Abstract

Public school teachers retire much earlier than comparable professionals. Pension rule changes affecting new teachers can be used to close this gap in the long run, but any effects will not be observed for decades and the implications for workforce quality are unclear. This paper considers targeted incentive policies designed to deter retirement among senior, experienced high-need science and math teachers, as a policy to staff classrooms with qualified teachers and improve workforce quality. We use structural estimates from a dynamic retirement model to simulate the workforce effects of targeted late-career salary bonuses and deferred retirement plans (DROPs) using administrative data from Missouri. Although both policies produce additional teaching years at relatively low costs, by forcing teachers to reveal work–retirement preferences, DROPs generally yield incremental teacher years at lower cost per year. More generally, this work highlights the utility of using structural retirement models to analyze fiscal and workforce effects of changes to public sector pension plans, since the effects of pension rule changes cumulate over many years.
1. INTRODUCTION

A large empirical literature finds substantial, persistent differences in teacher effectiveness within and between schools. High-quality teachers have large effects not only on test scores but also longer-term outcomes, such as matriculation to college and wages (Chetty, Friedman, and Rockoff 2014). This highlights the importance of recruiting, cultivating, and retaining better teachers, particularly in low-performing schools. Similar concerns have been raised about the science, technology, engineering, and mathematics (STEM) teaching workforce, with emphasis on high-need schools as well.

The literature on the labor supply elasticities of teachers in response to a variety of financial incentives is mixed. A recent IES-sponsored experiment providing incentives for highly effective teachers to move to low-performing schools found very inelastic responses (Glazerman et al. 2012). Similarly, studies by Dolton and Van der Klaauw (1995), Hanushek, Kain, and Rivkin (2004), and Feng (2009) imply modest labor supply elasticities. Alternatively, studies by Clotfelter et al. (2008), Feng and Sass (2016), and, to a lesser extent, Falch (2010), find larger labor-supply elasticities, ranging from 2 to 4 in absolute value.

Although research on the malleability of teachers’ labor supply decisions on the whole is mixed, studies of senior teachers consistently show a high degree of responsiveness to pension system incentives (Furgeson, Strauss, and Vogt 2006; Costrell and McGee 2010; Brown 2013; Fitzpatrick and Lovenheim 2014; Knapp et al. 2016; Ni and Podgursky 2016). Traditional teacher pension plans (i.e., final average salary defined–benefit) contain strong incentives designed to “pull” teachers to certain combinations of age or experience, and then “push” them into retirement. Retirement rates tend to spike at “full” and “early retirement” cells in age-experience grids. Moreover, when retirement incentives change across the cells in these grids, retirement rates change accordingly.

The elastic response of teachers to retirement incentives and the powerful “pull” and “push” incentives built into most teacher retirement plans suggests an alternative route to teacher staffing in high-need schools or fields—namely, enticing senior teachers to postpone retirement by altering the “push” incentives for retirement. To date, there seems to be little recognition of the potential for such policies, as the empirical research on the effects of pension incentives on teacher quality and school performance is limited.1 This is particularly relevant because available data suggest that, on average, teachers retire at relatively young ages compared with other professional workers (see, e.g., Harris and Adams 2007).

In most private sector firms and in other areas of government employment, retirement benefits are used to reshape the workforce and upgrade quality. For example, the United States armed services has for decades manipulated retirement incentives to reshape the workforce to meet manpower requirements (Warner and Pleeter 2001; Asch, Mattock, and Hosek 2015).2 In the private sector, professionals are primarily covered by defined contribution (DC) pension plans, which do not have the “push” incentives for retirement that are typical of teacher plans. Nonetheless, private sector firms

---

1. Exceptions include Koedel, Podgursky, and Shi (2013), who study the effects of push and pull incentives on workforce quality; Fitzpatrick and Lovenheim (2014), who examine the effect of an early retirement incentive in Illinois on student test scores; and Chingos and West (2015), who examine the relationship between teacher quality and preferences for retirement plan structure.

use bonuses and other tools to discourage or encourage retirements by senior professional staff.³

For administrators in traditional public schools, there is no ability to experiment with alternative retirement plans because educators in all of the school districts in a state are required to participate in the state’s teacher plan.⁴ This is in contrast to other dimensions of compensation policy (such as performance pay) where local experimentation is feasible. Although there is variation across states in the parameters or rules for teachers’ defined-benefit (DB) retirement plans, there is essentially no natural experimentation with alternative retirement compensation models or policies. Some states have adopted DC and/or hybrid plans (i.e., some combination of DB and DC plans) for teachers in recent years, but these new structures typically apply only for new hires and have not been in place long enough to assess their effects on retirement behavior.

In the absence of sufficient “regulatory space” to generate policy variation and data to undertake traditional evaluations, in this paper we take an alternative approach and use structural estimates from a Stock-Wise “option value” retirement model to simulate the workforce effects of alternative late-career compensation schemes and changes to pension plan rules. We study the state of Missouri and focus on high-need teachers, which we proxy by a STEM teaching field. This paper explores the efficacy of changing retirement incentives in a selective manner for a targeted group of STEM teachers. School districts, particularly those with mostly low–socioeconomic status students, regularly report difficulties in recruiting highly qualified STEM teachers (Podgursky 2010). Given that the median retirement age for Missouri STEM teachers is 57 years, one approach to reducing staffing pressures might be to lengthen the typical STEM-teacher career.

We consider the efficacy of policies designed to offset the powerful late career “push” incentives embedded in traditional teacher pensions plans. In particular, we focus on two policies that target STEM teachers: late career salary bonuses and deferred retirement plans (DROPs). The former are policies that can be implemented by individual school districts or statewide, while the latter would be statewide and require changes to the rules governing the pension plan. Our estimates suggest that DROPs are more cost-efficient than retention bonuses, but, with a careful design, the cost per incremental year of retained teaching by either policy may be justified if programs are targeted to effective and/or high-need teachers.

2. PATTERNS OF RETIREMENT AND PENSION PLAN RULES
Harris and Adams (2007) use data from the 1992–2001 Current Population Surveys to compare career attrition rates of teachers to those of accountants, nurses, and social workers. Despite the intensive focus on early-career teacher attrition in the literature (e.g., Ingersoll 2001; Goldhaber, Gross, and Player 2011), Harris and Adams show

---

⁴ Charter schools in fourteen states are allowed to opt out of state teacher plans. However, there has been no research to date on the effect of teacher retirement behavior in charter schools that exercise that option (Olberg and Podgursky 2011). A few cities (e.g., Chicago, New York City, St. Louis, Kansas City) have municipal teacher plans. In these cases, all educators in district-operated schools are required to participate in the municipal plan.
Table 1. Key Teacher Pension Plans Rules: Missouri Public School Retirement System (PSRS)

<table>
<thead>
<tr>
<th>Missouri PSRS</th>
</tr>
</thead>
</table>
| **Replacement Factor** | 2.5% if experience ≤ 30 years  
2.55% if experience > 30 years |
| **Eligibility - Regular** | Age 60 & experience ≥ 5 years, OR  
Experience 30 years, OR  
Age + experience ≥ 80 years |
| **Eligibility - Early** | Experience 25 years, OR  
Age 55 & Experience ≥ 5 years |
| **Social Security** | No |

Source: Pension plan reports. Note that Missouri PSRS does not cover teachers in Kansas City and St. Louis—who are covered by separate plans—but covers all other teachers in the state.

that early-career teachers behave similarly to early-career workers in comparable professions. The divergence between teachers and other professional workers with respect to separations comes later in the career: “teacher turnover is relatively high among older teachers reflecting the fact that they retire considerably earlier than other professionals” (Harris and Adams 2007, p. 326). What factors are driving high late-career attrition among teachers? Although individual workforce participation decisions are caused by a variety of factors, pension plan incentives surely play an important role.

