Teachers and the Evolution of Aggregate Inequality

Anson Zhou*

January 2024

Abstract

Teachers account for less than 5% of the workforce but have disproportionate impacts on the achievements of future generations. This paper studies how the reward structure of teachers affects income inequality at the aggregate level. I show that there is a two-way relationship between teacher quality and population-wide human capital distribution in an overlapping generations economy. On the one hand, the dispersion of human capital governs the occupation selection effect on teacher quality. On the other hand, deteriorating teacher quality elevates the dispersion itself through children's human capital formation. Exploiting empirical evidence on duty-to-bargain laws, I identify the model in closed form and quantify the proposed mechanism. Counterfactual results suggest that rewarding teachers' human capital, e.g., through performance-based compensations, has large dynamic spillover effects such as reducing aggregate income inequality and boosting intergenerational mobility.

JEL classification: I24, J24, J31, J45

Keywords: teachers, human capital, inequality

*Anson Zhou: Faculty of Business and Economics, The University of Hong Kong. I am indebted to Juan Carlos Cordoba, Chao Fu, Mike Ricks, Cailin Slattery, Uta Schöenberg, Michela Tincani, Joanna Venator, Alexander Willén, Alan Yang, Matt Zahn, Lichen Zhang, Xiaodong Zhu for helpful comments and suggestions. I also thank the participants at HK Junior Macro Group and HKEA Biennial Meeting.

1. Introduction

Teachers account for less than 5% of the workforce, but they play a fundamental role in shaping students' achievements and path to upward mobility (Rivkin, Hanushek and Kain 2005, Chetty, Hendren, Kline and Saez 2014b, Card, Domnisoru and Taylor 2022). It is widely acknowledged that teacher selection and compensation have profound implications not only on the equilibrium of teacher labor market but also on the next generation's outcomes (Hanushek 2011).

While recent research have made significant progress in measuring such implications (Jacob, Rockoff, Taylor, Lindy and Rosen 2018, Biasi, Fu and Stromme 2021, Lovenheim and Willén 2019, Lavy 2020, Tincani 2021), two important questions remain unanswered. First, how does the reward structure of teachers affect the dynamics of income inequality among non-teachers and the propagation of aggregate shocks such as the skill-biased technical changes? The answer to this question speaks to the unique role of teachers in the aggregate labor market and provides a connection between labor, education, and macroeconomics. Second, how does the magnitude of these effects evolve over time as the children being affected grow up, join the labor market, and (some of which) become teachers themselves? Will the one-generation estimates in the empirical literature be a lower or upper bound in a dynamic environment? This question is important for understanding the long-run general equilibrium impacts of education policies, especially when policymakers need to make cost-benefit analysis that is forward-looking by design in the context of teachers.

In this paper, I study these two questions in an overlapping-generations economy with heterogeneous agents and endogenous human capital formation. Building on the framework in Bénabou (2002), the model additionally considers occupation selection into teaching and non-teaching professions. The key insight of the model is that the selection effects of education policies depend on the dispersion of human capital in the population to begin with. Therefore, in an environment with human capital formation so that the level of dispersion is endogenous, *static* changes in teacher selection have large *dynamic* effects.

The mechanism can be succinctly summarized as follows: consider a scenario where decreasing rewards to human capital in the teaching profession pushes the most talented individuals into non-teaching occupations and thus reduces teacher quality. Deteriorating teacher quality reduces human capital in the next generation, with the effects being disproportionately larger among low-income families when teacher quality and parental investments are substitutes.[1] As a result, the economy has lower intergenerational mobility and a greater dispersion of human capital to begin with in the next generation. Importantly, this increased human capital dispersion reinforces the selection channel and leads to further reductions in teacher quality - a vicious cycle.

To quantify the mechanism, I show that the model provides analytical characterizations of optimal individual decisions as well as the transition path of the equilibrium human capital distribution. I also prove that the key parameters, especially the ones governing the human capital production function, are uniquely identified in closed form using cross-sectional moments and estimates on how teacher compensation scheme affects teacher labor market and children's outcomes. To that end, I utilize variations in duty-to-bargain laws across states and over time to measure the effects of teacher pay compression. In particular, I find that the enactment of such laws reduce both employment and earnings dispersion among teachers, while having mildly negative effects on the level of average earnings. I also borrow the corresponding estimates on children's earnings from Lovenheim and Willén (2019) where the authors find that duty-to-bargain laws reduce the earnings of children affected, with the effects being larger among chil-

[1]See Kotera and Seshadri (2017) and Yum (2023). I will also verify in this condition in the calibration results using empirical evidence from duty-to-bargain laws.

dren with parents from disadvantaged backgrounds. As non-targeted moments, the calibrated model also predicts teacher value-added that is very close to the estimates by Chetty, Friedman and Rockoff (2014a).

In counterfactual analyses, I show that even though teachers account for less than 5% of the total workforce, they have disproportionate impacts on the dynamics of inequality in the aggregate economy due to their fundamental roles in human capital formation. In particular, the two-way feedback mechanism proposed in this paper results in large amplifications of skill-biased technical changes (SBTC) along the transition path.

In the first counterfactual experiment, I raise the returns to human capital in the teaching profession by 5% to mimic a performance-based compensation reform. The magnitude of 5% is chosen to be the same as the initial impacts that duty-to-bargain laws had on the income dispersion among teachers in the empirical findings, albeit in a different direction. I find that only two-fifth of the initial impacts on teachers' income inequality, measured in the coefficient of variations, survive in the long-run steady state. Income inequality among non-teachers, however, reduces significantly by 3%, reflecting changes in the human capital dispersion itself. Intergenerational elasticity of earnings also decreases from 0.36 to 0.27 in the long run. In other words, changing returns to human capital in the teaching profession pushes the economy along the Great Gatsby curve (Corak 2013) towards less inequality and more intergenerational mobility. I also show that the magnitude of these effects varies along the transition path, approaching the new steady-state in about two generations. As a result, static or one-generation analyses overstate the effects of performance-based compensation reforms on income inequality within teachers and understate the spillover effects on intergenerational mobility and the income inequality among non-teachers.

In the second counterfactual, I study the role of teacher selection in propagating aggregate shocks, in particular on how the "convexification" of non-teaching occupations (Autor, Goldin and Katz 2020) affect income inequalities. I find that absent changes in teacher quality, the increase in income inequality among non-teachers would be 50% smaller in the long run. This is because increasing rewards to human capital in the non-teaching profession reduces teacher quality significantly, leading to a much larger dispersion of human capital in consecutive generations.

If the government wants to mitigate the adverse effects of the skill-biased technical change (SBTC) on aggregate inequalities, how much should it raise the returns to human capital among teachers? A model-based decomposition shows that to neutralize the effects of STBC, the returns to human capital among teachers actually need to rise *relatively more* than that in non-teaching occupations. The intuition for this interesting finding is that in addition to retaining high-quality teachers, the education system needs to address the fact that SBTC widens the gap in parental income and hence the private investments in children's human capital formation.

Related Literature

This paper builds on the literature that studies the determinants of inequality and intergenerational mobility. The literature has traditionally focused on factors that affect the demand for education, such as credit constraints (Lee and Seshadri 2019, Caucutt and Lochner 2020), information frictions (Hoxby and Turner 2015), and neighborhood effects (Chetty, Hendren, Kline and Saez 2014b, Durlauf and Seshadri 2018, Fogli and Guerrieri 2019, Chyn and Daruich 2022).[2] This paper contributes to the less studied literature regarding the supply side of education (e.g., Agostinelli, Luflade, Martellini et al. 2021 and Fu, Guo, Smith and Sorensen 2022) by showing that teacher market reforms could be powerful instruments that move

[2] See Blanden, Doepke and Stuhler (2023) for a recent summary of the literature.

the economy along the Great Gatsby curve toward lower inequalities and greater intergenerational mobility.

This paper is also related to the large body of empirical literature in labor and education economics that studies the structure of teacher labor market and its implications on students. The most related papers include Hoxby (1996) and Lovenheim (2009) on teacher unions, Bacolod (2007) on the importance of women's alternative employment opportunities, Lovenheim and Willén (2019), Lavy (2020), Biasi, Fu and Stromme (2021), and Tincani (2021) on collective bargaining and performance-based compensation, and Card, Domnisoru and Taylor (2022) on school quality and minimum teacher salary laws. In various institutional settings, the literature finds that rewarding more effective teachers raises teacher quality and result in improved children's outcomes, especially for those with disadvantaged backgrounds. To the best of my knowledge, this is the first paper to incorporate these one-generation estimates into an overlapping-generations general equilibrium setting and uncover large dynamic spillover effects into non-teaching professions.

