The Allocation of Teaching Talent and Human Capital Accumulation

Simeon D. Alder[*]   Yulia Dudareva[†]   Ananth Seshadri[‡]

March 30, 2022

PRELIMINARY AND INCOMPLETE!

Abstract

The educational landscape in the U.S. has gone through major changes since the end of World War II. Real expenditures per student have risen from approximately $2,100 to more than $10,000 by the turn of the century and more than $12,000 since the Great Recession. At the same time, the student-teacher ratio has fallen from a national average of almost 27 in 1955 to 16 by the 2010s. Yet despite the rise in expenditures and the reduction in class sizes, educational outcomes in the U.S. don't compare very favorably with countries at similar income levels. One aspect of U.S. education that has not garnered a lot of attention until fairly recently is occupational choice. We add an education sector to an otherwise standard Hsieh et al. (2018)-style model to explore the extent to which changes in career opportunities in other occupations affect the selection of workers into teaching careers. In our model, changes in the allocation of teaching talent have implications for the evolution of class size as well as quality of instruction and hence the accumulation of human capital during the workers' formative years. This gives rise to a trade-off between static and dynamic efficiency, which we quantify by way of a structural model. In order to discipline the parameterization of the model, we compare model-generated moments to their empirical

[*]University of Wisconsin - Madison
[†]University of Wisconsin - Madison
[‡]University of Wisconsin - Madison

1

counterparts in three longitudinal surveys: Project TALENT, the NLSY79, and NLSY97. Project TALENT surveys 377,000 U.S. high school students in 1960 and follows them for more than a decade. Together with the two NLSY's, it provides a detailed account of, among many others, the educational trajectories and career paths of young American adults over the course of more than four decades.

1   Introduction

The educational landscape in the U.S. has gone through major changes since the end of World War II. Real expenditure per student has risen from approximately $2,100 to more than $10,000 by the turn of the century and more than $12,000 since the Great Recession. At the same time, the student-teacher ratio has fallen from a national average of almost 27 in 1955 to 16 by the 2010s. Yet despite the rise in expenditures and the reduction in class sizes, educational outcomes in the U.S. don't compare very favorably with countries at similar income levels.

One aspect of U.S. education that has not garnered a lot of attention until fairly recently is occupational choice. We add an education sector to an otherwise standard Hsieh et al. (2018)-style model to explore the extent to which changes in career opportunities in other occupations affect the selection of workers into teaching careers. In our model, teaching is distinct from other occupations in the economy since it produces human capital embodied in students rather than a consumption good or service. As a result, changes in the allocation of teaching talent have implications for the evolution of class size as well as quality of instruction and hence the accumulation of human capital during the workers' formative years.

While the model has implications for the stationary static allocation of talent across occupations, we emphasize the dynamic effects driven by human capital accumulation in primary and secondary education in order to explore the deeper causes of America's education *malaise*.

Section 2 introduces the model. Section 4 reviews the data and we calibrate the model to match salient empirical moments in section 8. Section 9 concludes.

2   Model

2.1   Population and Endowments

The model is populated by a constant measure M of heterogeneous individuals. Each individual is born with an exogenous N-dimensional vector of skills $\vec{a}$, where the nth element characterizes her innate ability in occupation n. This vector of abilities is

drawn from a known joint distribution $F_a(\vec{a})$.

Individual agents live for two periods. They attend school when they are children ("young") and they work when they are adults ("old"). For simplicity, we abstract from retirement and instead assume that everyone exits the economy when they reach the end of adulthood.

Each cohort has measure $\frac{M}{2}$ and therefore the age distribution of the population is constant over time.

2.2 Technologies

Children have access to a teaching technology that produces occupation-specific human capital with four inputs: their own ability a^n, a time investment s, an investment in terms of the final good e, and the teacher's human capital h^T:

$$h^{n'}(a) = \left(h^T\right)^\beta a^\alpha s\left(a\right)^\phi e(a^n)^\eta N(h^T, \widetilde{H}^T)^{-\sigma},$$

where $\widetilde{H}^T = \int_0^\infty \left(h^T\left(a\right)\right)^{\frac{\beta}{\sigma}} f^T(a) da$ is an aggregate measure of human capital across all current teachers and $N(h^T, \widetilde{H}^T)$ denotes the class size of a teacher with human capital h^T. The density $f^T(a)$ characterizes the ability distribution of the current cohort of teachers.

Without loss of generality, occupation 1 is *teaching*. In this economy, teachers are distinct from other occupations in that their own human capital is an input into a production function that enables their students to accumulate human capital. In contrast, the human capital of workers in other occupations is used to produce units of a homogenous final (consumption) good. This kind of distinction is salient since we are assuming that the economy is populated by overlapping generations of workers who accumulate human capital when young, which then pays off when they are of working age (or "old").

2.3 Abilities and Distortions

As in Roy (1951), workers make occupational choices based on their full range of idiosyncratic abilities $\vec{a}$ and the corresponding "payoffs." To this basic mechanism we

add Hsieh et al. (2018)-style forces that distort the allocation of skill across occupations and a human capital formation technology with time, good, and teacher inputs.

There are two types of distortions: educational barriers labeled τ^e and labor market discrimination labeled τ^w. These distortions are specific to groups of individuals and we assume that there are G different groups and we will label them with superscripts $g \in \{1, \ldots, G\}$. The distortions evolve according to Markov chains with transition matrices $P_{\tau^{e\prime}, \tau^e} = \mathbb{P}^{e,n,g}\left(\tau_{t+1}^e = \tau^{e\prime} | \tau_t^e = \tau^e\right)$ and $P_{\tau^{w\prime}, \tau^w} = \mathbb{P}^{w,n,g}\left(\tau_{t+1}^w = \tau^{w\prime} | \tau_t^w = \tau^w\right)$, respectively, for each occupation n and group g.

The timing convention is similar to Hsieh et al. (2018). In the first period, prospective workers – students, that is – make human capital investments. Since there is no uncertainty, students have perfect foresight and their investment decision will depend on their fully anticipated occupational choice in period 2. They retire at the end of their working period and are replaced by a new cohort of students with measure $\frac{M}{2}$. The environment is stationary in the sense that the members of each cohort draw their ability vectors from the same distribution F_a.

The worker's utility depends on her consumption when she is old and her time devoted to human capital investment when she is young:

$$U_t = \max_{C_{t+1}, s_t, e_t} \ln C_{t+1} + \ln\left(1 - s_t\right) \tag{1}$$

where

$$C_{t+1} = \max\left\{(1 - \tau^{\omega,g,n})\omega_{t+1}(h_{t+1}^n) - e_t\right\}_{n \in N}, \tag{2}$$

$$h_{t+1}^n = g(h_t^T, a^n)s_t(a^{\hat{n}})^\phi e_t\left(a^{\hat{n}}\right)^\eta \left(N_t(h_t^T)\right)^{-\sigma}, \tag{3}$$

$$g(h_t^T, a^n) = (h_t^T)^\beta \left(a^n\right)^\alpha, \tag{4}$$

Need to add marginal tax rate t!

