A Quantitative Theory of Teacher Quality

Evidence from a Dynamic Structural Model

Estimated using an Incentive Pay Experiment

PRELIMINARY

Nirav Mehta

November 3, 2025

Teacher quality is an important determinant of the distribution of student achievement. Researchers have documented considerable cross-sectional variation in the quality of even novice teachers, but this variation is not well-explained by teachers' observable characteristics. Although the determinants of teacher quality are not yet well understood, researchers have documented the effects of output-based incentives on teacher quality and substantial growth in teacher quality over teachers' careers. I combine classic human capital and asymmetric information frameworks to develop an estimable dynamic model of teacher quality. Teacher human capital may be generated by on-the-job investments and/or by learning-by-doing. The model allows for variation in teachers' initial human capital (*hidden types*) as well as for teachers to make unobserved effort inputs (*hidden actions*). I then estimate the model using experimental variation from Muralidharan and Sundararaman (2011), a teacher incentive pay experiment in Andhra Pradesh. Preliminary results indicate that hidden human capital types explain the lion's share of quality variation, and that on-the-job investments are responsible for the majority of teacher quality growth. Under the incentive scheme in Muralidharan and Sundararaman (2011), the increase in teacher quality that took place during the intervention is attenuated by the reduction in future quality.

Keywords: human capital, teacher quality, on-the-job training, learning-by-doing

JEL codes: I2, I28, J2, J24, J45, C1

University of Western Ontario, `nirav.mehta@uwo.ca`, `https://niravmehtax.github.io`. I thank Tian Liu for research assistance, and Lance Lochner, Rory McGee, Abhijeet Singh, Jeff Smith, Todd Stinebrickner, and Chris Udry for their comments. I gratefully acknowledge funding from the SSHRC Insight Development Grant Program.

1. Introduction

Teacher quality, typically measured by a teacher's value added to student achievement, is an important determinant of student achievement and economic growth (Rivkin, Hanushek, and Kain 2005; Hanushek 2011; Chetty, Friedman, and Rockoff 2014). While the research and policy focus has typically been on cross-sectional variation in teacher quality, which is largely modeled as unobserved heterogeneity, researchers have also documented substantial growth in quality over teachers' careers (Harris and Sass 2011; Wiswall 2013; Papay and Kraft 2015).

While the growth in quality over teachers' careers is clearly important, little is known about *why* it occurs. Economists have developed two main theories to explain how workers accumulate human capital, which is then used to produce output: *On-the-Job Training* (OJT) and *Learning-by-Doing* (LBD). In the pure OJT model, workers (in the current application, this will be teachers) allocate work hours away from production (here, classroom teaching) to invest in their human capital (here, teachers' formal or informal professional development).[1] Notably, the pure OJT model implies a tradeoff between current and future production because of the *multitask problem* (Hölmstrom and Milgrom 1991). No such tradeoff is necessarily present under the pure LBD model, wherein workers accumulate human capital simply by producing output (i.e., classroom teaching).[2]

Given the importance of teacher human capital accumulation, it is important to distinguish between OJT and LBD. Unfortunately, the relative importance of these forces in determining teacher quality is unknown. Policies that provide incentives for teachers to augment their human capital could be more effective policy instruments than those currently popular among researchers and policymakers, which operate via selection of more productive teachers by selecting better "hidden types" (Rivkin, Hanushek, and Kain 2005; Hanushek 2011) or introduce performance-based incentive pay schemes, which explicitly incentivize teachers to increase student achievement by eliciting higher levels of "hidden actions" from teachers (Muralidharan and Sundararaman 2011; Barlevy and Neal 2012). For example, if OJT is important, a performance-based incentive pay scheme could result in an overall reduction of student achievement—*even if it increased current achievement*. This is because teachers' work hours would be diverted from human capital investments, reducing future teacher quality and thus reducing the achievement of future cohorts of students.

Prior empirical research has adopted "either-or", or "pure", specifications, only allowing for either OJT or LBD as the force generating human capital. As noted by Killingsworth (1982), there are few general testable implications of OJT vs. LBD models. This can largely be attributed to a lack of the right data. Intuitively, separating these forces requires a factor that differentially

[1] See Becker (1962) and Ben-Porath (1967). For an example of an OJT-intensive job, consider being a student in school.

[2] See Weiss (1972), Killingsworth (1982), and Shaw (1989). For an example of a LBD-intensive job, consider "chicken sexing", where workers develop an ability to determine the sex of baby chickens through their experience sorting chicks (McWilliams 2018).

affects the returns to investing in one's own human capital for different workers. It may not be surprising then, that attempts to distinguish between the OJT and LBD models have been hampered by equilibrium re-adjustments that occur in private employer-employee contexts (Rosen 1972; Heckman, Lochner, and Cossa 2002).

In this paper, I develop and estimate a structural model of teachers' life-cycle human capital accumulation, which nests the "either-or" OJT and LBD specifications used by prior researchers. The model also features unobservable heterogeneity in human capital initial conditions (hidden types) and unobserved time allocation choices (hidden actions), all in one coherent framework.[3] A structural approach is necessary to answer many important policy questions about the design of alternative teacher remuneration schemes without having to conduct innumerable experimental interventions, which are costly in terms of money and crucially, given the importance of teacher quality, the lag time until results are collected. In particular, the model allows for the two types of asymmetric information about teacher quality considered in the literature, hidden types ("adverse selection") and hidden actions ("moral hazard").

I estimate the model using data from Muralidharan and Sundararaman (2011), a teacher pay-for-performance experiment. Experimental introduction of performance-based pay for teachers provides the needed variation in returns to human capital investments. Further, labor prices are often fixed in the education sector (e.g., teachers are typically paid according to public salary schedules[4]) and worker-specific output is observed by researchers, even if noisily so (i.e., teacher value added), making teacher quality trajectories a particularly attractive context for discriminating between OJT and LBD forces.

This paper yields several important substantive contributions about the nature of teacher quality and the design of teacher payment contracts. First, I use a nested framework to quantify the importance of OJT and LBD components to human capital accumulation. In the likely case that both OJT and LBD are present, "either-or" models would necessarily be misspecified, under- or overstating the roles played by the forces. This would lead to incorrect policy guidance, so the results will be directly useful to policymakers seeking to increase student achievement with limited resources. Further, because this has not been done for any type of worker, this contribution is even more useful because the model could also be adapted for other applications.

This paper also provides a quantitative understanding of the relative importance of the two types of asymmetric information potentially underlying teacher quality: hidden types (i.e., "adverse selection") and hidden actions (i.e., "moral hazard"). This emerges naturally from the structural model developed to understand the dynamics of teacher quality, yet is novel and of considerably policy importance. For example, screening for more effective teachers (as

[3]In the contract theory literature, models with "hidden actions" may generate "moral hazard". While models with "hidden types" are capable of generating "adverse selection", I use the former term because I do not model selection into teaching. That being said, the distribution of hidden types will naturally affect the optimal reimbursement contract.

[4]See Podgursky and Springer (2011) for an extensive discussion of this point.

proposed by Hanushek 2011) is the appropriate type of mechanism if most of the asymmetric information about teacher quality is immutable (derived from hidden types), but output-based incentive pay will likely be more effective in the presence of unobserved effort inputs (i.e., hidden actions). Because both are likely present (see the discussion of Dinerstein and Opper 2022, below) it is important to understand the relative importance of these types of asymmetric information.

This paper also brings new evidence to the debate about the relative magnitudes of early and late-career growth in teacher quality. Recent research finding growth throughout teachers' careers (Harris and Sass 2011; Wiswall 2013; Papay and Kraft 2015) has relaxed functional form assumptions made by earlier research (see, e.g., Rivkin, Hanushek, and Kain (2005) or the summary in Hanushek and Rivkin (2006)), which typically found little evidence of growth beyond the first few years of a teacher's career. This debate may persist, however, because the findings are based on observational, not experimental, data. By using exogenous variation in incentives to invest in human capital (generated by the pay-for-performance experiment), this paper can provide credible estimates of the average growth in teacher quality over teachers' careers. Another substantive contribution made possible by this paper's innovative methodology is to quantify *heterogeneity* in growth rates of teacher quality. While such heterogeneity would be difficult to measure in the statistical frameworks used by prior literature, it emerges naturally from this paper's theoretical framework: If teachers possess differing levels of initial human capital, they should also face different incentives to invest in their human capital.

Most relevant literature

This paper builds on a rich literature measuring changes in teacher quality over their careers, notably Hanushek and Rivkin (2006), Wiswall (2013), and Papay and Kraft (2015). Conceptually, the most closely related work is the spare literature seeking to understand the importance of OJT and LBD in generating human capital for workers more generally, namely Heckman, Lochner, and Cossa (2002) and Blandin (2018). Both papers develop and estimate/calibrate quantitative lifecycle models allowing for either OJT or LBD; neither allows for both forces in a nested framework.[5]

This paper is also closely related to some recent work using models to better understand teacher quality dynamics. Dinerstein, Megalokonomou, and Yannelis (2022) use administrative variation in teachers' wait time until receiving job assignments and a production model to estimate teacher human capital depreciation and returns to experience for Greek teachers, assuming a linear, time-invariant relationship between experience and achievement. Linearity may be reasonable in their setting because most teachers in their data have fewer than five years

[5]This paper is also related to work using structural models to study human capital accumulation using either pure OJT models (see, e.g., Brown 1976; Haley 1976; Heckman 1976; Heckman, Lochner, and Taber 1998; Taber, Seshadri, and Fan 2024; Fu, Lin, and Tanaka 2021) or pure LBD models (see, e.g., Blinder and Weiss 1976; Shaw 1989; Imai and Keane 2004).

of experience, so the estimates summarize early-career effects rather than a full lifecycle profile. Also closely related is Dinerstein and Opper (2022), who study the effect of screening policies on teacher quality, within a novel multitasking framework where teachers are characterized by fixed, two-dimensional, types, which measure the returns to effort on targeted and untargeted tasks. Finally, Mehta (2025) develops a partial identification approach that leverages follow-up data from a teacher incentive pay experiment in Kenya (Glewwe, Ilias, and Kremer 2010) to yield qualitative findings about the importance of OJT investments and LBD in generating teacher human capital. The current paper goes beyond simple estimation of the technology of human capital accumulation by estimating a behavioral model, which can be used to predict responses to compensation scheme changes as well as allow for normative analysis.

A small but substantively important set of papers examines other channels underlying teachers' improvement over their careers. Ost (2014), which measures the returns to teachers' general and grade-specific experience, finds both to be important determinants of teacher quality growth. Cook and Mansfield (2016) extends this work to also allow for general and context-specific permanent components to teacher quality. The current paper nicely complements these papers by explicitly viewing teacher quality through the lens of the main conceptual frameworks for human capital, and by identifying and separating the OJT and LBD channels of human capital accumulation, which are not necessary to distinguish given these other papers' goals.[6] This paper also complements the extensive literature studying teacher quality more generally, recently discussed in Hanushek and Rivkin (2006); Jackson, Rockoff, and Staiger (2014); Strøm and Falch (2020), and the sub-literature on teacher incentives (see, e.g., Hanushek and Raymond 2005; Muralidharan and Sundararaman 2011; Imberman and Lovenheim 2015; Petronijevic 2016).[7]

2. Background

I use publicly available data from Muralidharan and Sundararaman (2011), which featured an RCT conducted in 2005-2006 in Andhra Pradesh to study output-based teacher bonus pay. I use data from the control and individual incentive (hereafter, "treatment") arms, which cover 501 teachers teaching 25,656 students at 196 schools. Muralidharan and Sundararaman (2011) found sizable effects of incentive pay on student achievement, and a battery of robustness checks was consistent with the increases in measured achievement not being due to "teaching to the test."

The data effectively provide a two-year snapshot for teachers at different stages of their

[6]These papers may be viewed as part of the substantial empirical literature studying task-specific human capital for more general worker contexts. For recent examples, see Poletaev and Robinson (2008); Sanders (2010); Yamaguchi (2012); Robinson (2018). Sanders and Taber (2012) contains a detailed discussion of this and other extensions of the one-dimensional human capital model.

[7]Although less tightly related, this paper also relates to the literature studying teacher labor markets (see, e.g., Dolton and Klaauw 1999; Stinebrickner 2001; Behrman et al. 2016; Tincani 2021; Biasi 2021; Bobba et al. 2021; Biasi, Fu, and Stromme 2021).

careers. The key outcome variables in control and treatment groups are measures of quality and labor supply for teachers during the intervention. A teacher's quality is measured by her value added to student achievement. A teacher's labor supply, or total work time, is measured by the fraction of site visits by research team during which she was present. This is similar to total hours at work, used in other papers estimating lifecycle models of human capital accumulation. Section 5 describes how the data are mapped to the model.

3. Model

This section considers decisions of a teacher at the outset of her teaching career, which lasts T periods (years). She enters the first period with initial human capital θ_1, which varies across teachers.[8] In each period t, the teacher notes her compensation schedule (which depends on her treatment status) and her class size, and then chooses how to allocate a unit of time between leisure, l_t, and work. When at work, she splits her time (and thus, her human capital) between: (i) production (of student achievement), h_t, (ii) on-the-job investment (OJT) in her own human capital, I_t, and (iii) "on-the-job leisure" (OJL), or *shirking*, s_t.

3.1. Technologies and Constraints

Human capital is generated via the production function

$$\theta_{t+1} = [1 - \delta_\theta]\theta_t + \Delta\theta(\underbrace{\theta_t[1 - l_t]I_t}_{\text{OJT}}, \underbrace{\theta_t[1 - l_t]h_t}_{\text{LBD}}), \tag{1}$$

where θ_t denotes human capital available for use in period t and $\delta_\theta \in [0, 1]$ denotes its depreciation rate. The amount of human capital produced in period t is given by $\Delta\theta(\cdot, \cdot)$, where $I_t \in [0, 1]$ denotes the share of current human capital allocated to on-the-job investment and $h_t \in [0, 1]$ denotes the share of current human capital allocated to production, which also represents the LBD input to human capital accumulation. Production, on-the-job investment, and on-the-job leisure are rival, as captured by the time constraint

$$h_t + I_t + s_t = 1. \tag{2}$$

As in the classic human capital frameworks (Ben-Porath 1967; Killingsworth 1982), a teacher's output (here, quality), in period t equals the amount of her human capital used for production:

$$q_t = \theta_t[1 - l_t]h_t, \tag{3}$$

which is commonly applied to all students in the class.

