Licensing and Occupational Sorting in the Market for Teachers

Matthew Wiswall

New York University
Department of Economics
mwiswall@nyu.edu

Abstract

This paper develops and estimates a dynamic occupational choice model with an endogenous licensing decision to examine the current debate about licensing policy for American public school teachers. The estimated model is used to simulate the effects of changes in teacher licensing policy. We find that higher licensing requirements reduce total teacher labor supply and average teacher quality, while increasing the average length of teaching careers. These findings suggest that policies that implement more stringent teacher licensing requirements, such as the *No Child Left Behind* legislation, may be counter-productive and not lead to an increase in teacher quality.

1 Introduction

Entry requirements in the form of occupational licenses are a prevalent feature of many occupations. The primary justification for an occupational license is to solve an information asymmetry where the quality of a worker cannot be known with certainty by the employer. Licenses solve this problem either by screening applicants through testing or by imposing a floor on occupational specific human capital through minimum training requirements.[1] Since at least Adam Smith, occupational licenses have been criticized as an unnecessary barrier to occupational entry that create rents for the already licensed incumbent workers.[2] The welfare analysis of occupational licensing centers on a trade-off between the cost to employers and consumers of decreased occupational labor supply versus the potential benefits of restricting occupational access to higher quality workers, especially in occupations, such as medicine, where the costs to consumers of low quality workers is high. Raising licensing requirements then is expected to increase labor costs, but with the benefit of improving the quality of the licensed workforce.

This paper develops a dynamic occupational choice model with an endogenous licensing decision that shows that on at least one dimension, higher licensing requirements can lower, rather than raise, the quality of the licensed workers. We demonstrate this result using the example of the American teacher labor market in which public school teachers are required to obtain a license to teach.[3] Using detailed occupation, licensing, and earnings information

---

[1] These two policies are fundamentally similar. Testing not only screens out low quality workers, but testing also forces individuals to learn something in order to pass the test, e.g. most lawyers need to attend law school to pass the bar exam in the US. Minimum training requirements can act as a screen since completing the required training can only be accomplished by some, e.g. low quality lawyers fail to graduate from law school.

[2] For a recent review of this literature, see Kleiner (2000, 2006). Some examples of this literature include Leland (1979), Dorsey (1980), and Shapiro (1986). Kleiner (2006) cites several historical works in economics, including Smith's *Wealth of Nations*, which criticized licensing in these terms.

[3] Throughout the paper, we use the term "license" to refer to a policy that in the US education litera-

for a nationally representative sample of college graduates, we estimate the parameters of a structural model for this market and find that higher licensing requirements reduces the incentives for high quality workers to enter teaching. This seemingly counter-intuitive result is due to the fact that current teacher licensing policy is largely a weak quality screen. Instead, the distinguishing feature of teacher licensing is a requirement for license holders to make a substantial investment in teacher specific training.

In our model, as in previous models of occupational licensing, higher licensing requirements raises the cost of entering teaching. The major distinction between our model and previous models (e.g. Shapiro 1986) is that in our model workers do not choose whether to provide services of various levels of quality to a single sector, but instead choose which sector to enter. Different licensing requirements across sectors changes the relative utility workers receive from sectors and affects the occupational sorting of workers across sectors. A key part of our model is that the costs of meeting licensing requirements varies across workers who are assumed to be heterogeneous in their endowments of general skills. In the dynamic occupational choice model for teachers developed here, the cost of acquiring a teaching license and making the required sunk investment in teaching specific training is larger for individuals with higher levels of general skills because they receive larger non-teaching wage offers and have a higher probability of exiting teaching. If general skills are a component of quality teaching, then more stringent and costly licensing requirements disproportionately reduces the attractiveness of teaching to higher quality potential teachers. The implication of this model is that not only do higher licensing requirements reduce the overall supply of teachers, licensing also affects the types of individuals who enter teaching.

---

ture is also referred to as teacher "certification." Teacher licenses in the US combine some elements of an occupational license and a traditional occupational certificate. Teacher licensing policy in the US is more accurately described as a partially enforced licensing policy.

In the United States, teacher licensing requirements are set at the state level and apply to the approximately 3 million teachers who teach in public schools in the United States. For most states, there are two main features of teacher licenses: a test of basic skills and a minimum training requirement. The testing component is generally a weak screen, as the minimum test scores are so low that typically more than 90 percent of all prospective teachers pass these tests (U.S. Dept. of Educ. 2002). Although the training requirements vary somewhat across states, on average these requirements are substantial and amount to about one year of pedagogy course work and generally unpaid teaching apprenticeships. The time costs of meeting these requirements constitutes the largest cost of licensing for teachers.

Two features of the teacher labor market make it a particularly attractive occupation to study occupational licensing. First, unlike occupations such as law or medicine, which also have substantial licensing requirements, teaching is one of the lowest paying occupations for college graduates. For the legal and medical professions, there is little concern about attracting quality entrants, as salaries in these occupations are high, and individuals are generally well rewarded for effort and performance. In contrast, compensation in teaching is relatively low, and little, if any, of teacher compensation is tied to performance. Given the low effective compensation for completing the licensing requirements in teaching, the costs of completing these requirements would be expected to have a much larger impact on the choice to enter teaching than in higher paying professions.

A second attractive reason to study teacher licenses is that the licensing regime in teaching is relatively porous. With “alternative” licensing programs, such as the *Teach for America* program, and a number of school districts granting licensing “waivers” to non-licensed teachers, about 15 percent of new teachers teach without completing the regular licensing requirements.[4] This partially enforced licensing regime is an attractive feature for empirical research as it allows an explicit comparison of non-licensed and licensed teachers. With perfectly enforced licenses, we would have little information to understand the behavior of non-licensed individuals if licensing restrictions were removed.

The influx of non-licensed teachers into the labor force has alarmed some education policy makers and set off a new debate about the value of teacher licenses. Previous research on teacher licensing compares student level outcomes (e.g. scores on standardized tests taken by students) to examine whether students who are taught by licensed teachers have better average outcomes than students taught by non-licensed teachers.[5] Some research finds that licensed teachers are associated with higher achieving students (Darling-Hammond 2000; Darling-Hammond, Chung, and Frelow 2002; and Darling-Hammond, Berry, and Thoreson 2001). Other research finds that licenses makes no difference (Goldhaber and Brewer 2000, 2001; and Walsh 2001). It is difficult to interpret these findings because teacher licensing is not a randomly assigned teacher characteristic. As we demonstrate with the occupational choice model developed here, individuals self-select into teaching and licensing based on heterogeneous skills and tastes. Depending on the sorting induced by licenses, licensed teachers may have a different distribution of general skills than non-licensed teachers. Whether the course work required by licenses has any practical value to teacher performance cannot be ascertained from student level outcomes given that individuals self-select into licensing along a dimension that also potentially affects teacher productivity.

The labor market for teachers has been one of the most intensely studied labor markets in the United States. A number of studies have documented a negative relationship between

---

[4] Authors' calculation from the Baccalaureate and Beyond (B&B) data for the 1992/93 college graduating class discussed below.

[5] This is commonly called an "education production function" approach in which the output is student learning or human capital, measured by student test scores, and inputs include characteristics of teachers.

measured academic or cognitive ability and the decision to become a teacher.[6] The average ability of teachers has only declined over time as wage growth in other professions has drawn some of the best teachers out of the profession (Bacolod 2003; Corcoran, Evans, and Schwab 2002; Murnane et al 1991). This is particularly important since recent evidence indicates that there are large quality differences among teachers, although the sources of the quality differences are not well understood (Hanushek, Kain, and Rivkin 2005; Rockoff 2004; Aaronson, Barrow, and Sander 2003). Recent studies to understand teacher entry and exit decisions include Stinebrickner (2001a, 2001b,2002), which estimates a model of teacher labor supply using data on a cohort of already licensed teachers who entered teaching in the 1970s. The major distinction between this work and previous studies is that this paper explicitly incorporates an endogenous licensing decision into a dynamic occupational choice model. This allows us to directly examine how changes in licensing policy affect occupational choices.

Previous empirical work on licensing exploits variation in licensing policy across geographic areas to identify the effects of licensing requirements on labor supply and wages. Kleiner and Kudrle (2000) use variation across US states in licensing requirements for dentists to estimate the effects of licensing requirements on the cost of dental services, quality of dental care, and the supply of dentists. Kugler and Sauer (2005) use a discontinuity in Israeli physician licensing requirements in an instrumental variables framework to examine the selection of individuals into licensing and the rents associated with licensing. Angrist and Guryan (2004) use variation in teacher testing requirements across US states and school

---

[6] See, for example, Manski (1987) using college freshman admission Scholastic Aptitude Test (SAT) scores and high school rank in the National Longitudinal Study of the High School Class of 1972; Ehrenberg and Brewer (1995) using the selectivity rankings of colleges and universities teachers attended in the High School and Beyond data set; and Ballou and Podgursky (1995) using SAT scores of college bound seniors. Data presented below from the Baccalaureate and Beyond survey also confirms this finding.

district employers to examine whether teacher testing requirements alone increase the costs of entering teaching if there are costs to preparing for the test. In their regression analysis, Angrist and Guryan find that school districts that required a test did not hire better quality teachers. This conclusion is not surprising since the minimum passing scores are so low, most prospective would not need to spend much time preparing for these tests. There are potentially far larger costs to completing the course work required by teacher licenses.

This paper takes a different empirical approach to examining licensing. Rather than use variation in teacher licensing laws across states, we leave the exact licensing policy unspecified and estimate the costs of licensing through the behavior of agents that are reacting to the policy. While US states do differ in their licensing policies, as discussed in more detail below, the exact policies in each state are difficult to summarize parsimoniously. Licensing policies have many dimensions corresponding to different rules for different levels of schooling, subjects, and student populations. It is difficult to model the menu of licensing regulations facing agents.[7] In addition, using state level variation in licensing policy as a "natural experiment" is problematic as states differ in many unobservable dimensions, which may also affect occupational choices.[8] We sidestep these issues by parameterizing and estimating a general cost to licensing without specifying the actual licensing requirements that create this cost. This is analogous to estimating other behavioral parameters, such as the cost of raising children or attending school, where the actual components of these costs are not specified. To the extent that agents face different licensing requirements, the licensing costs estimated in this paper using a nationally representative data sample can be interpreted as a

---

[7] Complicating matters further, several states have compacts with other states allowing individuals licensed in one state to teach in other states, with some exceptions and requirements for additional training.

[8] As some some evidence of this, the states with the lowest licensing requirements are states such as Alaska and Montana, which have relatively weak labor markets for college graduates. States with relatively higher licensing requirements, include California and New York, which have much stronger labor markets for college graduates.

national cost to licensing across all the various state requirements.[9] A benefit to the approach used here is that it provides a framework to examine the effects of policy changes for policies for which we have no experimental data, such as the *No Child Left Behind* legislation.

Using a detailed panel data set of college graduates, we estimate that the cost to obtaining a teaching license in the first year after college graduation is $29,030 in 2003 dollars, a cost which is slightly more than the average annual starting salary for college graduates. In a series of counterfactual policy experiments based on the estimated model, we find that eliminating licensing costs and allowing free entry into teaching would increase total teacher labor supply by about 3.4 percent and raise average teacher quality, as measured by foregone non-teaching wages, by about 2.2 percent. We also find substantial effects of licensing on occupational attachment. Eliminating licensing requirements would increase the number of low teaching attachment individuals entering teaching and reduce the mean career length of teachers by 16.9 percent. Increases in licensing costs from the current policy are found to have opposite effects. Higher licensing costs reduce teacher labor supply and reduce average teacher quality while increasing the average length of teaching careers. In general we find non-linear effects of licensing, as small increases in licensing costs from free entry (no licensing requirements) have much larger effects than increases in licensing costs from the current policy.

We also use the estimated model to evaluate a policy of increasing licensing enforcement. When we examine the counterfactual policy of fully enforcing licensing requirements and not allowing non-licensed individuals to teach, we find very small changes in teacher labor supply as most previously non-licensed teachers choose to obtain a license. However, mean career

---

[9] To the extent that individuals can move to different states to complete a different set of licensing requirements, the estimated licensing costs are an amalgam of the licensing requirements for various states and the costs for individuals to move to other states with lower requirements.

attachment falls by 16.2 percent as lower attachment teachers form an increasing share of the teacher labor force. Average teacher quality, as measured by foregone wages, is higher under the fully enforced policy than under the current policy, but still lower than under free entry.

The remainder of the paper is organized as follows. First, we provide some background on licensing in the US teacher labor market. Next, we develop the licensing and occupational choice model. We then discuss how this model can be used to evaluate changes in teacher licensing policy. The model solution, econometric estimator, data, and descriptive evidence are then presented. The paper concludes by using the estimated parameters in a series of policy simulations of counterfactual changes in teacher licensing policy.

2 Background on Teacher Licensing

Teacher licensing requirements in the United States are set at the state level and apply to teachers who teach at public schools.[10] Private schools and some specially regulated charter schools are exempt from licensing requirements.[11] State licenses are sometimes valid in other states if the states have a teacher licensing compact. In addition, teachers who move to other states can often take additional courses to meet their new state's requirement.

Generally, states have separate licensing requirements for elementary and secondary levels and, for some states, also separate requirements for course subjects (e.g. science, social studies). The licensing policy specifies the degree that the individual must hold (typically a bachelor degree but sometimes also a masters degree is required), whether a subject area

---

[10] More information on teacher licensing regulations can be found in Ballou and Podgursky (1999).

[11] Charter schools are publicly funded but operate free from many of the regulations governing public schools. Private schools hire many licensed teachers. In the Baccalaureate and Beyond data studied here about half of private school teachers have some kind of regular certification.

major or minor is required (e.g. mathematics teachers must have a math major), and the composition of the teacher preparation program. The teacher preparation program requirements typically include a requirement to complete a number of courses in pedagogy subjects (course titles include “Methods and Strategies of Teaching,” “Social Foundations and History of American Education,” and “Philosophy of Education”) and some period as a practice or student teacher.

