

# The Allocation of Teaching Talent and Human Capital Accumulation

Simeon D. Alder\*

Yulia Dudareva<sup>†</sup>

Ananth Seshadri<sup>‡</sup>

January 24, 2020

PRELIMINARY AND INCOMPLETE!

## Abstract

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, 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.

---

\*University of Wisconsin - Madison

<sup>†</sup>University of Wisconsin - Madison

<sup>‡</sup>University of Wisconsin - Madison

# 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 3 reviews the data and we calibrate the model to match salient empirical moments in section 4. Section 5 concludes.

# 2 Model

The model is populated by a constant measure  $M$  of workers who are employed in one of the  $N$  sectors of the economy. Each individual is born with an exogenous  $N$ -dimensional vector of skills  $\vec{a}$ , where each element characterizes her innate ability in any one of the  $n \in \{1, \dots, N\}$  occupations in the economy, each corresponding to one of the sectors or industries. This vector of abilities is drawn from a known joint distribution  $F_a(\vec{a})$ .

Without loss of generality, occupation 1 is *teaching*. In this economy, teachers are distinct from other occupations since their technology produces human capital rather than the homogeneous final good in the remaining  $N - 1$  sectors. 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 working age (or "old").

## 2.1 Workers

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 input.

The timing convention is similar to [Hsieh et al. \(2018\)](#). In the first period, prospective workers – or students – 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 and time devoted to human capital investment:

$$U = \max_{C, s, e} \sum_{t=c}^{c+1} \beta^{t-c} \ln C_t + \ln(1 - s_c) \quad (1)$$

where

$$C_t = (1 - \tau)w_t(h) - e_t, \text{ for } t \in \{c, c + 1\} \quad (2)$$

$$\vec{h}_{c+1} = g(h_{c+1}^T, \vec{a}) s_c^\phi e_c^\eta (N(h^T))^{-\sigma} \quad (3)$$

$$g(h_{c+1}^T, \vec{a}) = (h_{c+1}^T)^\beta \left[ a_1^\alpha \dots a_N^\alpha \right]. \quad (4)$$

$C_t$  is consumption of the homogeneous final good when young and old (i.e. when  $t = c$  and  $t = c + 1$ ),  $s_c$  is the time allocated to human capital formation in the first period of each cohort indexed by  $c$ , and  $e_c$  are units of the final good invested in human capital formation when workers are still students. The periodic time endowment is set to unity.  $1 - s_c$  is leisure time when young and labor is supplied inelastically when individuals reach working age.  $\tau$  is the constant marginal tax rate levied on labor income.

The joint distribution of  $\vec{a}$  is a multivariate Fréchet with c.d.f.

$$F(a_1, \dots, a_N) = \exp \left[ - \sum_{n=1}^N a_n^{-\theta} \right], \quad (5)$$

where the idiosyncratic draws are independent. We borrow this assumption from [Eaton and Kortum \(2002\)](#).

The production technologies in occupations  $2, \dots, N$  are linear in efficiency units of labor  $h$  with productivities  $\{A_n\}_{n=2}^N$ .

## 2.2 Occupational Choice

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  $h^T$ .

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. Clearly, then,  $e_{c+1} = s_{c+1} = 0$ .

Since we’re mainly interested in the choice between a teaching career and other occupations, we will first characterize the problem in the special  $N = 2$  case and label the occupations with  $T$  (teaching) and  $O$  (other). In a second step, we can review 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  $\beta$  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.

### 2.2.1 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  $h^{T'}$ :

$$N(h^{T'}) = \left( \frac{h^{T'}}{h^T} \right)^{\frac{\beta}{\sigma}} N(h^T) \quad (6)$$

### 2.2.2 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  $a$ .<sup>1</sup>

Now that we have a characterization of class size and student-teacher matching we can think about the value of a teacher's contribution and how it shapes the workers' occupational choices. To proceed we need some additional notation.

Let  $(h^*(h^O), h^O)$  denote the human capital vector across the two occupations such that the worker is indifferent between a teaching career and the alternative. Since the "other" production technology is linear in labor with productivity  $A^O$ , the period 2 income in  $O$  is  $w^O h^O = A^O h^O$  and indifference implies that the marginal teacher's income is  $\omega(h^*(h^O)) = w^O h^O$ . Starting with the marginal teacher we can build the complete wage profile  $\omega$ .

To do so we need to first characterize how much more human capital a supra-marginal teacher produces and how valuable that extra amount is in light of the students' occupation choices. At this point, we're only interested in a stationary equilibrium and our notation should be interpreted accordingly.