[8] For estimation I extend the model to accommodate multiple education levels. Naturally, a teacher's education level may affect her initial human capital, θ_1, as well as career length, T, as potential working years spent in formal education would shorten the horizon until retirement.

Quality is measured using standardized achievement tests, which a large body of research has found to contain considerable noise (see, e.g., Papay 2011; Rockoff et al. 2012; Jackson, Rockoff, and Staiger 2014). This noisy measure of teacher quality, observed by the Principal (who determines bonuses when there is incentive pay) and the econometrician, is

$$m_t = q_t + \bar{\epsilon}_t, \tag{4}$$

where $\bar{\cdot}$ denotes the class average and the distribution of $\bar{\epsilon}_t$, $F_{\bar{\epsilon}_t}$, depends on class size.[9]

The presence of a direct, albeit noisy, output measure departs from the typical OJT or LBD setup, wherein researchers use earnings to back out human capital and time allocated to production, given that skill prices have been identified.[10] Here, h_t cannot be backed out from the output measure even if θ_1 were degenerate and (s_t, l_t) were observed. This generates an informational asymmetry that is not present in classical models of human capital accumulation, a natural assumption in light of the considerable noise in achievement measures.

3.2. Preferences

Flow utility in period t depends on quality, compensation, and both forms of leisure:

$$u_t = u_q\big(\alpha_q q_t + \alpha_w^* w(m_t; d_t) + \alpha_\chi \chi_t\big) + u_l(l_t) + u_s(s_t, l_t), \tag{5}$$

where d_t indicates whether a teacher is actively treated in period t, α_w^* measures the teacher's valuation of additional compensation based on their performance, and $w(\cdot; d_t)$ captures how teacher pay depends on the measure of student achievement, which depends on treatment status. In the control group, where $d_t = 0$ for all t, we have $w' = 0$. The term χ_t represents the teacher's salary income. The (potential) dependence of teacher pay on the quality measure m_t introduces uncertainty via the measurement error.

Starting with the first argument of u, the coefficient α_q on quality allows teachers to directly value growth in student achievement. This is quite natural in the education setting, where workers (here, teachers) are thought to be at least in part intrinsically motivated to produce output (i.e., teach their students) (see, e.g., Dixit 2002; Besley and Ghatak 2005). Moreover, it can explain why teachers—who, like any other worker, should value leisure—choose positive h when not explicitly incentivized to do so.[11] If $\alpha_w^* > 0$, then teachers value the remuneration stemming from output bonuses, which is consistent with a positive effect from a pay-for-performance scheme wherein $w' > 0$.

[9] Section 5.2 shows how this measure relates to the test score data.

[10] See, e.g., Heckman (1976); Heckman, Lochner, and Taber (1998); Huggett, Ventura, and Yaron (2006). See also Bowlus and Robinson (2012), who argue that skill prices may not be fixed and are not trivial to identify, as well as Rosen (1972); Heckman, Lochner, and Cossa (2002) for discussions of the difficulties involved with identifying skill prices if they reflect learning opportunities under LBD.

[11] A positive h is evinced by the non-degenerate distribution of quality.

Moving on to the second and third arguments of u, the dependence of utility on leisure (via u_l) allows for labor supply responses (i.e., changes in $1 - l_t$).[12] Finally, the amount of on-the-job leisure (s_t) may also affect teacher utility (u_s). On-the-job leisure could rationalize changes in quality (conditional on leisure) in response to the introduction of incentives, among teachers already choosing negligible investment (for example, teachers at the ends of their careers). While leisure is observable, on-the-job leisure is not, meaning s_t acts as a static moral hazard effort variable. I allow the valuation of OJL to depend on "off-the-job" leisure, l_t.

It may be useful at this point to discuss how the specification for teacher preferences compares to commonly used specifications in the OJT literature. For simplicity, suppose leisure were fixed. In the typical OJT setup, workers pay for investments by reducing production, which reduces consumption of market goods.[13] Here, under the status quo (i.e., absent the intervention), remuneration does not depend on output. However, if teachers have an intrinsic valuation of students' learning an analogous tradeoff emerges, except that by investing the teachers are decreasing their "consumption" of student achievement growth. A similar tradeoff emerges when considering the shirking decision: by shirking the teacher enjoys "on-the-job leisure" at the cost of reduced "consumption" of student achievement growth.

Teachers in the treatment group in Muralidharan and Sundararaman (2011) during the intervention received a bonus linear in average student achievement growth that year, yielding $w(m_t; 1) = w'_{MS} \cdot m_t$ (where the coefficient on m_t measures the increase in earnings from a unit increase in average student achievement growth).[14] This means that when estimating the effect of the incentive scheme we can model $\alpha_w^* \cdot w(m_t; d_t)$ using $d_t \alpha_w m_t$ without loss of generality, where the reduced-form parameter $\alpha_w = \alpha_w^* w'_{MS}$ absorbs the incentive strength. Thus we can then write utility as

$$u_t = u_q\big([\alpha_q + d_t\alpha_w] \cdot q_t + d_t\alpha_w\bar{\epsilon}_t + \alpha_\chi\chi_t\big) + u_l(l_t) + u_s(s_t, l_t). \tag{6}$$

3.3. Teacher's Problem

The teacher maximizes her present discounted value of lifetime utility, from her first year of teaching ($t = 1$) until retirement (which occurs immediately after period T).

The teacher acts as if she has perfect foresight about her treatment status in future periods.

[12]Note that labor supply includes the share of work time during which the teacher chooses to engage in OJL.

[13]Workers may also pay for investments using market goods; because investments via market goods are difficult to directly observe, the observed earnings are assumed to be computed net of this (see, e.g., Heckman 1976), which may effectively operate as workers paying for investments solely in terms of foregone production (see Taber, Seshadri, and Fan 2024).

[14]The bonus was designed to avoid "threshold effects" that could emerge from discontinuities in marginal incentives, meaning it was effectively linear in average student achievement growth (Muralidharan and Sundararaman 2011, p. 51).

Let $\mathbf{d_t}$ denote a vector of indicator variables describing the teacher's beliefs about treatment status d_τ in each $\tau \geq t$. If the teacher was in the control group then $d_t = 0, \forall t$. The intervention was a surprise, so if the teacher was in the treatment group she acts as if $d_t = 0$, until the experiment occurs by surprise.[15] The period index t denotes the year of the teacher's career. This generates identifying variation because teachers from different cohorts will be in different t's in the first year of the intervention.

A teacher's salary and class size are also required to solve her problem. Teacher salaries χ_t are determined by published schedules, and are largely explained by a teacher's age.[16] Therefore, I use this relationship to generate the vector $\boldsymbol{\chi_t}$ collecting the teacher's salary χ_τ in each $\tau \geq t$. Class size (i.e., the number of observations per teacher-year) enters the teacher's problem by controlling the precision of the quality measure. As with treatment status, let $\mathbf{n_t}$ denote the vector describing the teacher's class size n_τ in each $\tau \geq t$. I assume n_τ equals a teacher's actual class size when available (which is only the case during the intervention) and otherwise sets $n_\tau = \bar{n}$, her average class size in the sample.[17] I collect these last two "state variables" in $\boldsymbol{\Omega_t} = \left[\boldsymbol{\chi_t}, \mathbf{n_t}\right]$.

It will be convenient to work with the share of work time not spent on OJT that is spent on OJL, $\pi_t \in [0, 1]$, which gives $s_t = [1 - I_t]\pi_t$ and, when combined with the budget constraint, $h_t = [1 - I_t][1 - \pi_t]$. The recursive formulation of the teacher's problem is

$$
\begin{aligned}
(7) \qquad V_t(\theta_t; \mathbf{d_t}, \boldsymbol{\Omega_t}) = & \max_{(I_t, l_t, \pi_t) \in [0,1]^3} \mathbb{E}_t\left[u_t + \rho V_{t+1}(\theta_{t+1}; \mathbf{d_{t+1}}, \boldsymbol{\Omega_{t+1}})\right] \\
& \text{s.t. for } t = 1, \ldots, T \\
& \quad \theta_1 \text{ given} \\
& \quad u_t = u_q\left([\alpha_q + d_t\alpha_w] \cdot q_t + d_t\alpha_w\bar{\epsilon}_t + \alpha_\chi\chi_t\right) + u_l(l_t) + u_s([1 - I_t]\pi_t, l_t) \\
& \quad \theta_{t+1} = [1 - \delta_\theta]\theta_t + \Delta\theta(\theta_t[1 - l_t]I_t, \theta_t[1 - l_t][1 - I_t][1 - \pi_t]) \\
& \quad q_t = \theta_t[1 - l_t][1 - I_t][1 - \pi_t] \\
& \quad \bar{\epsilon}_t \sim F_{\bar{\epsilon}_t} \\
& \quad V_{T+1} = 0,
\end{aligned}
$$

where ρ is the discount rate.

[15] For example, if the first year of the (two-year) intervention occurs in year t then $\mathbf{d_t} = [1, 1, 0, \ldots, 0]$.

[16] The R^2 of a linear regression of teachers' monthly salaries on teacher age is 0.63. Using the estimated relationship, the expression for predicted salary (in Indian rupees, Rs.) given teacher age is $-4249.5 + 342.3 * \text{age}_t$.

[17] In principle, one could explicitly model future class sizes as random variables. In practice, however, there are only two years of data available for each teacher, and the lion's share of class-size variation occurs between, not within, teachers. The specification used here allows beliefs about class size to differ across teachers, and for a teacher's realized class size to affect her decisions.

TABLE 1. Notation Summary

State variables	
θ_t	Teacher human capital at time t
d_t	Incentive group indicator
χ_t	Teacher salary
n_t	Teacher's class size
Choices	
l_t	Leisure (off-the-job) share of total time
I_t	On-the-job investment share of work time
π_t	Fraction of non-investment work time allocated to OJL
$h_t = [1 - I_t][1 - \pi_t]$	Production (LBD) share of work time
$s_t = [1 - I_t]\pi_t$	On-the-job leisure (OJL) share of work time
Preferences	
u_q	Valuation of composite quality/compensation index c_t
	$c_t := [\alpha_q + d_t\alpha_w] \cdot q_t + d_t\alpha_w\bar{\epsilon}_t + \alpha_\chi\chi_t$
u_l	Valuation of leisure time, l_t
u_s	Valuation of on-the-job leisure, s_t
α_q	Weight on quality in the quality/compensation index
α_w	Valuation/slope for performance pay used in estimation
α_χ	Weight on salary income (fixed to 12 as salary data are monthly)
ρ	Per-period discount factor
Technology	
δ_θ	Human capital depreciation rate
λ	OJT share in CES human-capital production function $\Delta\theta$
ξ, ν	Scale and substitution parameters in human-capital production
Outcomes	
$q_t = \theta_t[1 - l_t]h_t$	Teacher quality
$m_t = q_t + \bar{\epsilon}_t$	Teacher quality measure

3.3.1. Model Solution

The necessary conditions for interior solutions of (7) are

$$(8)\quad l_t : \frac{\partial u_l}{\partial l_t} = \underbrace{\theta_t[1 - I_t][1 - \pi_t]}_{-\frac{\partial q_t}{\partial l_t}} \cdot \frac{\partial\, \mathbb{E}_t\left[u_q | d_t\right]}{\partial q_t} + \frac{\partial u_s}{\partial l_t} + \rho \frac{\partial \Delta\theta}{\partial\left[[1 - l_t]\right]} \cdot \frac{\partial\, \mathbb{E}_t[V_{t+1}]}{\partial \theta_{t+1}}$$

(9) $$I_t : \rho \frac{\partial \Delta\theta}{\partial I_t} \cdot \frac{\partial\, \mathbb{E}_t[V_{t+1}]}{\partial \theta_{t+1}} = \underbrace{\theta_t[1-l_t][1-\pi_t]}_{-\frac{\partial q_t}{\partial I_t}} \cdot \frac{\partial\, \mathbb{E}_t\left[u_q | d_t\right]}{\partial q_t} + \underbrace{\pi_t}_{-\frac{\partial s_t}{\partial I_t}} \cdot \frac{\partial u_s}{\partial s_t} + \rho \frac{\partial \Delta\theta}{\partial h_t}[1-\pi_t] \cdot \frac{\partial\, \mathbb{E}_t[V_{t+1}]}{\partial \theta_{t+1}}$$

(10) $$\pi_t : \underbrace{[1-I_t]}_{\frac{\partial s_t}{\partial \pi_t}} \cdot \frac{\partial u_s}{\partial s_t} = \underbrace{\theta_t[1-l_t][1-I_t]}_{-\frac{\partial q_t}{\partial \pi_t}} \cdot \frac{\partial\, \mathbb{E}_t\left[u_q | d_t\right]}{\partial q_t} + \rho \underbrace{\frac{\partial \Delta\theta}{\partial h_t}[1-l_t][1-I_t]}_{-\frac{\partial \Delta\theta_t}{\partial \pi_t}} \frac{\partial\, \mathbb{E}_t[V_{t+1}]}{\partial \theta_{t+1}},$$

where an envelope condition returns the marginal value human capital:

(11) $$\frac{\partial\, \mathbb{E}_t[V_{t+1}]}{\partial \theta_{t+1}} = \frac{\partial\, \mathbb{E}_t\left[u_{t+1} | d_{t+1}\right]}{\partial q_{t+1}} \cdot \frac{\partial q_{t+1}}{\partial \theta_{t+1}} + \rho \frac{\partial \theta_{t+2}}{\partial \theta_{t+1}} \frac{\partial\, \mathbb{E}_t[V_{t+2}]}{\partial \theta_{t+2}},$$

for $t \leq T-1$.[18]

3.4. Discussion

The model can capture several salient features of the literature and current policy discussion of teacher quality. First, there is variation in teachers' unobserved human capital, which could generate cross-sectional variation in teacher quality even conditional on input choices. Second, the model allows teachers to respond to the introduction of a pay-for-performance scheme and allows me to quantify the importance of teacher compensation relative to intrinsic motivation. The model allows for OJL, i.e., a static moral hazard/effort margin, which means it can generate substantial quality responses even if OJL is negligible and labor supply is not responsive. Third, if there is an OJT component to human capital accumulation and on-the-job investments are positive, then incentivizing current production (as would occur with performance-based contracts) would naturally generate a dynamic version of the multitask problem exposited in Hölmstrom and Milgrom (1991); any such effect would be tempered by the presence of an LBD component to human capital accumulation.