Table 1 summarizes key pension plan rules in Missouri. The replacement factor is a multiplier that, when combined with years of service, gives the pension replacement rate. For example, a teacher in Missouri with 30 years on the job would retire with a 2.5 percent replacement factor, yielding a pension that replaces 75 percent (30 × 0.025) of the final average salary (FAS) in retirement (where the FAS is calculated as the average of the highest three years of earnings). In the Missouri plan there are three conditions under which teachers become eligible to collect unreduced pension benefits (i.e., eligible for “full retirement”): (1) the sum of age and experience is 80 or above (“Rule of 80”), (2) the number of in-system service years is 30 or above, or (3) age is 60 or above with at least five years of service. The modal age of entry into teaching in Missouri is age 24, which means that with continuous work a typical entrant would become eligible for full retirement benefits at age 52 via Rule-of-80.

Figure 1 shows the expected pension wealth accrual profile for a representative age-24 entrant in the Missouri teaching workforce over a career cycle assuming continuous work. Pension wealth captures the present-discounted-value of the stream of payments that have been earned up to a given point in the career, assuming the teacher quits covered employment at the indicated age. The figure shows the back-loading of

---

5. We restrict our analysis to teachers in the state retirement plan. This excludes teachers in the St. Louis and Kansas City school districts, who have their own municipal plans. The state plan covers more than 90 percent of Missouri teachers.
6. In this paper we will use “years of service” and “experience” interchangeably. In practice, they may differ if teachers exercise options to purchase service years or use unused sick days to buy service years. These practices are relatively limited in the Missouri teacher plan.
7. Pension wealth is a function of the following information about the teacher for the three years prior to exit: (1) age, (2) system experience, and (3) earnings or expected earnings. We project out future wages using a growth function that depends on teaching experience. The parameters of the growth function come from a regression...
pension-wealth accrual and highlights the sharp retirement incentives created by the plan rules. Pension wealth rises slowly at first because the pension annuity cannot be collected until age 60. However, once the teacher reaches age 44, she is eligible for early retirement immediately (i.e., she can collect an annuity that is less than the regular retirement rate). With further experience she now approaches regular retirement eligibility under “Rule of 80” (Age + Experience ≥ 80) at age 52. Pension wealth peaks for this representative teacher at the “Rule of 80” notch and declines thereafter (although there is a small bump in the profile due to an increase in the replacement factor that comes with attaining the thirty-first year—see table 1). Working past the peak yields a higher annual annuity but fewer years to collect it, with the latter dominating the
former, and thus lowering expected pension wealth after full retirement eligibility is attained.

As noted in the Introduction, this paper explores the efficacy of changing these “push out” incentives in a selective manner for a targeted group of STEM teachers. Obviously, one way to extend the careers of all current teachers would be to change the pension rules so as to incentivize later retirement (impose a minimum age for full benefits of 62 or 65, reduce the generosity of the formula factor, etc.), but legally it is nearly impossible to change pension rules for incumbent teachers (Monahan 2010). Thus, when reforms of state and local pension plans occur, they focus on changes for new teachers only. The effect of such policies on retirements will not be felt for several decades. We focus instead on a set of voluntary policies that could be enacted in the near term and provide incentives for selected teachers to postpone retirement. In order to assess the effects of such incentive policies we first need a behavioral model of retirement.

3. ANALYTIC FRAMEWORK AND PARAMETER ESTIMATES

Structural Model

We model teacher retirements following Ni and Podgursky (2016), who, in turn, use the general framework developed by Stock and Wise (1990) to estimate a structural model that explains the recurring decision to work or retire at later stages of the career cycle. The model incorporates the “option value” of continued work at any given point in the career. The term “option value” is used in this context because in a DB plan the retirement decision is made only once and cannot be reversed. Thus, each year a teacher compares the value of exercising the option (retiring), versus continuing to work and exercising the retirement option at a future date.

In the model, a teacher’s expected utility in period $t$ is a function of expected retirement in year $m$ (with $m = t, \ldots, T$, where $T$ is an upper bound on the teacher’s lifetime). In period $t$, the expected utility of retiring in period $m$ is the discounted sum of pre- and post-retirement expected utility:

$$E_t V_t(m) = E_t \left\{ \sum_{s=t}^{m-1} \beta^{s-t} [(k_s (1-c) Y_s)^\gamma + \omega_s] + \sum_{s=m}^{T} \beta^{s-t} [(B_s)^\gamma + \xi_s] \right\},$$  

(1)

where $0 < k_s < 1$ captures the disutility of working, $Y$ is income while working in real dollars, $B$ is income during retirement in real dollars (i.e., the pension benefit), and $c$ is the teacher’s contribution rate to the pension plan. Teacher preferences for current versus future income are captured by the discount parameter $\beta$ and risk aversion is reflected in $\gamma$. This specification assumes the disutility of work, $k_s$, changes monotonically with age: $k_s = \kappa \left( \frac{T_0}{\text{age}} \right)^{\kappa_1}$, where $T_0 = 60$ and $\kappa$, $\kappa_1$ are parameters to be estimated.

Uncertainty enters the model through the random variables $\omega_s$ and $\xi_s$, which reflect unobserved factors that change the utility of teaching and retirement. We assume that the unobserved innovations in preferences for teaching relative to retiring, $v_s = \omega_s - \xi_s$, follow an AR(1) process:

$$v_s = \rho v_{s-1} + \varepsilon_s,$$

(2)
where $\varepsilon_s$ is iid normal $N(0, \sigma^2)$, introducing two more parameters to be estimated ($\rho, \sigma$). The variance of $v_s$, which is expressed in dollars, reflects the many unmeasured personal and household factors that affect teachers’ propensities to retire. Because the error is additive to the utility of salary, the standard deviation of $v_s$ can be compared to the level of teacher pay raised to the power $\gamma$.

**Likelihood for Empirical Model**

The expected gain from retirement in period $m$ over retirement in the current period $t$ is given by

$$G_t(m) = E_t V_t(m) - E_t V_t(t) = g_t(m) + K_t(m) v_t,$$

where $g_t(m)$ is the difference in the expected utility from the flow of salary and pension benefits between retiring in period $m$ and current period $t$, $K_t(m) = \sum_{s=t}^{m-1} \pi(s|t)(\beta \rho)^{s-t}$ is a positive-valued function of the model parameters, and $K_t(m) v_t$ is the expected difference in the preference errors between retiring in period $m$ and period $t$ based on the information in period $t$. There are three sources of uncertainty in the model: uncertainty of future earnings and benefits (which we ignore here because teachers’ salary schedules are fixed), uncertainty of survival, and uncertainty in the aforementioned preference shocks. To make survival uncertainty explicit, for a teacher alive in period $t$ we denote the probability of survival to period $s > t$ as $\pi(s|t)$. We can then write

$$g_t(m) = \sum_{s=t}^{m-1} \pi(s|t) (k_s (1-c) Y_s)^{\gamma} + \sum_{s=m}^{T} \pi(s|t) (B_s)^{\gamma} - \sum_{s=t}^{T} \pi(s|t) (B_s)^{\gamma}.$$  

Let $m_t^\dagger = \arg\max g_t(m)/K_t(m)$, be the year among all possible choices of $m$ that maximizes $g_t(m)/K_t(m)$. Then the probability that the teacher retires in period $t$ ($G_t(m) \leq 0$ for all $m > t$) is: $\text{Prob}(g_t(m_t^\dagger)/K_t(m_t^\dagger) \leq -v_t)$. The maximum likelihood estimation of model parameters maximizes the joint probability of each teacher’s retiring in the year she is observed to retire.

