Teacher Quality Policy When Supply Matters

Jesse Rothstein

UC Berkeley & NBER

We want to raise teacher productivity. How?

Objectives

Policies

	Existing teachers work harder	Existing teachers work better	Higher-ability teacher workforce
Performance pay	Repeated failure (in US)		??
Performance-based non-retention	No evidence either way		??
Coaching and mentoring		++ (Taylor & Tyler 2011)	

Jumping off points

Safelite performance pay (Lazear 2000)

- Big effects of pay-for-performance on effort & selection.
- Lazear (2003) suggests same for education.

A long literature on performance measurement (AKA value added modeling).

- Ongoing debates about statistical properties in low-stakes settings.
- Little formal consideration of how the measures will be used.

Tennessee POINT (Springer et al. 2010). Performance bonuses up to $15,000 per year had no effect.

- Three-year experiment with volunteers.
- Gets at effort margin; selection impossible to study using RCTs

Staiger & Rockoff (JEP 2010). Model selection effects of performance-based firing rules.

- No model of the labor market -- only tradeoff is ability vs. experience.
- Optimal policy: Fire 80% of teachers after year 2.

This paper

- Goal: Understand potential selection effects of performance pay and performance-based retention, taking account of:
  - (Some of the) known imperfections of performance measures
  - Labor market effects (aka self selection contraints)
- Strategy: Develop dynamic model of occupation choice (teaching vs. other) & calibrate with plausible parameters.
  - Focus on role of limited information
  - Set aside influence activities / goal distortion / manipulation – treat performance measure as noisy but unbiased.
- Policy counterfactuals: New teacher contracts
  - Implemented by entire education sector
  - Established and permanent

The logic of the model

- Let τ be a teacher's true ability, and μ_t her perceived ability (posterior mean) after t years.
- Alternative contracts change the expected compensation as a function of τ. The more reliable the performance measure, the steeper the slope of $E[w \mid \tau]$.
- Incentives for selection depend on $E[E[w \mid \tau] \mid \mu_t]$. This flattens the slope if teachers do not have information about their own ability.
- But we care about selection on τ, not on μ. This creates more flattening if teachers' information is limited.
- In the model, teachers start with a small amount of private information and learn more from subsequent performance measures.

Contracts

1. Baseline: No firing, salaries rise with experience.
2. Bonuses
   - 20% bonuses if $0.5(y_1 + y_2) > y^{PP}$
   - 1st year teachers ineligible.
   - y^{PP} calibrated so 25% of current teachers qualify.
   - Not a tournament – more could qualify if ability distribution rose.
   - Base salaries reduced to yield same total number of teachers.
3. Firing
   - Teacher fired if district's posterior mean falls below a threshold y^F.
   - Posterior mean: $$\bar{y}\frac{\sigma_\tau^2}{\sigma_\tau^2 + \sigma_\varepsilon^2/t}$$
   - y^F calibrated so 10% of current teachers would be fired immediately.
   - Firing reduces future earnings distribution by 10%.
   - Salaries increased to yield same total number of teachers.

Dynamic optimization problem

$$V_t\left(\theta_{t-1};C\right)=E\left[u\left(w_t\right)+\delta\max\left(\omega_{t+1},V_{t+1}\left(\theta_t;C\right)\right)|\theta_{t-1}\right]$$

Value if remain in teaching in year t

State variables: Initial information, performance in years 1 through t-1

Contract

Dynamic optimization problem

$$V_t\left(\theta_{t-1};C\right)=E\left[u\left(w_t\right)+\delta\max\left(\omega_{t+1},V_{t+1}\left(\theta_t;C\right)\right)|\theta_{t-1}\right]$$

Wage & bonus in year t

Outside offer between t and t+1

Inside value if stay to t+1, as function of information at start of t+1.