3.4.1. Identification Intuition

Human capital production. The key variation for identifying OJT versus LBD comes from heterogeneous responses by teachers at different stages of their careers. For simplicity, fix the scale of $\Delta\theta$ and first consider the case without OJL. First suppose the OJT component dominates. This means that younger workers, who invest substantially more in their human capital, have more room for investment reductions. Hence, they will have positive contemporaneous quality improvements, given their labor supply. Older workers can mainly respond through decreasing their leisure time, as they have minimal human capital investments. If their leisure doesn't

[18] The marginal value of human capital in period T is zero.

respond to the intervention, there should be a negligible quality effect among older workers. Now suppose instead that the LBD component dominates. Because investment is negligible, the main response margin is through leisure time. Younger workers are already choosing leisure time to take into account LBD on future human capital, which means there would be a smaller total work time, and, hence, quality, response compared to older workers.

The essential intuition behind identification of the unobserved OJT input and its effect on human capital is as follows. The human-capital production function defines a stationary mapping from the current state (θ_t, I_t, l_t) to the next state θ_{t+1}. Although investment I_t is unobserved (and OJL is set aside for now), it is influenced by factors that do not enter the technology directly: (i) the incentive to invest declines with age, as the horizon shrinks, and (ii) the performance pay scheme raises the cost of investment through its effect on the quality-compensation index in u_q. Hu and Xin (2024) show that when an instrument z_t affects an unobserved discrete action but not the state transition, the joint distribution of the next state and the instrument can be used to recover both the transition rule and the conditional choice probabilities for the unobserved action. The number of identifiable investment levels corresponds to the support of the instrument z_t. In my setting, θ_t is itself unobserved, but I observe (rather, I can identify)[19] $r_t = \frac{q_t}{1-l_t} = \theta_t[1 - I_t]$, which varies with θ_t only through $(1 - I_t)$. Once the transition density $f(r_{t+1} \mid r_t, I_t)$ is recovered via mixture inversion of $f(r_{t+1} \mid r_t, z_t)$ (using instrument-induced shifts in $\Pr(I_t \mid r_t, z_t)$ and the exclusion of z_t from the transition), it can be mapped into the $\theta_{t+1} \mid \theta_t, I_t$ space using the identified distribution of $I_t \mid r_t, z_t$. In principle, non-experimental lifecycle variation in investment behavior could serve this role, as teacher age provides an excluded source of variation that shifts I_t without directly entering the technology. This differs from prior work on identifying the determinants of human capital accumulation (Heckman, Lochner, and Cossa 2002), where identification could not be obtained due to the endogeneity of prices. However, the experimental intervention is still useful, as it effectively doubles the number of identifiable actions by providing exogenous variation in both teacher age and treatment assignment.

On-the-Job Leisure (OJL). OJL can help fit any quality effects among the oldest teachers, conditional on labor supply, as teachers will not be undertake substantial OJT investments later in their careers.

4. Data

As in Muralidharan and Sundararaman (2011), I pool Math and Telugu test scores to measure student achievement. The raw score (% correct answers) has a mean of 35 percentage points and a standard deviation of 25 percentage points. Test scores were normalized to have mean

[19]Technically, I treat $r_t \equiv q_t/(1 - l_t)$ as identified (not observed) from (m_t, l_t, n_t) via the measurement model $m_t = q_t + \bar{\epsilon}_t$ with known $F_{\bar{\epsilon}_t}$ (which depends on n_t).

zero and SD one in control group. In the model, quality will be in terms of percent correct (i.e., quality needs to be nonnegative), but here I will present results in terms of the standardized measure (which has no bearing on the substantive results, as it is an affine transformation of quality).

4.1. Descriptive Statistics

Table 2 summarizes the student-year-level data. The data are balanced across years one and two of the intervention, as well as the control and treatment groups. The current outcome achievement test score ("s_t_nts_level_mean") has a standard deviation of around 1 and a mean of a little above zero (recall that this score was standardized to have a mean of zero and standard deviation of one in the control group); the prior score is "s_t_1_nts_level_mean". The variables "s_SC", "s_ST", and "s_OBC" indicate whether the student is a member of a scheduled caste, scheduled tribe, or "other backwards caste" (which is an official designation), all of which would represent access to fewer resources; in contrast, "hh_affluence_index" and "parent_literacy_index" reflect more available resources.

TABLE 2. Summary of Student-Year Data

	Obs.	Mean	St. Dev.	Min.	p25	p50	p75	Max.
year	71400	1.543	0.498	1	1.000	2.000	2.000	2
incentive	71400	0.497	0.500	0	0.000	0.000	1.000	1
s_t_nts_level_mean	49040	0.098	1.019	-2.277	-0.703	-0.015	0.788	4.720
s_t_1_nts_level_mean	50572	0.039	0.863	-2.111	-0.531	0.000	0.377	5.162
s_Male	57508	0.499	0.500	0	0.000	0.000	1.000	1
s_SC	57484	0.197	0.398	0	0.000	0.000	0.000	1
s_ST	57484	0.029	0.167	0	0.000	0.000	0.000	1
s_OBC	57484	0.588	0.492	0	0.000	1.000	1.000	1
hh_affluence_index	53786	3.530	1.425	0	3.000	4.000	5.000	7
parent_literacy_index	53740	1.114	1.035	0	0.000	1.000	2.000	4

There are 25,656 unique students.

Table 3 presents descriptive statistics for teachers' first years in sample. There is considerable variation in teacher age (t_age), which is crucial to separating OJT from LBD. In terms of education (t_educ_*), most teachers have at least a BA. Table 4 presents descriptive statistics for teachers, over both years. The main variables to note in this table are the number of times teachers were present during site visits (t_num_times_present_at_visit) and the day before site visits (t_num_times_present_day_before_visit). Compared to the number of site visits (t_visits_this_year), we can see teachers were present about 80% of the time.

TABLE 3. Summary of Teacher-Level Data (first year)

Statistic	N	Mean	St. Dev.
t_male	501	0.575	0.495
incentive	501	0.501	0.500
t_age	501	37.587	8.789
t_age_group_1	501	0.062	0.241
t_age_group_2	501	0.375	0.485
t_age_group_3	501	0.385	0.487
t_age_group_4	501	0.178	0.383
t_educ_years	501	4.756	1.741
t_educ_10th	501	0.032	0.176
t_educ_12th	501	0.152	0.359
t_educ_BA	501	0.601	0.490
t_educ_MS	501	0.204	0.403
t_educ_other	501	0.012	0.109

Teacher age group cutpoints are 0, 25, 35, 45, 60.

TABLE 4. Summary of Teacher-Level Data (both years)

	Obs.	Mean	St. Dev.	Min.	p25	p50	p75	Max.
year	935	1.48	0.50	1	1	1	2	2
t_age	935	38.21	8.48	18	32	38	43	58
t_male	935	0.58	0.49	0	0	1	1	1
t_educ_years	935	4.82	1.70	0	5	5	5	7
incentive	935	0.49	0.50	0	0	0	1	1
t_num_obs_year1	918	67.26	28.89	14	46	64	84	172
t_num_obs_year2	885	85.98	40.24	10	58	82	104	330
t_num_obs_avg_year	935	76.36	30.43	14	53	74	96	225
has_only_year_one	935	0.05	0.23	0	0	0	0	1
has_only_year_two	935	0.02	0.13	0	0	0	0	1
t_visits_this_year	934	5.00	1.03	1	4	6	6	6
t_num_times_present_at_visit	932	3.95	1.49	0	3	4	5	6
t_num_times_present_day_before_visit	934	4.10	1.37	0	3	4	5	6

There are 501 unique teachers.

Teacher quality. Table 5 describes quality in the control and treatment groups, for teachers of different ages.[20] Column (1) presents coefficients from a regression of student test scores on student characteristics and an indicator for being in the incentive group. Average quality, measured by the intercept, is (-0.292 x 25) + 35 = 28% correct, and the coefficient on incentive shows the average treatment effect on quality is 1/5 SD (recall the test scores had been standardized to have a standard deviation of 1), which corresponds to an increase of 0.204 x 25 = 5% correct.

Column (2) adds teacher age group indicators and incentive-teacher-age-group interactions, where control teachers in the youngest age group are the reference category. Consistent with other findings in the literature (see, e.g., Wiswall 2013), quality increases over careers by ~0.15 SD of test scores, which is around 1/3 SD of quality. Also consistent with the literature, and with the human capital framework (Becker 1962; Ben-Porath 1967), the increases are most pronounced earliest in teachers' careers, tapering off at the end. The incentive interactions with teacher age group show a large quality increase among the least-experienced teachers (the coefficient on incentive, which measures the effect among the youngest teachers, is 0.386 much higher than the average among all teachers presented in Col. (1)), and that the quality increases decline in teacher experience.

Labor supply. In contrast to quality, there is no clear pattern in labor supply among control teachers, nor is there a discernible average labor supply response. To maintain the focus on patterns to fit, the labor supply results are in the Appendix.[21] If anything, there is weak evidence of a reduction in labor supply among the oldest teachers.

Implications for modeling. Summing up, quality exhibits clear increases over teachers' careers as well as substantial treatment effects, with a noticeable gradient with respect to teacher age group. Neither type of pattern is evident for labor supply. Modeling labor supply is still important, however, because the LBD channel (when there is no OJL) operates solely through changes in labor supply. If the descriptive labor supply results miss a response that did indeed exist, assuming labor supply was fixed would unfairly penalize the "pure LBD" specification. Moreover, there could be heterogeneity in labor supply responses, not picked up by these descriptive statistics.

The descriptive results also suggest it might be prudent to allow for OJL. The average treatment effect on quality among the oldest teachers was essentially null. If there had indeed been a real reduction in these teachers' labor supply, these teachers would need another margin of adjustment while at work (otherwise they would have clear reductions in quality). OJL allows

[20] See Appendix Table A1 for the estimates using only the control group, which are essentially the same as those in Column (1) of Table 5.

[21] Table A2 shows there is no clear pattern with respect to teacher age in the control group. Table A3 shows there is also no discernible average labor supply response (e.g., the coefficient on "incentive" in Col. (1) is not significantly different from zero).

for this.[22]

5. Empirical Implementation

5.1. Empirical Specification

Utility. Preferences are CARA in the output/compensation index and isoelastic (CRRA) in the leisure terms:

$$u_t = -\exp\left(-\eta\left[[\alpha_q + d_t\alpha_w]q_t + d_t\alpha_w\bar{\epsilon}_t\right]\right) + \alpha_l\frac{l_t^{1-\gamma_l}}{1-\gamma_l} + \alpha_s\frac{[S(s_t, l_t)]^{1-\gamma_s}}{1-\gamma_s}, \tag{12}$$

where S is a CES aggregator of s_t and l_t, with scale fixed at 1, substitution parameter ν_s, and OJL share parameter λ_s. The assumed specifications yield a convenient closed-form expression for expected utility under a linear (e.g., the observed) bonus scheme. See Appendix A for details.

Initial human capital. Initial human capital θ_1 is treated as unobserved heterogeneity from the econometrician's perspective. The main identifying assumption is that $d_t \perp\!\!\!\perp \theta_1$, which is a natural, if stronger, assumption if one is convinced of the validity of the experimental intervention (which pertains to mean independence). The initial human capital distribution is parameterized according to

$$\ln\theta_1 \sim \mathrm{N}\left(\mu_{\ln\theta_1}, \sigma^2_{\ln\theta_1}\right).$$

The distribution of θ_1 is constant across cohorts.

Human capital production. The function governing growth in human capital, $\Delta\theta$, is CES, which results in the below law of motion for human capital:

$$\begin{aligned}\theta_{t+1} &= [1-\delta_\theta]\theta_t + \zeta\left[\lambda_I\cdot[\theta_t[1-l_t]I_t]^\nu + (1-\lambda_I)\cdot\left[\theta_t[1-l_t][1-I_t][1-\pi_t]\right]^\nu\right]^{\frac{\xi}{\nu}} \\ &= [1-\delta_\theta]\theta_t + \zeta\left[\theta_t[1-l_t]\right]^\xi\left[\lambda\cdot I_t^\nu + (1-\lambda)\cdot\left[[1-I_t][1-\pi_t]\right]^\nu\right]^{\frac{\xi}{\nu}},\end{aligned} \tag{13}$$

where human capital brought to work, $\theta_t[1-l_t]$, generates heterogeneity in the marginal products of human capital inputs.

[22] Indeed, Table A4 shows the gradient in achievement treatment effects by teacher age seems present, even conditioning on teacher labor supply. While conditioning on teacher presence, a choice variable, complicates interpretation of the coefficients, these estimates do seem to suggest the entire margin of older teachers' quality responses does not operate via labor supply. Therefore, it may be useful to allow for OJL. A different assessment uses percentage changes in quality and labor supply among the oldest teachers. Mean quality among the oldest teachers dropped by about 6%, while their labor supply dropped by almost twice as much (11%), suggesting the presence of a countervailing force.