C_{t+1} denotes consumption of the homogeneous final good when old, $\hat{n} \in N$ indexes the workers occupational choice when young, $s_t(\hat{n})$ is the time allocated to human capital formation, and $e_t(\hat{n})$ are units of the final good invested in human capital formation when workers are still students. These investments are occupation-specific (CAN WE RELAX THIS ASSUMPTION?) and indexed by the superscript n.[1] The

---

[1]Nothing prevents students from making investments in human capital in multiple occupations.

periodic time endowment is set to unity. $1 - s_t$ is leisure time when young and labor is supplied inelastically when individuals reach working age. $\tau^{\omega, g, n}$ is a labor market distortion (or wedge) that is specific to workers in group g in occupation n.

A worker's vector of occupation-specific abilities is drawn from a multivariate distribution with c.d.f.

The joint distribution of $\vec{a}$ is a multivariate Fréchet with c.d.f. $F(a_1, \ldots, a_N)$. The draws from this distribution are independent across workers.

The production technologies in occupations $2, \ldots, N$ are linear in the workers human capital. A worker with human capital h in occupation n produces

$$y^n = h^n A^n \tag{5}$$

units of output.

Since industry output is homogeneous, aggregate output is given by

$$Y = \sum_{n=2}^{N} \int \text{To Be Completed!} \tag{6}$$

Teaching is distinct from other occupations since it "transforms" students with skill endowment $\vec{a}$ into workers with human capital $\vec{h}$ based on equation (3). The acquisition of human capital depends on the teacher's input h^T, the student's talent $\vec{a}$, the student's time and good investments s and e, as well as class size $N(h^T)$, which we allow to vary with the teacher's h^T.

2.4 Occupational Choice

Individuals will choose the occupation that delivers the highest lifetime utility. Since human capital takes "time to build", the return to educational good and time investments is zero when working.

Since we are mainly interested in the choice between a teaching career and other occupations, we will first characterize the problem for the special $N = 2$ case and label the occupations with T (teaching) and O (other). In a second step, we can review

---

Since there is no uncertainty in future payoffs they will, however, select only one in equilibrium.

4

how changes in discrimination or educational barriers in these "other" occupations affect the selection and human capital composition of teachers over time.

Equation (3) highlights the distinctive feature of teaching compared to all other occupations: the elasticity of output with respect to the teacher's human capital (h^T) is β whereas production is linear in human capital in O, i.e. the output elasticity is 1. Due to the complementarity between student ability and teacher type, the efficient allocation would match higher-ability students with more skilled teachers.

2.4.1 Teachers and Class Size

This effect, however, is tempered by the teacher's span-of-control. $N(\cdot)$ denotes the class size and the coefficient $-\sigma$ captures the stylized fact that in larger classes teachers can pay less attention to individual students. Here, we allow class size to depend on h^T and we are focusing on $N(h^T)$ such that – for given e – students are indifferent between teachers with different levels of human capital, say h^T and $h^{T'}$:

$$N(h^{T'}) = \left(\frac{h^{T'}}{h^T}\right)^{\frac{\beta}{\sigma}} N(h^T). \tag{7}$$

2.4.2 Student-Teacher Matching

While this guarantees that students are indifferent between different teachers, it may be the case that better teachers prefer to work with higher ability students. This will be the case, for instance, if students make human capital investments in terms of goods as a function of (1) their own ability, (2) anticipated occupation, and (3) teacher quality.

For simplicity, we assume that students make investments as a function of a and occupation, but not h^T. The joint assumption $e(a)$ and $N(h^T)$ implies that the student-teacher matching is random since in this specification the human capital production function is modular in h^T and a.[2]

---

[2] The anecdotal evidence suggests that student-teacher assignments don't involve systematic sorting at the school level. There is, however, some evidence in support of sorting between families with highly educated parents and high-ability teachers into better funded schools or school districts. Identifying the empirically relevant extent of this type of sorting is working in progress and future versions of our model will allow for it, if salient.

2.5 Values

The structure we impose on class size and student-teacher matching simplifies the characterization of an equilibrium considerably. While idiosyncratic heterogeneity still matters, the aggregate state can be summarized by a single moment of the distribution of human capital among current teachers (which we discuss in more detail below). Thanks to this feature, the dimensionality of the dynamic program associated with this economy is finite and we can characterize stationary equilibria as well as full transition paths sharply.

The idiosyncratic state consists of the N-dimensional vector of abilities $\vec{a}$. Recall that we assume $N = 2$ for the time being, where $n = 1$ denotes teaching and $n = 2$ lumps together all other occupations. Instead of 1 and 2, we'll label the occupations by T (for teaching) and O (for other).

Students solve the following maximization problem which entails a discrete occupational choice as well as continuous human capital investments in terms of time and final goods. Formally, they solve:

$$V^g(a^T, a^O, \widetilde{H}^T) = \max_{\{s^O, s^T, e^O, e^T\}} \left\{ V^{O,g}(a^O, \widetilde{H}^T), V^{T,g}(a^T, \widetilde{H}^T) \right\}$$

where

$$
\begin{aligned}
V^{O,g}(a^O, \widetilde{H}^T) = {} & \ln\left(1 - s^O\left(a^O, \widetilde{H}^T\right)\right) \\
& + \mu \ln\left[h'^O A^O (1 - \tau'^{O,g,w}) \right. \\
& \left. - e^O(a^O, \widetilde{H}^T)(1 + \tau^{O,g,e}) - T\right], \\
V^{T,g}(a^T, \widetilde{H}^T) = {} & \ln\left(1 - s^T\left(a^T, \widetilde{H}^T\right)\right) \\
& + \mu \ln\left[\omega'(h'^T, \widetilde{H'}^T)(1 - \tau'^{T,g,w}) \right. \\
& \left. - e^T(a^T, \widetilde{H}^T)(1 + \tau^{T,g,e}) - T\right]
\end{aligned}
$$

To make further progress, we need to characterize an occupational choice threshold along which a student with abilities (a^T, a^O) is indifferent between teaching and production work. More formally, the threshold is a function

$$a^{T*}(a^O) = \bar{a}(a^O, \widetilde{H}^T)$$

such that

$$V^O(a^O, \widetilde{H}^T) = V^T\left(a^{T*}(a^O), \widetilde{H}^T\right), \text{ for all } a^O \in (0, \infty) \tag{8}$$

Unless it's necessary, we will omit the aggregate state argument $\widetilde{H}^T$ from $\bar{a}$ in order to keep the notation clean and tidy.

The optimization problem is subject to the following constraints and laws of motion:

$$T = \int_0^\infty \omega\left(h^{T'}(a)\right) f'^T(a) da \tag{9}$$

$$f^T(a) = \int_0^{\bar{a}^{-1}(a)} f(a, b) db \tag{10}$$

$$f^O(b) = \int_0^{\bar{a}(b)} f(a, b) da \tag{11}$$

Teachers are public sector employees and their compensation is funded by tax revenues. Equation (9) is the government's balanced budget requirement, which is a salient feature in the U.S. context. Equations (10) and (11) are the marginal densities of teachers and production workers after taking into account the endogenous occupational choice in (8).