The extensive teacher preparation requirements are the main cost to obtaining a teacher license. Figure 1 displays the number of teacher preparation courses required in 2001 by states for elementary and secondary licenses. The average number of courses required across all states is 12.2 courses for elementary licenses and 11.3 for secondary license. Consistent with these requirements, the college graduate with a teacher’s license have on average completed about 25 credits in education course work, 22 percent of total college credits. Non-licensed teachers by contrast have on average only a few credits in education courses.[12]

Given the extensive requirements for regular licenses and localized shortages of licensed teachers, many states now offer some form of “alternative” licensing. Although these licenses are generally counted among the regular licenses, alternative licenses require substantially less pedagogy course work. Some of these programs are intended for older individuals who want to transition from other occupations into teaching, while other programs are intended to attract recent college graduates. Alternative licensing programs were introduced in some states during the early 1980s as teacher shortages, especially in minority and poorer urban areas, forced school districts and states to re-examine their licensing requirements. By 2002 alternative licensing programs had been adopted by at least 45 states and the District of Columbia, an increase from 8 states in 1983 (Feistrizer 2002). For the nationally represen-

---

[12] State requirements from the federal Department of Education’s *Title 2* data collection. Courses taken by licensed and non-licensed teachers are from the B&B data sample described below.

tative sample of college graduates used here, about 15 percent of new teachers each year are licensed through alternative programs or teach without a license. Like regular licenses, alternative licensing requirements vary widely. Some states and school districts require some course work to be taken concurrently or just prior to teaching with an alternative license, while other states and districts mandate little additional training (Darling-Hammond 2000, U.S. Dept. of Educ. 2002).

The recent federal *No Child Left Behind* legislation sets new standards for all public schools to hire "highly qualified" teachers, which the legislation defines as teachers who have met their state's licensing requirements and have a subject area major in the subject they teach (e.g. mathematics teachers must have a bachelor's degree in mathematics). The legislation continues to allow individual states to set their own specific teacher licensing requirements, but requires all teachers to meet these requirements. While it is unclear to what extent the legislation will be enforced, the legislation as written contains both an increase in licensing requirements and more stringent enforcement of existing licensing requirements.[13]

3 Model

This section lays out the general model that will be estimated and used to evaluate teacher licensing policy in the United States. There are two sectors in the economy. The K-12 (kindergarten through high school) education sector consists of school district employers that employ teachers to teach students and produce student human capital. The alternative sector employs the remaining workers. The sectors are distinguished by their exogenously

---

[13] Information on the 2001 *No Child Left Behind* legislation can be found on the Department of Education's website: www.ed.gov.

determined wage offers, job offers, and licensing requirements. The education sector has a licensing policy that restricts job offers to non-licensed individuals according to a policy described below. The education sector includes both private and public schools. The presence of private school employers and public schools that hire non-licensed teachers is incorporated into the arrival rate of teaching job offers for non-licensed teachers. The alternative sector has no licensing restrictions and offers a job to all workers.

Individuals are heterogeneous in their endowments of skills and tastes in the two sectors. In each period, agents make occupation and licensing decisions in response to the wage offers, job offers, and licensing policy in the two sectors. The distribution of skills and accumulated work experience human capital among the individuals who choose to teach determines the level of student human capital production in the education sector. Teacher licensing policy affects the costs and benefits of teaching, the distribution of teaching relevant human capital in the teacher labor force, and therefore the production of education output.

3.1 Timing

Agents, indexed i, make decisions in a discrete number of periods from the initial period $t = 1$ until an exogenous retirement date in period $t = T$. The model applies to college graduates, periods are years, and the initial period ($t = 1$) is the first year following college graduation when college graduates are aged about 22-24. At the beginning of each period, agents first choose whether to obtain a teacher license. After the licensing decision, agents receive job offers from one or both of the sectors. After receiving job offers, agents choose from among three mutually exclusive activities: work as a teacher in the education sector (sector 1) if a teaching job is offered, work in the alternative sector (sector 2), or stay out of the labor market. Output in the education sector is determined by the distribution of teaching relevant human capital for those individuals who choose to teach in this period.

3.2 Human Capital Heterogeneity

Individual productivity in the education and alternative sectors is a function of the individual's own human capital. There are five types of human capital: informal teaching specific skills (a_i), general skills (g_i), formal teacher training (z_i), accumulated teaching experience by period t (x_{1it}), and accumulated alternative sector experience (x_{2it}). The human capital level of individual i by period t is given by the vector $H_{it} = [z_i, a_i, g_i, x_{1it}, x_{2it}]$.

Prior to the first period of the model (i.e. prior to college graduation), each agent is endowed with a level of informal teaching skills (a_i) and general skills (g_i). The level of formal teacher training (z_i) and experience in each of the sectors (x_{1it} and x_{2it}) are the result of endogenous licensing and occupation choices. Note that we distinguish between human capital from formal training required by licensing (e.g. knowledge of various methods of teaching) from informal teaching specific skills (e.g. ability to communicate teaching material to adolescents). General skills are the skills that make an individual productive both in teaching and in the alternative sector, such as intelligence or general communication or organizational skills.

3.3 Teacher Licenses

Teacher licenses are assumed to have the following features: 1) Prior to the first period ($t = 1$), no agents have a license. 2) Licenses are permanent; once an agent has obtained a license, the agent can never loose the license. 3) Agents with a license are always offered a job as a teacher. 4) The probability non-licensed agents receive a teaching job offer is $\gamma \in [0, 1]$. All licensed and non-licensed agents are offered a job in the alternative sector.

Licenses require an individual who obtains the license to acquire a level $z_i = z^0 \geq 0$ of formal teacher training. The one-time cost to formal training is given by the function $\kappa(z)$, where $\kappa(0) = 0$ and $\kappa'(z) > 0$ for all z. The cost of completing the licensing requirements, $\kappa(z^0)$, primarily represents the opportunity cost of time spent completing the formal teacher training, although $\kappa(z^0)$ can include any other fees or psychic costs. This cost must be paid prior to obtaining the license.[14]

The licensing policy is fully characterized by z^0 and γ. $z^0 = 0$ or $\gamma = 1$ corresponds to a policy of *free entry* into teaching. $\gamma = 1$ corresponds to a *fully enforced* licensing policy in which only license holders are allowed to teach. $\gamma < 1$ and $z^0 > 0$ corresponds to a *partially enforced* licensing policy in which non-licensed individuals are offered teaching jobs but at a lower rate than licensed holders.

3.4 Education Production Function

Each teacher teaches a group of students and produces education output. Following an extensive literature on education production functions (Hanushek, Kain, and Rivkin 2005; Rockoff 2004; Aaronson, Barrow, and Sander 2003), we define the per period contribution of the teacher to student learning as *teacher value added.* In the education production function literature, the output of the educational process is student learning or human capital production. The inputs into education production are various aspects of the educational system that affect student learning, such as teachers and school resources, and inputs from parents, peers, and the student's own endowment of abilities.

Teacher value added is modeled symmetrically for all teachers. We ignore issues that some

---

[14] For most licensed teachers, teacher licensing training occurs during college as part of an individual's college course work, and therefore occurs prior to the first period in the model. Although not modeled for simplicity, non-licensed teachers may also have positive levels of formal teacher training, some of which may be provided and paid for by school employers.

teachers teach more students than others and the possible presence of complementarities between teacher inputs and other education inputs. Because licensing policy can affect who in the pool of college graduates works as a teacher, our measure of teacher value added is defined for all individuals, whether they actually teach or not. Thus our concept of teacher value added is *potential* teacher value added.

Call potential teacher value added for individual i in period t, y_{1it}. Potential teacher value added is modeled as a function of individual human capital:

$$y_{1it} = h_1(H_{it}) = h_1(z_i, a_i, g_i, x_{1it}, x_{2it}). \tag{1}$$

We assume that education production is increasing in each of the component types of human capital.

3.5 Wage Offers

In a competitive equilibrium model, wage offers would reflect the marginal productivity of the human capital bundle possessed by each heterogeneous worker (e.g. Heckman and Sedlacek 1985). In this type of setup, wages for the two sectors indexed $k = 1, 2$ would be modeled as

$$w_k(H_{it}) = R_k' H_{it} \text{ and } R_k = P_k \frac{\partial h_k(H_{it})}{\partial H_{it}}, \tag{2}$$

where R_k is the vector of prices for each type of human capital, and P_k are the output prices for production in sector k. The wage setting mechanism in (2) is difficult to implement for the education sector for several reasons: 1) output prices are not easily defined in the education sector given that public school budgets are set by a political process, 2) competitive pressures on public schools in the United States are generally weak and education sector wage offers may not reflect the marginal productivity of a teacher's bundle of human capital, and 3) education sector employers may have imperfect knowledge about the education production function.

As an alternative to a more natural wage setting mechanism in (2), we adopt a partial equilibrium formulation and approximate the education sector wage setting mechanism by assuming that education sector wage offers (w_{1it}) are a Mincer type function of informal teaching skills (a_i), general skills (g_i), and accumulated teaching experience (x_{1it}):

$$w_{1it} = w_1(H_{it}) = \exp(\pi_1(a_i, g_i) + \beta_{11}x_{1it} + \beta_{12}x_{1it}^2 + \epsilon_{1it}), \tag{3}$$

where $\pi_1(a_i, g_i)$ is a scalar valued function indicating the mapping of informal teaching and general skills to education sector wage offers. β_{11} and β_{12} are wage offer parameters, and ϵ_{1it} is a mean zero i.i.d. education sector wage offer shock. Although we assume that school employers observe an individual's level of informal teaching skills and general skills, how these skills map into teacher wage offers we leave largely undefined through the $\pi_1(a_i, g_i)$ function.[15]

To maintain consistency with the education sector, we also model wage offers in the alternative sector as a Mincer type function of general skills (g_i) and alternative sector work experience (x_{2it}):

$$w_{2it} = w_2(H_{it}) = \exp(\pi_2(g_i) + \beta_{21}x_{2it} + \beta_{22}x_{2it}^2 + \epsilon_{2it}), \tag{4}$$

---

[15] The empirical section defines how we estimate the distribution of human capital endowments in the economy through this specification. Wage offers in the education sector are assumed not to be a function of formal training mandated by licensing. From a comparison of salaries between licensed and non-licensed teachers, there is little evidence that teacher salaries depend on licensing and formal training.

where $w_2(H_{it})$ is the wage setting policy in the alternative sector. $\pi_2(g_i)$ is the mapping between general skills and alternative sector wage offers. β_{21} and β_{22} are wage offer parameters, and ϵ_{2it} is a mean zero i.i.d. alternative sector specific wage offer shock.

3.6 Utility Functions

The flow utility in each period t is the sum of current consumption (C_{it}), current non-pecuniary benefits (B_{it}), and the one time licensing cost if a teacher license is obtained in this period:

$$U(C_{it}, B_{it}, L_{it}) = C_{it} + B_{it} - L_{it}\kappa(z^0).$$

$L_{it} = 1$ is an indicator variable for obtaining a license in period t, $L_{it} = 0$ otherwise. We normalize utility from the non-working, out of the labor force, alternative to 0.

With no saving or borrowing in the model, current consumption is current period labor market earnings for whichever sector the agent works in this period, if any:

$$C_{it} = d_{1it}w_{1it} + d_{2it}w_{2it}.$$

d_{1it} is a dummy variable equal to 1 if the agent teaches, and 0 otherwise. d_{2it} is defined similarly for the alternative sector. Choosing to stay out of the labor market corresponds to $d_{1it} = 0$ and $d_{2it} = 0$.

Non-pecuniary benefits embody the tastes for working in each sector:

$$B_{it} = d_{1it}b_{1it} + d_{2it}b_{2it}.$$

$b_{1it} \in R$ and $b_{2it} \in R$ indicate the relative tastes agents have for working in the two sectors.

b_{1it} and b_{2it} are modeled to have fixed and time varying individual components:

$$b_{1it} = \tau_{1i} + \eta_{1it} \text{ and } b_{2it} = \tau_{2i} + \eta_{2it}.$$

$\tau_{1i} \in R$ and $\tau_{2i} \in R$ are individual, time invariant components that reflect individual tastes for the two sectors. η_{1it} and η_{2it} are time varying shocks to tastes. Note that B_{it} can be negative to reflect the dis-utility from working relative to remaining out of the labor force.

3.7 Individual Decision Problem

Individuals choose a licensing and occupational choice sequence to maximize discounted expected utility:

$$\max_{\{L_{it},d_{1it},d_{2it}\}_{t=1}^{T}} \sum_{t=1}^{R} \delta^{t-1} E_{\nu,J}[U(C_{it}, B_{it}, L_{it})] \qquad (5)$$
$$\text{s.t. } d_{1it} \in J_{it}$$

where $J_{it} = 1$ is an indicator that a teaching job offer is made, $J_{it} = 0$ otherwise. $\delta \in [0, 1]$ is the discount rate.

Expectations are formed over the joint distribution of future wage shocks, taste shocks, and teaching job offers. Define the vector of wage and taste shocks as $\nu_{it} = [\epsilon_{1it}, \epsilon_{2it}, \eta_{1it}, \eta_{2it}]$. We assume ν_{it} is distributed multivariate Normal with mean 0_4 and covariance matrix Σ_ν. We further assume ν_{it} and job offers J_{it} are independent in each period and non-serially correlated across time. Agents know the distribution of these stochastic elements, including the arrival rate of teaching jobs for non-licensed agents (γ), but not the future sequence of realizations. Expectations for agents are over the joint distribution of these elements, which we denote $E_{\nu,J}$.

3.8 Dynamic Programming Representation

The dynamic program consists of two stages for each period t. In the first stage, ν_{it} is revealed, and the licensing decision is made. In the second stage, job offers are revealed, and the occupational choice is made. The licensing decision from the first stage (L_{it}) and teaching job offers in this period (J_{it}) are therefore state variables for the second stage occupation decision. Call the vector of remaining state variables S_{it}: $S_{it} = [x_{1it}, x_{2it}, z_i, g_i, a_i, \tau_{1i}, \tau_{2i}, t]$. Note that since the decision horizon is finite, t is also a state variable. Work experience in the two sectors evolves as $x_{1it+1} = x_{1it-1} + d_{1it}$ and $x_{2it+1} = x_{2it-1} + d_{2it}$. Initial conditions are $x_{1i1} = x_{2i1} = 0$.