An issue with our approach is that the forward-looking model described in equation 1 applies to the population of all employed teachers who make retirement decisions. In that population, we assume the preference uncertainty variable $v_s$ is centered on zero. However, if we draw a given sample of retirement-eligible or near-eligible teachers (in our case, teachers of age between 48 and 65 with at least five years’ teaching experience), we will oversample teachers with positive values of $v_s$ because teachers with negative values are more likely to have retired before reaching the sampling window. As a consequence, the model will tend to overpredict retirement rates in the early years of the panel. We correct for this in our analysis by modifying the distribution from which teachers’ values of $v_s$ are drawn to reflect the fact that at any given point in time, we only observe teachers who have elected not to retire up to that point. Technical details about our procedure can be found in Kong et al. (2018).

**Maximum Likelihood Estimates and In-Sample Fit**

We estimate the model using administrative data from the Missouri Department of Elementary and Secondary Education (MODESE). The data panel included 2,131 STEM
teachers aged 48–65 who have at least five years of experience in 2011. We track our sample of STEM teachers forward in time for three years, through 2014, to evaluate retirement behavior. Table 2 provides descriptive information about the sample. Over the three-year period, 31.3 percent of the STEM teachers in our sample retire.

We use data on teacher salaries, age, and experience—and the pension plan rules shown in table 1—to project forward values of Y and B for each teacher in each year. These projections facilitate estimation of the optimal retirement age for each teacher, which maximizes expected utility, per equation 1. We compare model predictions of retirements to observed outcomes.

Table 3 reports the parameter estimates for the model, which are similar to estimates based on a separate sample of teachers in Ni and Podgursky (2016). For example, the estimated subjective discount rate is 0.961, or 3.9 percent. The disutility of work

---

Table 2. Summary of Statistics: Missouri Science, Technology, Engineering, and Math Teachers

<table>
<thead>
<tr>
<th>Sample Year</th>
<th>Number of Teachers</th>
<th>Age (years)</th>
<th>Experience (years)</th>
<th>Male</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base year</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All 2011</td>
<td>2,131</td>
<td>54.3</td>
<td>20.3</td>
<td>0.31</td>
</tr>
<tr>
<td>Retirement year</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>238</td>
<td>57.3</td>
<td>24.2</td>
<td>0.31</td>
</tr>
<tr>
<td>2013</td>
<td>217</td>
<td>57.6</td>
<td>23.8</td>
<td>0.37</td>
</tr>
<tr>
<td>2014</td>
<td>213</td>
<td>58.3</td>
<td>25.6</td>
<td>0.35</td>
</tr>
<tr>
<td>Not retired by 2014</td>
<td>1,463</td>
<td>55.2</td>
<td>20.8</td>
<td>0.30</td>
</tr>
</tbody>
</table>

Source: Missouri administrative teacher records. Teachers aged 48–65 years in 2011.

Table 3. Maximum Likelihood Estimation Estimates of Structural Parameters

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Estimate</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>β</td>
<td>0.961</td>
<td>(0.011)</td>
</tr>
<tr>
<td>γ</td>
<td>0.730</td>
<td>(0.041)</td>
</tr>
<tr>
<td>σ</td>
<td>4,653.671</td>
<td>(2,288.916)</td>
</tr>
<tr>
<td>ρ</td>
<td>0.665</td>
<td>(0.017)</td>
</tr>
<tr>
<td>κ</td>
<td>0.624</td>
<td>(0.075)</td>
</tr>
<tr>
<td>κ₁</td>
<td>1.315</td>
<td>(0.571)</td>
</tr>
</tbody>
</table>

Log-likelihood: -1868.155

Notes: Parameters are estimated from administrative data of Missouri science, technology, engineering, and math teachers who are aged 48–65 years with at least five years of experience in 2011. The sample period is 2011–2014. For details regarding estimation methodology, see Ni and Podgursky (2016) and Kong et al. (2018).

---

9. In this study, STEM teachers are defined as teachers for whom the majority of their teaching minutes per day are in courses classified as science, math, or technology. STEM teachers constitute 11 percent of all senior teachers in the 48–65 years retirement window. The vast majority (90 percent) are employed in middle or high schools. The average salary of senior STEM teachers in 2011 was $51,452, which is nearly identical to that of senior non-STEM teachers ($51,316).
versus leisure (retirement) grows with age, teachers are notably risk-averse, and shocks to preferences strongly persist, which likely reflects persistent (but unmeasured) personal and household factors, such as health and marital status.

To illustrate the in-sample fit, actual and predicted survival rates for our sample are presented in figure 2. Overall, the model does a good job fitting employment survival for this group. In figure 3, panels A and B report the age distributions of retired and nonretired STEM teachers over the sample period. Nonretired teachers are those who had not retired by 2014. The model provides an excellent fit to the age distribution of both groups. Panels C and D report experience distributions for retirees and non-retirees. Here the fit is not as good as for age. The model overpredicts retirements in the 25–30 years' experience range and correspondingly underpredicts in the same range for non-retirees. There may be nonfinancial reasons for some teachers to avoid taking early retirement ("25 and out"). However, this misfit in the experience distribution does not have a large impact on the simulated effect of retention incentives because the latter is the difference between the predicted quantities. Bias in the prediction with the retention incentives is offset by similar bias without the incentives.\textsuperscript{10}

\textsuperscript{10} Specifically, we reestimated the model shutting off the "25 and out" option. This reduced the experience misfit. The resulting cost estimates for retention incentives were very similar to those reported below, but slightly higher. Because shutting off the early retirement option is ad hoc and cannot be justified theoretically—and does not affect the main findings of this paper—we report only the standard estimates. A related problem, which may act to inflate our cost estimates, is the failure to fully capture the spike at 31 years of experience.
4. THEORY AND SIMULATION OF RETENTION POLICIES

General Considerations

In this section we consider retention incentives offered to selected senior teachers. The incentives may change teachers’ retirement decisions. They also come with a direct cost, and alter pension benefits. We will consider the combined financial costs of the incentives (net of teacher salary), and compare them to the behavioral response generated. The policy objective is to increase the years of teaching for senior STEM teachers in a cost-effective manner.

For the design of a retention policy we need to consider: (1) who qualifies, (2) how much a qualified teacher receives, (3) whether the policy is a short-term or long-term policy, and (4) the specific form of the retention incentive (e.g., a retention bonus or DROP).

