ADS classified that these states adopted the implied-contract exception, while the WS did not. A potential explanation for these severe discrepancies is that the decisions of the courts in these states could be read somehow ambiguously. For example, in *Shaw v. S.S. Kresge*

[20] Quoted from the judge's opinion: "The authorities are in general agreement with the proposition that contracts which purport to give a person permanent employment through the use of such phrases as 'for life,' or 'for so long as the employee chooses,' or 'so long as the employee's work is satisfactory,' with nothing more, are for an indefinite period and are terminable at the will of either party."

[21] 171 Cal.Rptr. 917

[22] 477 So.2d 1352

Co.(1975)[23], the Indiana Court of Appeals stated that, as long as there is no promise made by an employer for the duration of an employment relationship, the employment relationship is terminable at will. This decision has not been overruled yet and is still in place[24]. However, in *Romak v. Public Service Co.* (1987)[25], the Supreme Court of Indiana admitted that oral assurance regarding job security made by an employer at hiring stage created an enforceable contract. Since the two conflicting cases are still in place, one might classify Indiana into either adopting or non-adopting group.

The same is true for the other states. On the one hand, the courts ruled that terms in personnel manual (*Jackson v. Action for Boston Community Development, Inc.*, 1988[26]); oral agreement that employment was permanent (*Webber v. M.W. Kellogg Co.*, 1986[27]); oral assurance for job security conditional on funding availability (*Miller v. SEVAMP Inc.*, 1987[28]) did not create an enforceable implied contract. However, on the other hand, the courts also stated that terms in hospital's bylaws (*Hobson v. McLean Hospital Corp.*, 1988[29]); oral agreements for termination only for good cause (*Johnson v. Ford Motor Co.*, 1985[30]), oral assurances that an employee would be fired only for cause (*Frazier v. Colonial Williamsburg Foundation*, 1983[31]) created an enforceable implied contract.

To the extent that the courts' opinion could be read somehow ambiguously, one might understand why the two legal classifications are inconsistent against each other for those four

[23] 328, N.E. 2d 775
[24] Source: the West Law database
[25] 511 N.E.2d 1024
[26] 525 N.E.2d 411, Supreme Judicial Court in Massachusetts
[27] 720 S.W.2d 124, Court of Appeals in Texas
[28] 362 S.E.2d 915, Supreme Court of Virginia
[29] 522 N.E.2d 975, Supreme Judicial Court in Massachusetts
[30] 690 S.W.2d 90, Court of Appeals in Texas
[31] 574 F.Supp. 318, United States District Court, E.D. Virginia

states. On the one hand, it could be said that these states did not recognize the implied-contract exception because the courts' position had not been clearly determined yet. On the other hand, however, it is still reasonable to say that the state courts applied the conventional employment-at-will doctrine less strictly, which could be interpreted as the implied-contract exception was recognized to some extent.

Given these differences between the two classifications, I choose to follow the characterization by Autor et al. (2006)[12] for my empirical analysis. Any economic effects of the changes in the common law would be better detected by the use of the legal classification trying to choose the "first" decision (the ADS scheme) than the classification which tries to locate the "most articulate" decision (the WS scheme). In order to mitigate concerns about subjectivity, all the empirical results below are also estimated using the alternative legal classification based on the WS scheme. The results are presented in the appendix. I find that my results are substantially robust to the different use of legal classifications.

1.5.3 Sample

Table 1.1 illustrates the timings of recognizing the implied-contract exception across different states. In order to secure at least four years of pre-adoption period and eight years of post-adoption period for all the observations in my 1978-1999 sample, I use observations from 30 states that recognized the implied-contract exception between 1982 and 1988[32]. As figure 1-2 suggests, this is also the period during which the U.S. state courts most actively recognized

[32] Any states that adopted the implied-contract exception between 1989 and 1991 also satisfy this condition and thus could have been included into my sample. However, according to the ADS legal classification, there is no state that recognized the implied-contract exception between 1989 and 1991.

the implied-contract exception. Having both pre-adoption and post-adoption periods for all the observations in my sample is necessary for estimating distributional effects of the implied-contract exception on the overall tenure distribution, which will be discussed later in this paper. Another reason for choosing these states is that they are likely to be influenced by the landmark 1981 Californian case, *Pugh v. See's Candies, Inc.* (1981), in which the court recognized an employee's "longevity of service" as a valid factor creating an enforceable implied contract.[33]

I also add observations from seven states where the implied-contract exception was never adopted to my sample. Consequently, my sample consists of 176,620 workers who are male, working in the private sector, not self-employed, aged 20 to 64 in either the 30 states[34] that recognized the implied-contract exception between 1982 and 1988 or the seven states never adopting the exception over the sample period, 1978-1999. The 37 sample states are marked with an asterisk in table 1.1. None of the 37 states in my sample adopted the implied-contract exception before 1982, while the exception was in place in all the 30 adopting states but none of the seven non-adopting states after 1988. For subsequent use, I denote the period of 1978-1981 as pre-treatment period, the period of 1982-1988 as treatment period, and the period of 1989-1999 as post-treatment period.

Table 1.2 summarizes descriptive statistics of my sample by each period. It appears that

[33] Adding the excluded states into my sample does not change my results in sections 1.6.1 and 1.6.2.

[34] I exclude Arizona and Missouri from my sample because these states repealed the implied-contract exception shortly after they had recognized the exception. The Court of Appeals of Arizona recognized the implied-contract exception in June 1983 (*Leikvold v. Valley View Community Hosp.*, 688 P.2d 201). However, only a few months later, the Supreme Court of Arizona vacated this decision (688 P.2d 170, April 1984). Similarly, the Missouri Court of Appeals recognized the implied-contract exception in January 1983 (*Arie v. Intertherm, Inc.* 648 S.W.2d 142). However, this ruling was overturned by the Supreme Court of Missouri in February 1988 (*Johnson v. McDonnell Douglas Corp.* 745 S.W.2d 661).

Table 1.1: Timing of Law Adoption and Sample of States

Year of Adoption	States	# of States
1972	California	1
1974	Illinois	1
1976	Oklahoma	1
1977	Maine, Idaho, Washington	3
1978	Oregon	1
1980	Michigan, New Mexico	2
1981	Tennessee	1
1982	New York*, Ohio*	2
1983	Minnesota*, Missouri, South Dakota* Nebraska*, Virginia*, Kentucky* Arizona, Colorado*, Nevada*, Alaska*	10
1984	North Dakota*, Kansas*, Arkansas*	3
1985	Vermont*, Connecticut*, New Jersey*, Wisconsin*, Maryland*, Texas*, Wyoming*	7
1986	West Virginia*, Utah*, Hawaii*	3
1987	Indiana*, Iowa*, South Carolina* Alabama*, Montana*	5
1988	New Hampshire*, Massachusetts*	2
1992	Mississippi	1
Never had laws	Rhode Island*, Pennsylvania* Delaware*, North Carolina*, Georgia* Florida*, Louisiana*	7
Total		50

Note. In order to secure at least four years of pre-treatment period and eight years of post-treatment period, I use the 30 states that recognized the implied-contract exception between 1982 and 1988 or the seven states that never adopted the exception. These states are marked with an asterisk. I exclude Arizona(recognized in 1983 and repealed in 1984) and Missouri(recognized in 1983 and repealed in 1988) because the implied-contract exception was repealed after being recognized.

job tenure has been declining over my sample period. Average tenure was about 8.28 years during the pre-treatment period but decreased to 7.73 and 7.34 years in the treatment and post-treatment periods, respectively. This pattern can be seen more clearly in figure 1-3. Panel B of figure 1-3 shows that the estimated cumulative distribution functions (CDF) of job tenure have been shifting to the left over my sample period.

Table 1.2 also shows that workers observed in later years tend to spend more time in education and are more likely to be minority and have Hispanic ethnicity than those observed in earlier years. To the extent that these worker characteristics are related with job tenure, the observed decline in tenure distribution might have been simply driven by the changes in worker composition, not reflecting any structural decline in job stability. Figure 1-4 demonstrates that the observed decline in tenure distribution is not merely compositional. In order to adjust for the changes in worker characteristics, I employ the reweighting procedure suggested by DiNardo et al. (1996)[37]. Particularly, I reweight tenure distributions in later years (1982-1988 and 1989-1999) using distribution of worker characteristics in 1978-1981 as weights and compare the reweighted distributions with the actual tenure distribution in 1978-1981.[35] Even after adjusting for the changes in worker characteristics, I still observe that the overall CDFs of tenure distribution has been shifting to the left. Interestingly, I also find that the shift of tenure distribution mainly took place between the treatment and post-treatment periods, suggesting that there was structural decline in job stability between the two periods[36].

[35] Details on this procedure are provided in the appendix.

[36] This result corresponds to the conclusion of Diebold et al. (1996)[35]. Using the CPS data on job tenure between 1973 and 1991, they computed retention rates for artificial cohorts and find that aggregate retention rates were stable over the 1980's.

Table 1.2: Summary Statistics

Variables:	Total 1978-1999	By Periods 1978-1981	1982-1988	1989-1999
Current employer tenure (years)				
Mean	7.65	8.28	7.73	7.34
25th percentile	1.43	1.52	1.44	1.39
Median	4.33	4.75	4.29	4.17
75th percentile	10.8	11.8	10.9	10.3
Individual-level covariates:				
Age (years)				
Mean	38.0	38.2	37.3	38.2
25th percentile	28	28	28	29
Median	36	36	35	37
75th percentile	47	48	46	46
Education (mean; yes=1)				
Less than 12 years	.161	.236	.180	.118
12 years	.388	.407	.408	.368
13 to 15 years	.217	.171	.182	.257
16 years or more	.234	.186	.230	.257
Race/ethnicity (mean; yes=1)				
Nonwhite	.118	.099	.110	.130
Hispanic	.068	.044	.057	.084
State-level covariates (mean) :				
Employment-to -population ratio	.717	.696	.708	.748
Manufacturing-to -nonmanufacturing employment ratio	.237	.312	.250	.202
Union coverage rate	.194	.257	.213	.165
Number of observations	176,620	46,098	45,686	84,836

Note. Employment-to-population ratio and the ratio of manufacturing employment to nonmanufacturing employment are calculated for male workers using the data file used for Autor et al. (2006) available at http://economics.mit.edu/faculty/dautor/data/autdonschw06. Union coverage rates are from the Union Membership and Coverage Database available at www.unionstats.com. Both of the data are originally calculated from the monthly Current Population Survey.

Figure 1-3: Changes in Tenure Distribution

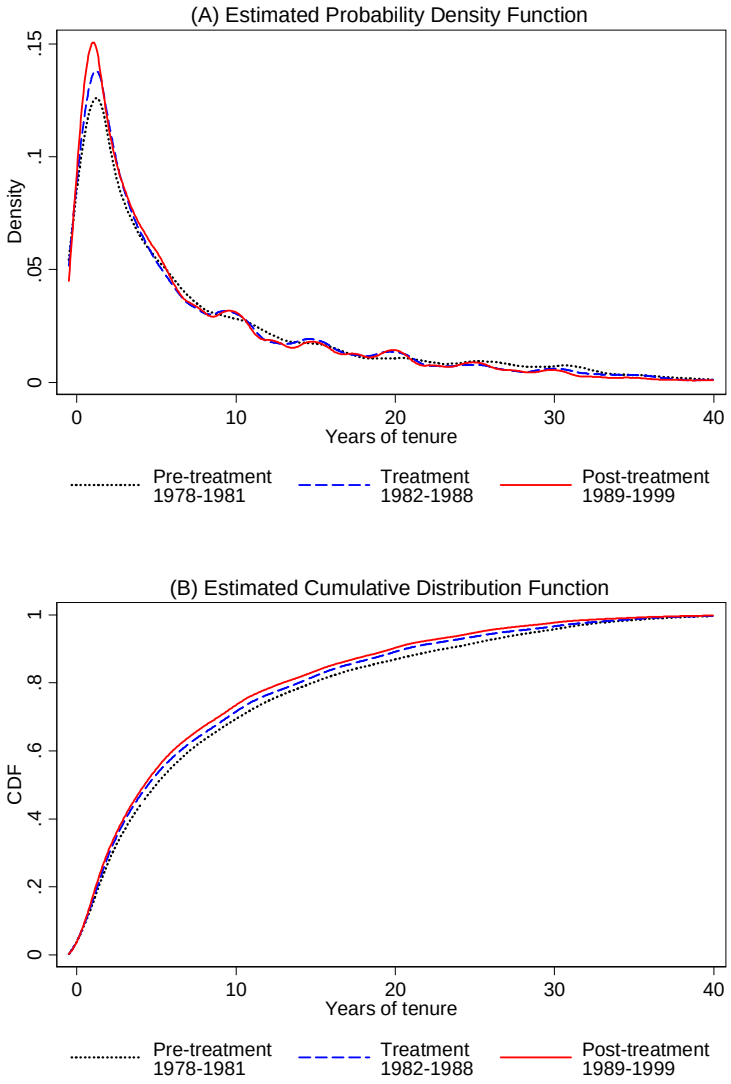
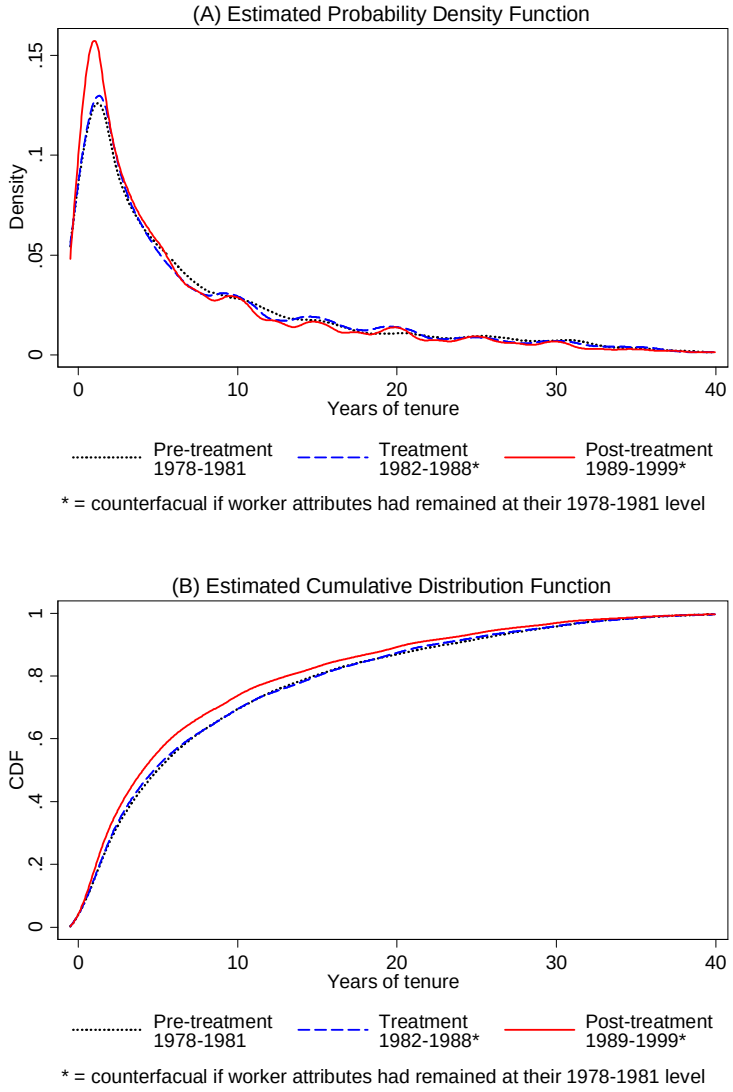


Figure 1-4: Changes in Tenure Distribution (Adjusted for Changes in Worker Characteristics)



Note. Changes in worker characteristics such as age, education, race, and ethnicity are adjusted when estimating tenure distributions in treatment (1982-1988) and post-treatment (1989-1999) periods using reweighting procedure suggested by DiNardo et al. (1996). Details on the reweighting procedure are provided in the appendix.

1.6 Empirical analysis

1.6.1 Difference-in-differences

Employing the repeated cross-sectional data on current employer tenure drawn from the 14 CPS supplements between 1978 and 1999, I estimate the following equation by the ordinary least square (OLS) method

$$\log(TEN_{ist}) = \beta_0 + \alpha_s + \gamma_t + \beta_1 ICE_{st} + X_{ist}\beta_2 + \varepsilon_{ist}, \tag{1.1}$$

where $\log(TEN_{ist})$ represents the logarithm of years of tenure of worker i in state s in year t; α_s is a set of state dummies (state-fixed effects); γ_t is a set of year dummies (year-fixed effects); ICE_{st} denotes an indicator for whether the implied-contract exception is in place in state s in year t; X_{ist} is a set of individual characteristics such as age, age squared, four dummies for categorized years of education, and dummies for race, ethnicity, and their interaction; and ε_{ist} is a random error term. Since the regressor of interest, ICE_{st}, varies only at a state and a year level, potential omitted variables, if any, must vary at a state and a year level, too. If those potential omitted variables can be captured by the sum of time-invariant state fixed effects, α_s, and year effects that are common across different states, γ_t, the "difference-in-differences" parameter, β_1, identifies the proportional mean effect of recognizing the implied-contract exception on job tenure.

The regression equation 1.1 measures the mean effect of the implied-contract exception. However, the legal intervention might affect the tenure distribution in ways not revealed by

an examination of mean. Particularly, the effects on the incidence of long-term employment relationships, which is also central for this study, will be captured by looking at upper tail of the tenure distribution. In order to examine the impacts of recognizing the implied-contract exception on the incidence of long-term employment relationships, I estimate the following logit model

$$\Pr\left[TEN_{ist} \geq C\right] = \Lambda\left(\sigma_0 + \pi_s + \lambda_t + \sigma_1 ICE_{st} + X_{ist}\sigma_2\right), \tag{1.2}$$

where $\Pr\left[TEN_{ist} \geq C\right]$ indicates probability of jobs lasting at least $C \in \{10, 20\}$ years; π_s and λ_t denote state and year fixed effects, respectively; and $\Lambda\left(\cdot\right)$ denotes the cumulative distribution function (CDF) of a logistic distribution. Following Farber (2008)[44], when estimating equation 1.2, I restrict my sample to workers aged at least 35 and 45 years for $C = 10$ and $C = 20$, respectively, because workers younger than those ages cannot accumulate substantial years of job tenure.

I augment equations 1.1 and 1.2 by controlling for a richer set of covariates. First, I add covariates related to state labor market conditions to the list of the covariates of equations 1.1 and 1.2. Job tenure tends to move in a counter-cyclical way because workers with shorter tenure are more likely to lose their jobs than those with longer tenure during recession (Abraham and Medoff, 1984[2]). The share of manufacturing employment has been declining for over the past sixty years[37] and there is a view that the shift of the U.S. labor market from manufacturing toward non-manufacturing employment might be related to the observed decline in job stability (Farber, 2008[44]). To the extent that labor union

[37]Farber (2008)[44] reported that the manufacturing share of non-farm employment was 30.9 percent in 1950 and fell to 10.4 percent in 2006.

provides job security with its members, the well-known de-unionization in the 1980s (Farber, 1990[41]) might account for the decline in job tenure. In order to control for these influences, I include employment-to-population ratio, the ratio of manufacturing employment to non-manufacturing employment, and union coverage rate of each state in each year to the list of my covariates.

Another potential source of bias that the state fixed effects and the year fixed effects in equations 1.1 and 1.2 might not control for is that there is a clear regional difference in the state courts' tendency toward adopting the implied-contract exception. Among the seven states that are classified as non-adopting states (i.e. states that have never adopted the implied-contract exception over the sample period), four states[38] are located in the Southern region[39] (Autor et al., 2004[11]). To the extent that there were any regional differences in the labor market conditions that might correlate with job tenure, the observed correlation between the law adoption and job tenure might simply arise from the differences in the labor market conditions between the Southern states and the non-Southern states. In order to control for this potential confounding factor, I also include interactions between nine Census region dummies and individual calendar year dummies into the list of controls. This allows the year effects in equations 1.1 and 1.2 can vary across different geographic regions in an unrestricted way (region-specific year fixed effects).

Tables 1.3, 1.4, and 1.5 summarize the OLS and logit estimation results of equations 1.1 and 1.2. In each of the three tables, column 1 reports the estimation results of the baseline

[38] The four states are North Carolina, Georgia, Florida, and Louisiana.

[39] When the Walsh and Schwarz (1996) classification is used, six states - Virginia, North Carolina, Georgia, Florida, Louisiana, Texas - out of the 11 control states are located in the Southern region.

specifications which control for individual characteristics. In column 2, I augment these baseline models by adding covariates related to state labor market conditions. Column 3 shows the estimation results when region-by-year dummies, but not the state-level covariates, are added to the list of controls of the baseline models. Finally, column 4 presents the estimation results when both the region-by-year dummies and the state-level covariates are controlled for. The estimated coefficients on the indicator for whether the implied-contract exception is in place (ICE_{st}) is presented on the first row of each table. Bertrand et al. (2004)[18] emphasized the potential risk of inconsistent standard errors when employing difference-in-differences estimation for many years of data. They recommended to allow for an arbitrary autocorrelation process when computing the standard errors if the number of states is sufficiently large. Following the suggestion by Bertrand et al. (2004)[18], all the standard errors are clustered at state level.

Table 1.3 summarizes the proportional effects of the implied-contract exception on mean tenure. Column 1 of table 1.3 reports the estimation result of equation 1.1. It appears that the law reduced mean tenure by about 2.1 percent but this estimate is statistically insignificant. In column 2, I observe that the estimated coefficients on the doctrine adoption indicator become larger in magnitude and also more precise when state-level labor market conditions are controlled. The doctrine appears to reduce mean tenure by a statistically significant 4.1 percent. Given the stark contrast in the tendency toward adopting the doctrine between the Southern states and the non-Southern states, I added interactions between nine Census region dummies and individual calendar year dummies to the list of controls in columns 3 and 4. If the observed correlation are simply driven by the tendency that the

Table 1.3: OLS Estimation Results: Effects on Mean Log Tenure

Dependent variable:	Specifications			
Log of years of job tenure	(1)	(2)	(3)	(4)
Implied-contract exception	-.021	-.041**	-.024	-.037**
(Standard error)	(.019)	(.015)	(.018)	(.018)
Covariates:				
Individual characteristics	Yes	Yes	Yes	Yes
State labor market conditions	No	Yes	No	Yes
State and year dummies	Yes	Yes	Yes	Yes
Region by year dummies	No	No	Yes	Yes
Age restriction	No	No	No	No
R-squared	.236	.237	.237	.237
Number of observations	176620	176620	176620	176620

Note. Dependent variable is the log of years of current employer tenure. Regressor of interest is an indicator that takes 1 if the implied-contract exception is in place and 0 otherwise. Individual characteristics include age, age squared, three dummies for education attainment: high school graduate, some college, more than college graduate where the omitted category is less than high school graduate, dummies for nonwhite, Hispanic origin, and an interaction between nonwhite and Hispanic. State labor market conditions include employment-population ratio, the ratio of manufacturing employment to non-manufacturing employment, and union coverage rate. Standard errors are clustered at state level. The F-test for joint significance of the region-by-year dummies rejects the null hypothesis of equal year fixed effects across different geographic regions at one percent of significance level in columns 3 and 4. p<.01 (***), p<.05 (**), p<.10 (*)

Table 1.4: Logit Estimation Results: Effects on Pr[10 years or longer]

Dependent variable:	Specifications			
Indicator for jobs lasting at least 10 years	(1)	(2)	(3)	(4)
Implied-contract exception	-.089**	-.117***	-.105***	-.123***
(Standard error)	(.043)	(.038)	(.038)	(.041)
[Marginal effect]	[-.022]	[-.029]	[-.026]	[-.030]
Covariates:				
Individual characteristics	Yes	Yes	Yes	Yes
State labor market conditions	No	Yes	No	Yes
State and year dummies	Yes	Yes	Yes	Yes
Region by year dummies	No	No	Yes	Yes
Age restriction	≥ 35	≥ 35	≥ 35	≥ 35
R-squared	.062	.063	.064	.064
Number of observations	100474	100474	100474	100474

Note. Dependent variable is an indicator for jobs with 10 years or longer current employer tenure. Regressor of interest is an indicator that takes 1 if the implied contract exception is in place and 0 otherwise. Individual characteristics include age, age squared, three dummies for education attainment: high school graduate, some college, more than college graduate where the omitted category is less than high school graduate, dummies for nonwhite, Hispanic origin, and an interaction between nonwhite and Hispanic. State labor market conditions include employment-population ratio, the ratio of manufacturing employment to non-manufacturing employment, and union coverage rate. Standard errors are clustered at state level. All marginal effects are evaluated at the mean values of the other covariates. The F-test for joint significance of the region-by-year dummies rejects the null hypothesis of equal year fixed effects across different geographic regions at one percent of significance level in columns 3 and 4. p<.01 (***), p<.05 (**), p<.10 (*)

Table 1.5: Logit Estimation Results: Effects on Pr[20 years or longer]

Dependent variable:	Specifications			
Indicator for jobs lasting at least 20 years	(1)	(2)	(3)	(4)
Implied-contract exception	-.037	-.042	-.011	-.037
(Standard error)	(.068)	(.060)	(.068)	(.062)
[Marginal effect]	[-.008]	[-.009]	[-.002]	[-.008]
Covariates:				
Individual characteristics	Yes	Yes	Yes	Yes
State labor market conditions	No	Yes	No	Yes
State and year dummies	Yes	Yes	Yes	Yes
Region by year dummies	No	No	Yes	Yes
Age restriction	≥ 45	≥ 45	≥ 45	≥ 45
Pseudo R-squared	.048	.048	.050	.050
Number of observations	53965	53965	53965	53965

Note. Dependent variable is an indicator for jobs with 20 years or longer current employer tenure. Regressor of interest is an indicator that takes 1 if the implied contract exception is in place and 0 otherwise. Individual characteristics include age, age squared, three dummies for education attainment: high school graduate, some college, more than college graduate where the omitted category is less than high school graduate, dummies for nonwhite, Hispanic origin, and an interaction between nonwhite and Hispanic. State labor market conditions include employment-population ratio, the ratio of manufacturing employment to non-manufacturing employment, and union coverage rate. Standard errors are clustered at state level. All marginal effects are evaluated at the mean values of the other covariates. The F-test for joint significance of the region-by-year dummies rejects the null hypothesis of equal year fixed effects across different geographic regions at one percent of significance level in columns 3 and 4. p<.01 (***), p<.05 (**), p<.10 (*)

employment protection was adopted mostly in the non-Southern states, the statistically significant correlation between the doctrine adoption and mean tenure would disappear with the inclusion of the region-specific year fixed effects. However, the result in column 4 confirms that the law reduced mean tenure by about 3.7 percent even after the region-by-year dummies are controlled. The F-test for joint significance of the region-by-year dummies rejects the null hypothesis of equal year fixed effects across different geographic regions at one percent of significance level. Based on these observations, I conclude that the doctrine adoption reduced mean tenure slightly by about four percent.

Tables 1.4 and 1.5 report the effects of the implied-contract exception on the incidence of long-term employment relationships defined as jobs lasting at least 10 or 20 years continuously. In both tables, all marginal effects of the law indicator are evaluated at the mean values of the other covariates. All the specifications in table 1.4 commonly demonstrate that recognizing the implied-contract exception decreased the probability of being in 10 years or longer employment relationships by about three percentage points relative to workers in the non-adopting states at that time. On the contrary, table 1.5 indicates that the doctrine did not affect the probability of jobs lasting 20 years or longer regardless of the choice of covariates. Smaller sample size due to more binding age restriction could explain relatively large standard errors to certain extent. However, the magnitude of the estimated coefficients on the law indicator is much smaller than those of table 1.4, which suggests the doctrine mostly affected jobs with 10 to 20 years of tenure but not those lasting beyond 20 years.

In order to see whether the data support this interpretation, I estimate equation 1.2 by using an indicator for jobs with 10 to 20 years of tenure as a dependent variable. For this

Table 1.6: Logit Estimation Results: Effects on Pr[10 to 20 years]

Dependent variable:	Specifications			
Indicator for jobs w/ 10-20 yrs of tenure	(1)	(2)	(3)	(4)
Implied-contract exception	-.072***	-.091***	-.109***	-.112***
(Standard error)	(.028)	(.034)	(.037)	(.036)
[Marginal effect]	[-.014]	[-.017]	[-.020]	[-.021]
Covariates:				
Individual characteristics	Yes	Yes	Yes	Yes
State labor market conditions	No	Yes	No	Yes
State and year dummies	Yes	Yes	Yes	Yes
Region by year dummies	No	No	Yes	Yes
Age restriction	≥ 35	≥ 35	≥ 35	≥ 35
Pseudo R-squared	.013	.013	.014	.014
Number of observations	100474	100474	100474	100474

Note. Dependent variable is an indicator for jobs with 10 to 20 years of current employer tenure. Regressor of interest is an indicator that takes 1 if the implied contract exception is in place and 0 otherwise. Individual characteristics include age, age squared, three dummies for education attainment: high school graduate, some college, more than college graduate where the omitted category is less than high school graduate, dummies for nonwhite, Hispanic origin, and an interaction between nonwhite and Hispanic. State labor market conditions include employment-population ratio, the ratio of manufacturing employment to non-manufacturing employment, and union coverage rate. Standard errors are clustered at state level. All marginal effects are evaluated at the mean values of the other covariates. The F-test for joint significance of the region-by-year dummies rejects the null hypothesis of equal year fixed effects across different geographic regions at one percent of significance level in columns 3 and 4. p<.01 (***), p<.05 (**), p<.10 (*)

estimation, I restrict my sample to workers aged at least 35 years because workers younger than those ages cannot accumulate substantial job tenure. The estimation results in table 1.6, in conjunction with the results of tables 1.4 and 1.5, confirms the prediction that the law mainly reduced the incidence of jobs with 10 to 20 years of tenure.

The contrasting effects of the law on the incidence of jobs with 10 to 20 years of tenure and those lasting beyond 20 years of tenure could be interpreted in the following way. As soon as the law is introduced, cost-minimizing employers would want to reduce the share of long-

tenured workers who are likely to receive the employment protection in their firms. However, they cannot reduce the share of long-tenured workers immediately because discharging long-tenured workers who are currently employed would be costly due to the protection. Instead, employers can prevent medium-tenured workers who are less likely to be protected by the law at that time from accumulating substantial job tenure and being eligible for the protection in the future. In order to prevent medium-tenured workers from accumulating substantial job tenure, employers would substitute medium-tenured workers with newly hired employees because they still need a certain number of workers. Consequently, the advancement of medium-tenured workers into long-tenured ones would be reduced over time due to the substitution of medium-tenured workers with newly hired employees by cost-minimizing employers.

There are at least two testable implications of this hypothesis. First, if the implied-contract exception induce employers to substitute medium-tenured workers with newly hired employees, the incidence of short-tenured jobs would become more common after the doctrine is recognized. In order to test whether the data support this prediction, I estimate equation 1.2 by using an indicator for jobs with five years or less tenure as a dependent variable. Again, when estimating the equation, I restrict my sample to workers aged at least 25 years old, because workers younger than those ages cannot accumulate effectively substantial job tenure.