To make further progress, we need to solve for the optimal human capital investment decisions in terms of time and goods. The former, as is standard in these models, doesn't depend on the student's talent or future occupation:

$$s = \frac{1}{1 + \frac{1-\eta}{\gamma\phi}}. \quad (7)$$

The latter, in contrast, depends on the student's occupational choice and the slope of the wage profile in each occupation. In  $O$ , the wage profile is linear and the marginal value of human capital is  $w^O = A^O$ . In teaching,  $\omega(h^T)$  can be non-linear and the choice of  $e$  depends on  $\omega'(h^T)$ :

$$e(a) = \begin{cases} [(1-\tau)w^O(a^O)^\alpha s^\phi \eta c^*]^\frac{1}{1-\eta}, & \text{if } a^T < \frac{h^*(h^O)}{h^O} \cdot a^O \\ [(1-\tau)\frac{d\omega}{dh}|_{h^T} \cdot (a^T)^\alpha s^\phi \eta c^*]^\frac{1}{1-\eta}, & \text{if } a^T \geq \frac{h^*(h^O)}{h^O} \cdot a^O \end{cases} \quad (8)$$

where  $h^{T'}$  is the student's human capital in teaching in the second period of her life (i.e. at  $c+1$ ) and  $c^* = (h^T)^\beta (N(h^T))^{-\sigma}$  is a constant. It is pinned down by the students' indifference condition as well as market clearing for students and teachers. We will return to the latter

---

<sup>1</sup>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.

once we have all the ingredients to describe the distribution of  $h^T$  among teachers.

At this point, it is worth emphasizing that  $e(a)$  depends on  $a$  and the student's future occupation. However, educational investments and the teacher's input into human capital production are such that *all* raw skills are “upgraded” proportionally and we have  $\frac{h^{T'}}{h^{O'}} = \frac{a^T}{a^O}$ .

Now we have all the elements to quantify how much more human capital a teacher with  $h^T > h^*(h^O)$  produces compared to the marginal teacher with  $h^*(h^O)$ . Holding  $a$  fixed, all teachers generate the same human capital “gain” per student. The increased value of a better teacher comes from class size and we can compute the additional human capital by integrating over the distribution of student ability (adjusted for the size of each cohort  $M/2$ ):

$$\begin{aligned} & [N(h^T) - N(h^*(h^O))] \int a^\alpha s^\phi e(a)^\eta \frac{dF(a)}{M/2} \\ &= N(h^*(h^O)) \left[ \left( \frac{h^T}{h^*(h^O)} \right)^{\frac{\beta}{\sigma}} - 1 \right] \int a^\alpha s^\phi e(a)^\eta \frac{dF(a)}{M/2}. \end{aligned} \quad (9)$$

Equation (9) describes the excess human capital production relative to a marginal teacher with human capital  $h^*(h^O)$ . Absolute class sizes must be such that the (inelastic) supply of students can be distributed among the endogenous supply of teachers, which is a function of the joint distribution of human capital in the two occupations and each workers occupational choice. Let  $f(a^O, a^T)$  denote the joint density over the students' skills in the two occupations and  $f^T(a^T) = \int_0^\infty f(a^O, a^T) da^O$  is the marginal density over  $a^T$ . We can then take advantage of the fact that  $\frac{h^{T'}}{h^{O'}} = \frac{a^T}{a^O}$  to derive the density  $\hat{f}^T$  over  $a^T$  among teachers only:

$$\hat{f}^T(a^T) = f^T(a^T) \Pr \left[ \frac{a^T}{a^O} \geq \frac{h^*(h^O)}{h^O} \right]. \quad (10)$$

Given (3), (7), and (8), we can describe the teachers' human capital as a function of their  $a^T$ . To close the model, the market for students and teachers must clear, i.e. we must find a reference class size  $N(h^*(h^O))$  for a marginal teacher with  $h^T(a^T) = h^*(h^O)$ :

$$N(h^*(h^O)) = \frac{M/2}{\int_0^\infty \left[ \left( \frac{h^T(a^T)}{h^*(h^O)} \right)^{\frac{\beta}{\sigma}} - 1 \right] \hat{f}^T(a^T) da^T}, \quad (11)$$

for any  $h^O$ .

**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 Data

[Discuss Project TALENT, NLSY79, and NLSY97.]

We rely on moments from three longitudinal surveys of young Americans to parameterize our structural model.

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 1,000 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 1963, 1967, and 1973. 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 survey. 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 ?? we summarize the descriptive statistics by gender, race, parental education, and parental economic status (all in 1960).

[It looks like race was reclassified the 5-year post-graduation survey. What do we know about this?] TO BE COMPLETED!

	Base Year	1-Year Post	5-Year Post	11-Year Post
Sample Size	368,938	188,138	128,163	92,962
Black (% of sample size)	6,488 (1.76%)	4,003 (2.13%)	4,856 (3.79%)	3,026 (3.26%)
Other / Unknown / Conflicting (% of sample size)	215,786	73,668	3,739	3,758
Female (% of sample size)	184,912	98,183	64,841	47,417
Parents: High School Degree (% of sample size)				
Parents: College Degree (% of sample size)				
Parents: High Income (% of sample size)				

Table 1: Descriptive Statistics

## 4 Quantitative Analysis

### 4.1 Calibration Strategy

TO BE COMPLETED!

### 4.2 Counterfactual Experiments

TO BE COMPLETED!

## 5 Conclusion

TO BE COMPLETED!



## References

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