

Teacher Pension Incentives and the Timing of Retirement ¹

Shawn Ni

University of Missouri - Columbia

Michael Podgursky

University of Missouri - Columbia;

Fellow, George W. Bush Institute, SMU

Abstract

The rising costs and large unfunded liabilities of defined benefit (DB) teacher retirement systems raise questions about their efficacy and viability. Reform of teacher pension plans depends critically on reliable predictions of behavioral responses to alternative pension rules. We estimate an option-value model of individual teacher retirement using administrative data for Missouri teachers. The model fits the observed aggregate retirement behavior very well. We use the estimated structural parameters to simulate retirement behavior under alternative pension rules. Our simulations show that on net the enhancements of Missouri teacher pension benefits in the 1990's lowered the average retirement age for teachers. Conversion from the current DB plan to a defined contribution (DC) plan would have the opposite effect, and would dampen "spikes" in teacher retirement timing. The 1990's enhancements raised welfare for all teachers, however, the DC plan that we simulate has a mixed welfare impact, raising welfare for teachers near retirement but reducing it for teachers with less experience.

Keywords: teacher pensions, school staffing, school finance.

JEL codes: H30, I22, J26, J38.

¹The authors wish to acknowledge the research assistance of Nilay Chandra and Wei Zhou, and research support from the Center for Analysis of Longitudinal Data in Education Research (CALDER) at the Urban Institute. We also thank Mark Ehlert for valuable support and comments. The usual disclaimers apply.

1 Introduction

Teacher retirement benefits represent a large and growing cost for public school districts. Most teacher pension funds have large reported unfunded liabilities. In many cases these will rise as the recent stock market decline fully works its way into pension fund annual reports. However, even before the recent stock market meltdown, employer (and teacher) contribution rates were rising. In March 2004, employer costs for retirement benefits averaged 11.9 percent of salary nationally for public school teachers. By March 2011 these had increased to 15.4 percent. For private sector managers and professionals these costs were 10.4 percent in March 2011, with no upward trend (Costrell and Podgursky, 2009, updated). Reform of teacher pensions has been widely discussed by pension administrators and legislators. Changes have been implemented in several states. However, reliable estimates of the fiscal and the staffing effects of such changes requires reliable estimates of the behavioral effect of current and alternative pension rules, which is the subject of the present study.

A large literature in labor economics has analyzed the effects of incentives in pension systems on the timing of retirement decisions, labor turnover, and workforce quality (Friedberg and Webb, 2005; Asch, Haider, and Aissimopoulos, 2005; Ippolito, 1997; Stock and Wise, 1990). However, little of this literature pertains to teachers. While there have been many studies of the effect of current compensation on teacher turnover and mobility (e.g., Murnane and Olsen, 1990; Stinebrickner, 2001; Hanushek, Kain, and Rivkin, 2004; Podgursky, Monroe, and Watson, 2004), the literature on teacher pensions and their labor market effects is slender, but growing (Ferguson, et.al., 2006; Brown, 2009; Costrell and McGee, 2010; Friedberg and Turner, 2010). This issue takes on particular importance because current research points to a major effect of teacher quality on student achievement (Rivkin, Hanushek, and Kain, 2005).

To date, however, none of the papers examining teachers estimate structural models that are standard in the empirical retirement literature (e.g., Stock and Wise, 1990; Berkovic and Stern, 1991). As noted above, given concerns about the fiscal state of the pension funds and staffing schools with qualified teachers, a study of the effect of teacher pension plan incentives on teacher retirement behavior has obvious policy relevance. A structural economic approach provides a firmer foundation for policy simulations concerning teacher labor markets. However, an analysis of teacher pensions has more general research interest as well. This is a large market - roughly 3.2 million public school teachers. In addition, other professional staff (e.g., counselors and administrators) are in the same systems, yielding a total closer to 3.7 million. While the rules of defined benefit pension systems vary from state to state, the general structure of these systems are similar, as are the teachers themselves. Thus there is reason to believe that the results of a single state study like this one would generalize to a much larger universe.

The administrative data about the teachers and their pension systems in state data systems are generally of high quality and an excellent resource for research on the behavioral effects of pension plans. The rules of the teacher pension systems are typically more complicated than those of private sector plans, but they are readily available to outside researchers. These pension rules subject teachers to exogenous and peculiar incentives that allow researchers to study behavioral responses. Moreover, these peculiar rules have changed over time in ways that are readily documented.² State administrative data files also provide highly reliable data on teacher employment histories, salaries, and the exact timing of retire-

²For example, “rule of 80” permits regular retirement when $\text{Age} + \text{Experience} \geq 80$. While one might expect experience and age to have effects on retirement, there is no reason to expect an effect of the sum of the two passing a threshold of 80 to affect retirement, independent of pension rules. There are other such rules which produce sharp peaks and troughs in pension wealth accrual. See Costrell and Podgursky (2009) for further discussion.

ment. These administrative panel data are of very high quality compared to the household survey data that have been used in some other studies.³

There are other interesting features of teacher pension decisions. In modeling retirement in other markets, a worker's information on future wages or salaries may substantially differ from that known to the researcher. However, the salaries of teachers are determined by salary schedules and are highly predictable. Thus it is more likely that the teacher data offer a better test of the decision models commonly used in retirement research.

Finally, these exercises are usefully carried out in the framework of a structural model rather than reduced form models, since the "Lucas critique" (originally directed to the use of reduced-form models in macroeconomics) is relevant. Empirical regularities observed in a reduced-form models are the outcome of responses to pension plan incentives. If those incentives change, the empirical regularities change, possibly in complicated ways. Identification of "deep parameters" is valuable if researchers or policy-makers seek to understand the effect of changes in key features of these pension plans.

In the following section we discuss an "option value" dynamic panel model of retirement versus work. We report estimates of model parameters and show that the model fits our data very well. We then use the estimated model to simulate the effect of pension enhancements enacted during the 1990's and a hypothetical defined contribution (DC) alternative. We also explore the welfare effects of these plan alternatives. A concluding section summarizes our findings and suggests directions for further work in the area.

³There are tradeoffs. These administrative data are rich in information about the teachers, their employers, and their work histories. Unfortunately, our data file has no information about the teacher's household. In particular, we have no information about spousal income, or even whether the teacher is married.

2 Institutional Background

Missouri public school teachers, like nearly all public school employees, are covered by a defined benefit (DB) pension system. In fact, Missouri public school teachers are in three different DB systems. Teachers in the St. Louis and Kansas City districts, less than ten percent of teachers statewide, are covered by Social Security and are in their own pension systems. The rest of the public school teachers in the state are not covered by the Social Security system (as teachers) and are in a state-wide educator plan (Public School Retirement System, PSRS).⁴ Our focus in this paper is on teachers in the statewide plan. From a research perspective this is attractive since the incentives from the PSRS plan are not mingled with those of the Social Security system. Under current rules, Missouri teachers become eligible for a full (undiscounted) pension if they meet one of three conditions: a) sixty years of age and at least five years of experience, b) thirty years of experience (and any age), or c) the sum of age and years of service equals or exceeds 80 (“rule of 80”). Benefits at retirement are determined by the following formula (some variant of which is nearly universal in teacher DB systems):

$$\text{Annual Benefit} = S \times FAS \times R \quad (1)$$

where S is service years (essentially years of experience in the system), FAS is final average salary calculated as the average of the highest three years of salary, and R is the replacement factor. Teachers earn 2.5 percent for each year of teaching service up to 30 years. Thus, a teacher with 30 years experience and a final average salary of \$60,000 would receive: Annual

⁴The BLS reports that 72 percent of public school are covered by Social Security. State and local employees were not covered by the 1935 Social Security Act. Amendments in the early 1950’s permitted these employees to enter the system. Some groups of teachers (as a group) chose to enter, whereas others did not. The result is complicated mosaic. Usually, all teachers in a state are in or out (e.g., California out, Florida in, see Costrell and Podgursky (2009)).