Table 1.7 shows the estimation results. Although the implied-contract exception could reduce the incidence of short-tenured jobs because employment protection in general is known to reduce new hiring (Bentolila and Bertola, 1990[17]; Hopenhayn and Rogerson, 1993[59]),

Table 1.7: Logit Estimation Results: Effects on Pr[5 years or less]

Dependent variable:	Specifications			
Indicator for jobs with 5 years or less tenure	(1)	(2)	(3)	(4)
Implied-contract exception	.048	.089***	.072*	.092***
(Standard error)	(.053)	(.034)	(.037)	(.036)
[Marginal effect]	[.012]	[.022]	[.018]	[.023]
Covariates:				
Individual characteristics	Yes	Yes	Yes	Yes
State labor market conditions	No	Yes	No	Yes
State and year dummies	Yes	Yes	Yes	Yes
Region by year dummies	No	No	Yes	Yes
Age restriction	≥ 25	≥ 25	≥ 25	≥ 25
Pseudo R-squared	.106	.106	.107	.107
Number of observations	154797	154797	154797	154797

Note. Dependent variable is an indicator for jobs with 5 years or less current employer tenure. Regressor of interest is an indicator that takes 1 if the implied contract exception is in place and 0 otherwise. Individual characteristics include age, age squared, three dummies for education attainment: high school graduate, some college, more than college graduate where the omitted category is less than high school graduate, dummies for nonwhite, Hispanic origin, and an interaction between nonwhite and Hispanic. State labor market conditions include employment-population ratio, the ratio of manufacturing employment to non-manufacturing employment, and union coverage rate. Standard errors are clustered at state level. All marginal effects are evaluated at the mean values of the other covariates. The F-test for joint significance of the region-by-year dummies rejects the null hypothesis of equal year fixed effects across different geographic regions at one percent of significance level in columns 3 and 4. p<.01 (***), p<.05 (**), p<.10 (*)

the estimated coefficients on the law indicator in table 1.7 are all positive, supporting the hypothesis that employers would react to the legal intervention by substituting medium-tenured jobs with short-tenured ones. When state labor market conditions are controlled in columns 2 and 4, the incidence of short jobs with five years or less tenure increased by about two percentage points due to the protection.

Although the estimation results in tables 1.6 and 1.7 show that the incidence of long jobs became less common, while that of short jobs became more common following the adoption of

the implied-contract exception, it might take time for employers to adjust their employment policy to a new legal circumstance and substitute long-tenured jobs with short-tenured ones. In order to explore the temporal pattern of the impact of the law, I employ the event study model by Jacobson et al. (1993)[63]. Particularly, I estimated:

$$\Pr\left[C_1 \leq TEN_{ist} \leq C_2\right] = \Lambda\left(\eta_0 + \rho_s + \omega_t + \sum_{\tau=-4}^{8} \delta^{\tau} ICE_{s,t-\tau} + X_{it}\eta_1\right), \tag{1.3}$$

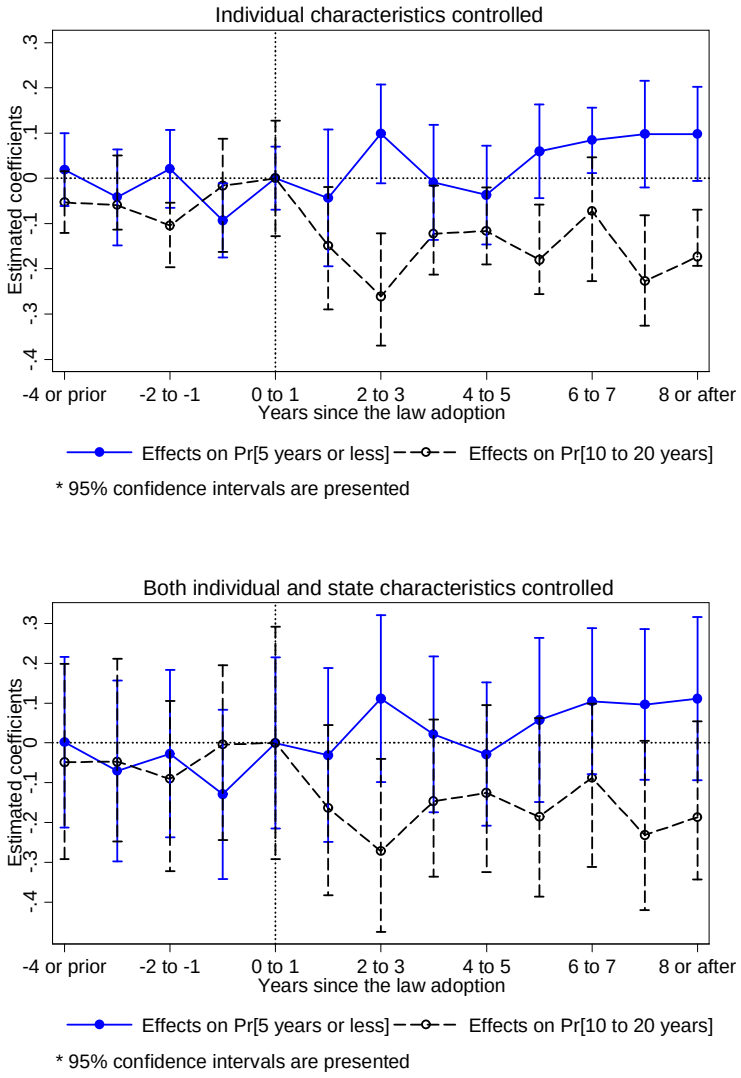
where I use two dependent variables, $\Pr\left[TEN_{ist} \leq 5\right]$ and $\Pr\left[10 \leq TEN_{ist} \leq 20\right]$; ρ_s and ω_t denote state and year fixed effects, respectively; $ICE_{s,t-\tau}$ is a set of event time dummies that take the value of 1 if state s adopts the implied-contract doctrine in year $t-\tau$ and 0 otherwise; and X_{it} is a vector of covariates. The set of parameters of interest, δ^{τ}s, captures the effects on the probability of long jobs with 10 to 20 years of tenure (when $\Pr\left[10 \leq TEN_{ist} \leq 20\right]$ is used as a dependent variable) or the effects on the probability of short jobs with five years or less tenure (when $\Pr\left[TEN_{ist} \leq 5\right]$ is used as a dependent variable) when τ years have passed since the law adoption. For example, $\tau \in \{-1, 0, 1\}$ represents the period between the 12th month before the law adoption and the month prior to the law adoption ($\tau = -1$), the period between the month of the law adoption and the 11th month since the law adoption ($\tau = 0$), the period between the 12th month and the 23rd month since the law adoption ($\tau = 1$), respectively. By comparing the estimates of δ^{τ}s across different values of τ, one can understand the temporal pattern of the law's effects on the incidence of the long and short jobs over time. I estimate equation 1.3 by a binary logit model using age restrictions of at least 35 and 25 years old for the dependent variables of $\Pr\left[10 \leq TEN_{ist} \leq 20\right]$ and

$\Pr[TEN_{ist} \leq 5]$, respectively.

Figure 1-5 plots the estimated $\widehat{\delta}^{\tau}$s in equation 1.3 against 12 event time slots, four leads and eight lags, each of which is spaced by 12 months from one another. When plotting the estimates, I normalize the coefficients on event time dummy for the year of law adoption ($\widehat{\delta}^{0}$) to be zero. In panel A of figure 1-5, I control for individual characteristics (X_{it}) such as age, education, race, and ethnicity as well as both state and year fixed effects. In panel B of figure 1-5, I add covariates related to state-level labor market conditions (Z_{st}) such as employment-to-population ratio, the ratio of manufacturing employment to non-manufacturing employment, and union coverage rate of each state in each year to the list of the covariates. Although the error bounds become larger with the addition of the state-level covariates, overall patterns of the estimates are similar across both panels of figure 1-5. The temporal patterns of the law's effects on the longer jobs and the shorter ones are overlapped with each other until the year of the law adoption, while they start diverging from each other since then. Perhaps more importantly, the gap between the incidence of the longer jobs and the shorter ones appear to be wider over time, suggesting that tenure distribution gradually shifted from long-tenured jobs toward short-tenured ones over time following the law adoption.

Another testable implication of my hypothesis is that, if the reason why the incidence of jobs with 20 years or longer tenure was not affected by the law (as shown in table 1.5) is because workers who already had such long years of tenure at the time of the law adoption could receive the protection, then it should follow that those workers could accumulate job tenure more easily than their peers in the states without the protection. This implies, in the

Figure 1-5: Time Patterns of the Law Impacts: Long Jobs vs. Short Jobs



Note. Individual characteristics include age, age squared, three dummies for education attainment: high school graduate, some college, more than college graduate where the omitted category is less than high school graduate, dummies for nonwhite, Hispanic origin, and an interaction between nonwhite and Hispanic. State characteristics include employment-population ratio, the ratio of manufacturing employment to non-manufacturing employment, and union coverage rate. Standard errors are clustered at state level.

fullness of time, the incidence of jobs with very long years of tenure would be more common in the adopting states than the non-adopting states. In order to test whether the data support this prediction, I estimate the logit model in equation 1.2 by using an indicator for jobs with 30 years or longer tenure as a dependent variable. When estimating the equation, I add nine more recent years (2000-2008) to the 1978-1999 sample because the sample period of 1978-1999 is not long enough to detect any potential increase in the incidence of jobs with 30 years or longer tenure under my hypothesis. In particular, workers who already had 20 years of tenure when the law was introduced could accumulate 30 years of tenure starting about a decade after the year of law adoption. Given that the latest year of the law adoption in my sample is 1988, my 1978-1999 sample might not include a large enough number of post-treatment observations to detect any increase in the incidence of 30 years or longer jobs. I also restrict my sample to workers aged 55 or older because those younger than the age cannot accumulate substantial years of job tenure.

Table 1.8 shows the estimation results. It reports that the estimated effects of recognizing the implied contract exception on the incidence of jobs with 30 years or longer tenure are all positive, suggesting that the observed zero effect of the implied-contract exception on the incidence of jobs with 20 years or longer tenure in table 1.5 might be because workers who had already accumulated 20 years or longer tenure at the time of the law adoption were protected by the law. However, the estimates are too imprecise to support this interpretation firmly. More importantly, since all the states in my sample adopted the implied-contract exception between 1982 and 1988, the estimates in table 1.8 implies that observations from more than two decades before and after the law adoption are used to make a pre-post comparison,

Table 1.8: Effects on Pr[30 years or longer] for 1978-2008 Sample

Dependent variable:	Specifications			
Indicator for jobs lasting at least 30 years	(1)	(2)	(3)	(4)
Implied-contract exception	.069	.095	.137*	.146*
(Standard error)	(.091)	(.090)	(.072)	(.079)
[Marginal effect]	[.010]	[.014]	[.014]	[.021]
Covariates:				
Individual characteristics	Yes	Yes	Yes	Yes
State labor market conditions	No	Yes	No	Yes
State and year dummies	Yes	Yes	Yes	Yes
Region by year dummies	No	No	Yes	Yes
Age restriction	≥ 55	≥ 55	≥ 55	≥ 55
Pseudo R-squared	.037	.038	.045	.045
Number of observations	32822	32822	32822	32822

Note. The CPS data on job tenure over extended period of time (1978-2008) are used for estimation. Dependent variable is an indicator for jobs with 30 years or longer current employer tenure. Regressor of interest is an indicator that takes 1 if the implied contract exception is in place and 0 otherwise. Individual characteristics include age, age squared, three dummies for education attainment: high school graduate, some college, more than college graduate where the omitted category is less than high school graduate, dummies for nonwhite, Hispanic origin, and an interaction between nonwhite and Hispanic. State labor market conditions include employment-population ratio, the ratio of manufacturing employment to non-manufacturing employment, and union coverage rate. Standard errors are clustered at state level. All marginal effects are evaluated at the mean values of the other covariates. The F-test for joint significance of the region-by-year dummies rejects the null hypothesis of equal year fixed effects across different geographic regions at one percent of significance level in columns 3 and 4. p<.01 (***), p<.05 (**), p<.10 (*)

which is not convincing. Given these limitations, I take the estimation results in table 1.8 as tentative evidence for my claim that the reason why the incidence of jobs with 20 year or longer tenure are not affected by the law is presumably because workers who already had such long years of tenure at the time of the law adoption were protected by the law and thus could more easily accumulate job tenure further relative to their peers in states without the protection.

1.6.2 Distributional difference-in-differences

The estimation results so far indicate that the implied-contract exception exerted differential impacts on different parts of the tenure distribution. The law reduced the incidence of long-term employment relationships, defined as jobs with 10 years or longer tenure, by about three percentage points. However, the incidence of very long jobs lasting beyond 20 years was rarely affected. On the other hand, the law increased the incidence of short-tenured jobs, defined as jobs with five years or less tenure by about two percentage points. Consequently, average tenure was reduced by about four percent. A reasonable interpretation of these results is that the law incentivized employers to gradually reduce the share of long-tenured jobs in their firms and utilize short-tenured ones more than they used to. To the extent that such substitution occurs over time, the duration of long-tenured jobs would become shorter, while that of short-tenured ones would become longer. In the end, the overall effects of the law on the tenure distribution would be negative on the upper part of the distribution, while the effects would be positive on the lower part of the distribution.

In order to test whether the data support this prediction, I attempt to estimate the causal effect of the implied-contract exception on the overall tenure distribution. Note that my sample consists of 30 states that recognized the implied-contract exception between 1982 and 1988 and seven states that have never adopted the exception. Hence, none of the 37 states in my sample adopted the implied-contract exception during 1978-1981, while the exception was in place in all the 30 adopting states but not in the seven non-adopting states during 1989-1999. For subsequent use, consider the implied-contract exception as a treatment; the 30 adopting states as a treatment group; the 7 non-adopting states as a

Table 1.9: Comparison of Tenure Distributions: Adopting States vs. Non-Adopting States

	Control group			Treatment group		
	Pre	Post	Post-Pre	Pre	Post	Post-Pre
10th percentile	0.51	0.52	0.01	0.53	0.52	-0.01
25th percentile	1.53	1.34	-0.18	1.60	1.38	-0.22
median	4.60	4.06	-0.54	4.86	4.08	-0.78
mean	7.99	7.17	-0.82	8.45	7.23	-1.17
75th percentile	11.26	10.10	-1.16	12.13	10.22	-1.92
90th percentile	21.56	19.61	-1.94	23.20	19.73	-3.47

Note. Control group represents 7 states that never adopted the implied-contract exception. Treatment group denotes 30 states that adopted the implied-contract exception between 1982 and 1988. Pre-treatment period is 1978-1981. Post-treatment period is 1989-1999.

control group; the years 1978-1981 as period 1 (pre-treatment period); and the years 1989-1999 as period 2 (post-treatment period). I assess the distributional effects of the law on job tenure by looking at how a tenure distribution of the treatment group changed between period 1 and period 2 relative to that of the control group.

Table 1.9 compares the changes in tenure distributions of the treatment and control groups between period 1 and period 2. Tenure distributions shifted to the left in both groups, however, the shift is slightly more pronounced among the treatment group. Of course, a simple comparison of tenure distributions between the treatment and control groups would not reflect any causal effect of the law adoption. In order to assess the causal effect, it needs to identify and estimate a *counterfactual* tenure distribution of the treatment group if the implied-contract exception had not been adopted in those states. Comparing the counterfactual tenure distribution with a realized tenure distribution of the treatment group would yield the causal effect of the legal intervention on the overall tenure distribution of the adopting states.

A question that naturally follows is how to identify such a counterfactual distribution. Recently, Bonhomme and Sauder (2011)[20] provided conditions under which the counterfactual distribution can be identified. In order to apply their econometric model to my study, I begin by estimating the following equation by the OLS method:

$$TEN_{ist} = \alpha_s + \gamma_t + X_{it}\beta + Z_{st}\pi + \varepsilon_{ist}, \tag{1.4}$$

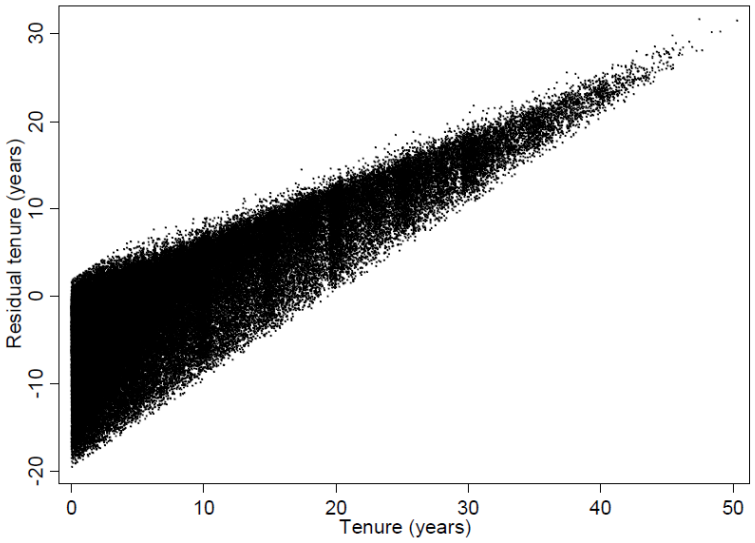
where TEN_{ist} denotes years of job tenure of worker i in state s in year t; α_s and γ_t denotes state and year fixed effects; X_{it} represents individual-level covariates of worker i in year t such as age, age squared, categorized educational attainment, dummies for race, Hispanic ethnicity, and their interaction; Z_{st} stands for state-level covariates such as employment-to-population ratio, the ratio of manufacturing employment to non-manufacturing employment, and union coverage rate of state s in year t.

Given the OLS estimates from equation 1.4, I calculate

$$U_{ist} = TEN_{ist} - \widehat{\gamma}_t - X_{it}\widehat{\beta} - Z_{st}\widehat{\pi}, \tag{1.5}$$

where $\widehat{\gamma}_t$, $\widehat{\beta}$, and $\widehat{\pi}$ are the OLS estimated coefficients of equation 1.4 for survey year dummies, covariates on individual characteristics, and covariates on state labor market conditions. By considering the residual U_{ist} as a new dependent variable, I control for any unwanted effects of those covariates. Figure 1-6 shows that the residual tenure (U_{ist}) is strongly correlated with the tenure (TEN_{ist}), which suggests that the effects of the law adoption on the tenure distribution would be similar to those for the residual tenure distribution.

Figure 1-6: Scatter Plot between Tenure and Residual Tenure



I re-label the time subscript of U_{ist} in equation 1.5 in the following way:

$$U_{ist} = U_{is1} \text{ if } t \text{ belongs to period 1 (pre-treatment, 1978-1981)}$$

$$U_{ist} = U_{is2} \text{ if } t \text{ belongs to period 2 (post-treatment, 1989-1999)}$$

Following the potential outcome framework by Rubin (1974)[93], let U^1_{is2} (or U^0_{is2}) denote potential years of the residual tenure of worker i in state s in period 2 if the implied-contract exception had been (or had not been) in place in the state. The object of interest is the quantile treatment effects on the treated (QTT) at each τth percentile point of the residual tenure distribution:

$$QTT(\tau) = F^{-1}_{U_{is2}|D_s=1}(\tau) - F^{-1}_{U^0_{is2}|D_s=1}(\tau), \ \tau \in (0,1), \tag{1.6}$$

where $F^{-1}_W(\cdot)$ denotes the inverse of the cumulative distribution function (CDF) of a random

variable W; D_s denotes an indicator variable that takes 1 if state s belongs to the treatment group and 0 otherwise. Since the distribution of $U_{is2}|D_s = 1$ is empirically observable, it is straightforward to estimate $F^{-1}_{U_{is2}|D_s=1}(\cdot)$ nonparametrically. The main challenge is how to identify $F^{-1}_{U^0_{is2}|D_s=1}(\cdot)$, which is counterfactual and hence unobservable from the data. In order to identify the counterfactual distribution of $U^0_{is2}|D_s = 1$, I employ the econometric model developed by Bonhomme and Sauder (2011)[20]. The following presentation of the model is heavily drawn from their paper.

I begin by modeling U_{is1} and U^0_{is2} as the sum of three components:

$$U_{is1} = \alpha_1 + \eta_s + v_{is1} \tag{1.7}$$

$$U^0_{is2} = \alpha^0_2 + \eta_s + v^0_{is2}, \tag{1.8}$$

where α_1 and α^0_2 are scalar parameters; η_s denotes state fixed effects that are likely to be correlated with D_s; v_{is1} and v^0_{is2} are idiosyncratic error terms that are allowed to correlate with each other in an arbitrary way.

I make two independence assumptions on equations 1.7 and 1.8.

Assumption 1.1 (DID assumption) v_{is1} and v^0_{is2} are independent of D_s.

Assumption 1.1 is analogous to the key identification assumption of the standard difference-in-differences (DID) model. The standard DID model assumes that treatment and control groups share common time-varying unobservables in outcome, hence such confounders can be differenced out by taking difference between pre-treatment and post-treatment means within each group. Analogously, assumption 1.1 requires that treatment ($D_s = 1$) and con-

trol ($D_s = 0$) groups share common time-varying unobservables (v_{is1} and v^0_{is2}) in the residual tenure.

Assumption 1.2 (FE assumption) v_{is1} and v^0_{is2} are independent of η_s given D_s.

Assumption 1.2 requires that v_{is1} and v^0_{is2} are independent of state fixed effects, η_s, within each of the treatment and control groups. Note that η_s is still allowed to be correlated with treatment indicator (D_s). In this respect, η_s is analogous to a fixed effect in the standard linear panel regression in that it is uncorrelated with time varying innovations (v_{is1} and v^0_{is2}) but may be correlated with a regressor of interest (D_s) (Bonhomme and Sauder, 2011[20]).

Both assumption 1.1 and assumption 1.2 impose certain restrictions on the time-varying unobservable factors (v_{is1} and v^0_{is2}) of the residual tenure. Assumption 1.1 requires that such factors are balanced between the treatment and control groups, while assumption 1.2 requires that they are also balanced across different states within each of the treatment and control groups. Given that job tenure depends on labor market conditions which are likely to differ across states, these assumptions seem to be unappealing. However, the residual tenure has been calculated by filtering out the effects of state-level labor market conditions such as employment, industrial composition, and unionization level from job tenure. To the extent that these state-level covariates account for potential state-specific trends in job stability, assumptions 1.1 and 1.2 may not be restrictive for my application.

Given the additive structure of the residual tenure in equations 1.7 and 1.8 and the two independence assumptions (assumption 1.1 and assumption 1.2), the probability density function (PDF) of the counterfactual $U^0_{is2}|D_s = 1$ is identified as (Corollary 1 of Bonhomme

and Sauder, 2011[20]):

$$f_{U^0_{is2}|D_s=1}(u) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\exp(-jhu) \frac{E[\exp(jhU_{is1})|D_s=1]}{E[\exp(jhU_{is1})|D_s=1]} E[\exp(jhU_{is2})|D_s=0]\,dh, \tag{1.9}$$

where $h \in \mathbf{R}$ and $j = \sqrt{-1}$. Details on this identification result are in the appendix. I estimate equation 1.9 with

$$\widehat{f}_{U^0_{is2}|D_s=1}(u) = \frac{1}{2\pi}\int_{-T_N}^{T_N}\exp(-jhu) \frac{\frac{1}{N_1}\sum_{i,D_s=1}\exp(jhU_{is1})}{\frac{1}{N_0}\sum_{i,D_s=0}\exp(jhU_{is1})} \frac{1}{N_0}\sum_{i,D_s=0}\exp(jhU_{is2})\,dh, \tag{1.10}$$

where N_1 and N_0 denote number of observations of the treatment and control groups, respectively; T_N is a trimming parameter for the integration in equation 1.10. Following Bonhomme and Sauder (2011)[20], I choose T_N by a rule of thumb method suggested by Diggle and Hall (1993)[36]. Details on this procedure are in the appendix.

Panel A of figure 1-7 plots the estimated PDFs of the *realized* residual tenure of the treatment group in period 2 ($U_{is2}|D_s = 1$; solid line) and the *counterfactual* residual tenure of the same group in the same period *if the implied-contract exception had not been in place* ($U^0_{is2}|D_s = 1$; dashed line) that is calculated from equation 1.10. Since I compare a *realized* distribution with a *counterfactual* one, any difference between the two distributions can be interpreted as a causal effect of the implied-contract exception. Although the estimated counterfactual density entails some bumpiness, comparing overall shape of the two distri-

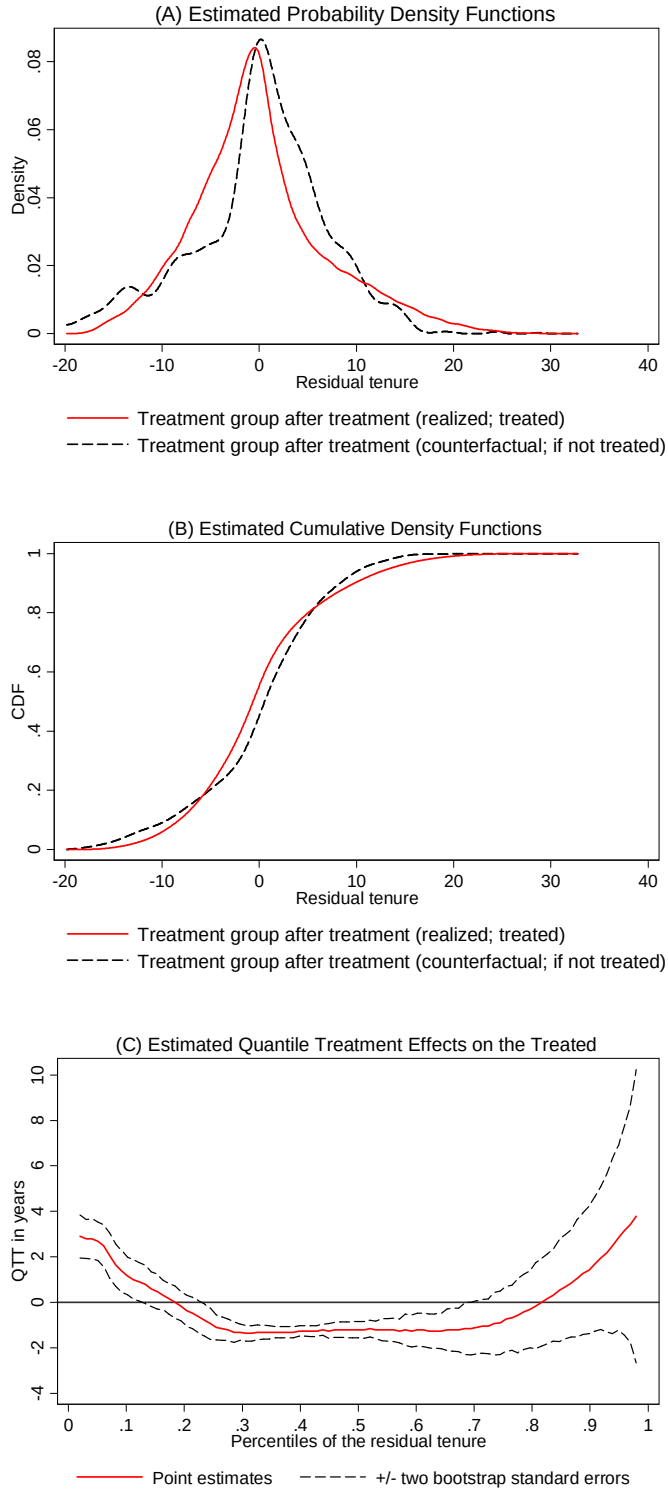
butions reveals that the implied-contract exception shifts the center of the residual tenure distribution slightly to the left.

Panel B of figure 1-7 shows the corresponding cumulative distribution functions (CDFs) calculated from integrating the estimated densities in panel A. The estimated CDFs demonstrate that the implied-contract exception increased both upper and lower parts of the distribution, while reducing the middle part of the distribution. In panel C, I invert the estimated CDFs in panel B and calculate the quantile treatment effects on the treated (QTT) given in equation 1.6. The result shows a clear U-shaped pattern of the quantile treatment effects: raising both upper and lower parts of the residual tenure distribution, while reducing the middle part of the distribution. In order to investigate statistical significance of the estimated quantile treatment effects, I also calculate error bounds by two times of bootstrap standard errors with 500 iterations. Following Hall (1992)[53] and Horowitz (2001)[61], when estimating the bootstrap confidence interval, I use a larger[40] trimming parameter (i.e. undersmoothing) than the one chosen to compute the point estimates in equation 1.10. The negative effects on the middle of the distribution and the positive effects on the lower decile of the distribution are statistically significant, while the effects on the upper tail are statistically insignificant.

Although the results for the residual tenure distribution cannot be directly translated to the effects on the tenure distribution, the results have some implications. As figure 1-6 suggests, the residual tenure is strongly correlated with job tenure in my sample. Hence, one can reasonably predict that the quantile treatment effects of the law adoption on the tenure

[40] I choose to use $T_N = 7.2$ for computing the bootstrap confidence interval, which is four times larger than $T_N = 1.8$ that I use to compute the density estimation.

Figure 1-7: Estimated Distributional Effects of the Implied-Contract Exception on Residual Tenure



distribution would be similar to those for the residual tenure distribution. This implies that the implied-contract exception is likely to increase the lower part of the tenure distribution but also decrease the upper part of the distribution, which is predicted to happen if employers substitute longer-tenured jobs with shorter-tenured ones in order to minimize employment cost under the implied-contract exception. In this respect, I take the estimation results in figure 1-7 as additional evidence supporting my substitution hypothesis.

1.7 Conclusion

Exploiting the cross-state and over-time variation in adopting the implied-contract exception to the conventional employment-at-will doctrine in the U.S. common law system, I examine the causal effect of the protection on mean tenure and the incidence of long-term employment relationships in the U.S. labor market during 1978 to 1999. I find that the implied-contract exception decreased mean tenure by about 4 percent and the probability of being in long-term employment relationships, defined as jobs lasting longer than 10 years, by about 3 percentage points. However, I find no evidence that the incidence of jobs with 20 years or longer tenure was affected by the law.

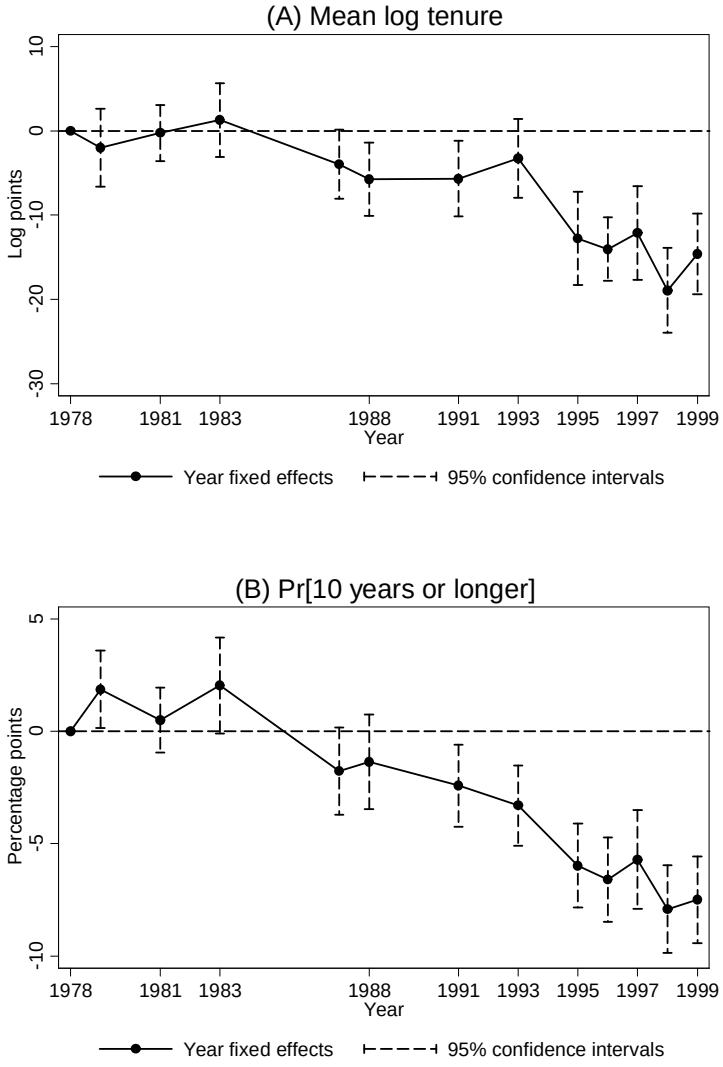
The observed labor market shift away from long-term employment relationships seems to have been driven by cost-minimizing employers' reaction to the legal intervention by gradually replacing long-tenured jobs, which are likely to be protected by the law, with short-tenured ones which are less likely to be protected. I observe that, as soon as the law had been introduced, long-tenured jobs became less and less common over time, while short-

tenured ones became more and more common. Comparing tenure distributions of states that adopted the implied-contract exception with that of non-adopting states across pre-adoption and post-adoption periods, I also find evidence that the law changed the tenure distribution by increasing the duration of short-tenured jobs but also decreasing that of longer-tenured ones, which is predicted to happen if long-tenured jobs were substituted with short-tenured ones.