The first stage value function for an individual without a license can be written as a nested maximization problem over the second stage value function:

$$\begin{aligned} V_1(S_{it}, L_{it-1} = 0) &= \max_{L_{it}} \{ L_{it} E_J[V_2(S_{it}, L_{it} = 1, J_{it}) - \kappa(z^0)] \\ &+ (1 - L_{it}) E_J[V_2(S_{it}, L_{it} = 0, J_{it})] \}. \end{aligned} \tag{6}$$

The first part of this value function is the expected value of obtaining a teacher license if the agent does not already have one. The second part of the value function is the expected value of not obtaining a license in this period and entering the second stage of the period without a license.

Given that teaching jobs for non-licensed agents arrive with probability γ, we can write (6) as

$$\begin{aligned} V_1(S_{it}, L_{it-1} = 0) = \max_{L_{it}} \{ L_{it}[V_2(S_{it}, L_{it} = 1, J_{it}) - \kappa(z^0)] \\ + \ (1 - L_{it})[\gamma V_2(S_{it}, L_{it} = 0, J_{it} = 1) + (1 - \gamma) V_2(S_{it}, L_{it} = 0, J_{it} = 0)] \}. \end{aligned} \tag{7}$$

Since licenses are permanent, the first stage value function for an already licensed agent is $V_1(S_{it}, L_{it-1} = 1) = V_2(S_{it}, L_{it} = 1, J_{it})$, where $V_2(S_{it}, L_{it} = 1, J_{it})$ is the second stage value function for an individual who has obtained a license by the first stage:

$$V_2(S_{it}, L_{it} = 1, J_{it}) = \max_{d_{1it}, d_{2it}} \{\widetilde{U}(C_{it}, B_{it}) + \delta E_\nu[V_1(S_{it+1}, L_{it} = 1)]\}. \tag{8}$$

$\widetilde{U}(C_{it}, B_{it}) = C_{it} + B_{it}$ is the utility flow for this period net of any licensing costs. S_{it+1} indicates the updated vector of state variables in the next period given the choice made in the current period, e.g. if $d_{1it} = 1$, then $S_{it+1} = [x_{1it} + 1, x_{2it}, g_i, a_i, \tau_{1i}, \tau_{2i}, t + 1]$.

The second stage value function for a non-licensed agent who is offered a job in the education sector is

$$V_2(S_{it}, L_{it} = 0, J_{it} = 1) = \max_{d_{1it}, d_{2it}} \{\widetilde{U}(C_{it}, B_{it}) + \delta E_\nu[V_1(S_{it+1}, L_{it} = 0)]\}, \tag{9}$$

The second stage value function for a non-licensed agent who is not offered a job in the education sector is

$$V_2(S_{it}, L_{it} = 0, J_{it} = 0) = \max_{d_{2it}} \{\widetilde{U}(C_{it}, B_{it}) + \delta E_\nu[V_1(S_{it+1}, L_{it} = 0)]\}. \tag{10}$$

4 Policy Evaluation

We next examine how the model can be used to evaluate changes in licensing policy. Our discussion of policy evaluation is in three parts. First, we examine how licensing policy affects the labor supply of teachers and the distribution of human capital among teachers. Second, we use the education production function framework to discuss how this changed distribution of human capital could affect education production. We discuss how economy-wide welfare could be evaluated using the tradeoff between education production and licensing costs incurred by license holders. Finally, we examine reduced form approaches to policy evaluation that use comparisons between the measured student outcomes for licensed and non-licensed teachers. We show that in general these approaches cannot identify the effects of policy changes.

4.1 Responding to Changes in Licensing Policy

The key implications of the model can be written in terms of the three component second stage value functions (8-10). For notational convenience, we write $V_2(S_{it}, L_{it} = a, J_{it} = b) = V_2(S_{it}, a, b)$.

Individual i obtains a license in period t if

$$V_2(S_{it}, 1, 1) - \kappa(z^0) \geq \gamma V_2(S_{it}, 0, 1) + (1 - \gamma) V_2(S_{it}, 0, 0). \tag{11}$$

An inspection of the maximization problems reveals that for any permissible values of the parameters, the value functions are ordered as

$$V_2(S_{it}, 1, 1) \geq V_2(S_{it}, 0, 1) \geq V_2(S_{it}, 0, 0).$$

The first inequality $V_2(S_{it}, 1, 1) \geq V_2(S_{it}, 0, 1)$ is because obtaining a license preserves the option value to enter the education sector in the future if a teaching job is not offered. The second inequality is because being offered a job in teaching expands the choice set for non-licensed agents.

Given this ordering, two implications are immediate. First, for $z^0 = 0$ (no teacher licensing requirements), all agents choose to obtain the license. Second, for $\gamma = 1$ (free entry into teaching) and $z^0 > 0$, no agents choose to obtain the license. To see the second implication, note that with $\gamma = 1$, $V_2(S_{it}, 1, 1) = V_2(S_{it}, 0, 1)$ since non-licensed agents are always offered teaching jobs. Therefore $V_2(S_{it}, 1, 1) - \kappa(z^0) > V_2(S_{it}, 0, 1)$ for all $z^0 > 0$.

Under licensing policies with non-zero training requirements ($z^0 > 0$) and restricted entry into teaching ($\gamma < 1$), the value of the license is increasing in the difference between the utility of having the option to work in the education sector and the utility from being constrained to work only in the alternative sector or stay out of the labor market. This utility difference for each heterogeneous agent depends on their relative tastes and wage offers in the two sectors. In general, licenses are more valuable for individuals with higher relative wage offers in the education sector and higher tastes for teaching.

With partially enforced licensing policies, increasing licensing requirements has two first order effects on licensing and occupational choices: 1) higher licensing requirements reduces the proportion of the population choosing to obtain a license, and 2) higher licensing requirements reduces the proportion of the population choosing to teach. The first effect is the result of the assumption that fulfilling licensing requirements is costly. The second effect follows from the first. As the proportion of the population obtaining a teacher license falls, the marginal individuals who choose to not obtain a license under the higher requirements become less likely to teach. This is because for all agents, the relative utility from teaching with a license is at least as great as the relative utility from teaching without a license. The magnitude of these effects depends on the model parameters and underlying distribution of population preferences and skills.

In addition to effects on the labor supply of teachers, licensing policy also affects the distribution of human capital in the teacher labor force. Higher formal training requirements increases the “wedge” between the types of individuals who choose to teach and those those that do not. The utility from teaching relative to other alternatives is decreasing in the licensing training requirements. As the cost of obtaining a teaching license rises, only those individuals with high relative utility in the education sector are willing to pay the higher cost of licensing. Raising licensing requirements causes individuals closer to the occupational margin to forego teaching, while those with higher relative education sector utility pay the higher cost of licensing and remain in the education sector. This implies that as licensing costs increase, occupational sorting increases, and teachers and non-teachers become more dissimilar in their endowment of skills and tastes.[16]

The increase in occupational sorting induced by the higher licensing costs has two implications for the distribution of human capital in the teacher labor force. First, higher training requirements and higher costs of obtaining a license can affect the career length and distribution of teaching experience human capital (x_{1it}) for those individuals who decide to teach. As licensing requirements increase, individuals with relatively lower teaching wage offers and tastes for teaching choose to either teach without a license or leave teaching altogether. Since these marginal teachers would have had shorter teaching careers, as these individuals leave teaching, the composition of the education sector labor force shifts toward teachers with longer career lengths.

A related second effect of higher training requirements is that the distribution of skills among teachers shifts toward teacher specific skills (a_i), which are only productive in the education sector, and away from general skills (g_i), which are productive in both sectors. Individuals with high levels of general skills receive higher alternative sector wage offers

---

[16] The sorting possible in the model is in fact more complex in the sense there can be some individuals who under lower licensing requirements would obtain a license and teach, but under higher requirements would teach without a license, while there other individuals who would leave teaching altogether.

relative to those individuals with lower levels of general skills. Therefore individuals with relatively higher levels of general skills have higher opportunity costs of teaching and become relatively more likely to forego teaching as training requirements increase.

4.2 Evaluating Licensing Policy

The model provides a mapping from licensing policy to teacher labor supply and the distribution of human capital among teachers. We next consider the welfare implications of various types of licensing policies. The main tradeoff in the welfare analysis is between the effects of licensing policy on education production through changing the size and composition of the teacher labor force versus the costs of formal teacher training borne by license holders. While our estimated model provides the first step in the welfare analysis by providing the link between licensing policy and occupational sorting, a necessary second step for a complete welfare analysis is estimating the exact form of the education production function (1). Without any education outcome data, we leave the education production function unspecified and evaluate licensing policies in the context of several possible education production function cases.

Consider how a social planner would choose licensing policy parameters z^0 and γ to maximize total education output net of total licensing costs:[17]

$$\max_{z^0,\gamma} \lambda_1 Y^*(z^0,\gamma) - \lambda_2 \kappa^*(z^0,\gamma). \tag{12}$$

$Y^*(z^0,\gamma)$ is the total education output or teacher value added produced by a college cohort in their lifetime given the licensing policy defined by z^0 and γ. $\kappa^*(z^0,\gamma)$ is the total licensing

---

[17] We assume that the social planner takes as given the current wage setting policies in the education sector (3) and the alternative sector (4).

costs incurred by a college cohort under this licensing policy. $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ are Pareto weights representing how much value is placed on education output relative to licensing costs. For example, $\lambda_1 > 0$ and $\lambda_2 = 0$ represents preferences in which only education output matters to the social planner.

We can write the total education production produced by a college cohort as

$$Y^*(z^0, \gamma) = \sum_{t=1}^{T} \{ \sum_i 1\{d_{1it} = 1 | z^0, \gamma\} \int h_1(H_{it}) f_t(H_{it} | d_{1it} = 1, z^0, \gamma) dH_{it} \}.$$

The first term is the sum of individuals choosing to teach in each period. The second term is the average education value added for these teachers. Both the number of individuals choosing to teach and the distribution of human capital among teachers $f_t(H_{it}|d_{1it} = 1, z^0, \gamma)$ depend on the licensing policy parameters z^0 and γ.[18] Similarly, we can define the total licensing costs incurred by license holders as

$$\kappa^*(z^0, \gamma) = \sum_{t=1}^{T} \sum_i 1\{L_{it} = 1 | z^0, \gamma\} \kappa(z^0)$$

where $1\{L_{it} = 1|z^0, \gamma\}$ is an indicator variable for the individuals who obtain a license given the licensing policy z^0 and γ.

Using this simple social planner framework, we next consider optimal licensing policy in the context of several types of education production functions.

Case 1: "Formal Teacher Training is Invaluable"

---

[18] The model is short-run in the sense that we ignore the issue that changes in licensing policy may affect the initial college graduate distribution of skill (a_i, g_i) and taste endowments (τ_{1i} and τ_{2i}). When we examine the effects of policy changes in a series of policy simulations, we assume that prior licensing requirements are unchanged for current teachers, and these new policies only apply to the new cohort of college graduates. Over time, as these new cohorts of teachers replace retiring teachers, who entered teaching under prior policies, the composition of all teachers will resemble the composition of the recent cohort. The analysis is therefore of the new steady state in which all cohorts of teachers have entered under the new policy.

In this case, a level of formal training below some minimum z' results in zero or negative teacher value added:

$$y_{it} = \begin{cases} h_1(z_i, \widetilde{H}_{it}) > 0 & \text{if } z_i \geq z' \\ h_1(z_i, \widetilde{H}_{it}) \leq 0 & \text{if } z_i < z' \end{cases}$$

If the welfare costs of meeting this minimum are less than the benefit of at least some education production, then the optimal licensing policy is to set the training requirement at $z_0 \geq z'$ and $\gamma = 0$, a fully enforced policy. If the social planner only values education production ($\lambda_2 = 0$), the optimal licensing policy would be to set $z^0 \geq z'$, regardless of the total licensing costs borne by license holders. Even higher training requirements may be desirable depending on the education production function and the effect licensing policy has on the distribution of human capital among teachers.

Case 2: "Formal Teacher Training is Worthless"

In this case, formal training is assumed to provide no teacher value added:

$$\frac{\partial h_1(z_i, \widetilde{H}_{it})}{\partial z_i} = 0.$$

Interestingly, this case does not necessarily imply that free entry ($z^0 = 0$ or $\gamma = 1$) is the optimal licensing policy. Given that higher training requirements can change the distribution of other types of human capital (informal teaching skills, general skills, teaching experience), non-zero training requirements may be optimal. This would be the case if, for example, higher training requirements cause lower attachment teachers to choose to not enter teaching, thereby increasing the average level of teaching experience in the the teacher labor force. This would be socially desirable if the education productivity of teaching experience is high.

Case 3: "General Skills are Invaluable"

In this case, only general skills g_i are assumed productive in producing teacher value added.

$$\frac{\partial h_1(H_{it})}{\partial g_i} > 0 \text{ and } \frac{\partial h_1(H_{it})}{\partial K} = 0, \text{for } K = \{z_i, a_i, x_{1it}, x_{2it}\}.$$

In this case, the optimal licensing policy is free entry as the highest level of education output can be reached by making teaching as attractive as possible to all individuals. Imposing entry costs on the education sector in the form of licensing requirements makes teaching less attractive to individuals with high levels of general skills who have relatively higher wage offers in the alternative sector.

4.3 Reduced Form Approaches to Policy Evaluation

Previous research on teacher licensing uses student level measures of teacher value added to compare licensed and non-licensed teachers. The intention of this research is to ascertain whether the training required by licensing improves student outcomes, that is, measure the educational productivity of formal teacher training. We can summarize this methodology by considering within the context of our behavioral model the difference in mean teacher value added between teachers who are licensed and have obtained a $z_i = z^0$ level of formal teacher training with non-licensed teachers who have no formal teacher training ($z_i = 0$):

$$\Delta_t(z^0, \gamma) = E[h_1(z^0, \widetilde{H}'_{it})|d_{1it} = 1, L_{it} = 1, z^0, \gamma]$$

$$-E[h_1(0, \widetilde{H}_{it})|d_{1it} = 1, L_{it} = 0, z^0, \gamma],$$

where $\widetilde{H}_{it}$ is the vector of human capital excluding formal teacher training: $\widetilde{H}_{it} = [a_i, g_i, x_{1it}, x_{2it}]$. The expectations are with respect to the distribution of human capital for licensed and non-licensed teachers under the current licensing policy (z^0, γ):

$$\Delta_t(z^0, \gamma) = \int h_1(z^0, \widetilde{H}_{it}) f_t(\widetilde{H}_{it} | d_{1it} = 1, L_{it} = 1, z^0, \gamma) d\widetilde{H}_{it}$$

$$- \int h_1(0, \widetilde{H}_{it}) f_t(\widetilde{H}_{it} | d_{1it} = 1, L_{it} = 0, z^0, \gamma) d\widetilde{H}_{it}$$

where $f_t(\widetilde{H}_{it} | d_{1it}, L_{it}, z^0, \gamma)$ is the distribution of human capital in period t conditional on the occupational choices, licensing choices, and licensing policy in the economy at that date.