The aggregate laws of motion for human capital in teaching ($\widetilde{H}^T$) and "other" (H^O) are:

$$H'^O = \int_0^\infty \left(\frac{2\widetilde{H}^T}{M}\right)^\sigma a^\alpha s^O\left(a, \widetilde{H}^T\right)^\phi e^O(a, \widetilde{H}^T)^\eta f^O(a) da \tag{12}$$

$$\widetilde{H'}^T = \int_0^\infty \left(\left(\frac{2\widetilde{H}^T}{M}\right)^\sigma a^\alpha s^T\left(a, \widetilde{H}^T\right)^\phi e^T(a, \widetilde{H}^T)^\eta\right)^{\frac{\beta}{\sigma}} f^T(a) da \tag{13}$$

Equation (12) is the aggregation over h^O of the current cohort of students while equation (13) is the $\left(\frac{\beta}{\sigma}\right)^{\text{th}}$ moment of the distribution of h^T of the current cohort of students.

The only restriction on the teachers' wage profile is that it's proportional to her class size. We revisit this aspect in section 5 where we impose some additional structure on this profile. While ω is linear in class size, it is non-linear in the teacher's human capital and this will turn out to be important when we analyze the effects associated

with changes in educational barriers and labor market discrimination over time.[3]

3 Equilibrium

Let's first characterize the optimal human capital investment for a prospective teacher with ability a.[4]

The F.O.C.s for s^T and e^T, respectively, after a few steps of algebra are:

$$(1 - \tau^e)e^t \left(\frac{\sigma}{\beta\eta} - 1 \right) - T = \left(1 - s^T\right)\mu(1 - \tau^w)\frac{\partial\omega'}{\partial s^T} \tag{14}$$

$$(1 - \tau^w)\frac{\partial\omega'}{\partial e^T} = 1 - \tau^e, \tag{15}$$

where

$$\frac{\partial\omega'}{\partial s^T} = \frac{\beta\phi}{\sigma}\lambda' \left(\frac{M}{2H^T} \right)^{-\beta} \frac{M}{2H'^T} \left[a^\alpha(s^T)^\phi(e^T)^\eta \right]^{\frac{\beta}{\sigma}} \frac{1}{s^T} \tag{16}$$

$$\frac{\partial\omega'}{\partial e^T} = \frac{\beta\eta}{\sigma}\lambda' \left(\frac{M}{2H^T} \right)^{-\beta} \frac{M}{2H'^T} \left[a^\alpha(s^T)^\phi(e^T)^\eta \right]^{\frac{\beta}{\sigma}} \frac{1}{e^T} \tag{17}$$

The F.O.C.s for prospective production workers are analogues of (14) and (15) with the simplification that the wage profile in production (as opposed to teaching) is linear in h^O.

In the special case with $\beta = \sigma$, the model reverts to a standard occupational choice model à la *Roy* and the wage profile is linear in human capital on *both* occupations.

Proposition 1 If $\beta = \sigma$, then $\frac{d\omega}{dh}$ and $\frac{h^*(h^O)}{h^O}$ are constant, i.e. the occupational cutoff satisfies $h^* = \frac{h^T}{h^O}$ for all h^O.

---

[3]The key implication is that, in contrast to a standard Roy model, occupational choice depends on an agent's comparative as well as absolute advantage.

[4]A stickler might object that it should really be a^T, but we're dropping the superscript to economize on notation whenever possible.

4 Data

We parameterize the model with data from longitudinal surveys that cover American youth at three different times after World War II: Project TALENT, NLSY 79, and NLSY 97. While the two NSLY surveys have been used extensively in the literature, Project TALENT is somewhat less known and we present it in some detail below.

4.1 Project TALENT

Project TALENT surveyed approximately 377,000 students enrolled in high school in 1960. The vast majority of these students were in the 14-18 year age range with a few outliers outside this bracket (on both sides). The sample of surveyed students was representative and accounted for 5 percent of the high school population that year. The project was funded by the United States Office of Education (the precursor of the Department of Education) and its goal was to collect high quality data on the achievements, aptitudes, and interests of high school students in the United States, and to examine whether, and how, these data predicted future outcomes in terms of educational attainment, career, and wellbeing.

In the initial wave of surveys, students were subjected to a two-day battery of aptitude, ability, and achievement tests. In addition, students answered questions about dispositional traits, their interests (with respect to occupations and activities), as well as personal and family information. In total, the surveyed students answered more than one thousand questions.

Follow-up surveys were then carried out one, five, and eleven years after the students' expected high school graduation. For instance, students who were in eleventh grade in 1960 were expected to graduate in 1961 and were contacted again in 1962, 1966, and 1972. Ninth graders were contacted in 1964, 1968, and 1975. A lack of federal funding stopped the project during the planning stage for a 17-year follow-up, which never took place.

51 percent of the original participants participated in the 1-year post-graduation surveys. The five and eleven-year follow-ups successfully contacted 35 and 25 percent of the original participants, respectively. One concern with attrition rates of this magnitude is that the survey sample of waves two through five is subject to selection along

one or several attributes.

In Table 1 we summarize the descriptive statistics by gender, race, parental education, and parental economic status (all in 1960). The racial composition of the sample exhibits some anomalies. The racial classification in the 1960 base year is problematic since almost 60 percent of the sample are in the "other", "unknown", or "conflicting" category. One conceivable explanation for this unusual pattern as well as the *rise* of black respondents when the sample size is *shrinking* is a retroactive racial reclassification of all the participants in the 5-year post survey.[5]

	Base Year	1-Year Post	5-Year Post	11-Year Post
Sample Size	368,938	188,138	128,163	92,962
Black	6,488	4,003	4,856	3,026
(% of sample size)	(1.8%)	(2.1%)	(3.8%)	(3.3%)
Other / Unknown / Conflicting	215,786	73,668	3,739	3,758
(% of sample size)	(58.5%)	(39.1%)	(2.9%)	(4.0%)
Female	184,912	98,183	64,841	47,417
(% of sample size)	(50.1%)	(52.2%)	(50.6%)	(51.0%)
Parents: High School Degree	TBC	TBC	TBC	TBC
(% of sample size)				
Parents: College Degree	TBC	TBC	TBC	TBC
(% of sample size)				
Parents: High Income	TBC	TBC	TBC	TBC
(% of sample size)				

Table 1: Project TALENT Demographic Composition

The rationale for using panel data from Project TALENT and the two NLSYs is that all three surveys allow us to connect scores from aptitude tests to labor market outcomes. Project TALENT contains a number of tests administered in 1960 when the respondents were attending high school, though in different grades.[6] We mainly use composite scores built from a battery of underlying tests measuring verbal and math-

---

[5]The classification of respondents into race categories in the base year suffers from inconsistencies (e.g. self-reported vs. pollster's classification). Moreover, the sample proportion of black high school students differs from that reported in the 1960 Census.