In order to calculate how much my primary findings can account for the observed decline in job tenure, figure 1-8 depicts the time series of mean tenure and the incidence of long-term employment relationships during 1978 to 1999 conditioning on workers' demographic characteristics. Panel A of figure 1-8 indicates that mean tenure declined by about 20 percent between 1978 and 1999. Panel B of figure 1-8 shows that the probability of being in long-term employment relationships with at least 10 years of tenure decreased by about 8 percentage points over the same period. This indicates that, roughly speaking, the estimated negative four percent of the law impact on mean tenure accounts for about 20 (=4/20) percent of the observed decline in the average tenure. Also, the estimated negative three percentage points of the law's effect on the incidence of long-term employment relationships accounts for about 37.5 (=3/8) percent of the observed decline in the incidence of the long-term employment relationships.

Based on these results, I conclude that the observed decline in job stability in the U.S. can be, at least partly, explained by the changes in legal environments during the 1980s, an institutional factor which has rarely been discussed by the previous literature. It seems that the implied-contract exception changed the tenure distribution from one state where long-

Figure 1-8: Decline in Mean Tenure and the Incidence of Long-Term Employment Relationships



Note. Panel A plots estimated coefficients of year dummies from the OLS regression of the logarithm of years of job tenure on survey year dummies, age, age squared, categorized education dummies, and dummies for nonwhite, Hispanic origin, and their interaction. Panel B plots estimated marginal effects of survey year dummies from a logit regression of an indicator for jobs with at least 10 years of tenure on the same covariates used for the OLS regression in panel A. The marginal effects are evaluated at the mean values of the other covariates.

term employment relationships are more common to another state where such relationships are less common and employers rely on short-tenured jobs more than they used to. Such findings also add to a line of research on employment protection (Acemoglu and Angrist, 2001[3]; Autor, 2003[10]; Yoo and Kang, 2012[110]) in that employment protection could decrease employment of its protected worker group by increasing firing costs of those workers.

Chapter 2

Does Religious Participation Cause Better Health Outcomes?

2.1 Introduction

Does religious participation bring about healthier lives? It has been widely discussed in various fields of social science that religious participation may improve health conditions[1]. The logic is that religious involvement helps people have better psychological conditions, more healthy behaviors, and more social connections, which in turn would enhance their health conditions as well (Koenig et al., 2012[68][2]). For example, people commonly seek religious belief to relieve anxiety and stress in their daily lives. Risky behaviors such as drug use, dangerous sexual activity, cigarette smoking, overconsumption of alcohol or food,

[1] See Ellison and Levin (1998)[39] and Koenig et al. (2012)[68] for a survey of this literature.

[2] Koenig et al. (2012)[68] conceptualized potential channels through which religious participation can affect physical health as psychological, behavioral, and social pathways.

and suicide are discouraged or even prohibited in most religions. Religion also encourages people to have strong social connections by emphasizing virtues such as love, forgiveness, altruism, gratefulness, friendship, and marital stability. To the extent that these influences have impacts on health conditions, religious participation would lead to healthier lives.

This hypothesis is supported by a large amount of empirical studies reporting positive correlations between religious participation and various measures of health conditions. Hummer et al. (1999)[62] analyzed the National Health Interview Survey and found that those who never attend a church service tend to display higher risks of mortality than those who regularly attend a church service, even after controlling for demographic, socioeconomic, and baseline health conditions. Koenig et al. (2012)[68] presented an extensive survey on empirical studies reporting evidence of potential health benefits of religious participation. Counting studies that used a frequency of attending a religious service as a measure of the extent of religious participation and were published in last 10 years among their list of studies, people with a higher level of religious participation tend to face lower risks of various diseases such as heart disease (Lutgendorf et al., 2004[76]; Obisesan et al., 2006[86]; Schnall et al., 2010[95]; Horne et al., 2008[60]; Buzazeri et al., 2008[24]), hypertension (Gillum and Ingram, 2006[47]; Yeager et al., 2006[109]; Gorman and Sivaganesan, 2007[49]), stroke (Obisesan et al., 2006[86]), cognitive impairment (Van Ness and Kasl, 2003[107]; Koenig et al., 2004[69]; Yearger et al., 2006[109]; Murray-Swank et al., 2007[83]; Reyes-Ortiz et al., 2008[91]; Corsentino et al., 2009[30]), immune-related disorders (Sephton et al., 2001[96]; Lutgendorf et al., 2004[76]), cancer (Oman et al., 2002[87]; Kinney et al., 2003[67]), physical disability (Benjamins, 2004[16]; Reyes-Ortiz et al., 2006[14]; Park et al., 2008[89]), and somatic symptoms

(Krause and Ellison, 2003[70]).

Most of these studies use the regression-control strategy, which basically compares health conditions of people who attend a religious service regularly with that of those who do not after controlling for potential confounders such as demographic information, socioeconomic variables, and baseline health measures (for studies using panel data). Although the regression-control strategy is a useful way of summarizing a statistical relationship between religious participation and health outcomes, it is still open to question as to whether and to what extent the observed correlations reflect a causal relationship. In order to interpret the correlations as causal relationships, it is necessary to assume that the levels of religious participation are as good as randomly assigned across different people conditional on their observable characteristics, which is often called selection-on-observables assumption. This assumption is hard to justify because religious people are likely to be different from those who are not in many ways not explained by only observable information. For example, it is possible that people with healthier life styles are also those who attend a religious service more frequently. At the same time, it is also possible that people suffering from serious diseases are more likely to have a religious belief to cope with the diseases. One might argue that controlling for more covariates could mediate this problem. However, it is known that adding too many control variables could exacerbate any potential selectivity bias (DiNardo and Lee, 2010[38]).

This essay aims to test whether or not the observed correlations between religious participation and health conditions reflect a causal relationship. Using data from the General Social Survey between 1972 and 2001, which is virtually the only source of data that collects

information on religious participation and health conditions of nationally representative U.S. adults over several decades, I find that the extent of religious participation of the respondents, measured by their frequencies of attending a church service, are strongly correlated with various measures of their health conditions, as having been reported in the previous religion-health studies. Specifically, I observe that people who attend a church service at least once a week are more likely to report excellent health conditions, experience less days of poor health and/or activity limitation, face a lower risk of mortality, and live longer than those who never attend a church service.

In order to figure out whether the observed strong correlations reflect causal relationships, I use the repeals of the government regulations in the U.S., the so-called "blue laws", as a source of exogenous variation in religious participation. The blue laws refer to laws that restrict certain secular activities such as labor or commerce on Sundays in order to enforce the observance of Sunday as a day of worship (Gruber and Hungerman, 2008[50]). The blue laws had been widely implemented in the U.S. until the 1950s, while most states repealed the laws starting in the early 1960s (Goos, 2005[48]; Gruber and Hungerman, 2008[50]). Gruber and Hungerman (2008)[50] found that religious participation, measured by the frequency of church attendance among Christians, fell significantly after the repeals of the blue laws because the repeals allowed people to work or go shopping on Sundays rather than attending church services.

If religious participation were indeed causally related with health conditions, the decreased levels of religious participation due to the repeals of the blue laws would lead to decreased levels of health conditions as well. To test this hypothesis, I estimate the impacts

of the repeals of the blue laws on religious participation of respondents and their health conditions using a difference-in-differences framework. I find that the law repeals decreased the probability of attending a church service at least once a week by about 6 to 8 percentage points, which accounts for about 1/5 to 1/4 of the predicted probability that an average respondent would attend a church service with such a frequency. In spite of the substantial impacts of the blue laws on religious participation of respondents, I do not find any evidence that the law repeals are associated with their health conditions, which would be observed if religious participation were indeed causally related with health conditions. Taking these reduced-form results as indirect evidence suggesting that there is no causal relationship between religious participation and health conditions, I also directly estimate a causal relationship between the two by using the blue law status of the GSS respondents' states of residence as an instrumental variable (IV) for their levels of religious participation. When the potential endogeneity of the levels of religious participation of the respondents is corrected by the IV strategy, the strong correlations between religious participation and health conditions are no longer observed. Based on these findings, I conclude that the previous studies reporting strong relationships between religious participation and health outcomes are likely to have been driven by selectivity bias rather than reflecting causal relationships.

2.2 Previous literature

Earlier literature on the effects of the blue laws interpreted the repeals of the blue laws as deregulation on shop opening hours. Tullock (1975)[105] predicted that forcing stores to close

on Sundays will reduce operating costs of the stores more than it will reduce sales revenue of them, which will result in a temporary rise in profits. However, the increase in profits will be competed away by new entrants over the long run. At equilibrium, the number of stores will increase and the profit levels of all the stores will be back to normal. The prediction by Tullock, however, was challenged by Moorehouse (1984)[80] which reported that the blue laws reduced number of stores in a given market. Goos (2005)[48] also found that the blue laws decreased number of stores as well as employment and revenue of the stores.

More recent literature interpreted the repeals of the blue laws as an increase in the opportunity cost of religious participation. Gruber and Hungerman (2008)[50] found that the blue law repeals reduced religious participation. Gerber et al. (2008)[46] argued that religious participation enhances voter turnout by showing that the repeals of the blue laws are associated with decline in religious participation as well as voter turnout. Cohen-zada and Sander (2011)[29] used the blue laws as an IV for religious participation and found that religious participation has a positive effect on happiness of women.

2.3 Data

The data of this study are from the General Social Survey (GSS). The GSS is an ongoing cross-sectional survey conducted by the National Opinion Research Center (NORC). In almost every year since 1972, the GSS has been asking demographic, behavioral, and attitudinal questions to about 1,500 to 3,000 nationally representative individuals in the U.S.

Table 2.1: Sample of States and Timing of the Blue Law Repeals

Year of Repeal	States	# of States
1955	Iowa	1
1965	Kansas	1
1966	Washington	1
1969	Florida	1
1973	Ohio, Utah	2
1975	Virginia	1
1977	Indiana, South Dakota	2
1978	Pennsylvania	1
1981	Tennessee	1
1982	Vermont	1
1985	Minnesota, South Carolina, Texas	3
1991	North Dakota	1
Never had laws	Arizona, California, Colorado, Idaho Nevada, New Mexico, Oregon, Wyoming	8
	Total	24

Note. The source of the legal classification is Gruber and Hungerman (2008)[21]. The other 26 states are excluded from the sample because (1) the information on the blue laws are not available (20 states: Alaska, Alabama, Arkansas, Connecticut, Delaware, Georgia, Hawaii, Kentucky, Louisiana, Maryland, Michigan, Mississippi, Missouri, Montana, Nebraska, New Hampshire, New Jersey, North Carolina, Oklahoma, and Rhode Island) and (2) it is unclear to locate the exact timing of the repeals because the laws were gradually changed over time (six states: Illinois, Massachusetts, Maine, New York, West Virginia, and Wisconsin) (Gruber and Hungerman, 2008) [21].

Table 2.2: Summary Statistics

Variables:	N	Mean	S.D
Church attendance:			
Never (yes=1)	17,735	.105	.307
Less than once a year (yes=1)	17,735	.080	.272
Once a year (yes=1)	17,735	.132	.339
Several times a year (yes=1)	17,735	.132	.338
Once a month (yes=1)	17,735	.077	.267
Two to three times a month (yes=1)	17,735	.097	.296
Nearly every week (yes=1)	17,735	.062	.241
Every week (yes=1)	17,735	.225	.418
More than once a week (yes=1)	17,735	.088	.284
Linear index (Never = 1; More than once a week = 9)	17,735	5.19	2.61
Subjective health condition:			
Poor (yes=1)	13,660	.057	.232
Fair (yes=1)	13,660	.183	.386
Good (yes=1)	13,660	.445	.497
Excellent (yes=1)	13,660	.316	.465
Linear index (Poor = 1; Excellent = 4)	13,660	2.02	.854
Covariates:			
Age (years)	17,857	46.0	17.8
Female (yes=1)	17,912	.581	.493
Nonwhite (yes=1)	17,912	.133	.340
Years of education	17,873	12.5	3.05
Real income ($1,000; 1986 base)	16,354	29.9	25.7
Married (yes=1)	17,911	.581	.493
Living in the 100 largest metropolitan areas (yes=1)	17,912	.223	.416
Living in the 100 largest suburban areas (yes=1)	17,912	.245	.430
Employed (yes=1)	17,912	.603	.489
Unemployed (yes=1)	17,912	.025	.158
Out of labor force (yes=1)	17,912	.371	.483
Catholic (yes=1)	16,752	.255	.436
Fundamentalist (yes=1)	16,604	.353	.478

who are aged 18 or older[3]. A unique advantage of using the GSS data for this study is that the core part of its questionnaire contains questions on both religious participation and health conditions, which have been asked in a consistent way over several decades. Since its first round in 1972, the GSS has been asking its respondents how often they attend religious services during the survey year and how they would describe conditions of their own health. The nine possible answers to the question on religious participation are "never", "less than once a year", "once a year", "several times a year", "once a month", "two to three times a month", "nearly every week", "every week", and "more than once a week". The four possible responses to the question on health conditions are "excellent", "good", "fair", and "poor".

For the timings of each state's repeals of the blue laws, I refer to the legal classification by Gruber and Hungerman (2008)[50]. Following Gruber and Hungerman (2008)[50], I restrict my sample to those who live in 24 states where valid information on the timings of repealing the blue laws is available. The other 26 states that are excluded from the sample consist of 20 states[4] where the information on the blue laws are not available and six states[5] where the laws were changed gradually over time and hence it is unclear to locate the exact timing of the repeals (Gruber and Hungerman, 2008[50]).

Table 2.1 lists the 24 states used in this study. The first blue law repeal was in 1955 in Iowa, while the last one was in 1991 in North Dakota. Given that the last repeal occurred in 1991, I choose my sample period to be from 1972, the earliest year of the GSS data, to 2001[6],

[3]I refer to the GSS website (http://www3.norc.org/GSS+Website/) for this information.

[4]These states are Alaska, Alabama, Arkansas, Connecticut, Delaware, Georgia, Hawaii, Kentucky, Louisiana, Maryland, Michigan, Mississippi, Missouri, Montana, Nebraska, New Hampshire, New Jersey, North Carolina, Oklahoma, and Rhode Island.

[5]These states are Illinois, Massachusetts, Maine, New York, West Virginia, and Wisconsin.

[6]In fact, my sample includes observations between 1972 and 2000 because the GSS was not conducted in

a decade after the last repeal took place in North Dakota, because a pre-post comparison beyond a decade after repealing the blue laws would not be useful much. Consequently, my sample includes 10 states where the blue laws were repealed during my sample period (1972-2001), four states where the laws had already been repealed before the earliest year of my sample (1972), and eight states that never had such laws.

Following the previous studies (Gruber and Hungerman, 2008[50]; Gerber et al., 2008[50]; Cohen-Zada and Sander, 2012[29]), I also restrict my sample to those who are Catholic or Protestant because they are those whose behaviors are most likely to be affected by the blue laws. These people are about 86.25 percent of the 1972-2001 GSS sample. Table 2.2 summarizes descriptive statistics of the sample.

2.4 Empirical analysis

2.4.1 Correlation between religious participation and health condition

In order to confirm whether the strong positive correlations between religious participation and health conditions reported in the previous literature are also observed in my sample, I begin by examining a cross-sectional relationship between the two variables using the 1972-2001 GSS data. Recall that the GSS measures a respondent's level of religious participation as their frequency of church attendance which consists of nine ordinal categories: "never", "less than once a year", "once a year", "several times a year", "once a month", "two to

2001.

three times a month", "nearly every week", "every week", and "more than once a week". Using this information, I create a new ordinal variable that has four categories. Specifically, I choose those who answer "never" as a reference group; combine the responses of "less than once a year", "once a year", and "several times a year" into a new category named *"infrequently"*; merge the responses of "once a month", "two to three times a month", and "nearly every week" into another category and name it as *"occasionally"*; and group the responses of "every week" and "more than once a week" into the other category defined as *"weekly or more often"*. Consequently, the new categorical variable on church attendance has four levels of "*never*", "*infrequently*", "*occasionally*", and "*weekly or more often*".

Table 2.3 shows a cross tabulation between religious participation and health conditions of the GSS respondents. In the table, panel A reports proportion of each level of health condition at a given level of religious participation (row proportion), while panel B shows proportion of each level of religious participation at a given level of health condition (column proportion). Both panels indicate that religious participation is positively correlated with health conditions, markedly at the highest and lowest levels of the two variables. In panel A, the fraction of respondents reporting an excellent health condition is highest for those who go to church weekly or more often (34.3%) but lowest for those who never go to church (25.8%). On the contrary, the fraction of respondents reporting a poor health condition is lowest for the weekly-churchgoers (4.96%) but highest for the non-churchgoers (10.9%). Similarly, in panel B, the fraction of the weekly churchgoers is highest for those reporting an excellent health condition (34.0%) but lowest for those reporting a poor health condition (27.3%). By contrast, the fraction of the non-churchgoers is lowest for those with excellent

Table 2.3: Cross Tabulation between Religious Participation and Subjective Health Condition

	Subjective health condition				
	Poor	Fair	Good	Excellent	Total
A. Row proportion (%)					
Church attendance					
Never	10.9	19.7	43.6	25.8	100
Infrequently	5.17	18.3	46.2	30.4	100
Occasionally	5.07	19.8	43.0	32.2	100
Weekly or more often	4.96	16.9	43.8	34.3	100
Total	5.69	18.3	44.4	31.5	100
B. Column proportion (%)					
Church attendance					
Never	20.2	11.3	10.3	8.61	10.5
Infrequently	31.6	34.6	36.1	33.4	34.7
Occasionally	20.9	25.3	22.7	23.9	23.5
Weekly or more often	27.3	28.9	30.9	34.0	31.3
Total	100	100	100	100	100

Note. "Never" refers to those who do not go to church at all. "Infrequently" refers to those who go to church but less than once a month. "Occasionally" refers to those who go to church at least once a month but not every week. "Weekly or more often" refers to those who go to church at least once a week.

health conditions (8.61%) but highest for those with poor health conditions (10.9%). A chi-square test of independence also rejects its null hypothesis of independence at 1% level of significance. The value of the chi-square test statistics is about 118, which greatly exceeds the critical value of 21.7.

Of course, people who attend church services more frequently are not directly comparable with those who attend services less often because both groups are different from each other in their demographic and socioeconomic characteristics. For example, Azzi and Ehrenberg (1975)[15] summarized empirical regularities that those who are female, older, high-income,

black, and residents in rural areas are more likely to attend a religious service than those who are not. In order to control for these demographic and socioeconomic characteristics that are likely to be correlated with religious participation, I estimate the following ordered probit model:

$$\Pr\left[Health_{it} = j\right] = \Phi\left(\lambda^{j} < AttendDummies_{it}\beta_1 + X_{it}\beta_2 < \lambda^{j+1} \right), \tag{2.1}$$

where $Health_{it} \in \{1, 2, 3, 4\}$ denotes a subjective current health condition of individual i in year t coded as "poor (=1)", "fair (=2)", "good (=3)", and "excellent (=4)"; $AttendDummies_{it}$ is a vector of dummy variables indicating a frequency of church attendance of individual i in year t, which consists of three dummies for "infrequently", "occasionally", and "weekly or more often" where the omitted category is "never"; X_{it} denotes a vector of individual characteristics such as categorical age, race, gender, years of education, real income, marital status, type of localities in which respondents live, and dummies for Catholic and fundamentalist of individual i in year t; $\lambda = \left\{\lambda^1, \lambda^2, \lambda^3, \lambda^4, \lambda^5\right\}$ is a set of threshold parameters with $\lambda_1 = -\infty$ and $\lambda_5 = \infty$; and $\Phi(\cdot)$ is the cumulative distribution function (c.d.f.) of the standard normal distribution.

Table 2.4 shows the ordered probit estimation results of equation 2.1[7]. In column 1, I control for individual characteristics such as five-year age dummies, a dummy for female, a dummy for nonwhite, years of education, household real income (1986 base), a dummy for being married, dummies for living in the 100 largest standard metropolitan statistical areas

[7]In 1982 and 1987, the GSS oversampled black people. In order to adjust for the oversampling of black people, all estimates in this paper are weighted in a way that establishes a nationally representative sample.

(SMSA) and the 100 largest suburban areas, and dummies for Catholic and fundamentalist. The result shows that the estimated coefficients of the church attendance dummies are all positive and increasing in magnitude as the frequency of church attendance goes up, suggesting that those who attend a church service more often tend to report a better health condition. In column 2, I add a set of survey year dummies and state of residence dummies in order to control for any unobserved heterogeneity across different survey samples and geographic regions. There is a line of literature arguing that self-reported health conditions may have measurement errors correlated with labor market status of respondents (Anderson and Burkhauser, 1985[6]; Butler et. al., 1987[25]; Bound, 1991[21]). In order to control for these influences, I also include dummies for being employed and unemployed, where the omitted category is being out of the labor force, into the list of covariates in column 3. Controlling for these additional covariates does not change the results much.

Although the estimation result in table 2.4 shows a positive correlation between religious participation and health condition, interpreting the magnitudes of the estimated coefficients needs an extra work. In table 2.5, I calculate predicted probabilities and marginal effects for each of the four values of $Health_{it}$. All the predicted probabilities and marginal effects are evaluated at the mean values of the covariates. I find that the correlations between religious participation and health conditions are particularly substantial for the "excellent" and "poor" health conditions. People who go to church at least once a week are about 10 percentage points more likely to report an "excellent" health condition than those who never go to church. On the other hand, the weekly churchgoers are about 2 percentage points less likely to report a "poor" health condition than the non-churchgoers. These probability gaps

Table 2.4: Ordered Probit Estimation Results: Religious Participation and Subjective Health Condition

Dependent variable:	Specifications		
Subjective health condition	(1)	(2)	(3)
Church attendance			
Infrequently	.057	.065	.042
(Standard error)	(.037)	(.037)	(.037)
Occasionally	.108***	.118***	.099**
(Standard error)	(.039)	(.039)	(.039)
Weekly or more often	.271***	.280***	.266***
(Standard error)	(.038)	(.039)	(.039)
Covariates:			
Individual characteristics	Yes	Yes	Yes
Year and state dummies	No	Yes	Yes
Labor market status	No	No	Yes
Log Likelihood	-13630	-13593	-13501
Number of observations	12255	12255	12255

Note. Dependent variable is subjective current health conditions coded as "poor (=1)", "fair (=2)", "good (=3)" and "excellent (=4)". The omitted category for the church attendance is "never". Covariates include (1) individual-level characteristics: five-year age dummies, years of education, real household income (USD 1,000; base year: 1986), and dummies for nonwhite, female, Catholic, and fundamentalist; (2) survey year dummies and state of residence dummies; and (3) labor market status dummies: dummies for employed and unemployed where the reference group is those who are out of the labor force. $p<.01$ (***), $p<.05$ (**), $p<.10$ (*)

Table 2.5: Predicted Probabilities and Marginal Effects after the Ordered Probit Estimation in Table 2.4

Dependent variable:	Specifications		
Subjective health condition	(1)	(2)	(3)
A. Subjective health condition = excellent			
Predicted probability	.303	.303	.301
Marginal effects			
Infrequently	.020 (.013)	.023 (.013)	.015 (.013)
Occasionally	.038*** (.014)	.042*** (.014)	.035** (.014)
Weekly or more often	.097*** (.014)	.100*** (.014)	.095*** (.014)
B. Subjective health condition = good			
Predicted probability	.486	.488	.491
Marginal effects			
Infrequently	-.004 (.002)	-.004* (.003)	-.003 (.002)
Occasionally	-.008** (.003)	-.009*** (.003)	-.007** (.003)
Weekly or more often	-.022*** (.004)	-.023*** (.004)	-.022*** (.004)
C. Subjective health condition = fair			
Predicted probability	.176	.175	.176
Marginal effects			
Infrequently	-.012 (.008)	-.014 (.008)	-.009 (.008)
Occasionally	-.022*** (.008)	-.025*** (.008)	-.021*** (.008)
Weekly or more often	-.056*** (.008)	-.058*** (.008)	-.055*** (.008)
D. Subjective health condition = poor			
Predicted probability	.035	.034	.033
Marginal effects			
Infrequently	-.004 (.003)	-.005* (.003)	-.003 (.003)
Occasionally	-.008*** (.003)	-.008*** (.003)	-.007*** (.003)
Weekly or more often	-.019*** (.003)	-.019*** (.003)	-.018*** (.002)
Covariates:			
Individual characteristics	Yes	Yes	Yes
Year and state dummies	No	Yes	Yes
Labor market status	No	No	Yes

Note. All the predicted probabilities and the marginal effects are evaluated at mean values of the covariates. The omitted category for the church attendance is "never". Standard errors are in parenthesis. p<.01 (***), p<.05 (**), p<.10 (*)

in reporting "excellent" and "poor" health conditions between the weekly churchgoers and the non-churchgoers amount to roughly 1/3 and 1/2 of the predicted probabilities that an average respondent would describe their own health conditions as "excellent" and "poor", respectively.

The estimation results in tables 2.4 and 2.5 suggest that self-reported current health conditions of respondents are strongly correlated with their levels of religious participation. However, it is also possible that the observed correlations between the two variables simply result from the subjective nature of $Health_{it}$ used in equation 2.1. Kahneman and Krueger (2006)[65] argued that individuals' responses to a survey question on subjective well-being may vary with various circumstances surrounding the respondents. For example, people who attend a religious service regularly may have different understandings of what an "excellent health condition" means in the survey question than those who never attend a religious service. If such a reporting error in $Health_{it}$ is correlated with church attendance behaviors of respondents, a positive correlation between the two variables would be observed even in the absence of any underlying relationship between religious participation and health condition.

There are several alternative health variables in the GSS which measure health conditions of respondents in more concrete and objective ways than $Health_{it}$ that I mainly use in this study. In 2002 and 2006, the GSS asked its respondents to count the number of days during which they had poor health conditions or activity limitations within a month before their surveys. Unfortunately, I cannot use these alternative health variables for my main analysis because they are only available for the years of 2002 and 2006 which are out of my sample period (1972-2001). However, I can still use the information to assess whether and to

what extent $Health_{it}$ is a reliable measure of health conditions of respondents by looking at how $Health_{it}$ correlates with those alternative, and also more objectively measured, health variables. For this purpose, I estimate the following equation by the ordinary least square (OLS) method:

$$Days_{it} = HealthDummies_{it}\gamma_1 + X_{it}\gamma_2 + \varepsilon_{it}, \tag{2.2}$$

where $Days_{it}$ denotes number of days of poor health condition (or activity limitation) during the past 30 days of individual i in year t; $HealthDummies_{it}$ is a vector of dummy variables indicating subjective current health conditions of individual i in year t, which consists of three dummies for "fair", "good", and "excellent" where the omitted category is "poor"; and X_{it} denotes a vector of the covariates used in equation 2.1.

Although $Days_{it}$ in equation 2.2 is more objectively measured than $Health_{it}$ in equation 2.4, they are both self-reported measures and hence are not perfectly free from the issue of subjectivity. Death records, which are likely to convey some information on health conditions of the deceased while they were alive, do not have such a problem. In 2011, National Opinion Research Center (NORC), the survey conductor of the GSS, linked 18 waves of the GSS (1978-2002) to death certificate data from 1979 to 2008 (Muennig et al., 2011[81]). The resulting 2011 General Social Survey - National Death Index (GSS-NDI) data provides information on whether the 1978-2002 GSS respondents had been dead as of 2008 and, if dead, the timings and causes of the death. Using the 1978-2002 GSS-NDI data, I construct two outcome variables: an indicator variable for the vital status as of 2008 of respondent i in year t and their ages at death. When constructing these variables, I exclude observations

with external causes of death such as accidents, homicides, legal intervention, or injury from war operations because those death records are not likely to reflect health conditions of the deceased while they were alive. To the extent that healthier people have a lower mortality risk and longer life expectancy, the vital status of the respondents and their ages at death can be used as alternative health variables which do not have any subjectivity problem[8].

Using the 2011 GSS-NDI data, I estimate the following probit model:

$$\Pr\left[Death_{it} = 1\right] = \Phi\left(HealthDummies_{it}\delta_1 + X_{it}\delta_2 + \upsilon_{it}\right), \tag{2.3}$$

where $Death_{it}$ is an indicator variable that takes 1 if respondent i in year t is dead as of 2008 or 0 otherwise; $HealthDummies_{it}$ and X_{it} are subjective health dummies and individual characteristics used in equation 2.2.

Similarly, I estimate the following equation:

$$AgeDeath^*_{it} = HealthDummies_{it}\eta_1 + X_{it}\eta_2 + \nu_{it}, \tag{2.4}$$

where $AgeDeath^*_{it}$ is the age at death of respondent i in year t. Note that $AgeDeath^*_{it}$ is only observable for those who are dead as of 2008. For those who are alive as of 2008, all I can observe is that their $AgeDeath^*_{it}$, which has not been realized as of 2008, are greater than their ages in 2008, denoted as $Age2008_{it}$. Thus, I estimate equation 2.4 by the censored regression model using $AgeDeath_{it} = \min\left\{AgeDeath^*_{it},\ Age2008_{it}\right\}$ as a dependent variable

[8]The vital status and ages at death of the GSS respondents may not be ideal for this study because they are available as early as 1978 and hence do not have observations from most years in the 1970s during which the blue laws were most actively repealed. For this reason, I use these variables as supplementary outcome variables to the subjective health condition that I mainly use for this study.

and assuming the normality of the error term (ν_{it}).

The first three columns of tables 2.6, 2.7, 2.8, and 2.9 summarize the estimation results of equations 2.2 (tables 2.6 and 2.7), 2.3 (table 2.8), and 2.4 (table 2.9). The results commonly show that people reporting higher levels of subjective health conditions tend to display better health conditions in terms of the objective health variables as well, suggesting that $Health_{it}$ is a reliable measure of the true health conditions of the respondents. For example, people who describe their current health conditions as "excellent" tend to experience less poor health conditions by about 16 days within a month (table 2.6); less activity limitations by about 8 days within a month (table 2.7); are about 13 percentage point less likely to die between their survey years and 2008 (table 2.8); and live longer by about 5 years (table 2.9) than those who describe their health conditions as "poor".