Benefit = $30 \times \$60,000 \times 0.025 = \$45,000$. There are several other minor adjustments to the formula in equation (1). First, in order to provide teachers with assistance in purchasing health insurance, the average district contribution to individual teacher health insurance is included in FAS. Thus, if the average of the highest three salary years was \$60,000 and the average contribution to health insurance was \$3,000 annually, then FAS would equal \$63,000. Second, there is a “25 and out” option that permits retirement at a reduced rate if teachers have 25 or more years of experience. Finally, the value of R used in formula (1) is 2.5 for experience up to 30 years and 2.55 for experience of 31 or more years.

3 A Teacher’s Retirement Decision

Our focus is on the timing of retirement. We assume that an experienced educator who is teaching in the current year has two choices: teach next year or retire.⁵ In modeling the timing of retirement, Stock and Wise (1990) (SW) develop an structural model based on the “option value” method, in which a worker’s expected utility in period t is a function of expected retirement in year r (with $r = t, \dots, T$ and T is an upper bound on age). In period t , the expected utility of retiring in period r is the discounted sum of pre- and post retirement expected utility

$$\mathbb{E}_t V_t(r) = \mathbb{E}_t \left\{ \sum_{s=t}^{r-1} \beta^{s-t} [(k_s Y_s)^\gamma + w_s] + \sum_{s=r}^T \beta^{s-t} [(B_s)^\gamma + \xi_s] \right\},$$

where $0 < k_s < 1$ captures disutility of working, Y is income, and B is the pension benefit. The unobserved innovations in preferences are AR(1):

$$w_s = \rho w_{s-1} + \epsilon_{ws}, \quad \xi_s = \rho \xi_{s-1} + \epsilon_{\xi s}.$$

⁵In this context “retire” can also mean stop teaching and collect a pension at a future date rather than immediately.

This specification assumes that the disutility of work, k_s , does not depend on age. This is a problematic assumption that, as will be seen below, is at variance with our data. SW relax this assumption by allowing k_s to change monotonically with age: $k_s = \kappa(\frac{60}{age})^{\kappa_1}$. The standard specification emerges if $\kappa_1 = 0$. We adopt this more general approach.

In period t there is only one decision the worker needs to make: to retire or continue teaching. The retirement decision is irreversible. Once a teacher retires, she cannot return to the same job.⁶ Because the future is uncertain and the teacher is not risk neutral, there is a value associated with keeping the option retirement open, hence this is termed an “option value” model.

The expected gain from retirement at age r over retirement in the current period is

$$\begin{aligned} G_t(r) &= \mathbb{E}_t V_t(r) - \mathbb{E}_t V_t(t) \\ &= \mathbb{E}_t \sum_{s=t}^{r-1} \beta^{s-t} (k_s Y_s)^\gamma + \mathbb{E}_t \sum_{s=r}^T \beta^{s-t} (B_s)^\gamma - \mathbb{E}_t \sum_{s=t}^T \beta^{s-t} (B_s)^\gamma + \mathbb{E}_t \sum_{s=t}^{r-1} \beta^{s-t} (w_s - \xi_s). \end{aligned}$$

There are three sources of uncertainty: the uncertainty of future earnings and benefits, uncertainty of survival, and uncertainty in aforementioned preference shocks. To make survival uncertainty explicit, for a teacher alive in year t we denote the probability of survival to period $s > t$ as $\pi(s|t)$. Now

$$\begin{aligned} G_t(r) &= \sum_{s=t}^{r-1} \pi(s|t) \beta^{s-t} \mathbb{E}_t (k_s Y_s)^\gamma + \sum_{s=r}^T \pi(s|t) \beta^{s-t} \mathbb{E}_t (B_s)^\gamma - \sum_{s=t}^T \pi(s|t) \beta^{s-t} \mathbb{E}_t (B_s)^\gamma \\ &\quad + \sum_{s=t}^{r-1} \pi(s|t) \beta^{s-t} \mathbb{E}_t (w_s - \xi_s). \end{aligned}$$

The sum of the first three terms is a function of current salary and experience, and is denoted $g_t(r)$. The last term $\sum_{s=t}^{r-1} \pi(s|t) (\beta \rho)^{s-t} (w_t - \xi_t)$ is unobservable and is denoted $K_t(r) = \sum_{s=t}^{r-1} \pi(s|t) (\beta \rho)^{s-t}$ (which depends on unknown parameters) times an error term

⁶Thus, we are ruling out the option of a teacher retiring and returning to a PSRS-covered job (“double-dipping”). PSRS rules make it very difficult to return to full time covered employment and collect a pension, although part-time teaching employment (less than half time) is an option.

$\nu_t = w_t - \xi_t$, which follows $\nu_t = \rho\nu_{t-1} + \epsilon_t$ where ϵ_t is assumed to be $N(0, \sigma^2)$. The retirement decision can thus be formulated as choosing $r = t, \dots, T$ that maximizes

$$G_t(r) = g_t(r) + K_t(r)\nu_t.$$

Let

$$r_t^\dagger = \operatorname{argmax} g_t(r)/K_t(r),$$

the probability that teacher retires in period t ($G_t(r) \leq 0$ for all $r > t$) is $\operatorname{Prob}(\frac{g_t(r_t^\dagger)}{K_t(r_t^\dagger)} \leq -\nu_t)$.

The likelihood can be specified under a normality assumption on ν_t and given rules for predicting future earnings. We assume salary is predictable under an estimated nonlinear (a third order polynomial) function of experience.⁷ For estimation of the model, if a teacher $i \in \{1, \dots, I\}$ retires in period t , $d_{it} = 1$, otherwise $d_{it} = 0$. After retirement the teacher is dropped out of the sample (after T_i for teacher i). For cross-section data with a teacher i observed only in period t , the likelihood is

$$L(\gamma, k, \beta, \sigma, \rho \mid \mathbf{Y}, \mathbf{B}, \mathbf{D}) \propto \prod_{i=1}^I \Phi\left(\frac{g_t(r_t^\dagger)}{K_t(r_t^\dagger)} / \sigma_\nu\right)^{d_{it}} (1 - \Phi\left(\frac{g_t(r_t^\dagger)}{K_t(r_t^\dagger)} / \sigma_\nu\right))^{1-d_{it}},$$

where $\Phi(\cdot)$ is the cumulative density function of standard normal and σ_ν is the standard deviation of ν_t . For panel data the likelihood is made more complicated by the serial correlation of ν_t . Suppose a teacher is observed for period $t, t+1, \dots, t+n$ and she retired in $t+n$, then the likelihood is the probability of the joint event $\pi\left(\frac{g_t(r_t^\dagger)}{K_t(r_t^\dagger)} > -\nu_t, \dots, \frac{g_{t+n-1}(r_{t+n-1}^\dagger)}{K_{t+n-1}(r_{t+n-1}^\dagger)} >$

⁷Missouri teachers, like nearly all public school teachers, are paid according to salary schedules that set pay based on years of teaching experience and education credentials (frequently terminating in an MA). Thus it is not unrealistic to treat teacher pay as a function of teaching experience, assuming all teachers move from the BA column on the schedule over to the MA column with the passage of time. Because we focus on late-career teachers, the degree-related salary adjustment is largely absent in the sample. The fairly deterministic advancement over well-defined district salary schedules underlies the salary growth assumption in the text.