There are two issues with this approach. First, an estimator for $\Delta_t(z^0, \gamma)$ does not identify the "causal effect" of formal teacher training on student outcomes. In the causal effects framework, the *average treatment effect* of licensing is

$$ATE_t(z^0, \gamma) = E[y_{1it} | L_{it} = 1] - E[y_{1it} | L_{it} = 0]$$

$$= \int h_1(z^0, \widetilde{H}_{it}) f_t(\widetilde{H}_{it} | z^0, \gamma) d\widetilde{H}_{it} - \int h_1(0, \widetilde{H}_{it}) f_t(\widetilde{H}_{it} | z^0, \gamma) d\widetilde{H}_{it}$$

In general, $ATE_t(z^0, \gamma) \neq \Delta_t(z^0, \gamma)$ if the distribution of human capital is not independent of occupation and licensing decisions:

$$f_t(\widetilde{H}_{it} | z^0, \gamma) \neq f_t(\widetilde{H}_{it} | d_{1it} = 1, L_{it} = 1, z^0, \gamma) \neq f_t(\widetilde{H}_{it} | d_{1it} = 1, L_{it} = 0, z^0, \gamma)$$

The non-identification of the average treatment effect of training is due to the non-randomness of licensing choices. If individuals self-select into licensing based on factors that also affect the distribution of other components of teacher value added, including teacher specific skills, general skills, teaching experience, etc., we cannot consistently estimate the causal effect of formal teacher training using estimated differences in outcomes for licensed and non-licensed teachers. The ideal social experiment is where some randomly chosen college graduates are assigned to take the education courses required by teacher licenses and a control group is not. Because of self-selection, what we observe in data from comparisons of licensed and non-licensed teachers is a combination of the productivity of training and the productivity of other sources of teacher human capital.

A second, and distinct, issue is that even if we are able to consistently estimate the average treatment effect of teacher training, this would not identify the effects of policy. A crucial part of the policy question is how the distribution of human capital among teachers is affected as the licensing policy is changed. There is no reason to believe that the average treatment effect under the current licensing policy, $ATE_t(z^0, \gamma)$, is the same as under an alternative policy, $ATE_t(z^{0\prime}, \gamma')$. This issue is one of extrapolation. What we cannot determine from a comparison of licensed and non-licensed teachers under the current licensing policy is how the distribution of teaching human capital and student outcomes would be affected as the licensing policy is changed.

5 Estimation

The model is solved through a backwards recursion of the dynamic program and estimated on a nationally representative panel of college graduates using a simulated method of moments estimator.

5.1 Empirical Specification

The model is specified up to the model parameters and the distribution of the population heterogeneity in tastes (τ_{1i} and τ_{2i}) and skills (a_i and g_i). We do not observe actual training requirements z^0. Instead we identify the cost of acquiring the level of teacher formal training necessary to obtain a license, $\kappa(z^0)$. We generalize the licensing costs specification to include a time varying component. In the empirical specification, licensing costs for period t are modeled as $\kappa(z_0) = \kappa_t = \kappa_0 + \kappa_1(t-1)$, where $\kappa_0 \geq 0$ and $\kappa_1 \geq 0$ are parameters to be estimated. This specification is intended to capture the psychic costs or the costs of career interruption to completing licensing requirements later in life.

We set the last period of the model, the exogenous retirement date, to $T = 33$ when agents are approximately aged 55-57. The discount rate is assumed $\delta = 0.95$. The time varying shocks $\nu_{it} = [\epsilon_{1it}, \epsilon_{2it}, \eta_{1it}, \eta_{2it}]$ are assumed uncorrelated with variances $V[\nu_{it}] = [\sigma_{\epsilon 1}, \sigma_{\epsilon 2}, \sigma_{\eta 1}, 1]$, where the taste shock for the non-teaching occupation is normalized to 1. We assume the discount rate is $\delta = 0.95$. The full set of common parameters to be estimated are then $\theta_1 = [\beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}, \kappa_0, \kappa_1, \gamma, \sigma_{\epsilon 1}\sigma_{\epsilon 2}, \sigma_{\eta 1}]$.

We do not observe in data tastes in the two occupations τ_{1i}, τ_{2i}, general skills g_i, and specific teaching abilities a_{1i}. Because we can only infer skill levels from data on accepted wages in the education and non-education sectors, we cannot separately identify the contribution of various unobserved skills to wages. Instead we attempt to identify the distribution of the two scalar terms $\pi_{1i} = \pi_1(g_i, a_{1i})$ and $\pi_{2i} = \pi_2(g_i)$. Heterogeneity in the population is then fully determined by the population distribution of the vector $A_i = [\pi_{1i}, \pi_{2i}, \tau_{1i}, \tau_{2i}]$.

A_i is assumed to have a discrete distribution with four points of support, where these points can be thought of as distinct agent "types." Each of the four types, characterized by one of the vectors $A^j = [\pi_1^j, \pi_2^j, \tau_1^j, \tau_2^j]$, for $j = 1, 2, 3, 4$, has probability mass p_j, where $p_j \geq 0$ for all $j = 1, 2, 3, 4$, and $p_4 = 1 - \sum_{j=1}^{3} p_j$. The 19 heterogeneity parameters to be estimated are given by $\theta_2 = [A^1, A^2, A^3, A^4, p_1, p_2, p_3]$.[19]

---

[19] This characterization of the unobserved heterogeneity is similar to Heckman and Singer (1984) and widely used in more recent empirical work, e.g. Keane and Wolpin (1997). Including more points of support

5.2 Dynamic Program Solution

Starting with period $t = T$, we first compute the expectation of the value function at each of the feasible state points. For each S_{iT} feasible state point, there are two licensing states: $L_{iT-1} = 1$ and $L_{iT-1} = 0$. For notational convenience, let $Q_{iT} = \max_{d_{1iT}, d_{2iT}} \{\widetilde{U}(C_{iT}, B_{iT})\}$ and $\widetilde{Q}_{iT} = \max_{d_{2iT}} \{\widetilde{U}(C_{iT}, B_{iT})\}$. The nested maximization problem (7) implies that we need to calculate:

$$E_\nu V_1(S_{iT}, L_{iT-1} = 1) = \int Q_{iT} \phi_\nu d\nu, \text{ and}$$

$$E_\nu V_1(S_{iT}, L_{iT-1} = 0) = \int \max\{(Q_{iT} - \kappa_T), (\gamma[Q_{iT}] + (1-\gamma)[\widetilde{Q}_{iT}])\} \phi_\nu d\nu,$$

where $\phi(\nu)$ is the four dimensional Normal density. Given the high dimensionality of this integral, we use Monte Carlo methods to approximate these integrals:

$$\widehat{E}_\nu V_1(S_{iT}, 1) = \frac{1}{R} \sum_r Q_{iTr}, \text{ and} \tag{13}$$

$$\widehat{E}_\nu V_1(S_{iT}, 0) = \frac{1}{R} \sum_r \max\{(Q_{iTr} - \kappa_T), (\gamma Q_{iTr} + (1-\gamma)\widetilde{Q}_{iTr})\}, \tag{14}$$

where $Q_{iT,r}$ and $\widetilde{Q}_{iTr}$ are the solutions to the maximization problems given a vector of draws ν_{iTr} from $\phi(\nu)$.

With a Monte Carlo approximation for the value functions (13) and (14), we can then

---

becomes increasingly computationally burdensome. We found that 4 points of support (with 19 associated parameters) was sufficient to provide reasonable estimates for the other parameters. This number of points of support is in the range often used in other work, see Mroz (1999) and Cameron and Taber (1998).

move backwards to period $T-1$ and calculate

$$\widehat{E}_\nu V_1(S_{iT-1},1)=\frac{1}{R}\sum_r Q_{iT-1r}+\delta\widehat{E}_\nu V_1(S_{iT},1)$$

$$\widehat{E}_\nu V_1(S_{iT-1},0)=\frac{1}{R}\sum_r \max\{(Q_{iT-1r}+\delta\widehat{E}_\nu V_1(S_{iT},1)-\kappa_{T-1}),$$
$$(\gamma[Q_{iT-1r}+\delta\widehat{E}_\nu V_1(S_{iT},0)]+(1-\gamma)[\widetilde{Q}_{iT-1r}+\delta\widehat{E}_\nu V_1(S_{iT},0)])\}$$

We can continue to move backward in this fashion to solve for the expectation of all first stage value functions $t=2,\ldots,T$.

5.3 Estimator

The data described below contain observations of $L_{it}, d_{1it}, d_{2it}, w_{1it}, w_{2it}, x_{1it}, x_{2it}$ for a panel of sample size N. Under the assumed data generating process the following M moment conditions hold at the true parameters θ_1^* and θ_2^*:

$$E[\widehat{Y}-E_{A|\theta_2^*}E_\nu m(A,\nu,\theta_1^*)]=0_M,$$

where $\widehat{Y}=\frac{1}{N}\sum_{i=1}^N Y_i$ is a $Mx1$ vector of aggregate observed moments calculated from the data, and $E_{A|\theta_2}E_\nu m(A,\nu,\theta_1)$ is a $Mx1$ vector of moments calculated from the assumed data generating process given θ_1 and θ_2 parameters.

The simulated method of moments estimator for θ_1^* and θ_2^* solves

$$\min_{\theta_1,\theta_2}\; g(\theta_1,\theta_2)'Wg(\theta_1,\theta_2),$$

where

$$g(\theta_1, \theta_2) = \widehat{Y} - \sum_{k=1}^{K} p_k \frac{1}{NS} \sum_{ns=1}^{NS} m(A^k, \nu_{ns}, \theta_1),$$

and W is an MxM weighting matrix. The estimation algorithm proceeds as follows. Given a candidate set of parameters, θ_1 and θ_2, we first solve the dynamic program for each of the $K = 4$ points of support on the discrete distribution. Given the solution to the dynamic program, we then draw NS $\nu_{ns} = \nu_{ns,1}, \ldots, \nu_{ns,T}$ sequences from the distribution of multivariate Normal shocks. For each of the K types, we then solve the model forward from $t = 1$ for each sequence of ν_{ns} draws. Moments are calculated for each of the K types. Unconditional moments are computed based on weighting the type specific moments using the candidate $p_1, \ldots, p_K$ probabilities. The moments used in the estimation include the probability of licensure for each period of observed data, probabilities of occupational choices for each period by licensure, first and second moments of wages in both occupations by licensure and period, transition probabilities between occupations, and between period differences in wages by occupation. The weighting matrix is the inverse of the variance the estimated data moments with off-diagonal elements ignored.[20]

---

[20] Identification of structural parameters is through functional form and distributional assumptions. Licensing policy parameters κ_0, κ_1, and γ are identified through the observed behavior of agents whose actions are taken in response to these parameters. Heuristically, a positive proportion of non-licensed agents in the data implies that current policy licensing costs are positive. On the other hand, licensing costs cannot be too high or we would observe no one obtaining a license. γ, along with other parameters, determines the rate at which non-licensed individuals are observed entering teaching. γ is separately identified because of its relationship to teaching choices for non-licensed agents. A low value of γ implies that no non-licensed individuals teach, whereas a high value implies at least some do teach.

6 Data

6.1 B & B Data

The main source of data is the Baccalaureate and Beyond (B&B) survey. This longitudinal survey provides detailed information on the college experiences and post-graduation employment for the 1992-93 college graduating class. In 1993, the B&B surveyed a nationally representative sample of college graduates from American universities who received their bachelor degree during the 1992-93 academic year (July 1992 to June 1993). After graduation, subsequent B&B surveys in 1994, 1997, and 2003 re-surveyed these same respondents. 11,192 college graduates responded to the initial survey. The survey collected information on freshman admissions test scores (SAT and ACT scores), college majors and course taking while in college, and employment and earnings during each interviews. Most importantly for this study, the survey collected rich information about teaching employment, including the type and date of teacher licenses obtained.

We use employment information for each of the three post-graduation interviews to classify respondents into one of three mutually exclusive activities: working as a K-12 teacher, working in a non-teaching occupation, and out of the labor force.[21] For the 1994 and 1997 interviews, the survey asks employment information for the month of April of that year. The 2003 interview asks for current employment information. If the individual reports working either full or part-time, we classify them as employed. We then use self-reported occupations to classify the respondent as working as a K-12 teacher or working in a non-teaching occupation. The remaining respondents are classified as out of the labor force.

Full time, full year annual salaries are constructed based on reported pay by multiplying

---

[21] We make no distinction between teachers working in public or private schools. About 10 percent of all K-12 teachers work in private schools.

reported pay by 12 if the respondent reported being paid by the month, 50 if paid by the week, 250 if paid by the day, and 2000 if paid by the hour. All salaries are reported in 2003 dollars.

Individuals are classified as having a teacher license if the respondent received a "probationary," "regular," or "advanced" license by the date of the employment information (April for the 1994 or 1997 interviews or by the interview date for the 2003 interview). Most states grant "probationary" licenses (certification) for individuals who have completed the "regular" or "advanced" requirements but have not met a minimum teaching experience requirement (typically 1 or 2 years). Respondents are classified as not having a license if they report no license or report having obtained a type of alternative license ("other," "emergency," or "temporary").[22]

The sample used in the estimation is limited to the sample who responded to all three waves of the survey and provided non-missing employment, occupation, licensing, earnings, race, college major, and test score data. The sample is further limited to respondents who were aged 24 years and younger at the time of college graduation. This excludes about 23 percent of the sample who were 25 and older at graduation. We further exclude about 10 percent of the sample which reported annual salaries either lower than \$10,000 or higher than \$150,000 in 2003 dollars. The final sample consists of a balanced panel of 4,095 respondents, with three post-college observations 1994 (when individuals are aged about 24), 1997 (aged about 27), and 2003 (aged about 33).[23]

---

[22] In the empirical specification, licenses are assumed not to expire. In reality, many states require additional course work to renew a license every few years. All individuals who reported having a license in early interview are assumed to have a license in later periods.