This paper also contributes to the literature that uses structural models to study the aggregate impacts of education policies (Abbott, Gallipoli, Meghir and Violante 2019, Fuchs-Schündeln, Krueger, Ludwig and Popova 2022). A key parameter in this class of models is the elasticity of substitution between parental and public inputs in children's human capital production. This parameter governs the crowding-out effects and distributive consequences of government policies. While the literature either assume it to be infinity (Daruich 2018) or calibrate it using cross-sectional variations of education expenditures as a share of income (Kotera and Seshadri 2017), this paper proposes a new identification strategy utilizing empirical evidence on the differential impacts of education policies. The results provide additional evidence showing that the two kinds of inputs are imperfect substitutes. Furthermore, this paper also provides a rare example where all key parameters in the quantitative general equilibrium model are transparently identified in closed form.

Last, this paper is related to the literature that studies the aggregate impacts of reward structure in some special occupations, such as government officials (Murphy, Shleifer and Vishny 1991, Acemoglu 1995) and entrepreneurs (King and Levine 1993, Baumol 1996). This paper contributes to the literature by studying another pivotal occupation in the economy – teachers. The framework developed here could be fruitfully applied to other occupations if the same amplifying mechanism applies and the set of identifying moments are available in the data.

The rest of the paper is organized as follows. In Section 2, I present the model and discuss its key assumptions and mechanisms. In Section 3, I present the identification proof and the empirical results on how duty-to-bargain laws affect teachers and students. Section 4 contains the main results on policy counterfactuals. Section 5 discusses several robustness checks. Section 6 concludes.

2. Model

I study an overlapping-generations economy populated by agents that live for two periods - children and adults. Children do not make any decisions. Their human capital is formed through a production function that takes teacher quality and parental investments as inputs. Adults with different levels of human capital supply labor inelastically and select into two occupations: teachers and workers (non-teachers). After occupation selection, adult parents choose child investments that maximizes their utility from consumption and preferences on their children's human capital.

2.1 Occupation Selection

In the beginning of each period, adults with human capital h make occupation choice after observing idiosyncratic preference shock ν that follows a Gumbel distribution with location parameter κ and scale parameter θ. They solve the following occupation choice problem:

$$\max_{j\in\{1,2\}} \mathbb{1}_{j=1}(\underbrace{\log(wh^{\psi_1}) + \nu}_{\text{teachers}}) + \mathbb{1}_{j=2}\underbrace{\log(h^{\psi_2})}_{\text{workers}}$$

where w is the relative wage across occupations and ψ_j is the occupation-specific returns to human capital.

Assume that the underlying human capital distribution of adults $F(h)$ follows a log-normal distribution

$$\log(h) \sim F(h) = \mathcal{N}(\mu, \sigma^2)$$

The log-normality of the human capital distribution is a condition that I will verify in the equilibrium. Because human capital is endogenously formed, the values of μ and σ^2 are both equilibrium objects in the steady state.

I define teacher quality ξ as the z-score of average (log) human capital of teachers in the population, i.e.,

$$\xi \equiv \frac{\overline{\log(h_1)} - \mu}{\sigma} \tag{1}$$

where h_1 denotes the human capital of individuals that choose to be teachers. Throughout the paper, I use $\overline{x}$ to denote the average value of variable x.

2.2 Human Capital Formation

After the occupation selection stage, adult workers become parents and have children.[3] Parents solve the following maximization problem

$$\max_{c,e\geq 0} \ \log(c) + \beta \mathbb{E}_{\epsilon} \log(h')$$

subject to budget constraint

$$c + e = y = h^{\psi_2}$$

and human capital production function

$$\log(h') = A + \underbrace{\log(\epsilon)}_{\text{normal dist.}} + \underbrace{\delta_0 \xi + \delta_1 \log(e) + \delta_2 \log(e)\xi}_{\text{translog}} + \underbrace{\rho \log(h)}_{\text{residual persistence}} \tag{2}$$

where c is consumption, e is education investments, h' is children's human capital, A is a normalizing constant, and ϵ is an idiosyncratic shock that follows log-normal distribution

$$\log(\epsilon) \sim \mathcal{N}(-\sigma_\epsilon^2/2, \sigma_\epsilon^2)$$

so that $\mathbb{E}(\epsilon) = 1$.

I assume that average teacher quality ξ and parental investments e interact in the human capital production function (2) in a translog fashion.[4] Parameter δ_2 governs whether they are gross complements or substitutes. I also allow for a direct spillover effect of parents' human capital to the next generation governed by ρ.

[3] I assume that only workers have children for analytical tractability of the model. Including teachers' children in this stage will not change the results significantly because teachers account for less than 5% of the adult population.

[4] One can also view the translog production function as the Kmenta (1967) approximation to the more commonly-used CES production function.

2.3 Discussions

The model presented above is rather stylized. Here, I discuss several important assumptions in the paper regarding human capital production function, occupation selection, and teacher quality in order to provide a rationale for my modeling choices. The payoff of these assumptions will become clear when I discuss the model solution and identification in Sections 2.4 and 3.

First, the human capital production function (2) should be interpreted as a reduced-form way of modeling how parental inputs and the pool of teachers jointly determine children's outcome. This approach abstracts away from the matching process between heterogeneous teachers and students in (potentially frictional) education markets.[5] Thus, parameters $\{\delta_0, \delta_1, \delta_2\}$ should be interpreted broadly. For example, beyond purchasing books and computers, the increase in children's human capital coming from $\delta_1 \log(e)$ also reflects parents' efforts in matching their own children to better teachers in the economy. Likewise, the interaction term $\delta_2 \log(e)\xi$ captures the idea that need to compete for better teachers varies as the total amount of teaching resource changes. An explicit micro-foundation of the assignment problem between heterogeneous students and teachers (e.g., Seshadri 2000, Biasi, Fu and Stromme 2021) is interesting and important in its own right but *not* the main focus of this paper.[6]

Second, I have assumed that individuals differ by a one-dimensional human capital h and an idiosyncratic taste shock ν. Human capital here summarizes a broad set of personal traits that affect productivities, including but not limited to cognitive ability, non-cognitive ability, experience, and education. A direct implication of this assumption is that changing opportunities in non-teaching occupations affects the quality of teachers. This prediction is consistent with empirical

[5] I thank Michela Tincani for pointing this out.

[6] For example, Dupuy (2012) provides conditions where a CES production function can be micro-founded in an assignment model with heterogeneous workers to heterogeneous jobs.

observations by Bacolod (2007) and Corcoran, Evans and Schwab (2004).

An alternative modeling choice is to use a multi-dimensional human capital model à la Roy (1952) where individuals have comparative advantage in one of the occupations (Bacolod 2007, Tincani 2021). With strong positive correlation between different abilities, the Roy-style model offers the same intuition on the selection effects. Such a model, however, requires the modeling of separate production functions for each dimension of human capital and the corresponding mappings to productivity across occupations. Here, I restrict to the one-dimensional case for tractability.

An important observation, however, is that a highly positive correlation between teaching and non-teaching abilities is a *sufficient* but not *necessary* condition for the model mechanism to work. As long as the skills that are more important in the teaching profession (e.g., communication and interpersonal relation skills) are equally malleable as the ones utilized in other occupations, the changing dispersion of teaching abilities mirrors that of non-teaching abilities.

To explore this concern further, I collect data on the importance of five cross-functional skills by occupation from O*NET, including complex problem solving skills, resource management skills, social skills, system skills, and technical skills. I find that for each skill requirement, the value of teachers are all within one standard deviation of the values across other non-teaching occupations (see Table 1). In other words, teacher is not an outlier occupation that utilizes a skill combination that is drastically different from non-teachers. Therefore, although extending the framework to a full-fledged Roy model is potentially interesting, I do not expect including multi-dimensional human capital will change the results significantly.

Third, the two parameters ψ_1 and ψ_2 provide sufficient statistic of the returns to human capital in teaching and non-teaching occupations respectively. It is important to note that these parameters reflects both technological differences across occupations and also policy choices. For example, as I show in the empirical anal-

Table 1: Skill Importance by Occupation

	Teacher	Non-Teachers	
	Value	Mean	Std
Complex Problem Solving Skills	3.53	3.19	0.50
Resource Management Skills	2.40	2.39	0.66
Social Skills	3.27	2.89	0.52
System Skills	3.17	2.85	0.60
Technical Skills	1.45	1.93	0.79

Notes: This table displays the importance of each cross-functional skills by teachers and non-teachers in the O*NET dataset.

ysis, ψ_1 falls when the teachers' union compresses the wage distribution through collective bargaining. Likewise, ψ_1 rises when school districts decide to adopt performance-based compensations. By the same token, the "convexification" of non-teaching occupations (Autor, Goldin and Katz 2020) due to skill-biased technical changes manifests itself in the model as a rise of ψ_2. Thus, I will change the value of ψ_1 and ψ_2 when I use the model to study counterfactuals. Moreover, while I identify ψ_1 and ψ_2 using aggregate moments in Section 3, I provide direct measures regarding the relative magnitude of these two key parameters using micro-level data in Section 5.3.