On question 1, note that the policy will generate a weak behavioral response if the teachers who are offered an incentive would have continued working in its absence. Thus, for the policy to be cost effective, it must be the case that a group of teachers with a sufficiently high probability of retirement is targeted, otherwise the incentive payments will largely accrue to Non-Switcher Recipient teachers (see below) who would have kept working anyway.

It must also be the case that the retirement behavior of at least some of these retirement-prone teachers can be changed, that is, there must be some teachers at the margin such that a retention incentive can convince them to continue teaching. Note
that a retention incentive not only affects teachers directly targeted in the retirement window but also teachers who may enter the retirement window in the future.

To summarize, a retention incentive or policy has three possible effects when offered to a currently-working teacher who is eligible to receive it in year $R$, where $R$ is some future year:

- **Switcher**: A teacher switches from planned separation before year $R$ in the absence of the incentive, to taking the incentive and working to year $R$ or beyond;
- **Non-Switcher Recipient**: A teacher takes the incentive and works in year $R$ (and perhaps beyond), but would have worked in year $R$ anyway even in the absence of the incentive;\[11\]
- **Non-Recipient**: A teacher chooses to separate before year $R$ and thus foregoes the incentive. Note that in most such cases there would not be a change in separation timing, and thus no change in years of teaching or pension wealth. However, some forward-looking teachers may change retirement plans in response to the incentive, but unforeseen circumstances intercede prior to year $R$ (e.g., shocks in $v_t$), in which case the incentive would still have changed separation behavior.

An incentive that increases teaching years while minimizing the financial cost needs to target Switcher and avoid Non-Switcher Recipient teachers. In a static setting, if we want to offer an incentive to teachers with different probabilities of retiring, then it can be shown that to maximize the ratio $N_{\text{Switcher}} / N_{\text{Non-Switcher}}$, the incentive should be offered to teachers who are most likely to retire. Intuitively, if an incentive is offered to senior teachers who are unlikely to retire (say those with twenty-four years of experience) to induce them to stay for one more year, such teachers will have low ratio of $N_{\text{Switcher}} / N_{\text{Non-Switcher}}$; that is, almost all teachers will take the incentive because they would plan to continue teaching anyway. This inefficient incentive does not result in additional teaching years but is quite costly. On the other hand, maximizing this ratio should not be the only criterion for policy makers. Offering an incentive to teachers who are almost surely retiring (say those with forty years of experience) likely generates mostly Switcher cases, but there are few teachers in this category. A useful incentive should balance efficient targeting as well as the aggregate impact on teaching years.

Regarding question 2—on the dollar value of the incentive—as this value becomes larger more teachers will respond. However, the higher the value of the incentive, the more expensive each Switcher and Non-Switcher Recipient becomes. We will experiment with incentives of different sizes in the analytic work that follows.

Regarding question 3, the effect of a bonus for a given experience level differs depending on the time horizon. A bonus policy put in place for a longer period will have a larger effect than a short-term “single-shot” bonus. For example, a bonus offered at a particular experience level for just a single year will only be relevant for a small subset of teachers at that experience level at the time of the bonus offering; a long-term bonus at the same experience level has the potential to influence the behavior for teachers over

\[11\] In conventional microeconomics terminology, Switchers would be “marginal” teachers and Non-Switcher Recipients would be “inframarginal” teachers.
a much wider range of experience levels for an extended period of time. This example helps illustrate why it would be misleading to use a short-term experiment to study the effect of a bonus policy. Our structural analysis uses a time horizon long enough so that all teachers in the cohort we study retire and can be potentially affected.

We devote the next subsection to question 4—the design of the incentive plan.

Two Retention Incentives: Overview

We evaluate two different types of incentive policies: retention bonuses and DROPs. For both retention bonuses and DROPs, we consider the “who,” “how much,” and “how long” questions discussed above. As the preceding discussion makes clear, there are several factors that determine the efficiency of the incentives. We conduct numerical simulations over different incentive sizes and various levels of experience at which teachers are offered the incentive.

The foundation of the numerical experiments with the two policies is the option-value model described in section 3. Under each policy, we simulate STEM teacher retirements using the model. The results are presented in tables 4–6, where we examine changes in teaching years and financial costs under various versions of the different policies. We use predicted retirements for the 2011 cohort under the current pension rules as the baseline condition for all of our calculations. For the counterfactual scenarios we assume the incentive program is announced in 2011 and track our cohort forward in time until all teachers retire. We compare the predicted retirement outcomes under the baseline condition (no incentives) to predicted retirement outcomes the various policy counterfactuals.

To assess the performance of the counterfactual policies we focus on two considerations. First is the dollar cost per additional generated year of teaching, which is straightforward. Second is the scope of the policy effect, for example, a policy with a very low cost per additional year of teaching but that generates very few additional years (perhaps by targeting very old teachers, almost none of whom are Switchers) and is of little policy value. We further explore the efficiency of the incentives by decomposing the sources of changes in teaching years. This allows us to identify the channels through which the incentives operate.

To be more specific, denote $p^*_1$ as the number of teachers who retire before reaching the threshold ($R$) for incentive eligibility with a retention incentive in place; and $p^*_2$ as the number who retire at or after the threshold. Further, denote $p_1$ and $p_2$ as the numbers of retirees before and after reaching the threshold in the absence of the retention incentive. The measures $p_1 + p_2$ define the policy-relevant population. By definition $p^*_1 + p^*_2 = p_1 + p_2$. If a retention incentive induces “switching,” then $p^*_2 - p_2 > 0$. Connecting these constructs to the teacher types discussed above, the number of Switchers given by $p^*_2 - p_2$; the number of Non-Recipient teachers by $p_1$; and the number of Non-Recipient teachers by $p^*_1$.

Further, denote $Y_1$ as the total years of teaching by teachers in our sample who ultimately retire with experience less than $R$, and $Y_2$ as the total years taught by teachers who retire with experience $R$ or above. The total number of teaching years from our sample is estimated to be 14,424 in the absence of retention incentives. At its most basic level, our question is by how much this number rises, and at what cost, under different policy alternatives.
We will use the following statistics in order to answer this question. First, the policy effect on additional teaching years can be written as \( \Delta Y = (Y_1^* + Y_2^*) - (Y_1 + Y_2) \), where \( Y_1^* \) and \( Y_1 \) are defined analogously to \( p_1^* \) and \( p_1 \) to indicate total years of teaching by teacher type (retiring before or after \( R \)) and policy period. This value can be decomposed into (a) gains in teaching years among those who retire before \( R \), given by \( [(Y_{1}^{*}/p_{1}^{*}) - (Y_{1}/p_{1})]p_{1}^{*} \), and (b) gains in years among those who retire at or after \( R \), which itself is the sum of two components. First is \( [(Y_{2}^{*}/p_{2}^{*}) - (Y_{2}/p_{2})]p_{2} \), which is the change in the number of years worked per individual among those who retire at or after \( R \) even without the incentive. Second is the change in years worked attributable to the compositional shift of some teachers from \( p_2 \) to \( p_2^* \) due to the policy, \( [(Y_{2}^{*}/p_{2}^{*}) - (Y_{1}/p_{1})]\Delta p \), where \( \Delta p = p_2^* - p_2 \). One can verify that these two components can be expressed together as \( (Y_{2}^{*} - Y_{2}) - Y_{1}(\Delta p/p_{1}) \). This combined expression shows that the policy effect due to switchers (Switcher teachers) is equal to the increase in total years of teaching for teachers who retire at or after \( R \), with a corrective term subtracted to account for the shift of some individuals from the pre-\( R \) to post-\( R \) retirement group.