$-\nu_{t+n-1}, \frac{g_{t+n}(r_{t+n}^\dagger)}{K_{t+n}(r_{t+n}^\dagger)} < -\nu_{t+n}$). By the definition of conditional probability, one can view this joint probability as products of a sequence of conditionals:

$$\begin{aligned}
& \pi\left(\frac{g_t(r_t^\dagger)}{K_t(r_t^\dagger)} > -\nu_t, \dots, \frac{g_{t+n-1}(r_{t+n-1}^\dagger)}{K_{t+n-1}(r_{t+n-1}^\dagger)} > -\nu_{t+n-1}, \frac{g_{t+n}(r_{t+n}^\dagger)}{K_{t+n}(r_{t+n}^\dagger)} < -\nu_{t+n}\right) \\
= & \pi\left[\left(\frac{g_{t+n}(r_{t+n}^\dagger)}{K_{t+n}(r_{t+n}^\dagger)} < -\nu_{t+n} \mid \frac{g_t(r_t^\dagger)}{K_t(r_t^\dagger)} > -\nu_t, \dots, \frac{g_{t+n-1}(r_{t+n-1}^\dagger)}{K_{t+n-1}(r_{t+n-1}^\dagger)} > -\nu_{t+n-1}\right)\right] \\
& \times \pi\left[\left(\frac{g_{t+n-1}(r_{t+n-1}^\dagger)}{K_{t+n-1}(r_{t+n-1}^\dagger)} > -\nu_{t+n-1} \mid \frac{g_t(r_t^\dagger)}{K_t(r_t^\dagger)} > -\nu_t, \dots, \frac{g_{t+n-2}(r_{t+n-2}^\dagger)}{K_{t+n-2}(r_{t+n-2}^\dagger)} > -\nu_{t+n-2}\right)\right] \\
& \dots \\
& \times \pi\left[\left(\frac{g_{t+1}(r_{t+1}^\dagger)}{K_{t+1}(r_{t+1}^\dagger)} > -\nu_{t+1} \mid \frac{g_t(r_t^\dagger)}{K_t(r_t^\dagger)} > -\nu_t\right)\right] \\
& \times \pi\left[\frac{g_t(r_t^\dagger)}{K_t(r_t^\dagger)} > -\nu_t\right].
\end{aligned}$$

Denote $\boldsymbol{\nu}_{t,t+n} = (\nu_t, \dots, \nu_{t+n})'$. The event $\left(\frac{g_t(r_t^\dagger)}{K_t(r_t^\dagger)} > -\nu_t, \dots, \frac{g_{t+n-1}(r_{t+n-1}^\dagger)}{K_{t+n-1}(r_{t+n-1}^\dagger)} > -\nu_{t+n-1}, \frac{g_{t+n}(r_{t+n}^\dagger)}{K_{t+n}(r_{t+n}^\dagger)} < -\nu_{t+n}\right)$ which can be expressed as $\boldsymbol{\nu}_{t,t+n} \in A_{t,t+n}$ in space R^n . The marginal distribution of $\nu_t \sim N(0, \sigma_\nu^2)$ where $\sigma_\nu^2 = \frac{\sigma^2}{1-\rho^2}$. Given $\nu_t = \rho\nu_{t-1} + \epsilon_t$, the covariance of $\boldsymbol{\nu}_{t,t+n}$ is given by

$$\boldsymbol{\Sigma} = \frac{\sigma^2}{1-\rho^2} \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{n-1} & \rho^{n-1} \\ \rho & 1 & \rho & \dots & \rho^{n-2} & \rho^{n-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{n-3} & \rho^{n-3} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \rho^{n-1} & \rho^{n-2} & \dots & \dots & \rho & 1 \end{pmatrix}.$$

The log likelihood is

$$\log L(\gamma, k, \beta, \sigma, \rho \mid \mathbf{Y}, \mathbf{B}, \mathbf{D}) = \sum_{i=1}^I \log \pi_i(\boldsymbol{\nu}_{t,t+n} \in A_i) = \sum_{i=1}^I \log \int_{A_i} \phi(\boldsymbol{\nu}_{t,t+n}) d\boldsymbol{\nu}_{t,t+n}, \quad (2)$$

where for teacher i retiring in period $t+n$, $\boldsymbol{\nu}_{t,t+n} \in A_i$ if $\frac{g_t(r_t^\dagger)}{K_t(r_t^\dagger)} > -\nu_t, \dots, \frac{g_{t+n-1}(r_{t+n-1}^\dagger)}{K_{t+n-1}(r_{t+n-1}^\dagger)} > -\nu_{t+n-1}, \frac{g_{t+n}(r_{t+n}^\dagger)}{K_{t+n}(r_{t+n}^\dagger)} < -\nu_{t+n}$, and $\phi(\cdot)$ denotes multivariate normal density distribution of

$N(0, \mathbf{\Sigma})$. An obstacle to evaluating the likelihood is the large computational time of n - dimensional integration. Even for a moderate size n (say 5), deterministic methods for numerical integration can be prohibitively costly. In this study, we solve the problem through Monte Carlo simulation. The covariance matrix $\mathbf{\Sigma}$ permits a Cholesky decomposition $\mathbf{\Sigma} = \mathbf{V}\mathbf{V}'$. The algorithm for computing $\int_{A_i} \phi(\mathbf{v}_{t,t+n}) d\mathbf{v}_{t,t+n}$ is as follows: (1) Draw $\mathbf{e}^{\{m\}}$ from $N(0, I)$ ($m = 1, \dots, M$) and let $\mathbf{v}_{t,t+n}^{\{m\}} = \mathbf{V}\mathbf{e}^{\{m\}}$. (2) Use the frequency $\frac{1}{M} \sum_{m=1}^M I(\mathbf{v}_{t,t+n}^{\{m\}} \in A_i)$ to approximate $\int_{A_i} \phi(\mathbf{v}_{t,t+n}) d\mathbf{v}_{t,t+n}$. $I(\mathbf{v}_{t,t+n}^{\{m\}} \in A_i) = 1$ if $\mathbf{v}_{t,t+n}^{\{m\}} \in A_i$ and $I(\mathbf{v}_{t,t+n}^{\{m\}} \in A_i) = 0$ otherwise. This sampling method is more efficient in regions of high likelihood and less so at the tails of the distribution. It is therefore suitable for MLE.

The option value model above assumes that a teacher chooses the year of retirement that maximizes the expected present value of the utility of the salary and benefit flows given current information. In a dynamic programming setting, a teacher evaluates the expectation of the value of salary and benefit flow under present and future optimal choices. Hence the option value model does not take into account the value of options in the future. The gain from this is a simpler derivation of the empirical model. Stern (1997) shows that the option value model may yield different results from those obtained by dynamic programming. Lumsdaine, Stock and Wise (1992) argue that it is not obvious that the more sophisticated dynamic programming model is more realistic for modeling actual retirement decisions. They find that the predictive performance of the option value model is comparable to that of a dynamic programming approach. As we will see below, the SW option value model fits our data very well.

A “peak value” approach has been used in some applied retirement studies (e.g., Coile and Gruber, 2007; Friedburg and Webb, 2005). It can be treated as a special case of the SW model, in which the teacher chooses the timing of retirement to maximize the present value of her expected pension wealth. This implies the following restrictions: $\kappa = 0, \gamma =$

1, $\sigma = \rho = 0$. Setting the discount rate β to be the inverse of one plus the nominal interest rate, the peak value model corresponds to the SW model and the objective function becomes $\mathbb{E} \sum_{s=t}^T \frac{B_s}{(1+r)^{s-t}}$, where the expectation is with respect to survival probability.