In the next three columns of the tables, I also examine how the objective health variables are associated with church attendance behaviors of respondents. If church attendance are not associated with the objectively measured health variables, it would be doubtful whether the observed correlations between subjective health conditions and church attendance indeed reflect a real health gap between churchgoers and non-churchgoers. For this purpose, I estimate equations 2.2, 2.3, and 2.4 by using dummies for church attendance as regressors of interest instead of dummies for subjective health conditions. The estimation results show that regular churchgoers tend to have a better health condition than non-churchgoers in terms of the objective health measures as well. For example, people who attend church services at least once a week tend to experience less poor health days by about 1.3 days within a month; less activity limitations by about a day within a month; are about 4 percentage points less

Table 2.6: OLS Estimation Results: Subjective Health Conditions, Religious Participation, and Number of Poor Health Days

Dependent variable:	Specifications					
Days of poor health	(1)	(2)	(3)	(4)	(5)	(6)
Subjective health						
Fair	-10.6***	-10.8***	-10.8***			
(Standard error)	(2.16)	(2.08)	(2.08)			
Good	-14.2***	-14.5***	-14.5***			
(Standard error)	(2.10)	(2.02)	(2.02)			
Excellent	-15.7***	-16.0***	-16.0***			
(Standard error)	(2.09)	(2.02)	(2.02)			
Church attendance						
Infrequently				-.738	-.682	-.682
(Standard error)				(.456)	(.463)	(.463)
Occasionally				-1.36***	-1.29***	-1.29***
(Standard error)				(.463)	(.470)	(.470)
Weekly or more				-1.37***	-1.34***	-1.34***
(Standard error)				(.473)	(.477)	(.477)
Covariates:						
Individual characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Year and state dummies	No	Yes	Yes	No	Yes	Yes
Labor market status	No	No	Yes	No	No	Yes
R-squared	.168	.198	.198	.025	.049	.049
# of observations	2004	2004	2004	3042	3042	3042

Note. Dependent variable is the number of poor health days within a month prior to survey dates. The omitted categories for the subjective health condition and the church attendance are "poor" and "never", respectively. Covariates include (1) individual characteristics: five-year age dummies, years of education, real household income (USD 1,000; base year: 1986), and dummies for nonwhite, female, Catholic, and fundamentalist; (2) survey year dummies and state of residence dummies; and (3) labor market status dummies: dummies for employed and unemployed where the reference group is those who are out of the labor force. Observations from states where blue law information is not available are added in order to improve precision of estimates. p<.01 (***), p<.05 (**), p<.10 (*)

Table 2.7: OLS Estimation Results: Subjective Health Conditions, Religious Participation, and Number of Activity Limitation Days

Dependent variable: Days of activity limitation	Specifications (1)	(2)	(3)	(4)	(5)	(6)
Subjective health						
Fair	-5.89***	-5.96***	-5.96***			
(Standard error)	(1.95)	(1.90)	(1.90)			
Good	-7.28***	-7.41***	-7.41***			
(Standard error)	(1.92)	(1.87)	(1.87)			
Excellent	-8.01***	-8.12***	-8.12***			
(Standard error)	(1.92)	(1.87)	(1.87)			
Church attendance						
Infrequently				-.246	-.180	-.180
(Standard error)				(.341)	(.349)	(.349)
Occasionally				-.434	-.388	-.388
(Standard error)				(.352)	(.364)	(.364)
Weekly or more				-.909***	-.830**	-.830**
(Standard error)				(.343)	(.351)	(.351)
Covariates:						
Individual characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Year and state dummies	No	Yes	Yes	No	Yes	Yes
Labor market status	No	No	Yes	No	No	Yes
R-squared	.075	.099	.099	.016	.032	.032
# of observations	2004	2004	2004	3042	3042	3042

Note. Dependent variable is the number of activity limitation days within a month prior to survey dates. The omitted categories for the subjective health condition and the church attendance are "poor" and "never", respectively. Covariates include (1) individual characteristics: five-year age dummies, years of education, real household income (USD 1,000; base year: 1986), and dummies for nonwhite, female, Catholic, and fundamentalist; (2) survey year dummies and state of residence dummies; and (3) labor market status dummies: dummies for employed and unemployed where the reference group is those who are out of the labor force. Observations from states where blue law information is not available are added in order to improve precision of estimates. $p<.01$ (***), $p<.05$ (**), $p<.10$ (*)

Table 2.8: Probit Estimation Results: Subjective Health Condition, Mortality Risk, and Religious Participation

Dependent variable:	Specifications					
Dead as of 2008	(1)	(2)	(3)	(4)	(5)	(6)
Subjective health						
Fair	-.281***	-.269***	-.247***			
(Standard error)	(.053)	(.057)	(.057)			
[Marginal effect]	[-.083]	[-.076]	[-.070]			
Good	-.444***	-.422***	-.391***			
(Standard error)	(.051)	(.054)	(.055)			
[Marginal effect]	[-.138]	[-.125]	[-.116]			
Excellent	-.469***	-.495***	-.462***			
(Standard error)	(.053)	(.057)	(.058)			
[Marginal effect]	[-.138]	[-.138]	[-.129]			
Church attendance						
Infrequently				.023	-.030	-.020
(Standard error)				(.034)	(.036)	(.036)
[Marginal effect]				[.008]	[-.009]	[-.006]
Occasionally				.009	-.063*	-.054
(Standard error)				(.036)	(.038)	(.038)
[Marginal effect]				[.003]	[-.020]	[-.017]
Weekly or more				-.028	-.138***	-.130***
(Standard error)				(.035)	(.037)	(.037)
[Marginal effect]				[-.009]	[-.043]	[-.041]
Covariates:						
Individual characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Year and state dummies	No	Yes	Yes	No	Yes	Yes
Labor market status	No	No	Yes	No	No	Yes
Log likelihood	-7879	-7204	-7196	-11292	-10356	-10342
# of observations	16454	16454	16454	22819	22819	22819

Note. Dependent variable is a dummy for vital status as of 2008 that takes 1 if dead and 0 otherwise. The omitted categories for the subjective health condition and the church attendance are "poor" and "never", respectively. Covariates include (1) individual characteristics: five-year age dummies, years of education, real household income (USD 1,000; base year: 1986), and dummies for nonwhite, female, Catholic, and fundamentalist; (2) survey year dummies and state of residence dummies; and (3) labor market status dummies: dummies for employed and unemployed where the reference group is those who are out of the labor force. Observations with external causes of death such as accidents, homicides, legal intervention, or injury from war operations are excluded from the sample, while observations from states where blue law information is not available are added in order to improve precision of estimates. All marginal effects are calculated at mean values of the other covariates. p<.01 (***), p<.05 (**), p<.10 (*)

Table 2.9: Censored Regression Results: Subjective Health Condition, Life Expectancy, and Religious Participation

Dependent variable:	Specifications					
Age at Death	(1)	(2)	(3)	(4)	(5)	(6)
Subjective health						
Fair	2.81***	2.80***	2.61***			
(Standard error)	(.452)	(.434)	(.434)			
Good	4.23***	4.21***	3.93***			
(Standard error)	(.439)	(.421)	(.424)			
Excellent	5.11***	4.98***	4.68***			
(Standard error)	(.476)	(.457)	(.460)			
Church attendance						
Infrequently				.337	.158	.050
(Standard error)				(.338)	(.323)	(.323)
Occasionally				.764**	.530	.424
(Standard error)				(.355)	(.340)	(.341)
Weekly or more often				1.52***	1.23***	1.13***
(Standard error)				(.343)	(.330)	(.331)
Covariates:						
Individual characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Year and state dummies	No	Yes	Yes	No	Yes	Yes
Labor market status	No	No	Yes	No	No	Yes
Log likelihood	-21263	-21171	-21162	-31362	-31208	-31187
# of censored observations	11789	11789	11789	15969	15969	15969
# of observations	16442	16442	16442	22803	22803	22803

Note. Dependent variable is age at death which is right-censored at age in 2008 for those who were alive as of 2008. The omitted categories for the subjective health condition and the church attendance are "poor" and "never", respectively. Covariates include (1) individual characteristics: five-year age dummies, years of education, real household income (USD 1,000; base year: 1986), and dummies for nonwhite, female, Catholic, and fundamentalist; (2) survey year dummies and state of residence dummies; and (3) labor market status dummies: dummies for employed and unemployed where the reference group is those who are out of the labor force. Observations with external causes of death such as accidents, homicides, legal intervention, or injury from war operations are excluded from the sample, while observations from states where blue law information is not available are added in order to improve precision of estimates. $p<.01$ (***), $p<.05$ (**), $p<.10$ (*)

likely to die between their survey years and 2008; and live longer by about 1.1 years than those who never go to church.

Collectively, the estimation results in tables 2.6, 2.7, 2.8, and 2.9 support that $Health_{it}$ contains some important information of the true health conditions of respondents and the observed correlations between $Health_{it}$ and $AttendDummies_{it}$ reported in tables 2.4 and 2.5 are likely to reflect real difference in health conditions across people with varying levels of religious participation rather than to have been driven by potential reporting errors in $Health_{it}$ that might be correlated with church attendance behaviors of respondents. As having been widely discussed in the previous religion-health studies, religious participation seems to have health benefits regardless of the choice of different health measures. In the following sections, I discuss whether and to what extent the observed health benefits from religious participation can be attributable to a causal effect.

2.4.2 Effects of the blue laws

In order to figure out whether the observed strong correlation between religious participation and health conditions reflect a causal relationship as discussed in the previous studies, I exploit the repeals of the blue laws as a source of exogenous variation in the opportunity cost of religious participation. Previous studies on the blue laws found that the repeals of the blue laws reduced a frequency of religious participation significantly by relaxing regulations against secular activities on Sundays and hence increasing the opportunity cost of attending a church service (Gruber and Hungerman, 2008[50], Gerber et al., 2008[46], Cohen-Zada and Sander, 2011[29]). In order to confirm whether such impacts are also observed in my sample,

I begin by estimating the following ordered probit model:

$$\Pr\left[Attend_{ist} = j\right] = \Phi\left(\omega_j < \pi_1 Law_{st} + X_{it}\pi_2 + \alpha_s + \eta_t < \omega_{j+1} \right), \tag{2.5}$$

where $Attend_{it} \in \{1, 2, 3, 4\}$ denotes ordered responses on the frequency of attending a church service of individual i in year t coded as "never (=1)", "infrequently (=2)", "occasionally (=3)", and "weekly or more often (=4)"; Law_{st} denotes an indicator variable that takes 1 if the blue law was in place in state s in year t and 0 otherwise; X_{it} denotes a vector of individual characteristics such as categorical age, race, gender, years of education, real income, marital status, type of localities in which respondents live, and dummies for Catholic and fundamentalist of individual i in year t; α_s represents state fixed effects (a set of state dummies); η_t denotes year fixed effects (a set of year dummies); $\omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5\}$ is a set of threshold parameters with $\omega_1 = -\infty$ and $\omega_5 = \infty$; and $\Phi(\cdot)$ is the cumulative distribution function (c.d.f.) of the standard normal distribution. When estimating equation 2.5, I exclude observations from the years during which the blue laws were repealed because it is not clear whether Law_{st} takes 1 or 0 for those observations.

Table 2.10 reports the estimation results of equation 2.5. In column 1, the estimated coefficient on Law_{st} is positive and statistically significant, suggesting that the blue laws induced people to attend a church service more frequently. A key identification assumption underlying the causal interpretation is that states under the blue laws and those without the laws share a common time trend in church attendance behaviors of their residents. If the blue laws were repealed in states where church attendance was in decline (relative to

Table 2.10: Orderded Probit Estimation Results: Effects of Blue Laws on Religious Participation

Dependent variable:	Specifications		
Church attendance	(1)	(2)	(3)
Blue law dummy	.160***	.206***	.232***
(Standard error)	(.036)	(.056)	(.066)
Covariates:			
Individual characteristics	Yes	Yes	Yes
Year and state dummies	Yes	Yes	Yes
State-specific linear time trends	No	Yes	Yes
State-specific quadratic time trends	No	No	Yes
Log likelihood	-19688	-19657	-19648
Number of observations	15666	15666	15666

Note. Dependent variable is the frequency of attending a church service, which is coded as "never (=1)", "infrequently (=2)", "occasionally (=3)" and "weekly or more often (=4)". Individual characteristics include five-year age dummies, years of education, real household income (USD 1,000; base year: 1986), and dummies for nonwhite, female, Catholic, and fundamentalist. Observations from the years during which the blue laws were repealed are excluded. p<.01 (***), p<.05 (**), p<.10 (*)

Table 2.11: Predicted Probabilities and Marginal Effects after the Orderded Probit Estimation in Table 2.10

Dependent variable:	Specifications		
Church attendance	(1)	(2)	(3)
A. Church attendance = weekly or more often			
Predicted probability	.299	.298	.298
Marginal effect of blue law	.057***	.074***	.084***
(Standard error)	(.013)	(.021)	(.025)
B. Church attendance = occasionally			
Predicted probability	.252	.253	.253
Marginal effect of blue law	.006***	.006***	.007***
(Standard error)	(.001)	(.001)	(.001)
C. Church attendance = infrequently			
Predicted probability	.359	.359	.360
Marginal effect of blue law	-.039***	-.050***	-.057***
(Standard error)	(.009)	(.014)	(.017)
D. Church attendance = never			
Predicted probability	.090	.089	.089
Marginal effect of blue law	-.024***	-.030***	-.033***
(Standard error)	(.005)	(.007)	(.008)
Covariates:			
Individual characteristics	Yes	Yes	Yes
Year and state dummies	Yes	Yes	Yes
State-specific linear time trends	No	Yes	Yes
State-specific quadratic time trends	No	No	Yes

Note. All predicted probabilities and marginal effects are calculated at mean values of the other covariates. p<.01 (***), p<.05 (**), p<.10 (*)

the non-repealing states) anyway, the observed correlation between the blue laws and the frequency of church attendance would be spurious rather than reflecting a causal relationship. Adding state-specific time trends in the frequency of church attendance to the list of controls would relax, although in a limited way, the arguably restrictive assumption that the blue-law and the non-blue-law states follow the same time trend in religious participation. In columns 2 and 3, I included state-specific linear and quadratic time trends in the frequency of church attendance as additional covariates. The estimated coefficients on Law_{st} become even larger with the additions of the time trends, indicating that the blue laws tend to have been repealed in states where there were increasing trends in religious participation.

In order to interpret the magnitudes of the estimated coefficients in table 2.10, I compute predicted probabilities for each of the four ordered responses on church attendance and marginal effects of Law_{st} for each of the four responses in table 2.11. All the predicted probabilities and marginal effects are calculated at the mean values of covariates. I find that the blue laws exerted substantial impacts on the highest-ordered ("weekly or more often") and lowest-ordered ("never") responses relative to the other two responses in the middle ("occasionally" and "infrequently"). In column 1, the marginal effect for the highest-ordered response, "weekly or more often", is .057, indicating that an average respondent in blue-law states is 5.7 percentage points more likely to attend a church service at least once a week relative to their peers in non-blue-law states. This is a substantial difference: given that the predicted probability for the outcome is .299, it follows that about 1/5 ($.057/.299 \approx .191$) of the predicted probability is attributable to the blue laws. Similarly, the marginal effect for the lowest-ordered outcome ("never") is -.024, which accounts for about 1/4 ($.024/.090 \approx .267$)

of the predicted probability for the response. On the other hand, the marginal effects for the other two outcomes – "occasionally" and "infrequently" – only account for about 1/50 (.006/.252 $\approx$ .024) and 1/10 (.039/.359 $\approx$ .106) of the predicted probabilities for those outcomes, respectively. A possible interpretation of these results is that the blue laws shifted the entire distribution of church attendance to the right by inducing people in each of the ordinal responses to move to the next higher level. For example, the law induced people who "never" go to church to move to the next higher level, "infrequently"; people who attend church services "infrequently" to attend "occasionally"; and those who go to church services "occasionally" to the highest "weekly or more often". In the end, the law would have a substantial positive effect on the highest response ("weekly or more often") and a substantial negative impact on the lowest one ("never"). On the other hand, the effects would be relatively smaller for the other two outcomes in the middle ("occasionally" and "infrequently") because the inflow from the nearby lower-level response and the outflow into the next higher-level one would offset each other[9].

The estimation results in tables 2.10 and 2.11 indicate that the blue laws increased levels of religious participation significantly. Hence, if there is indeed a causal link between religious participation and physical health as having been discussed in the previous studies, the blue laws would also improve health conditions through inducing people to attend a religious service more frequently. For example, attending a church service may help people manage anxiety and stress in their daily lives, refrain from risky behaviors such as drinking, smoking, and dangerous sexual intercourse, and have more social relationships, all of which

[9]Since the GSS sample that I use in this study is not longitudinal, it is not feasible to formally test the validity of this interpretation.

would in turn enhance their health conditions. In order to test whether the data support this prediction, I estimate the impacts of the blue laws on various health measures. In particular, I regress each of the subjective health conditions ($Health_{it}$), the vital status indicator ($Death_{it}$), and the age at death ($AgeDeath^*_{it}$) on the blue law indicator (Law_{st}) controlling for state and year fixed effects as well as individual characteristics (X_{it}). For estimation, I use ordered probit, univariate probit, and censored regression models for each of the dependent variables: $Health_{it}$, $Death_{it}$, and $AgeDeath^*_{it}$. As in the estimation of equation 2.5, I drop the observations from the years when the blue laws were repealed because it is unclear to determine whether the blue law indicator should take either 1 or 0 for those years.

Tables 2.12, 2.13, and 2.14 summarize the estimated reduced-form effects of the blue laws on health outcomes of respondents. For each of the tables, column 1 shows the results for the basic specification which controls for state and year dummies as well as individual characteristics. In columns 2 and 3, I include state-specific time trends as additional control variables to the basic model. If health were indeed causally related with religious participation, positive estimates of the coefficients on the blue law dummy would be observed because the blue laws would improve health outcomes through inducing people to participate in a church service more frequently as reported in tables 2.10 and 2.11. Regardless of the choice of the model specifications, however, I fail to reject that the estimated effects of the blue laws are statistically different from zero, suggesting that the hypothetical causal link between religious participation and physical health is not likely to exist.

The reduced-form estimation results in tables 2.12, 2.13, and 2.14 provide indirect evi-

Table 2.12: Orderded Probit Estimation Results: Effects of the Blue Laws on Subjective Health Condition

Dependent variable:	Specifications		
Subjective health condition	(1)	(2)	(3)
Blue law dummy	-.012	-.096	-.103
(Standard errors)	(.042)	(.068)	(.082)
Covariates:			
Individual characteristics	Yes	Yes	Yes
Year and state dummies	Yes	Yes	Yes
State-specific linear time trends	No	Yes	Yes
State-specific quadratic time trends	No	No	Yes
Log likelihood	-13433	-13417	-13404
Number of observations	12073	12073	12073

Note. Dependent variable is respondents' subjective health conditions coded as poor (=1), fair (=2), good (=3) and excellent (=4). Individual characteristics are five-year age dummies, years of education, real household income (USD1,000; base year: 1986), and dummies for nonwhite, female, Catholic, and fundamentalist. Observations from the years during which blue laws were repealed are excluded. Standard errors are in parenthesis. $p<.01$ (***), $p<.05$ (**), $p<.10$ (*)

Table 2.13: Predicted Probabilities and Marginal Effects after the Ordered Probit Estimation in Table 2.12

Dependent variable:	Specifications		
Subjective health condition	(1)	(2)	(3)
A. Subjective health condition = excellent			
Predicted probability	.304	.304	.303
Marginal effect of blue law	-.004	-.033	-.035
(Standard error)	(.015)	(.023)	(.027)
B. Subjective health condition = good			
Predicted probability	.486	.486	.487
Marginal effect of blue law	.001	.005	.005
(Standard error)	(002)	(.002)	(.003)
C. Subjective health condition = fair			
Predicted probability	.176	.176	.176
Marginal effect of blue law	.003	.021	.022
(Standard error)	(.009)	(.015)	(.018)
D. Subjective health condition = poor			
Predicted probability	.035	.034	.034
Marginal effect of blue law	.001	.008	.008
(Standard error)	(.003)	(.006)	(.007)
Covariates:			
Individual characteristics	Yes	Yes	Yes
Year and state dummies	Yes	Yes	Yes
State-specific linear time trends	No	Yes	Yes
State-specific quadratic time trends	No	No	Yes

Note. All predicted probabilities and marginal effects are calculated at mean values of the other covariates. p<.01 (***), p<.05 (**), p<.10 (*)

Table 2.14: Effect of Blue Laws on Vital Status and Life Expectancy

	Specifications		
	(1)	(2)	(3)
A. Dependent variable: vital status as of 2008 (death=1)			
Estimation method	Probit		
Blue law dummy	.033	.167	.185
(Standard error)	(.071)	(.111)	(.139)
[Marginal effect]	[.011]	[.055]	[.062]
Covariates:			
Individual characteristics	Yes	Yes	Yes
Year and state dummies	Yes	Yes	Yes
State-specific linear time trends	No	Yes	Yes
State-specific quadratic time trends	No	No	Yes
Log Likelihood	-5576	-5563	-5556
Number of observations	12098	12098	12098
B. Dependent variable: age at death (right-censored)			
Estimation method	Censored regression		
Blue law dummy	.016	-.940	-1.24
(Standard error)	(.648)	(.992)	(1.26)
Covariates:			
Individual characteristics	Yes	Yes	Yes
Year and state dummies	Yes	Yes	Yes
State-specific linear time trends	No	Yes	Yes
State-specific quadratic time trends	No	No	Yes
Log likelihood	-16589	-16578	-16570
Number of censored observations	8504	8504	8504
Number of observations	12087	12087	12087

Note. In panel A, the dependent variable is a dummy for vital status as of 2008 that takes 1 if dead and 0 otherwise. In panel B, the dependent variable is age at death which is right-censored at age in 2008 for those who were alive as of 2008. In both panels, individual characteristics include five-year age dummies, years of education, real household income (USD 1,000; base year: 1986), and dummies for nonwhite, female, Catholic, and fundamentalist. Observations with external causes of death such as accidents, homicides, legal intervention, or injury from war operations are excluded from the sample. $p<.01$ (***), $p<.05$ (**), $p<.10$ (*)

Table 2.15: Effects of Blue Laws on Labor Supply

	Specifications		
	(1)	(2)	(3)
A. Dependent variable: labor force participation (yes=1)			
Estimation method	Probit	Probit	Probit
Blue law dummy	-.015	-.037	-.009
(Standard error)	(.049)	(.074)	(.090)
[Marginal effect]	[-.005]	[-.013]	[-.003]
Log likelihood	-7778	-7760	-7740
Number of observations	15773	15773	15773
B. Dependent variable: employed (yes=1)			
Estimation method	Probit	Probit	Probit
Blue law dummy	.036	.085	.071
(Standard error)	(.048)	(.073)	(.087)
[Marginal effect]	[.013]	[.031]	[.026]
Log likelihood	-8289	-8273	-8252
Number of observations	15773	15773	15773
C. Dependent variable: usual working hours in a week			
Estimation method	Tobit	Tobit	Tobit
Blue law dummy	1.39	2.99*	2.29
(Standard error)	(1.05)	(1.59)	(1.87)
Log likelihood	-51357	-51342	-51322
Number of observations	15728	15728	15728
Covariates:			
Individual characteristics	Yes	Yes	Yes
Year and state dummies	Yes	Yes	Yes
State-specific linear time trends	No	Yes	Yes
State-specific quadratic time trends	No	No	Yes

Note. All predicted probabilities and marginal effects are evaluated at the means of the following covariates. Individual characteristics include age, years of education, real household income (USD1,000; base year: 1986), and dummies for nonwhite, female, Catholic, and fundamentalist. p<.01 (***), p<.05 (**), p<.10 (*)

Table 2.16: Probit Estimation Results: Effects of Blue Laws on Drinking and Smoking Behaviors

	Specifications		
	(1)	(2)	(3)
A. Dependent variable: ever drank (yes=1)			
Blue law dummy	.007	-.056	-.129
(Standard error)	(.076)	(.121)	(.143)
[Marginal effect]	[.002]	[-.019]	[-.044]
Log likelihood	-3524	-3508	-3499
Number of observations	6828	6828	6828
B. Dependent variable: ever drunk too much (yes=1)			
Blue law dummy	-.049	-.076	.036
(Standard error)	(.089)	(.139)	(.167)
[Marginal effect]	[-.018]	[-.028]	[.013]
Log likelihood	-2887	-2881	-2869
Number of observations	4731	4731	4731
C. Dependent variable: smoker (yes=1)			
Blue law dummy	-.029	.001	-.089
(Standard error)	(.072)	(.113)	(.135)
[Marginal effect]	[-.010]	[.000]	[-.031]
Log likelihood			
Number of observations	6826	6826	6826
D. Dependent variable: ever smoked (yes=1)			
Blue law dummy	-.139	-.252	-.199
(Standard error)	(.104)	(.168)	(.198)
[Marginal effect]	[-.048]	[-.085]	[-.068]
Log likelihood	-4129	-4119	-4112
Number of observations	4103	4103	4103
Covariates:			
Individual characteristics	Yes	Yes	Yes
Year and state dummies	Yes	Yes	Yes
State-specific linear time trends	No	Yes	Yes
State-specific quadratic time trends	No	No	Yes

Note. All predicted probabilities and marginal effects are evaluated at the means of the following covariates. Individual characteristics include age, years of education, real household income (USD1,000; base year: 1986), and dummies for nonwhite, female, Catholic, and fundamentalist. $p<.01$ (***), $p<.05$ (**), $p<.10$ (*)

dence suggesting that religious participation levels are not likely to be causally related with physical health. In order to find a direct relationship between religious participation and health conditions, I am going to use the blue law as a source of exogenous variation in levels of religious participation in the next section. A key identification assumption for the instrumental variable (IV) strategy is that the blue laws may affect health conditions but only through their impacts on religious participation, which is often called an exclusion restriction. Although the blue laws may affect health conditions through many other channels, the two most important channels would be through affecting individuals' labor supply and consumption behaviors. As mentioned in section 2.1, the blue laws were implemented in practice by prohibiting all kinds of labor and commerce on Sundays. Hence, the repeals of the blue laws are most likely to induce people to work and go shopping more among others. The potential changes in labor supply and consumption behaviors could affect health conditions directly. It has been widely discussed that working hours are closely related with health outcomes (see Sparks et al., 1997[99] for a review of this literature). Deregulation on shop-opening days may create higher demands for unhealthy products such as liquor and cigarettes as well as other products. To the extent these changes in labor and consumption behaviors driven by the repeals of the blue laws affect health conditions directly other than through affecting religious participation, the validity of the exclusion restriction would be questionable.

However, I find that the potential effects of the blue laws on physical health through affecting labor supply and consumption behaviors of respondents, if any, are not likely to be substantial enough to invalidate my use of the blue laws as an exogenous source of variation

in religious participation. Table 2.15 summarizes the impacts of the blue laws on labor supply decisions. For each of panels A, B, and C, I regress a dummy for labor force participation, a dummy for employment, and usual hours of work in a week, respectively, on the blue law dummy controlling for state and year dummies as well as individual characteristics. Regardless of the choice of model specifications, the estimated coefficients are not statistically different from zero, suggesting that the potential impacts of the blue laws on labor supply decisions are negligible. Similarly, table 2.16 shows how the blue laws affect drinking and smoking behaviors. Again, I find no evidence that the blue laws changed consumption behaviors of liquor and cigarettes.

2.5 Instrumental variable strategy

Given that the potential threats to the exclusion restriction are not a major concern, I attempt to estimate the causal effect of religious participation, measured by a frequency of attending a church service of respondents on their current health conditions by using the blue law as an IV for church attendance. Unfortunately, it is not straightforward to implement this strategy because both the dependent variable ($Health_{it}$) and the endogenous regressor ($Attend_{it}$) are ordinal responses while only a single dummy instrument (Law_{st}) is available. A common way to deal with an ordered endogenous regressor is to include a set of dummies each of which represents an ordered category of the regressor. However, since I have only one instrumental variable (Law_{st}), this would result in underidentification problem. In addition, it is necessary to use a nonlinear instrumental variable model to deal

with an ordinal dependent variable. However, it is very hard to precisely estimate a nonlinear instrumental variable model using a single binary instrumental variable in practice.

A simple way to avoid this econometric issue is to dichotomize both the dependent and the independent variables. This allows me to employ the bivariate probit model which has been discussed in Heckman (1978)[58], Evans and Schwab (1995)[40], Angrist (2001)[7], and Angrist and Pischke (2009)[8]. To apply this approach, I dichotomize both the outcome variable and the endogenous regressor of interest:

$$Y_{it} = \begin{Bmatrix} 1 \text{ if } Health_{it} = 4 \\ 0 \text{ if } Health_{it} < 4 \end{Bmatrix}, \tag{2.6}$$

$$D_{it} = \begin{Bmatrix} 1 \text{ if } Attend_{it} = 4 \\ 0 \text{ if } Attend_{it} < 4 \end{Bmatrix}, \tag{2.7}$$

where $Health_{it} \in \{1, 2, 3, 4\}$ denotes subjective health condition of individual i in year t coded as "poor (=1)", "fair (=2)", "good (=3)", and "excellent (=4)" and $Attend_{it} \in \{1, 2, 3, 4\}$ is an ordered response on the frequency of church attendance of individual i in year t coded as "never (=1)", "infrequently (=2)", "occasionally (=3)", and "weekly (=4)".

Following Agrist and Pischke (2009)[8][10], suppose the dichotomous outcome variable, an indicator for "excellent health condition", is determined in the following way:

$$Y_{it} = 1\left[\beta_1^* D_{it} + X_{ist}\beta_2^* > \varepsilon_{ist}\right], \tag{2.8}$$

[10] The illustration of the bivariate probit model in this section is mainly drawn from section 4.6.3 in Angrist and Pischke (2009)[8].

where X_{ist} denotes individual characteristics such as years of education, real income and dummies for five-year age categories, race, gender, marital status, type of localities in which respondents live, Catholic, fundamentalist, and state of residence and survey year of individual i in state s in year t; and ε_{ist} represents unobservable random components that might affect respondents' health conditions.

The potential outcomes defined by equation 2.8 are

$$Y_{it}^{0} = 1\left[X_{ist}\beta_{2}^{*} > \varepsilon_{ist}\right], \tag{2.9}$$

$$Y_{it}^{1} = 1\left[\beta_{1}^{*} + X_{ist}\beta_{2}^{*} > \varepsilon_{ist}\right], \tag{2.10}$$

where Y_{it}^{1} (or Y_{it}^{0}) denotes potential health condition of individual i in year t in the state of the world where the respondent were a weekly-or-more-frequently churchgoer (or not). The object of interest is the average treatment effect (ATE):

$$ATE = E\left[Y_{it}^{1} - Y_{it}^{0}\right] = E\left[1\left[\beta_{1}^{*} + X_{ist}\beta_{2}^{*} > \varepsilon_{ist}\right] - 1\left[X_{ist}\beta_{2}^{*} > \varepsilon_{ist}\right]\right], \tag{2.11}$$

In order to identify the ATE, the latent parameters β_{1}^{*} and β_{2}^{*} need to be identified. If D_{it} is uncorrelated with ε_{ist} conditional on X_{ist} in equation 2.8, the β_{1}^{*} and β_{2}^{*} can be consistently estimated by a univariate probit regression of Y_{it} on D_{it} and X_{ist} under the assumption that ε_{ist} follows the standard normal distribution. Due to the normality of ε_{ist}, the ATE in equation 2.11 can be computed as:

$$ATE = E\left[\Phi\left(\beta_{1}^{*} + X_{ist}\beta_{2}^{*}\right) - \Phi\left(X_{ist}\beta_{2}^{*}\right)\right] \tag{2.12}$$

where $\Phi(\cdot)$ denotes the c.d.f. of standard normal distribution.

However, D_{it} is likely to be correlated with ε_{ist} because people who attend church services at least once a week ($D_{it} = 1$) are likely to be different from those who do not ($D_{it} = 0$) in many ways not captured by the observable characteristics (X_{ist}), which may be related with their health conditions (Y_{it}). To allow for this possibility, I model a respondent's church attendance behavior as:

$$D_{it} = 1\left[\gamma_1^* Z_{st} + X_{ist}\gamma_2^* > \upsilon_{ist}\right], \tag{2.13}$$

where Z_{st} denotes the blue law dummy (equivalent to Law_{st} in equation 2.5) that takes 1 if the blue law was in place in state s in year t or 0 otherwise; and υ_{ist} denotes unobservable random components that might affect respondents' church attendance behaviors (D_{it}).

For identification, I assume that ε_{ist} and υ_{ist} follow the bivariate normal distribution with mean zero, unit variance, and correlation coefficient of ρ which may take a non-zero value. In this way, I allow the random components of health conditions (ε_{ist}) and church attendance (υ_{ist}) to be correlated with each other. In addition, I assume the random components are independent of Z_{st} conditional on X_{ist}. This condition can be understood as an exclusion restriction in that the instrumental variable (Z_{st}) can affect the outcome (Y_{it}) but only through affecting the endogenous regressor (D_{it}).