[23] Because the original sample was collected first as a cluster sample of schools and then stratified according to a number of demographic characteristics, sample weights are used in the descriptive statistics and estimation of structural parameters.

6.2 Descriptive Statistics

Table 1 reports basic descriptive statistics for the entire B&B sample. Across the three years of data, 8-10 percent of the sample teaches, about 3/4 is employed in a non-teaching occupation, and 12-14 percent of the sample is out of the labor force. Table 2 reports descriptive statistics for the sample broken down by occupation and licensure. In 1994, when the teachers are aged about 24, about 85 percent of these beginning teachers taught with a license, and 15 percent without a license. By 2003, when the respondents in the sample are aged approximately 33, the fraction of teachers with a licensure increases to about 92 percent. It is important to note that about 4 to 5 percent of individuals reporting that they worked in a non-teaching occupation also reported holding a teacher license. This is not inconsistent with the theoretical model, as individuals may obtain a license because of the option value of having access to the teacher labor market in the future. In terms of demographic composition, the gender composition by licensure stands out. Overall, 78 percent of teachers are women. However, women are more likely to obtain a license as 80 percent of licensed teachers are women compared to 2/3 of non-licensed teachers. This is consistent with the greater propensity for female college graduates to choose teaching as a long term career.

6.3 CPS Data

In our initial estimation attempt, the estimated model over-predicted wage growth over later ages. To capture more reasonable life-time wage growth, we use additional cross-sectional moments from the 2003 March Current Population Survey (CPS). In particular, we include first and second moments of annual earnings for individuals aged 51-55 by occupation. Using occupation and education codes in the 2003 March CPS, we construct a sample of college graduates and classify each graduate as a teacher if they report an occupation as an elementary or middle school, secondary, or special education teacher. We then apply the variable definitions, as closely as possible, used for the B&B data to construct annual salaries from reported weekly earnings in the CPS. When comparing mean wages for individuals aged 33 in the March CPS and the B&B sample, we found that mean wages were about 13.5 percent higher in the B&B sample than in the March CPS sample. This difference may be due to greater levels of mis-reported education in the March CPS or differences in the earnings measures. To make the samples as comparable as possible, we adjusted earnings in the March CPS upward 13.5 percent thereby making wages for respondents aged 33 equal in both samples.

7 Descriptive Evidence

7.1 Attachment to Teaching

The theoretical model predicts that licensed teachers would have a longer attachment to teaching than non-licensed teachers. Figure 2 graphs the difference in the distribution of years taught by age 33 for licensed and non-licensed teachers. As found in other research (e.g. Stinebrickner 2001, 2002), there is considerable attrition from teaching overall. However, we also find that there is a distinct pattern by licensing status. During the first 9 years of the sample, non-licensed teachers teach on average 2.5 years, whereas licensed teachers teach about 5 years on average. The longer attachment to teaching for licensed teachers is due to two factors. First, mechanically, teachers with a license are more likely to be offered future teaching jobs. Second, individuals with greater preferences for teaching or lower outside wage offers are more likely to obtain a teaching license.

7.2 Sorting on General Skills

The theoretical model predicts that individuals choose to obtain a license based on relative tastes and skill endowments. We use scores on university admissions examinations, in particular Scholastic Aptitude Test (SAT) scores, as a measure of general skills.[24] Previous research has found a negative relationship between a college graduate's SAT score and the probability the graduate works as a teacher (e.g. Manski 1987, Ballou and Podgursky 1997).[25] This relationship has been interpreted as indicating that high SAT college graduates have higher general skills and a larger opportunity cost of teaching due to higher non-teaching wage offers. Indeed, in the first year of the B&B panel, the correlation between SAT scores and non-teaching earnings is 0.12, whereas the correlation between SAT scores and teaching earnings is -0.02. What has not received much attention in previous research is that the distributions of SAT scores for licensed and non-licensed teachers are quite distinct.

Table 3 divides the sample into four groups by whether the respondent ever obtained a teaching license and ever taught in a K-12 school. Consistent with the prior research, mean SAT scores for all teachers are 59 points lower than mean SAT for all non-teachers. However, the difference between non-licensed teachers and non-licensed non-teachers is near zero. On the other hand, mean SAT scores for licensed teachers are about 63 points lower than for non-licensed teachers. This difference in SAT scores is consistent with the model prediction that licensing costs causes individuals to sort on general skills.[26]

---

[24] Some respondents reported scores from another admissions test, the American College Testing (ACT) examination. Because more of the sample took the SAT than the ACT, we use SAT scores in the analysis below. For the 23 percent of the sample that only took the ACT, we impute their SAT score by calculating the sample SAT percentile score corresponding to their sample percentile on the ACT. The SAT and ACT scores from the B&B sample come from the testing companies and college transcripts.

[25] As describe above, a similar relationship has been found for other observable measures of ability, e.g. Ehrenberg and Brewer 1994.

[26] In a series of robustness checks (available on request), we estimate the difference in SAT scores for teachers including additional variables as controls. Including demographic variables for race and gender, variables for the type of school taught at (public or private), and variables for level of school taught at

7.3 Occupational Transitions and Wage Dynamics

Table 4 provides evidence on occupational transitions and wage dynamics. Consistent with the finding of lower attachment to teaching for non-licensed teachers, 36 percent of non-licensed teachers in 1994 (approximately age 24) also teach in 2003 (approximately age 33), compared to 62 percent of licensed teachers. 53 percent of non-licensed teachers in 1994 are working outside of teaching in 2003, compared to only 17 percent of licensed teachers.

In 1994, the first year following college graduation, there are relatively small differences in salaries across occupations or licensing status. By 2003, much larger differences in salaries are apparent. By 2003, the mean salary for the majority of the sample, which neither taught in 1994 or 2003 nor obtained a teacher license, is \$58,482. For licensed teachers who taught in both 1994 and 2003, the mean teaching salary is 35 percent lower at \$38,157. Mean teaching salaries for non-licensed teachers in 2003 are nearly identical at \$40,703. Examining the differences in salaries between 2003 and 1994 indicates that larger salary changes occurred for those individuals who left teaching. The mean 2003 salary for individuals who taught in 1994, but by 2003 left teaching, is substantially larger than the salary for individuals who remained in teaching, both for individuals with and without a license. The groups who left teaching and experienced this higher salary growth also had higher mean SAT scores. These patterns in salary growth and SAT scores suggest that the occupational transitions, in particular the attrition from teaching, are non-randomly based on general skills and non-teaching wage offers.

(elementary or secondary) has little effect on the estimated difference in SAT scores between licensed and non-licensed teachers.

8 Estimation Results

8.1 Parameter Estimates

Table 5 reports the parameter estimates for the common parameters, and Table 6 reports the estimated distribution of unobserved heterogeneity. We estimate that the linear return to experience is 7.8 percent in the teaching occupation and 8.2 percent in the non-teaching occupation. The major distinction between the two occupations is that we estimate the standard deviation in wage shocks for the non-teaching occupation to be twice as high as in the teaching occupation: $\sigma_{\epsilon 1} = 0.18$ versus $\sigma_{\epsilon 2} = 0.36$.

The fixed cost of obtaining a license for a college graduate in the first period is estimated at \$29,030. Given the value for κ_1, the licensing cost is estimated to increase as the individual ages. We estimate that the cost to obtaining a license increases by \$2,000 each period an individual delays in obtaining the license. The arrival rate of teaching job offers for non-licensed agents is estimated at $\gamma = 0.87$. This value is inclusive of non-licensed job offers from private school employers as well as public school employers.

8.2 Sample Fit

Table 7 examines the within sample fit of the parameter estimates. Unsurprisingly, in general, the model fit is not as good for lower probability events, such as teaching without a license, which encompasses less than 1.5 percent of the total college graduate sample, as it is for the events that have much higher probabilities, such as working outside of teaching. The estimated model over-predicts the fraction of the sample obtaining a teaching license in the initial period following college graduation (1994) by about 13.8 percent, and under-predicts the fraction obtaining a license in 2003 by about 11.3 percent.

The model performs better in predicting occupational choices. The model predicts that 6.8 percent of individuals teach with a license in the initial period, compared to the actual of 6.5 percent. The model predicts 1.2 percent teach without a license, compared to the actual of 1.1 percent. The model under-predicts by about half the number of individuals working in the non-teaching sector with a license. But the model is fairly close in predicting that 74.2 percent agents initially work outside of teaching without a license, compared to the actual fraction 74.9 percent. For 2003, the model predicts that only 6.7 percent of individuals teach with a license, compared to the actual of 8.9 percent. The fraction teaching without a license is predicted at 1.4 percent, compared to the actual of 0.8 percent. The model does make a close prediction of the fraction choosing to work outside of teaching without a license at 74.0 percent, compared to the actual of 72.8 percent.

The estimated model's predictions for mean wages in the initial period (1994) are relatively close for each of the occupational and licensing states. Teaching salaries both for 2001 and ages 51-55 (CPS data) are relatively well predicted by the estimated model. However, one of the major absences of within sample fit is the failure of the estimated model to capture the growth in non-teaching wages by 2003. For 2003, the model predicts mean non-teaching salaries are \$45,900, whereas the actual is \$59,000. This lack of model fit may be due to the inability to fit the rapidly rising wage profile for college graduates using a standard Mincerian two parameter log wage equation (linear and quadratic terms in experience).

9 Policy Simulations

This section uses the estimated model parameters to examine the effects of changes in licensing policy. The simulations are based on 2,000 simulated draws for each of the four agent types using the estimated parameters in Tables 5 and 6. We consider changes in the cost to obtaining a license (κ_0 and κ_1), which represents changes in training requirements, and changes in the arrival rate of teaching job offers (γ), which represents changes in the enforcement of the licensing requirements. The policy simulations change the licensing policy for the new cohort of college graduates and simulates their behavior forward until their retirement. Over time, as these new cohorts of teachers replace retiring teachers, who entered teaching under prior licensing policies, the composition of all teachers will resemble the composition of the recent cohort. Therefore the results presented here indicate what the new long-term steady state would be under alternative licensing policies. Five outcomes are examined: 1) total lifetime labor supply of teachers, 2) the fraction of teachers who are licensed, 3) the average length of teaching careers, 4) the occupational sorting of individuals with different types of endowments and opportunity costs of teaching, and 5) the per capita costs of licensing.

9.1 Teacher Labor Supply

Teacher labor supply is measured as the per capita number of periods all agents work as a teacher during their lifetime:

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T}\sum_{t=1}^{T}1\{d_{1it}=1\}$$

where N is the total number of simulated individuals. Figure 3 graphs the change in teacher labor supply over different licensing costs. The horizontal axis measures licensing costs relative to the estimated current cost. At a value of 1, the licensing cost is set to the estimated parameters $\kappa_0 = \widehat{\kappa}_0$ and $\kappa_1 = \widehat{\kappa}_1$. A value less than 1 is a reduction in licensing costs relative to the estimated value, e.g. a value of 0.5 is a 50 percent reduction in licensing cost ($\kappa_0 = 0.5\widehat{\kappa}_0$ and $\kappa_1 = 0.5\widehat{\kappa}_1$). The origin is free entry into teaching ($\kappa_0 = 0$ and

$\kappa_1 = 0$).[27]

Under the current policy, total lifetime teacher labor supply is estimated at 2.67 years. At free entry, the simulations indicate that total lifetime teacher labor supply is 2.76 years, 3.4 percent higher than under the current licensing policy. By definition, all of this labor supply under free entry is from licensed teachers. As licensing costs increase, some agents choose to teach without a license, and the fraction of teacher labor supply from non-licensed teachers increases. At the same time, the labor supply from licensed teachers also declines as licensing costs increase. At the estimated parameters, total lifetime labor supply is 2.21 years. Eliminating licensing requirements is predicted to increase total teacher labor supply by 24.9 percent. A doubling of licensing costs ($\kappa_0 = 2\widehat{\kappa}_0$ and $\kappa_1 = 2\widehat{\kappa}_1$) reduces total labor supply to 2.04 years, a 7.7 percent reduction from the current licensing policy.

9.2 Proportion of Teachers with a License

Figure 4 graphs the proportion of lifetime teacher labor hours supplied by non-licensed teachers. By definition, at free entry all teacher labor is supplied by licensed teachers. As licensing costs increase, the proportion of teacher labor supplied by licensed teachers declines. The reduction in the fraction of licensed teachers is fairly rapid with small increases in licensing costs from free entry. As shown below, Type 2 and Type 4 individuals, which have relatively high comparative advantage in the non-teaching occupation, enter teaching under no or low licensing costs, but are far less likely to enter teaching under even modest levels of licensing costs. In contrast, the proportion of teachers without a license falls less rapidly under increases in licensing costs from the current policy. As costs double from the

---

[27] Although not considered here, one could evaluate other changes in licensing policy, such as lowering costs of licensing for older agents who want to make mid-career entry into teaching. This would involve a change in κ_1 while leaving κ_0 fixed.

estimated current level, the proportion of teachers with a license falls from about 83 percent to 77 percent.

9.3 Attachment to Teaching

The theoretical model predicts that as licensing costs increase, individuals with shorter anticipated career lengths in teaching increasingly choose not to work as teachers or obtain a teaching license. We measure attachment to teaching using the mean number of years taught for those individuals who teach:

$$\frac{1}{N_1}\sum_{i\in I_1}\frac{1}{T}\sum_{t=1}^{T}1\{d_{1it}=1\},$$

where I_1 is the set of simulated individuals that teach (either licensed or non-licensed teachers, depending on the measure), and N_1 is the total number of individuals who teach.[28]

Figure 5 graphs the mean career length for licensed and non-licensed teachers over a range of potential licensing costs. Under free entry, all teachers are licensed, and the mean career length for licensed teachers is 11.36 years. Under a licensing cost of 10 percent of the current cost ($\kappa_0 = 0.1\widehat{\kappa}_0$ and $\kappa_1 = 0.1\widehat{\kappa}_1$), non-licensed teachers teach on average for about 1.03 years, compared to 4.34 years under the current policy. The average career length for non-licensed teachers increases 10 percent to 4.78 years under a doubling of licensing costs. For licensed teachers, the average career length increases to 17.89 years, or a 14.5 percent increase from the current policy, under a doubling of the licensing cost.