TABLE 5. Teacher Quality Treatment Effects

	s_t_nts_level_mean	
	(1)	(2)
(Intercept)	-0.292***	-0.436***
	(0.018)	(0.033)
s_t_1_nts_level_mean	0.536***	0.537***
	(0.005)	(0.005)
s_Male	0.020*	0.019*
	(0.009)	(0.009)
s_SC	-0.009	-0.001
	(0.015)	(0.015)
s_ST	-0.049	-0.047
	(0.033)	(0.033)
s_OBC	0.032*	0.038**
	(0.013)	(0.013)
hh_affluence_index	0.054***	0.051***
	(0.003)	(0.003)
parent_literacy_index	0.041***	0.042***
	(0.005)	(0.005)
incentive	0.204***	0.386***
	(0.009)	(0.047)
t_age_group: 2		0.149***
		(0.030)
t_age_group: 3		0.151***
		(0.030)
t_age_group: 4		0.178***
		(0.032)
t_age_group: 2 & incentive		-0.072
		(0.049)
t_age_group: 3 & incentive		-0.202***
		(0.049)
t_age_group: 4 & incentive		-0.458***
		(0.052)
Estimator	OLS	OLS
N	34,849	34,849
R^2	0.263	0.270

Note: OLS regressions of standardized test score, pooled over years and subjects. Regressions include student characteristics and, in column (2), teacher age group indicators and interactions with incentive. Teacher age group cutpoints are 0, 25, 35, 45, and 60.

Other variables. Teacher i's horizon, T_i, depends on her years of formal education. Teachers with 10 or 12 years of education (i.e., not having more than a high school degree) are assumed to start teaching at age 18, and retire after age 60. Additional years of education shorten the horizon.

5.2. Mapping the Data to the Model

Quality / Student achievement. Quality is denominated in the percent correct answers, and is measured via a value added model for student achievement (Hanushek 1971; Murnane 1975):

$$y_{ijt} = q_{it} + x'_{jt}\beta_x + \epsilon_{ijt} \tag{14}$$

where y_{ijt} measures student j's current achievement, q_{it} is teacher i's quality in period t, x_{jt} measures student j's characteristics in year t (and includes prior y), and ϵ_{ijt} is an IID, ex-post, shock with distribution $\mathrm{N}\left(0, \sigma_\epsilon^2\right)$.[23] Averaging over the teacher's n_{it} students yields the noisy measure of quality entering the teacher's problem:

$$m_{it} \equiv \bar{y}_{it} - \bar{x}'_{it}\beta_x = q_{it} + \bar{\epsilon}_{it}, \tag{15}$$

where $\bar{\cdot}$ denotes the average, and the distribution of $\bar{\epsilon}_{it}$, $F_{\bar{\epsilon}_{it}}$, depends on the number of students according to $\bar{\epsilon}_{it} \sim \mathrm{N}\left(0, \sigma_\epsilon^2/n_{it}\right)$.

Labor supply / Teacher presence at site visits. The number of times the teacher is present during the random site visits in each year is modeled as a binomial random variable, with success probability equal to the labor supply, $1 - l_{it}$.[24]

5.3. Estimation

I estimate parameters using an estimation criterion function that is the weighted sum of the log likelihood and an indirect inference criterion. The likelihood is comprised of the likelihoods of observed student achievement and teacher presence at site visits. The likelihood of the observed (standardized) test score for teacher i's student j in year t is

$$\mathcal{L}^y_{ijt} = \frac{1}{\left[\frac{\sigma_\epsilon}{25}\right]}\phi\left(\frac{y_{ijt} - \left[\frac{q^*_{it}-35}{25} + x'_{jt}\left[\frac{\beta_x}{25}\right]\right]}{\left[\frac{\sigma_\epsilon}{25}\right]}\right),$$

[23] To assess the appropriateness of this specification, I first estimate β_x using a regression model with teacher-year fixed effects. The most important student characteristics are prior score, followed by household affluence and parental literacy indices. Coefficient estimates without teacher FE are virtually identical to those with teacher FE. Therefore, the estimates hereafter will control for student characteristics but not teacher-year FE.

[24] I allow the probability for being present the day before a site visit to scale this probability via a parameter.

where ϕ denotes the standard normal density and q^*_{it} is the model quality for teacher i in year t, which has been demeaned (the mean was 35 percent) and divided through by the standard deviation in the control group (25 percentage points).

The probability teacher i was present at pa_{it} out of visits$_{it}$ site visits in year t is

$$\mathcal{L}^{pa}_{it} = \binom{\text{visits}_{it}}{pa_{it}} \left[1 - l_{it}\right]^{pa_{it}} \times l_{it}^{\text{visits}_{it} - pa_{it}},$$

and the probability the teacher was present pb_{it} times the day before the site visits is analogous, but with an extra parameter ϖ_b governing the mapping from labor supply to the number of successes:

$$\mathcal{L}^{pb}_{it} = \binom{\text{visits}_{it}}{pb_{it}} \left[\varpi_b[1 - l_{it}]\right]^{pb_{it}} \times \left[1 - [\varpi_b[1 - l_{it}]]\right]^{\text{visits}_{it} - pb_{it}}.$$

The log likelihood function is the product of the likelihoods for observed achievement and number of times the teacher was present the day of and the day before site visits. The indirect inference portion of the estimation criterion has as its main targets the treatment effects by teacher age group. The final estimates will be obtained by minimizing an indirect inference criterion that has been augmented to include the score of the log likelihood function as targets.[25]

5.4. Computation and solution

Numerical solution. The leisure, investment, and OJL share of post-investment human capital choices are modeled as continuous. As is well-known, the possible existence of an LBD component leads to a nonconcave lifecycle problem. Therefore, I solve each teacher's problem using a global optimization routine.[26] I first draw teacher types from the initial conditions distribution. Then I solve for the path $\{l_{it}, I_{it}, \pi_{it}\}_{t=1,\ldots,T_i}$ for all teachers i, and keep only t overlapping with 2005-2006 for estimation.[27]

In practice, the optimal path can be obtained by optimizing (7) or, when relevant, by solving the necessary conditions in the system (8)-(10). There is only one shock, to measured quality. This simplifies the model solution, as the expectation operator does not need to be carried through to V_{t+1}.

5.4.1. Other computational and data details

Almost 90% of teachers are present for both years of the intervention (the difference in being present for both years by the treatment indicator is insignificant, which is consistent with the low

[25]I avoid doing this until the estimates have converged, as the score is very expensive to calculate.

[26]I solve and estimate the model using Julia (Bezanson et al. 2017) and R (R Core Team 2022). I use the NLopt optimization library (Johnson 2007), and a combination of L-BFGS (Liu and Nocedal 1989) and (the NLopt variant of) Subplex (Rowan 1990) optimization algorithms to solve the model, and a combination of Subplex (Rowan 1990), BOBYQA (Powell 2009), and controlled random search (Price 1983) for estimation.

[27]I simulate the model 33 times per teacher.

attrition level and lack of differential attrition reported in the original study (Muralidharan and Sundararaman 2011, p. 54)). For teachers observed in only the first year of the intervention (10% of teachers), the model is solved as if the teacher had expected to have been in the same treatment status in the second year. For teachers observed in only the second year of the intervention (3% of teachers), the model is solved as if the teacher had been in the same treatment status the previous year.

6. Estimation Results

The estimation results show the model can capture key patterns in the data, and can also yield some policy guidance.

6.1. Parameter Estimates

This section discusses selected preliminary parameter estimates (see Table 6 for the remaining estimates). I start with the human capital production function, reproduced here for convenience:

$$\theta_{t+1} = [1-\delta_\theta]\theta_t + \zeta\big[\theta_t[1-l_t]\big]^{\xi}\Big[\lambda\cdot I_t^{\nu} + [1-\lambda]\cdot\big[[1-I_t][1-\pi_t]\big]^{\nu}\Big]^{\frac{\xi}{\nu}}.$$

I estimate the OJT share in the technology to be substantial: $\hat{\lambda} = 0.62$ and the OJT and LBD inputs to be complements ($\hat{\nu} = -12.2$). I also estimate depreciation in teacher human capital of $\hat{\delta}_\theta = 0.025$, which is consistent with numbers for human capital depreciation for workers more generally (see, e.g., Kuruscu 2006). For comparison, Dinerstein, Megalokonomou, and Yannelis (2022) estimate 4.3% annual depreciation using variation in non-teaching spells, which would represent a lower bound to the extent that teachers invest in their human capital while awaiting assignment.

Moving on to teacher preferences, also reproduced here for convenience:

$$u_t = -\exp\Big(-\eta\big[[\alpha_q + d_t\alpha_w]q_t + d_t\alpha_w\bar{\varepsilon}_t + \alpha_\chi\chi_t\big]\Big) + \alpha_l\frac{l_t^{1-\gamma_l}}{1-\gamma_l} + \alpha_s\frac{\Big[\big[\lambda_s\cdot s_t^{\nu_s} + [1-\lambda_s]\cdot\big[[1-I_t][1-\pi_t]\big]^{\nu_s}\big]^{\frac{1}{\nu_s}}\Big]^{1-\gamma_s}}{1-\gamma_s}.$$

I estimate the intrinsic weight on quality to be $\hat{\alpha}_q = 502$, which implies a valuation of average quality for teachers in the control group of $\hat{\alpha}_q \times \bar{q}_{\text{cont.}}$ = ₹14,910, which is considerable when compared to the teachers' average annual salary of around ₹100,000.

The estimate of the incentive slope in this paper's quality units is $\hat{\alpha}_w = 665$, which implies the average quality improvement among teachers in the incentive group increased income by ₹4,227.

Moving on to valuation of nonwork and on-the-job leisure, I estimate that labor supply is inelastic ($\hat{\gamma}_l = 11.7$), which is consistent with the small average labor supply response in the data, while OJL is more responsive ($\hat{\gamma}_s = 5.1$).

TABLE 6. Parameter estimates

Parameter	Estimate
Preferences	
η	2.44223E-08
α_q	501.7702631
α_w	665.141633
α_l	2.21657E-13
α_s	3.50861E-08
γ_l	11.68967281
γ_s	5.059281318
ν_s	0.620825952
λ_s	0.758996941
Technology	
δ_θ	0.025379782
ζ	0.234874019
ξ	0.726357132
ν	-12.1936743
λ	0.619816799
Initial human capital	
$\mu_{\ln\theta_1}$	3.706381434
$\sigma^2_{\ln\theta_1}$	0.228141094

6.2. Model Fit

Table 7 shows that the model does a reasonable job of fitting the average quality level in the control group (left two columns) overall, and by teacher age group. It also matches the average treatment effect on quality (though it is somewhat overstated), as well as the declining gradient in the treatment effect on quality with respect to teacher age group (right two columns).

Table 8 shows that the model can capture the probability a teacher will be present during a site visit both overall and by teacher age group for the control group (left two columns), but overstates the treatment effect of the intervention on total work hours (right two columns).[28] That being said, labor supply fits better than it did in a specification without any OJL.

7. Quantitative Findings

The estimated model can provide a variety of novel positive and normative results about teacher quality and compensation. I first use the model in Section 7.1 to quantify the importance of labor supply, investment choices, and OJL in explaining lifecycle quality under the status quo

[28] I find a similar pattern for teachers' being present the day before a site visit.

TABLE 7. Teacher Quality

	Control group		Treatment effect	
	Data	**Sim**	**Data**	**Sim**
overall	28.357	29.787	5.106	6.034
age_group 1	24.62	24.114	9.662	11.909
age_group 2	28.388	28.628	7.861	8.703
age_group 3	28.389	31.953	4.613	4.246
age_group 4	29.077	29.061	-1.789	2.088

Note: Quality is denominated in % correct, which has a mean of 35 and SD of 25 in the control group in year 1. Teacher age group cutpoints are 25, 35, and 45.

TABLE 8. Teacher Present During Site Visit

	Control group		Treatment effect	
	Data	**Sim**	**Data**	**Sim**
overall	0.784	0.84	-0.005	0.01
age_group 1	0.773	0.836	0.071	0.018
age_group 2	0.759	0.839	0.029	0.009
age_group 3	0.793	0.843	-0.005	0.01
age_group 4	0.811	0.838	-0.086	0.011

Note: Teacher age group cutpoints are 25, 35, and 45.

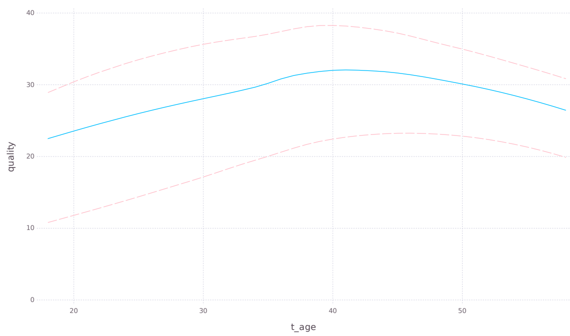
(control group), and then examine how the determinants of quality respond to the observed incentives in MS2011 (Section 7.2). I then use the estimated model to decompose teacher quality into within- and between-teacher components, as well as a portion attributable to the observed incentive scheme (Section 7.3).

I then use the model to perform out-of-sample analysis, which can inform policymakers about effects of other incentive schemes or scenarios of interest. In Section 7.4 I use the model to examine potential dynamic effects of MS2011 incentive cessation. Finally, in Section 7.5 I illustrate how informational asymmetries result in suboptimal allocations by comparing outcomes with those under the assumption of full information.

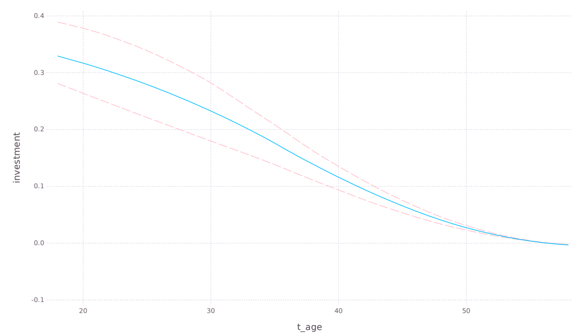
7.1. Understanding Status Quo (Control Group) Quality

I start by describing how key model variables behave along the lifecycle for teachers in the control group. Figure 1A presents quality by teacher age. First looking at the mean (in blue), teacher quality peaks in the early 40s. Notably, quality variation remains substantial over teachers' careers (the 25th and 75th percentiles are in red), a finding that emerges naturally from current framework, which is quite novel and not possible to analyze when studying average growth in teacher quality (as is done in the literature). Of course, since quality is only a fraction of human capital, the quality distribution is compressed compared to the distribution of teacher human capital.