Given the exclusion restriction and the bivariate normality of the error terms, the latent parameters of equations 2.8 and 2.13 and the correlation coefficient ρ can be jointly estimated

by maximizing the following log likelihood function (Agrist and Pischke, 2009[8]):

$$\sum \left\{ \begin{array}{c} Y_{it} \log \left[\Phi_2 \left(\beta_1^* D_{it} + X_{ist} \beta_2^*, \gamma_1^* Z_{st} + X_{ist} \gamma_2^*, \rho \right) \right] \\ + (1 - Y_{it}) \log \left[1 - \Phi_2 \left(\beta_1^* D_{it} + X_{ist} \beta_2^*, \gamma_1^* Z_{st} + X_{ist} \gamma_2^*, \rho \right) \right] \end{array} \right\}, \tag{2.14}$$

where $\Phi_2 (\cdot, \cdot, \rho)$ denotes the bivariate normal distribution function with mean zero, unit variance, and correlation coefficient of ρ. Similarly to the univariate probit case, the ATE can be computed by substituting the estimated β_1^* and β_2^* into equation 2.12 by the virtue of the normality of ε_{ist}.

Table 2.17 summarizes the estimation results when an indicator for having an excellent health condition is used as an outcome variable. In columns 1 and 2, I estimate equation 2.8 by the univariate probit model ignoring the potential endogeneity of the regressor D_{it}. The estimated latent parameter, β_1^*, is positive and statistically significant and the estimated ATE[11] is about .05, indicating that those who attend religious services at least once a week are about five percentage points more likely to report excellent health conditions relative to those who do not. Given that D_{it} is likely to be correlated with ε_{ist} in equation 2.8, I estimate equations 2.8 and 2.13 jointly by the bivariate probit model which allows the unobservable random components of the equations are correlated with each other. Columns 3 and 4 report the estimation results. The standard errors of the estimated β_1^* and the ATE are too large to reject that they are statistically different from zero, suggesting that health conditions and religious attendance levels are not likely to be causally related with each other. In panel B,

[11] The standard errors of the ATE in table 2.17 are computed by bootstrap method. The number of the bootstrap iteration is 200.

I repeat the analysis above using an indicator for a respondent's vital status as of 2008 as a supplementary dependent variable. The results are similar to the case when the indicator for excellent health condition is used as a dependent variable. When the endogeneity of church attendance is ignored, weekly churchgoers are about 2.8 percentage points less likely to die between their survey years and 2008 relative to those who do not (columns 1 and 2). However, when the endogeneity is controlled by the bivariate probit model, the mortality gaps between the two groups are statistically insignificant. (columns 3 and 4).

2.6 Summary and conclusion

It has been widely reported in various fields of social science that higher levels of religious participation are strongly associated with better health outcomes (Ellison and Levin, 1998[39]; Deaton, 2011[33]; Koenig et al., 2012[68]). However, it is still an unanswered question as to whether such correlations reflect a causal relationship. In order to shed some light on this question, I examine whether religious participation is causally related with health conditions using the blue laws as a source of exogenous variation in religious participation. Using the GSS data between 1972 and 2001, I confirm that higher levels of religious participation are strongly correlated with better health outcomes, as having been reported in the previous religion-health literature. I also find evidence that the blue laws induced people to participate in religious services more frequently. In spite of the positive impacts of the blue laws on religious participation, however, I do not find any evidence that the blue laws are associated with health conditions of the GSS respondents, which would be observed if religious partic-

Table 2.17: IV Estimation Results: Weekly Church Attendance and Excellent Health Conditions

	Univariate Probit		Bivariate Probit	
	(1)	(2)	(3)	(4)
A. Dependent variable = excellent health condition (yes=1)				
Weekly churchgoer (yes=1)	.172***	.175***	-.281	-.285
(Standard error)	(.028)	(.028)	(.327)	(.336)
Average treatment effect	.049***	.050***	-.088	-.089
(Standard error)	(.007)	(.007)	(.089)	(.095)
Estimated correlation coefficient			.271	.276
(Standard error)			(.195)	(.201)
Covariates				
Individual characteristics	Yes	Yes	Yes	Yes
Labor market status	No	Yes	No	Yes
State and year dummies	Yes	Yes	Yes	Yes
Log likelihood	-7056	-7036	-13799	-13777
Number of observations	11961	11961	11961	11961
B. Dependent variable = dead as of 2008 (yes=1)				
Weekly churchgoer (yes=1)	-.116***	-.117***	-.444	-.458
(Standard error)	(.031)	(.031)	(.381)	(.379)
Average treatment effect	-.028***	-.028***	-.112	-.115
(Standard error)	(.007)	(.007)	(.083)	(.082)
Estimated correlation coefficient			.199	.207
(Standard error)			(.234)	(.234)
Covariates				
Individual characteristics	Yes	Yes	Yes	Yes
Labor market status	No	Yes	No	Yes
State and year dummies	Yes	Yes	Yes	Yes
Log likelihood	-5654	-5638	-12519	-12496
Number of observations	12072	12072	12072	12072

Note. In panel A, the dependent variable is a dummy for having an excellent health condition. In panel B, the dependent variable is a dummy for vital status as of 2008 that takes 1 if dead and 0 otherwise. In both panels, the covariates include (1) individual characteristics: five-year age dummies, years of education, real household income (USD 1,000; base year: 1986), and dummies for nonwhite, female, Catholic, and fundamentalist; (2) labor market status dummies: dummies for employed and unemployed where the reference group is those who are out of the labor force; (3) survey year dummies and state of residence dummies. In panel B, observations with external causes of death such as accidents, homicides, legal intervention, or injury from war operations are excluded from the sample. The standard errors for the ATT are computed by bootstrapping with 200 iterations. p<.01 (***), p<.05 (**), p<.10 (*)

ipation were indeed causally related with health. Taking this as indirect evidence against the existence of the causal link between religious participation and health, I also estimate the causal effect of weekly church attendance on health outcomes using the blue law status of states of the GSS respondents as an IV for their probability of weekly church attendance. When the potential endogeneity of the church attendance behaviors is corrected by the IV strategy, the health benefits of weekly churchgoers are no longer observed. Based on these results, I conclude that the previous studies reporting strong relationships between religious participation and health outcomes are likely to have been driven by selectivity bias rather than a causal relationship.

I acknowledge that my findings should be taken as suggestive and only as tentative evidence against a causal relationship between religious participation and health for the following reasons. First, there may be a reasonable doubt about the validity of the exclusion restriction which is crucial for the IV strategy in section 2.5. Although I have checked for the two arguably most important channels, labor supply and consumption behavior, that might invalidate my analysis, it is still possible that there may be other channels through which the blue laws may affect health outcomes directly. Second, as Hummer et. al. (1999)[62] pointed out, there could be many other, perhaps very important, dimensions of religious participation that cannot be measured by the frequency of religious service attendance. For example, it is entirely possible that the blue laws did not in fact change levels of religious participation because people reacted to the repeals of the blue laws by substituting service attendance with prayers and meditations. I leave these issues for future research. Presumably, these problems might be why Deaton (2011)[33] mentioned that he had not found any credible

way of distinguishing causality between religion and health.

Chapter 3

The Impact of Private Tutoring on Academic Performance of Students

3.1 Introduction

Education is often considered to be one of the most important policy tools to enhance economic growth and social mobility. For this reason, economists have long been interested in measuring the effectiveness of monetary educational investments on student outcomes. Previous studies in this literature has mostly focused on investments made in formal educational sectors such as public schools (Card and Krueger, 1996[27]; Krueger, 2003[71]; Hanushek, 1986[55], 1997[56], 2003[57]) and private schools (Evans and Schwab, 1995[40]; Neal, 1997[84]; Altonji et al., 2005[4]). Recently, a small but growing number of studies are turning attention to measuring the effectiveness of non-formal educational investments, private tutoring, which are widespread across the world these days (Bray and Kwok, 2003[22];

Dang and Rogers, 2008[32])[1] but has been less explored than the educational investments made in the public and private schooling systems.

Private tutoring, often referred to as "shadow education", can be defined as "a set of educational activities outside formal schooling that are designed to improve a student's chances of successfully moving through educational allocation process" (Stevenson and Baker, 1992[100]). It can be further differentiated from formal schooling in that: (1) tutors are motivated by a financial gain, (2) tutees and their parents set higher expectation for the tutor than for a formal school teacher (otherwise, they would not have spent any time and money on the tutoring service), and (3) the objective of its education is to help students be successful in a curriculum of formal education and hence it does not stand alone as an independent educational sector (Tansel and Bircan, 2006[104]; Dang and Rogers, 2008[32]). These distinctive features of private tutoring highlight why understanding the effect of private tutoring is particularly important compared with that of formal schooling. If private tutoring can improve academic outcomes of students and if those outcomes are tightly linked to educational opportunity, income, and social status in their future, private tutoring may limit educational equity and intergenerational income and social mobility (Gurun and Millimet, 2008[51]). Also, to the degree that tutees and their parents believe that quality of instruction is higher in private tutoring classes than it is in their regular school classes, private tutoring can interfere with

[1] A survey of incidence of private tutoring in selected countries is in table A.1 in Dang and Rogers (2008)[32]. Their list of countries includes Azerbaijan, Bangladesh, Canada, Cambodia, Cyprus, Egypt, Greece, Hong Kong, Japan, Kenya, Korea, Lithuania, Mauritius, Romania, Singapore, Sri Lanka, Tanzania, Taiwan, Turkey, Morocco, Romania, Ukraine, the United Kingdom, the United States, Vietnam, and Zimbabwe. Another survey of scale of private tutoring industry in selected countries is in section 2.2 in Bray and Kwok (2003). Their list of countries includes Egypt, India, Japan, Kenya, Malta, Romania, Korea, and Taiwan. Both studies commonly point out that private tutoring is particularly prevalent in East Asia, while it is also growing in other parts of the world.

educational processes in formal schooling (Bray and Kwok, 2003[22]).

In spite of its importance, however, little has been known about whether and to what extent private tutoring can affect academic performance of students. Lack of official data on private tutoring, partly because of its shadowy nature[2], may be one reason why its effect has been less examined relative to that of formal schooling (Tansel and Bircan, 2006[104], Bray and Kwok, 2003[22]). Nonetheless, there are a few studies trying to measure the effectiveness of private tutoring for student outcomes. The results are fairly mixed. Some studies (Stevenson and Baker, 1992[100]; Tansel and Bircan, 2005[103]; Ha and Harphan, 2005[52]; Dang, 2007[31]; Ono, 2007[88]) reported strong positive impacts, while others (Briggs, 2001[23]; Kang, 2007[66]; Gurun and Millimet, 2008[51]; Ryu and Kang, 2013[94]) found that the effects are close to zero or even negative (Lee et al., 2004[74]; Cheo and Quah, 2005[28]).

Given the lack of previous research and the conflicting views among the existing studies, we try to contribute to the literature by estimating the causal effect of private tutoring on achievement test scores of students in three academic subjects – Korean, English, and math – using longitudinal survey data on nationally representative middle schoolers in Korea during 2006-2007. In order to control for potential differences in observable (e.g. student, family, and school backgrounds) and time-invariant unobservable (e.g. cognitive ability) educational inputs of students that may affect both test score outcomes and private tutoring decisions of the students, we apply an empirical model developed by Bonhomme and Sauder (2011)[20]

[2]Bray and Kwok (2003)[22] pointed out that tutors may be unwilling to expose their tutoring activities for tax avoidance reasons. They also argued that pupils and their parents may be unwilling to expose details on their private tutoring expenditures because large amounts of expenditures might seem to confer an unfair advantage in competition with the students' peers and the parents' lack of confidence in school teachers.

to our study. A unique advantage of using this model is that it allows us to estimate not only the average effect of private tutoring but also its distributional effect at each percentile of a test score distribution. While estimating the mean effect is crucial for understanding how private tutoring affects test scores, it may miss some important heterogeneity of the effect of private tutoring across students located at a different point of a test score distribution. For example, students at the bottom of a test score distribution may benefit more from tutors than those who are already in good standing by themselves. Simultaneously, it is also possible that students at the top of the distribution can learn more from private tutoring classes than their peers as they do in their regular school classes. Estimating the distributional effects can shed light on these potential heterogeneity of the effectiveness of private tutoring across students with different levels of pre-determined academic quality, which has not been discussed in the previous studies.

After controlling for observable and time-invariant unobservable characteristics of students, we find that private tutoring has no effect on Korean test scores, while it has modest effects on English and math test scores. For the Korean test scores, both mean and distributional effects of private tutoring are statistically insignificant from zero, which implies the effect of private tutoring in Korean is homogenously zero across students at different points of the test score distribution. On the other hand, we find that the average effects of private tutoring are about 10 and 16 percent of a standard deviation of test scores in English and math, respectively. Looking at the distributional effects, we find that the effects are positive in the upper half of the distributions, while they are statistically insignificant from zero in the lower half of the distributions. At their peaks around the 70th to the 80th percentile,

the effects of private tutoring amount to roughly 30 and 35 percent of a standard deviation of English and math test scores, respectively. Collectively, these results suggest that the effectiveness of receiving private tutoring, if any, is at most modest but the extent of its effectiveness is likely to vary substantially across students with different levels of academic quality. Specifically, it seems that students in good standing tend to benefit more from private tutoring than those at the bottom of the test score distributions. This suggests that private tutoring is not an effective remedial educational measure for students left behind, although it seems to have a modest effect of facilitating learning processes of students in good standing.

3.2 Previous literature

There are a few studies on the effects of private tutoring on students' academic performance. Stevenson and Baker (1992)[100], Tansel and Bircan (2005)[103], and Ha and Harphan (2005)[52], and Ono (2007)[88] examined data from Japan, Turkey, Vietnam, Japan and reported strong positive effects of private tutoring, respectively. On the other hand, Briggs (2001)[23] found a negligible effect in the U.S., while Lee et al (2004)[74] and Choe and Quah (2005)[28] even reported negative effects of private tutoring in Korea and Singapore, respectively. However, these studies did not explicitly deal with potential endogeneity of private tutoring and hence their results should be interpreted with cautions.

More recent studies tried to address the potential endogeneity of private tutoring in various ways. Dang (2007)[31] examined the effect of private tutoring expenditures on self-

reported academic performance of students using nationally representative household survey data in Vietnam during 1997-1998. In order to address the potential endogeneity of private tutoring expenditures, Dang developed and estimated a simultaneous equation system consisting of a Tobit model for private tutoring expenditures and an ordered probit model for the self-reported academic performance. He found that private tutoring has a significant impact on students' academic performance. However, the measure for academic performance used by Dang (2007)[31] is a self-reported variable, which has four ordered responses of "excellent", "good", "average", and "poor". This raises a concern about whether the self-reported measure can reflect a student's true academic performance. Also, it should be noted that the joint Tobit-ordered probit model used by Dang (2007)[31] relies on strong parametric assumptions.

While Dang (2007)[31] dealt with the potential endogeneity of private tutoring by explicitly modeling the process generating private tutoring expenditures and academic performance of students using a simultaneous framework, Kang (2007)[66] tried to find an exogenous variation that affects private tutoring decisions of parents but does not affect academic outcomes of students. In particular, Kang used a student's birth order as an instrumental variable (IV) for private tutoring expenditures. The logic is that parents tend to have more concerns about a first-born child's education and invest more for the student's private tutoring than for the later-borns, while a student's birth order is determined exogenously by nature and is unlikely to directly affect academic performance of the student. Indeed, in the data used by Kang (2007)[66], Korean parents were found to spend on average about 30 percent more tutoring expenditures for their first-borns relative to the later-borns. Using the first-born indicator as

an IV for tutoring expenditures, Kang found evidence of a modest effect of private tutoring on the national college-entrance exam scores for high school graduates in Korea: a 10 percent increase in tutoring expenditure was found to improve only about .56 percentile point in the test scores.

Although the IV used by Kang (2007)[66] is fairly strong and novel in the literature, reasonable doubts arise about the validity of the exclusion restriction. If parents indeed are more concerned about their first-born children's education and invest more for the first-borns' private tutoring, other parental inputs, for example helping children with their homework, would be also greater for the first-borns than the later-borns. To the extent that these parental inputs are related with academic outcomes of students, the validity of the IV strategy would be questionable. Gurun and Millimet (2008)[51] pointed out this issue and took a different approach. Given that a valid exclusion restriction is unavailable in their data, Gurun and Millimet tried to assess the importance of the potential endogeneity of private tutoring by using the bivariate probit framework suggested by Altonji et al. (2005[4], 2008[5]). In particular, they consider

$$Y_i = 1\left[\beta_1 D_i + X_i\beta_2 + \varepsilon_i > 0\right], \tag{3.1}$$

$$D_i = 1\left[X_i\beta_3 + \lambda_i > 0\right], \tag{3.2}$$

where Y_i denotes a dummy variable that takes 1 if student i enters a university and 0 otherwise; D_i is a dummy variable that takes 1 if student i receives private tutoring and 0 otherwise; X_i is a vector of observable characteristics of student i; and ε_i and λ_i are

error terms which follow a bivariate normal distribution of zero means, unit variances, and correlation coefficient ρ. The source of the potential endogeneity of D_i in equation 3.1 is potential correlation between ε_i and λ_i, which is represented by $\rho \neq 0$. Following Altonji et al. (2005[4], 2008[5]), Gurun and Millimet (2008)[51] performed the following sensitivity analysis. First, they constrained ρ to be 0 and found a large positive estimate for ${\beta_1}^3$. Then, they constrained ρ to take different values and looked at how the corresponding estimates for β_1 change over different values of ρ. They found that the estimated effect of private tutoring (represented by estimated β_1) becomes statistically insignificant and even falls below zero when only a moderate level of endogeneity (represented by ρ) is allowed. Given these results, they concluded that the strong positive effects of private tutoring often reported in the previous studies may have been driven by the potential endogeneity problem.

Recently, Ryu and Kang (2013)[94] extended Kang (2007)[66]'s study by employing an alternative empirical strategy. Specifically, they employed the monotone instrumental variable (MIV) strategy by Manski and Pepper (2000)[77] and tried to estimate the sharp *bounds* of the causal effect of private tutoring on test scores of students. In the standard IV strategy, an IV is not allowed to affect an outcome variable directly (i.e. exclusion restriction). By contrast, in the MIV strategy by Manski and Pepper (2000)[77], a MIV may affect an outcome variable directly but only *monotonically* (i.e. either positively or negatively). A child's first-born status, which was used as an IV in Kang (2007)[66], may not be a valid IV for private tutoring expenditures because the first-borns may receive a larger parental support than the later-borns in many other forms than private tutoring. However, the first-born indicator

[3]This is basically what Tansel and Bircan (2005)[103] found. Tansel and Bircan used the same data as Gurun and Millimet (2008)[51] did but they assumed that private tutoring is exogenously determined.

can be still used as a MIV of Manski and Pepper (2000)[77] as long as a child's birth order is *monotonically* related with the student's academic performance. Using a child's first-born status as a MIV for private tutoring expenditures, Ryu and Kang (2013)[94] found that the estimated upper bound for the causal effect of private tutoring expenditures is small, which implies the causal effect is likely to be close to zero.

The empirical strategy of Ryu and Kang (2013)[94] is more advanced than that of Kang (2007)[66] in that it draws the same conclusion while relaxing the restrictive IV assumption to the less restrictive MIV assumption. However, in order to derive the sharp bounds, Ryu and Kang (2013)[94] had to employ two additional monotonicity assumptions: (1) the effect of private tutoring expenditures on test scores is non-negative for all students (monotone treatment response: MTR) and (2) the non-random selection into a larger amount of private tutoring expenditures is positive on average (monotone treatment selection: MTS). Given that there are studies reporting evidence of negative effects of private tutoring (Lee et al., 2004[74]; Choe and Quah, 2005[28]) and negative selection into a larger amount of private tutoring expenditures (Gurun and Millimet, 2008[51]), the MTS and MTR assumptions seem to be restrictive and the validity of the causal inference based on those assumptions will be limited.

The conflicting results from the previous studies may be partly because they used data from different periods and countries. However, a perhaps more important reason may be because it is very hard to deal with the potential endogeneity of private tutoring. Unlike the studies on the effectiveness of educational investments in formal schooling sectors where experimental (e.g. the STAR experiment in Krueger and Whitmore, 2001[72]) and/or quasi-

experimental (e.g. private school voucher programs in McEwan, 2004[78]) data are often available, the previous studies on private tutoring had to rely on observational data for a causal inference. To overcome the challenge, they tried to either model the data generating process econometrically (Dang, 2007[31]; Gurun and Millimet, 2008[51]) or try to find an exclusion restriction that affects private tutoring behaviors but not (or in a limited way) students' outcomes (Kang, 2007[66]; Ryu and Kang, 2013[94]), while most of the literature simply proceed with an unrealistic assumption that private tutoring is exogenously determined.

As in the previous studies, we do not find any plausible source of exogenous variation in students' (or their parents') private tutoring decisions in our data. Given this limitation, we attempt to estimate the mean and distributional effects of private tutoring controlling for students' observable and time-invariant unobservable educational inputs by using the empirical model developed by Bonhomme and Sauder (2011)[20], which will be described in section 3.4.

3.3 Data

The data for this study are from the Korea Education Longitudinal Study (KELS). The KELS is an annual longitudinal survey whose basic structure is similar to that of the National Educational Longitudinal Studies of the U.S. (Ryu and Kang, 2013[94]) A nationally representative sample of 6,908 seventh graders[4] (age 13) was first surveyed in 2005 and fol-

[4]In terms of the Korean educational system, these students are in the first year of their middle schools.

lowed every year since then. The KELS also surveyed parents, teachers, and school principals of the 6,908 students in order to collect information on family and school characteristics of the students.

We measure academic performance of the students by their achievement test scores in three academic subjects – Korean, English, and math – which are available in the first three waves (years 2005, 2006, and 2007) of the KELS data. Among the three waves, information on private tutoring experience of students is available in the second and the third waves (years 2006 and 2007)[5], which we mainly use for this paper. In each year of 2006 and 2007, the KELS asked parents of the 6,908 students whether or not the students had received private tutoring during the survey year for each subject of Korean, English, and math, separately. Based on the parents' responses to this question, we divide students into two groups, treatment and control groups, based on whether they have received private tutoring in 2007 (treatment group) or not (control group). In order to compare test scores of students in the treatment and control groups at a common baseline, we need to control for students' pre-determined academic quality as well as their other characteristics. For this reason, we control for test scores and private tutoring status of students as well as their individual, parental, and school characteristics observed in 2006 throughout this study.

Tables 3.1, 3.2, and 3.3 show summary statistics of our Korean, English, and math sam-

[5]In fact, the first wave (year 2005) of the KELS data also includes information on private tutoring. However, a careful reading of the parental questionnaires of the KELS shows that the reference period of questions on private tutoring changed between the first wave (year 2005) and the other two waves (years 2006 and 2007). In 2005, the KELS asked a student's private tutoring experience during the survey month (Oct. 2005). On the other hand, in 2006 and 2007, it asked a student's private tutoring experience during the entire survey year (2006 or 2007). Since pupils may receive different amounts of tutoring in different seasons (Bray and Kwok 2003[22]), we choose to focus on the second and third waves of the KELS during which the survey collected the information on private tutoring in a consistent way.

Table 3.1: Summary Statistics: Korean

	(1) Total		(2) Treatment		(3) Control	
Variables:	Mean	S.D.	Mean	S.D.	Mean	S.D.
Treatment (in 2007):						
Private tutoring (yes=1)	.569	.495	1.00	0.00	0.00	0.00
Outcome (in 2007):						
Test scores	57.7	20.8	58.4	20.5	56.7	21.2
Baseline characteristics (in 2006):						
Student characteristics:						
Test scores	60.2	18.8	60.9	18.6	59.4	19.0
Private tutoring (yes=1)	.542	.498	.728	.445	.297	.457
Female (yes=1)	.481	.500	.431	.495	.548	.498
First born (yes=1)	.500	.500	.529	.499	.461	.499
Number of siblings	1.21	.723	1.17	.677	1.26	.777
Handicapped (yes=1)	.019	.137	.019	.136	.020	.139
Parental characteristics:						
Average age	42.3	4.08	42.1	3.69	42.6	4.08
Average years of education	12.8	2.23	13.0	2.10	12.6	2.36
Married (yes=1)	.896	.305	.931	.254	.851	.357
Monthly income (1000 KRW)	3379	2393	3642	2459	3032	2257
Having a religion (yes=1)	.686	.464	.698	.459	.670	.470
School characteristics:						
Large city (yes=1)	.467	.499	.477	.500	.454	.498
Medium city (yes=1)	.445	.497	.456	.498	.431	.495
Rural area (yes=1)	.088	.283	.067	.251	.115	.319
Private school (yes=1)	.205	.404	.199	.399	.212	.409
Coed school (yes=1)	.640	.480	.642	.479	.636	.481
Boys-only school (yes=1)	.188	.391	.204	.403	.167	.373
Girls-only school (yes=1)	.172	.378	.154	.361	.197	.398
Class size (# of students)	35.4	5.46	35.8	5.00	35.0	5.99
Number of observations	4126		2346		1780	

Table 3.2: Summary Statistics: English

	(1) Total		(2) Treatment		(3) Control	
Variables:	Mean	S.D.	Mean	S.D.	Mean	S.D.
Treatment (in 2007):						
Private tutoring (yes=1)	.758	.429	1.00	0.00	0.00	0.00
Outcome (in 2007):						
Test scores	55.6	26.7	59.9	26.4	42.2	23.2
Baseline characteristics (in 2006):						
Student characteristics:						
Test scores	57.3	24.4	61.0	24.1	45.8	21.4
Private tutoring (yes=1)	.749	.434	.868	.338	.374	.484
Female (yes=1)	.485	.500	.472	.499	.525	.500
First born (yes=1)	.502	.500	.530	.499	.416	.493
Number of siblings	1.20	.706	1.17	.653	1.31	.842
Handicapped (yes=1)	.018	.134	.017	.131	.021	.142
Parental characteristics:						
Average age	42.3	3.98	42.2	3.65	42.7	4.85
Average years of education	12.9	2.21	13.2	2.13	12.0	2.19
Married (yes=1)	.903	.296	.937	.243	.795	.404
Monthly income (1000 KRW)	3440	2331	3740	2376	2501	1899
Having a religion (yes=1)	.690	.463	.700	.459	.658	.475
School characteristics:						
Large city (yes=1)	.471	.499	.491	.500	.407	.491
Medium city (yes=1)	.445	.497	.440	.496	.463	.499
Rural area (yes=1)	.084	.277	.069	.253	.130	.337
Private school (yes=1)	.200	.400	.199	.399	.204	.403
Coed school (yes=1)	.640	.480	.646	.478	.618	.486
Boys-only school (yes=1)	.188	.391	.188	.390	.189	.392
Girls-only school (yes=1)	.172	.378	.166	.372	.192	.394
Class size (# of students)	35.6	5.39	36.0	4.96	34.2	6.37
Number of observations	4617		3498		1119	

Table 3.3: Summary Statistics: Math

	(1) Total		(2) Treatment		(3) Control	
Variables:	Mean	S.D.	Mean	S.D.	Mean	S.D.
Treatment (in 2007):						
Private tutoring (yes=1)	.764	.425	1.00	0.00	0.00	0.00
Outcome (in 2007):						
Test scores	52.9	26.3	57.2	26.0	39.1	22.3
Baseline characteristics (in 2006):						
Student characteristics:						
Test scores	52.7	24.6	56.3	24.5	41.1	21.0
Private tutoring (yes=1)	.764	.425	.862	.345	.372	.484
Female (yes=1)	.487	.500	.478	.500	.517	.500
First born (yes=1)	.505	.500	.530	.499	.426	.495
Number of siblings	1.20	.698	1.16	.643	1.31	.842
Handicapped (yes=1)	.019	.138	.019	.135	.022	.146
Parental characteristics:						
Average age	42.3	3.96	42.1	3.59	42.7	4.96
Average years of education	12.9	2.23	13.2	2.13	12.0	2.25
Married (yes=1)	.904	.294	.936	.245	.801	.399
Monthly income (1000 KRW)	3479	2392	3785	2454	2491	1861
Having a religion (yes=1)	.693	.461	.705	.456	.655	.476
School characteristics:						
Large city (yes=1)	.471	.499	.495	.500	.394	.489
Medium city (yes=1)	.445	.497	.441	.497	.458	.498
Rural area (yes=1)	.084	.278	.064	.245	.148	.355
Private school (yes=1)	.201	.401	.201	.401	.201	.401
Coed school (yes=1)	.641	.480	.649	.477	.615	.487
Boys-only school (yes=1)	.186	.389	.186	.389	.188	.391
Girls-only school (yes=1)	.173	.378	.165	.371	.198	.398
Class size (# of students)	35.6	5.39	36.1	4.93	34.1	6.44
Number of observations	4697		3589		1108	

ples, respectively. After removing observations with missing information on the variables we use in this study, we have 4,126, 4,617, and 4,697 valid observations for the Korean, English, and math samples, respectively. In each of the three tables, columns 1, 2, and 3 show summary statistics of the entire sample, the treatment group, and the control group, respectively. In all of the three subjects, students in the treatment groups tend to report better academic performance than their peers in the control groups. In the Korean achievement test, students who receive private tutoring score slightly higher on average than their peers who do not receive private tutoring by about 1.62 points. This is about 8 percent of a standard deviation of Korean test scores of the treatment group. By contrast, students in the treatment groups achieve on average 17.7 and 18.0 points higher in the English and math tests than those not receiving private tutoring, which amount to about 67 and 69 percent of a standard deviation of the treatment groups, respectively. The tables also show that the treatment and control groups are also very different from each other in terms of other dimensions as well. For example, students in the treatment group tend to score higher in the achievement tests and be more likely to receive private tutoring than their peers in the control group even before the treatment is realized. The two groups also show a large difference in their student, parental, and school characteristics. In the following section, we discuss how to control for these differences in observable characteristics and potential differences in *time-invariant* unobservable characteristics between the two groups.

3.4 Empirical analysis

3.4.1 Mean effects of private tutoring

For each of the Korean, English, and math samples, we have data on

$$(Y_{i1}, Y_{i2}, D_i, X_i), \tag{3.3}$$

where Y_{i1} and Y_{i2} denote test scores of student i in 2006 (denoted by period 1 hereafter) and 2007 (denoted by period 2 hereafter) periods, respectively; D_i is our treatment indicator that takes 1 if student i receives private tutoring in period 2 and 0 otherwise; and X_i is a vector of individual, parental, and school characteristics of student i.

In order to examine whether the improvement in test scores is greater for students in the treatment group than those in the control group, we begin by estimating the following equation by the ordinary least square (OLS) method:

$$Y_{i2} - Y_{i1} = \beta_0 + \beta_1 D_i + X_i \beta_2 + \varepsilon_i, \tag{3.4}$$

where the dependent variable is a change in test scores of student i which is calculated by subtracting a pre-treatment test score of student i from a post-treatment test score of the student ($Y_{i2} - Y_{i1}$). The error term, ε_i, denotes unobserved and unmeasured characteristics of student i that might affect academic achievement of the student. If private tutoring (D_i) is as good as randomly assigned among students who share the same observable baseline characteristics (X_i) and hence D_i would be uncorrelated with ε_i conditional on X_i, the OLS

estimation of equation 3.4 would yield a consistent estimate for β_1.

Table 3.4 shows the OLS estimation results of equation 3.4 for each of the Korean, English, and math samples. In column 1, as a benchmark, we do not control for any observable characteristics and report the simple mean differences in the changes in test scores between the treatment and control groups. In columns 2, 3 and 4, we add student, parental and school characteristics as additional covariates to the list of controls in columns 1, 2 and 3, respectively. Regardless of the choice of the covariate specifications, we find evidence that students benefit from receiving private tutoring in English and math but not in Korean. When the individual, parental, and school characteristics of students are controlled, test scores of students who receive private tutoring tend to improve more than those of students who do not by about 3.3 and 3.8 points in English and math, respectively. However, we do not find any statistically significant difference in the changes in test scores between the two groups in Korean.