The forward looking retirement decision model is in contrast to reduced-form models such as commonly used probit or logit models, where the retirement decision is made in a static setting. In a probit model, assume that the iid error term ϵ_{it} follows a normal distribution $N(0, \eta^2)$, and that a teacher with characteristics \mathbf{x}_{it} retires if $y_{it}^* = \mathbf{x}_{it}'\boldsymbol{\theta} + \epsilon_{it} > 0$. Then

$$L(\boldsymbol{\theta}|\mathbf{X}, \mathbf{d}) = \prod_{i=1}^I \prod_{t=1}^T \Phi(\mathbf{x}_{it}'\boldsymbol{\theta}/\eta)^{d_{it}} [1 - \Phi(\mathbf{x}_{it}'\boldsymbol{\theta}/\eta)]^{1-d_{it}}. \quad (3)$$

The logarithm of likelihood (3) takes the form similar to (2) with the exception that in (2) the variables inside the normal CDF are nonlinear and depict a series of correlated decisions.

While reduced form logit or probit models may fit the retirement data reasonably well, they depict an empirical relationship between the covariates (constant, experience, and age) and retirement decisions under the prevailing pension rules. Changing these rules potentially changes the empirical relationship. Hence, these reduced form estimates are vulnerable to the ‘‘Lucas critique’’ and are thus problematic for analyzing counterfactual outcomes under alternative pension regimes. The forecasts of structural models, on the other hand, are determined by pension rules and ‘‘deep’’ parameters characterizing the teacher preferences. The deep parameters do not change with the pension rules, which permits experiments with pension policies.

4 Data and Estimates

Our data consists of a cohort of 9605 Missouri teachers aged 50-55 at the beginning of the 2002-03 school year. We tracked this cohort of teachers forward to the 2008-09 school year. Table 1 shows that roughly eighty percent of teachers in the sample are female. Over the six

year period from the 02-03 to 07-08 school year, roughly half of the teachers in the cohort retired.

The first column of Table 2 reports maximum-likelihood estimates of the structural parameters in the retirement model: $\kappa, \kappa_1, \beta, \gamma, \sigma, \rho$. We begin with the pooled estimates in the first two columns. All of the parameter estimates are statistically significant, of the right sign, and of reasonable magnitude. The parameter β reflects the rate of time preference for the teacher, in our case suggesting a roughly four percent discount rate. The parameter k measures the value of work versus leisure time. The disutility of working is modeled as $k_s = \kappa(\frac{60}{age})^{\kappa_1}$. If $k_s = 1$ then there is no disutility associated with teaching. Our estimates are $\kappa = 0.654$ and $\kappa_1 = 1.763$, which imply that the disutility of working (i.e., the retirement benefit equivalent of one dollar of salary) is 0.901 at age 50, 0.653 at age 60 and 0.510 at age 70. Figure 1 plots the relationship between estimated disutility of teaching (relative to pension income) and age. Allowing for age-dependency in the disutility of teaching substantially improves the fit of the model. The second column of Table 2 shows that restricting the measure of disutility k_s to be a constant (by setting κ_1 to 0) results in a large drop in the likelihood (relative to the test statistic of $\chi^2(1)$.)

The estimate of the discount factor β is about 0.96. The point estimate of γ is significantly less than unity, indicating risk aversion. The large value of σ indicates a good deal of heterogeneity in preferences. This is not surprising since we have no covariates in the model. One might expect various household and personal factors such as spouse's pension, teacher health, and the teacher's preference for teaching to affect the timing of retirement. These and other factors are picked up in σ . In addition, these omitted factors tend to persist over time, as indicated by large and significant values for ρ .

Table 2 also reports estimates for males and females separately. The point estimates are fairly similar, with the exception of κ_1 . In both cases the data support the model with age-

dependent disutility of working. The preference parameter κ_1 of male teachers is 3.316 while that for female teachers is 1.681. This suggests that as male teachers age, their disutility of teaching relative to retirement rises more quickly than for female teachers. In addition, the mortality rate of general population of males is higher (0.748% at age 55) than that of females (0.434% at the same age). These factors predict earlier retirement of male teachers. The other estimates are similar for male and female teachers.

As noted above, a number of articles in the literature have estimated peak value models rather than a full structural model. We also consider the peak value constrained version of the model: $k = 0, \gamma = 1, \sigma = \rho = 0, \beta = 0.95$. With these constraints, the log-likelihood under the peak value model is -73812.829 . Two times the difference in the log-likelihood between SW and peak value models can be used for hypothesis testing. The critical value of $\chi^2(5)$ at the 1% level is 15.09. The peak-value model is rejected overwhelmingly.

4.1 Goodness of fit

For all teachers aged 50-55 in 2002 (our baseline sample), we use the estimated parameters of the structural model and the information on these teachers in 2002 to generate the probability that each teacher took one of the following 7 actions: retired in year 2003, retired in 2004, ..., retired in 2008, and remained in teaching workforce in 2008. The probabilities are obtained through Monte Carlo simulation. Specifically, for each teacher in the 2002 sample, regardless of the actual retirement decision the teacher took, we draw 6 serially correlated error terms ϵ_t ($t = 2002, \dots, 2007$). If according to the SW model, with the realized error terms of ϵ_{2002} and given the age, salary, and experience, the teacher should choose to retire in 2002, then for that draw the teacher is recorded as retired in 2003. If the model predicts that teacher chooses not to retire in 2002, then we project the 2003 salary and add one year to the age and experience. If the model predicts retirement given the ϵ_{2003} draw and the new state variables, then the

teacher is recorded as retired in 2004. We repeat the process to 2007. If model predicts the teacher chooses not to retire up to 2007, then the teacher is recorded as a non-retiree at the end of the sample. For each teacher we replicate the above experiment a large number of times (100,000, changing it to 1,000,000 produces the same results). The frequency of the simulated retirement decisions give rise to the predicted probabilities. We aggregate the probabilities over the teachers in the 2002 sample to obtain the aggregate predicted retirement. We present aggregated predicted and actual retirement by age, experience, and age by experience. Comparisons of the observed and predicted distributions of the retirees (at the time of separation) and non-retirees (who do not separate until 2008) are used to gauge the fit of the model.

Figure 2 shows the percentage of teachers in the 50-55 age cohort in 2002 who remain in teaching in each year until 2008. The predicted survival rate is fairly close to the observed rate except in the first year. The predicted 2003 retirement of the cohort is larger than the observed rate. We believe that this is related to our choosing (due to data constraints) a fixed five-year window of teachers nearing retirement. Table 1 shows that teachers who retired in different years have different characteristics. Those who choose to retire in 2002 (and retired in 2003) tend to be more experienced in 2002 but retire at relatively younger ages (the average is just over 53 years old.) The young retirement age in 2002 is a consequence of selecting a sample of 50-55 year-olds. We do not include the teachers who retired in 2002 and who are older than 55. Hence the number of retirees in 2003 in our panel substantially understates actual retirement in that year.

Suppose the data are generated by the option value model, where each teacher is associated with an unobserved preference shock ν_t in year t . Recall that the marginal condition for retirement is $\frac{g_t(r^\dagger)}{K_t(r^\dagger)} \leq -\nu_t$, and ν_t has a positive serial correlation and unconditional mean of zero. Those with a high value of $-\nu_{2002}$ also have high values of $-\nu_{2001}$ and are more likely

to retire before 2002. The remaining teachers who are eligible for retirement but who remain in the teaching workforce are more likely to have a low value of $-\nu_{2001}$, thus low $-\nu_{2002}$. The model simulation assumes zero mean in ν_{2002} , hence it over-estimates the probability of retirement in 2003. In the later part of the sample period such sample selection bias is greatly reduced.