**Modeling Selective Retention or Longevity Bonuses**

We now introduce the retention bonus to the option value model in section 3. The bonus, \( b \), offered in period \( q, 1 < q < m \), modifies expected utility in equation 1 as follows:

\[
E_t V_t(m) = E_t \left\{ \sum_{s=1, s \neq q}^{m-1} \beta^{s-t} [(k_s(1-c)Y_s)^\gamma + \omega_s] + \beta^{q-t} [(k_q((1-c)Y_q + b))^\gamma + \omega_q] + \sum_{s=m}^{T} \beta^{s-t} [(B_s)^\gamma + \xi_s] \right\}.
\]

(5)

All repeating variables and parameters in equation 5 are as defined in equation 1. The new term, \( (k_q((1-c)Y_q + b))^\gamma \), captures the dollar value of the retention bonus in the year of eligibility.

An appeal of the bonus policies we consider is that they can be implemented independently by a single district, or statewide. They would not need to be coordinated with the pension plan. We assume that bonuses are for one year only and do not enter base pay, which means that they do not enter into the calculation of the retirement annuity. The policy experiments we consider vary the point in the career at which the bonus is offered and the size of the bonus, \( b \).

**Modeling Selective DROPs**

DROPs are an alternative way to retain teachers who are eligible for retirement. They permit employees to retire and begin collecting all or part of their pension annuities, and continue working for a limited period of time. From the date the teacher enters the program forward she no longer contributes to the pension plan, nor does she accrue additional service or benefits. The annuity payments are usually put into an escrow account while the teacher continues to work, and become available with interest when she stops (i.e., discontinues covered employment). Under a DROP, a teacher who takes
a partial pension from period $m$ to $m + n - 1$, then leaves the teaching force in period $m + n$, has the expected utility:

$$
E_t V_t(m) = E_t \left\{ \sum_{s=t}^{m-1} \beta^{s-t}\left[ (k_s (1 - c) Y_s)^\gamma + \omega_s \right] + \sum_{s=m}^{m+n-1} \beta^{s-t}\left[ (k_s (Y_s + \alpha B_s))^\gamma + \omega_s \right] 
\right. 
\left. + \sum_{s=m+n}^{T} \beta^{s-t}(B_s)^\gamma + \xi_s \right\}.
$$

(6)

Again, all repeating variables and parameters in equation 6 are as defined in equation 1. The addition of the middle term, including $(k_s (Y_s + \alpha B_s))^\gamma$, captures the nature of the DROP—namely, income is equal to the value of the annuity during years of work under the DROP, $B_s$, multiplied by the replacement rate, $\alpha$, and a salary during DROP-covered years from which no pension contributions are taken. Note that equations 5 and 6 do not introduce any new parameters to the model, just new income values during work and retirement.

Although not all states have DROPs, several have implemented them in plans covering teachers. For example, teachers in Arkansas for many years have had the option of participating in a DROP where they receive roughly 70 percent of their pension annuity and can continue to work full time for up to ten years. Similarly, Florida teachers can retire, have the annuity deposited in an escrow account, and continue in full-time employment for up to five years, after which they terminate employment and collect their accumulated annuity payments plus interest. The Louisiana teacher retirement system also provides a DROP option for retirement-eligible teachers for up to three years. These DROP programs are entirely voluntary and, as such, there should be no legal problems with this type of pension reform, as the experience with DROPs in many states indicates.

Although several states offer or have offered this option for educators, in every case of which we are aware it is an untargeted program. That is, it is available to all retirement eligible teachers. We are aware of no cases of a targeted DROP but in principle there is no reason a targeted DROP would not be possible. We consider DROPs for STEM teachers with different levels of annuity replacement while working. All of the plans we consider permit recipients to collect annuities for one year when they hit certain experience benchmarks, after which they must retire. We consider cases where teachers collect 30, 70, and 100 percent of the regular annuity. The former cases produce savings to the pension plan, but may remain attractive to teachers because they permit significantly higher incomes during DROP years because partial retirement payments can be collected at the same time as salaries. A key parameter of interest in the policy simulations is the experience benchmarks at which the DROP is offered.

**Bonus Simulation Results**

Table 4 reports simulated changes in teaching years and costs per additional teaching year for single-year $10,000 bonus offers to our cohort of STEM teachers at experience benchmarks ranging from 26 to 35 years. The average salary for our 2011 cohort of STEM teachers is $51,542, thus the bonus is roughly equivalent to 20 percent of salary (this and all subsequent values are in 2011 dollars). We use the structural model to
Table 4. Effects on Science, Technology, Engineering, and Math Teacher Retention: Search Over Various Years of Service (Bonus = $10,000)

<table>
<thead>
<tr>
<th>Year of receiving bonus (R)</th>
<th>26</th>
<th>27</th>
<th>28</th>
<th>29</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΔY (additional teaching years)</td>
<td>94</td>
<td>97</td>
<td>96</td>
<td>92</td>
<td>93</td>
</tr>
<tr>
<td>Switchers</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Non-Switcher Recipients</td>
<td>501</td>
<td>503</td>
<td>479</td>
<td>438</td>
<td>389</td>
</tr>
<tr>
<td>Non-Recipients</td>
<td>987</td>
<td>1,103</td>
<td>1,218</td>
<td>1,330</td>
<td>1,433</td>
</tr>
<tr>
<td>ΔY before R</td>
<td>29</td>
<td>33</td>
<td>35</td>
<td>37</td>
<td>37</td>
</tr>
<tr>
<td>ΔY at or after R</td>
<td>65</td>
<td>64</td>
<td>61</td>
<td>56</td>
<td>56</td>
</tr>
<tr>
<td>ΔPensionWealth/ΔY</td>
<td>$17,619</td>
<td>$16,852</td>
<td>$16,247</td>
<td>$15,583</td>
<td>$13,329</td>
</tr>
<tr>
<td>Bonus cost/ΔY</td>
<td>$54,671</td>
<td>$53,292</td>
<td>$51,359</td>
<td>$48,900</td>
<td>$43,078</td>
</tr>
</tbody>
</table>

Total Cost/ΔY (sum of previous two rows) | $72,290 | $70,144 | $67,606 | $64,483 | $56,407 |