Dynamic optimization problem

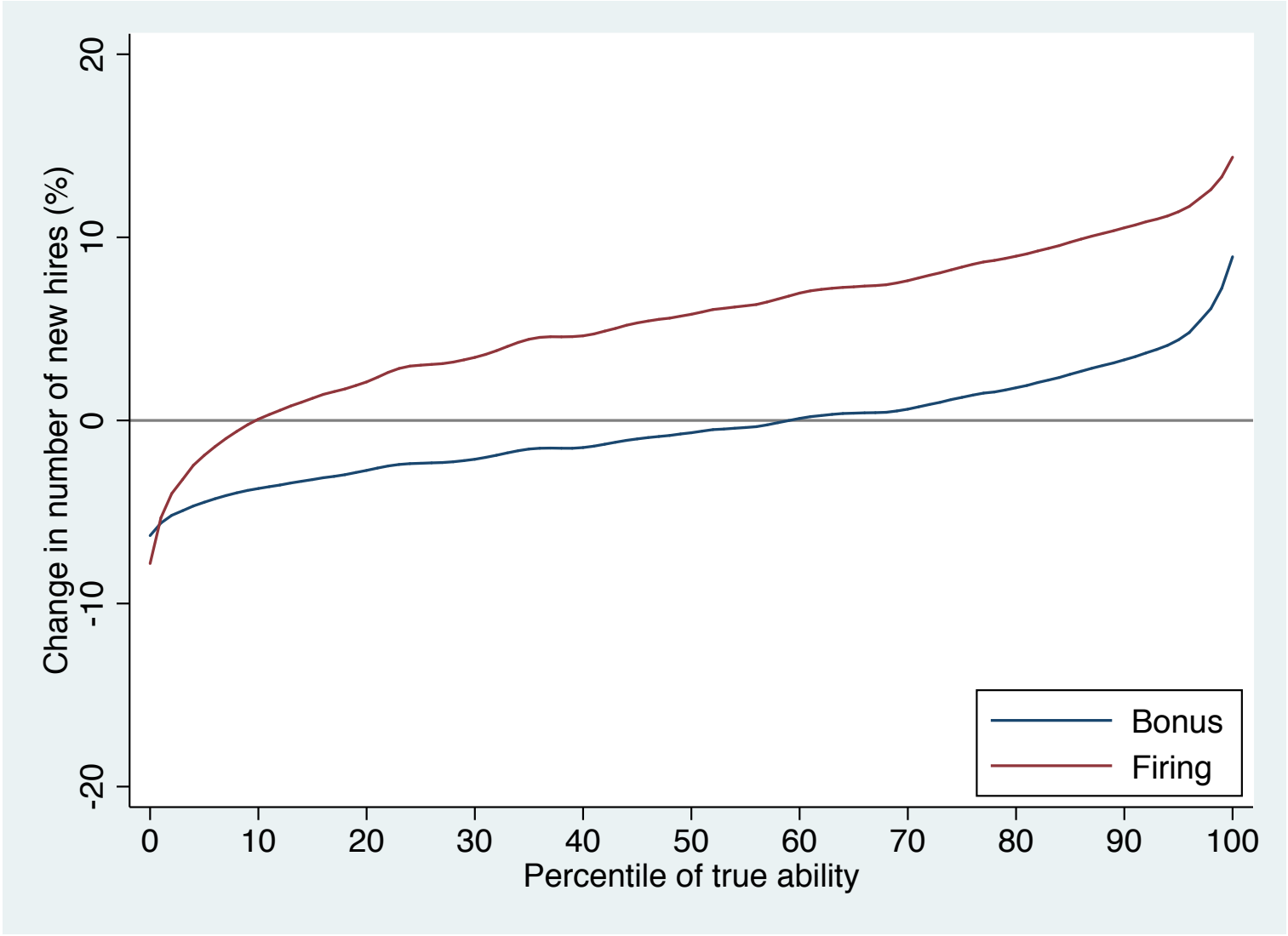
$$V_t\left(\theta_{t-1};C\right)=E\left[u\left(w_t\right)+\delta\max\left(\omega_{t+1},V_{t+1}\left(\theta_t;C\right)\right)|\theta_{t-1}\right]$$

Depends on performance in years 1…t.