Fourth, in our definition teacher quality ξ reflects the *relative* position of teachers' human capital in the population. In other words, parallel shifts in the population-wide human capital distribution, i.e., changes in μ, *does not* affect teacher quality ξ. I use this definition to emphasize the selection channel and its amplification mechanisms. Our results will be stronger if teacher quality also depends on the absolute level of teachers' human capital because there will be an additional feedback loop where teacher quality feeds back on average human capital in the population and so on – an channel that leads to economic growth. This intuition is closely related to the "social learning" models presented in Lucas and Moll (2014) and Lucas (2015). Quantifying the impacts of teachers on growth, however, requires additional empirical evidence and is also not the main focus of this paper.

Fifth, the model assumes that all individuals are eligible to become teachers. Given that a bachelor's degree is considered a minimum requirement for full-time teachers in K-12 schools in the U.S., the population being studied fits best to the ones with such level of education. To better match the model to data, I restrict the sample to the corresponding sub-population in the quantitative analyses.[7]

Lastly, I adopt a stylized model of occupation selection. To focus on the propagation mechanism through endogenous human capital formation, I abstract away from some important margins in the context of teacher selection and quality such as the role of risk preferences (Cortes and Pan 2018), dynamic learning (Wiswall 2013, Rothstein 2015), the allocation of teachers across schools or districts (Biasi, Fu and Stromme 2021), public versus private schools (Tincani 2021), and teacher efforts (Bold, Filmer, Martin, Molina, Stacy, Rockmore, Svensson and Wane 2017). I view these possibilities as intriguing potential extensions.

2.4 Model Solution

The share of adults that becomes teachers by human capital h is given by

$$l(h) = \frac{(w\kappa h^{\psi_1-\psi_2})^\theta}{1+(w\kappa h^{\psi_1-\psi_2})^\theta} \approx (w\kappa h^{\psi_1-\psi_2})^\theta$$

where the dispersion of idiosyncratic preference shock θ governs the elasticity of occupation choice to differences in income.

Integrating $l(h)$ across the human capital distribution $F(h)$, the aggregate share

[7]One can think about the workers without college education as coming from a separate subgroup whose human capital is also affected by teacher quality ξ. Then, the implication of changing teacher quality extends to the whole population in the data.

of teachers π is

$$\pi = \int l(h)\, dF(h) = (w\kappa)^{\theta} \cdot \exp\left(\mu(\psi_1 - \psi_2)\theta + \frac{((\psi_1 - \psi_2)\sigma\theta)^2}{2}\right)$$

Using the definition in Equation (1), teacher quality ξ can be solved as

$$\xi = (\psi_1 - \psi_2) \cdot \sigma \cdot \theta \tag{3}$$

Equation (3) is important as it shows that teacher quality depends on three objects. First, the relative skill bias across sectors $(\psi_1 - \psi_2)$ represents the selection mechanism. If the relative return to human capital between teaching and non-teaching occupation rises, teacher quality ξ improves. Second, teacher quality is proportional to the dispersion of human capital among potential teachers σ. This captures the idea that selection is more powerful when agents become increasingly heterogeneous. Last, teacher quality also depends on θ which reflects the elasticity that individuals move across occupations in response to changes in valuations.

A notable property of the model, as shown in Equation (3), is that teacher quality ξ is independent of relative wage w. This seems to run in contradiction to the common idea that more generous compensation attracts higher-quality teachers. To resolve this paradox, it is worth pointing out that there are two ways to increase income for teachers in the model – one is to raise w and the other is to raise ψ_1. An increase in w leads to a *uniform shift* of the $l(h)$ profile, raising employment share π and the relative wage of teachers to non-teachers, but leaving teacher quality ξ unchanged. An increase in ψ_1, on the other hand, leads to a *rotation* of the $l(h)$ profile, resulting in higher teacher quality ξ as well as ambiguous effects on the employment share π and the relative wage of teachers to non-teachers.

The novel part of this framework that distinguishes it from a typical static occupation selection model is the idea that in a dynamic environment, the dispersion of human capital σ is an endogenous object that responds to changes in occupation selection and parental investments. In other words, changes in teacher quality lead to dynamic effects through endogenous σ. In light of this, Equation (3) also provides a decomposition formula that will be used in later analysis:

$$\underbrace{d\log(\xi)}_{\text{change in teacher quality}} = \underbrace{d\log(\psi_1 - \psi_2)}_{\text{change in selection}} + \underbrace{d\log(\sigma)}_{\text{change in h.c. dispersion}} \tag{4}$$

Now I turn to optimal individual decisions. For interior solutions of e, the optimal investment to income ratio is

$$\frac{e}{y} = \frac{\beta(\delta_1 + \delta_2\xi)}{1 + \beta(\delta_1 + \delta_2\xi)} \approx \beta(\delta_1 + \delta_2\xi)$$

Substitute e back into the human capital production function, we have

$$\log(h') = A + \log(\epsilon) + f(\xi; \vec{\delta}) + (\rho + \psi_2(\delta_1 + \delta_2\xi)) \log(h) \tag{5}$$

where $\vec{\delta} = \{\delta_0, \delta_1, \delta_2\}$ and

$$f(\xi; \vec{\delta}) = \delta_0\xi + (\delta_1 + \delta_2\xi) \cdot \log(\beta(\delta_1 + \delta_2\xi))$$

The dynamics of Equation (5) indicate that when $\log(h)$ is normal, $\log(h')$ stays normal. The transition path of human capital distribution is therefore analytically characterized:

$$\begin{cases} \mu' = A + f(\xi; \vec{\delta}) + (\rho + \psi_2(\delta_1 + \delta_2\xi)) \cdot \mu - \sigma_\epsilon^2/2 \\ (\sigma')^2 = (\rho + \psi_2(\delta_1 + \delta_2\xi))^2 \cdot \sigma^2 + \sigma_\epsilon^2 \\ \xi = (\psi_1 - \psi_2) \cdot \sigma\theta \end{cases} \tag{6}$$

The system in (6) has a close relationship to Bénabou (2002) which exploits the properties of the normal distribution to study the relationship between tax and education policies. The key difference here is the third equation in (6) where teacher quality ξ is an endogenous object and enters the evolution of human capital. As a result, changing ψ_1 or ψ_2 has dynamic implications through ξ and endogenous human capital distribution. This mechanism is absent in Bénabou (2002).

Equation (6) also highlights the key role that δ_2 plays in determining whether teacher quality is amplified or dampened when the dispersion of human capital σ is endogenous. Suppose $\psi_1 < \psi_2$ in the baseline economy so that teacher quality $\xi < 0$, then consider a small decrease of ξ and focus on how σ' changes. If $\delta_2 < 0$, then the intergenerational persistence $(\rho + \psi_2(\delta_1 + \delta_2\xi))$ is higher and we have higher σ'. This feeds back to the determination of teacher quality in the next period and leads to further declines in ξ'. On the other hand, σ' will be smaller when teacher quality decreases when $\delta_2 > 0$. Therefore, relating back to the human capital production function in Equation (2), endogenous human capital amplifies (dampens) changes in selection if and only if teacher quality is gross substitute (complement) to parental investments.

3. Identification and Calibration

In this section, I carry out the calibration of the model by proving identification and finding empirical counterparts to the model's predictions.

3.1 Model Identification

The objects that need to be identified include

$$\underbrace{\delta_0, \delta_1, \delta_2, A, \rho, \sigma_\epsilon^2,}_{\text{h.c. technologies}} \underbrace{\kappa, \theta, \beta}_{\text{preferences}}, \underbrace{\psi_1, \psi_2}_{\text{labor market}}, \underbrace{\xi, w, \mu, \sigma^2}_{\text{endogenous objects}}.$$

This list of variables includes both deep parameters and equilibrium objects such as μ and σ^2. While some objects are endogenously determined in the equilibrium, Proposition 1 shows that they serve as key auxiliary variables that facilitates the identification.

The steady-state of the model can be summarized in the following set of equations.

As discussed in Section 2.4, the optimal occupation choice of individuals result in aggregate teacher share

$$\pi = (w\kappa)^{\theta} \cdot \exp\left(\frac{\mu\xi}{\sigma} + \frac{\xi^2}{2}\right) \tag{7}$$

and a formula of teacher quality

$$\xi = (\psi_1 - \psi_2) \cdot \theta\sigma. \tag{8}$$

The ratio of average income between teachers and non-teachers is

$$\frac{\overline{y_1}}{\overline{y_2}} = w \cdot \exp\left(\mu(\psi_1 - \psi_2) + \frac{\sigma^2}{2}(\psi_1 - \psi_2)(\psi_1 + \psi_2 + 2\psi_1\theta)\right). \tag{9}$$

Owing to the log-normality of the human capital distribution, income inequality within occupations is given by

$$\text{CV}(y_1) = \sigma\psi_1 \qquad \text{and} \tag{10}$$

$$\text{CV}(y_2) = \sigma\psi_2. \tag{11}$$

We use coefficient of variations (CV) as the main measure of inequality in this paper because it provides analytical clarity that highlights the key mechanism of the paper. Moreover, as shown by Bendel, Higgins, Teberg and Pyke (1989), when the underlying distribution is log-normal, coefficient of variation and the widely-used

Gini coefficient are highly and positively correlated.

