Teacher Quality Policy When Supply Matters: Online Appendix

Jesse Rothstein

July 24, 2014

A Search model

Each teacher draws a single outside job offer each year. If she accepts the offer, she exits teaching forever. The outside offer arrives after the teacher learns her previous year's performance (and is paid on that basis).

Outside offers are indexed by the continuation value that they provide, ω. I assume that the outside offer received prior to year $t > 1$, ω_t, has a censored Pareto distribution:

$$
F_t(\omega_t) = \begin{cases} 0 & \text{if } \omega_t \leq V_t^0 \lambda_0^{\frac{1}{\zeta_t'}} \\ 1 - \lambda_0 \left(\frac{V_t^0}{\omega_t} \right)^{\zeta_t'} & \text{if } V_t^0 \lambda_0^{\frac{1}{\zeta_t'}} < \omega_t < HV_t^0 \\ 1 & \text{if } HV_t^0 \leq \omega_t. \end{cases} \tag{1}
$$

Here, V_t^0 is the continuation value if the teacher remains in teaching in year t under the baseline, single salary contract (which is constant across teachers), λ_0 is the annual exit hazard under this contract, and H is the maximum outside wage, expressed as a fraction of the inside continuation value.[1] Importantly, the distribution of ω_t is independent of the teacher's ability as a teacher, τ_i. Thus, as the teacher learns about τ_i she does not simultaneously learn about her future outside options (and vice versa).

Under the outside distribution (1), the probability that a teacher who would obtain continuation value $V_t \in \left[V_t^0 \lambda_0^{\frac{1}{\zeta_t'}}, HV_t^0 \right]$ in teaching will instead exit is

$\lambda_t(V_t) = \Pr\{\omega_t > V_t\} = \lambda_0 \left(\frac{V_t^0}{V_t} \right)^{\zeta_t'}$, with $\frac{\partial \ln \lambda_t(V_t)}{\partial \ln V_t} = -\zeta_t'$. The model in the main text is developed in terms of the negative of the elasticity of the exit hazard with respect to the inside wage under the baseline contract, $\zeta \equiv -\frac{\partial \ln \lambda_t}{\partial \ln w^0} = -\frac{\partial \ln \lambda_t}{\partial \ln V_t} * \frac{\partial \ln V_t}{\partial \ln w^0} = \zeta_t' * \frac{\partial \ln V_t}{\partial \ln w^0}$. The latter fraction varies with t. I thus solve

---

[1]The use of a censored distribution ensures that V_t is finite for any ζ_t'. It has no effect on the results so long as the censoring point is high enough that offers at that point are always accepted. I set $H = 2$, satisfying this criterion for the contracts under consideration here.

recursively for this elasticity – which depends on ζ'_s, $s > t$, but not on ζ'_t itself – and use it to define the elasticity parameter in (1) as $\zeta'_t \equiv \zeta * \left(\frac{\partial \ln V^0_t}{\partial \ln w^0} \right)^{-1}$.

The distribution of the initial non-teaching offer, ω_1, is similar to that of offers later in the career, though here the shape parameter is computed as $\zeta'_t \equiv \eta \left(\frac{\partial \ln V^0_1}{\partial \ln w^0} \right)^{-1}$.

B Solving the model

Equation (4) in the main text does not have a closed-form solution, but for any specified contract it can be solved recursively. Under the learning model developed above, the distribution of period-t performance measure given θ_{t-1} is

$$
y_t \,|\, \theta_{t-1} \sim \mathcal{N} \left(\hat{\tau}_{t-1}, \; \frac{1}{\frac{1}{(1-h)\sigma^2_\tau} + \frac{t-1}{\sigma^2_\epsilon}} + \sigma^2_\epsilon \right). \tag{2}
$$

This is a univariate distribution that can easily be computed for any specified value of $\hat{\tau}_{t-1}$. Given $\hat{\tau}_{t-1}$ and y_t, computation of $\hat{\tau}_t$ is trivial.