The total number of years taught by all teachers (licensed and non-licensed teachers) rises as licensing costs increase, but not as rapidly as the mean career length for licensed

---

[28] An individual can count as both a licensed and non-licensed teacher if the individual teaches without a license first and later obtains a a license.

teachers. Mean career length for all teachers increases 20.4 percent from free entry to the current policy, and increases a further 8.6 percent from the current policy to a doubling of licensing costs. Recall that as licensing costs increase, the proportion of teachers who teach without a license rises. As licensing costs rise, the composition of the teacher labor force becomes more weighted toward lower attachment non-licensed teachers, thus lowering the mean attachment for teachers overall.

It is important to note that occupational sorting affects the level of attachment for both licensed and non-licensed teachers. As the licensing cost increases, some low attachment individuals, who, under low licensing cost policies, would decide to obtain a license and teach for some relatively short period of time, under higher licensing costs, decide to exit teaching entirely. On the other hand, some higher attachment individuals decide not to obtain a license as the cost increases, but still teach for some periods. The influx of these higher attachment types increases the mean career length for non-licensed teachers.

9.4 Occupational Sorting

Figures 6 and 7 graph the distribution of each of the four types among the non-licensed and licensed teachers, respectively. Under low licensing costs, three types have significant representations among the non-licensed teachers. However, as licensing costs increase, the non-licensed teachers become dominated by Type 2 individuals. A similar pattern of occupational sorting is displayed for licensed teachers in Figure 7. Under low licensing costs, about 20 percent of licensed teachers are Type 3 and about 80 percent of licensed teachers are Type 1. As licensing costs increase, licensed teachers become dominated by Type 1 individuals.

The implications of this occupational sorting for the teacher labor force is that as licensing costs change, the types of individuals who become teachers also change. As discussed above, we have no direct measure of teacher quality or specific skills that contribute to teacher value added. Instead, we measure the quality of teachers using foregone non-teaching earnings.[29] Define the average lifetime opportunity cost of teaching as

$$\frac{1}{N_1}\sum_{i\in I_1}\frac{1}{T_{1i}}\sum_{t=1}^{T}1\{d_{1it}=1\}w_{2it},$$

where T_{1i} is the total number of periods this individual teaches: $T_{1i}=\sum_{t=1}^{T}1\{d_{1it}=1\}$. w_{2it} is the non-teaching wage each individual forgoes. To the extent that these foregone wages are based on general skills and are a productive education input, the average opportunity cost of teachers provides a measure teacher quality.

Figure 8 graphs the ratio of the opportunity cost measure for non-licensed vs. licensed teachers. Under free entry, all teachers have a license, and the ratio is 1 under this policy. If licensing does not affect occupational sorting, we would expect that this ratio would remain at 1 as licensing costs increase. Using the estimated parameters, we find that for all levels of licensing costs, non-licensed teachers have a higher level of opportunity cost of teaching than licensed teachers. This is consistent with the difference in SAT scores between licensed and non-licensed teachers documented above. As the licensing cost increases, the gap between the licensed and non-licensed teachers widens. At the current level of licensing costs, the ratio of average foregone wages for non-licensed to licensed teachers is about 1.8. If licensing costs are doubled, this ratio increases to about 1.9.

Figure 9 graphs the opportunity cost measure for non-licensed, licensed, and all teachers. The most noticeable feature is the sharp increase in the opportunity cost measure for non-licensed teachers with only a modest increase in licensing costs from free entry. The sharpness

---

[29] It is unlikely that teacher wages provide much indication of teacher quality. As noted above, non-teaching wages are positively correlated with SAT scores, while teaching wages are not.

of this increase is somewhat an artifact of the discrete nature of the assumed underlying distribution of heterogeneity with 4 points of support. As seen in Figure 6, higher licensing costs push Type 2 and Type 4 individuals from non-licensed teaching, and non-licensed teachers become dominated by Type 3 individuals. In contrast to the pattern for the non-licensed teachers, the quality of licensed teachers, as measured by foregone wages, declines as licensing costs increase. As licensing costs increase, only individuals with relatively low opportunity costs of teaching are willing to pay the cost of licensing. The overall quality of the teacher labor also declines with increasing licensing costs, but not as rapidly due to the increasing proportion of the relatively higher quality non-licensed teachers. Overall, the mean lifetime opportunity cost for all teachers declines from \$31,083 under free entry to \$30,489 under a doubling of licensing costs, about a 2 percent decline.

9.5 Changes in Licensing Enforcement

Table 8 summarizes the results from the main policy experiments. In addition, Table 8 also displays results from changes in licensing enforcement (γ). The first three columns report the results for the estimated current policy ($\kappa_0 = \widehat{\kappa}_0$, $\kappa_1 = \widehat{\kappa}_1$), free entry ($\kappa_0 = 0$, $\kappa_1 = 0$), and a doubling of licensing costs ($\kappa_0 = 2\widehat{\kappa}_0$, $\kappa_1 = 2\widehat{\kappa}_1$). The fourth column considers a policy in which the licenses are fully enforced ($\gamma = 0$) and only individuals with licenses are allowed to teach. Note that under this fully enforced licensing policy, the cost of the license remains at the estimated values ($\kappa_0 = \widehat{\kappa}_0$, $\kappa_1 = \widehat{\kappa}_1$). The remarkable aspect of this policy is that the total teacher labor supply is nearly the same as under free entry. Per capita total teacher labor supply under the fully enforced policy is actually higher than under the current policy (2.758 periods versus 2.675 periods). The reduction in the arrival rate of teaching jobs to non-licensed agents has caused individuals who previously taught without a license now to choose to obtain a license. The change in the composition of the licensed teacher labor force with the entry of these previously non-licensed reduces the mean number of periods taught for licensed teachers to 11.46 years compared to 15.62 years under the current policy, a 26.7 percent reduction in average career length. Comparing the mean career length for all teachers under the current policy with that under the fully enforced policy, we find that mean career length falls 16.2 percent (from 13.68 to 11.46 years).

The fifth column of Table 8 considers a combination of policies: a fully enforced policy ($\gamma = 0$) and a doubling of licensing costs ($\kappa_0 = 2\widehat{\kappa}_0$, $\kappa_1 = 2\widehat{\kappa}_1$). These results are similar to those for the fully enforced policy without the increase in licensing costs reported in the fourth column. Total per capita teacher labor supply is slightly reduced from 2.76 periods under free entry to 2.73 periods, but this is still slightly higher than under the current policy. The mean number of periods taught increases marginally under this combined policy to 11.63 average periods.

The last rows of Table 8 examine the opportunity cost measure of teacher quality under the various policies. Under the fully enforced policy, the mean teaching opportunity cost is \$31,081. This is higher than under the current policy, but slightly lower than under free entry. For the combined policy of fully enforcing the license and doubling licensing costs, the mean opportunity cost measure is smaller at \$30,873. Interpreting the opportunity cost measure as an indication of teacher quality, both of these fully enforced policies increase the quality of the teacher labor force compared to the current policy. The intuition for this result is that the higher level of enforcement of the license causes some individuals who previously would have taught without a license to obtain the teacher license. Because these individuals have been induced to obtain a teacher license, their share of the teacher labor supply grows both because they are always offered a teaching job and because teaching experience human capital becomes more valuable to them. The group of licensed teachers is then more heavily weighted toward the higher opportunity cost and higher quality previously non-licensed teachers. This results in a higher quality overall teacher labor relative to the current policy.

9.6 Total Licensing Costs

As discussed above, a fuller welfare analysis of licensing policy would consider the costs of completing licensing training requirements borne by license holders. To measure the costs of licensing, we calculate a per capita total licensing cost under each policy:

$$\frac{1}{N}\sum_{i=1}^{N}(\kappa_0 + \kappa_1 t)1\{L_{it} = 1\}.$$

This measure is similar to the total teacher labor supply measure as it measures the per capita cost of licenses to all college graduates. The last row of Table 8 reports the per capita licensing cost across the different policies. Under the current policy, the per capita licensing cost is \$4,220. The total cost of licensing in the economy is then $N^* * \$4,220$, where N^* is the number of college graduates. Under free entry, by definition, the cost to licensing is zero. With a doubling of licensing costs, the per capita cost of licensing increases to \$6,620. Note that this is not a doubling of costs from the current policy since fewer individuals choose to obtain a teacher license under the higher licensing cost. Under the fully enforced licensing policy, the per capita cost to licensing is even higher at \$7,010. With a doubling of licensing costs and the fully enforced licenses, the per capita cost is \$15,320. Although fully enforcing licenses increases the number of individuals obtaining a license and increases the quality of teachers slightly, the per capita licensing cost is about 2/3 higher.

10 Conclusions

With the proliferation of studies documenting the importance of teacher to student learning, a number of policies have been considered to improve the quality of the teacher labor force: increasing teacher pay to attract better teachers, instituting merit pay to reward better performing teachers, and changing teacher licensing requirements. A reduced form approach of comparing licensed to non-licensed teacher does not in general identify the effect of formal teacher training on student outcomes or identify the effect of changes in licensing policy. As an alternative, we develop a dynamic teacher labor supply model with an endogenous licensing decision and estimate this model using panel data on a recent cohort of college graduates. The estimated model shows how licensing requirements for teachers affect the aggregate supply of teachers, the sorting of individuals into teaching based on tastes and skills, and the rate of exit from teaching. These results provide an important first step in evaluating licensing policy.

Future research needs to consider how to value different teacher characteristics in the production of student learning. Are good teachers "made" through formal training and experience or are they "born," and we need to create policies that make teaching attractive to them. If recruiting teachers with high general skills is paramount, relaxing licensing requirements would bring more of these teachers into the profession. From the perspective of public schools, this is a relatively inexpensive approach since teacher quality could be raised without salary increases. In addition, lowering requirements to teach would also increase the number of college graduates willing to teach. This may result in downward supply pressure on wages and raise teacher to student ratios.

On the other hand, if the training provided by teacher preparation programs is vital to effective teaching, then maintaining or raising requirements is the better policy.[30] However, the policy of raising licensing standards would lower the number of college graduates willing to teach and would drive away many potentially skilled college students from the profession. But, with higher licensing requirements, the teachers that remain would have more formal training and stay in teaching longer. Understanding these tradeoffs is the first step in developing informed teacher licensing policy.

---

[30] As one recent defender of the current licening system argues, “Just because one is an Albert Einstein does not mean that he or she can successfully teach seventh grade algebra to middle schoolers.” From a 2002 press release from Arthur E. Wise, President of the National Council for Accreditation of Teacher Education.

References

Aaronson, Daniel, Lisa Barrow, and William Sander, "Teachers and Student Achievement in the Chicago Public High Schools," *Federal Reserve Bank of Chicago, working paper*, February 2003.

Angrist, Joshua and Jonathon Guryan, "Teacher Testing, Teacher Education, and Teacher Characteristics," *American Economic Review, Papers and Proceedings*, January 2004.

Bacolod, Marigee, "The Role of Alternative Opportunities in the Female Labor Market in Teacher Supply and Quality: 1940-90," *Working Paper*, November 2001.

Ballou, Dale and Michael Podgursky, "Recruiting Smarter Teachers," *Journal of Human Resources*, Spring 1995, *30* (2), 326–38.

___ **and** ___ , "Teacher Training and Licensure: A Layman's Guide," in Marci Kanstoroom and Jr. Chester E. Finn, eds., *Better Teachers, Better Schools*, Fordham Foundation, July 1999.

Cameron, Stephen V. and Christopher R. Taber, "Evaluation and Identification of Semiparametric Maximum Likelihood Models of Dynamic Discrete Choice," *Working Paper*, November 1998.

Corcoran, Sean P., William N. Evans, and Robert S. Schwab, "Changing Labor Market Opportunities for Women and the Quality of Teachers, 1957-2000," *American Economic Review, Papers and Proceedings*, May 2004, *94* (2), 230–40.

Darling-Hammond, Linda, "Teacher Quality and Student Achievement: A Review of State Policy Analysis," *Educational Policy Analysis Archives*, January 2000, *8* (1), 1–46.

___ **, Barnett Berry, and Amy Thoreson**, "Does Teacher Certification Matter? Evaluating the Evidence," *Educational Evaluation and Policy Analysis*, Spring 2001, *23* (1), 57–77.

___ **, Ruth Chung, and Fred Frelow**, "Variation in Teacher Preparation: How Well Do Different Pathways Prepare Teachers to Teach?," *Journal of Teacher Education*, September/October 2002, *53* (4), 286–302.

Dorsey, Stuart, "The Occupational Licensing Queue," *Journal of Human Resources*, Summer 1980, *15* (3), 424–34.

Ehrenberg, Ronald G. and Dominic J. Brewer, "Do School and Teacher Characteristics Matter? Evidence from High School and Beyond," *Economics of Education Review*, 1994, *13* (1), 1–17.

___ **and** ___ , "Did Teachers' Verbal Ability and Race Matter in the 1960s? Coleman Revisited," *Economics of Education Review*, 1995, *14* (1), 1–21.

Feistritzer, C. Emily and David T. Chester, *Alternative Teacher Certification: A State-by-State Analysis 2002*, Washington D.C.: National Center for Education Information, 2002.

Goldhaber, Dan D. and Dominic J. Brewer, "Does Teacher Certification Matter? High School Teacher Certification Status and Student Achievement," *Educational Evaluation and Policy Analysis*, Summer 2000, *22* (2), 129–45.

___ **and** ___ , "Evaluating the Evidence on Teacher Certification: A Rejoinder," *Educational Evaluation and Policy Analysis*, Spring 2001, *23* (1), 79–86.

Hanushek, Eric A., John F. Kain, and Steven G. Rivkin, "Teachers, Schools, and Academic Achievement," *Econometrica*, March 2005, *73* (2), 417–458.

Heckman, J. and B. Singer, "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica*, March 1984, *52* (2), 271–320.

Heckman, James J. and Guilherme Sedlacek, "Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market," *Journal of Political Economy*, December 1985, *93* (6), 1077–1125.

Keane, Michael P. and Kenneth I. Wolpin, "The Career Decisions of Young Men," *Journal of Political Economy*, 1997, *105* (3), 473–522.