Using the solution to the parent's maximization problem, education expenditure to income ratio is given by

$$\frac{e}{y} = \beta(\delta_1 + \delta_2\xi). \tag{12}$$

For children who become workers when they grow up, the intergenerational elasticity of income (IGE) is given by

$$\text{IGE} = \rho + \psi_2(\delta_1 + \delta_2\xi). \tag{13}$$

Last, utilizing the human capital dynamics in Equation (6), the steady-state mean and variance of human capital in the population can be expressed as

$$\sigma^2 = \frac{\sigma_\epsilon^2}{1 - \text{IGE}^2} \qquad \text{and} \tag{14}$$

$$\mu = \frac{A + \delta_0\xi + (\delta_1 + \delta_2\xi)\cdot\log\left(\frac{e}{y}\right) - \sigma_\epsilon^2/2}{1 - \text{IGE}}. \tag{15}$$

I can measure $\overline{y_1}, \overline{y_2}, \text{CV}(y_1), \text{CV}(y_2), \pi, \text{IGE}, e/y$ in the data - more details in the calibration section. Because the unit of human capital is undetermined, I have the freedom to make one normalization. After normalizing $\psi_2 = 1$, I need more information because there are 14 unknowns with 9 equations.[8]

The key insight that allows me to identify the model in closed form is that I can gather additional information from the model's comparative statics regarding a small change in ψ_1. When the economy is in the steady state, an incremental change in ψ_1 has two sets of effects that can be measured in the data.

First, Equations (16) and (17) quantify how a change in ψ_1 affects individual's income in the next generation. On the one hand, Equation (16) captures the (per-

[8]Normalizing $\psi_2 = 1$ is without loss of generality. It is equivalent to normalizing ψ_1, σ_ϵ, or A.

centage) effects on the average income of workers in the next generation y'. Note that it combines a direct effect coming from changes in teacher quality ξ and an indirect effect capturing parents' responses to changing teacher quality through endogenous investments e. On the other hand, Equation (17) expresses how the effects on children's income depend on parents' current income status y. The expression highlights that the key parameter δ_2 is identified from the differential impacts between rich and poor households.

$$\frac{\partial \overline{\log(y')}}{\partial \psi_1} = \psi_2 \cdot \underbrace{\sigma\theta}_{\frac{\partial \xi}{\partial \psi_1}} \left[\underbrace{\delta_0 + \delta_2 \log(e/y) + \psi_2 \delta_2 \mu}_{\text{direct}} + \underbrace{\beta \delta_2 \cdot \frac{\delta_1 + \delta_2 \xi}{e/y}}_{\text{indirect}} \right] \tag{16}$$

$$\frac{\partial^2 \log y'}{\partial \psi_1 \partial \log y} = \sigma\theta \cdot \psi_2 \cdot \delta_2 \tag{17}$$

The second set of equations quantify how a change in ψ_1 affect income and employment share of contemporaneous teachers, In particular, Equations (18) displays the effect on average income among teachers $\overline{y_1}$. Equation (19) expresses the effect on the teacher share of the labor force π.

$$\frac{\partial \overline{y_1}}{\partial \psi_1} = \overline{y_1} \cdot \big(\mu + \sigma^2(\psi_1 + \theta(2\psi_1 - \psi_2))\big) \tag{18}$$

$$\frac{\partial \pi}{\partial \psi_1} = \pi \cdot (\theta\mu + \sigma\theta\xi) \tag{19}$$

With these additional information, the identification proof is presented next.

Proposition 1: The model is identified up to the calibration of ρ if we observe $\overline{y_1}/\overline{y_2}$, CV$(y_1)$, CV$(y_2)$, π, IGE, e/y and measure the left-hand-sides of Equations (16)-(19).

Proof. Given that $\psi_2 = 1$, Equation (11) identifies σ; then Equation (10) identifies ψ_1; Equation (14) identifies σ_ϵ. Combining Equations (18) and (19) by substituting out μ identifies θ. Then, Equation (8) identifies ξ; Equation (19) identifies μ; Equation (9) identifies w; Equation (7) identifies κ; Equation (17) identifies δ_2; Equation (13) identifies δ_1 given that I calibrated ρ; Equation (12) identifies β; and lastly Equation (15) identifies A. □

3.2 Empirical Evidence

I measure the left-hand-sides of Equations (16)-(19) using the variations coming from the passage of duty-to-bargain (DTB) laws across states between 1960 and 1996.

Prior to 1960, teachers' unions in the United States played little role in the negotiation of contracts between teachers and school district and collective bargaining took place in only some large, urban school districts. Beginning with Wisconsin in 1960, however, states start to pass pubic sector duty-to-bargain (DTB) laws, which mandated that districts have to negotiate in good faith with a union that has been elected for the purposes of collective bargaining. These laws gave considerable power to teachers' unions in the collective bargaining process, leading to a sharp rise in teacher unionization and collectively bargained contracts.

Lovenheim and Willén (2019) examine the (differential) effects of DTB laws on the next generation's human capital and income using individual-level data from the American Community Survey (ACS). They find that a 10-year exposure to teachers under DTB laws reduces annual earnings by 2.36% (averaging men and women). Moreover, compared with Black and Hispanic children, the negative effects are 4.9 percentage points smaller among White and Asian children, whose parents had 60% higher income after adjusting for the number of children (calculated using data from the Current Population Survey (CPS) March Supplement). Similar findings are also reported in Tincani (2021) where the authors finds that tying public school teacher wages to skills and introducing minimum competency requirement for teacher reduce the achievement gap between the poorest and richest 25% of students by a third.

Complementing their findings on children's earnings, I run additional regressions to measure Equations (18) and (19). In particular, I collect data on full-time workers from the CPS-ASEC, categorize whether individuals are teachers or not based on occupation codes, and use standard two-way fixed effects methods to examine the how duty-to-bargain laws affect employment share and teacher earnings statistics

$$Y_{\text{state,year}} = \alpha_2 \cdot \text{DTB}_{\text{state,year}} + \text{State FE} + \text{Year FE} + \varepsilon \tag{20}$$

where observations are at the state-year level. The regression results are shown in Table 2.

Table 2: Regression Results

	(1) Teacher share	(2) CV(teacher income)	(3) Average teacher income
DTB	-0.351	-0.0292	-591.3
	(0.110)	(0.0138)	(425.4)
# Observations	1378	1364	1378

Notes: This table displays the results of regression (20). Standard errors in parentheses. Average teacher earnings is measured in year 2000 dollars.

I find that DTB laws have negative (but statistically insignificant) effects on average earnings of teachers $\overline{y_1}$ while reducing the coefficient of variations of teacher earnings by 0.03 from a base of 0.56. In addition, DTB laws reduce teacher employment share by 0.35 percentage points from a base of 4.6%. These findings are consistent with the literature that uses the recent expiration of collective bargaining agreements in some states as a source of exogenous variation. In particular, Biasi, Fu and Stromme (2021) shows that introducing flexible pay scheme attracted teachers from other school districts, raised the salaries of high-quality teachers, increased teacher pay dispersion measured by the coefficient of variation, and resulted in higher teacher quality.

Overall, the regression results are consistent with the interpretation that these laws mainly affect teacher market by changing the returns to human capital ψ_1. One potential worry is that variations in employment and wages could also reflect other effects of DTB laws, such as changes in relative wages w or entry barriers κ. I provide a detailed discussion about how this possibility affects the identification and our results in Section 5.

To utilize these estimates, I revisit Equations (16)-(19) and rewrite them as:

$$\frac{\Delta\overline{\log(y')}}{\Delta\psi_1} = \text{CV}(y_2)\cdot\theta\left[\delta_0 + \delta_2\log\left(e/y\right) + \psi_2\delta_2\mu + \beta\delta_2\cdot\frac{\delta_1+\delta_2\xi}{e/y}\right] \tag{16'}$$

$$\frac{\Delta^2\log y'}{\Delta\psi_1\Delta\log y} = \theta\delta_2\cdot\text{CV}(y_2) \tag{17'}$$

$$\frac{\Delta\log(y_1)}{\Delta\psi_1} = \mu+\sigma^2(\psi_1+\theta(2\psi_1-\psi_2)) \tag{18'}$$

$$\frac{\Delta\pi}{\pi\Delta\psi_1} = \theta\mu+\sigma\theta\xi \tag{19'}$$

The regression results in Table 2, together with the estimates in Lovenheim and Willén (2019), inform the left-hand-sides of these equations. In addition, the magnitude of changes in ψ_1 can be measured using changes in the coefficient of variations in teachers' earnings after the DTB laws were imposed:

$$\Delta\psi_1 = \psi_1\cdot\frac{\Delta\text{CV}(y_1)}{\text{CV}(y_1)} \tag{21}$$

This equation plays an important role as it provides a direct mapping from duty-to-bargain laws in the data to the magnitude of $\Delta\psi_1$ in the model.