The recursive solution thus has three steps. First, I compute $w^C_T (y_1, \ldots, y_T)$, the final period wage under contract C as a function of the performance signals to date. Second, I compute the value of remaining in teaching in period T, $V_T (\theta_{T-1}; C)$, as a function of θ_{T-1}, by integrating w^C_T over the conditional distribution of y_T given by (2). Third, for each $t < T$, given estimates of $V_{t+1} (\theta_t; C)$ as a function of θ_t, I compute $w^C_t (y_1, \ldots, y_t)$ as a function of y_t and θ_{t-1}, then integrate each over the distribution of y_t (and therefore of θ_t) given θ_{t-1} to obtain $V_t (\theta_{t-1}; C)$.

The state space θ_t is of dimension $t + 1$, creating a dimensionality problem for careers of reasonable length. Note, however, that each of the contracts considered above reduces the state space for computation of w^C_t from the t-dimensional distribution $\{y_1, \ldots, y_t\}$ to a one- or two-dimensional distribution: $\{y_{t-1}, y_t\}$ for the bonus contract and $\{\bar{y}_t\}$ for the tenure and alternative firing contracts. Meanwhile, the teacher's assessment of her own ability at the end of period $t - 1$ can be summarized either by the single variable $\hat{\tau}_{i,t-1}$ or by the pair $\{\mu, \bar{y}_{t-1}\}$. I can thus focus on state spaces of only two dimensions, $\tilde{\theta}_{t-1} = \{\hat{\tau}_{t-1}, y_{t-1}\}$ for the bonus contract or $\check{\theta}_{t-1} = \{\mu, \bar{y}_{t-1}\}$ for the tenure and firing contracts. I approximate the joint distributions of these two-dimensional state variables and y_t with grids of 149^3 points spaced to have equal probability mass.

Having computed $V_t (\theta_{t-1}, C)$ for each t, θ_{t-1}, and C, I simulate the impact of policies by drawing potential teachers from the $\{\mu, \tau\}$ distribution, then drawing performance measures $\{y_1, \ldots, y_T\}$ for each. For each career, I compute θ_{t-1} and V_t at each year t, and use these to compute the effects of contract C on the probability of entering the profession and, conditional on entering, on surviving to year t. Note that I need not model the distribution of $\{\mu, \tau\}$ in the *population* of potential teachers – under my constant elasticity assumptions, changes in the returns to teaching induce proportional changes in the amount

of labor supplied to teaching by each type that do not depend on the number of people of that type in the population.

C   Market clearing

Alternative contracts may yield greater or lesser entry or persistence in aggregate. For example, adding performance bonuses without reducing base pay will yield more entry from high-μ teachers and greater persistence of high-$\hat{\tau}_t$ teachers, without offsetting reductions from teachers with low μ or $\hat{\tau}_t$. Under each alternative contract, I compute the steady-state size of the teacher workforce, assuming that the contract has been in place for at least T years and that the same number of entering teachers have been hired in each year.

I consider two scenarios for labor demand. First, I assume that the education system will require the same number of teachers under the alternative contracts as are required under the baseline contract; where my computation yields a larger or smaller workforce than in baseline, I assume that the base salary is adjusted upward or downward to yield the appropriate number of teachers. The α^B and α^F parameters listed for the "fixed quantity" scenario in Table 1 are the adjustments required given the other parameters listed there; these are found via a numerical search algorithm. My second scenario assumes instead that the system's total budget is fixed, so that increases in average teacher salaries must be offset by reductions in the number of teachers (and therefore by increases in class size). The "fixed budget" scenario rows in Table 1 show the α parameters that balance the district's budget, given suitable changes in class size.