The OLS estimator in table 3.4 has some weak points that can be improved on. First, the estimator requires a parametric functional form of equation 3.4. Second, it is not clear what kind of average causal effect the OLS estimate for β_1 would identify. Third, perhaps most importantly, in order for the OLS estimate for β_1 to have a causal interpretation, we need the so-called selection-on-observable assumption which presumes students who receive private tutoring are comparable to those who do not as long as differences in observable characteristics between the two groups are controlled for. Although this assumption has been used extensively in the previous studies (Stevenson and Baker, 1992[100]; Briggs, 2001[23]; Lee et al., 2004[74]; Tansel and Bircan, 2005[103]; Ha and Harphan, 2005[52]; Choe and

Table 3.4: OLS Estimation Results: Private Tutoring and Change in Test Scores

Dependent variable:	Specifications			
Change in test scores	(1)	(2)	(3)	(4)
A. Subject: Korean				
Private tutoring	.163	.610	.530	.574
(Standard error)	(.546)	(.610)	(.616)	(.612)
Covariates:				
Student characteristics	No	Yes	Yes	Yes
Parental characteristics	No	No	Yes	Yes
School characteristics	No	No	No	Yes
R-squared	.000	.002	.004	.015
Number of observations	4126	4126	4126	4126
B. Subject: English				
Private tutoring	2.52	3.15	3.20	3.28
(Standard error)	(.650)	(.803)	(.820)	(.823)
Covariates:				
Student characteristics	No	Yes	Yes	Yes
Parental characteristics	No	No	Yes	Yes
School characteristics	No	No	No	Yes
R-squared	.003	.006	.007	.011
Number of observations	4617	4575	4272	4262
C. Subject: Math				
Private tutoring	2.90	3.64	3.60	3.80
(Standard error)	(.702)	(.832)	(.857)	(.853)
Covariates:				
Student characteristics	No	Yes	Yes	Yes
Parental characteristics	No	No	Yes	Yes
School characteristics	No	No	No	Yes
R-squared	.003	.006	.007	.020
Number of observations	4697	4612	4236	4203

Note. Outcome variable is the change in achievement test scores of students between 2006 and 2007. Treatment variable (private tutoring) is an indicator that takes 1 if a student receives private tutoring in 2007 and 0 otherwise. Covariates include (1) student characteristics: a dummy for receiving private tutoring in 2006, a dummy for female, a dummy for being handicapped, number of siblings; (2) parental characteristics: parents' average age, parents' average years of education, a dummy for being married, parents' average monthly income, and a dummy for having a religion; and (3) school characteristics: a dummy for being located in metropolitan area, a dummy for being located in suburban area, a dummy for private school, a dummy for boy-only school, a dummy for girl-only school, and class size.

Quah, 2005[28]; Ono, 2007[88]), it is easy to think of why this assumption may not hold. Students receiving private tutoring are likely to be different from their peers in many unobservable, but perhaps very important, ways. For example, it is possible that students with higher motivation levels are also more likely to use private tutoring to enhance their academic achievement further. On the other hand, it is also possible that students with lower levels of cognitive ability are more likely to seek private tutoring lest they should be left behind in their school classes. To the extent that there are unobservable characteristics that are correlated with academic performance of students but unbalanced between the treatment and control groups, the OLS estimates in table 3.4 would not reflect a causal relationship.

To address, at least partly, these problems, we apply a semiparametric estimation technique developed by Bonhomme and Sauder (2011)[20] to our study[6]. We begin by explicitly specifying the average causal effect of interest using the potential outcome framework by Rubin (1974)[93]. Note that our treatment (D_i) is realized in period 2. Hence, the observed test score of student i in period 2 (Y_{i2}) would be either of two potential test scores depending on whether the student receives the treatment or not. Specifically, Y_{i2} can be written as

$$Y_{i2} = D_i Y_{i2}^1 + (1 - D_i) Y_{i2}^0 \tag{3.5}$$

where Y_{i2}^1 (or Y_{i2}^0) denote potential test scores of student i in period 2 had they received (or not received) private tutoring during the period.

The causal effect of receiving private tutoring on achievement test scores of student i can

[6] Our illustration of the method is heavily drawn from sections II and III of Bonhomme and Sauder (2011)[20].

be measured by $Y_{i2}^1 - Y_{i2}^0$, the difference between the two potential test score outcomes of the student. Since we can only observe either Y_{i2}^1 or Y_{i2}^0 for a given student, however, the causal effect is fundamentally unidentifiable at individual level. Instead, we try to identify the average values of the causal effects for students who receive private tutoring in 2007, which is often referred to as the average treatment effect on the treated or ATT:

$$\begin{aligned} ATT &= E\left[Y_{i2}^1 - Y_{i2}^0 \,|\, D_i = 1\right] && (3.6) \\ &= E\left[Y_{i2} \,|\, D_i = 1\right] - E\left[Y_{i2}^0 \,|\, D_i = 1\right] && (3.7) \end{aligned}$$

In equation 3.7, $E\left[Y_{i2} \,|\, D_i = 1\right]$ is empirically observable, while $E\left[Y_{i2}^0 \,|\, D_i = 1\right]$ is counterfactual and hence unobservable. To identify the counterfactual $E\left[Y_{i2}^0 \,|\, D_i = 1\right]$, we model potential post-treatment test scores of student i in the state of world where they do not receive private tutoring (Y_{i2}^0) and their baseline test scores of student i (Y_{i1}) as the sum of three components (Bonhomme and Sauder, 2011[20]):

$$\begin{aligned} Y_{i1} &= f_1\left(X_i\right) + \eta_i + v_{i1} && (3.8) \\ Y_{i2}^0 &= f_2^0\left(X_i\right) + \eta_i + v_{i2}^0, && (3.9) \end{aligned}$$

where X_i denotes student, parental, and school characteristics of student i whose effects on test scores are flexibly modeled as arbitrary functions of $f_1\left(\cdot\right)$ and $f_2^0\left(\cdot\right)$; η_i represents unobservable characteristics of student i which are fixed between the two periods (e.g. cognitive ability); v_{i1} and v_{i2}^0 represent time-varying unobservable shocks to test scores (e.g.

physical conditions on the exam day) that are allowed to be correlated with each other in an arbitrary way. These equations may be viewed as an educational production function (Hanushek, 1986[55]) that relates observable (X_i) and unobservable (η_i; v_{i1}, v_{i2}^0) educational inputs to test score outputs (Y_{i1}, Y_{i2}^0) in each period. Except for its additive structure, the educational production function is fairly flexible in that it does not impose any distributional or functional-form restrictions on its three components.

Given that the educational production function takes the forms of equations 3.8 and 3.9, we impose the following assumption in order to identify the ATT:

Assumption 3.1: (v_{i1}, v_{i2}^0) are independent of D_i conditional on X_i

Assumption 3.1 requires that the treatment and control groups share the common time-varying unobservables (v_{i1}, v_{i2}^0) in their test score outcomes among those who have the same observable characteristics (X_i). Note that the educational production functions in equations 3.8 and 3.9 divide unobservable educational inputs into two parts: time-invariant unobservables (η_i) and time-varying ones (v_{i1}, v_{i2}^0). In this respect, assumption 3.1 may be referred to as *selection-on-observable-and-time-invariant-unobservable* assumption in that it restricts the determinants of potential non-random selection into private tutoring to be observables or time-invariant unobservables or both.

Under assumption 3.1, the ATT in equation 3.6 can be identified as (Abadie, 2005[1];

Bonhomme and Sauder, 2011[20])

$$ATT = \frac{1}{\Pr[D_i = 1]} E\left[\left\{\frac{D_i - \Pr[D_i = 1 \mid X_i]}{1 - \Pr[D_i = 1 \mid X_i]}\right\}(Y_{i2} - Y_{i1})\right], \tag{3.10}$$

for which we need a usual common support assumption:

Assumption 3.2: $\Pr[D_i = 1] > 0$ and $\Pr[D_i = 1 \mid X_i] < 1$ with probability 1

Details on how to derive equation 3.10 is provided in the appendix. Following Bonhomme and Sauder (2011)[20][7], we estimate the ATT in equation 3.10 by

$$\widehat{ATT} = \frac{1}{\frac{1}{N}\sum_{i=1}^{N} D_i} \frac{1}{N} \sum_{i=1}^{N} \left[\frac{D_i - \widehat{\Pr}[D_i = 1 \mid X_i]}{1 - \widehat{\Pr}[D_i = 1 \mid X_i]}(Y_{i2} - Y_{i1})\right] \tag{3.11}$$

Since X_i consists of many covariates including continuous variables, we estimate $\widehat{\Pr}[D_i = 1 \mid X_i]$ in equation 3.11 by a logit regression of D_i on X_i in order to avoid the curse of dimensionality problem. When computing the $\widehat{ATT}$ in equation 3.11, we restrict our estimation sample to observations with $.05 < \widehat{\Pr}[D_i = 1 \mid X_i] < .95$ in order to make sure that the common support assumption (assumption 3.2) would hold. We compute standard errors of the $\widehat{ATT}$ by bootstrapping with 500 iterations.

Table 3.5 reports estimation results for $\widehat{ATT}$ in equation 3.11. In Korean, the estimated mean effects are statistically insignificant from zero regardless of the choice of covariate specifications, suggesting that students on average do not benefit from receiving private

[7] When estimating equation 3.11, we follow Bonhomme and Sauder (2011)[20] for the choices described in the next paragraph.

Table 3.5: Estimated Mean Effects of Private Tutoring under Selection on Observables and Time-Invariant Unobservables

Dependent variable:	Specifications		
Test scores in 2007	(1)	(2)	(3)
A. Subject: Korean			
Estimated ATT	.550	.810	.839
(Standard error)	(.724)	(.733)	(.739)
Covariates:			
Student characteristics	Yes	Yes	Yes
Parental characteristics	No	Yes	Yes
School characteristics	No	No	Yes
Number of observations	4126	4126	4126
B. Subject: English			
Estimated ATT	3.15	2.99	2.50
(Standard error)	(.943)	(1.17)	(1.10)
Covariates:			
Student characteristics	Yes	Yes	Yes
Parental characteristics	No	Yes	Yes
School characteristics	No	No	Yes
Number of observations	4617	4458	4443
C. Subject: Math			
Estimated ATT	3.65	3.99	4.21
(Standard error)	(1.07)	(1.14)	(1.21)
Covariates:			
Student characteristics	Yes	Yes	Yes
Parental characteristics	No	Yes	Yes
School characteristics	No	No	Yes
Number of observations	4697	4457	4442

Note. Outcome variable is achievement test scores of students in 2007. Treatment variable is an indicator that takes 1 if a student receives private tutoring in 2007 and 0 otherwise. Covariates include (1) student characteristics: a dummy for receiving private tutoring in 2006, a dummy for female, a dummy for being handicapped, number of siblings; (2) parental characteristics: parents' average age, parents' average years of education, a dummy for being married, parents' average monthly income, and a dummy for having a religion; and (3) school characteristics: a dummy for being located in metropolitan area, a dummy for being located in suburban area, a dummy for private school, a dummy for boy-only school, a dummy for girl-only school, and class size. The size of estimation sample gets smaller as more variables are added to the list of controls because we restrict the estimation sample to observations with $.05<\Pr[D=1|X]<.95$ in order to make sure that assumption 3.2 holds. We compute standard errors by bootstrap and report them in parenthesis. The number of the bootstrap iterations is 500.

tutoring in the subject. In English and math, when all of the student, parental, and school characteristics are controlled, we find mean effects of 2.50 and 4.21 points, which are roughly 10 and 16 percent of a standard deviation of test scores, respectively. This suggests that private tutoring has a small positive effect on academic performance of students in those subjects.

3.4.2 Distributional effects of private tutoring

The empirical model in section 3.4.1 evaluates the *average* effects of receiving private tutoring on test scores. We find that private tutoring improves test scores of students on average by about 2.50 and 4.21 points in English and math, respectively, while it exerts no effect in Korean. Although estimating the average effect of private tutoring is crucial for understanding how private tutoring affects academic performance of students, the estimates may miss some important features of the effects of private tutoring. For example, private tutoring may be more helpful to students who are left behind in their school classes than those who are already in good standing by themselves. Simultaneously, it is also possible that more advanced students can make better use of private tutoring to enhance their academic performance further and, hence, private tutoring may be most effective for students at top of a test score distribution. In order to account for this potential heterogeneity of how private tutoring affects academic achievement of students, we employ the model by Bonhomme and Sauder (2011)[20] and estimate the *distributional* effect of receiving private tutoring on test scores at each percentile point of a test score distribution. In particular, our object of interest

in this section is the quantile treatment effect on the treated (QTT) which is defined as:

$$QTT(\tau) = F^{-1}_{Y_{i2}|D_i=1}(\tau) - F^{-1}_{Y^0_{i2}|D_i=1}(\tau), \ \tau \in (0,1), \tag{3.12}$$

where $\tau \in (0,1)$ represents a percentile point of a test score distribution and $F^{-1}_W(\cdot)$ denotes the inverse of the cumulative distribution function (CDF) of a random variable W. Note that the distribution of $Y_{i2}|D_i = 1$ – post-treatment test scores of students who receive private tutoring – is empirically observable and hence it is straightforward to estimate $F^{-1}_{Y_{i2}|D_i=1}(\cdot)$ nonparametrically. The key issue is how to estimate $F^{-1}_{Y^0_{i2}|D_i=1}(\cdot)$ because $Y^0_{i2}|D_i = 1$ – *potential* post-treatment test scores of students who receive private tutoring had they not received private tutoring – is counterfactual and hence unobservable. Bonhomme and Sauder (2011)[20] provided conditions under which the counterfactual distribution of $Y^0_{i2}|D_i = 1$ can be identified and estimated. Following their approach, we maintain all the assumptions that we made in section 3.4.1 (the educational production functions of equations 3.8 and 3.9 and assumptions 3.1 and 3.2). On top of those assumptions, we further assume

Assumption 3.3 (v_{i1}, v^0_{i2}) are independent of η_i conditional on X_i and D_i,

which presumes that the temporal shocks to test scores (v_{i1} and v^0_{i2}) are independent of time-invariant unobservables (η_i) among students who share the same observable characteristics (X_i and D_i). For example, this assumption excludes the possibility that students with different levels of motivation and cognitive ability (represented by η_i) face different patterns of physical conditions on their exam days (represented by v_{i1} and v^0_{i2}) conditional on their

observable characteristics (X_i and D_i). Given that educational production function takes the additive structure of equations 1.7 and 1.8 and that assumptions 3.1, 3.2, and 3.3 hold, the probability density function (PDF) of the counterfactual $Y_{i2}^0|D_i = 1$ is identified as (Bonhomme and Sauder, 2011[20])

$$\begin{aligned} f_{Y_{i2}^0|D_i=1}(y) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\exp(-jty) \\ &\quad \frac{1}{\Pr[D_i=1]}E\left[\omega(t\mid X_i)(1-D_i)\exp(jtY_{i2})\right]dt, \end{aligned} \tag{3.13}$$

where $j=\sqrt{-1}$, $t\in\mathbf{R}$, and

$$\omega(t\mid X_i) \equiv \frac{E[D_i exp(jtY_{i1})\mid X_i]}{E[(1-D_i)exp(jtY_{i1})\mid X_i]} \tag{3.14}$$

The details on the identification procedure is in the appendix. Following Bonhomme and Sauder (2011)[20][8], we estimate the counterfactual density in equation 3.13 with

$$\begin{aligned} \widehat{f}_{Y_{i2}^0|D_i=1}(y) &= \frac{1}{2\pi}\int_{-T_N}^{T_N}\exp(-jty) \\ &\quad \frac{1}{\frac{1}{N}\sum_{i=1}^{N}D_i}\left(\frac{1}{N}\sum_{i}^{N}\widehat{\omega}(t\mid X_i)(1-D_i)\exp(jtY_{i2})\right)dt, \end{aligned} \tag{3.15}$$

where

[8] When estimating equation 3.15, we follow Bonhomme and Sauder (2011)[20] for the choices described in the next paragraph.

$$\widehat{\omega}\left(t \mid X_i\right) = \frac{\widehat{E}\left[D_i exp(jtY_{i1}) \mid X_i\right]}{\widehat{E}[(1-D_i)exp(jtY_{i1}) \mid X_i]} \tag{3.16}$$

Since X_i consists of many covariates including continuous variables, we approximate the conditional expectations in the numerator and denominator in equation 3.16 with linear projections of $D_i exp(jt\rho Y_{i1})$ and $(1-D_i)exp(jt\rho Y_{i1})$ onto X_i, respectively, in order to avoid the curse of dimensionality. We choose the trimming parameter T_N in equation 3.15, which is analogous to choosing a bandwidth in nonparametric density estimation, by the rule of thumb method suggested by Diggle and Hall (1993)[36][9]. We compute the integration using the trapezoid rule with 200 equidistant nods.

The estimation results of equation 3.15 for each of the three subjects are presented in the left columns of figures 3-1, 3-2, and 3-3. In each of the three figures, we report estimation results for three different covariate specifications. In plot A, we include only student characteristics in X_i of equation 3.15, while, in plots B and C, we add parental and school characteristics to the list of covariates in plots A and B, respectively. In all of the plots, solid lines represent the kernel density estimates[10] for the realized test score distribution of students who receive private tutoring $(Y_{i2}|D_i = 1)$, while dashed lines represent the counterfactual test score distribution of students who receive private tutoring had they not received it $(Y^0_{i2}|D_i = 1)$. Since we compare a realized test score distribution of students receiving private tutoring with a counterfactual test score distribution of the same group of students, any difference between the two distributions can be attributable to the causal

[9] Details on how we determine the values of T_N are in the appendix.

[10] When estimating the density of $Y_{i2}|D_i = 1$, we use the Gaussian kernel with the rule of sum bandwidth suggested by Silverman (1986)[98].

Figure 3-1: Estimated Probability Density Functions (Left) and Quantile Treatment Effects on the Treated (Right): Korean

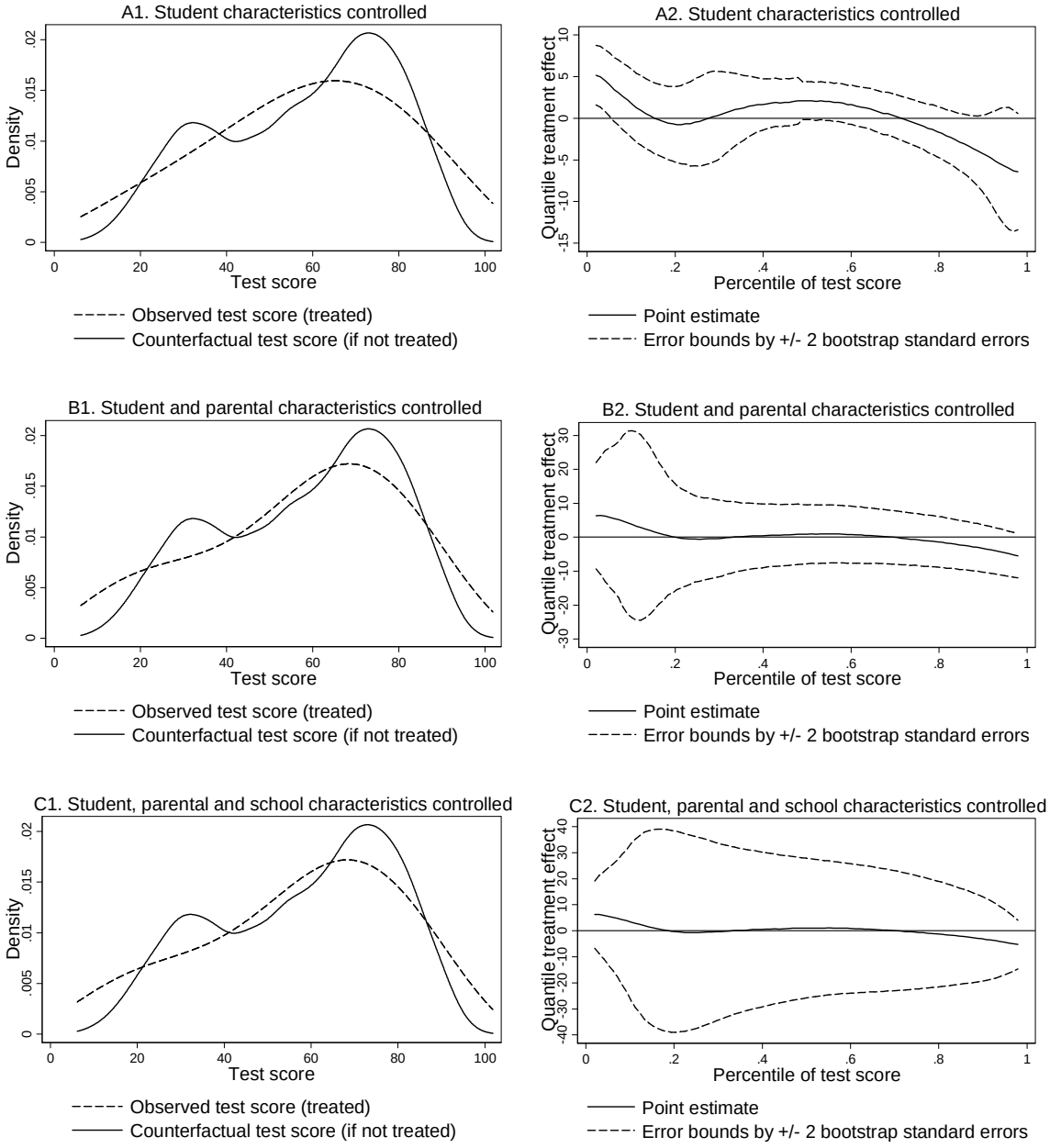


Figure 3-2: Estimated Probability Density Functions (Left) and Quantile Treatment Effects on the Treated (Right): English

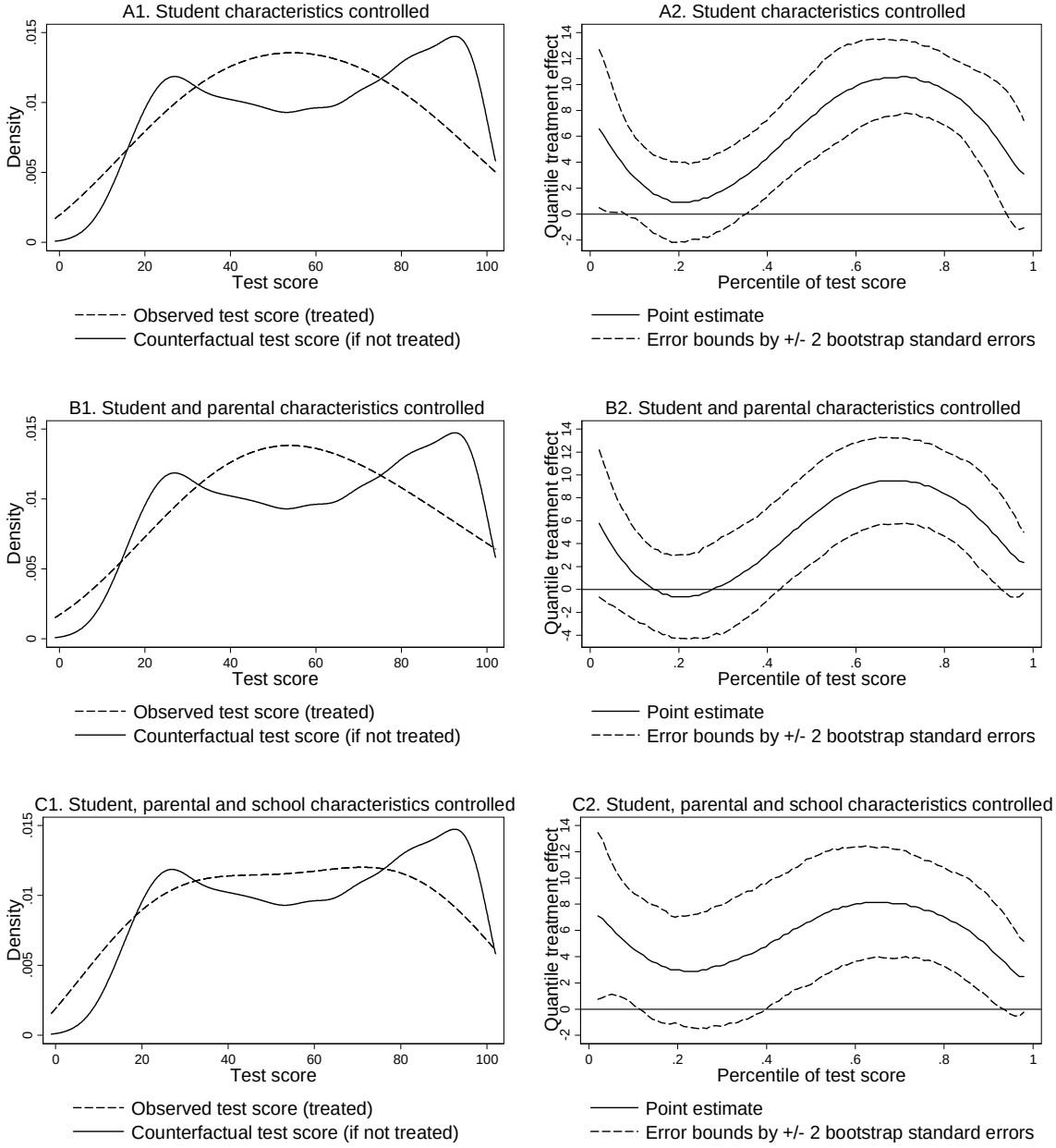
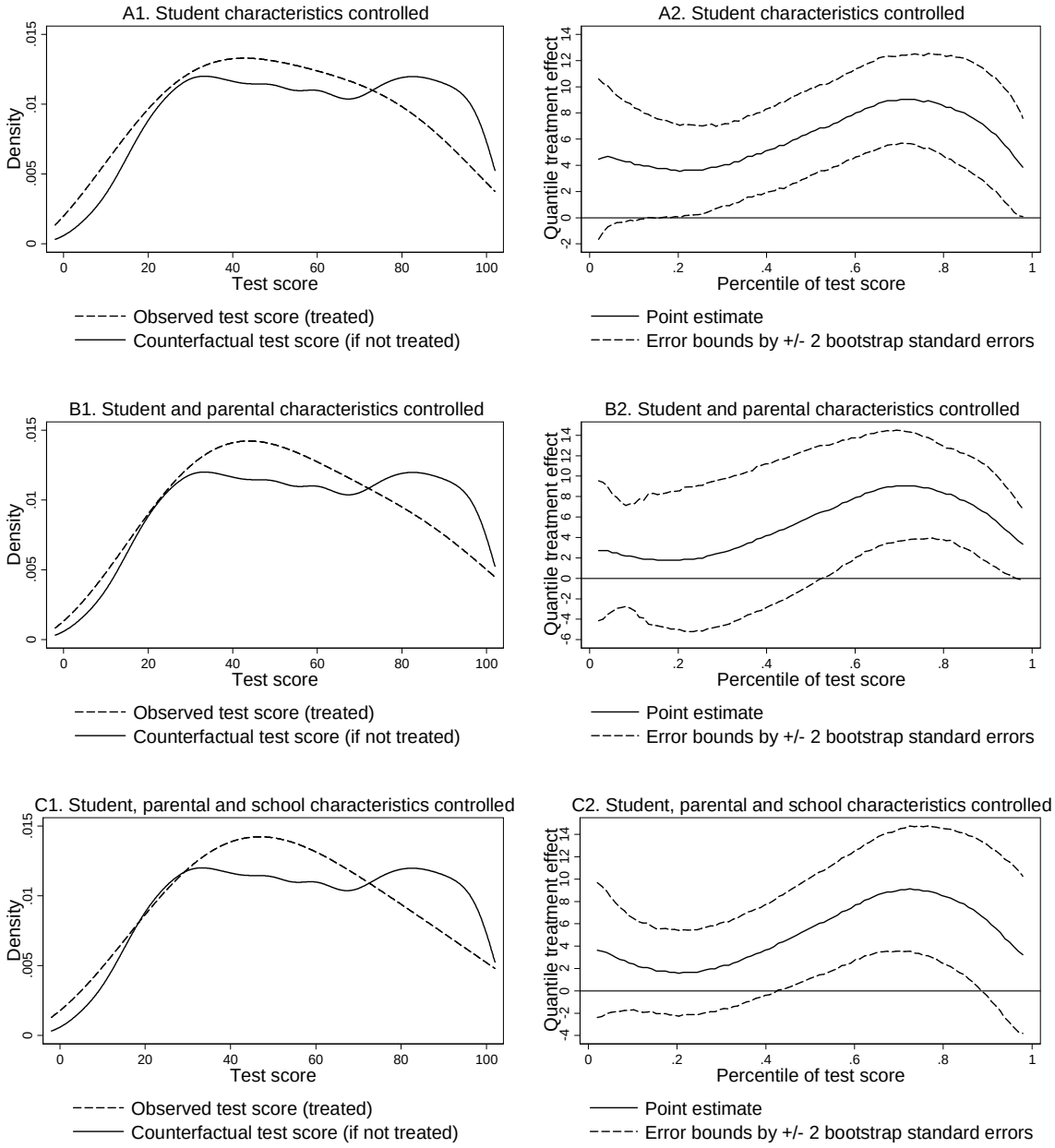


Figure 3-3: Estimated Probability Density Functions (Left) and Quantile Treatment Effects on the Treated (Right): Math



effect of private tutoring.

By integrating the estimated densities in figures 3-1, 3-2, and 3-3 over the range of test scores, we calculate the estimated CDFs of $Y_{i2}|D_i = 1$ and $Y_{i2}^0|D_i = 1$ for each of the three subjects. Given the results, we estimate the quantile treatment effects on the treated (QTT) in equation 3.12 by

$$\widehat{QTT}(\tau) = \widehat{F}^{-1}_{Y_{i2}|D_i=1}(\tau) - \widehat{F}^{-1}_{Y^0_{i2}|D_i=1}(\tau),\ \tau \in (0,1), \tag{3.17}$$

where $\widehat{F}^{-1}_W(\cdot)$ denotes the inverse of the estimated CDF of a random variable W. We compute standard errors of the $\widehat{QTT}(\tau)$ by bootstrapping[11] with 1000 iterations.

The estimation results for equation 3.17 are in the right columns of figures 3-1, 3-2, and 3-3. As in the mean effects, the patterns of the distributional effects in Korean are different from those in English and math. Recall that the mean effects of receiving private tutoring on Korean test scores are statistically insignificant from zero. Similarly, we do not find any statistically significant effect throughout the Korean test score distribution, either. This suggests that private tutoring has no effect on Korean test scores homogenously across students with different levels of pre-determined academic quality.

As opposed to the results for Korean, the estimation results for English and math reveal some important heterogeneity of the effects of private tutoring which is not captured by simply looking at its mean effects. Regardless of the choice of the covariate specifications,

[11] Following Hall (1992)[53] and Horowitz (2001)[61], when estimating the bootstrap standard errors, we use a four-times larger trimming parameter (i.e. undersmoothing) than the one chosen to compute the point estimates in equation 3.15.

the quantile treatment effects are at most modest or statistically insignificant at lower tails of the test score distributions. However, the effects rise to be positive and statistically significant in the middle of the distributions and become largest around the 70th to the 80th percentiles where they amount to about 8 to 9 points. After reaching their peaks, the effects revert to modest levels or be statistically insignificant from zero as moving up to upper tails of the distributions. This may be because students at top of the distributions have already scored almost 100 points, the maximum possible points of the achievement tests, before they receive the treatment, and hence any potential positive treatment effects for these top students cannot be captured by their achievement test scores. For example, the 90th percentiles of the baseline test scores of the students in the treatment group are 94 and 90 points in English and math, respectively. In sum, the QTT estimation results imply that private tutoring is helpful for students at upper half of the test score distributions but not for those at lower half of the distributions. This suggests that private tutoring mainly facilitates learning processes of students in good standings rather than serving as a remedial educational measure for students who are left behind.