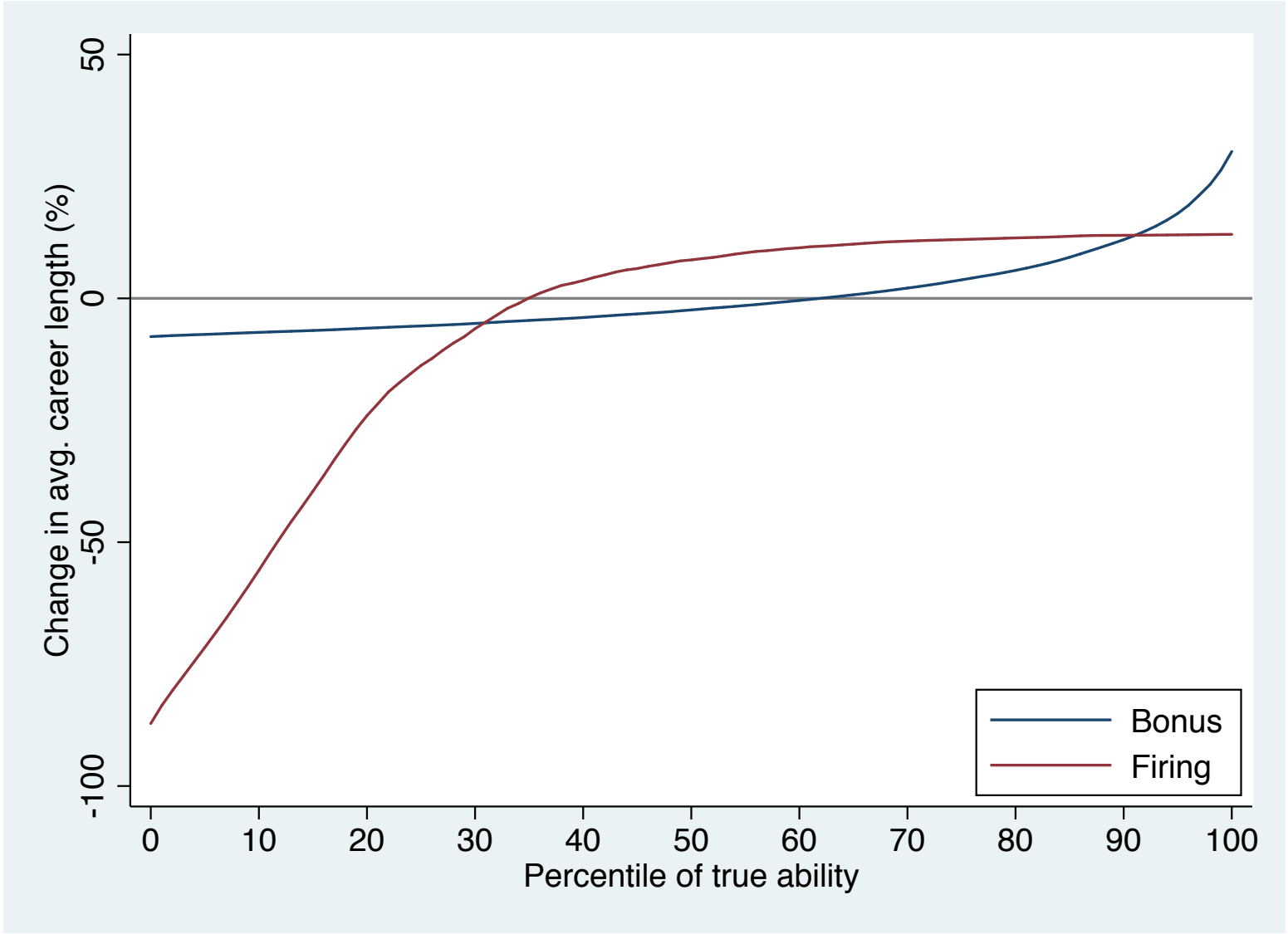
Depends on initial information, performance through t, expected performance after t.

Expectation is over posterior distribution of own ability, distribution of performance in t, t+1, …, T given ability.