D   Optimal firing thresholds

In Section IV.B of the main paper, I consider the optimal choice of thresholds (i.e., cutoff values of $\bar{y}_t$) for firing teachers at each year t. I compute these thresholds as the solution to the district's dynamic optimization problem, assuming that the district pays a cost of firing a teacher that is proportional to the expected number of years remaining in the teacher's career and that the district is myopic about potential labor supply responses. Specifically, let x_t represent the number of years that a teacher with t years of experience can be expected to remain in teaching given an annual exit probability of λ_0 and certain retirement after year T. It can be shown that

$$x_t = \frac{1 - \lambda_0}{\lambda_0} \left(1 - (1 - \lambda_0)^{T-t} \right).$$

Let $W_t \left(\bar{y}_t; c^{fire} \right)$ represent the value of retaining a teacher (i.e., offering her employment for the next year) after year $t < T$ if her average performance to date is $\bar{y}_t$ and the firing cost is c^{fire}; let $W_0 \left(c^{fire} \right)$ represent the value of hiring a new teacher from the baseline ability distribution; and let $Z_t \left(c^{fire} \right) \equiv$

3

$W_0\left(c^{fire}\right) - c^{fire}x_t$ represent the value of firing a teacher after year t. Then the continuation value of retaining a teacher after year $t = T - 1$ is:

$$W_t\left(\bar{y}_t; c^{fire}\right) = \lambda_0 W_0\left(c^{fire}\right) + (1 - \lambda_0)\left(\phi_t\bar{y}_t + r(t+1) + \delta W_0\left(c^{fire}\right)\right), \quad (3)$$

where $\phi_t = \frac{\sigma_\tau^2}{t^{-1}\sigma_\epsilon^2 + \sigma_\tau}$ and thus $E\left[\tau|\bar{y}_t\right] = \phi_t\bar{y}_t$. ($\delta$ is the discount rate.) The continuation value of retaining a teacher after year $t < T - 1$ is:

$$
\begin{aligned}
W_t\left(\bar{y}_t; c^{fire}\right) &= \lambda_0 W_0\left(c^{fire}\right) + (1 - \lambda_0)\left(\phi_t\bar{y}_t + r(t+1)\right) \\
&\quad + \delta E\left[\max\left\{W_{t+1}\left(\bar{y}_{t+1}; c^{fire}\right), Z_{t+1}\left(c^{fire}\right)\right\}\middle|\bar{y}_t\right],
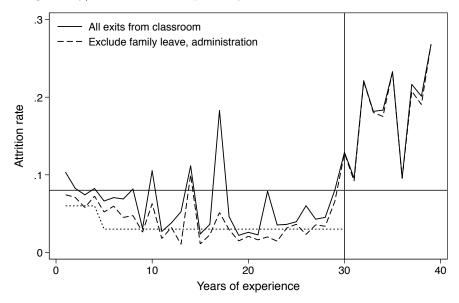\end{aligned}
\quad (4)
$$

Thus, the value of hiring a new teacher must be

$$W_0\left(c^{fire}\right) = 0 + r(0) + \delta E\left[\max\left\{W_1\left(y_1; c^{fire}\right), Z_t\left(c^{fire}\right)\right\}\right]. \quad (5)$$

Given a choice of c^{fire} and a hypothesized value for W_0, one can use (3), (4), and (5) recursively to solve for the implied value of W_0. The fixed point for this is the value $W_0\left(c^{fire}\right)$. Moreover, the firing thresholds at year t are the values of $\bar{y}_t$ that equate $W_t\left(\bar{y}_t; c^{fire}\right)$ with $Z_t\left(c^{fire}\right)$, and these can be used to compute the share of entering teachers who will be fired at some point in their careers. The estimates in Figure 7 in the main paper are obtained by choosing a range of values for c^{fire}; using a numerical search algorithm to find the fixed point W_0 given c^{fire}; computing the firing thresholds implied by these values and the resulting firing shares; and then solving the labor supply model given these thresholds for the market-clearing wages and average productivity levels.