3.3 Calibration Results

I calibrate $\rho = 0.6 \times$ IGE following the results in Lefgren, Sims and Lindquist (2012). It represents a "measure of ignorance" in our understandings of the determinants of intergenerational mobility.[9] Other target moments, including $\frac{\overline{y_1}}{\overline{y_2}} = 0.96$, CV$(y_1) = 0.56$, CV$(y_2) = 0.77$, $\pi = 4.6\%$, IGE $= 0.36$, and $e/y = 0.05$, are derived from the Current Population Survey (CPS), the Panel Study of Income Dynamics (PSID), and the past literature on parental investments in children's education (Lee and Seshadri 2019, Daruich 2018) for years closest to 2000. Because the model is just identified, these data moments will be exactly matched in the calibration procedure following Proposition 1.

Table 3: Calibration Results

	Human capital production	Value		Preference	Value
δ_0	teacher effect (level)	-0.48	κ	teacher cost	0.21
δ_1	investment effect	-0.42	θ	taste shock dispersion	3.07
δ_2	teacher effect (gradient)	-0.88	β	weight on children's h.c.	0.35
A	TFP of h.c. production	1.12			
ρ	residual persistence	0.22			
σ_ϵ	ability shock dispersion	0.72		**Equilibrium objects**	**Value**
			ξ	teacher quality	-0.64
	Labor market	**Value**	w	relative wage	2.18
ψ_1	skill bias (teachers)	0.73	μ	average h.c.	1.15
ψ_2	skill bias (non-teachers)	1	σ	h.c. dispersion	0.77

Notes: This table displays the calibration results. See text for the identification proof and targeted moments.

[9] In fact, parameter ρ can also be identified jointly in closed form if I target teacher-value added in the data. As discussed after calibration, the model generates teacher-value added that is consistent with the empirical estimates by Chetty, Friedman and Rockoff (2014a) as a non-targeted moment.

The parameters of particular interest are those in the human capital production function. Combining the results on $\{\delta_0, \delta_1, \delta_2\}$ with the observed level of e/y, I find that a one standard deviation change of teacher quality throughout ones' education leads to a rise in human capital by 1.6 percent. The implied result from the structural model here is slightly larger but broadly consistent with Chetty, Friedman and Rockoff (2014a), who conclude that "a one standard deviation increase in teacher quality in a single grade raises annual earnings by 1.3 percent." The result that $\delta_2 < 0$ is of particular importance because as discussed before, it implies that the endogenous human capital formation serves as an amplifying mechanism of the teacher selection channel and that public schooling serves as the "great equalizer" in the economy (Agostinelli, Doepke, Sorrenti and Zilibotti 2022). Using the approximation formula in Kmenta (1967), the implied elasticity of substitution between parental investment and teacher quality is around 2 in the calibrated model. This value is consistent with existing estimates, such as 1.92 by Blankenau and Youderian (2015) and 2.43 by Kotera and Seshadri (2017).

4. Counterfactual Analyses

In this section, I analyze two counterfactual to better understand the role of teacher quality and human capital formation in shaping the dynamics of inequality and intergenerational mobility.

4.1 Performance-Based Compensation

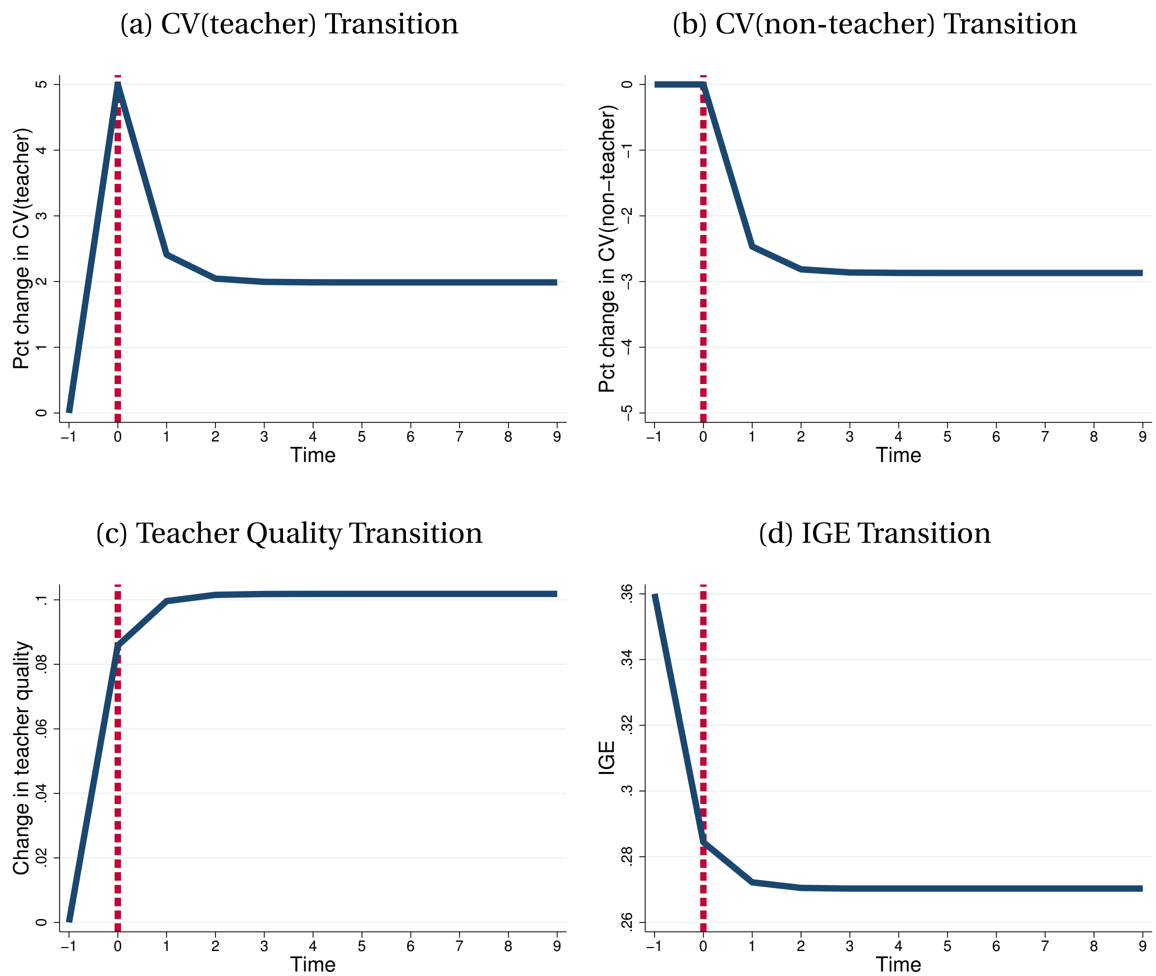
In the first counterfactual, I permanently increase the returns to human capital by 5% in the teaching profession to mimic a performance-based compensation. The magnitude of 5% is chosen to be the same as the initial impacts that duty-to-bargain laws had on the income dispersion among teachers in the empirical findings, albeit in a different direction. This policy rotates the wage-human capital profile among teachers, raising income for teachers with high human capital and reducing income for teachers with low human capital. The policy change occurs at $t = 0$. Figures 1a to 1d shows the transition path of income dispersion among teachers, income dispersion among non-teachers, teacher quality ξ, and IGE from the original economy to the new steady state.

I find that inequality within the teaching profession jumps initially. But in the long-run steady-state, only two-fifth of the initial impacts on income dispersion among teachers survive. In contrast, income dispersion among non-teachers decreases drastically by 3%, reflecting changes in the dispersion of human capital σ. Intergenerational elasticity of earnings falls from 0.36 to 0.27, due to the fact that higher teacher quality diminishes the role of private parental investments and hence intergenerational persistence of human capital.

The results also indicate that the magnitude of these effects vary along the transition path. In particular, empirical analyses of performance-based compensations are only assessing the effects on teacher labor markets at $t = 0$ and the effects on student outcomes at $t = 1$. Results in Figures 1a to 1d show that such estimates will

- overstate the long-run effects on inequality among teachers by more than 100%,
- understate the long-run effects on inequality among non-teachers by 15%,
- understate the long-run improvement of teacher quality by 15%, and
- understate the long-run reduction in IGE by 0.01.

Figure 1: Increasing Return to Human Capital Among Teachers



Notes: These figures plot the transition path of CV(teacher), CV(non-teacher), teacher quality ξ, and IGE following a 5% permanent increase of ψ_1 in the steady-state economy.

4.2 Skill-Biased Technical Change

In the second counterfactual, instead of changing ψ_1, I let ψ_2 rise by 1% permanently to mimic a small skill-biased technical change (SBTC), i.e., the "convexification" of the non-teaching labor market (Autor, Goldin and Katz 2020). Like the case before, the rise in ψ_2 occurs at $t = 0$. Figures 2a to 2d shows the transition path of income dispersion among teachers, income dispersion among non-teachers, teacher quality ξ, and IGE from the original economy to the new steady state.

I find that when ψ_2 increases, it not only raises inequalities within non-teachers but also those within teachers, reflecting changes in human capital dispersion σ. Intuitively, higher returns to human capital in non-teaching occupations attract the most talented workers and reduce teacher quality, with the effects being much larger in the long run (see Figure 2a). Interestingly, intergenerational elasticity of earnings (IGE) increases for two reasons. The first reason is the same as that in the previous counterfactual - reduced teacher quality raises the role of parental investments and hence intergenerational persistence. The second reason is that the dispersion of parental income is also greater due to higher ψ_2, magnifying disparities in parental investments.