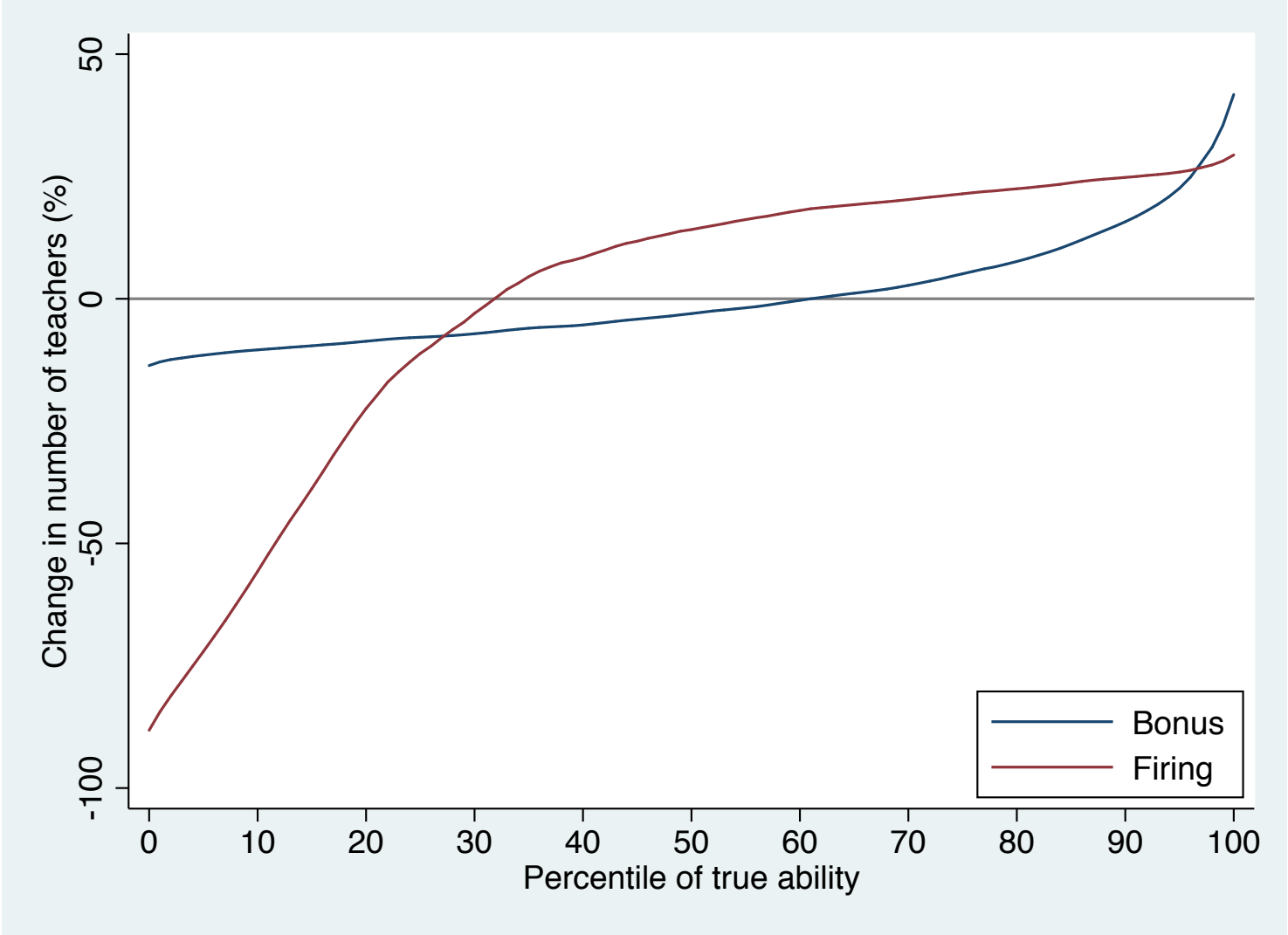
Entrants by ability under bonus & firing contracts