<table>
<thead>
<tr>
<th>Year of receiving bonus (R)</th>
<th>31</th>
<th>32</th>
<th>33</th>
<th>34</th>
<th>35</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΔY (additional teaching years)</td>
<td>112</td>
<td>63</td>
<td>55</td>
<td>49</td>
<td>43</td>
</tr>
<tr>
<td>Switchers</td>
<td>18</td>
<td>12</td>
<td>11</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>Non-Switcher Recipients</td>
<td>368</td>
<td>309</td>
<td>265</td>
<td>225.92</td>
<td>189</td>
</tr>
<tr>
<td>Non-Recipients</td>
<td>1,516</td>
<td>1,620</td>
<td>1,701</td>
<td>1,771.98</td>
<td>1,832</td>
</tr>
<tr>
<td>ΔY before R</td>
<td>44</td>
<td>23</td>
<td>18</td>
<td>15.79</td>
<td>13</td>
</tr>
<tr>
<td>ΔY at or after R</td>
<td>68</td>
<td>40</td>
<td>37</td>
<td>32.71</td>
<td>30</td>
</tr>
<tr>
<td>ΔPensionWealth/ΔY</td>
<td>$12,439</td>
<td>$9,771</td>
<td>$7,141</td>
<td>$5,580</td>
<td>$5,145</td>
</tr>
<tr>
<td>Bonus cost/ΔY</td>
<td>$34,422</td>
<td>$51,101</td>
<td>$48,137</td>
<td>$48,668</td>
<td>$46,192</td>
</tr>
</tbody>
</table>

Total Cost/ΔY (sum of previous two rows) | $46,861 | $60,872 | $57,278 | $54,248 | $51,337 |

Notes: All values are rounded to nearest dollar or integer, thus totals may not sum due to rounding. R refers to the year in which the bonus is paid. The sample size is 2,131. All estimates are in 2011 dollars.

Simulate retirements over thirty years under permanent, single-year bonus policies. Because the bonus policies are permanent, all teachers know that they will receive the bonus if they reach the relevant threshold in the future. This allows the implementation of the bonus to affect work/retirement decisions for teachers leading up to the incentivized experience level. The first row of table 4 reports the total additional years of teaching generated by the bonus payment. Values range between about 94 and 112 years, and the maximum occurs at R = 31. In the baseline sample there are 1,901 teachers with less than 31 years of experience, so this amounts to less than 0.1 years per eligible teacher. The last row of the table (Total Cost / ΔY) reports the cost per additional year of teaching. The cost per incremental year of teaching varies greatly by R. The per-year cost is at a minimum of $46,861, also when R equals 31.

We now consider the cost factors at R = 31. In the absence of a bonus, 1,533 of the 2,131 teachers in our sample are predicted to retire before completing 31 years of service.

As noted above, an advantage of the bonus policy is that it can be implemented by individual districts, but we assume it is implemented statewide in the simulations to capture the full effect of teacher responses. A policy at a single district could have a smaller effect on teachers leading up to the bonus year because teachers would factor in the likelihood of leaving the district; of course, the bonus could also incentivize district-level retention. Also note that relatively few senior teachers transfer districts—this behavior is much more common among inexperienced teachers—which makes this issue less of a concern.
(p_1), 368 teachers are predicted to retire at or after 31 years (p_2), and others are policy-irrelevant population (i.e., teachers who have 31 or more years of service in the initial period of our sample). With the bonus, the composition of the same policy-relevant population changes to 1,516 and 385 teachers in these groups, respectively. There are 18 Switchers, 1,516 Non-Recipients who exit prior to completing the thirty-first year and thus do not receive the bonus, and 368 Non-Switcher Recipients.

The total number of additional years of teaching gained from the bonus is 112. However, not all of these additional years come from Switcher teachers. The total additional years can be decomposed into the increase in years among those retiring before attaining the bonus, 44, and the increase in years among those who attain the bonus, 68. Thus, roughly 40 percent of the additional years are due to the increase in years by teachers who retire without becoming eligible to collect the bonus. These are forward-looking teachers whose pre-R retirement decisions are modified by the presence of the bonus policy, but who are shocked into retirement prior to bonus collection for some reason. Note that although these teachers account for a substantial fraction of the total increase in years of experience generated by the bonus, the per-teacher effect among this group is small—that is, the total increase in teaching years of 44 is spread among the 1,533 teachers who separate prior to R, giving a per-teacher effect of just 0.03 years (equivalent to having 3 percent of the teachers who retire before attaining the bonus delay their retirement by one year).

The analysis in table 4 suggests that giving the bonus to teachers at R = 31 is the most attractive policy option, but we also consider the broad channels that drive the effects and costs of the bonus policy throughout the table. First, on the policy-effect side, as the threshold value of R in table 4 increases from 26 to 35, the number of Non-Switcher Recipient teachers drops from 501 to 189, highlighting an unsurprising reduction in the delivery of “wasted” bonuses as R increases. On the other hand, the number of switchers—that is, Switcher teachers—does not change monotonically with R, and reaches the maximum with R equal to 31. This is intuitive because if R is well before the pension peak, then teachers taking the bonus are likely to continue teaching in its absence, so the number of Switchers is low. Alternatively, as R moves past the retirement peak fewer teachers are working and eligible for the bonus, hence the number of Switchers is also low.

Table 4 also sheds light on the cost per additional year. The cost is net of salary and reported in the table per additional year of teaching. The full cost includes the cost of the bonus and the change in pension wealth. The total cost can be reduced in many ways. Table 4 makes clear that a particularly important dimension of the problem is reducing inframarginal bonuses to Switcher teachers. The cost of bonuses to Switcher teachers is highest at low values of R because teachers are very likely to work during low-R years regardless of the bonus. Costs per year of additional teaching are also reduced by polices that generate more Switchers, more additional years from teachers

---

13. By net of salary we mean the cost per incremental year estimates do not include teacher salaries. If we assume the size of the teaching workforce does not change as a result of the bonus policy, a one-year delay in retirement means a one-year delay in hiring a replacement. Most replacements will be novices or teachers with low seniority, hence with lower average salaries. However, we have not included such “costs” in this calculation, because novice teachers are, on average, less effective. The short-term salary cost savings would come at some expense in quality.
Table 5. Effects on Science, Technology, Engineering, and Math Teacher Retention of Different Bonus Levels (R = 31)

<table>
<thead>
<tr>
<th>Bonus</th>
<th>$10,000</th>
<th>$20,000</th>
<th>$30,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year of receiving bonus (R)</td>
<td>31</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>ΔY (additional teaching years)</td>
<td>112</td>
<td>226</td>
<td>340</td>
</tr>
<tr>
<td>Switchers</td>
<td>18</td>
<td>35</td>
<td>51</td>
</tr>
<tr>
<td>Non-Switcher Recipients</td>
<td>368</td>
<td>368</td>
<td>368</td>
</tr>
<tr>
<td>Non-Recipients</td>
<td>1516</td>
<td>1499</td>
<td>1482</td>
</tr>
<tr>
<td>ΔY before R</td>
<td>44</td>
<td>90</td>
<td>135</td>
</tr>
<tr>
<td>ΔY at or after R</td>
<td>68</td>
<td>136</td>
<td>205</td>
</tr>
<tr>
<td>Δ Pension Wealth/ΔY</td>
<td>$12,439</td>
<td>$12,479</td>
<td>$12,519</td>
</tr>
<tr>
<td>Bonus cost/ΔY</td>
<td>$34,422</td>
<td>$35,610</td>
<td>$36,967</td>
</tr>
</tbody>
</table>

Total Cost/ΔY (sum of previous two rows) $46,861 $48,089 $49,486

Notes: All values are rounded to nearest dollar or integer, thus totals may not sum due to rounding. The number of Non-Switcher Recipients is not affected by the size of the bonus if R is fixed (i.e., teachers who were going to work 31 or more years even in the absence of the bonus). All estimates are in 2011 dollars.

who ultimately do not collect the bonus, and by reducing pension liabilities by extending teaching careers past the pension peak. With regard to the latter, higher values of R result in the largest reductions in pension liabilities (which is consistent with figure 1). The “optimal choice” of R requires striking a balance of all of these factors; again, R = 31 is the most efficient choice over the grid search performed in table 4.