E Appendix Figures

Figure A1: Empirical one-year attrition hazards from the 1999/00 Schools and Staffing Survey/Teacher Follow-Up Survey

Notes: Solid line shows fraction of teachers at each experience level in the 1999-2000 Schools and Staffing Survey who are not teaching at the time of the one-year Teacher Follow-Up Survey. Dashed line codes as non-exits (a) individuals caring for family members at the time of the follow-up who say they plan to return to teaching within a year and (b) individuals who continue to work for state & local governments in non-teaching jobs in elementary and secondary education (e.g., as principals). Vertical line corresponds to the assumed retirement date ($T = 30$) used in simulations here. Horizontal lines correspond to the assumed annual attrition hazards used in the main (solid line; $\lambda = 0.08$) simulations and in the alternative simulation in Table 3, Row 13 (dotted line; $\lambda = 0.06$ for experience ¡ 5 and $\lambda = 0.03$ thereafter). Sample sizes average 122 teachers per year of experience below 30.
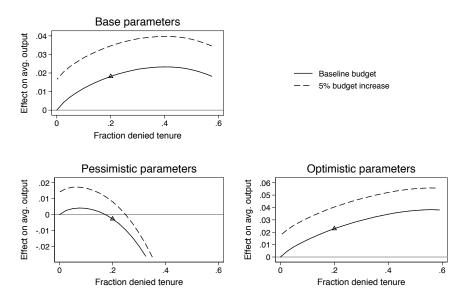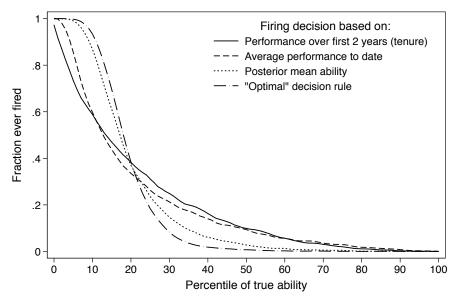
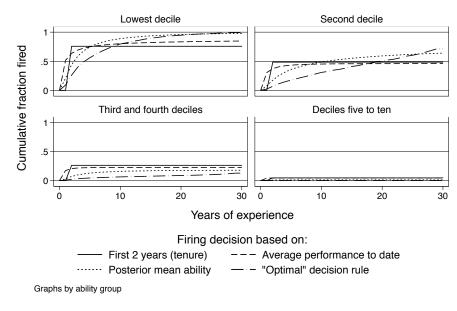Figure A2: Joint effects of tenure contracts and budget increases

Notes: Panels show changes in average output, relative to the single salary contract under the baseline budget and scaled in student-level standard deviations, associated with alternative tenure denial rates and/or budget allocations. Parameters are as indicated in Table 1; base wages are assumed set to fix the total district budget. Marked points indicate the contract parameters (20% denied tenure, with decisions after the second year) used for Table 3, Row 1. Dashed line models a 5% budget increase.

Figure A3: Probability of ever being fired over a 30-year career under different decision rules, by true ability

Notes: See Section IV.C for description of the decision rules. Each rule is set so that, given the current ability distribution, the unconditional probability of being fired before the end of a 30 year career, equals 20%. Figure shows probabilities conditional on ability, τ.

Figure A4: Cumulative firing probability by true ability decile and experience under different decision rules

Lowest decile | **Second decile**

Third and fourth deciles | **Deciles five to ten**

Cumulative fraction fired

Years of experience

Firing decision based on:
—— First 2 years (tenure)   – – – Average performance to date
········ Posterior mean ability   — · — "Optimal" decision rule

Graphs by ability group

Notes: See Section IV.C for description of the decision rules. Each rule is set so that, given the current ability distribution, the unconditional probability of being fired before the end of a 30 year career, equals 20%. Figure shows the probability that a teacher will be fired on or before year t under each decision rule, given ability in the indicated group and assuming no voluntary exit.