3.5 Conclusion

We try to estimate the causal effect of receiving private tutoring on academic performance of nationally representative middle school students in Korea during 2006-2007. We use a binary indicator that takes 1 is a student receives private tutoring and 0 otherwise as our treatment variable. In this respect, our estimation results may be understood as measuring

the overall effect of receiving private tutoring. Our measures for the academic performance are achievement test scores of the students in three academic subjects: Korean, English, and math.

Using the data, we first examine the average effect of private tutoring. The raw mean test score gaps between students who receive private tutoring and those who do not are 1.62, 17.7, and 18.0 points in Korean, English, and math, respectively, where the maximum possible points of the tests are 100. When the differences in observables (student, parental, and school characteristics) and time-invariant unobservables (e.g. cognitive ability) between the two groups are controlled, the estimated mean effects reduce to .839 (statistically insignificant), 2.50, and 4.21 points in Korean, English, and math, respectively. These results suggest that the observed test score gaps between students who receive private tutoring and those who do not are largely driven by a positive selection into private tutoring in terms of both observables and time-invariant unobservable characteristics.

We also try to contribute to the literature by estimating the distributional effect of private tutoring, which has not been discussed in the previous studies. In Korean, we cannot reject no effect throughout the entire test score distribution, which implies that the effect of private tutoring is homogenously zero across students at different points of the test score distribution. In English and math, however, we find a positive effect in the upper half of the test score distributions but no effect in the lower half of the distribution. The effects reach their peaks around the 70th to the 80th percentile of the distributions where they amount to roughly 8 to 9 points in the both subjects. We find statistically insignificant effects around the top of the distributions. However, this seems to be simply because the top students have

already scored close to 100, the maximum possible points of the achievement tests, at their baselines and hence their test scores have no margin to be improved.

We should admit that our estimation results rely on a strong assumption that the time-varying unobservables are unconfounded (assumption 3.1). This assumption rules out the possibility that unobservable student-specific temporal shocks to test scores might affect students' decisions whether or not to receive private tutoring. Lessons from the program evaluation literature indicate that this is unlikely. For example, Ashenfelter (1978)[9] found that participants in job training programs tend to experience a decline in their pre-training earnings (Ashenfelter's dip). Similarly, if negative temporal shocks to pre-treatment test scores induce students to seek private tutoring, this would bias our estimation results. However, if there is a mean reversion in the unobservable temporal shocks, we can expect that the student would be more likely to face a positive temporal shock to his post-treatment test scores, which would affect our estimates to be upward-biased. Under this conjecture, we expect that our estimates for the effects of private tutoring are likely to *overestimate* the true effect.

Based on these results, we conclude that the effectiveness of receiving private tutoring, if any, is at most modest. In this sense, the results of this study are in line with those of Briggs (2001)[23], Kang (2007)[66], Gurun and Millimet (2008)[51], and Ryu and Kang (2013)[94] but not with the previous studies reporting strong positive effects of private tutoring such as Stevenson and Baker (1992)[100], Tansel and Bircan (2005)[103], Ha and Harphan (2005)[52], Ono (2007)[88], and Dang (2007)[31]. As opposed to these studies, however, we also contribute to the literature by finding that the extent of the effectiveness of private tutoring

varies substantially across students with different levels of academic quality. Particularly, it seems that students with good academic quality tend to benefit more from private tutoring than those at the bottom of the test score distributions. This suggests that private tutoring in Korea may not be an effective remedial educational measure for students left behind, although it seems to have a modest effect of facilitating learning processes of students in good standing.

Appendix A

Appendix for Chapter 1

A.1 Details on reweighting procedure in figure 1-4

This section summarizes the reweighting procedure, originally suggested by DiNardo et al. (1996)[37], that I applied to the CPS job tenure data in figure 1-4 in section 1.5. The goal is to calculate the counterfactual tenure distributions in the years 1982-1988 and 1989-1999, had the worker characteristics remained at the 1978-1981 level. Comparing such counterfactual distributions with the realized 1978-1981 tenure distribution observed in the data can be understood as investigating how the tenure distribution has changed over time controlling for the changes in the worker characteristics. In the following presentation of the method, I follow discussions in DiNardo et al. (1996)[37] and Shimizutani and Yokoyama (2009)[97].

For this section, let us denote years 1978-1981, 1982-1988, and 1989-1999 as period 1, 2, and 3, respectively. The probability density function (PDF) of job tenure in period 1 can be

written as

$$f\left(y|T=1\right)=\int f\left(y|X=x,T=1\right)g\left(x|T=1\right)dX \tag{A.1}$$

where Y and X represent a workers' job tenure and demographic characteristics whose PDF is given as $f\left(\cdot\right)$ and $g\left(\cdot\right)$, respectively; and $T\in\left\{1,2,3\right\}$ denotes the period in which the observation is recorded. Similarly, the PDF of job tenure in period 2 is

$$f\left(y|T=2\right)=\int f\left(y|X=x,T=2\right)g\left(x|T=2\right)dX \tag{A.2}$$

Note that any difference between $f\left(y|T=2\right)$ and $f\left(y|T=1\right)$ might result from two reasons. First, the relationship between job tenure(Y) and worker characteristics (X) has changed over time. More precisely, the joint distribution between Y and X had changed over time, hence, $f\left(Y|X,T=1\right)\neq f\left(Y|X,T=2\right)$. Second, although the relationship between job tenure and worker characteristics are the same between the two periods, i.e. $f\left(Y|X,T=1\right)=f\left(Y|X,T=2\right)$, the composition of workers has changed, or more precisely, the distribution of worker characteristics has changed over time: $g\left(X|T=1\right)\neq g\left(X|T=2\right)$. Depending on which factor mainly account for the difference between $f\left(y|T=2\right)$ and $f\left(y|T=1\right)$, the observed change in tenure distribution can be understood as reflecting either the structural change in job stability (if the former is the main reason) or simply the changes in worker composition (if the latter one is the main reason).

In order to investigate whether the shift of tenure distribution toward long jobs observed in my sample is structural or compositional, I consider the counterfactual PDF of job tenure in period 2, had the distribution of worker characteristics remained at the same level as

period 1:

$$\widetilde{f}\left(y|T=2\right)=\int f\left(y|X=x,T=2\right)g\left(x|T=1\right)dX \tag{A.3}$$

Note that the counterfactual density, $\widetilde{f}\left(y|T=2\right)$, is calculated from reweighting the conditional density of Y given X in period 2, $f\left(Y|X,T=2\right)$, using the distribution of X in period 1, $g\left(X|T=1\right)$, as weights. Since the distribution of X is controlled between equation A.1 and A.3, any difference in the two densities can be understood as reflecting $f\left(Y|X,T=1\right)\neq f\left(Y|X,T=2\right)$: the structural relationship between tenure and worker characteristics has changed, or in other words, change in job stability.

Unfortunately, estimating equation A.3 directly is challenging due to the curse of high dimensionality. DiNardo et al. (1996)[37] suggested that equation A.3 can be estimated semiparametrically by a simple use of logit model. Particularly, equation A.3 can be rewritten as

$$\begin{aligned}
\widetilde{f}\left(y|T=2\right) &= \int f\left(y|X=x,T=2\right)g\left(x|T=1\right)dX && \text{(A.4)}\\
&= \int f\left(y|X=x,T=2\right)g\left(x|T=2\right)\frac{g\left(x|T=1\right)}{g\left(x|T=2\right)}dX && \text{(A.5)}\\
&= \int f\left(y|X=x,T=2\right)g\left(x|T=2\right) && \text{(A.6)}\\
&\quad \frac{g\left(T=1|X=x\right)\Pr\left[T=2\right]}{g\left(T=2|X=x\right)\Pr\left[T=1\right]}dX\\
&= \int f\left(y|X=x,T=2\right)g\left(x|T=2\right)\omega\left(x\right)dX && \text{(A.7)}
\end{aligned}$$

where equation A.6 holds by the Bayes rule. Following DiNardo et al. (1996)[37], I calculate the reweighting function $\omega\left(X\right)\equiv\frac{g(T=1|X)\Pr[T=2]}{g(T=2|X)\Pr[T=1]}$ by estimating $g\left(T=1|X\right)$ and $g\left(T=2|X\right)$

using logit model.

Similarly, I reweight period 3 tenure density using the distribution of worker characteristics in period 1 and calculate $\widetilde{f}(y|T=3)$. I plot the kernel density estimates of $f(y|T=1)$, $\widetilde{f}(y|T=2)$, and $\widetilde{f}(y|T=3)$ as well as their cumulative distribution functions in figure 1-4 in section 1.5.

A.2 Details on deriving equation 1.9

This section illustrates how I apply the model by Bonhomme and Sauder (2011)[20] to my study and derive equation 1.9. I refer to section II.B of Bonhomme and Sauder (2011)[20] for the following illustration of their method.

Intuitively, the model by Bonhomme and Sauder (2011)[20] can be understood as a logical extension of the standard difference-in-differences (DID) model. In the standard DID model, the average treatment effect (ATE) can be identified by comparing the change in the mean outcome of the treatment group between pre-treatment and post-treatment periods with the contemporaneous change in the mean outcome of the control group. Analogously, in the distributional DID model by Bonhomme and Sauder (2011)[20], the quantile treatment effects (QTE) can be identified by comparing the change in the outcome distribution of the treatment group between pre-treatment and post-treatment periods with the contemporaneous change in the outcome distribution of the control group. Just as a counterfactual mean is identified by adding and subtracting realized, hence empirically observable, means in the standard DID model, a counterfactual density is identified by manipulating realized densities

in the distributional DID model. Of course, manipulating densities is not as simple as manipulating means. However, the seemingly challenging calculation can be easily performed by a simple use of the characteristic function, which is defined as:

Definition The characteristic function of a random variable W, $\Psi_W(h)$, is defined as $\Psi_W(h) \equiv E[\exp(jhW)]$ where $j = \sqrt{-1}$ and $h \in \mathbf{R}$.

For any real-valued random variable, it is known that the PDF of the random variable can be achieved by the Fourier inversion of its characteristic function:

$$f_W(w) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-jhw) \Psi_W(h) \, dh \tag{A.8}$$

Thus, instead of identifying the PDF of the counterfactual $U_{is2}^0 \mid D_s = 1$ directly, I can identify its characteristic function first and then transform it to its density by equation A.8.

Under the assumptions 1.1 and 1.2 and the additive structure of U_{is2}^0 and U_{is1} given in equations 1.7 and 1.8, the characteristic function of the *counterfactual* $U_{is2}^0 | D_s = 1$ is identified as a function of three characteristic functions of the *observable* $U_{is2} | D_s = 0$, $U_{is1} | D_s = 1$,

and $U_{is1}|D_s = 0$ as follows (Bonhomme and Sauder, 2011[20]):

$$
\begin{aligned}
& \Psi_{U^0_{is2}|D_s=1}(h) \\
= \;& E\left[\exp\left(jhU^0_{is2}\right)|D_s=1\right] && \text{(A.9)} \\
= \;& E\left[\exp\left(jh\left\{\alpha^0_2+\eta_s+v^0_{is2}\right\}\right)|D_s=1\right] && \text{(A.10)} \\
= \;& \exp\left(jh\alpha^0_2\right)E\left[\exp\left(jh\eta_s\right)|D_s=1\right]E\left[\exp\left(jhv^0_{is2}\right)|D_s=1\right] && \text{(A.11)} \\
= \;& \exp\left(jh\alpha^0_2\right)E\left[\exp\left(jh\eta_s\right)|D_s=1\right]E\left[\exp\left(jhv^0_{is2}\right)|D_s=0\right] && \text{(A.12)} \\
= \;& \exp\left(jh\alpha^0_2\right)E\left[\exp\left(jh\eta_s\right)|D_s=0\right]E\left[\exp\left(jhv^0_{is2}\right)|D_s=0\right] && \text{(A.13)} \\
& \frac{E\left[\exp\left(jh\eta_s\right)|D_s=1\right]}{E\left[\exp\left(jh\eta_s\right)|D_s=0\right]} \\
= \;& E\left[\exp\left(jhU_{is2}\right)|D_s=0\right]\frac{E\left[\exp\left(jhU_{is1}\right)|D_s=1\right]}{E\left[\exp\left(jhU_{is1}\right)|D_s=0\right]} && \text{(A.14)} \\
= \;& \Psi_{U_{is2}|D_s=0}(h)\frac{\Psi_{U_{is1}|D_s=1}(h)}{\Psi_{U_{is1}|D_s=0}(h)} && \text{(A.15)}
\end{aligned}
$$

where equations A.9 and A.15 are from the definition of characteristic function; equation A.10 is from the additive structure of $U^0_{is2} = \alpha^0_2+\eta_s+v^0_{is2}$ given in equation 1.8; equation A.11 holds because α^0_2 is non-random and η_s is independent of v^0_{is2} conditional on D_s (assumption 1.2); equation A.12 holds because v^0_{is2} is independent of D_s (assumption 1.1.); finally, equation A.14 holds because

$$
\begin{aligned}
& \exp\left(jh\alpha^0_2\right)E\left[\exp\left(jh\eta_s\right)|D_s=0\right]E\left[\exp\left(jhv^0_{is2}\right)|D_s=0\right] \\
= \;& E\left[\exp\left(jh\left\{\alpha^0_2+\eta_s+v^0_{is2}\right\}\right)|D_s=1\right] && \text{(A.16)} \\
= \;& E\left[\exp\left(jhU_{is2}\right)|D_s=0\right] && \text{(A.17)}
\end{aligned}
$$

and also because

$$\begin{aligned}
&\frac{E\left[\exp\left(jh\eta_s\right)|D_s=1\right]}{E\left[\exp\left(jh\eta_s\right)|D_s=0\right]} \\
= &\frac{\exp\left(jh\alpha_1\right)E\left[\exp\left(jh\eta_s\right)|D_s=1\right]E\left[\exp\left(jhv_{1s1}\right)|D_s=1\right]}{\exp\left(jh\alpha_1\right)E\left[\exp\left(jh\eta_s\right)|D_s=0\right]E\left[\exp\left(jhv_{1s1}\right)|D_s=1\right]} &\text{(A.18)}\\
= &\frac{\exp\left(jh\alpha_1\right)E\left[\exp\left(jh\eta_s\right)|D_s=1\right]E\left[\exp\left(jhv_{1s1}\right)|D_s=1\right]}{\exp\left(jh\alpha_1\right)E\left[\exp\left(jh\eta_s\right)|D_s=0\right]E\left[\exp\left(jhv_{1s1}\right)|D_s=0\right]} &\text{(A.19)}\\
= &\frac{E\left[\exp\left(jh\left\{\alpha_1+\eta_s+v_{is1}\right\}\right)|D_s=1\right]}{E\left[\exp\left(jh\left\{\alpha_1+\eta_s+v_{is1}\right\}\right)|D_s=0\right]} &\text{(A.20)}\\
= &\frac{E\left[\exp\left(jhU_{is1}\right)|D_s=1\right]}{E\left[\exp\left(jhU_{is1}\right)|D_s=0\right]} &\text{(A.21)}
\end{aligned}$$

Substituting $\Psi_{U^0_{is2}|D_s=1}(h)=\Psi_{U_{is2}|D_s=0}(h)\frac{\Psi_{U_{is1}|D_s=1}(h)}{\Psi_{U_{is1}|D_s=0}(h)}$ into equation A.8 yields equation 1.9.

A.3 Details on choosing the truncation parameter in equation 1.10

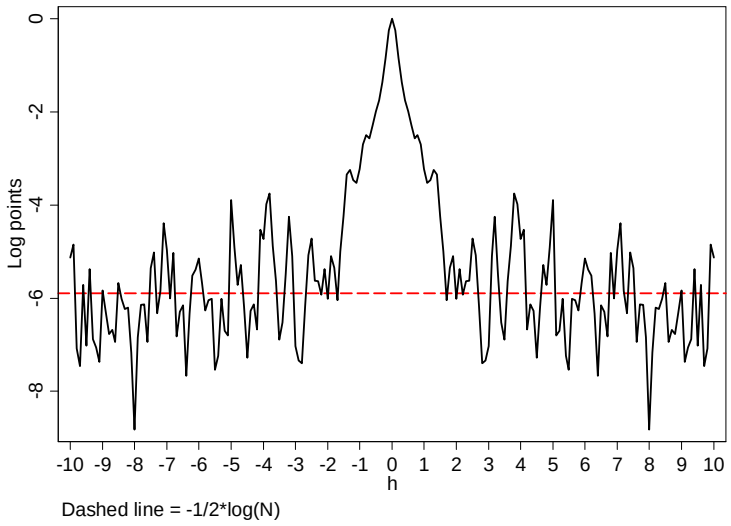
According to Diggle and Hall (1993)[36], an optimal T_N must solve (Bonhomme and Sauder, 2011[20])

$$\log\left|\widehat{\Psi}_{U^0_{is2}|D_s=1}(T_N)\right|=-\frac{1}{2}\log N \tag{A.22}$$

Figure A-1 shows that there are many solutions for equation A.22. Among them, I choose $T_N=1.8$ for my estimation results. Following Bonhomme and Sauder (2009[19], 2011[20]), I also estimate equation A.22 using alternative values of $T_N=5$, 10, and 20 for robustness check. Figure A-2 compares estimated densities based on the different use of truncation

parameters with the result from my choice $T_N = 1.8$. Although bumpiness of the estimated densities increase with larger truncation parameters (i.e. undersmoothing), overall shape of the densities are similar. This indicates that cumulative distributions and quantile treatment effects, which are calculated by integrating the densities, would not be affected much by the choice of truncation parameters.

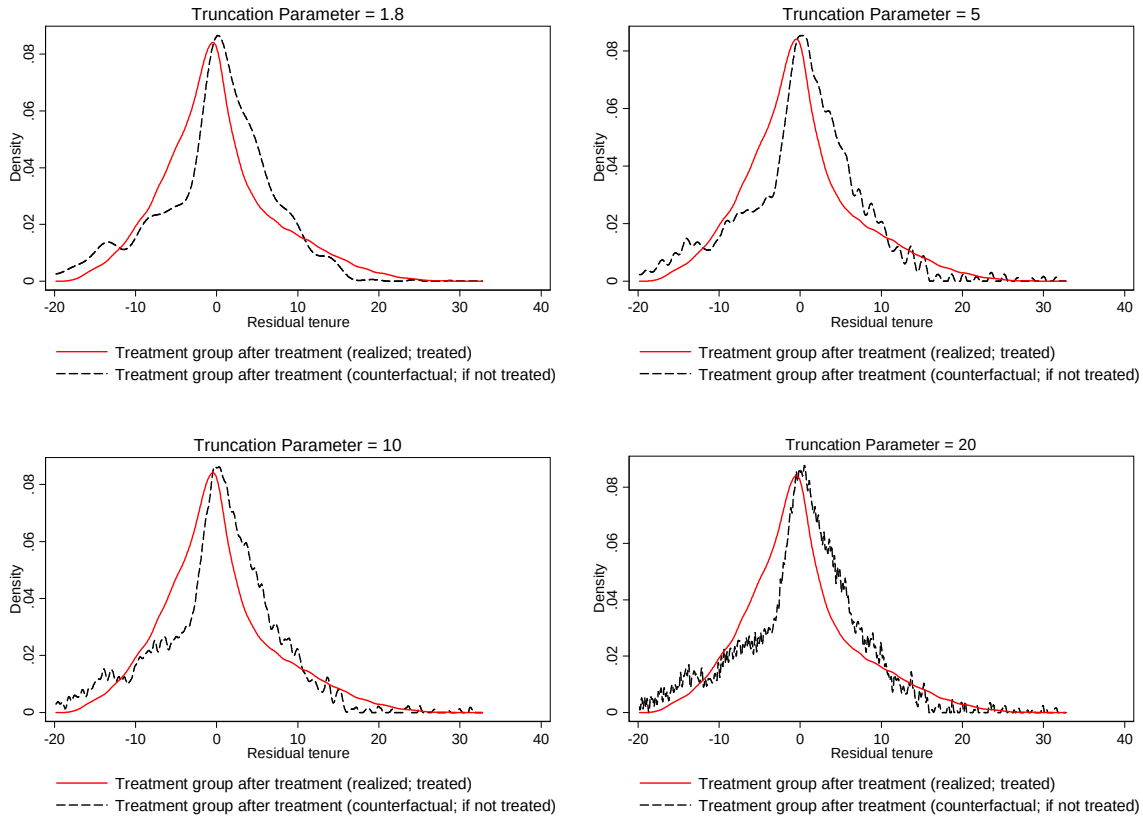
Figure A-1: Log of Absolute Value of Estimated Characteristic Function



A.4 Results based on the alternative legal classification

Table A.1 lists the timing of the law adoption based on the Walsh and Schwarz (1996)[108] legal classification. Among the 50 states excluding the District of Columbia, 17 states are coded inconsistently between the legal classification by Autor et al. (2006)[12] that I used for my main results and the classification by Walsh and Schwarz (1996)[108]. Those 17 states are marked with an asterisk in table A.1. For robustness check, I modify only what

Figure A-2: Robustness Check for Alternative Use of Truncation Parameters



I view the three leading discrepancies, which are marked boldfaced in table A.1, out of the 17 states in the legal classification by Walsh and Schwarz (1996)[108] and replicate my main estimation results.

Table A.1: Timing of Law Adoption by Alternative Legal Classification)

Year of Adoption	States	# of States
1976	Oklahoma	1
1978	Oregon	1
1980	New Mexico, Michigan	2
1981	California*, Utah*, Tennessee	3
1982	New York	1
1983	Missouri, South Dakota Minnesota, Alaska, Arizona Nevada, Kentucky, Nebraska	8
1984	North Dakota, Colorado* Washington*, Kansas	4
1985	Maine*, Maryland, New Jersey Wisconsin, Ohio* Wyoming, Connecticut, Mississippi*	8
1986	West Virginia, Idaho*, Vermont*, Hawaii	4
1987	**Illinois***, Arkansas*, Montana South Carolina, Alabama	5
1988	New Hampshire	1
1989	Iowa*	1
Never had laws	**Indiana***, **Texas*** Massachusetts*, Virginia*, Rhode Island, Pennsylvania, Delaware, North Carolina, Georgia, Florida, Louisiana	11
Total		50

Table A.2: Robustness Check: Effects on Mean Log Tenure

Variables	Specification (1)	(2)	(3)	(4)
Implied-contract exception	-.023	-.042**	-.028	-.035*
(Standard error)	(.022)	(.016)	(.020)	(.018)
Covariates:				
Individual characteristics	Yes	Yes	Yes	Yes
State labor market conditions	No	Yes	No	Yes
State and year dummies	Yes	Yes	Yes	Yes
Region by year dummies	No	No	Yes	Yes
R-squared	.232	.232	.233	.233
Number of observations	246112	246112	246112	246112

Table A.3: Robustness Check: Effects on Pr[10 years or longer]

Variables	Specification (1)	(2)	(3)	(4)
Implied-contract exception	-.105***	-.126***	-.112***	-.114***
(Standard error)	(.036)	(.031)	(.034)	(.033)
[Marginal effect]	[-.026]	[-.031]	[-.028]	[-.029]
Covariates:				
Individual characteristics	Yes	Yes	Yes	Yes
State labor market conditions	No	Yes	No	Yes
State and year dummies	Yes	Yes	Yes	Yes
Region by year dummies	No	No	Yes	Yes
Age restriction	≥35	≥35	≥35	≥35
Pseudo R-squared	.061	.062	.063	.063
Number of observations	139094	139094	139094	139094

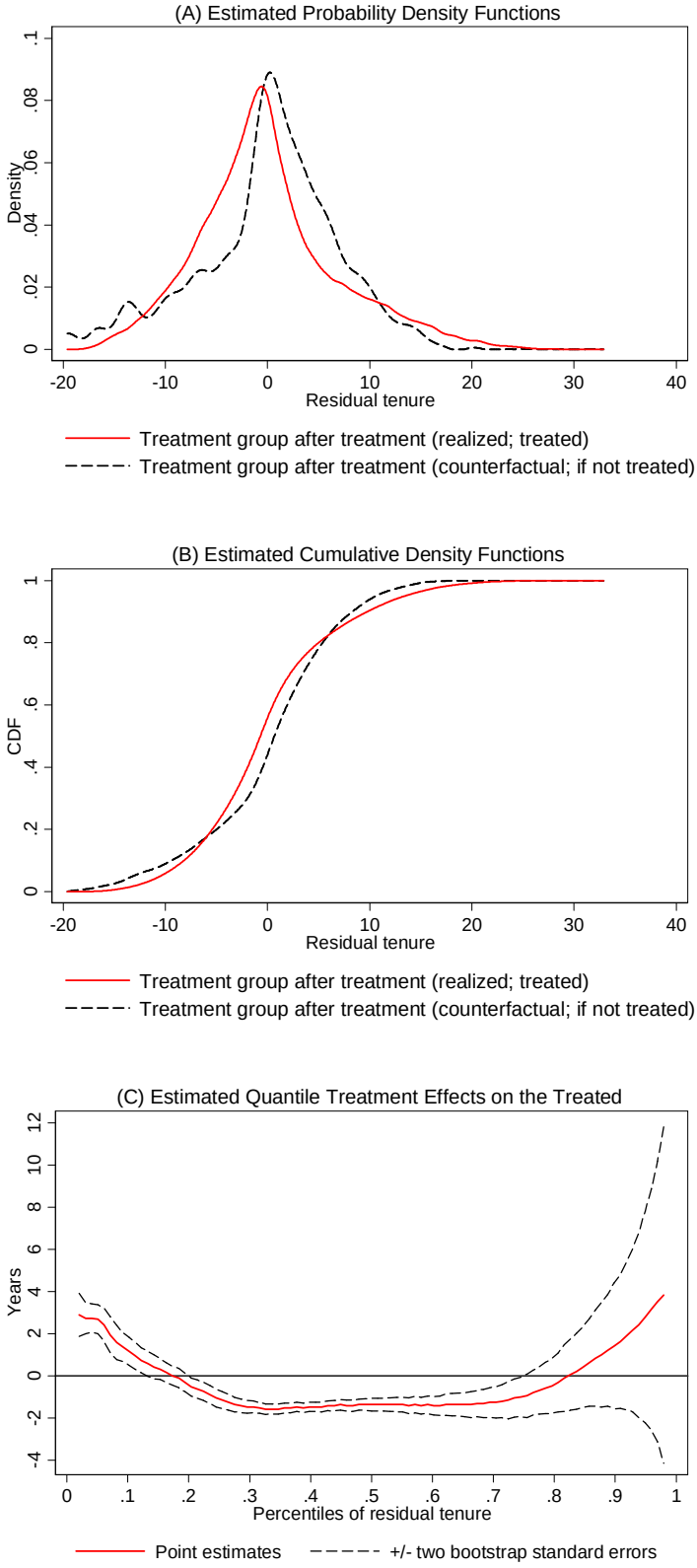
Table A.4: Robustness Check: Effects on Pr[20 years or longer]

Variables	Specification (1)	(2)	(3)	(4)
Implied-contract exception	-.074	-.064	-.058	-.069
(Standard error)	(.050)	(.046)	(.055)	(.052)
[Marginal effect]	[-.015]	[-.013]	[-.012]	[-.014]
Covariates:				
Individual characteristics	Yes	Yes	Yes	Yes
State labor market conditions	No	Yes	No	Yes
State and year dummies	Yes	Yes	Yes	Yes
Region by year dummies	No	No	Yes	Yes
Age restriction	≥45	≥45	≥45	≥45
Pseudo R-squared	.047	.048	.049	.049
Number of observations	74175	74175	74175	74175

Table A.5: Robustness Check: Effects on Pr[five years or less]

Variables	Specification (1)	(2)	(3)	(4)
Implied-contract exception	.072	.107***	.090***	.101***
(Standard error)	(.041)	(.028)	(.034)	(.031)
[Marginal effect]	[.018]	[.027]	[.023]	[.025]
Covariates:				
Individual characteristics	Yes	Yes	Yes	Yes
State labor market conditions	No	Yes	No	Yes
State and year dummies	Yes	Yes	Yes	Yes
Region by year dummies	No	No	Yes	Yes
Age restriction	≥25	≥25	≥25	≥25
Pseudo R-squared	.114	.115	.115	.115
Number of observations	215329	215329	215329	215329

Figure A-3: Robustness Check: Effects on the Residual Tenure Distribution



Appendix B

Appendix for Chapter 3

B.1 Details on deriving equation 3.10

This section illustrates how I apply the model by Bonhomme and Sauder (2011)[20] to my study and derive equation 3.10. I refer to section II.B of Bonhomme and Sauder (2011)[20] for the following illustration of their method.

The ATT can be written as

$$\begin{aligned} ATT &= E\left[Y_{i2}^1 - Y_{i2}^0 \,|\, D_i = 1\right] && \text{(B.1)} \\ &= \int E\left[Y_{i2}^1 - Y_{i2}^0 \,|\, X_i,\, D_i = 1\right] dP\left(X_i \,|\, D_i = 1\right) && \text{(B.2)} \\ &= \int E\left[Y_{i2} \,|\, X_i,\, D_i = 1\right] - E\left[Y_{i2}^0 \,|\, X_i,\, D_i = 1\right] dP\left(X_i \,|\, D_i = 1\right), && \text{(B.3)} \end{aligned}$$

where $E\left[Y_{i2}^0 \,|\, X_i,\, D_i = 1\right]$ needs to be identified. By the additive structure of the educational

production function in equations 1.7 and 1.8, we have

$$\begin{aligned} E\left[Y_{i2}^{0} \mid X_i,\, D_i = 1\right] &= E\left[f_2^0\left(X_i\right) + \eta_i + v_{i2}^{0} \mid X_i,\, D_i = 1\right] && \text{(B.4)} \\ &= f_2^0\left(X_i\right) + E\left[\eta_i \mid X_i,\, D_i = 1\right] + E\left[v_{i2}^{0} \mid X_i,\, D_i = 1\right] && \text{(B.5)} \\ &= f_2^0\left(X_i\right) + E\left[\eta_i \mid X_i,\, D_i = 0\right] + E\left[v_{i2}^{0} \mid X_i,\, D_i = 1\right] && \text{(B.6)} \\ &\quad + f_1\left(X_i\right) + E\left[\eta_i \mid X_i,\, D_i = 1\right] + E\left[v_{i1} \mid X_i,\, D_i = 1\right] \\ &\quad - f_1\left(X_i\right) - E\left[\eta_i \mid X_i,\, D_i = 0\right] - E\left[v_{i1} \mid X_i,\, D_i = 1\right], \end{aligned}$$

Under the selection on observables and time-invariant unobservables assumption (assumption 3.1), we also have

$$\begin{aligned} E\left[v_{i1} \mid X_i,\, D_i = 1\right] &= E\left[v_{i1} \mid X_i,\, D_i = 0\right] && \text{(B.7)} \\ E\left[v_{i2}^{0} \mid X_i,\, D_i = 1\right] &= E\left[v_{i2}^{0} \mid X_i,\, D_i = 0\right] && \text{(B.8)} \end{aligned}$$

Substituting equations B.7 and B.8 into equation B.6, $E\left[Y_{i2}^{0} \mid X_i,\, D_i = 1\right]$ is identified as

$$E\left[Y_{i2}^{0} \mid X_i,\, D_i = 1\right] = E\left[Y_{i2} | X_i, D_i = 0\right] + E\left[Y_{i1} | X_i, D_i = 1\right] - E\left[Y_{i1} | X_i, D_i = 0\right] \quad \text{(B.9)}$$

Substituting B.9 into B.3, the ATT is identified as the following difference-in-differences estimand:

$$ATT = \int \left\{E\left[Y_{i2} - Y_{i1} \mid X_i,\, D_i = 1\right] - E\left[Y_{i2} - Y_{i1} \mid X_i,\, D_i = 0\right]\right\} dP\left(X_i \mid D_i = 1\right) \quad \text{(B.10)}$$

Estimating equation B.10 nonparametrically is infeasible due to the curse of dimensionality problem. To proceed, we use the Lemma 3.1 in Abadie (2005)[1] following Bonhomme and Sauder (2011)[20]. Under the selection on observables and time-invariant unobservables assumption (assumption 3.1) and the common support assumption (assumption 3.2), Abadie (2005)[1] has shown that

$$E\left[Y_{i2}^{1} - Y_{i2}^{0} \mid X_i,\, D_i = 1\right] = E\left[\omega_i \left(Y_{i2} - Y_{i1}\right) \mid X_i\right], \tag{B.11}$$

where we have denoted as

$$\omega_i = \frac{D_i - P\left[D_i = 1 \mid X_i\right]}{P\left[D_i = 1 \mid X_i\right] P\left[D_i = 0 \mid X_i\right]} \tag{B.12}$$

Summing B.11 over X_i, the ATT can be written as (Abadie, 2005[1])

$$\begin{aligned} ATT &= \int E\left[\omega_i \left(Y_{i2} - Y_{i1}\right) \mid X_i\right] dP\left(X_i \mid D_i = 1\right) && \text{(B.13)} \\ &= \int E\left[\omega_i \left(Y_{i2} - Y_{i1}\right) \mid X_i\right] \frac{\Pr\left[D_i = 1 \mid X_i\right]}{\Pr\left[D_i = 1\right]} dP\left(X_i\right) && \text{(B.14)} \\ &= \frac{1}{\Pr\left[D_i = 1\right]} \int E\left[\omega_i \Pr\left[D_i = 1 \mid X_i\right] \left(Y_{i2} - Y_{i1}\right) \mid X_i\right] dP\left(X_i\right) && \text{(B.15)} \end{aligned}$$

Substituting B.12 into B.15 yields (Abadie, 2005[1])

$$
\begin{aligned}
& ATT \\
= & \frac{1}{\Pr\left[D_i = 1\right]} \int E\left[\frac{D_i - \Pr\left[D_i = 1 \,|\, X_i\right]}{\Pr\left[D_i = 0 \,|\, X_i\right]} \left(Y_{i2} - Y_{i1}\right) | X_i\right] dP\left(X_i\right) && \text{(B.16)} \\
= & \frac{1}{\Pr\left[D_i = 1\right]} E\left[\frac{D_i - \Pr\left[D_i = 1 \,|\, X_i\right]}{\Pr\left[D_i = 0 \,|\, X_i\right]} \left(Y_{i2} - Y_{i1}\right)\right], && \text{(B.17)}
\end{aligned}
$$

which is equation 3.10.