The observed percentage of the non-retired teachers in 2008 is 49.9%, and the estimated rate is 50.1%. The aggregate behavior at the end of the sample is better matched by the model than that at the beginning of the sample. We noted above that sample selection adversely affects the fit in early years of the sample. In addition, the teachers who retire in the later part of the sample period or who remain in teaching workforce to the end of the sample period are observed more times than those who retired early (and thus dropped out of the sample). The former thus carry a heavier weight in the likelihood function. The MLE is obtained through matching the sample observations, thus the parameters are selected so the behavior of those with heavier weight are better matched by the model.

The remaining charts explore model fit by age and experience. The left panel of Figure 3 plots the observed (black solid line) and predicted (dotted line) age distribution of the teachers who choose to retire. The two distributions are very closely matched, with the exception that the model predicts more earlier retirement (at age 50) than observed. In the right panel of the figure we plot the age distributions of non-retirees. The predicted and observed distributions are matched almost perfectly (note that the youngest non-retiree in 2007 is 55 years old).

Figure 4 plots the observed (black solid line) and predicted (dotted line) experience distributions at the separation year (for the retired) or the end of the sample period (for the non-retired). The two distributions in the left panel of Figure 4 are reasonably matched. The predicted distribution overpredicts the spikes corresponding to the 25-and-out rule, and

under-predicts the spike at 30 years experience. By comparison, the experience distribution of the predicted non-retired is more closely matched to that of the observed.

The above comparisons are based on marginal distributions (e.g., for each age, retirement rates were averaged over all levels of experience.) Because age and experience are correlated, a more complete picture of model fit is in Figure 5, where we compare the observed joint age-experience distribution versus simulated counterparts. Overall the estimated model fits the joint age-experience distribution of retirement very well. Figure 5 shows that the teachers who are predicted to retire at lower ages by the SW model tend to have high experience (with 25 years or more). Hence, the over- prediction of retirement for the low age/high (more than 25 years) experience group may be due to the sample selection issue discussed above.

The structural model is not merely useful for explaining history. A potentially important use is as a tool for studying alternative pension policies. We turn to that task in the next section.

5 Evaluation of Pension Plan Changes

In this section we use the structural estimates to explore the behavioral and welfare effects of pension plan changes. Given the lively policy debate in this area, there are many options one might explore. We restrict our attention to two. First, we examine the effect of pension plan enhancements enacted during the 1990's. Like many states, Missouri substantially increased the generosity of the teacher plan during the bull market of the 1990's. We examine the effect of these enhancements on teacher retirement behavior. In response to rising costs for maintaining the DB pension plans, several states have closed the traditional plans and put new teachers into DC plans (or introduced "hybrid" plans, which combine a DB and DC plan). Many other states are considering such reforms. Thus, we consider a DC alternative

plan.

5.1 Effect of recent pension enhancements

As noted, the Missouri legislature passed a series of pension enhancements during the 1990's. Nearly every year between 1992-93 through 2000-01, when the last significant enhancement passed, one or more rule changes were implemented which increased teacher pension wealth (i.e., by either increasing the value of the pension annuity or increasing the number of years over which it can be collected). Ni, Podgursky, and Ehlert (2009) estimate that these enhancements raised the aggregate peak value pension wealth of the 2007 teaching workforce by roughly four billion dollars. The most expensive of these included an early retirement provision ("25 and out"), increases in the multiplier from 2.3 to 2.5 percent, and introduction of "rule of 80" for regular retirement. Rather than examine all of these enhancements individually, we simplify the analysis by simply "turning back to clock" to 1995.

5.2 DB to DC conversion

Both theory and empirical evidence suggest that the current defined benefit (DB) plans create strong incentives for retirement around the years associated with peak value of pension wealth. This pension-induced incentive can "pull" teachers who do not enjoy teaching to stay for additional years to collect the pension and "push" teachers who enjoy teaching into retirement sooner than they would otherwise prefer. Some states are considering a switch to DC plans, in total, or partially in "hybrid" plans, to reduce fiscal exposure as well as eliminate these perverse incentives.⁸

⁸Several studies in the general literature on retirement have examined the effect of transitions from DB to DC plans, e.g., Friedburg and Webb(2005), Poterba, Rauh, Venti, Wise (2007).

We consider the following DC plan: teachers contribute a mandatory fixed percent of salary ($c = 10\%$), matched by an equivalent employer contribution into each teacher’s account each year. A teacher’s account accumulates with annual contributions and nominal investment returns of $R - 1 = 4\%$ on the fund balance. We treat this as a guaranteed return (e.g., as with TIAA or a “cash balance” pension plan). The inflation rate is assumed to be $i = 3\%$. The account is portable and teachers can withdraw from the account at any age without penalty. When a teacher retires, the contribution to the account stops and an insurance company provides an actuarially fair annuity (B) equal to the cash value in the teacher’s account.

For a teacher with a DC account value W_t , age a in year t , and a maximum life of 101 years, let the expected nominal flow of the annuity be B_{t+n} in the n -th year of retirement. The retiree survives to $t + n$ with conditional probability $\pi(t + n|t)$. The expected account value and the expected payment evolve as

$$W_{t+n} = W_{t+n-1}R - B_{t+n}, \quad B_{t+n} = \pi(t + n|t)(1 + i)^n B.$$

We set $W_{101} = 0$. It follows that

$$B = \frac{W_t}{\sum_{n=1}^{101-a} \pi(t + n|t) \left(\frac{1+i}{R}\right)^n}. \quad (4)$$

The remaining question is how to determine the DC account value for a teacher who is in the current DB plan. We consider the following scenario. All teachers in the DB plan in 2002 have cash balance W based on the current rules of the DB plan.⁹ Further accrual of pension wealth under the old plan is halted. Going forward the value in this account grows by the nominal interest rate (on the fund balance) and further annual contributions from teachers and districts.

⁹We compute the cash value of the DB accruals using a 2 percent real discount rate and standard life tables. For references on the choice of the real discount rate, see Grissom, Koedel, Ni, Podgursky, (2011).

With this initial value in the DC plan, the teacher considers whether to retire or continue to work as in the SW model: a worker's expected utility in period t is a function of expected retirement in year r (with $r = t, \dots, T$ and $T = 101$ is an upper bound on age). In period t , the expected utility of retiring in period r is the discounted sum of pre- and post retirement expected utility

$$\mathbb{E}_t V_t(r) = \mathbb{E}_t \left\{ \sum_{s=t}^{r-1} \beta^{s-t} [(k_s(1-c)Y_s)^\gamma + w_s] + \sum_{s=r}^T \beta^{s-t} [(B_s)^\gamma + \xi_s] \right\},$$

where Y is income and B is pension benefit and the unobserved errors are AR(1): $w_s = \rho w_{s-1} + \epsilon_{ws}$, $\xi_s = \rho \xi_{s-1} + \epsilon_{\xi s}$.

The benefit $B_s = B$ given in (4) with W_t replaced by the real value $\frac{W_r}{(1+i)^{r-t}}$. Note the account nominal value in year $r > t$ is the value of accumulation of contributions plus the compound return of the wealth in period t : $W_r = W_t R^{r-t} + \sum_{k=t+1}^r 2cY_k R^{r-k}$.