To better understand the relative importance of these different channel, I conduct a model-based decomposition of CV(y_2) in the following steps:

1. First, because CV$(y_2) \equiv \sigma\psi_2$, I decompose CV$(y_2)$:

$$\underbrace{\mathrm{CV}(y_2)' - \mathrm{CV}(y_2)}_{\text{changes in inequality}} = \underbrace{\sigma'(\psi_2' - \psi_2)}_{\text{direct effect}} + \underbrace{(\sigma' - \sigma)\psi_2}_{\text{indirect h.c. dist.}}$$

   I find that the direct effect of changing ψ_2 accounts for 48% of total changes in CV(y_2) in the long run, with the rest being explained by changing σ.

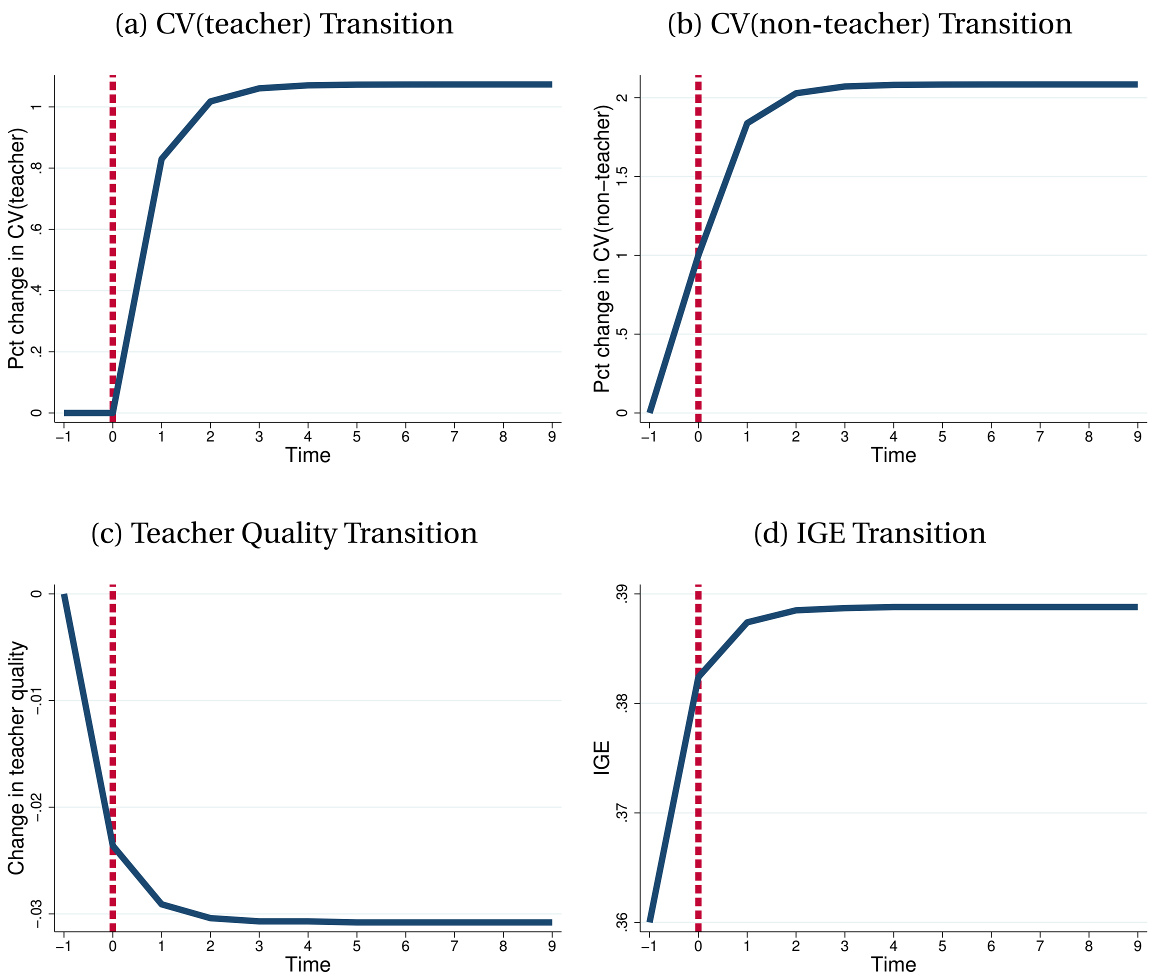
2. To understand what drives changes in σ, I conduct a decomposition of IGE because in the steady-state σ is proportional to $\sqrt{1 - \mathrm{IGE}^2}$. The formula $\mathrm{IGE} = \rho + \psi_2(\delta_1 + \delta_2\xi)$ motivates the following decomposition:

$$\underbrace{\psi_2'(\delta_1 + \delta_2\xi') - \psi_2(\delta_1 + \delta_2\xi)}_{\text{changes in IGE}} = \underbrace{\psi_2\delta_2(\xi' - \xi)}_{\text{teacher quality}} + \underbrace{(\delta_1 + \delta_2\xi')(\psi_2' - \psi_2)}_{\text{parental income}} \tag{22}$$

   I find that teacher quality channel accounts for 94% of changes in IGE with the rest being explained by changes in parental income dispersion.

These decomposition results lead to an intriguing policy implication. Suppose

Figure 2: Skill-Biased Technical Change



Notes: These figures plot the transition path of CV(teacher), CV(non-teacher), teacher quality ξ, and IGE following a 1% permanent increase in ψ_2 in the steady-state economy.

ψ_2 increases by 1%, the government can in principle counteract its contemporaneous effects on teacher quality by also raising ψ_1 by 1%. But as parental income become more dispersed, Equation (22) shows that simply raising ψ_1 by 1% is not enough to maintain IGE at $t = 0$. With a rise in IGE, human capital dispersion σ will still be higher in future generations, kicking off the vicious cycle with deteriorating teacher quality ξ in the future. Therefore, to fully neutralize the effects of the SBTC, the returns to human capital in the teaching profession (ψ_1) needs to rise *relatively more* than that in the non-teaching profession (ψ_2). If instead, policymakers or teacher unions shield teachers from rising inequality elsewhere by reducing ψ_1, income inequalities among non-teachers will become even more severe. In fact, the attempt to maintain income inequality among teachers through a reduction in ψ_1 would also be self-defeating. This is because the dispersion of human capital in the population where teachers are selected from will be rising in future generations.

4.3 Summary

To conclude, the counterfactual analyses highlights the trade-off in income inequalities between teaching and non-teaching labor markets. This trade-off emerges due to the feedback effects through occupation selection and human capital formation channels. While it is not obvious what the optimal policy should be, these findings suggest that the reward structure of teachers matter not only for those who are currently employed in the education system, but also for the public in large, including people who are not yet born.

Moreover, the results also show that teacher labor market serves as a key propagation channel of aggregate shocks to skill returns. While the tax system is perhaps the most direct way to address such shocks, the findings in this paper suggest that another powerful, albeit longer-term, policy instrument is to change the allocation of talents between occupations that produce human capital (teachers) and those that utilizes human capital in production (workers).

5. Robustness

In this section, I discuss the robustness of the main results.

5.1 Mapping DTB Laws to the Model

As discussed before, one potential worry is that variations in employment and wages following DTB laws could also reflect changes in relative wages w or entry barriers κ. In that case, instead of Equations (18') and (19'), we have

$$\begin{cases} \Delta \log(y_1) = \Delta\psi_1[\mu + \sigma^2(\psi_1 + \theta(2\psi_1 - \psi_2))] + \Delta w \\ \frac{\Delta\pi}{\pi} = \Delta\psi_1 \cdot (\theta\mu + \sigma\theta\xi) + \theta(\Delta w + \Delta\kappa) \end{cases} \tag{23}$$

Note that potential changes in w and κ does not affect the interpretation of other equations including (16'), (17'), and (21) because they do not alter teacher quality ξ in the model. They, however, do affect the calibration of other parameters. I divide the analysis into three cases.

First, suppose DTB laws affect entry barriers κ but not relative wage w. Then Equation (18') survives but equation (19') is contaminated because I do not have a direct measure of $\Delta\kappa$ in the data. In this case, I need to calibrate θ outside of the model and choose a value from the literature. An important insight here is that different value of θ does not affect the calibration of other variables except $\{\delta_0, \delta_2, \xi, w, \kappa\}$, and the results in all counterfactuals remain unchanged.

To see this, suppose θ is doubled. To make Equation (17') hold, δ_2 needs to halve. Similarly, δ_0 needs to halve to make Equation (16') hold. ξ needs to double because $\xi = (\text{CV}(y_1) - \text{CV}(y_2)) \cdot \theta$. Lastly, w and κ needs to adjust to maintain Equations (9) and (7). Beyond that, no other parameters need to change because δ_0, δ_2, and ξ enter the rest of the model, especially human capital dynamics, in a multiplicative fashion and these changes exactly cancel out.