Table 5 expands on the analysis of the R = 31 bonus by considering different levels of bonus payments. The table shows that when the dollar value of the bonus triples from $10,000 to $30,000, the increase in STEM teaching years roughly triples as well, from 112 to 340, so the average cost per additional year rises only slightly.

DROP Simulation Results

Recall that a DROP allows a teacher to retire, collect her pension, and continue full-time work for a limited period of time, after which she must stop teaching. It is the latter feature that differentiates a DROP from a bonus policy, because with the bonus policy teachers can collect the bonus and are free to continue to work as many additional years as they choose. This has important consequences. In this section, we explore the effects of a simple DROP that permits a straightforward comparison with a one-year bonus. In this case, the STEM teacher who reaches R years of service is eligible to participate in a one-year DROP during that year. She is allowed to work full time and collect her regular salary, and also collect some fraction of her retirement annuity (denoted by α). But if she enters the DROP, she must exit teaching in the system in year R + 1.14

14. The DROP analyzed here is simpler than those that have been implemented in practice in that it is for one year only. The typical DROP allows participation for more than one year. If a DROP allows for a period of voluntary participation (e.g., in Arkansas for up to ten years), in practice it is operating like a multiyear retention bonus. Analysis of multiyear DROPs is more complex and beyond the scope of this paper. However, our analysis of the simple one-year plan provides useful insights regarding the design of longer-term plans.
Table 6. Effects of a One-Year Deferred Retirement Plan (DROP) on Science, Technology, Engineering, and Math Teacher Retention

<table>
<thead>
<tr>
<th>Year of Receiving DROP Payment (R)</th>
<th>31 Years of Teaching</th>
<th>32 Years of Teaching</th>
<th>33 Years of Teaching</th>
<th>34 Years of Teaching</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\alpha = 0.3$</td>
<td>$\alpha = 0.7$</td>
<td>$\alpha = 1.0$</td>
<td>$\alpha = 0.3$</td>
</tr>
<tr>
<td>$\Delta Y$ (additional teaching years)</td>
<td>81</td>
<td>148</td>
<td>269</td>
<td>124</td>
</tr>
<tr>
<td>Switchers</td>
<td>87</td>
<td>104</td>
<td>124</td>
<td>93</td>
</tr>
<tr>
<td>Non-Switcher Recipients</td>
<td>368</td>
<td>368</td>
<td>368</td>
<td>309</td>
</tr>
<tr>
<td>Non-Recipients</td>
<td>1447</td>
<td>1429</td>
<td>1409</td>
<td>1539</td>
</tr>
<tr>
<td>$\Delta P_{\text{Pension Wealth}} / \Delta Y$</td>
<td>$-32,805$</td>
<td>$-3,435$</td>
<td>$9,612$</td>
<td>$-16,598$</td>
</tr>
<tr>
<td>DROP / $\Delta Y$</td>
<td>$67,160$</td>
<td>$44,781$</td>
<td>$28,432$</td>
<td>$43,212$</td>
</tr>
<tr>
<td>Total Cost / $\Delta Y$</td>
<td>$34,355$</td>
<td>$41,346$</td>
<td>$38,044$</td>
<td>$26,613$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year of Receiving DROP Payment (R)</th>
<th>31 Years of Teaching</th>
<th>32 Years of Teaching</th>
<th>33 Years of Teaching</th>
<th>34 Years of Teaching</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\alpha = 0.3$</td>
<td>$\alpha = 0.7$</td>
<td>$\alpha = 1.0$</td>
<td>$\alpha = 0.3$</td>
</tr>
<tr>
<td>$\Delta Y$ (Additional Teaching Years)</td>
<td>82</td>
<td>218</td>
<td>348</td>
<td>64</td>
</tr>
<tr>
<td>Switchers</td>
<td>65</td>
<td>107</td>
<td>127</td>
<td>47</td>
</tr>
<tr>
<td>Non-Switcher Recipients</td>
<td>265</td>
<td>265</td>
<td>265</td>
<td>226</td>
</tr>
<tr>
<td>Non-Recipients</td>
<td>1647</td>
<td>1605</td>
<td>1586</td>
<td>1735</td>
</tr>
<tr>
<td>$\Delta P_{\text{Pension Wealth}} / \Delta Y$</td>
<td>$-21,690$</td>
<td>$3,559$</td>
<td>$11,488$</td>
<td>$-21,056$</td>
</tr>
<tr>
<td>DROP / $\Delta Y$</td>
<td>$47,143$</td>
<td>$27,200$</td>
<td>$19,782$</td>
<td>$43,145$</td>
</tr>
<tr>
<td>Total Cost / $\Delta Y$</td>
<td>$25,453$</td>
<td>$30,759$</td>
<td>$31,270$</td>
<td>$22,089$</td>
</tr>
</tbody>
</table>

Notes: All values are rounded to nearest dollar or integer, thus totals may not sum due to rounding. The number of Non-Switcher Recipients is not affected by the size of the bonus if $R$ is fixed (i.e., teachers who were going to work 31 or more years even in the absence of the bonus). $R$ is the year in which the teacher collects both salary and part or all of pension. All estimates are in 2011 dollars. Numbers in bold represent teachers who receive the DROP payment at their thirty-second year but join at the thirty-first year.

The results for the one-year DROP at different experience benchmarks, and different pension replacement rates during the DROP year, $\alpha$, are summarized in table 6. The DROPs generate the most additional years of teaching when provided to teachers in their thirty-second year (the portion of the table highlighted in boldface—they receive the DROP payment at their thirty-second year but join at the thirty-first year). The cost per incremental year is similar if the DROP is offered for the thirty-third year, but with a smaller total impact on $Y$. The scope for impact is further reduced at the thirty-fourth year, but with some cost savings per incremental year. One reason the DROP is relatively less effective in changing behavior—and less cost-effective—at $R = 31$ is that under the DROP teachers do not receive credit for the thirty-first year of service (recall that work during the DROP year does not count toward service years by design), which in Missouri is associated with a formula-factor increase per table 1 (also see Koedel, Ni, and Podgursky 2014). This dulls the behavioral response. In contrast, under the bonus policy the thirty-first year would still count toward system service and, thus, the bonus enhances the system incentive. This issue exemplifies the importance of modeling the types of policies we consider within the context of pre-existing plan incentives.

Table 6 also shows differential effects under different annuity replacement rates, given by $\alpha$. Unsurprisingly, holding the value of $R$ fixed increases in $\alpha$ correspond to increases in the number of years of additional teaching generated by the policy. The cost per incremental year is not a monotonic function of the annuity replacement rate.
Generally speaking, comparing the results in tables 4 and 6 shows that a DROP is less expensive to implement than a bonus. This is because in contrast to a bonus, the DROP forces some Non-Switcher Recipient teachers (i.e., those who planned to stay past 32 years anyway) to self-identify and reject the DROP. To see why, imagine there is a group of teachers who strongly prefer teaching over retirement (i.e., these teachers have a large, positive value of $v_s$ in equation 2). These teachers would all collect the retention bonus. However, some of them would reject the DROP because it forces them to begin collecting the annuity at the thirty-first year of service and quit teaching after the following year. The DROP thus more effectively targets marginal teachers, although this comes at a cost in the form of reduced flexibility at the local level. For example, in the more flexible DROPs implemented in practice, district leaders may have more discretion with regard to how long specific individuals participate. The DROPs we simulate do not have this flexibility, but our analysis highlights the important role in terms of cost control played by the strong commitment mechanism. Our results suggest that plans with long, open-ended retirement windows pay a large premium for this flexibility.