Because the DC rules are simpler than the DB rules, we are able to formalize the marginal condition for retirement time under the DC rules and thereby gain some intuition about the tradeoff between work and retirement. Suppose in the absence of unobserved preference shifters the teacher with salary Y_t and pension wealth W_t is indifferent of retiring in year t (with a constant real pension flow of B after t) or $t + 1$ (with pension flow \tilde{B} at and after t .) Then

$$(k_t(1-c)Y_t)^\gamma + \sum_{s=t+1}^T \beta^{s-t} \pi(s|t) B^\gamma = \tilde{B}^\gamma + \sum_{s=t+1}^T \beta^{s-t} \pi(s|t) \tilde{B}^\gamma, \quad (5)$$

where $B = \frac{W_t}{\sum_{s=t+1}^T \pi(s|t) (\frac{1+i}{R})^{s-t}}$ and $\tilde{B} = \frac{W_t R + 2cY_t}{1 + \sum_{s=t+1}^T \pi(s|t) (\frac{1+i}{R})^{s-t}}$. Denote the constants $b_1 = \sum_{s=t+1}^T \pi(s|t) (\frac{1+i}{R})^{s-t}$, and $b_2 = \sum_{s=t}^T \beta^{s-t} \pi(s|t)$, then condition (5) can be written as

$$\left[b_1 k_t (1-c) \frac{Y_t}{W_t} \right]^\gamma + b_2 \left(R + 2c \frac{Y_t}{W_t} \right)^\gamma = (1 + b_2) \left(\frac{b_1}{1 + b_1} \right)^\gamma. \quad (6)$$

(6) implies that for a given age, under the DC plan a teacher chooses to retire when the ratio of salary to pension wealth is lower than a constant. The dynamics of the pension

wealth/salary ratio is given by $\frac{W_t}{Y_t} = R\left(\frac{W_{t-1}}{Y_{t-1}}\right)\left(\frac{Y_{t-1}}{Y_t}\right) + 2c\frac{Y_{t-1}}{Y_t}$. The pension wealth/salary ratio is increasing in the return to savings and increases over time as the real salary growth slows down at the later stage of teachers' careers. At some point the ratio $\frac{W_t}{Y_t}$ is large enough to render the LHS of (6) lower than the RHS of (6).

5.3 Policy simulations

We now examine the effect of the 1990's enhancements and our DC alternative plan on retirement behavior. In examining the goodness of fit of the model in the previous section, we were constrained to the seven year window of our panel data. In simulating the effect of these policies, there is no reason to restrict our time horizon so narrowly, thus we extend the forecast horizon to 20 years, by which time nearly all of these teachers will have retired.

Because the pension annuity B is increasing in initial pension wealth, the level at which pension wealth is set in the year of initial conversion from DB to DC plans affects the retirement decision. We set this initial level of pension wealth by discounting the flow of annuity payments at a two percent real rate and assuming a standard life table for survival rates. For teachers at or near the "peak value" of pension wealth, this can be an attractive option, the DC plan eliminates the penalty on working after reaching the peak value under the current rules (i.e., the "pushing out" effect of the current rules.) However, the DC plan does not necessarily postpone retirement. For some teachers, it is optimal to retire earlier under the DC than under the current rules. Whether this is case depends on the teacher's age, experience, and the initial 2002 pension wealth lump sum payment. As condition (6) shows, under the DC plan a teacher retires when the salary/pension wealth ratio is below a threshold. The higher the initial pension wealth, the earlier the retirement under the DC plan.

We begin by comparing survivor functions under the three regimes in Figure 6. Here we

see that under the 1995 rules the teachers remain employed at a higher rate over the period 2003-2010 than under the current rules and the gap narrows to insignificance by 2011-2021. Conversion to the DC plan would result in slightly higher retirement in the year 2003 than under the 1995 rules but possibly lower or higher than that under the current rules depending on initial pension wealth. Under the DC plan, the teaching survival rate declines much more slowly than under the two DB plans, regardless of the retirement in initial years. At the end of the forecast period, about 7% of the teachers are predicted to remain teaching under the DC plan, while around 1% are predicted to do so under the DB rules.

We plot the marginal distributions by age and experience under the current rules, the 1995 rules, and the DC plan in Figures 7 to 9. Note that these figures characterize the distribution of the teachers predicted to retire within 20 years. Because under the DC plan more teachers remain teaching at the end of the predicted period than under the DB plans, the magnitude of the curves are not entirely comparable. Figure 7 shows that under the 1995 rules retirement would have occurred at an older age. Under these rules, the retirement age distribution peaks at 60, and that there is a lower percentage of retirement occurring at age 50 and 51.¹⁰ Under the DC plan, Figure 7 also confirms what we learn from the predicted survival rates, that for those who are predicted to retire, retirement comes earlier than under the 1995 rules but later than under the current rules. There is a substantial shift toward later retirement. This illustrates the well understood proposition that DC plans eliminate the “pull” that makes teachers stay longer as they approach the peak of pension wealth, and the “push” that makes teachers retire early as they pass the peak of pension wealth under the DB plan). Hence the age distribution under a DC plan is less concentrated.

¹⁰The distribution of retirement age and/or experience depends on the forecasting horizon. In a shorter time frame, say 6 years, the retirement age distribution under the 1995 rules are more similar to that under the current rules than Figure 8 depicts.

The experience distributions of retirement also differ under the 1995 rules and the current rules. Figure 8 shows that under the 1995 rules retirement experience is more concentrated at 30 and there is no spike at 25 years of experience. Under the 1995 rules, retirement is more likely to occur later in a career. The primary reasons are absence of “25 and out” and the “rule of 80.” Under the DC plan the retirement experience distribution is flatter than under the current rules. Again, this confirms what we learned from the age distributions, that the DC plan eliminates the “pull” and “push” incentives.

Figure 9 shows the joint age-experience distributions of the retirement years under the three regimes. Under the current rules, there is a ridge reflecting the “rule 80”. Under the 1995 rules, the retirement age-experience distribution is more concentrated at age 60 or experience of 30 years, forming an “L.” The DC plan smooths out the retirement distribution making retirement less concentrated on arbitrary age and experience levels.

Using the structural estimates, we can examine effect of pension rules on teachers’ welfare. Let the expected utility of a teacher in 2002 under the current rules, the 1995 rules, and the DC rules be $U_{current}$, U_{1995} , U_{DC} . Suppose after we multiply the stream of salary and pension benefits under the current rules by a constant τ to make the utility under the current rules equal to that under the 1995 rules with the actual salary in 2002. The constant τ equals a ratio of expected utility to the power of $\frac{1}{\gamma}$: $\tau = (\frac{U_{1995}}{U_{current}})^{\frac{1}{\gamma}}$. Under the DC rules $\tau = (\frac{U_{DC}}{U_{current}})^{\frac{1}{\gamma}}$. A τ lower than unity means the teacher is better off under the current rules relative to alternative rules.

The relationship between relative welfare and experience is illustrated by Figure 10. The experienced teachers do best under the DC rules and worst under the 1995 rules. Figure 11 plots τ for subsamples of teachers who retired in different years. The chart shows that teachers who retired in 2003 would gain under the DC plan, because they would gain access to the same pension wealth with an added option of continuing to teach and letting the

pension wealth accumulate. For the 2003 retirees with long service experience, the option of teaching is particularly valuable. Those who are not retired in 2008, who tend to have short service experience in 2002, under the DC plan that starts with 2002 pension wealth would find their utility reduced. This is because under the current rules, the pension wealth is a nonlinear function of experience, accelerating as the teacher approaches the late stage of her career. Stopping the service experience clock in 2002 disproportionately reduced the pension wealth of the low experience teachers. This discussion suggests that, in order to convert from a DB plan to a DC plan without reducing the welfare of teachers and without raising the cost, there needs to be experience-dependent compensation. Figure 12 shows that the relationship between gender and welfare under alternative rules. Males do better under the DC rules and do worse under the 1995 rules. Figure 12 is based on the estimates for female and male samples (the third column of Table 2 for females and fifth column of Table 2 for males.) Using the estimates of the full sample (column 1 of Table 2) produces qualitatively similar plots.