[6]Roughly one quarter of the sample was in each high school grade (i.e. 9 through 12) in the baseline survey.

ematical aptitude. The former includes literature information, vocabulary, and English. The latter contains math information, arithmetic reasoning, introductory high school math, and advanced high school math. Table 2 shows summary statistics for the 88,330 individuals who participated in the 11-year follow-up survey. We focus on the 11-year *post* wave since it contains the bulk of occupational information at a time when the respondents are approximately thirty years old.[7]

Panel A: 5-Year Post			
	Teachers	Other	Not in Labor Force
Female	9,392	32,337	17,974
(% of subsample)	(74.01%)	(39.06%)	(79.46%)
Math composite mean	95.31	74.10	71.94
(st.dev.)	(31.78)	(33.43)	(33.19)
Verbal composite mean	126.26	109.63	112.67
(st.dev.)	(16.44)	(22.09)	(21.08)
Observations	12,690	82,798	22,620
(% of total sample)	(10.74%)	(70.10%)	(19.15%)

Panel B: 11-Year Post			
	Teachers	Other	Not in Labor Force
Female	3,747	19,527	21,983
(% of subsample)	(74.05%)	(38.95%)	(79.48%)
Math composite mean	90.72	75.19	69.85
(st.dev.)	(33.56)	(34.17)	(30.24)
Verbal composite mean	123.28	108.92	112.82
(st.dev.)	(18.19)	(22.67)	(20.36)
Observations	6,217	58,835	23,278
(% of total sample)	(7.04%)	(66.61%)	(26.35%)

Table 2: Project TALENT Test Scores by Occupation

Approximately three quarters of teachers in the sample are female, in the 5-year post

---

[7]The 5-year post-graduation survey contains occupational information but we have not made use of it yet.

as well as the 11-year post surveys. In non-teaching occupations the share of women is 39 percent and women account for 80 percent of respondents who report that they are not in the labor force. On average, teachers have higher mathematics and verbal test scores than respondents in other occupations and those who are not in the labor force. Together, these stylized facts are consistent with the hypothesis that high-ability women are disproportionately choosing teaching jobs and careers and we will return to this point below.

If the labor market discriminated more against certain groups (gender, race,...) in non-teaching occupations in 1960, we expect these groups to favor teaching over other occupations or to stay out of the labor force altogether. "Not in the labor force" is more a reflection of the absolute level of labor market discrimination against particular group while the occupational choice between teaching or other occupations reflects differences between group-specific barriers. To the extent that occupations require specific human capital investments we allow for educational barriers to affect occupational choices separately in the model. If both of these barriers declined after 1960, we would expect a narrowing of the educational achievement gap and more talented females should move from teaching into other occupations. This reallocation would lower the average quality of teachers and, therefore, we expect to observe more significant differences in the ability distributions between teachers and other occupations in cohorts who entered the labor market in the 60s and 70s (i.e. our Project TALENT population) compared to younger individuals (i.e. the NLSY79 and NLSY97 respondents).

Figure 1 shows the distribution of test scores for math and verbal composite by occupations in Project TALENT. The distribution of math scores in panel 1a is skewed to the right for "other" compared to the distribution of math scores among teachers. Panel 1b shows a similar pattern in terms of relative skewness. The distribution of scores of respondents in non-teaching occupations and those who report not being in the labor force are quite similar. Overall, individuals in teaching occupations tend to have higher scores than non-teachers. One possible explanation for these differences is that individuals select into teaching along a non-skill dimension and an obvious selection candidate is gender. Approximately three quarters of all teachers are female (see the first row in the two panels of Figure 2). If women exhibited a distinct distribution of scores this may explain the discrepancy between teaching and other occupations. The data, however, do not suggest that women are innately more or
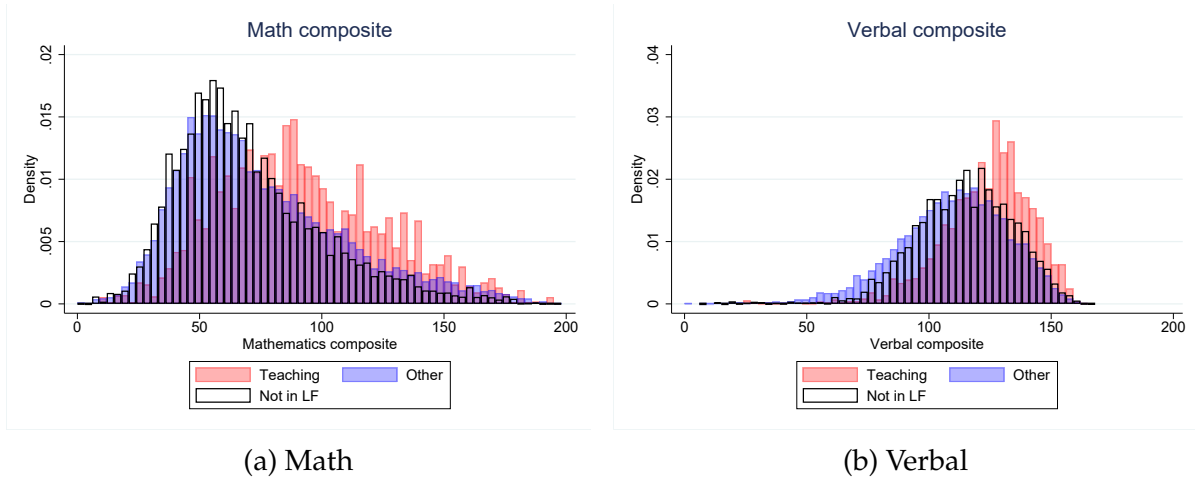
(a) Math   (b) Verbal

Figure 1: Composite Scores by Occupation in Project TALENT

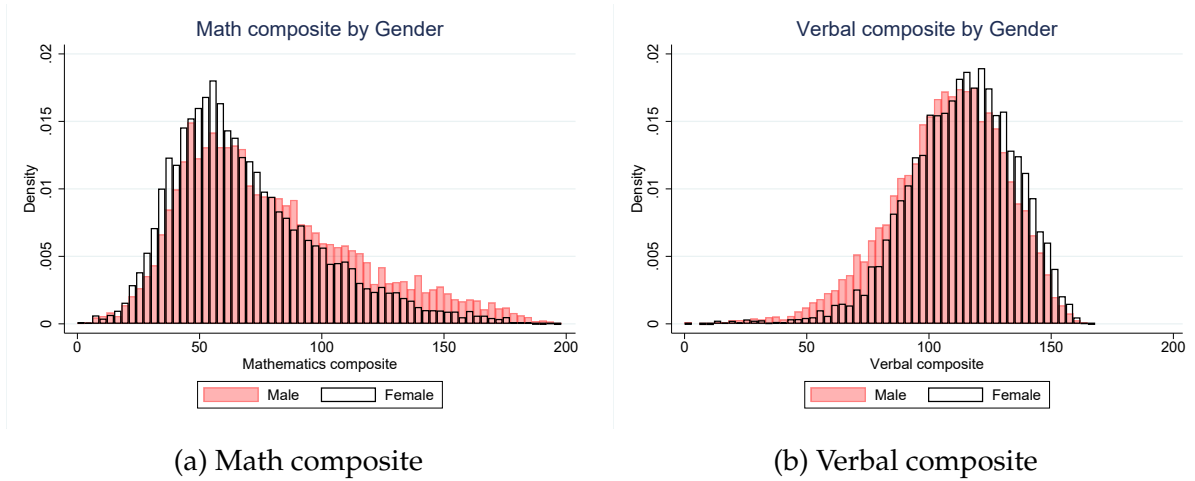(a) Math composite   (b) Verbal composite