Second, suppose DTB laws affect w but not κ. I can multiply the first equation in (23) by θ and subtract the second equation to cancel out Δw. Then, I can proceed like the previous case by calibrating θ exogenously. The results in counterfactuals are not affected.

Last, suppose DTB laws affect w as well as κ. Then the empirical results on how DTB laws affect teacher labor markets are not informative about parameters and I need to find additional calibration targets. One option is to equate the effects of a one standard deviation increase in teacher quality on children's income, which is $(\delta_0 + \delta_2 \log(e/y)) \cdot \psi_1$ in the model, to existing empirical estimates (e.g., Chetty, Friedman and Rockoff 2014a). Then, I can calibrate θ exogenously, and the counterfactual results are again not affected. As discussed before, when I compare teacher value added in the model to Chetty, Friedman and Rockoff (2014a) after calibration, the two values are quite similar.

To sum up, while potential changes in w and κ threaten the identification of θ, the counterfactual results will not be affected if I calibrate θ exogenously outside of the model.

5.2 Heterogeneous Preference

In the baseline model, I have assumed that the idiosyncratic preference shock ν is not correlated with human capital h. Suppose instead there is a systematic relationship between taste and human capital, consider

$$\nu = \psi_3 \log(h) + \tilde{\nu}$$

where ψ_3 governs the degree of correlation and $\tilde{\nu}$ is an idiosyncratic shock.

Then, individual's choice problem becomes

$$\max_{j \in \{1,2\}} \mathbb{1}_{j=1}(\log(wh^{\psi_1} \cdot \kappa \cdot h^{\psi_3}) + \tilde{\nu}) + \mathbb{1}_{j=2} \log(h^{\psi_2})$$

This leads to changes in the formulas of π and ξ where ψ_1 needs to be replaced by $\psi_1 + \psi_3$, which also implies that there is an additional parameter ψ_3 to be calibrated.

One solution is to equate the effects of a one standard deviation increase in teacher quality on children's income, which is $(\delta_0 + \delta_2 \log(e/y)) \cdot (\psi_1 + \psi_3)$ in this case, to existing empirical estimates (e.g., Chetty, Friedman and Rockoff 2014a). This gives an additional condition that allows us to separately identify ψ_3 from ψ_1 (which is identified using the dispersion of teachers' earnings). As discussed before, when I compare teacher value added in the model to Chetty, Friedman and Rockoff (2014a) after calibration, the two values are quite similar. This implies that the main results will not be significantly affected if I allow this form of correlation between preference and human capital.

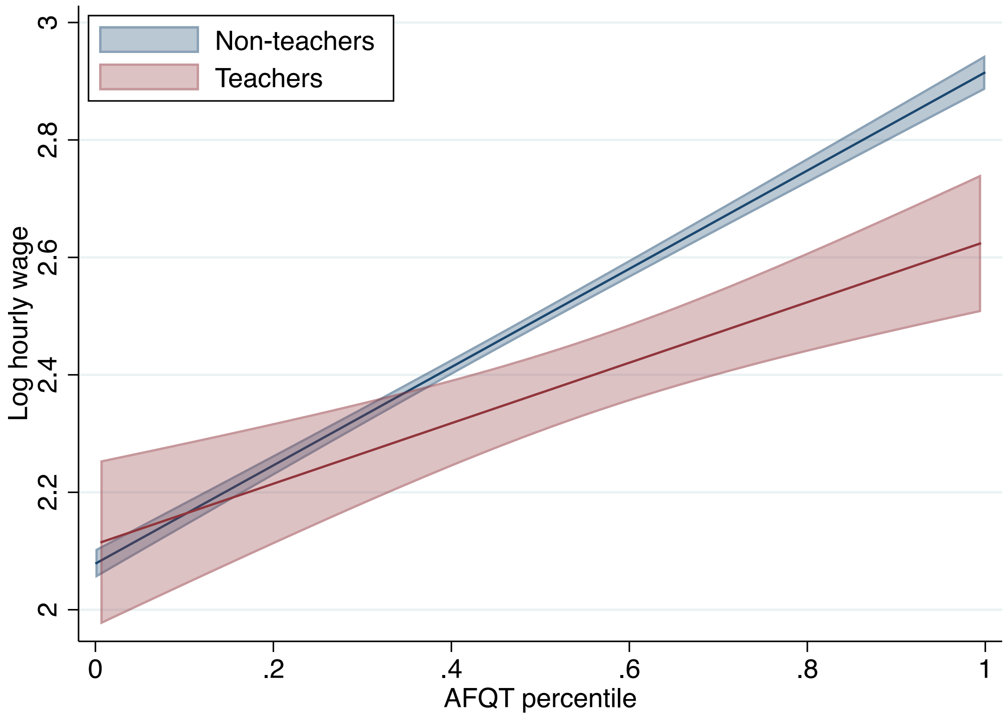
5.3 Quantifying Skill Bias Across Occupations

In Section 3, the two important parameters governing skill-bias across occupations ψ_1 and ψ_2 are identified using aggregate moments on the income dispersion in teaching and non-teaching occupations respectively. In this section, I take a more micro-level approach by directly regressing wage on measures of cognitive ability using the Armed Forces Qualification Test (AFQT) percentile score from the National Longitudinal Study of Youth 1979 (NLSY79) data.

The NLSY79 survey tracks a cohort of individuals aged 14 to 22 when they were initially interviewed in 1979. The survey collects their labor market history, including weeks worked, occupation codes, and hourly wages. For each individual, their cognitive ability was assessed in 1980 through ten intelligence tests known as the Armed Services Vocational Aptitude Battery (ASVAB), and a summarizing measure known as the Armed Forces Qualification Test (AFQT) percentile score was computed. With some caveats, the AFQT score was a commonly-used measure of cognitive skills in the literature (e.g., Neal and Johnson 1996). I restrict the sample to college-educated individuals who worked at least 30 weeks on the primary job last year with an hourly wage of at least one dollar.

Figure 3 plots the relationship between the AFQT score and hourly wage of individuals in the data. As can be seen, for both teachers and non-teachers, hourly wage is positively correlated with the AFQT score, but the correlation is stronger among non-teachers.

Figure 3: Relationship between the AFQT Score and Hourly Wage



Notes: This figure plots the relationship between the AFQT percentile score and log hourly wage in year 1996 across occupations for individuals in the NLSY79 sample. The line plots the best linear fitted value. The shaded area plots the 90% confidence interval around the fitted value.

To show this pattern more systematically, I run the following regressions:

$$Y_{i,t} = \alpha_{j,t} + \Psi_{j,t} \cdot \text{AFQT}_i + \varepsilon_{i,t} \tag{24}$$

where i indexes individuals, $j \in \{1, 2\}$ indexes teachers and non-teachers respectively, t represents survey year, and $Y_{i,t}$ is the log of hourly wage. I use the notation $\Psi_{j,t}$ because the independent variable AFQT_i denotes skill percentiles instead of skill levels, hence the interpretation of the coefficient is a little different from the occupation-specific skill bias $\psi_{j,t}$ in the model.

Table 4: Regression Results

	$t = 1996$		$t = 2006$	
	$j = 1$	$j = 2$	$j = 1$	$j = 2$
$\Psi_{j,t}$	0.515	0.837	0.827	0.927
	(0.113)	(0.024)	(0.122)	(0.029)
# Observations	240	2490	227	2193

Notes: This table displays the results of regression (24). Subscript $j \in \{1, 2\}$ indexes teachers and non-teachers respectively. Standard errors in parentheses.

Table 4 reports the regression results. As can be seen, AFQT percentiles is strongly correlated with hourly wage. For example, a one percentile increase in the ranking of AFQT score is correlated with a 0.515% higher hourly wage. Importantly, the regression results suggest that the coefficient is larger in non-teaching occupations ($j = 2$) than that among teacher ($j = 1$). Across the two waves of data, the ratio $\Psi_{1,t}/\Psi_{2,t}$ is on average 0.754 which is quite close to the ratio of skill bias $\psi_1/\psi_2 = 0.73$ in the model.

To sum up, using more direct measures of ability, empirical estimates using micro-level wage data confirms that teachers have a more compressed wage distribution than non-teachers, translating to a smaller skill bias in the model.

6. Conclusion

In this paper, I study how teacher selection affects inequality and intergenerational mobility. In a model with occupation choice and child investments, I investigate the feedback mechanism between teacher quality and human capital distribution. Human capital dispersion amplifies the selection channel in determining teacher quality, and teacher quality affects the degree of dispersion itself through human capital formation. The model is identified using empirical evidence on how duty-to-bargain laws affect teacher labor market and children's long-run outcomes. In counterfactuals, I find that tying compensation to performance in the teaching profession raises income inequality among teachers but reduces inequality among non-teachers and improves intergenerational mobility. Teachers also play an important role in propagating aggregate shocks. In particular, I find that increasing returns to human capital among teachers could alleviate the effects of skill-biased technical change on inequalities.