Under the 1995 rules, the teachers are uniformly worse off than under the current rules, the τ ratio is 92 – 95% on average, depending the retirement year under the current rules. More experienced teachers suffer the largest losses under the 1995 rules. Under the DC rules, the welfare impact varies. Averaging over all teachers in the 2002 cohort, the mean of τ is 91% under the DC plan. The welfare comparison depends on the details of the rules. If we allow for more generous provision of the initial pension wealth under the DC plan, the predicted behavior pattern under the DC plan does not change but more teachers will have a τ ratio above unity.

6 Conclusion

Policy discussions about teacher quality and teacher “shortages” often focus on recruitment and retention of young teachers. However, attention has begun to focus on the incentive effects of teacher retirement benefit systems, particularly given their rising costs and their large unfunded liabilities. In this paper we estimate a structural model of teacher retirement for Missouri teachers and use it to estimate the effect of pension rules on the timing of retirement. The model fits the data very well, and nicely mimics the sharp spikes associated with certain age and experience combinations. We used the structural model to evaluate the effect of pension enhancements enacted during the 1990’s and a hypothetical DC pension plan. Enhancements during the 1990’s are found to have lowered the retirement age for teachers. A DC (or cash balance) alternative plan would greatly ameliorate the spikes and smooth out retirements. The 1990’s enhancements increased the welfare of all teachers, although the benefits vary with age and experience. DC plans have a mixed effect, improving welfare for some teachers and lowering it for others. Male teachers in our sample on average benefit more from a DC conversion than female teachers.

As states consider reform of teacher pension plans, behavioral econometric models can be of great value in estimating labor market and fiscal effects of reforms. Moreover, thanks to substantial financial support from the U.S. Department of Education, many states have built extensive longitudinal databases on educators and students that can support this type of research.

References

Asch, Beth, Steven J. Haider, and Julie Zissimopoulos. (2005). "Financial Incentives and Retirement: Evidence from Federal Civil Service Workers." *Journal of Public Economics*. 89 (2-3), 427-440.

Berkovec, James and Stern, Steven. (1991). "Job Exit Behavior of Older Men", *Econometrica*, 59 (1), 189-210.

Brown, Kristine. (2009). "The Link Between Pensions and Retirement Timing: Lessons from California Teachers." National Center on Performance Incentives. Vanderbilt University.

Coile, Courtney and Jonathan Gruber. (2007). "Future Social Security Entitlements and the Retirement Decision." *Review of Economics and Statistics*. 89(2), 234-246.

Costrell, Robert, and Joshua McGee. (2010). "Teacher Pension Incentives, Retirement Behavior, and Potential for Reform in Arkansas." *Education Finance and Policy*. 5(4), 492-518.

Costrell, Robert, and Michael Podgursky. (2009). "Teacher Retirement Benefits." *Education Next*. 9(2) (Spring), 58-63.

Friedberg, Leora and Sarah Turner. (2010). "Labor Market Effects of Pensions and Implications for Teachers." *Education Finance and Policy*. 5(4), 463-491.

Friedberg, Leora, and Anthony Webb. (2005). "Retirement and the Evolution of Pension Structure." *Journal of Human Resources*. 40(2), 281-308.

Furgeson, Joshua, Robert Strauss, and William Vogt. (2006). "The Effects of Defined Benefit Pension Incentives and Working Conditions on Teacher Retirement Decisions." *Education Finance and Policy* 1(3), 316-48.

Grissom, Jason A., Cory Koedel, Shawn Ni, and Michael Podgursky. (2011). "Pension-Induced Rigidities in the Labor Market for School Leaders." Working paper, University of Missouri.

Hanushek, Erik A. Kain, John F. and Rivkin, Steven G. (2004). "Why Public Schools Lose

Teachers.” *Journal of Human Resources*. 39(2), 326-354

Ippolito, Richard A. (1997). *Pension Plans and Employee Performance: Evidence, Analysis, and Policy*. Chicago: University of Chicago Press.

Lumsdaine, Robin. L, Stock, James H., and Wise, David A. (1992). “Pension Plan Provisions and Retirement: Men and Women, Medicare, and Models”, NBER Working Paper 4201.

Murnane, Richard J., and Olsen, Randall J. (1990). “The effects of salaries and opportunity costs on length of stay in teaching: Evidence from North Carolina.” *Journal of Human Resources*, 25(Winter), 106-124.

Ni, Shawn, Michael Podgursky, and Mark Ehlert (2009). “Teacher Pension Incentives and Labor Market Behavior: Evidence from Missouri Administrative Teacher Data.” Vanderbilt University: National Center on Performance Incentives.

Podgursky, Michael, Ryan Monroe, and Donald Watson. (2004). “Teacher Pay, Mobility, and Academic Quality.” *Economics of Education Review*. 23, 507-518.

Poterba, James, Joshua Rauh, Steven Venti, and David Wise. (2007). “Defined Contribution Plans, Defined Benefit Plans, and the Accumulation of Retirement Wealth.” *Journal of Public Economics*. 91, 2062-2086.

Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. (2005). “Teachers, Schools, and Academic Achievement.” *Econometrica*. 73(2), 417-458.

Stinebrickner, Todd. (2001). “A dynamic model of teacher labor supply.” *Journal of Labor Economics*, 19(1) (January), 196-230.

Stock, James and David Wise. (1990). “Pensions, the Option Value of Work, and Retirement.” *Econometrica*. 58(5) (September), 1151-1180.

Stern, Steven. (1997). “Approximate Solutions to Stochastic Dynamic Programs.” *Econometric Theory*. 13, 392-405.

Table 1: Sample Averages by the Year of Retirement

Retirement Year	number of teachers	age	experience	male
All 2002	9605	56.04	24.15	0.20
2003	640	53.09	28.53	0.29
2004	862	53.58	28.60	0.26
2005	972	54.42	28.01	0.24
2006	822	55.39	27.73	0.21
2007	778	56.39	26.86	0.20
2008	719	57.44	26.37	0.19
Not Retired	4812	57.04	20.61	0.17

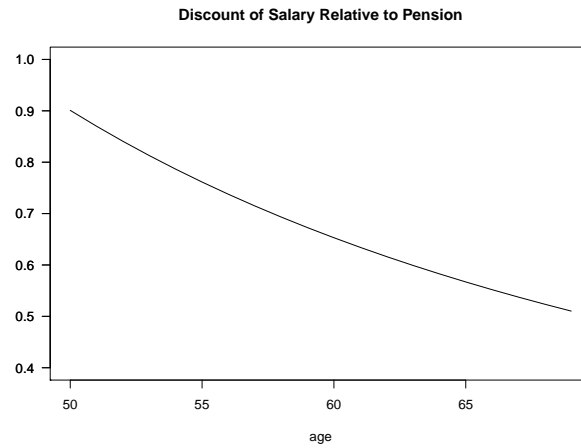
Note: The sample of Missouri PSRS teachers are age 50-55 in 2002. “All 2002” denotes the total cohort of 9605 in the base year. The rows with year labels are averages over all teachers who retired in that year. “Not Retired” are teachers who remained employed in 2008 at the end of the panel. The average age and experience are at the year of separation. Male=1 for male teachers.