Figure 2: Composite Scores by Gender in Project TALENT

less talented than men. Women's verbal scores slightly dominate the men's while the opposite is true for the math scores (see Figure 2b).

The distributions of composite IQ scores for Project TALENT respondents split by gender are quite similar (see Figure 3). What reconciles Figures 1-3 is the selection of more talented women into teaching, which we plot in Figure 3b. Due to the smaller numbers of individuals in Figure 3b compared to the three previous figures (the number of men is particularly low), the plot contains more noise and one should draw any conclusion with appropriate caution. With this in mind, it does appear that teaching attracts higher-ability individuals and that the selection of women along this dimen-

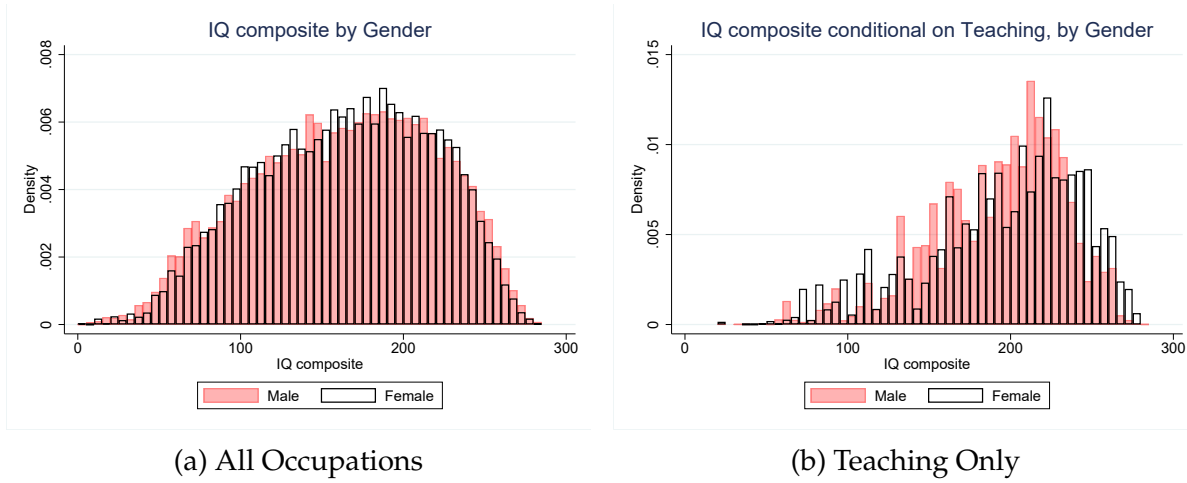|	
(a) All Occupations	(b) Teaching Only

Figure 3: Composite IQ Scores by Gender in Project TALENT

sion is somewhat stronger, plausibly due to the lack of career opportunities in other occupations.

The 11-year follow-up survey provides a snapshot of young workers' occupational choices together with the pre-labor market aptitude profile in the early 1970s. For later years, we use date from the two NLSYs and we will give a brief overview in the next section.

4.2 NLSY79 and NLSY97

In contrast to Project TALENT, the two NLSYs initiated in 1979 and 1997 are widely used and many readers are already familiar with the data. Cooksey (2018), among many others, provides a comprehensive overview of the two surveys. For the purposes of our paper, the *Armed Forces Qualification Test* (AFQT) scores are of particular interest since we need proxies to capture differences in the individuals' raw skills, which is the crucial source of exogenous variation in our model. The score consists of four sections of the *Armed Services Vocational Aptitude Battery* (ASVAB): arithmetic reasoning, word knowledge, paragraph comprehension, and numerical operations. The tests took place in 1980 and 1999 for NLSY79 and NLSY97 cohorts, respectively.

Since the tests were administered at different ages (and times) and used different formats, NLS program staff re-normed the scores in 2006, in order to facilitate cross-

cohort comparisons.[8] Table 3 shows summary statistics for the individuals in our sample, divided into three categories: (1) teachers, (2) non-teachers/others, (3) not in labor force. Since an individual usually has multiple employment spells, she might be in all these categories over the course of the panel. We define an individual as a teacher if she has at least one spell in teaching that is longer than 9 months (one academic year). An individual is out of labor force if she reported to be out of labor force for 208 consecutive weeks for a reason *other* than schooling.[9]

	Panel A: NLSY79		
	Teachers	Other	Not in Labor Force
Female	505	2,739	1,006
(% of subsample)	(74.93%)	(43.82%)	(65.80%)
AFQT mean	53,160.36	42,859.81	31,842.39
(st.dev.)	(29,157.36)	(28,518.21)	(31,842.39)
Observations	674	6,251	1,066
(% of total sample)	(7.98%)	(73.97%)	(12.61%)

	Panel B: NLSY97		
	Teachers	Other	Not in Labor Force
Female	577	2,580	331
(% of subsample)	(71.50%)	(45.1%)	(63.65%)
AFQT mean	62,229.16	44,239.26	33,108.32
(st.dev.)	(26,349.47)	(28,681.16)	(27,507.65)
Observations	807	5,721	520
(% of total sample)	(11.51%)	(81.60%)	(7.42%)

Table 3: NLSY AFQT Scores by Occupation

If indeed labor market barriers are shrinking over time, we would expect to see some evidence of this in the occupational choice of women in more recent cohorts and

[8]Altonji et al. (2009) also propose a method to enable a comparison of AFQT scores across the two NLSY samples. Note that we cannot construct equivalence scores that would allow as to compare results from aptitude tests in Project TALENT with those from the two NLSYs.
[9]After four years, "not in the labor force" is effectively an absorbing state. The probability of returning to the labor force is less than 1%.

hence in the NLSY79 and NLSY97. In these younger cohorts, the share of women among teachers continues to be high. The gender composition of teachers in NLSY 79 is comparable to the one we observe in Project TALENT. In the more recent NLSY 97, the share is a few percentage points lower. In contrast, the share of female workers in other occupations is rising, from approximately 39 percent in Project TALENT to 44 percent in NLSY79, and 45 percent in NLSY97. The gender composition of those who are out of the labor force is heading in the opposite direction. It dropped to 66 percent among NLSY79 respondents and 64 percent among NLSY97 respondents, compared to almost 80 percent in the 11-year post graduation survey from Project TALENT.