B.2 Details on deriving equation 3.13

This section illustrates how I apply the model by Bonhomme and Sauder (2011)[20] to my study and derive equation 3.13. I refer to section II.B of Bonhomme and Sauder (2011)[20] for the following illustration of their method.

For any real-valued random variable W, it is known that its probability density function can be obtained by the inverse Fourier transformation of its characteristic function, $\Psi_W\left(t\right) \equiv E\left[\exp\left(jtW\right)\right]$:

$$
f_W\left(w\right) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\left(-jtw\right) \Psi_W\left(t\right) dt, \tag{B.18}
$$

where $j = \sqrt{-1}$ and $t \in R$. Hence, identifying the characteristic function of $Y_{i2}^0 | D_i = 1$ suffices the identification of its density.

Let $\Psi_{Y_{i2}^0|D_i=1}(t)$ denote the characteristic function of $Y_{i2}^0|D_i=1$. By definition,

$$\begin{aligned}\Psi_{Y_{i2}^0|D_i=1}(t) &= E\left[\exp\left(jtY_{i2}^0\right)|D_i=1\right] &\text{(B.19)}\\ &= \int \Psi_{Y_{i2}^0|D_i=1,X_i}(t|X_i)dP\left(X_i|D_i=1\right) &\text{(B.20)}\end{aligned}$$

Bonhomme and Sauder (2011[20]; Theorem 2) has shown that, when educational production functions take the forms of equations 3.8 and 3.9 and assumptions 3.1, 3.2, and 3.3 hold, the conditional characteristic function of the counterfactual $Y_{i2}^0|D_i=1$ is identified as a function of three conditional characteristic functions of the realized $Y_{i2}|D_i=0$, $Y_{i1}|D_i=1$, and $Y_{i1}|D_i=0$:

$$\Psi_{Y_{i2}^0|D_i=1,X_i}(t|x)=\frac{\Psi_{Y_{i1}|D_i=1,X_i}(t|x)}{\Psi_{Y_{i1}|D_i=0,X_i}(t|x)}\Psi_{Y_{i2}|D_i=0,X_i}(t) \quad \text{(B.21)}$$

Substituting B.21 into B.20 yields:

$$\begin{aligned}\Psi_{Y_{i2}^0|D_i=1}(t) &= \int \frac{\Psi_{Y_{i1}|D_i=1,X_i}(t|x)}{\Psi_{Y_{i1}|D_i=0,X_i}(t|x)}\Psi_{Y_{i2}|D_i=0,X_i}(t|x)\,dP\left(X_i|D_i=1\right) &\text{(B.22)}\\ &= \int \frac{\Psi_{Y_{i1}|D_i=1,X_i}(t|x)}{\Psi_{Y_{i1}|D_i=0,X_i}(t|x)}\Psi_{Y_{i2}|D_i=0,X_i}(t|x) &\text{(B.23)}\\ &\quad \frac{P\left(D_i=1|X_i\right)}{P\left(D_i=1\right)}dP\left(X_i\right)\\ &= \frac{1}{P\left(D_i=1\right)}\int \frac{P\left(D_i=1|X_i\right)}{P\left(D_i=0|X_i\right)}\frac{\Psi_{Y_{i1}|D_i=1,X_i}(t|x)}{\Psi_{Y_{i1}|D_i=0,X_i}(t|x)} &\text{(B.24)}\\ &\quad \Psi_{Y_{i2}|D_i=0,X_i}(t|x)\,P\left(D_i=0|X_i\right)dP\left(X_i\right)\end{aligned}$$

Note that it holds

$$\Psi_{Y_{i2}|D_i=0,X_i}(t|x)\,P(D_i=0|X_i) \tag{B.25}$$

$$= E\left[\exp(jtY_{i2})\,|D_i=0,X_i\right]P(D_i=0|X_i)$$

$$= E\left[(1-D_i)\exp(jtY_{i2})\,|X_i\right] \tag{B.26}$$

Similarly,

$$\Psi_{Y_{i1}|D_i=0,X_i}(t|x)\,P(D_i=0|X_i) = E\left[(1-D_i)\exp(jtY_{i1})\,|X_i\right] \tag{B.27}$$

$$\Psi_{Y_{i1}|D_i=1,X_i}(t|x)\,P(D_i=1|X_i) = E\left[D_i\exp(jtY_{i1})\,|X_i\right] \tag{B.28}$$

Substituting B.26, B.27, and B.28 into B.24 yields

$$\Psi_{Y_{i2}^0|D_i=1}(t) = \frac{1}{p_D}\int\omega(t|X_i)\,E\left[(1-D_i)\exp(jtY_{i2})\,|X_i\right]dP(X_i) \tag{B.29}$$

$$= \frac{1}{p_D}E\left[\omega(t|X_i)\,(1-D_i)\exp(jtY_{i2})\right] \tag{B.30}$$

where we have denoted as

$$\omega(t|X_i) \equiv \frac{E[D_i exp(jtY_{i1})|X_i]}{E[(1-D_i)exp(jtY_{i1})|X_i]} \tag{B.31}$$

Substituting equation B.30 into equation A.8 yields equation 3.13.

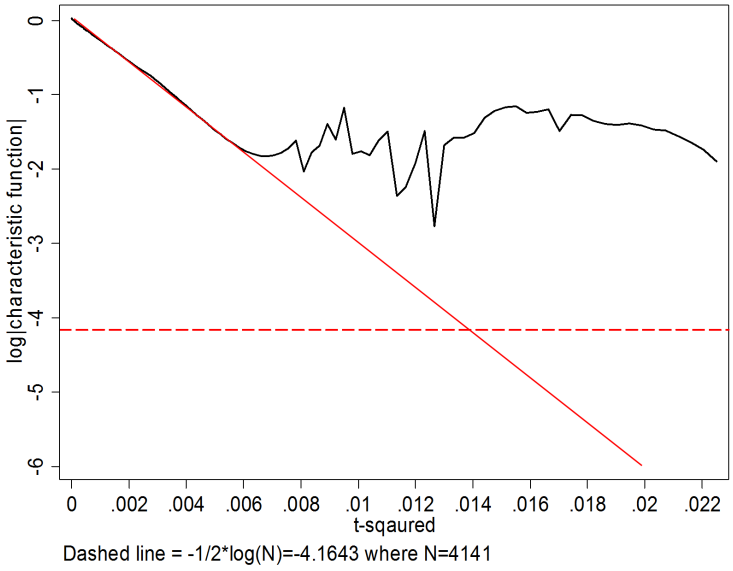
B.3 Details on chooing the truncation parameter in equation 3.15.

According to Diggle and Hall (1993)[36], an optimal T_N must satisfy (Bonhomme and Sauder, 2011[20])

$$\log \left| \widehat{\Psi}_{Y_{i2}^0|D_i=1}(T_N) \right| = -\frac{1}{2} \log N \tag{B.32}$$

Figure A-1 plots the estimated $\log \left| \widehat{\Psi}_{Y_{i2}^0|D_i=1}(t) \right|$ against t^2 for the Korean sample when student, parental, and school characteristics are used as covariates. Following Bonhomme and Sauder (2009[19], 2011[20]), we extrapolate the (almost) linear part of $\log \left| \widehat{\Psi}_{Y_{i2}^0|D_i=1}(t) \right|$ and find the value of t where the extrapolated line crosses $-\frac{1}{2} \log N$. This yileds $t \approx \sqrt{.014} \approx .118$, which we use as the T_N. For other subject samples and covariate specifications, we determine T_N in a similar way.

Figure B-1: Log of Absolute Value of Estimated Characteristic Function



Bibliography

[1] A. Abadie. Semiparametric Difference-in-Differences Estimators. *The Review of Economic Studies*, 72(1):pp. 1–19, 2005.

[2] K.G. Abraham and J.L. Medoff. Length of Service and Layoffs in Union and Nonunion Work Groups. *Industrial and Labor Relations Review*, 38(1):pp. 87–97, 1984.

[3] D. Acemoglu and J. Angrist. Consequences of Employment Protection? The Case of the Americans with Disabilities Act. *Journal of Political Economy*, 109(5):pp. 915–957, 2001.

[4] J.G. Altonji, T.E. Elder, and C.R. Taber. Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy*, 113(1):pp. 151–184, 2005.

[5] J.G. Altonji, T.E. Elder, and C.R. Taber. Using Selection on Observed Variables to Assess Bias from Unobservables When Evaluating Swan-Ganz Catheterization. *The American Economic Review*, 98(2):345–350, 2008.

[6] K.H. Anderson and R.V. Burkhauser. The Retirement-Health Nexus: a New Measure of an Old Puzzle. *The Journal of Human Resources*, 20(3):pp. 315–330, 1985.

[7] J.D. Angrist. Estimation of Limited Dependent Variable Models with Dummy Endogenous Regressors. *Journal of Business & Economic Statistics*, 19(1):pp. 2–16, 2001.

[8] J.D. Angrist and J. Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, 2009.

[9] O. Ashenfelter. Estimating the Effect of Training Programs on Earnings. *The Review of Economics and Statistics*, 60(1):pp. 47–57, 1978.

[10] D.H. Autor. Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing. *Journal of Labor Economics*, 21(1):pp. 1–42, 2003.

[11] D.H. Autor, J.J. Donohue III, and S.J. Schwab. The Employment Consequences of Wrongful-Discharge Laws: Large, Small, or None at All? *The American Economic Review*, 94(2):pp. 440–446, 2004.

[12] D.H. Autor, J.J. Donohue III, and S.J. Schwab. The Costs of Wrongful-Discharge Laws. *The Review of Economics and Statistics*, 88(2):pp. 211–231, 2006.

[13] D.H. Autor, W.R. Kerr, and A.D. Kugler. Does Employment Protection Reduce Productivity? Evidence from US States. *The Economic Journal*, 117(521):pp. F189–F217, 2007.

[14] H. Ayele, T. Mulligan, D.V. Espino, I.M. Berges, and K.S. Markides. Higher Church Attendance Predicts Lower Fear of Falling in Older Mexican-Americans. *Aging & Mental Health*, 10(1):pp. 13–18, 2006.

[15] C. Azzi and R. Ehrenberg. Household Allocation of Time and Church Attendance. *The Journal of Political Economy*, 83(1):pp. 27–56, 1975.

[16] M.R. Benjamins. Religion and Functional Health among the Elderly: Is There a Relationship and Is It Constant? *Journal of Aging and Health*, 16(3):pp. 355–374, 2004.

[17] S. Bentolila and G. Bertola. Firing Costs and Labour Demand: How Bad Is Eurosclerosis? *The Review of Economic Studies*, 57(3):pp. 381–402, 1990.

[18] M. Bertrand, E. Duflo, and S. Mullainathan. How Much Should We Trust Differences-in-Differences Estimates? *The Quarterly Journal of Economics*, 119(1):pp. 249–275, 2004.

[19] S. Bonhomme and U. Sauder. Accounting for Unobservables in Comparing Selective and Comprehensive Schooling. CEMFI Working Paper No. 0906, August 2009.

[20] S. Bonhomme and U. Sauder. Recovering Distributions in Difference-in-differences Models: A Comparison of Selective and Comprehensive Schooling. *Review of Economics and Statistics*, 93(2):pp. 479–494, 2011.

[21] J. Bound. Self-Reported Versus Objective Measures of Health in Retirement Models. *Journal of Human Resources*, 21(1):pp. 106–138, 1991.

[22] M. Bray and P. Kwok. Demand for Private Supplementary Tutoring: Conceptual Considerations, and Socio-Economic Patterns in Hong Kong. *Economics of Education Review*, 22(6):pp. 611–620, 2003.

[23] D.C. Briggs. The Eeffect of Admissions Test Preparation: Evidence from NELS: 88. *Chance*, 14(1):pp. 10–18, 2001.

[24] G. Burazeri, A. Goda, and J.D. Kark. Religious Observance and Acute Coronary Syndrome in Predominantly Muslim Albania: a Population-Based Case-Control Study in Tirana. *Annals of Epidemiology*, 18(12):pp. 937–945, 2008.

[25] J.S. Butler, R.V. Burkhauser, J.M. Mitchell, and T.P. Pincus. Measurement Error in Self-Reported Health Variables. *The Review of Economics and Statistics*, 69(4):pp. 644–650, 1987.

[26] P. Cappelli. Examining the Incidence of Downsizing and its Effect on Establishment Performance. In D. Neumark, editor, *On the Job: Is Long Term Employment a Thing of the Past?*, chapter 13, pages pp. 463–513. The Russell Sage Foundation, 2000.

[27] D. Card and A.B. Krueger. School Resources and Student Outcomes: An Overview of the Literature and New Evidence from North and South Carolina. *The Journal of Economic Perspectives*, 10(4):pp. 31–50, 1996.

[28] R. Cheo and E. Quah. Mothers, Maids and Tutors: An Empirical Evaluation of Their Effect on Children's Academic Grades in Singapore. *Education Economics*, 13(3):pp. 269–285, 2005.

[29] D. Cohen-Zada and W. Sander. Religious Participation Versus Shopping: What Makes People Happier? *Journal of Law and Economics*, 54(4):pp. 889–906, 2011.

[30] E.A. Corsentino, N. Collins, N. Sachs-Ericsson, and D.G. Blazer. Religious Attendance Reduces Cognitive Decline among Older Women with High Levels of Depressive Symptoms. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 64(12):pp. 1283–1289, 2009.

[31] H. Dang. The Determinants and Impact of Private Tutoring Classes in Vietnam. *Economics of Education Review*, 26(6):pp. 683–698, 2007.

[32] H. Dang and H.F. Rogers. The Growing Phenomenon of Private Tutoring: Does it Deepen Human Capital, Widen Inequalities, or Waste Resources? *The World Bank Research Observer*, 23(2):pp. 161–200, 2008.

[33] A. Deaton. Aging, Religion, and Health. In David A.W., editor, *Explorations in the Economics of Aging*, chapter 8, pages pp. 237–267. University of Chicago Press, 2011.

[34] J.N. Dertouzos and L.A. Karoly. *Labor Market Responses to Employer Liability.* RAND, Institute for Civil Justice, 1992.

[35] F.X. Diebold, D. Neumark, and D. Polsky. Comment on Kenneth A. Swinnerton and Howard Wial," Is Job Stability Declining in the U.S. Economy?". *Industrial and Labor Relations Review*, 49(2):pp. 348–352, 1996.

[36] P.J. Diggle and P. Hall. A Fourier Approach to Nonparametric Deconvolution of a Density Estimate. *Journal of the Royal Statistical Society. Series B (Methodological)*, 55(2):pp. 523–531, 1993.

[37] J. DiNardo, N.M. Fortin, and T. Lemieux. Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach. *Econometrica*, 64(5):pp. 1001–1044, 1996.

[38] J. DiNardo and D.S. Lee. Program Evaluation and Research Designs. In O. Ashenfelter and D. Card, editors, *Handbook of Labor Economics*, volume 4, chapter 5, pages pp. 463–536. Elsevier, 2011.

[39] C.G. Ellison and J.S. Levin. The Religion-Health Connection: Evidence, Theory, and Future Directions. *Health Education & Behavior*, 25(6):pp. 700–720, 1998.

[40] W.N. Evans and R.M. Schwab. Finishing High School and Starting College: Do Catholic Schools Make a Difference? *The Quarterly Journal of Economics*, 110(4):pp. 941–974, 1995.

[41] H.S. Farber. The Decline of Unionization in the United States: What can be Learned from Recent Experience? *Journal of Labor Economics*, 8(1):pp. S75–S105, 1990.

[42] H.S. Farber. Alternative and Part-Time Employment Arrangements as a Response to Job Loss. *Journal of Labor Economics*, 17(4):pp. S142–S169, 1999.

[43] H.S. Farber. Is the Company Man an Anachronism? Trends in Long-Term Employment in the U.S., 1973-2005. In S.H. Danziger and C.E. Rouse, editors, *The Price of*

Independence: The Economics of Early Adulthood, chapter 3, pages pp. 56–83. The Russell Sage Foundation, 2006.

[44] H.S. Farber. Employment Insecurity: The Decline in Worker-Firm Attachment in the United States. Working Paper No. 530, Industrial Relations Section, Princeton University, 2008.

[45] H.S. Farber. Job Loss and the Decline in Job Security in the United States. Working Paper No. 520, Industrial Relations Section, Princeton University, 2008.

[46] A. Gerber, J. Gruber, and D.M. Hungerman. Does Church Attendance Cause People to Vote? Using Blue Laws' Repeal to Estimate the Effect of Religiosity on Voter Turnout. National Bureau of Economic Research Working Paper No. 14303, 2008.

[47] R.F. Gillum and D.D. Ingram. Frequency of Attendance at Religious Services, Hypertension, and Blood Pressure: the Third National Health and Nutrition Examination Survey. *Psychosomatic Medicine*, 68(3):pp. 382–385, 2006.

[48] M. Goos. The Impact of Shop Closing Hours on Labor and Product Markets. London School of Economics Working Paper, 2005.

[49] B.K. Gorman and A. Sivaganesan. The Role of Social Support and Integration for Understanding Socioeconomic Disparities in Self-Rated Health and Hhypertension. *Social Science & Medicine*, 65(5):pp. 958–975, 2007.

[50] J. Gruber and D.M. Hungerman. The Church versus the Mall: What Happens When Religion Faces Increased Secular Competition? *The Quarterly Journal of Economics*, 123(2):pp. 831–862, 2008.

[51] A. Gurun and D.L. Millimet. Does Private Tutoring Payoff? IZA Discussion Paper No. 3637, 2008.

[52] T. Ha and T. Harpham. Primary Education in Vietnam: Extra Classes and Outcomes. *International Education Journal*, 6(5):pp. 626–634, 2005.

[53] P. Hall. Effect of Bias Estimation on Coverage Accuracy of Bootstrap Confidence Intervals for a Probability Density. *The Annals of Statistics*, 20(2):pp. 675–694, 1992.

[54] R. Hall. The Importance of Lifetime Jobs in the US Economy. *The American Economic Review*, 72(4):pp. 716–724, 1982.

[55] E.A. Hanushek. The Economics of Schooling: Production and Efficiency in Public Schools. *Journal of Economic Literature*, 24(3):pp. 1141–1177, 1986.

[56] E.A. Hanushek. Assessing the Effects of School Resources on Student Performance: An Update. *Educational Evaluation and Policy Analysis*, 19(2):pp. 141–164, 1997.

[57] Eric A Hanushek. The Failure of Input-Based Schooling Policies. *The Economic Journal*, 113(485):pp. F64–F98, 2003.

[58] James J Heckman. Dummy Endogenous Variables in a Simultaneous Equation System. *Econometrica*, 46(4):pp. 931–959, 1978.

[59] H. Hopenhayn and R. Rogerson. Job Turnover and Policy Evaluation: A General Equilibrium Analysis. *Journal of Political Economy*, 101(5):pp. 915–938, 1993.

[60] B.D. Horne, H.T. May, J.L. Anderson, A.G. Kfoury, B.M. Bailey, B.S. McClure, D.G. Renlund, D.L. Lappé, J.F. Carlquist, P.W. Fisher, et al. Usefulness of Routine Periodic Fasting to Lower Risk of Coronary Artery Disease in Patients Undergoing Coronary Angiography. *The American Journal of Cardiology*, 102(7):pp. 814–819, 2008.

[61] J.L. Horowitz. The Bootstrap. In J.J. Heckman and E.E. Leamer, editors, *Handbook of Econometrics*, volume 5, chapter 52, pages pp. 3159–3228. Elsevier, 2001.

[62] R.A. Hummer, R.G. Rogers, C.B. Nam, and C.G. Ellison. Religious Involvement and US Adult Mortality. *Demography*, 36(2):pp. 273–285, 1999.

[63] L.S. Jacobson, R.J. LaLonde, and D.G. Sullivan. Earnings Losses of Displaced Workers. *American Economic Review*, 83(4):pp. 685–709, 1993.

[64] D.A. Jaeger and A.H. Stevens. Is Job Stability in the United States Falling? Reconciling Trends in the Current Population Survey and Panel Study of Income Dynamics. *Journal of Labor Economics*, 17(S4):pp. S1–S28, 1999.

[65] D. Kahneman and A.B. Krueger. Developments in the Measurement of Subjective Well-Being. *The Journal of Economic Perspectives*, 20(1):pp. 3–24, 2006.

[66] C. Kang. Does Money Matter? The Effect of Private Educational Expenditures on Academic Performance. National University of Singapore, Department of Economics Working Paper No. 0704, 2007.

[67] A.Y. Kinney, L.E. Bloor, W.N. Dudley, R.C. Millikan, E. Marshall, C. Martin, and R.S. Sandler. Roles of Religious Involvement and Social Support in the Risk of Colon Cancer among Blacks and Whites. *American Journal of Epidemiology*, 158(11):pp. 1097–1107, 2003.

[68] H. Koenig and V.B. Carson. *Handbook of Religion and Health*. Oxford University Press, 2012.

[69] H.G. Koenig, L.K. George, and P. Titus. Religion, Spirituality, and Health in Medically Ill Hospitalized Older Patients. *Journal of the American Geriatrics Society*, 52(4):pp. 554–562, 2004.

[70] N. Krause and C.G. Ellison. Forgiveness by God, Forgiveness of Others, and Psychological Well-Being in Late Life. *Journal for the Scientific Study of Religion*, 42(1):pp. 77–93, 2003.

[71] A.B. Krueger. Economic Considerations and Class Size. *The Economic Journal*, 113(485):pp. F34–F63, 2003.

[72] A.B. Krueger and D.M. Whitmore. The Effect of Attending a Small Class in the Early Grades on College-Test Taking and Middle School Test Results: Evidence from Project STAR. *The Economic Journal*, 111(468):pp. 1–28, 2001.

[73] A.D. Kugler and G. Saint-Paul. How Do Firing Costs Affect Worker Flows in a World with Adverse Selection? *Journal of Labor Economics*, 22(3):pp. 553–584, 2004.

[74] J. Lee, Y. Kim, and C. Yoon. The Effects of Pre-Class Tutoring on Student Achievement: Challenges and Implications for Public Education in Korea. *KEDI Journal of Educational Policy*, 1(1):pp. 25–42, 2004.

[75] A.R. Levenson. Long-Run Trends in Part-Time and Temporary Employment: Toward an Understanding. In D. Neumark, editor, *On the Job: Is Long Term Employment a Thing of the Past?*, chapter 10, pages pp. 335–397. The Russell Sage Foundation, 2000.

[76] S.K. Lutgendorf, D. Russell, P. Ullrich, T.B. Harris, and R. Wallace. Religious Participation, Interleukin-6, and Mortality in Older Adults. *Health Psychology*, 23(5):pp. 465–475, 2004.

[77] C.F. Manski and J.V. Pepper. Monotone Instrumental Variables: With an Application to the Returns to Schooling. *Econometrica*, 68(4):pp. 997–1010, 2000.

[78] P.J. McEwan. The Potential Impact of Vouchers. *Peabody Journal of Education*, 79(3):pp. 57–80, 2004.

[79] T.J. Miles. Common Law Exceptions to Employment at Will and US Labor Markets. *Journal of Law, Economics, and Organization*, 16(1):pp. 74–101, 2000.

[80] J.C. Moorhouse. Is Tullock Correct about Sunday Closing Laws? *Public Choice*, 42(2):pp. 197–203, 1984.

[81] P. Muennig, G. Johnson, J. Kim, T. Smith, and Z. Rosen. The General Social Survey-National Death Index: an Innovative New Dataset for the Social Sciences. *BMC Research Notes*, 4(1):pp. 385–391, 2011.

[82] C.J. Muhl. The Employment-at-Will Doctrine: Three Major Exceptions. *Monthly Labor Review*, 124:pp. 3–11, 2001.

[83] A. Murray-Swank, R. Goldberg, F. Dickerson, D. Medoff, K. Wohlheiter, and L. Dixon. Correlates of Religious Service Attendance and Contact with Religious Leaders among Persons with Co-Occurring Serious Mental Illness and Ttype 2 Diabetes. *The Journal of Nervous and Mental Disease*, 195(5):pp. 382–388, 2007.

[84] D.A. Neal. The Effects of Catholic Secondary Schooling on Educational Achievement. *Journal of Labor Economics*, 15(1):pp. 98–123, 1997.

[85] D. Neumark. *On the Job: Is Long-Term Employment a Thing of the Past?* Russell Sage Foundation Publications, 2000.

[86] T. Obisesan, I. Livingston, H.D. Trulear, and F. Gillum. Frequency of Attendance at Religious Services, Cardiovascular Disease, Metabolic Risk Factors and Dietary Intake in Americans: an Age-Stratified Exploratory Analysis. *The International Journal of Psychiatry in Medicine*, 36(4):pp. 435–448, 2006.

[87] D. Oman, J.H. Kurata, W.J. Strawbridge, and R.D. Cohen. Religious Attendance and Cause of Death over 31 Years. *The International Journal of Psychiatry in Medicine*, 32(1):pp. 69–89, 2002.

[88] H. Ono. Does "Examination Hell" Pay off? A Cost–Benefit Analysis of?Ronin?and College Education in Japan. *Economics of Education Review*, 26(3):pp. 271–284, 2007.

[89] C.L. Park, B. Moehl, J.R. Fenster, DP Suresh, and D. Bliss. Religiousness and Treatment Adherence in Congestive Heart Failure Patients. *Journal of Religion, Spirituality & Aging*, 20(4):pp. 249–266, 2008.

[90] H.H. Perritt. *Employee Dismissal: Law and Practice*. Aspen Publishers, 1997.

[91] C.A. Reyes-Ortiz, I.M. Berges, M.A. Raji, H.G. Koenig, Y.F. Kuo, and K.S. Markides. Church Attendance Mediates the Association between Depressive Symptoms and Cognitive Functioning among Older Mexican Americans. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 63(5):pp. 480–486, 2008.

[92] S.J. Rose. Declining Job Security and the Professionalization of Opportunity. National Commission for Employment Policy, Research Report 95-04, 1995.

[93] D.B. Rubin. Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies. *Journal of Educational Psychology*, 66(5):pp. 688–701, 1974.

[94] D. Ryu and C. Kang. Do Private Tutoring Expenditures Raise Academic Performance? Evidence from Middle School Students in South korea. *Asian Economic Journal*, 27(1):pp. 59–83, 2013.

[95] E. Schnall, S. Wassertheil-Smoller, C. Swencionis, V. Zemon, L. Tinker, O.S. Mary Jo, L. Van Horn, and M. Goodwin. The Relationship between Religion and Cardiovascular

Outcomes and All-Cause Mortality in the Women's Health Initiative Observational Study. *Psychology and Health*, 25(2):pp. 249–263, 2010.

[96] S.E. Sephton, C. Koopman, M. Schaal, C. Thoresen, and D. Spiegel. Spiritual Expression and Immune Status in Women with Metastatic Breast Cancer: an Eexploratory Study. *The Breast Journal*, 7(5):pp. 345–353, 2002.

[97] S. Shimizutani and I. Yokoyama. Has Japan's Long-Term Employment Practice Survived? Developments since the 1990s. *Industrial and Labor Relations Review*, 62(3):pp. 313–326, 2009.

[98] B.W. Silverman. *Density Estimation for Statistics and Data Analysis*, volume 26. Chapman & Hall, 1986.

[99] K. Sparks, C. Cooper, Y. Fried, and A. Shirom. The Effects of Hours of Work on Health: A Meta-Analytic Review. *Journal of Occupational and Organizational Psychology*, 70(4):pp. 391–408, 1997.

[100] D.L. Stevenson and D.P. Baker. Shadow Education and Allocation in Formal Schooling: Transition to University in Japan. *American Journal of Sociology*, (6):pp. 1639–1657, 1992.

[101] J. Stewart. Recent Trends in Job Stability and Job Security: Evidence from the March CPS. Bureau of Labor Statistics Working Paper No. 356, 2002.

[102] K.A. Swinnerton and H. Wial. Is Job Stability Declining in the U.S. Economy? *Industrial and Labor Relations Review*, 48(2):pp. 293–304, 1995.

[103] A. Tansel and F. Bircan. Effect of Private Tutoring on University Entrance Examination Performance in Turkey. IZA Discussion Paper, No. 1609, 2005.

[104] A. Tansel and F. Bircan. Demand for Education in Turkey: A Tobit Analysis of Private Tutoring Expenditures. *Economics of Education Review*, 25(3):pp. 303–313, 2006.

[105] G. Tullock. The Transitional Gains Trap. *The Bell Journal of Economics*, 6(2):pp. 671–678, 1975.

[106] R. Valletta. Declining Job Security. *Journal of Labor Economics*, 17(4):pp. S170–S197, 1999.

[107] P.H. Van Ness and S.V. Kasl. Religion and Cognitive Dysfunction in an Elderly Cohort. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 58(1):pp. S21–S29, 2003.

[108] D.J. Walsh and J.L. Schwarz. State Common Law Wrongful Discharge Doctrines: Update, Refinement, and Rationales. *American Business Law Journal*, 33(4):pp. 645–690, 1996.

[109] D.M. Yeager, D.A. Glei, M. Au, H.S. Lin, R.P. Sloan, and M. Weinstein. Religious Involvement and Health Outcomes among Older Persons in Taiwan. *Social Science & Medicine*, 63(8):pp. 2228–2241, 2006.

[110] GJ. Yoo and CH. Kang. The Effect of Protection of Temporary Workers on Employment Levels: Evidence from the 2007 Reform of South Korea. *Industrial and Labor Relations Review*, 65(3):pp. 578–606, 2012.