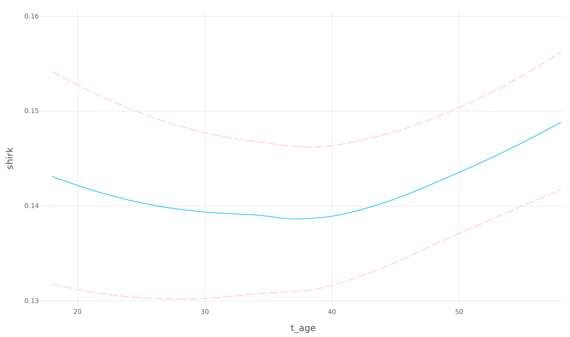
FIGURE 1. Model Outcomes for Teachers in the Control Group, by Teacher Age



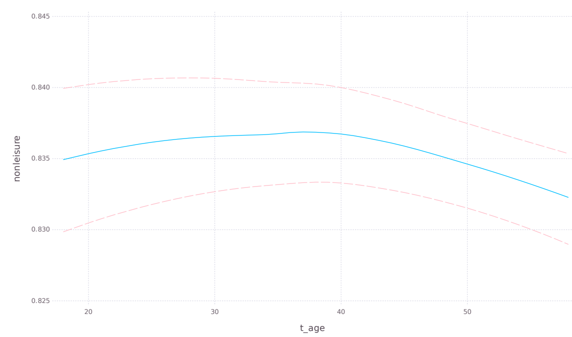
A. Quality



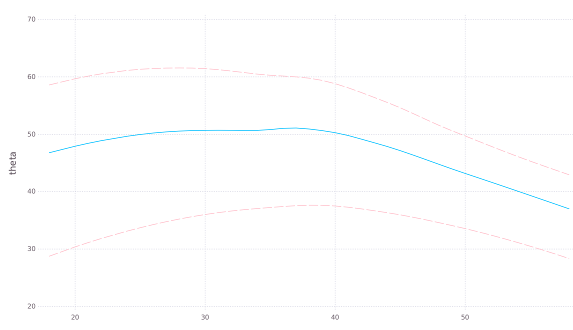
B. Investment share of work time I_t



C. Shirking/OJL share of work time s_t



D. Labor supply



E. Human capital

Notes: Figure plots simulated model outcomes for teachers in the control group. Investment is the share of labor supply, or total work time, spent on investment (I_{it}), on-the-job leisure (OJL) is the share of labor supply, or total work time, spent enjoying OJL (s_{it}), total work time is the share of total time not spent on leisure ($[1 - l_{it}]$), and quality ($\theta_{it}[1 - l_{it}][1 - I_{it}][1 - \pi_{it}]$) is denominated in % correct answers, which has a mean of 35 and SD of 25 in the control group in year 1. The red dashed lines denote the 25th and 75th percentiles.

FIGURE 2. How work human capital is allocated over the career

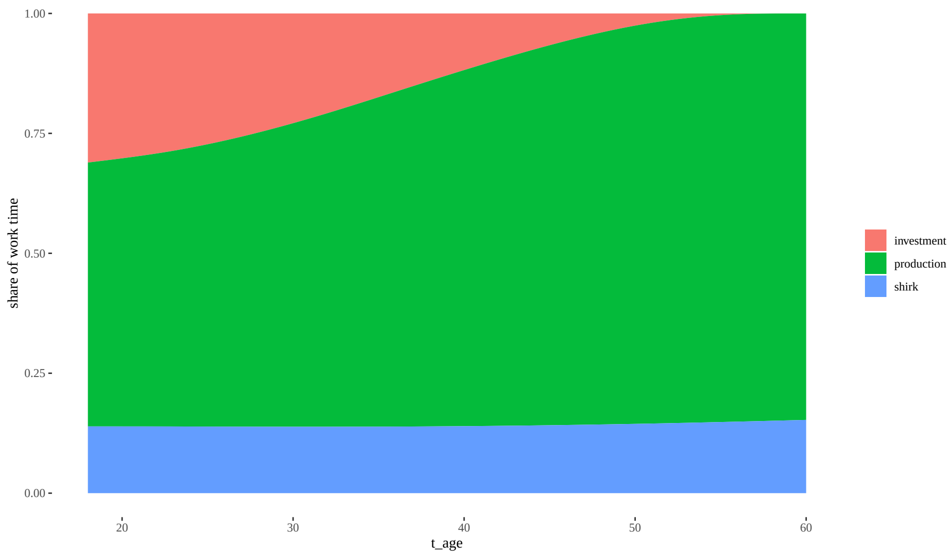


Figure 2 shows the average allocations of work human capital among I_{it}, h_{it}, s_{it} over teachers' careers. As expected, OJT investment I_{it} (shaded red) decreases over teachers' careers. Investment for the youngest teachers is about a third of work human capital, and naturally decreases over the career. While investment is of course unobservable, one can compare this with the estimate of between 50-60% in Heckman, Lochner, and Taber (1998). Figure 1B shows there is much more variation in investment behavior between ages than within ages.

The blue region of Figure 2 shows how the share of work human capital (un)allocated to OJL, s_{it}, slightly increases on average over teachers' careers, and Figure 1C show that it exhibits considerable variation throughout careers. Finally, the green region of Figure 2 shows the share of work human capital allocated to production increases with age, due to the decline in OJT investment.

Figure 1D shows total work time or labor supply $1 - l_{it}$ by age, again for teachers in the control group. Time at work increases and peaks early in teachers' careers, but is fairly flat. The red dashed lines denote the 25th and 75th percentiles here and in other figures, and show that variation in total work time is small, relative to its level. Finally, Figure 1E shows how human capital evolves over teachers' careers. Human capital increases slowly, and peaks before quality, again, due to declining OJT investment. We can also see that variation in human capital decreases as OJT investments become negligible toward the end of teachers' careers.

7.2. Understanding Quality Treatment Effects

I now describe simulated treatment effects, which are computed within teacher. First, incentive pay increased mean quality by 6.3 percent, or ≈ 45% SD of control group quality, which corresponds to about 1/4 SD of control group achievement (recall the SD of achievement among control group teachers was 25 percent correct). This slightly overstates but is largely consistent with the estimated effect from Table 5 (0.204 SD), and also that presented in the original paper by Muralidharan and Sundararaman (2011) (0.22 SD). For reference, quality variation among youngest teachers in control group is about 0.6 SD of achievement.

While it may not be surprising that the model can fit average achievement treatment effects, as well as the gradient, a notable benefit from the structural approach taken in this paper is in understanding how teacher quality is generated, and how changes in teacher inputs lead to quality variation. Incentive pay increased mean production time by 0.16 (from a base of 0.72). About two-thirds of this comes from a reduction in OJT investments: on average, I_t decreases by 0.12 (from a base of 0.14), while reductions in OJL account for most of the remainder, as it decreases on average by 0.025 (from a base of 0.14). There is also a small increase in mean labor supply, as $[1 - l_t]$ increases by 0.011 (from a base of 0.84).

We can also use the model to compute treatment effects with respect to teacher age. Figure 3A plots treatment effects for teacher quality, by age. Quality increases are largest for younger teachers, and the treatment effect variance (indicated by the red lines, denoting the 25th and 75th percentiles), is the largest here, too. Unobserved heterogeneity is most salient in teachers' careers. Figure 3B plots treatment effects for investment, by age. OJT investment decreases over career, which means the largest reductions are among early career teachers. In contrast, OJL increases slightly over careers of teachers in the control group, so the treatment effects pictured in Figure 3C move in the opposite direction to those for investment, with respect to age. Figure 3D plots treatment effects for labor supply, by age. While under the preliminary estimates the model generates a counterfactual increase in labor supply, these responses are not large, and are in particular small when compared to changes in investment and OJL. For example, the absolute value of the percentage change in labor supply is about a tenth that for OJL.

7.3. Decomposition of Teacher Quality

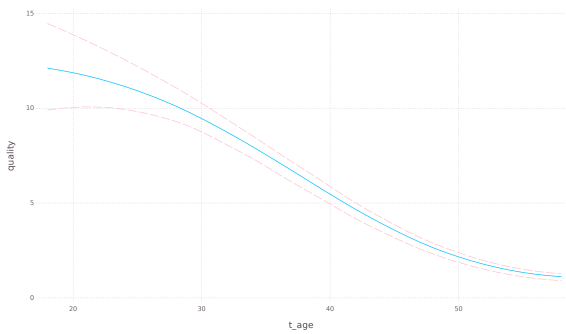
To quantify the importance of different courses of heterogeneity, I decompose lifecycle quality simulated under the status quo (i.e., control scenario) using

$$q_{it} = \mu_i + \vartheta_{it},$$

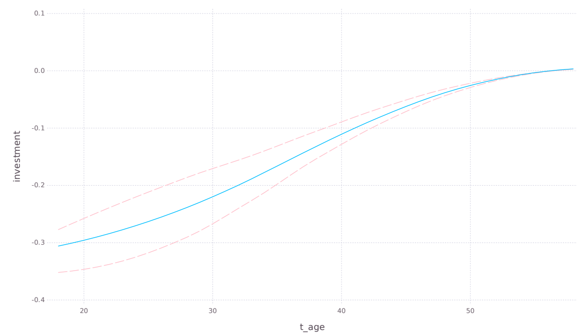
where μ_i is teacher i's mean quality and ϑ_{it} captures i's period deviations from their mean. The between-teacher variation μ_i comes from variation in initial human capital, θ_1, while variation in the deviations ϑ_{it} comes from lifecycle variation in investment share I_t, OJL s_t, and also leisure l_t.

Overall, 68% of quality variation under the status quo is between teachers. The top panel of Figure 4 shows the density of teacher quality, by teacher age group, with the youngest teachers at the bottom and oldest at the top. The vertical line in each density indicates the mean quality

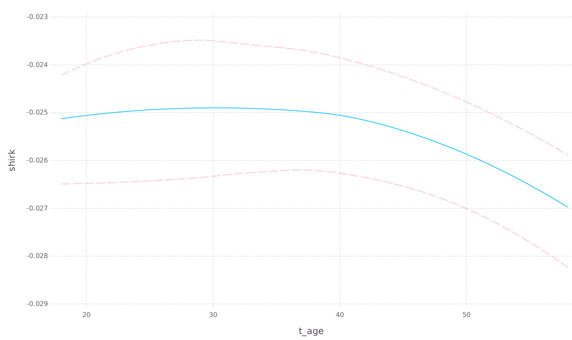
FIGURE 3. Treatment Effects on Model Outcomes, by Teacher Age



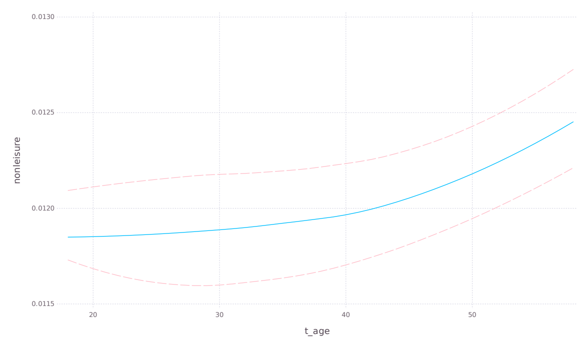
A. Quality



B. Investment share of work time I_t



C. Shirking/OJL share of work time s_t



D. Labor supply

Notes: Figure plots simulated treatment effects. Investment is the share of labor supply, or total work time, spent on investment (I_{it}), on-the-job leisure (OJL) is the share of labor supply, or total work time, spent enjoying OJL (s_{it}), total work time is the share of total time not spent on leisure ($[1-l_{it}]$), and quality ($\theta_{it}[1-l_{it}][1-I_{it}][1-\pi_{it}]$) is denominated in % correct answers, which has a mean of 35 and SD of 25 in the control group in year 1. The red dashed lines denote the 25th and 75th percentiles.

for teachers in that age group. It is immediately apparent that the difference in locations of the means is much smaller than the amount of heterogeneity in quality among teachers in any age group. Teacher age accounts for 18% of quality variation, the majority of the remainder after having removed between-teacher variation, leaving interactions between θ_1 and teacher age (via different input choices and their effects) to account for the remaining 14% of quality variation.

To gauge the importance of model components in generating observed variation in measured quality, I next also include teachers in the treatment group (under the MS2011 incentive strength), and again decompose quality into within- and between-teacher components, but now the status quo and treatment scenarios contribute to within-teacher variation (as well as between-teacher variation, by changing within-teacher means).

The blue densities in the bottom panel of Figure 4 depict the distribution of quality under the MS2011 incentives for teachers in each age group (the red depicts the same, for the control group), with the blue vertical lines indicating age-group-specific mean qualities when treated. Overall, the variation of quality varies much more between teachers than due to stage of lifecycle or status quo vs. treatment scenario. For example, while 10% of quality variation is due to incentive scenario among youngest teachers, this shrinks to 1% among those in the oldest group.