One significant difference between the Project TALENT and NLSY respondents are the aptitude scores for those who are not in the labor force. While their average score was *higher* than the average score of those in *Other* in Project TALENT, the ranking is reversed in the two NLSYs. This suggests that the exit rate from "not in the labor force" is correlated with the respondents' ability. High and low ability individuals appear to respond differentially to changes in labor market discrimination and this is further corroborated in a set of plots we present below.

Figure 4 plots the NLSY analogues of the math and verbal scores in Project TALENT shown in Figure 1. Clearly, teaching still attracts higher-ability individuals than "other" occupations or those who are not in the labor force. While the ability distribution of teachers still first-order dominates the scores in "other" occupations in the 1979 survey, there is a remarkable shift in the talent distribution of those who are not in the labor force. It is for more right-skewed than the distribution of workers in "other" occupations. Unfortunately, the 1997 data is considerably more noisy mostly due to the smaller sample size and we're refraining from reading too much into any particular feature of the data. This being said, it is worth highlighting that the definition of "not in the labor force" is *not* identical in the TALENT data on the one hand and the NLSY data on the other.[10] It is, however, suggestive of a trend whereby high-ability individuals are pulled into the labor force more "aggressively" compared to low-skill workers over time.

Figures 5-6 are the NLSY analogues of Figures 2-3.

Unfortunately, the composite score distributions aren't comparable between Project

---

[10]The difference stems from the fact that we observe Project TALENT participants relatively infrequently and the respondents' labor market histories are less complete.

(a) Math Knowledge (NLSY79)

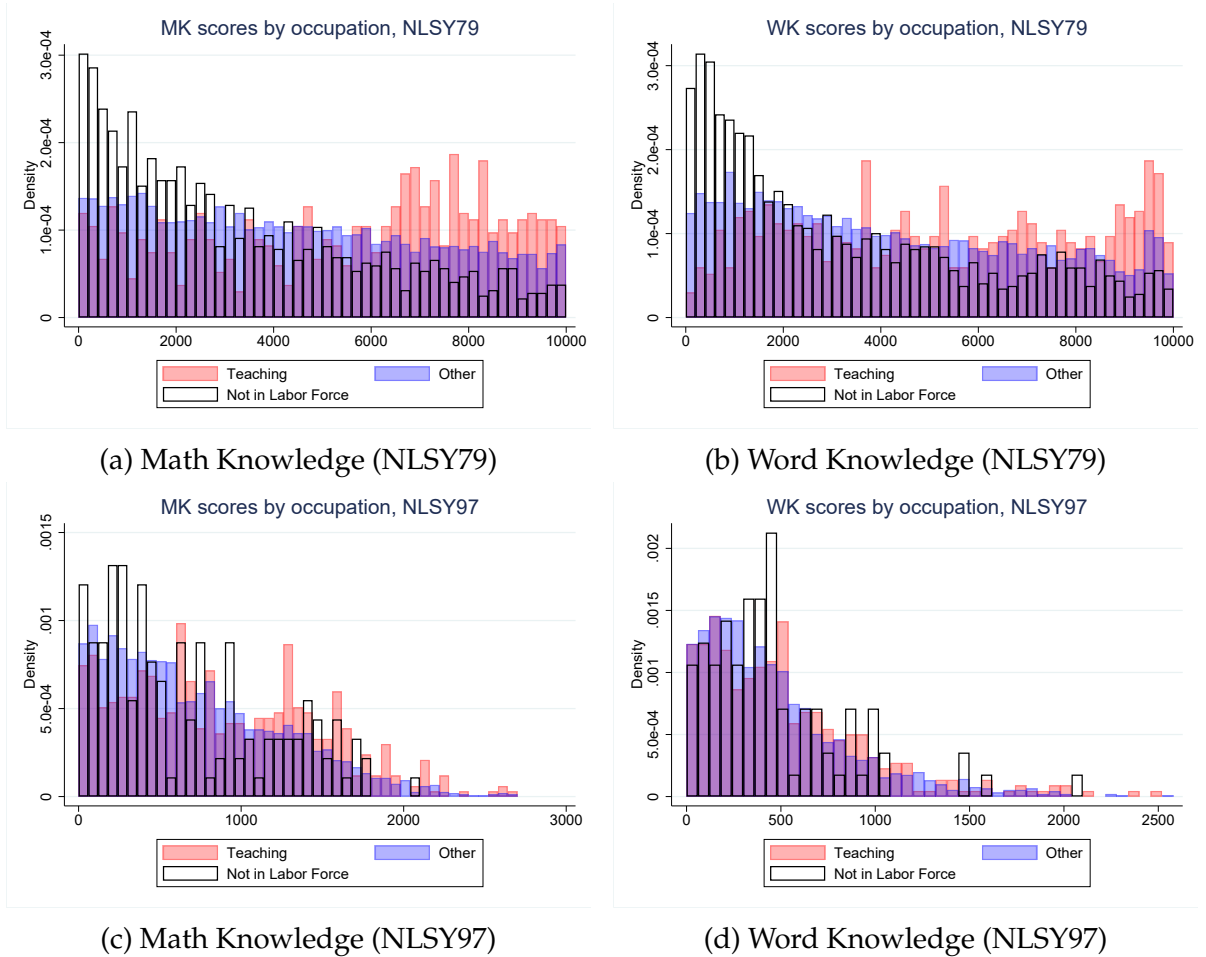(b) Word Knowledge (NLSY79)