Career length by ability



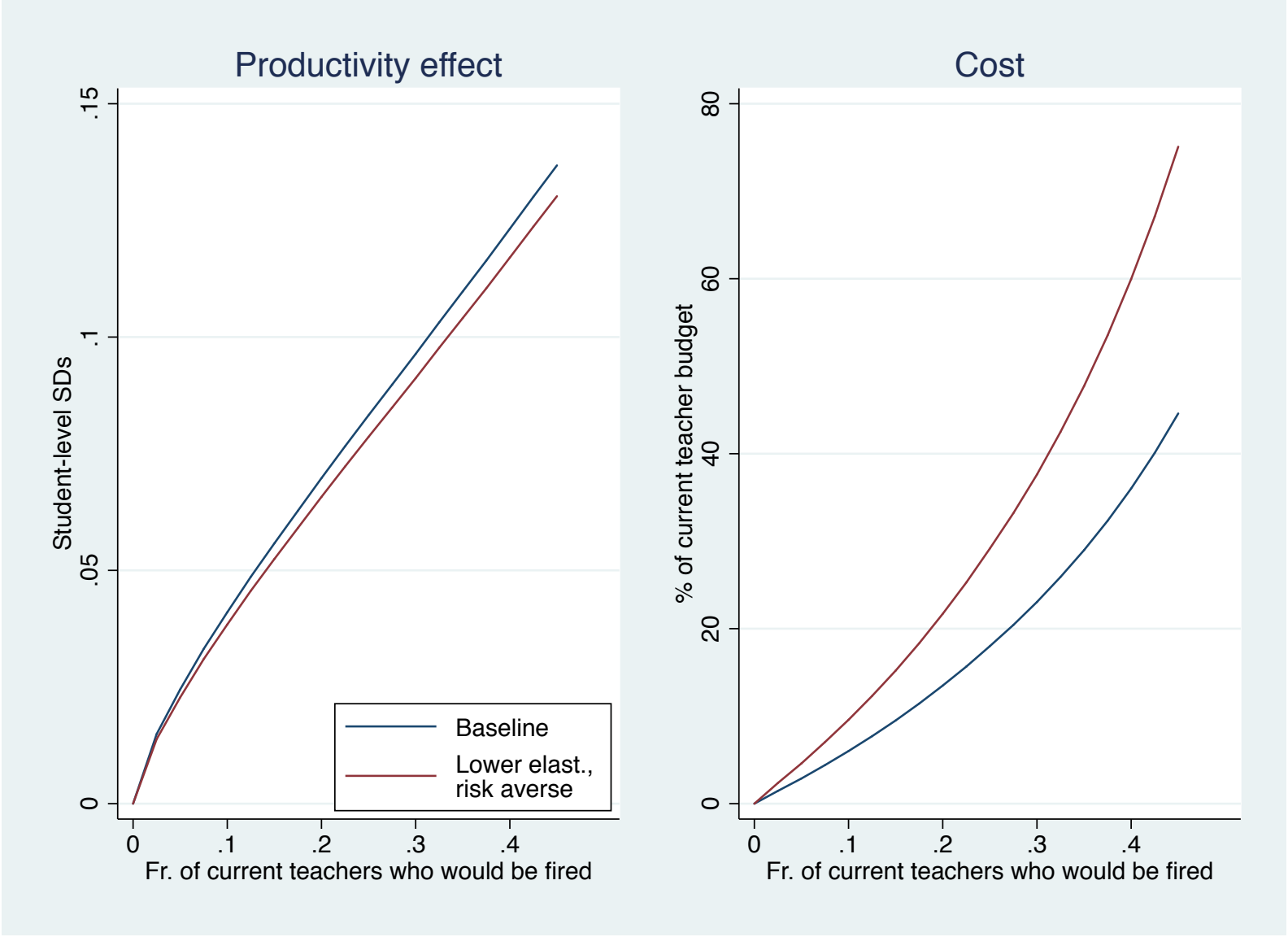
Total # of teachers by ability



Impact of alternative contracts

	Base	Bonuses	Firing
Ability: Mean (SD)	0.000 [0.150]	0.015 [0.153]	0.040 [0.130]
Fr. 1st year	8.0%	8.0%	8.1%
Fr. 1st three years	30.9%	30.8%	31.0%
Avg. experience	8.8	8.9	9.1
Impacts: Mean (SD)	-0.011 [0.151]	0.004 [0.155]	0.029 [0.134]
Base salary (rel. to baseline)		-3.6%	+5.4%
Total wage bill (rel. to baseline)		**+1.8%**	**+5.9%**
Net impact on effectiveness		**+0.015**	**+0.041**

Varying the firing rate



Multiple tasks & influence activities

- Multiple tasks
  - Suppose two dimensions of output, A & B, with corr(τ_A, τ_B) = 0.4.
  - Reward A but want B.
- Influence activities
  - Can raise measured performance by E, at personal cost $c(E) = kE^2$.
  - Calibrate k: $c(0.137) = 0.2$. [$SD(\tau_A \mid \tau_B)=0.137$].
  - Option A: E doesn't affect dimension-B output.
  - Option B: E comes at expense of dimension B.

Impact of firing policy with multiple tasks & influence activities

	Measured effect	True effect on measured output	True effect on unmeasured output	Cost
Baseline	**+0.042**	+0.042	n/a	+6.1%
Multiple outputs	+0.042	+0.042	+0.020	+6.1%
Influence activity – nondistorting	+0.037	+0.030	+0.013	+3.9%
Influence activity – distorting	+0.037	+0.030	**+0.005**	+3.9%

Conclusions

- Can't predict effect of changing the teaching contract without accounting for the teacher labor market.
- When labor market responses are incorporated:
  - Both bonuses & firing policies have positive effects.
  - Both are expensive (but still pass cost benefit test).
  - Plausible effects are not enormous.
  - Effects evaporate if allow for multiple outputs & influence activities.
- Caveat 1: Model is cooked to make the policies look good.
  - Highly elastic labor supply
  - Lots of private information
  - Little risk aversion
  - Unbiased performance measure
- Caveat 2: Many key parameters are made up.
  - Traditional program evaluations / social experiments will be uninformative.
  - Keys: Labor supply, private information, potential to screen on entry, outside labor market return to teaching experience, impact of firing, potential for goal distortion.