7.4. Dynamic effects of incentives on future quality

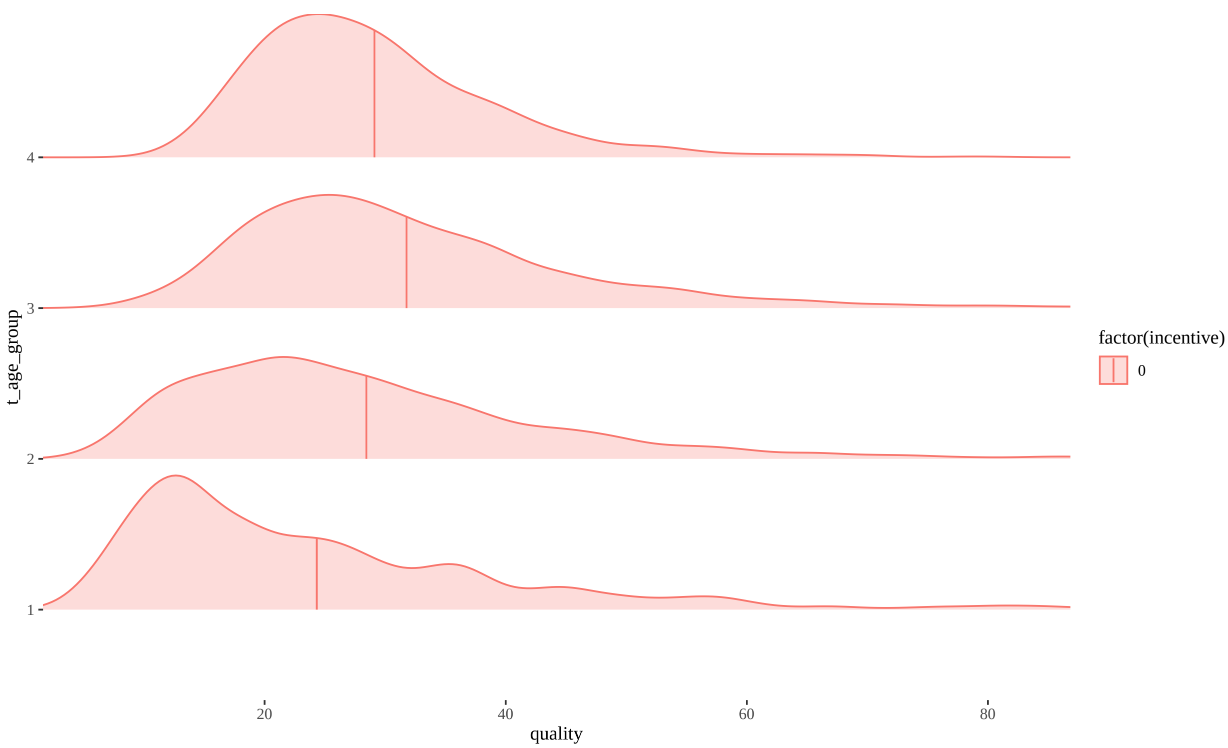
Having estimated a human capital production allowing for both OJT and LBD channels for accumulation, a natural question is how did the static incentives in MS2011, which ended after two years, dynamically affect teacher quality.

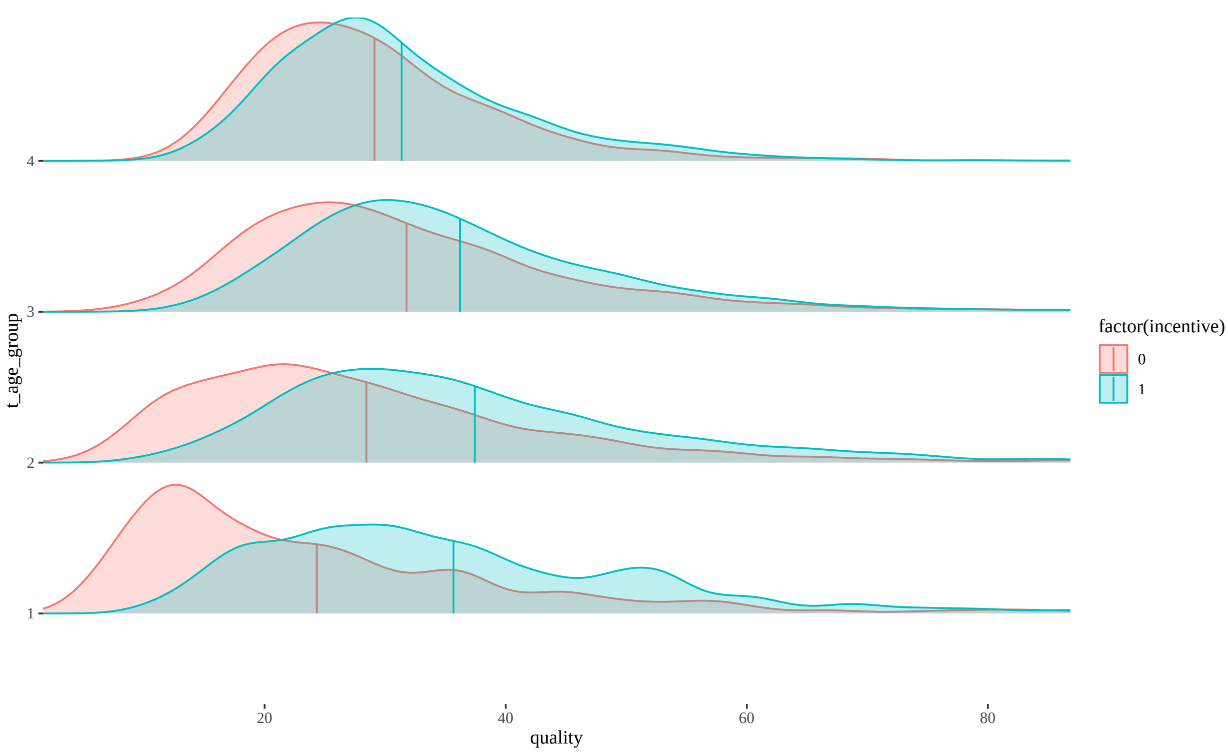
I examine this by computing for each teacher[29] their quality under the status quo and MS2011 incentive pay scenarios, and then presenting each teacher's quality for each year of their career under each scenario, starting from the year of the intervention, relative to (divided by) their quality the year before the intervention (i.e., under the status quo). The results are in Figure 5, which depicts individual teachers' quality relative to their pre-intervention quality under the control (red) and treatment (blue) scenarios as dots, with averages among all teachers of that age and scenario indicated by "x"'s. At the top-left, the figure for the youngest teachers, we can see they have largest relative increase in quality during the intervention, but also that the reductions in OJT investments during the intervention lower post-intervention human capital and quality. The similar pattern is evident among teachers in the older age groups, though the smaller reductions in OJT generate smaller dynamic reductions in future quality. To gauge the overall effect, I sum the discounted quality increases through the end of each teacher's career at age 60.[30] While the MS2011 intervention substantially increased quality during the intervention,

[29]These results have been computed using teachers assigned to the treatment group in MS2011, which should be similar to results computed using the full sample.

[30]I use a 3% discount rate (the same as teachers' individual discount rates).

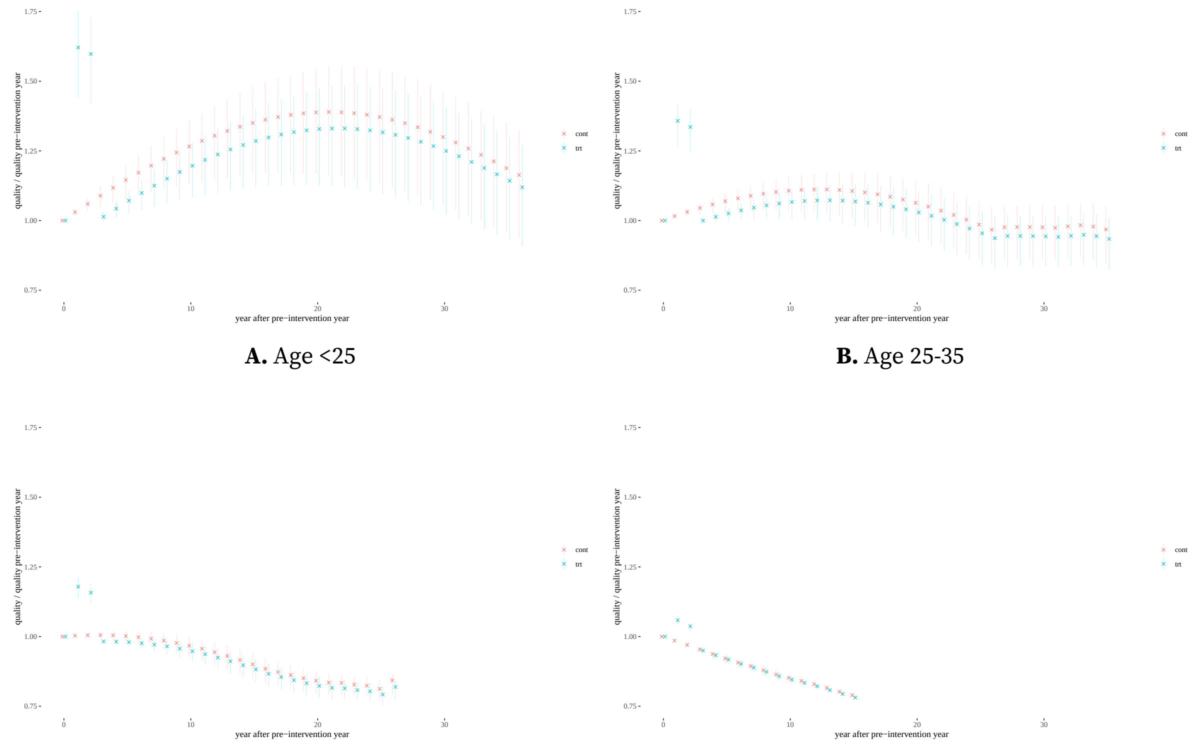
FIGURE 4. Quality decomposition under the status quo and MS2011 incentives





the sum of present-discounted quality is negative for teachers in the youngest two age groups, and positive for those in the older two groups, essentially canceling each other out.

FIGURE 5. Dynamic treatment effects: Quality relative to 2004 quality, by teacher age group



A. Age <25 **B.** Age 25-35

C. Age 35-45 **D.** Age 45+

Notes: Figure plots simulated quality for teachers by their age group in the first year of the intervention, under the status quo (red) and MS2011 incentive pay program (blue). The values are relative to teachers' quality the year before the intervention, where the "x"'s denote averages among teachers in that scenario and year of their career and the vertical lines denote the 25th and 75th percentiles.

7.5. Full-information Benchmark

Under full information, the problem of the principal is to choose a sequence of transfers $\boldsymbol{\psi} = \{\psi_t\}_{t=1}^{T}$ and teacher choice variables to solve

$$\max_{\substack{(I_t, l_t, \pi_t) \in [0,1]^{3T} \\ \boldsymbol{\psi}}} \sum_{t=1}^{T} \rho_g^{t-1} \left[\alpha_g q_t - \psi_t\right] \tag{16}$$

s.t.

$$V_1^{\mathrm{FI}}(\theta_1, \psi) \geq \sum_{t=1}^{T} \rho^{t-1} \underline{u}, \tag{17}$$

where q_t is the teacher quality and α_g is the government's valuation of quality, in ₹, and $\underline{u}$ is the flow utility of the teacher's outside option. The transfer takes place of any potential bonus pay, meaning the first argument of the teacher's utility function inside V_1^{FI} becomes $\alpha_q q_t + \alpha_\chi \chi_t + \psi_t$. The voluntary participation constraint (17) reflects that the teacher must be induced to participate at the start.[31]

To solve for the full-information allocation I must fix values for ρ_g, $\underline{u}$, and α_g. I set $\rho_g = \rho = 0.97$, and $\underline{u}$ to be that teacher's expected value of their optimal program at the start of their career, under the status quo control group (i.e., $\underline{u}(\theta_1) = V_1(\theta_1, \mathbf{d_1} = \mathbf{0})$). Finally, for α_g I use the calculations used in MS2011 for the effect of the intervention on the present discounted value of student earnings.

Figure 6 presents outcomes for teachers with different levels of initial human capital under full information (red), with their status quo outcomes in blue for comparison. For all teachers, leisure and shirking would be lower under full information, although the Principal would not find it worthwhile to eliminate either. Notably, these variables would also vary less under full information than under the status quo, showing the advantage of being able to precisely target unobserved heterogeneity.

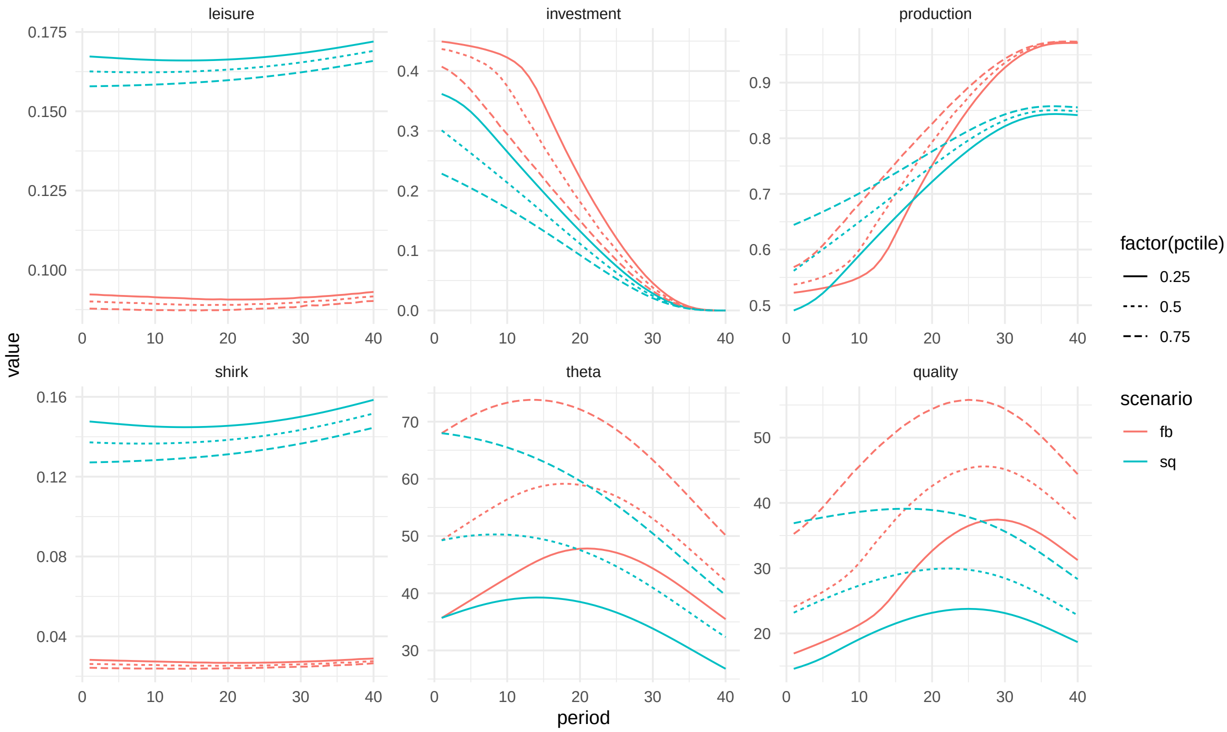
Investment would be much higher for all initial skill levels, which causes the share of work human capital used in production to tilt upwards so that it is initially lower under full information than under the status quo, crossing the share under the status quo in the second decade of teachers' careers. Overall, quality would be much higher under full information.