(c) Math Knowledge (NLSY97)
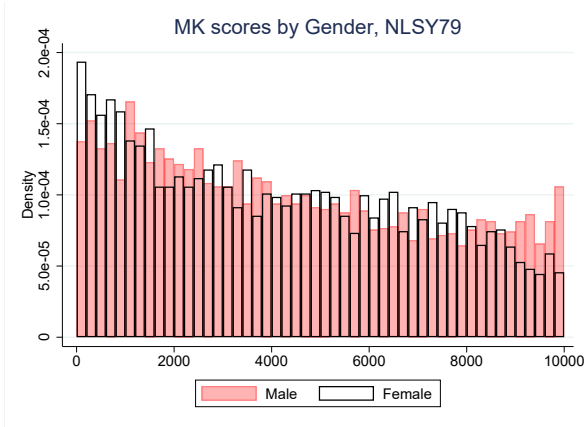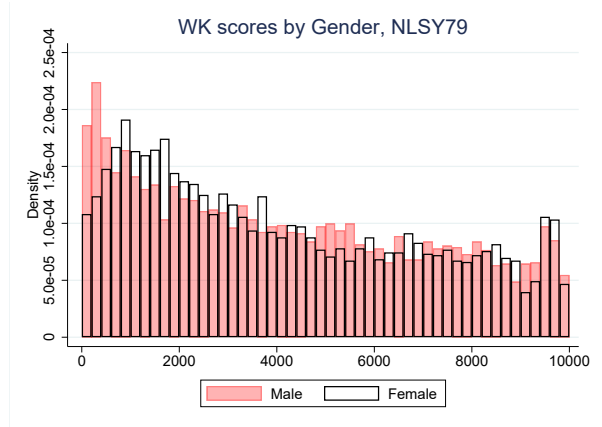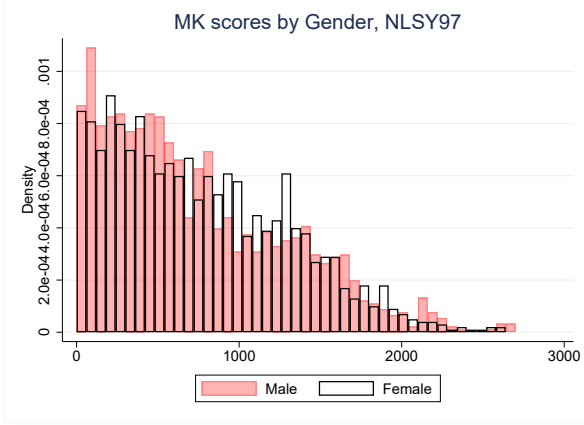
(d) Word Knowledge (NLSY97)
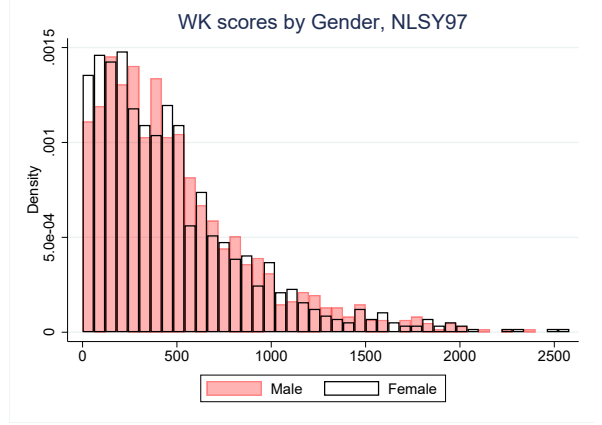
Figure 4: AFQT Scores by Occupations

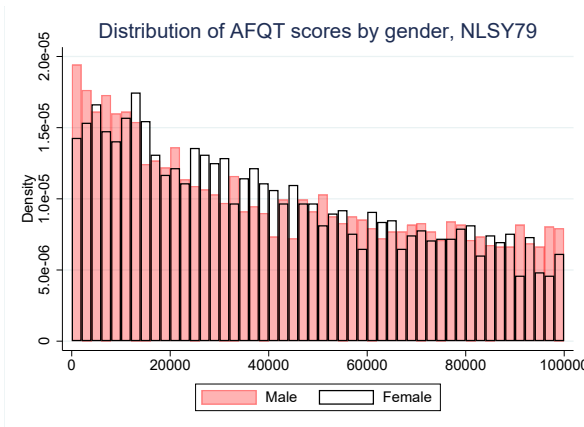(a) Math Knowledge (NLSY79) (b) Word Knowledge (NLSY79)
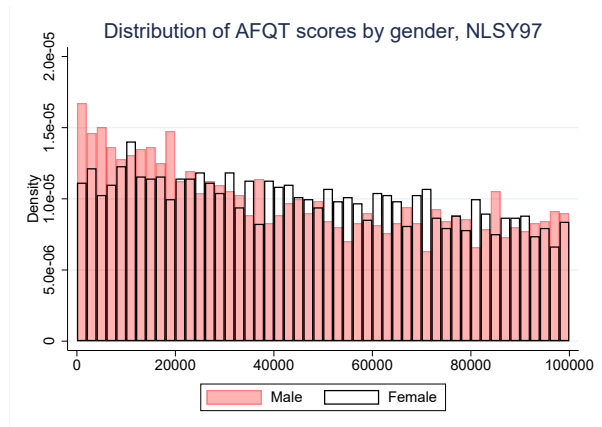
(c) Math Knowledge (NLSY97) (d) Word Knowledge (NLSY97)

Figure 5: Topic Scores by Gender

(a) AFQT Composite (NLSY79) (b) AFQT Composite (NLSY97)

Figure 6: AFQT Scores by Gender

TALENT and NLSY cohorts for several reasons. Although the ability tests try to capture math and verbal abilities, the tests themselves are very different. Therefore, it is impossible to compare changes in score distributions over time. Despite this limitation we can still compare the composition (by gender or occupation) *within* cohorts across the three surveys. In fact, the particular *Roy* flavor of our model suggests that this is where the interesting action is.

Differences in the frequency of the surveys further complicate comparisons between Project TALENT and the NLSY's. The NLSY79 and NLSY97 survey the respondents far more frequently and this enables us to construct more detailed employment histories and occupation choices over the course of panels. While Project TALENT provides data on the respondents' *current* occupation in the 1, 5, and 11-year post waves, it misses some of the higher-frequency churn since it doesn't ask any retrospective career questions. Moreover, the probability of being a teacher one year after high school graduation is vanishingly small (0.08% of the sample). In the figures and tables presented so far we report the occupations reported in the 11-year follow-up study. While certainly less complete than the NLSY data, Project TALENT still offers insights into the career decisions of young adults at a time for which there is only a very limited supply of nationally representative longitudinal survey data.[11]

Project TALENT, NLSY79, and NLSY97 provide us with the data moments that we match to model-generated moments in order to quantify the static and dynamic effects of the implied educational barriers and labor market barriers. While the discussion so far emphasized the contrast between men and women, it is worth emphasizing that our model framework can accommodate more richness in terms of the salient groups of workers. In ongoing work we are incorporating race and ethnicity to more fully capture the salient educational and labor market frictions in the data.[12]

The final element in our calibration element relies on O*Net's characterization of occupation-specific skill requirements. We use a three-dimensional job characterization: mathematics, verbal, and social. We decompose the requirements of each job into an absolute and a comparative dimension in a way that aligns naturally with the

---

[11]In ongoing work, we synthetically impose some of the data limitations in Project TALENT on the NLSY in order to get a better sense for underlying changes in labor market conditions and career decisions across cohorts. Evidently we're giving up some valuable details in the process for the benefit of more cross-survey comparability.

[12]Since the racial categorization is somewhat problematic in Project TALENT, this extension poses additional challenges.

notion of absolute and comparative advantage.

We characterize the comparative requirements by a triple of weights based on the dimension-by-dimension skill ranking of each job. To account for the possibility that jobs with similar comparative profiles are characterized by different overall skill intensities, we also create an absolute profile based on the weighted sum of skill-specific rankings. Given the non-linearity of payoffs in our model, the distinction between comparative and absolute advantage is critical for our calibration strategy[13]

Last but not least, we use these skill profiles to project the survey respondents' three-dimensional abilities into an N-dimensional vector of job-specific abilities.