Lastly, I would like to suggest that while this paper studies teachers, the same argument holds for other occupations, such as doctors and elected officials, as long as worker quality spills over to the productivity of future generations. These effects can be quantified in the same way through combining structural modeling and empirical analysis. I leave this as interesting future avenues of research.

References

Abbott, Brant, Giovanni Gallipoli, Costas Meghir, and Giovanni L Violante, "Education policy and intergenerational transfers in equilibrium," *Journal of Political Economy*, 2019, *127* (6), 2569–2624.

Acemoglu, Daron, "Reward structures and the allocation of talent," *European Economic Review*, 1995, *39* (1), 17–33.

Agostinelli, Francesco, Margaux Luflade, Paolo Martellini et al., "On the spatial determinants of educational access," Technical Report 2021.

_ , Matthias Doepke, Giuseppe Sorrenti, and Fabrizio Zilibotti, "When the great equalizer shuts down: Schools, peers, and parents in pandemic times," *Journal of Public Economics*, 2022, *206*, 104574.

Autor, David, Claudia Goldin, and Lawrence F Katz, "Extending the race between education and technology," in "AEA Papers and Proceedings," Vol. 110 American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203 2020, pp. 347–351.

Bacolod, Marigee P, "Do alternative opportunities matter? The role of female labor markets in the decline of teacher quality," *The Review of Economics and Statistics*, 2007, *89* (4), 737–751.

Baumol, William J, "Entrepreneurship: Productive, unproductive, and destructive," *Journal of Business Venturing*, 1996, *11* (1), 3–22.

Bénabou, Roland, "Tax and education policy in a heterogeneous-agent economy: What levels of redistribution maximize growth and efficiency?," *Econometrica*, 2002, *70* (2), 481–517.

Bendel, RB, SS Higgins, JE Teberg, and DA Pyke, "Comparison of skewness coefficient, coefficient of variation, and Gini coefficient as inequality measures within populations," *Oecologia*, 1989, *78*, 394–400.

Biasi, Barbara, Chao Fu, and John Stromme, "Equilibrium in the market for public school teachers: District wage strategies and teacher comparative advantage," Technical Report, National Bureau of Economic Research 2021.

Blanden, Jo, Matthias Doepke, and Jan Stuhler, "Educational inequality," in "Handbook of the Economics of Education," Vol. 6, Elsevier, 2023, pp. 405–497.

Blankenau, William and Xiaoyan Youderian, "Early childhood education expenditures and the intergenerational persistence of income," *Review of Economic Dynamics*, 2015, *18* (2), 334–349.

Bold, Tessa, Deon Filmer, Gayle Martin, Ezequiel Molina, Brian Stacy, Christophe Rockmore, Jakob Svensson, and Waly Wane, "Enrollment without learning: Teacher effort, knowledge, and skill in primary schools in Africa," *Journal of Economic Perspectives*, 2017, *31* (4), 185–204.

Card, David, Ciprian Domnisoru, and Lowell Taylor, "The intergenerational transmission of human capital: Evidence from the golden age of upward mobility," *Journal of Labor Economics*, 2022, *40* (S1), S39–S95.

Caucutt, Elizabeth M and Lance Lochner, "Early and late human capital investments, borrowing constraints, and the family," *Journal of Political Economy*, 2020, *128* (3), 1065–1147.

Chetty, Raj, John N Friedman, and Jonah E Rockoff, "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood," *American Economic Review*, 2014, *104* (9), 2633–2679.

_ , Nathaniel Hendren, Patrick Kline, and Emmanuel Saez, "Where is the land of opportunity? The geography of intergenerational mobility in the United States," *The Quarterly Journal of Economics*, 2014, *129* (4), 1553–1623.

Chyn, Eric and Diego Daruich, "An equilibrium analysis of the effects of neighborhood-based interventions on children," Technical Report, National Bureau of Economic Research 2022.

Corak, Miles, "Income inequality, equality of opportunity, and intergenerational mobility," *Journal of Economic Perspectives*, 2013, *27* (3), 79–102.

Corcoran, Sean P, William N Evans, and Robert M Schwab, "Changing labor-market opportunities for women and the quality of teachers, 1957–2000," *American Economic Review*, 2004, *94* (2), 230–235.

Cortes, Patricia and Jessica Pan, "Occupation and gender," *The Oxford Handbook of Women and the Economy*, 2018, pp. 425–452.

Daruich, Diego, "The macroeconomic consequences of early childhood development policies," *FRB St. Louis Working Paper*, 2018, (2018-29).

Dupuy, Arnaud, "A microfoundation for production functions: Assignment of heterogeneous workers to heterogeneous jobs," *Economica*, 2012, *79* (315), 534–556.

Durlauf, Steven N and Ananth Seshadri, "Understanding the Great Gatsby curve," *NBER Macroeconomics Annual*, 2018, *32* (1), 333–393.

Fogli, Alessandra and Veronica Guerrieri, "The end of the american dream? inequality and segregation in us cities," Technical Report, National Bureau of Economic Research 2019.

Fu, Chao, Junjie Guo, Adam J Smith, and Alan Sorensen, "Students' heterogeneous preferences and the uneven spatial distribution of colleges," *Journal of Monetary Economics*, 2022, *129*, 49–64.

Fuchs-Schündeln, Nicola, Dirk Krueger, Alexander Ludwig, and Irina Popova, "The long-term distributional and welfare effects of Covid-19 school closures," *The Economic Journal*, 2022, *132* (645), 1647–1683.

Hanushek, Eric A, "The economic value of higher teacher quality," *Economics of Education Review*, 2011, *30* (3), 466–479.

Hoxby, Caroline M and Sarah Turner, "What high-achieving low-income students know about college," *American Economic Review*, 2015, *105* (5), 514–517.

Hoxby, Caroline Minter, "How teachers' unions affect education production," *The Quarterly Journal of Economics*, 1996, *111* (3), 671–718.

Jacob, Brian A, Jonah E Rockoff, Eric S Taylor, Benjamin Lindy, and Rachel Rosen, "Teacher applicant hiring and teacher performance: Evidence from DC public schools," *Journal of Public Economics*, 2018, *166*, 81–97.

King, Robert G and Ross Levine, "Finance, entrepreneurship and growth," *Journal of Monetary economics*, 1993, *32* (3), 513–542.

Kmenta, Jan, "On estimation of the CES production function," *International Economic Review*, 1967, *8* (2), 180–189.

Kotera, Tomoaki and Ananth Seshadri, "Educational policy and intergenerational mobility," *Review of Economic Dynamics*, 2017, *25*, 187–207.

Lavy, Victor, "Teachers' pay for performance in the long-run: The dynamic pattern of treatment effects on students' educational and labour market outcomes in adulthood," *The Review of Economic Studies*, 2020, *87* (5), 2322–2355.

Lee, Sang Yoon and Ananth Seshadri, "On the intergenerational transmission of economic status," *Journal of Political Economy*, 2019, *127* (2), 855–921.

Lefgren, Lars, David Sims, and Matthew J Lindquist, "Rich dad, smart dad: Decomposing the intergenerational transmission of income," *Journal of Political Economy*, 2012, *120* (2), 268–303.

Lovenheim, Michael F, "The effect of teachers' unions on education production: Evidence from union election certifications in three midwestern states," *Journal of Labor Economics*, 2009, *27* (4), 525–587.

_ and Alexander Willén, "The long-run effects of teacher collective bargaining," *American Economic Journal: Economic Policy*, 2019, *11* (3), 292–324.

Lucas, Robert E, "Human capital and growth," *American economic review*, 2015, *105* (5), 85–88.

_ and Benjamin Moll, "Knowledge growth and the allocation of time," *Journal of Political Economy*, 2014, *122* (1), 1–51.

Murphy, Kevin M, Andrei Shleifer, and Robert W Vishny, "The allocation of talent: Implications for growth," *The Quarterly Journal of Economics*, 1991, *106* (2), 503–530.

Neal, Derek A and William R Johnson, "The role of premarket factors in black-white wage differences," *Journal of political Economy*, 1996, *104* (5), 869–895.

Rivkin, Steven G, Eric A Hanushek, and John F Kain, "Teachers, schools, and academic achievement," *Econometrica*, 2005, *73* (2), 417–458.

Rothstein, Jesse, "Teacher quality policy when supply matters," *American Economic Review*, 2015, *105* (1), 100–130.

Roy, Andrew Donald, "Safety first and the holding of assets," *Econometrica*, 1952, pp. 431–449.

Seshadri, Ananth, "Specialization in education," Technical Report, mimeo, University of Wisconsin 2000.

Tincani, Michela M, "Teacher labor markets, school vouchers, and student cognitive achievement: Evidence from Chile," *Quantitative Economics*, 2021, *12* (1), 173–216.

Wiswall, Matthew, "The dynamics of teacher quality," *Journal of Public Economics*, 2013, *100*, 61–78.

Yum, Minchul, "Parental time investment and intergenerational mobility," *International Economic Review*, 2023, *64* (1), 187–223.