8. Concluding Remarks

Preliminary results indicate that OJT is the primary source of human capital growth among teachers, and that, at the incentive strength in Muralidharan and Sundararaman (2011), hidden types are responsible for the majority of the heterogeneity in teacher quality. I have also found that the incentives used in Muralidharan and Sundararaman (2011) reduced dynamic teacher quality growth by lowering human capital investments to negligible amounts. Results from the full-information scenario show informational asymmetries lead to suboptimal variation in teacher inputs and quality, resulting in substantially lower quality.

[31] Alternatively, the participation constraint could be modeled as $V_t^{\mathrm{FI}}(\theta_t, \psi) \geq \sum_{\tau=t}^{T} \rho^{\tau-1} \underline{u}$ for $t = 1, \ldots, T$, which would require that the teacher must be induced to remain in the arrangement in each period. While this specification may seem appealing, as it accounts for the possibility that the teacher may cease to participate, the growth in teacher salaries under the status quo would introduce another dynamic element, in addition to the one present due to human capital accumulation.

FIGURE 6. Full-information outcomes for teachers with 25th, 50th, and 75th percentile initial human capital



References

Barlevy, Gadi, and Derek Neal. 2012. "Pay for Percentile." *American Economic Review* 102 (5): 1805–31.

Becker, Gary S. 1962. "Investment in Human Capital: A Theoretical Analysis." *Journal of Political Economy* 70 (5, Part 2): 9–49.

Behrman, Jere R, Michela M Tincani, Petra E Todd, and Kenneth I Wolpin. 2016. "Teacher Quality in Public and Private Schools Under a Voucher System: The Case of Chile." *Journal of Labor Economics* 34 (2): 319–362.

Ben-Porath, Yoram. 1967. "The Production of Human Capital and the Life Cycle of Earnings." *Journal of Political Economy* 75 (4): 352–365.

Besley, Timothy, and Maitreesh Ghatak. 2005. "Competition and Incentives with Motivated Agents." *American Economic Review* 95 (3): 616–636.

Bezanson, Jeff, Alan Edelman, Stefan Karpinski, and Viral B Shah. 2017. "Julia: A Fresh Approach to Numerical Computing." *SIAM review* 59 (1): 65–98.

Biasi, Barbara. 2021. "The Labor Market for Teachers under Different Pay Schemes." *American Economic Journal: Economic Policy* 13 (3): 63–102.

Biasi, Barbara, Chao Fu, and John Stromme. 2021. "Equilibrium in the Market for Public School Teachers: District Wage Strategies and Teacher Comparative Advantage."Technical report, National Bureau of Economic Research.

Blandin, Adam. 2018. "Learning by Doing and Ben-Porath: Life-cycle Predictions and Policy Implications." *Journal of Economic Dynamics and Control* 90: 220–235.

Blinder, Alan S, and Yoram Weiss. 1976. "Human Capital and Labor Supply: A Synthesis." *Journal of Political Economy* 84 (3): 449–472.

Bobba, Matteo, Tim Ederer, Gianmarco Leon-Ciliotta, Christopher Neilson, and Marco G Nieddu. 2021. "Teacher Compensation and Structural Inequality: Evidence from Centralized Teacher School Choice in Perú."

Bowlus, Audra J, and Chris Robinson. 2012. "Human Capital Prices, Productivity, and Growth." *American Economic Review* 102 (7): 3483–3515.

Brown, Charles. 1976. "A Model of Optimal Human-Capital Accumulation and the Wages of Young High School Graduates." *Journal of Political Economy* 84 (2): 299–316.

Chetty, Raj, John N Friedman, and Jonah E Rockoff. 2014. "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood." *American Economic Review* 104 (9): 2633–2679.

Cook, Jason B, and Richard K Mansfield. 2016. "Task-specific Experience and Task-specific Talent: Decomposing the Productivity of High School Teachers." *Journal of Public Economics* 140: 51–72.

Dinerstein, Michael, Rigissa Megalokonomou, and Constantine Yannelis. 2022. "Human Capital Depreciation and Returns to Experience." *American Economic Review* 112 (11): 3725–62.

Dinerstein, Michael, and Isaac M Opper. 2022. "Screening with Multitasking: Theory and Empirical Evidence from Teacher Tenure Reform."Technical report, National Bureau of Economic Research.

Dixit, Avinash. 2002. "Incentives and Organizations in the Public Sector: An Interpretative Review." *Journal of Human Resources*: 696–727.

Dolton, Peter, and Wilbert van der Klaauw. 1999. "The Turnover of Teachers: A Competing Risks Explanation." *Review of Economics and Statistics* 81 (3): 543–550.

Fu, Chao, Hsuan-Chih (Luke) Lin, and Atsuko Tanaka. 2021. "College Loans and Post-Schooling Skill Accumulation." *Working Paper*.

Glewwe, Paul, Nauman Ilias, and Michael Kremer. 2010. "Teacher Incentives." *American Economic Journal: Applied Economics* 2 (3): 205–227.

Haley, William J. 1976. "Estimation of the Earnings Profile from Optimal Human Capital Accumulation." *Econometrica*: 1223–1238.

Hanushek, Eric A. 1971. "Teacher Characteristics and Gains in Student Achievement: Estimation Using Micro Data." *American Economic Review* 61 (2): 280–288.

Hanushek, Eric A. 2011. "The Economic Value of Higher Teacher Quality." *Economics of Education Review* 30 (3): 466–479.

Hanushek, Eric A., and Margaret E. Raymond. 2005. "Does School Accountability Lead to Improved Student Performance?" *Journal of Policy Analysis and Management* 24 (2): 297–327.

Hanushek, Eric A., and Steven G. Rivkin. 2006. "Teacher Quality." *Handbook of the Economics of Education* 2: 1051–1078.

Harris, Douglas N, and Tim R Sass. 2011. "Teacher Training, Teacher Quality and Student Achievement." *Journal of Public Economics* 95 (7): 798–812.

Heckman, James J. 1976. "A Life-Cycle Model of Earnings, Learning, and Consumption." *Journal of Political Economy* 84 (4): S11–44.

Heckman, James J., Lance Lochner, and Ricardo Cossa. 2002. "Learning-by-Doing vs. On-the-Job Training: Using Variation Induced by the EITC to Distinguish Between Models of Skill Formation." *NBER Working Paper Series*: 9083.

Heckman, James J., Lance Lochner, and Christopher Taber. 1998. "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents." *Review of Economic Dynamics* 1 (1): 1–58.

Hölmstrom, Bengt, and Paul Milgrom. 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, and Organization* 7: 24–52.

Hu, Yingyao, and Yi Xin. 2024. "Identification and Estimation of Dynamic Structural Models with Unobserved Choices." *Journal of Econometrics* 242 (2): 105806.

Huggett, Mark, Gustavo Ventura, and Amir Yaron. 2006. "Human Capital and Earnings Distribution Dynamics." *Journal of Monetary Economics* 53 (2): 265–290.

Imai, Susumu, and Michael P Keane. 2004. "Intertemporal Labor Supply and Human Capital Accumulation." *International Economic Review* 45 (2): 601–641.

Imberman, Scott A., and Michael F. Lovenheim. 2015. "Incentive Strength and Teacher Productivity: Evidence from a Group-Based Teacher Incentive Pay System." *Review of Economics and Statistics* 97 (2): 364–386.

Jackson, C Kirabo, Jonah E Rockoff, and Douglas O Staiger. 2014. "Teacher Effects and Teacher-related Policies." *Annual Review of Economics* 6 (1): 801–825.

Johnson, Steven G. 2007. "The NLopt nonlinear-optimization package." `https://github.com/stevengj/nlopt`.

Killingsworth, Mark R. 1982. "'Learning by Doing' and 'Investment in Training': A Synthesis of Two 'Rival' Models of the Life Cycle." *Review of Economic Studies* 49 (2): 263–271.

Kuruscu, Burhanettin. 2006. "Training and Lifetime Income." *American Economic Review* 96 (3): 832–846.

Liu, Dong C., and Jorge Nocedal. 1989. "On the Limited Memory BFGS Method for Large Scale Optimization." *Mathematical Programming* 45: 503–528.

McWilliams, James. 2018. "The Lucrative Art of Chicken Sexing." *Pacific Standard*.

Mehta, Nirav. 2025. "A Partial Identification Approach to Identifying the Determinants of Human Capital Accumulation: An Application to Teachers." *Journal of Applied Econometrics* 40 (6): 639–654.

Muralidharan, Karthik, and Venkatesh Sundararaman. 2011. "Teacher Performance Pay: Experimental Evidence from India." *Journal of Political Economy* 119 (1): 39–77.

Murnane, Richard J. 1975. *The Impact of School Resources on the Learning of Inner City Children.*: Cambridge, MA: Ballinger.

Ost, Ben. 2014. "How Do Teachers Improve? The Relative Importance of Specific and General Human Capital." *American Economic Journal: Applied Economics* 6 (2): 127.

Papay, John P. 2011. "Different Tests, Different Answers: The Stability of Teacher Value-Added Estimates

across Outcome Measures." *American Educational Research Journal* 48 (1): 163–193.
Papay, John P, and Matthew A Kraft. 2015. "Productivity Returns to Experience in the Teacher Labor Market: Methodological Challenges and New Evidence on Long-term Career Improvement." *Journal of Public Economics* 130: 105–119.
Petronijevic, Uros. 2016. "Incentives and Productivity in the Economics of Education." PhD dissertation, University of Toronto (Canada).
Podgursky, Michael, and Matthew Springer. 2011. "Teacher Compensation Systems in the United States K-12 Public School System." *National Tax Journal* 64 (1): 165–192.
Poletaev, Maxim, and Chris Robinson. 2008. "Human Capital Specificity: Evidence from the Dictionary of Occupational Titles and Displaced Worker Surveys, 1984–2000." *Journal of Labor Economics* 26 (3): 387–420.
Powell, M. J. D. 2009. "The BOBYQA Algorithm for Bound Constrained Optimization without Derivatives."Technical Report NA2009/06, Department of Applied Mathematics and Theoretical Physics, Cambridge University. Cambridge, UK.
Price, WL. 1983. "Global Optimization by Controlled Random Search." *Journal of Optimization Theory and Applications* 40 (3): 333–348.
R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria.
Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. 2005. "Teachers, Schools, and Academic Achievement." *Econometrica* 73 (2): 417–458.
Robinson, Chris. 2018. "Occupational Mobility, Occupation Distance, and Specific Human Capital." *Journal of Human Resources* 53 (2): 513–551.
Rockoff, Jonah E, Douglas O Staiger, Thomas J Kane, and Eric S Taylor. 2012. "Information and Employee Evaluation: Evidence from a Randomized Intervention in Public Schools." *American Economic Review* 102 (7): 3184–3213.
Rosen, Sherwin. 1972. "Learning and Experience in the Labor Market." *Journal of Human Resources*: 326–342.
Rowan, Thomas Harvey. 1990. "Functional Stability Analysis of Numerical Algorithms." PhD dissertation, Department of Computer Science, University of Texas at Austin. Austin, TX.
Sanders, Carl. 2010. "Skill Uncertainty, Skill Accumulation, and Occupational Choice." *Working Paper*.
Sanders, Carl, and Christopher Taber. 2012. "Life-Cycle Wage Growth and Heterogeneous Human Capital." *Annu. Rev. Econ.* 4 (1): 399–425.
Shaw, Kathryn L. 1989. "Life-cycle Labor Supply with Human Capital Accumulation." *International Economic Review*: 431–456.
Stinebrickner, Todd R. 2001. "A Dynamic Model of Teacher Labor Supply." *Journal of Labor Economics* 19 (1): 196–230.
Strøm, Bjarne, and Torberg Falch. 2020. "The Role of Teacher Quality in Education Production." In *The Economics of Education*, 307–319: Elsevier.
Taber, Christopher, Ananth Seshadri, and Xiaodong Fan. 2024. "Estimation of a Life-Cycle Model with Human Capital, Labor Supply, and Retirement." *Journal of Political Economy* 132 (1): 48–95.
Tincani, Michela Maria. 2021. "Teacher Labor Markets, School Vouchers and Student Cognitive Achievement: Evidence from Chile." *Quantitative Economics* 12: 173–216.
Weiss, Yoram. 1972. "On the Optimal Lifetime Pattern of Labour Supply." *Economic Journal* 82 (328): 1293–1315.
Wiswall, Matthew. 2013. "The Dynamics of Teacher Quality." *Journal of Public Economics* 100: 61–78.
Yamaguchi, Shintaro. 2012. "Tasks and Heterogeneous Human Capital." *Journal of Labor Economics* 30 (1): 1–53.