THIS IS CURRENTLY WORK IN PROGRESS!

5 Decentralized Competitive Equilibrium (No Frictions)

In order to highlight the mechanisms that govern the individuals' occupational choice we first solve for a decentralized equilibrium that implements the optimal allocation of adults to "teaching" and "other" in a single location. Put differently, we solve the model for a wage profile $\omega^*(h^T)$ that pays teachers with human capital h^T their marginal product.[14]

How can we characterize the teacher's marginal product? Recall that only workers produce the final (consumption) good using a linear production function with productivity A^O. Thanks to stable random matching of students to teachers the ability distribution of students is identical across class rooms and a teacher's marginal value only varies with the number of students in her class room, denoted by $N(h^t)$. Clearly, then, the wage profile is linear in $N(\cdot)$. Since N is not linear in h^T in general this implies that the wage profile is not linear in h^T either.

Since teachers are compensated for the incremental human capital in "other", the

---

[13]In this respect, our model differs from a plain-vanilla occupational choice model à la Roy (1951).

[14]In section 2 we specified the model for an arbitrary profile ω. Here we impose further restrictions on this function.

wage profile has the following functional form:

$$
\omega(h^T, \widetilde{H}^T) = \frac{\overbrace{H'^O A'^O}^{\text{total output (next period)}}}{\underbrace{\int_0^\infty f^O(a)da}_{\substack{\text{fraction of prospective} \\ \text{production workers in class}}}} \quad \underbrace{\frac{\left(h^T\right)^{\frac{\beta}{\sigma}}}{\widetilde{H}^T}}_{\text{fraction of students taught}}
\tag{18}
$$

where

$$
\widetilde{H'}^T = \int_0^\infty h^{T\frac{\beta}{\sigma}} f^T(a)da
$$

is the $\left(\frac{\beta}{\sigma}\right)^{\text{th}}$ moment of the distribution of human capital among current teachers.

The second term in equation (18) captures the idea that a teacher's wage depends on class size, which is a non-linear function of her human capital h^T. Recall from equations (16) and (17) that a prospective teacher's optimal s^T and e^T depends on the quality of today's teachers (i.e. their human capital) and on the slope of the wage profile next period and hence on both today's $\widetilde{H}^T$ as well as tomorrow's $\widetilde{H'}^T$.

6 Decentralized Competitive Equilibrium with Educational Barriers and Labor Market Discrimination

To highlight the effect of educational barriers and labor market discrimination we add group- and occupation-specific distortions to the model in section 5. This is mostly a heuristic exercise since parsimony of our state space enables us to characterize the full transition path from one set of distortions to another. We highlight this feature of the model here since there have arguably been secular changes in educational barriers and labor market distortions for groups other than white males. To fix ideas, however, we simply compare steady states associated with particular wedges τ indexed by ω (for labor market discrimination), e (for educational barriers), and g for groups (race, gender,...).

We can show that this economy has a steady state in H^T and we can characterize it in

a single equation:

$$\text{To Be Inserted!} \tag{19}$$

Moreover, the occupational choice boundary, denoted $\bar{a}^T(a^O)$ depends on the magnitude of the educational barriers and labor market frictions, but not current or future aggregate human capital in teaching. The boundary can be characterized implicitly by the following condition:

$$\left(\frac{1-\tau'^{O,w}}{1-\tau'^{T,w}}\right)\left(\frac{1-\eta}{1-\frac{\eta\beta}{\sigma}}\right)\left(\frac{1-s^O}{1-s^T}\right)^{\frac{1}{\mu}}\left(\frac{(a^O)^{\frac{\alpha}{1-\eta}}}{\left(\bar{a}^T(a^O)\right)^{\frac{\alpha}{\frac{\sigma}{\beta}-\eta}}}\right)$$
$$=\left(\frac{\int_0^\infty a^{\frac{\alpha}{1-\eta}}f^O(a)da}{\int_0^\infty f^O(a)da}\right)\left(\int_0^\infty a^{\frac{\alpha}{\frac{\sigma}{\beta}-\eta}}f^T(a)da\right)^{-1} \tag{20}$$

The calibration in section 8 will go beyond the simple comparative statics here and match moments along the secular transition from the late 1960s to the more recent past.

7 Model Extension with $L > 1$ Locations

In section 2 we motivated lump-sum taxation by the stylized fact that school districts are, to a considerable extent, funded by local property taxes. It is therefore natural to consider multiple locations with distinct taxes $\{T_1, \ldots, T_L\}$. In this version of the model, young agents not only decide which career path they would like to pursue, but they also choose a location for their education. The assumption of a single location with (stable) random matching between students and teachers had the benefit of tractability. The aggregate state of the economy could be summarized by a single scalar, $\widetilde{H}^T$.

Abandoning the assumption of random student assignment implies that the aggregate state is now given by an entire distribution of $\widetilde{h}^T$ rather than a single moment of that same distribution. This turns out be a rather challenging generalization. What appears to be more relevant empirically is the introduction of multiple locations, differentiated by local (lump sum) tax levels and possibly some (as of yet) unspec-

ified amenity. As long as we maintain the assumption of effectively random student-teacher assignment within locations, we can rely on tools from Lucas and Moll (2014) and more recently used in Martellini (2019) to keep the state space small enough and thus the dynamic program tractable.

This model extension enables us to capture the quantitative effects associated with the salient stylized fact of increasing socio-economic segregation of students – and plausibly even teachers – between inner cities and suburbs.

THIS IS CURRENTLY WORK IN PROGRESS!

8 Quantitative Analysis

8.1 Calibration Strategy

THIS IS CURRENTLY WORK IN PROGRESS!

8.2 Counterfactual Experiments

TO BE COMPLETED!

9 Conclusion

TO BE COMPLETED!

References

ALTONJI, J., P. BHARADWAJ, AND F. LANGE (2009): "Constructing AFQT Scores that are Comparable Across the NLSY79 and the NLSY97," Data Note.

COOKSEY, E. C. (2018): "Using the National Longitudinal Surveys of Youth (NLSY) to Conduct Life Course Analyses," in *Handbook of Life Course Health Development*, ed. by N. Halfon, C. B. Forrest, R. M. Lerner, and E. M. Faustman, Cham: Springer International Publishing, 561–577.

EATON, J. AND S. KORTUM (2002): "Technology, Geography, and Trade," *Econometrica*, 70, 1741–1779.

HSIEH, C.-T., E. HURST, C. I. JONES, AND P. J. KLENOW (2018): "The Allocation of Talent and U.S. Economic Growth," *Working Paper*, 1–72.

LUCAS, R. E. AND B. MOLL (2014): "Knowledge Growth and the Allocation of Time," *Journal of Political Economy*, 122, 1–51.

MARTELLINI, P. (2019): "The City-Size Wage Premium: Origins and Aggregate Implications," *Working Paper*, 1–76.

ROY, A. (1951): "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers New Series*, 3, 135–146.