Kleiner, Morris M., "Occupational Licensing," *Journal of Economic Perspectives*, Autumn 2000, *14* (5), 189–202.

___ , *Licensing Occupations: Ensuring Quality or Restricting Competition*, Kalamazoo, Mich.: W.E. Upjohn Institute for Employment Research, 2006.

___ **and Robert T. Kudrle**, "Does Regulation Affect Economic Outcomes? The Case of Dentistry," *Journal of Law and Economics*, October 2000, *43* (2), 547–582.

Kugler, Adrian D. and Robert M. Sauer, "Doctors without Borders? Relicensing Requirements and Negative Selection in the Market for Physicians," *Journal of Labor Economics*, 2005, *23* (3), 437–65.

Leland, Hayne E., "Quacks, Lemons, and Licensing: A Theory of Minimum Quality Standards," *Journal of Political Economy*, December 1979, *87* (6), 1328–46.

Manski, Charles F., "Academic Ability, Earnings, and the Decision to Become a Teacher: Evidence from the National Longitudinal Study of the High School Class of 1972," in David A. Wise, ed., *Public Sector Payrolls*, Chicago: University of Chicago Press, 1987.

Meeting the Highly Qualified Teachers Challenge: The Secretary's Annual Report on Teacher Quality, Washington, D.C.: U.S. Department of Education, 2002.

Mroz, Thomas A., "Discrete Factor Approximations for Use in Simultaneous Equation Models: Estimating the Impact of a Dummy Endogenous Variable on a Continuous Outcome," *Journal of Econometrics*, October 1999, *92* (2), 233–274.

Murnane, Richard J. et al., *Who Will Teach? Policies that Matter*, Cambridge, Mass.: Harvard University Press, 1991.

Rockoff, Jonah E., "The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data," *American Economic Review, Papers and Proceedings*, January 2004.

Shapiro, Carl, "Investment, Moral Hazard, and Occupational Licensing," *Review of Economic Studies*, October 1986, *53* (5), 843–62.

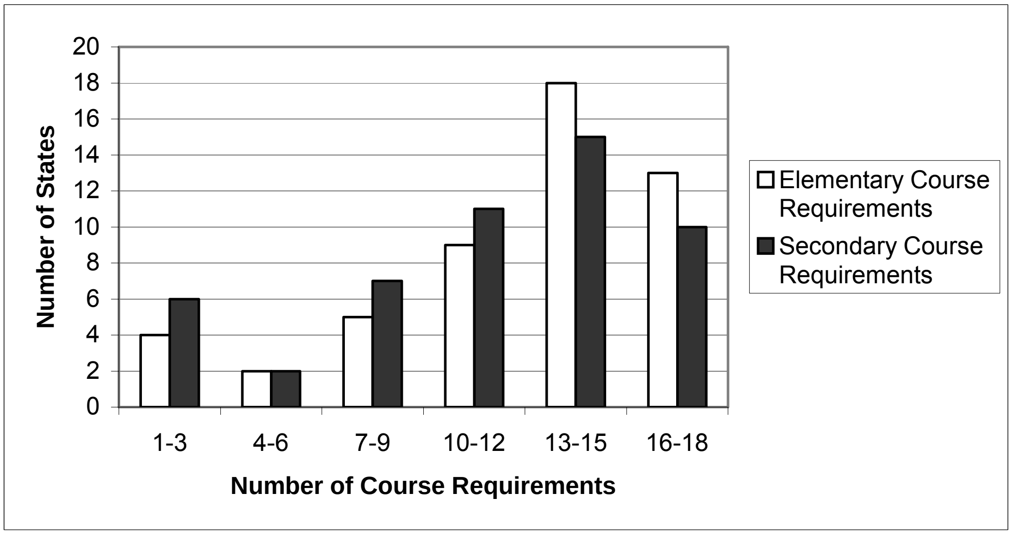
Stinebrickner, Todd R., "A Dynamic Model of Teacher Labor Supply," *Journal of Labor Economics*, January 2001a, *19* (1), 196–230.

———, "Compensation Policies and Teacher Decisions," *International Economic Review*, August 2001b, *42* (3), 751–779.

———, "An Analysis of Occupational Change and Departure from the Labor Force: Evidence of the Reasons that Teachers Leave," *Journal of Human Resources*, Winter 2002, *37* (1), 192–216.

Walsh, Kate, *Teacher Certification Reconsidered: Stumbling for Quality*, Abell Foundation, 2001.

Figure 1: Teacher Licensing Education Course Requirements by State



Notes: States include all 50 states and the District of Columbia.
Source: US Department of Education Title 2 Data Collection, 2001.

Table 1: Descriptive Statistics for Full Sample

Fraction Male	0.44
Fraction White	0.86
1994	
Teachers	0.08
Employed Non-Teaching	0.78
Out of Labor Force	0.14
1997	
Teachers	0.10
Employed Non-Teaching	0.78
Out of Labor Force	0.12
2003	
Teachers	0.10
Employed Non-Teaching	0.77
Out of Labor Force	0.13
Observations	4095

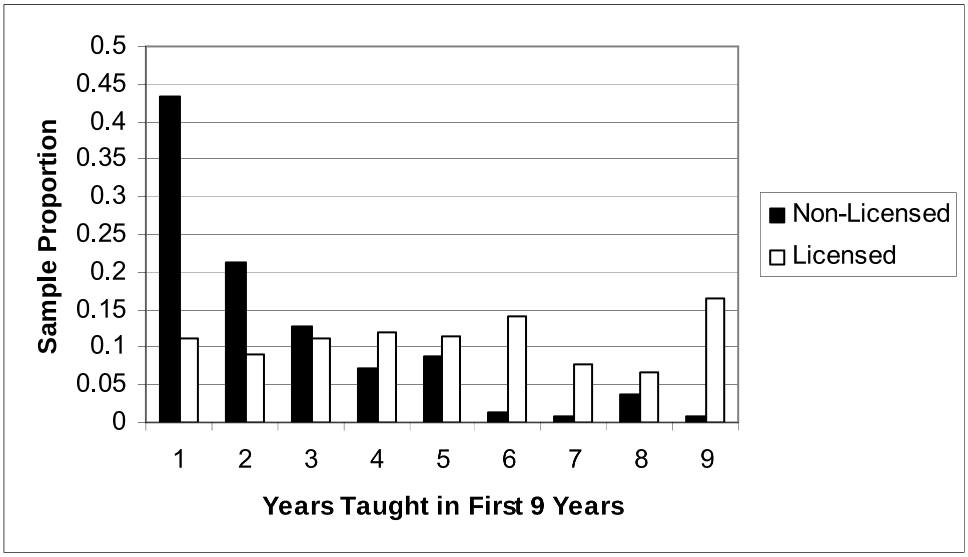
Notes: Statistics calculated using sample weights.
Source: Baccalaureate and Beyond (B&B) Survey of 1992/93 college graduates.

Table 2: Descriptive Statistics by Occupation and Licensing

		1994 (~age 24)			**2003 (~age 33)**	
	All	**Teacher License**	**No License**	**All**	**Teacher License**	**No License**
Teachers						
Number in Sample	372	314	58	441	405	36
Frac. with License	0.85	1	0	0.92	1	0
Frac. Male	0.22	0.20	0.33	0.24	0.22	0.42
Frac. White	0.89	0.92	0.72	0.86	0.88	0.64
Employed Non-Teachers						
Number in Sample	3090	156	2934	3071	188	2883
Frac. with License	0.04	1	0	0.05	1	0
Frac. Male	0.45	0.25	0.46	0.51	0.39	0.52
Frac. White	0.87	0.90	0.87	0.86	0.90	0.85

Notes: The omitted group are respondents out of the labor force in 1994 or 2003. "Frac. with License" refers to individuals with a teaching license by 1994 or 2003. Fractions are calculated using sample weights.
Source: Baccalaureate and Beyond (B&B) Survey of 1992/93 college graduates.

Figure 2: Distribution of Years Teaching (1994-2003) by Licensing Status



Notes: Sample includes only individuals who taught for at least one year. Distribution calculated using sample weights. 1994-2003 period is when the B&B cohort is approximately aged 24-33.
Source: Baccalaureate and Beyond (B&B) Survey of 1992/93 college graduates.

Table 3: Mean SAT Scores by Teaching and Licensing

	All	Licensed	Non-Licensed	Licensed – Non-Licensed
Ever Teach	956.4	949.3	1012.7	-63.4
	(8.56)	(9.09)	(25.81)	(27.26)
Never Teach	1015.7	999.5	1016.2	-16.8
	(3.88)	(18.57)	(3.97)	(18.92)
Difference	-59.2	-50.1	-3.5	
	(9.39)	(20.63)	(25.98)	

Notes: Standard errors in parentheses. All statistics calculated using sample weights.
Source: Baccalaureate and Beyond (B&B) Survey of 1992/93 college graduates.

Table 4: Transition Probabilities and Salaries by Occupation and Licensing

Occup./License in 1994	Teach in 2003	N	Trans. Prob.	Mean Salary 1994	Mean Salary 2003	Difference in Salaries (2003 - 1994)	Mean SAT Score
1) Teachers, Non-Licensed	Yes	20	0.36	24,612	40,703	16,091	885.4
				(2,840)	(2,310)	(3,621)	(57.62)
	No	31	0.53	24,043	44,700	20,658	1021.9
				(1,855)	(4,785)	(4,990)	(36.30)
2) Teachers, Licensed	Yes	193	0.62	25,075	38,157	12,216	907.4
				(629)	(783)	(976)	(12.38)
	No	51	0.17	23,439	49,747	23,279	1037.1
				(929)	(4,230)	(4,202)	(34.58)
3) Non-Teachers, Non-Licensed	Yes	132	0.04	21,428	37,179	15,751	956.6
				(742)	(1,393)	(1,774)	(19.43)
	No	2,429	0.84	28,267	58,482	30,215	1008.8
				(287)	(679)	(647)	(4.55)
4) Non-Teachers, Licensed	Yes	59	0.32	20,859	37,042	16,183	918.5
				(936)	(1,403)	(1,360)	(22.61)
	No	66	0.45	22,171	42,845	20,674	997.4
				(1,174)	(3,464)	(3,613)	(33.63)

Notes: The sample is divided into four groups based on occupation and licensing status in 1994. “Trans. Prob.” is the estimated probability an individual from each of these groups works as a teacher in 2003 (“Teach in 2003”: “Yes”) or works in a non-teaching occupation (“Teach in

2003": "No"). The remaining fraction of the sample reports being out of the labor force. Standard errors are in parentheses. All statistics are calculated using sample weights. Source: Baccalaureate and Beyond (B&B) Survey of 1992/93 college graduates.

Table 5: Parameter Estimates for Common Parameters

	Point Estimate	Standard Error
Teaching Wage Offer		
β_{11}	0.0783	0.0328
β_{12}	0.0012	0.0023
$\sigma_{\varepsilon 1}$	0.1752	0.1000
Non-Teaching Wage Offer		
β_{21}	0.0817	0.0118
β_{22}	0.0016	0.0004
$\sigma_{\varepsilon 2}$	0.3556	0.1125
Non-Pecuniary Benefits in Teaching		
$\sigma_{\eta 1}$	0.9965	0.2790
Licensing Costs		
κ_0	0.2903	0.1151
κ_1	0.0209	0.0342
Arrival Rate of Teaching Job Offers for Non-Licensed Agents		
γ	0.8686	0.0820

Table 6: Heterogeneity Distribution Parameters

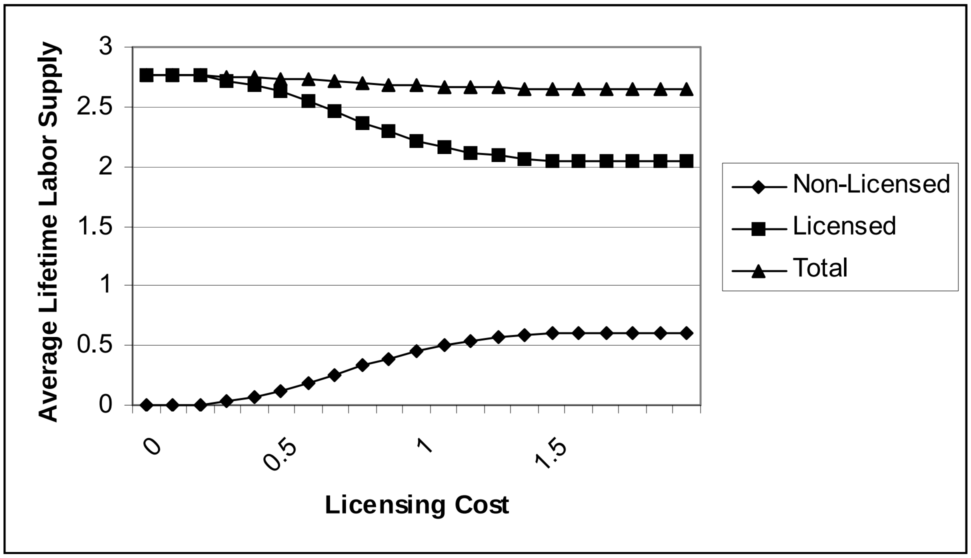
	π_1	τ_1	π_2	τ_2	Probability
Type 1	-1.6232	-1.0245	-1.3908	-0.3528	0.1140
Type 2	-1.3996	-0.8190	-1.5894	-2.8720	0.0104
Type 3	-1.3936	0.0107	-1.3366	-0.7684	0.1265
Type 4	-1.3911	0.6636	-1.3062	-3.5389	0.7491

Table 7: Within Sample Fit

	Actual	Predicted
Proportion with License by 1994 (~age 24)	0.109	0.124
Proportion with License by 1997 (~age 27)	0.140	0.137
Proportion with License by 2003 (~age 33)	0.159	0.141
Teaching with License in 1994	0.065	0.068
Teaching without License in 1994	0.011	0.012
Employed Non-Teaching with a License in 1994	0.034	0.016
Employed Non-Teaching without a License in 1994	0.749	0.742
Teaching with License in 2003	0.089	0.067
Teaching without License in 2003	0.008	0.014
Employed Non-Teaching with a License in 2003	0.042	0.032
Employed Non-Teaching without a License in 2003	0.728	0.740
Mean Wage for Teachers with License in 1994	0.247	0.257
Mean Wage for Teachers without License in 1994	0.236	0.277
Mean Wage for Non-Teachers with License in 1994	0.222	0.224
Mean Wage for Non-Teachers without License in 1994	0.277	0.270
Mean Wage for Teachers with License in 2003	0.387	0.354
Mean Wage for Teachers without License in 2003	0.319	0.291
Mean Wage for Non-Teachers with License in 2003	0.451	0.323
Mean Wage for Non-Teachers without License in 2003	0.590	0.459
Mean Wage for All Teachers Ages 51-55	0.602	0.596
Mean Wage for All Non-Teachers Ages 51-55	0.673	0.721