Appendix A. Empirical Specification Details

First, note that we have $-\eta\left[\alpha_\chi \chi_t + [\alpha_q + d_t\alpha_w]q_t + d_t\alpha_w\bar{\epsilon}_t\right] \sim \mathrm{N}\left(-\eta\left[\alpha_\chi\chi_t + [\alpha_q + d_t\alpha_w]q_t\right], \frac{[\eta d_t\alpha_w]^2\sigma_\epsilon^2}{n_{it}}\right)$. Expected utility in period t is then[32]

$$\mathbb{E}_t\left[u_t|d_t\right] = \underbrace{-\exp\left(-\eta\left[\alpha_\chi\chi_t + [\alpha_q + d_t\alpha_w]q_t - \frac{1}{2}\frac{\eta[d_t\alpha_w]^2\sigma_\epsilon^2}{n_{it}}\right]\right)}_{u_q} + \underbrace{\alpha_l\frac{l_t^{1-\gamma_l}}{1-\gamma_l}}_{u_l} + \underbrace{\alpha_s\frac{[S([1-I_t]\pi_t, l_t)]^{1-\gamma_s}}{1-\gamma_s}}_{u_s}.$$

We can also write the derivatives of expected utility used to solve the model:

$$\frac{\partial\,\mathbb{E}_t\left[u_q|d_t\right]}{\partial q_t} = -\eta[\alpha_q + d_t\alpha_w]\,\mathbb{E}_t\left[u_q|d_t\right]$$
$$\frac{\partial u_l}{\partial l_t} = \frac{\alpha_l}{l_t^{\gamma_l}}$$
$$\frac{\partial u_s}{\partial S} = \frac{\alpha_s}{S^{\gamma_s}}.$$

[32]Recall that if $a \sim \mathrm{N}\left(\mu_a, \sigma_a^2\right)$ then $\mathbb{E}\left[\exp(a)\right] = \exp(\mu_a + \frac{1}{2}\sigma_a^2)$.

Appendix B. Other tables

Table A4 presents coefficients of achievement treatment effects, conditional on teachers' being present on the day of the site visit (column 1), the day before the site visit (column 2), and present at or before the site visit (column 3). Although the conditioning on teacher presence, a choice variable, complicates interpretation of the coefficients, it shows that the gradient in achievement treatment effects by teacher age seems present, even conditioning on teacher labor supply.

Table A5 presents results from regressions of control-group achievement on teacher age, sex, and education, controlling for student characteristics. Columns (1)-(3) show that teacher quality increases with experience, consistent with the results presented in Table A1. We can also see that male teachers have higher average achievement (t_male). Finally, the coefficients on teacher education generally show an initially increasing, concave, relationship between teacher education and quality. There is no clear linear relationship when estimating using teachers in all age groups (column (1)), but the concave relationship is borne out in columns (2) and (3), which respectively model teacher education using a quadratic and discrete categories. Turning our attention to the youngest teachers, we can see that quality is increasing linearly in education (column (4)) and is also increasing when education is measured using discrete levels (column (5)). The youngest teachers are valuable because they are most informative about the distribution of initial human capital. Note that because teachers also accumulate human capital while working, the positive coefficient on education for young teachers indicates that education is even more productive than working/investing while a teacher.

Appendix C. Details for Normative Results

C.1. Calibrating the Government Weight on Achievement, α_g

Object and units.. Let q denote the class-average quality, (i.e., value-added (VA)) in percentage points (pp) of percent-correct. I set α_g so that a 1 pp increase in a teacher's class-average q maps to present-value (PV) rupees at the time of the intervention. Thus, α_g has units "₹per pp per class".

Assumptions (MS2011 conventions)..

- Returns to achievement (per SD): $r_{\text{SD}} = 0.18$ (average of 0.16 for math and 0.20 for language).
- Score scale: one standard deviation equals 25 pp; hence 1 pp = 1/25 SD = 0.04 SD.
- Benchmark annual wage: $\bar{w} = ₹28,000$.
- Discounting of the 40-year earnings flow at 10% (the amount used in MS2011):

$$A_{40} \equiv \frac{1-(1.1)^{-40}}{0.10} \approx 9.779.$$

TABLE A1. Teacher Quality in the Control Group

	s_t_nts_level_mean	
	(1)	(2)
(Intercept)	-0.266***	-0.415***
	(0.024)	(0.035)
s_t_1_nts_level_mean	0.548***	0.550***
	(0.007)	(0.007)
s_Male	0.019	0.019
	(0.012)	(0.012)
s_SC	-0.023	-0.017
	(0.020)	(0.021)
s_ST	-0.032	-0.022
	(0.044)	(0.044)
s_OBC	0.052**	0.056**
	(0.017)	(0.017)
hh_affluence_index	0.039***	0.039***
	(0.004)	(0.004)
parent_literacy_index	0.057***	0.056***
	(0.006)	(0.006)
t_age_group: 2		0.151***
		(0.028)
t_age_group: 3		0.151***
		(0.028)
t_age_group: 4		0.178***
		(0.030)
Estimator	OLS	OLS
N	17,758	17,758
R^2	0.291	0.293

Note: OLS regressions of standardized test score for teachers in control group, pooled over years and subjects. Regressions include student characteristics and, in column (2), teacher age group indicators. Teacher age group cutpoints are 0, 25, 35, 45, and 60.

TABLE A2. Present at/before visit in the Control Group

	present at			present before			avg. of present at & before		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(Intercept)	0.784***	0.773***	0.761***	0.817***	0.881***	0.801***	0.801***	0.827***	0.781***
	(0.011)	(0.046)	(0.017)	(0.009)	(0.041)	(0.015)	(0.008)	(0.036)	(0.013)
t_age_group: 2		-0.014			-0.092*			-0.053	
		(0.050)			(0.044)			(0.039)	
t_age_group: 3		0.021			-0.044			-0.012	
		(0.049)			(0.043)			(0.039)	
t_age_group: 4		0.039			-0.070			-0.016	
		(0.052)			(0.046)			(0.041)	
!(isless(int(t_age_group), 3))			0.039			0.027			0.033
			(0.022)			(0.019)			(0.017)
Estimator	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
N	473	473	473	473	473	473	473	473	473
R^2	0.000	0.008	0.007	-0.000	0.016	0.004	-0.000	0.012	0.008

Note: "present at" and "present before" respectively denote the shares of site visits during which the teacher was present at the school the day of the visit and the day before the visit.

TABLE A3. Present at/before visit Treatment Effects

	present at			present before			avg. of present at & before		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(Intercept)	0.784***	0.773***	0.761***	0.817***	0.881***	0.801***	0.801***	0.827***	0.781***
	(0.011)	(0.046)	(0.017)	(0.010)	(0.041)	(0.015)	(0.008)	(0.036)	(0.013)
incentive	-0.005	0.071	0.032	-0.001	0.026	0.029	-0.003	0.048	0.030
	(0.015)	(0.074)	(0.024)	(0.014)	(0.066)	(0.021)	(0.012)	(0.058)	(0.019)
t_age_group: 2		-0.014			-0.092*			-0.053	
		(0.049)			(0.044)			(0.039)	
t_age_group: 3		0.021			-0.044			-0.012	
		(0.049)			(0.044)			(0.039)	
t_age_group: 4		0.039			-0.070			-0.016	
		(0.052)			(0.046)			(0.041)	
t_age_group: 2 & incentive		-0.042			0.008			-0.017	
		(0.078)			(0.069)			(0.061)	
t_age_group: 3 & incentive		-0.076			-0.028			-0.053	
		(0.077)			(0.069)			(0.061)	
t_age_group: 4 & incentive		-0.157			-0.097			-0.127*	
		(0.082)			(0.073)			(0.064)	
!(isless(int(t_age_group), 3))			0.039			0.027			0.033
			(0.022)			(0.019)			(0.017)
!(isless(int(t_age_group), 3)) & incentive			-0.061*			-0.049			-0.055*
			(0.031)			(0.028)			(0.024)
Estimator	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
N	932	932	932	934	934	934	932	932	932
R^2	0.000	0.011	0.005	0.000	0.024	0.003	0.000	0.021	0.006

Note: "present at" and "present before" respectively denote the shares of site visits during which the teacher was present at the school the day of the visit and the day before the visit.

TABLE A4. Quality Treatment Effects, Given Present

	s_t_nts_level_mean		
	(1)	(2)	(3)
(Intercept)	-0.494***	-0.595***	-0.585***
	(0.037)	(0.039)	(0.040)
s_t_1_nts_level_mean	0.536***	0.535***	0.535***
	(0.005)	(0.005)	(0.005)
s_Male	0.018	0.019*	0.018
	(0.009)	(0.009)	(0.009)
s_SC	0.000	0.003	0.001
	(0.015)	(0.015)	(0.015)
s_ST	-0.042	-0.043	-0.040
	(0.033)	(0.033)	(0.033)
s_OBC	0.040**	0.040**	0.040**
	(0.013)	(0.013)	(0.013)
hh_affluence_index	0.051***	0.051***	0.051***
	(0.003)	(0.003)	(0.003)
parent_literacy_index	0.042***	0.043***	0.043***
	(0.005)	(0.005)	(0.005)
t_age_group: 2	0.154***	0.169***	0.165***
	(0.030)	(0.030)	(0.030)
t_age_group: 3	0.154***	0.162***	0.160***
	(0.030)	(0.030)	(0.030)
t_age_group: 4	0.180***	0.196***	0.189***
	(0.033)	(0.033)	(0.033)
incentive	0.384***	0.392***	0.386***
	(0.047)	(0.047)	(0.047)
t_age_group: 2 & incentive	-0.067	-0.081	-0.071
	(0.049)	(0.049)	(0.049)
t_age_group: 3 & incentive	-0.197***	-0.207***	-0.199***
	(0.049)	(0.049)	(0.049)
t_age_group: 4 & incentive	-0.456***	-0.457***	-0.451***
	(0.053)	(0.052)	(0.052)
present_at_visit	0.071***		
	(0.021)		
present_day_before_visit		0.176***	
		(0.023)	
present_at_or_day_before_visit			0.174***
			(0.026)
Estimator	OLS	OLS	OLS
N	34,709	34,831	34,709
R^2	0.270	0.272	0.271

TABLE A5. Teacher Characteristics and Quality, Control Group

	all teachers			t_age_group=1	
	s_t_nts_level_mean				
	(1)	(2)	(3)	(4)	(5)
(Intercept)	-0.421***	-0.532***	-0.544***	-0.704***	-0.778***
	(0.037)	(0.042)	(0.052)	(0.156)	(0.163)
s_t_1_nts_level_mean	0.548***	0.548***	0.548***	0.365***	0.377***
	(0.007)	(0.007)	(0.007)	(0.029)	(0.030)
s_Male	0.018	0.019	0.019	0.095	0.096
	(0.012)	(0.012)	(0.012)	(0.054)	(0.054)
s_SC	-0.022	-0.033	-0.032	-0.094	-0.102
	(0.021)	(0.021)	(0.021)	(0.113)	(0.113)
s_ST	-0.037	-0.045	-0.045	-0.524**	-0.489**
	(0.044)	(0.044)	(0.044)	(0.180)	(0.182)
s_OBC	0.056**	0.053**	0.053**	0.009	0.019
	(0.017)	(0.017)	(0.017)	(0.103)	(0.103)
hh_affluence_index	0.039***	0.038***	0.038***	0.044*	0.038
	(0.004)	(0.004)	(0.004)	(0.020)	(0.020)
parent_literacy_index	0.056***	0.054***	0.054***	0.092***	0.086**
	(0.006)	(0.006)	(0.006)	(0.028)	(0.028)
t_age_group: 2	0.136***	0.120***	0.120***		
	(0.030)	(0.030)	(0.030)		
t_age_group: 3	0.120***	0.102***	0.103***		
	(0.030)	(0.030)	(0.031)		
t_age_group: 4	0.146***	0.132***	0.133***		
	(0.031)	(0.031)	(0.031)		
t_male	0.086***	0.083***	0.083***	0.177**	0.140*
	(0.013)	(0.013)	(0.013)	(0.065)	(0.069)
t_educ_years	-0.004	0.074***		0.088***	
	(0.004)	(0.015)		(0.018)	
t_educ_years ** 2		-0.009***			
		(0.002)			
t_educ_level: 2			0.126**		0.312**
			(0.045)		(0.096)
t_educ_level: 5			0.148***		0.510***
			(0.043)		(0.100)
t_educ_level: 7			0.075		
			(0.045)		
Estimator	OLS	OLS	OLS	OLS	OLS
N	17,758	17,758	17,758	945	945
R^2	0.294	0.296	0.296	0.198	0.200

Note: OLS regressions of standardized test score for students in control group, pooled over years and subjects. Regressions include student characteristics. Columns (1)-(3) present results for all teacher age groups, while columns (4)-(5) presents results for teachers in the youngest age group. Teacher age group cutpoints are 0, 25, 35, 45, and 60.

- Lag to labor-market entry: discount back 12 years from age 20 to age ≈ 8 by $(1.1)^{-12} \approx 0.3186$.
- Class size: $\bar{n} = 38$ students per teacher.

Step 1: PV per student for a 1 pp gain.. A 1 pp VA gain raises log wages by

$$r_{\text{pp}} \equiv r_{\text{SD}} \times \frac{1}{25} = 0.18 \times 0.04 = 0.0072.$$

The implied annual rupee gain is $r_{\text{pp}}\bar{w} = 0.0072 \times 28{,}000 = ₹201.6$. I next calculate the present-discounted value of this annual gain over 40 years, and discount this amount to bring it back 12 years:

$$\text{PV}_{\text{student, 1 pp}} = 201.6 \times A_{40} \times (1.1)^{-12} \approx 201.6 \times 9.779 \times 0.3186 \approx ₹628.$$

Step 2: Convert to the teacher/class unit (pp per class).. Finally, multiply by class size to cast in terms of model's q, which is per-class:

$$\alpha_g \;=\; \text{PV}_{\text{student, 1 pp}} \times \bar{n} \;=\; 628 \times 38 \;\approx\; ₹23{,}900 \text{ per pp per class.}$$