Table 6 also shows that the DROP generates substantially more savings in terms of pension wealth for each year of induced teaching at lower values of $\alpha$. This is because teachers do not accrue any additional pension benefits during the DROP year, and agree to receive just a fraction of the full pension payment during the DROP year in compensation. Figure 4 summarizes the per year costs and expected total teaching years gained from the bonus and the more remunerative (and thus likely more appealing to educator organizations) DROPs. This illustrates the lower per year costs and the larger labor supply impact from the DROPs.
Discussion

It should be noted that in the case of either a retention bonus or selective DROP, targeting of the retention incentives is critical to cost efficacy. There are two aspects of targeting that are important. The first is targeting the incentives to selected teachers rather than all teachers. Retention bonuses aimed at all teachers are a very expensive way to retain STEM teachers, or any other subgroup that policymakers identify as particularly valuable. Senior STEM teachers account for roughly 12 percent of all senior teachers. If the policy goal is to retain STEM teachers, then providing a retention bonus to all teachers would raise STEM per year costs roughly eight-fold. This is an important caveat as policies like the ones we consider have historically been implemented without clear targeting toward teachers in high demand.

The second aspect of targeting, discussed above, is with respect to retirement peaks inherent to the system under intervention. In Missouri, retirements are concentrated most strongly at the thirty-first year; plan rules in other states will generate different retirement patterns. The results above illustrate the importance of appropriate targeting. For example, table 4 shows that a $10,000 retention bonus provided at 31 years of experience in Missouri has a cost of about $47,000 per additional year of teaching generated. The same bonus provided at twenty-six years of experience has a cost of $72,000 due to the greater incidence of Non-Switcher Recipient teachers.\(^{15}\)

In this regard, it is interesting to note a recent study of a $5,000 retention bonus policy in Tennessee designed to retain high-quality teachers (i.e., teachers with a high score on a state teacher evaluation plan) in low-performing, high-need public schools (Springer, Swain, and Rodriguez 2016). These researchers found no significant effects overall but some positive effects for teachers in tested subjects. The authors are unable to calculate the cost per incremental year that our structural approach permits. However, they estimate that if the ratio of Switcher to Non-Switcher Recipient teachers was 1:4 (considerably higher than ours), then the cost per incremental year of teaching would be roughly $30,000, which is similar to our DROP and less than bonus estimates.

Finally, although these retention incentive schemes affect teacher retention for a number of years prior to the year in which the bonus is provided—hence the need for a dynamic analysis—it should also be noted that we estimate that these policies would have virtually no effect on turnover of early career teachers. By extension, we would not expect these types of plans to affect the teacher applicant pool, either. This is consistent with studies suggesting that new teachers have relatively little understanding of (or perhaps interest in) their retirement plans (DeArmond and Goldhaber 2010; Ettema 2011), or that they attach a low value to deferred relative to current benefits (Fitzpatrick 2015).

5. CONCLUSION

Traditional DB pension plans create strong “pull” and “push” incentives that concentrate teacher retirements at relatively early ages. Legal and political factors have

\(^{15}\) Modeling a DROP at 26 years of experience is complicated by the fact that most teachers would not be eligible for regular retirement at this milestone. Thus, we do not simulate such an option.
prevented reforms of these systems for incumbent teachers and the limited reforms
that have occurred are focused on new teachers. The benefits of reforms for new hires
will not be seen for many decades. In the meantime, policy makers have ignored the
potential for improving the teaching workforce by neutralizing the strong “push” incentives
built into these plans in a selective manner through targeted retention bonuses
for late career teachers. Using administrative microdata from Missouri, we simulate
the effects of two types of policies designed to postpone retirement for STEM teachers.
We show that selectively neutralizing the “push” incentives for STEM teachers via a
targeted DROP at their thirty-second year can yield additional teaching years by senior
teachers for about $30,000 per year, net of salary. Retention bonuses are somewhat
costlier, at roughly $47,000 per incremental year.

Are such policies worth the cost? If the bonuses are targeted to the most effective
STEM teachers then these policies would comfortably pass a cost–benefit test, particularly
since the counterfactual would be replacing a year of instruction by a highly effective
senior teacher with a year of instruction from a novice. For example, based on the
estimates of Chetty, Friedman, and Rockoff (2014), one year of instruction by a teacher
in the 95th percentile (as compared to the mean) yields a discounted benefit of roughly
$212,000. If targeting based on teacher quality is not possible, the cost effectiveness
of the policies we consider is less clear. Because roughly one-eighth of senior teachers
Teach STEM, the cost per additional STEM-year with untargeted retention schemes
would be roughly eight times what we report. However, even with a quality independ-
dent policy, available evidence showing that experienced teachers are more effective
on average than novices (e.g., Clotfelter, Ladd, and Vigdor 2006; Wiswall 2013) could
support some of the lower-cost DROP options we consider. If a policy were targeted
to improve equity as well—for example, by focusing on experienced STEM teachers in
high-poverty schools—the value of the equity benefits would further improve a cost-
effectiveness calculation.

The problem that the bonus and DROP programs we consider are designed
to ameliorate—that teachers retire at young ages relative to their professional
counterparts—is the product of strong incentives built into teachers’ final average salary
DB pension plans across the nation (Costrell and Podgursky 2009). Despite major fiscal
problems in these plans in many states, and policy shifts in a handful of states, there
is currently no indication of a broad shift away from this structure of retirement ben-
efits for teachers. Moreover, even in states where reforms have been undertaken, they
do not affect incumbents. Thus, working within the constraints of current plans, states
might consider exploiting the powerful retirement incentives in a strategic manner to
improve workforce quality. This strategy has been used for decades by the U.S. military
and might have benefits for public schools as well.

Aside from the specifics of this particular policy simulation, this exercise shows the
value of using structural retirement models to assess the short- and long-run workforce
effects of pension system reforms. Given the long time horizon over which changes
to pension plans can reshape the workforce, and the fact that existing pension plans
provide almost no scope for experimentation, research assessing the workforce effects
of pension reforms cannot make use of the traditional tools of policy evaluation. Mi-
crosimulation exercises such as the one in this paper can help policy makers make
more informed decisions with regard to pension plan reform.
ACKNOWLEDGMENTS

The authors wish to thank the Missouri Department of Elementary and Secondary Education for allowing use of their teacher data and Xiang Li for excellent research assistance. We gratefully acknowledge research support from the Laura and John Arnold Foundation and CALDER, which is funded by a consortium of foundations (for more information about CALDER funders, see www.caldercenter.org/about-calder). The views expressed here are those of the authors and should not be attributed to their institutions, data providers, or the funders. Any and all errors are attributable to the authors.

REFERENCES