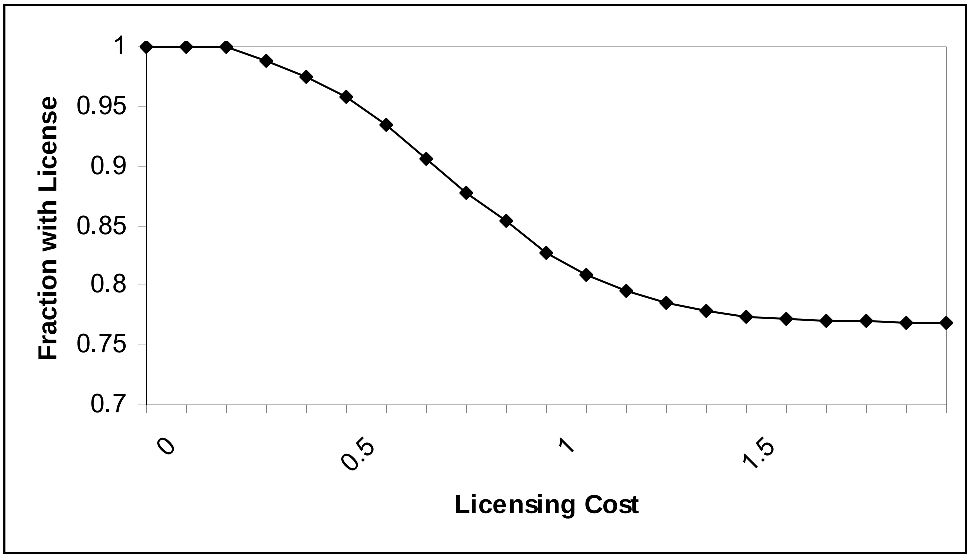
Notes: Predicted is from data simulation using estimated parameters. All actual statistics except last two rows are from the Baccalaureate and Beyond (B&B) Survey of 1992/93 college graduates. Mean Wage for Ages 51-55 is from 2003 March CPS. All actual statistics calculated using sample weights.

Figure 3: Policy Simulation: Teacher Labor Supply



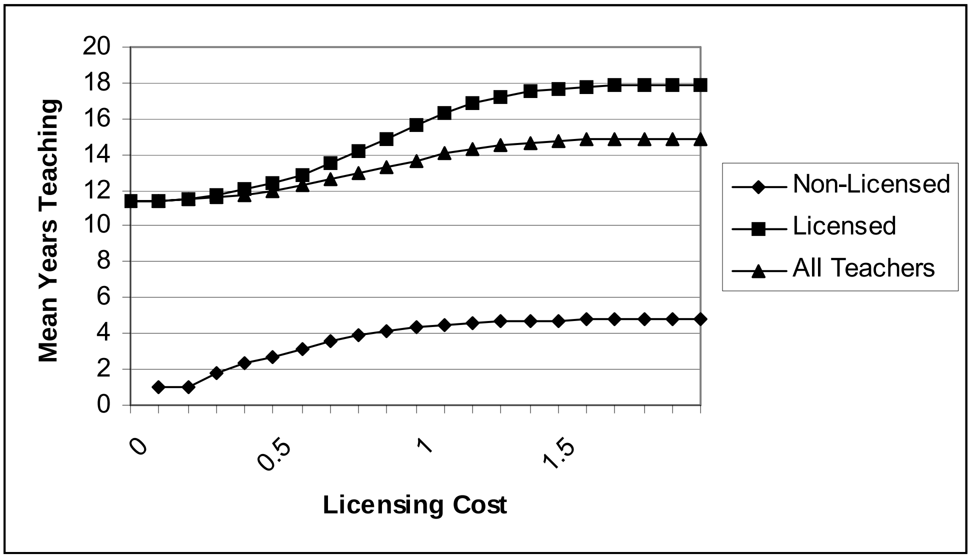
Notes: Licensing cost of 0 is free entry (no requirements), 1 is current policy, and 1.5 is 50 percent larger licensing costs than under the current policy.

Figure 4: Policy Simulation: Fraction of Teacher Labor Supply with a License



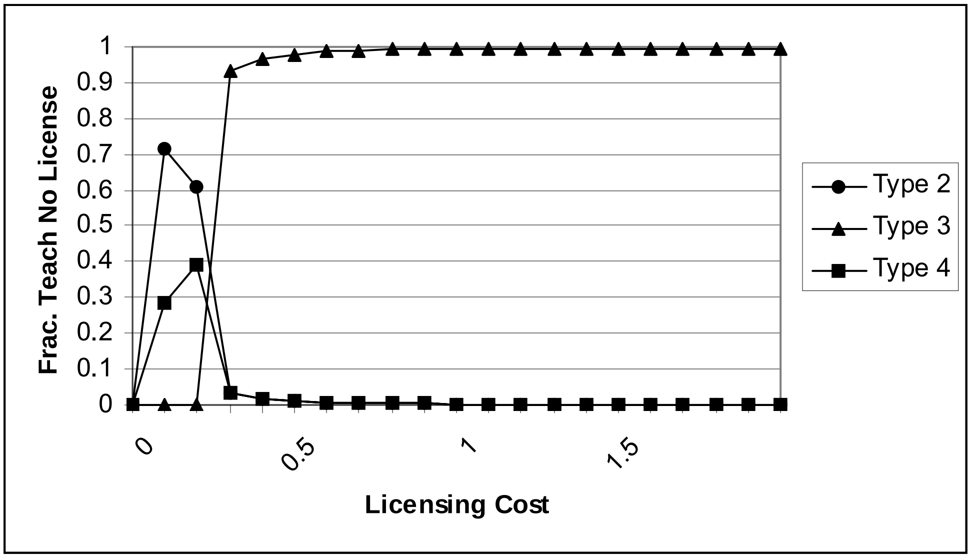
Notes: Licensing cost of 0 is free entry (no requirements), 1 is current policy, and 1.5 is 50 percent larger licensing costs than under the current policy.

Figure 5: Policy Simulation: Attachment to Teaching



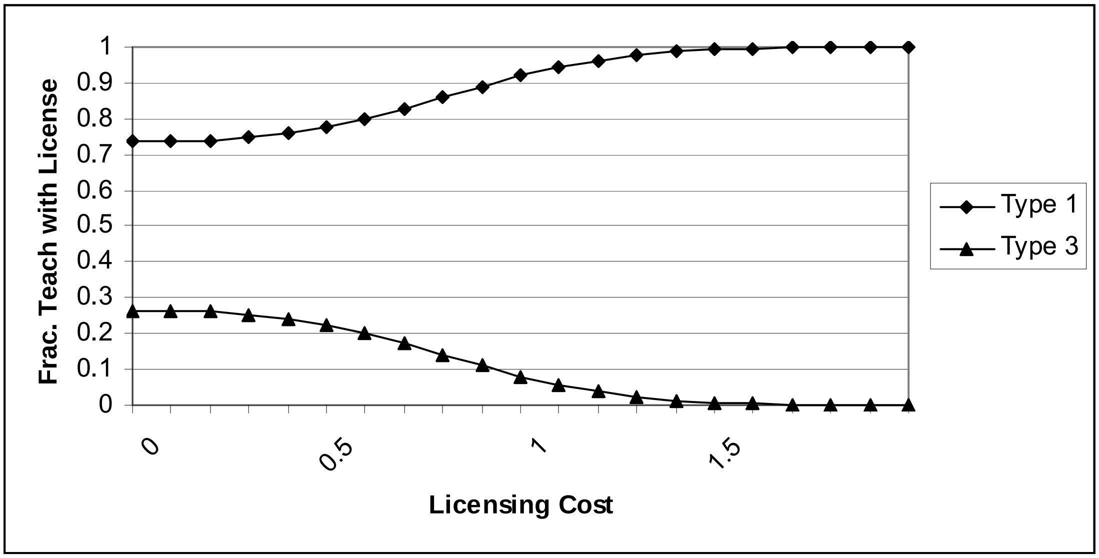
Notes: Licensing cost of 0 is free entry (no requirements), 1 is current policy, and 1.5 is 50 percent larger licensing costs than under the current policy.

Figure 6: Policy Simulation: Proportion of Each Type Among Non-Licensed Teachers



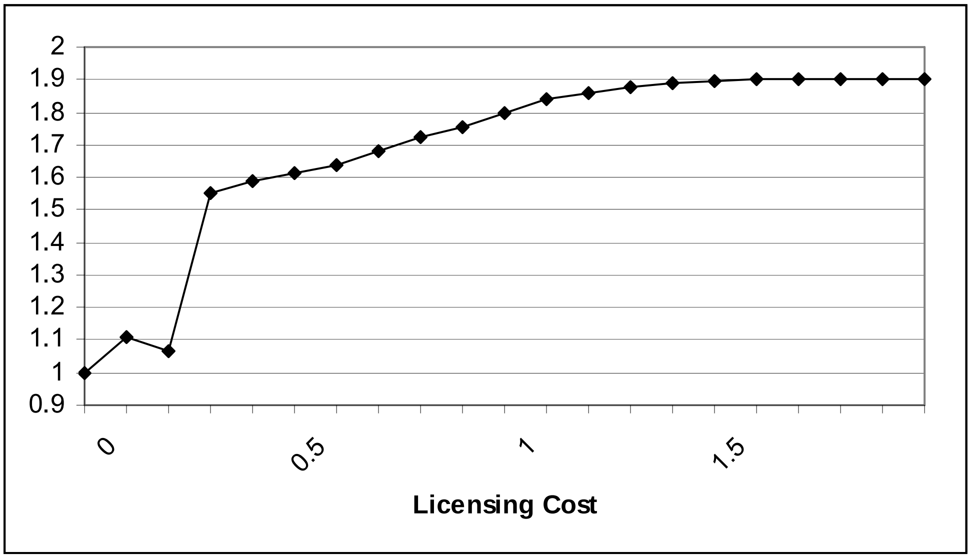
Notes: In the simulations, there are no Type 1 individuals who choose to teach without a license for any of the licensing costs. Licensing cost of 0 is free entry (no requirements), 1 is current policy, and 1.5 is 50 percent larger licensing costs than under the current policy.

Figure 7: Policy Simulation: Proportion of Each Type among Licensed Teachers



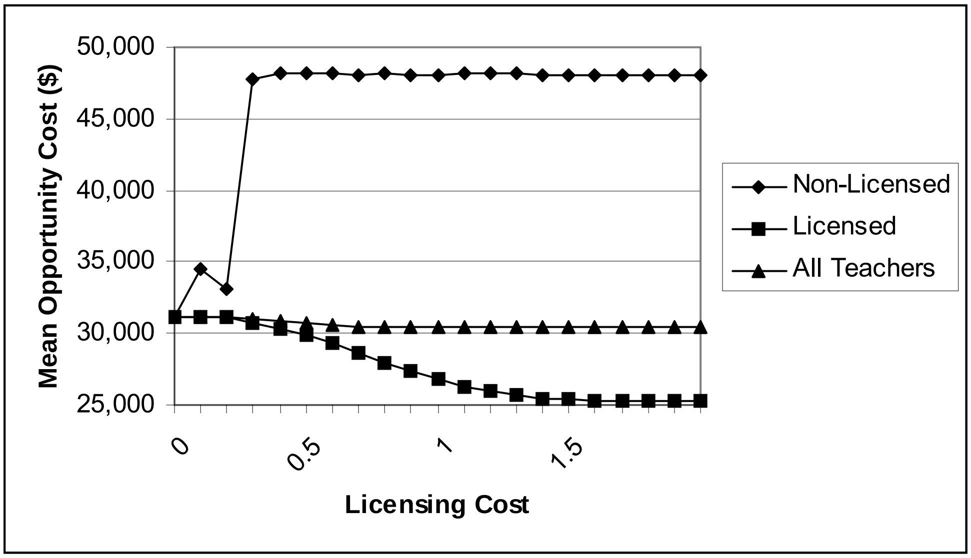
Notes: In the simulations, the fraction of Type 2 and Type 4 individuals choosing to teach with a license is less than 1 percent for all licensing costs. Licensing cost of 0 is free entry (no requirements), 1 is current policy, and 1.5 is 50 percent larger licensing costs than under the current policy.

Figure 8: Policy Simulation: Ratio of Non-Teaching Opportunity Cost for Employed Teachers (Licensed vs. Non-Licensed)



Notes: Licensing cost of 0 is free entry (no requirements), 1 is current policy, and 1.5 is 50 percent larger licensing costs than under the current policy.

Figure 9: Policy Simulation: Opportunity Cost of Teaching



Notes: Licensing cost of 0 is free entry (no requirements), 1 is current policy, and 1.5 is 50 percent larger licensing costs than under the current policy.

Table 8: Summary of Policy Simulations

	Current Policy	Free Entry	Double Licensing Costs	Fully Enforced	Fully Enforced and Double Licensing Costs
Teacher Labor Supply (No License)	0.46	0	0.61	0	0
Teacher Labor Supply (License)	2.214	2.760	2.038	2.758	2.732
Teacher Labor Supply (All)	2.675	2.760	2.649	2.758	2.732
Frac of Teachers w/ License	0.83	1	0.77	1	1
Mean Periods Teach (No License)	4.34	NA	4.77	NA	NA
Mean Periods Teach (License)	15.62	11.36	17.88	11.46	11.63
Mean Opportunity Cost (No License)	$48,076	NA	48,032	NA	NA
Mean Opportunity Cost (License)	$26,750	31,083	25,229	31,081	30,873
Mean Opportunity Cost (All)	$30,423	31,083	30,489	31,081	30,873
Per Capita Total Licensing Cost	$4,220	0	6,620	7,010	15,320