Table 2: MLE Estimates of Structural Parameters

Parameters	Whole Sample		Female		Male	
	Full Model	Constant κ	Full Model	Constant κ	Full Model	Constant κ
κ	0.653 (0.009)	0.675 (0.009)	0.654 (0.012)	0.686 (0.013)	0.711 (0.028)	0.717 (0.029)
κ_1	1.765 (0.086)	0	1.681 (0.069)	0	3.316 (0.456)	0
β	0.959 (0.028)	0.967 (0.031)	0.959 (0.035)	0.967 (0.042)	0.957 (0.077)	0.965 (0.077)
γ	0.670 (0.034)	0.668 (0.040)	0.669 (0.045)	0.672 (0.057)	0.670 (0.111)	0.674 (0.070)
σ	2799.201 (136.972)	2729.768 (175.244)	2760.441 (214.142)	2716.468 (293.034)	3765.290 (811.277)	3590.589 (415.888)
ρ	0.509 (0.041)	0.573 (0.042)	0.512 (0.033)	0.628 (0.087)	0.478 (0.09)	0.509 (0.070)
log-likelihood	-12833.539	-12939.789	-10050.364	-10133.786	-2776.343	-2802.565

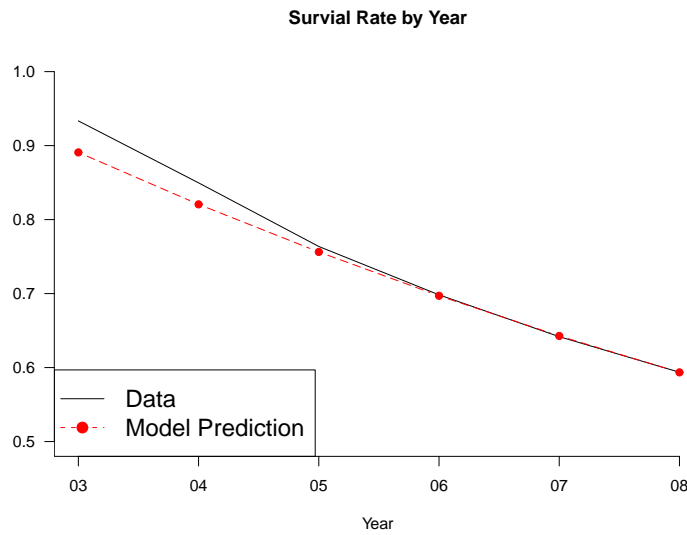
Note: Standard errors in parentheses.

Figure 1: Value of Salary relative to Pension and Survival Rate from Retirement



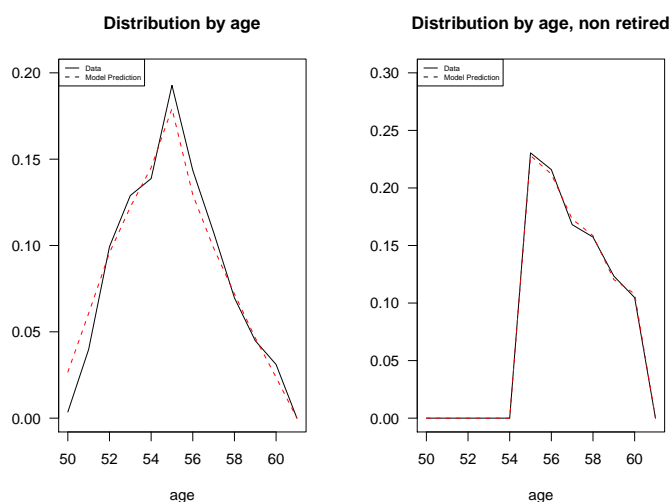
Note for Figure 1: The figure plots the age-dependency of disutility of work, $k_s = \kappa(\frac{60}{age})^{\kappa_1}$, based on estimates of κ and κ_1 in the first column of Table 2.

Figure 2: Survival Rate: Estimated and Observed



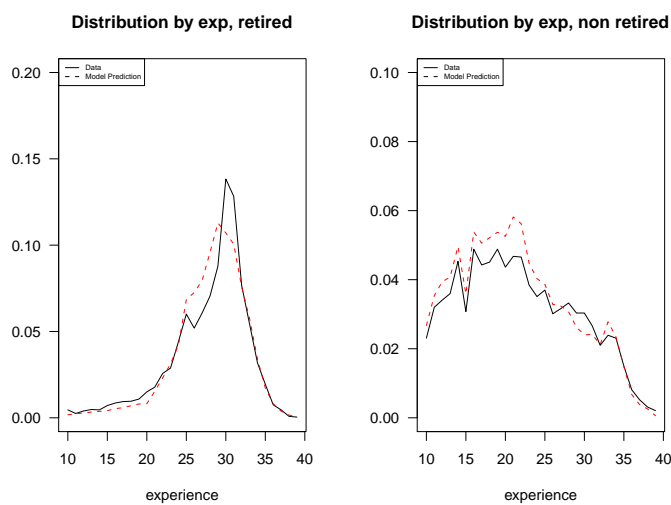
Note for Figure 2: The observed survival rate pertains to all teachers in the sample year. The simulated survival rate are based on the estimates in the first column of Table 2.

Figure 3: Observed and Simulated Age Distributions of Retired and Non-retired Teachers



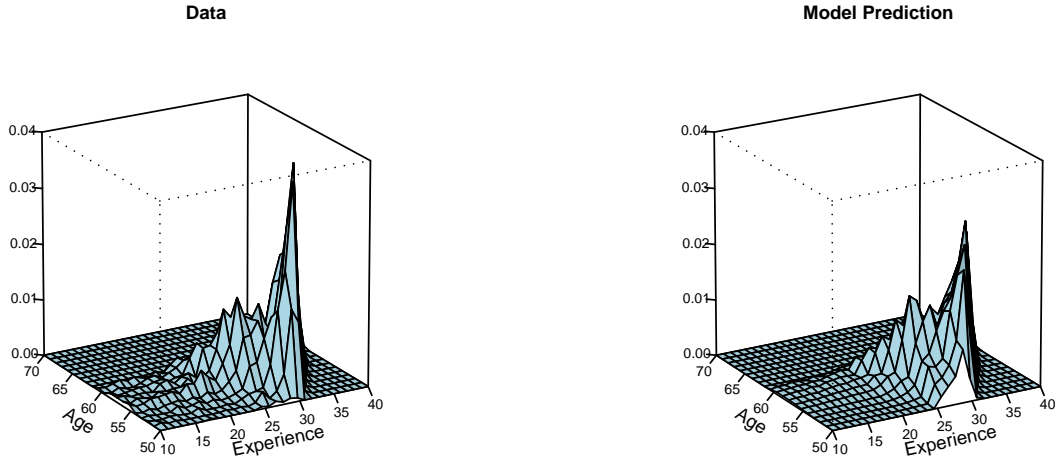
Note for Figure 3: The observed age pertains to all teachers at the year of retirement (for the left panel) or the non-retired at the end of the sample period (for the right panel). The simulation is based on the estimates in first column of Table 2.

Figure 4: Observed and Simulated Experience Distributions of Retired Teachers



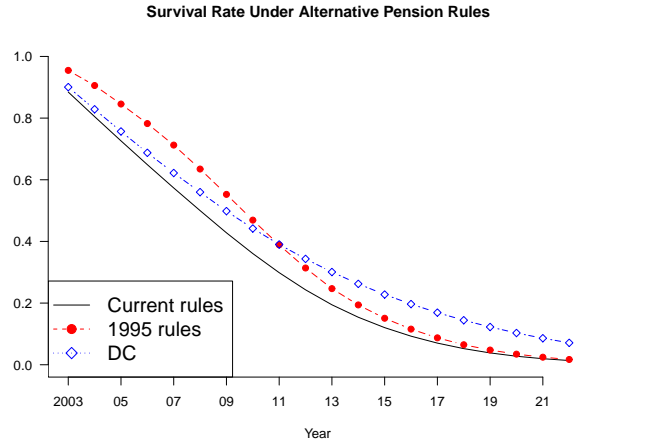
Note for Figure 4: The observed experience pertains to all teachers at the year of retirement (for the left panel) or the non-retired at the end of the sample period (for the right panel). The simulation is based on the estimates in the first column of Table 2.

Figure 5: Observed and Simulated Joint Retirement Age-Experience Distribution for Retired Teachers



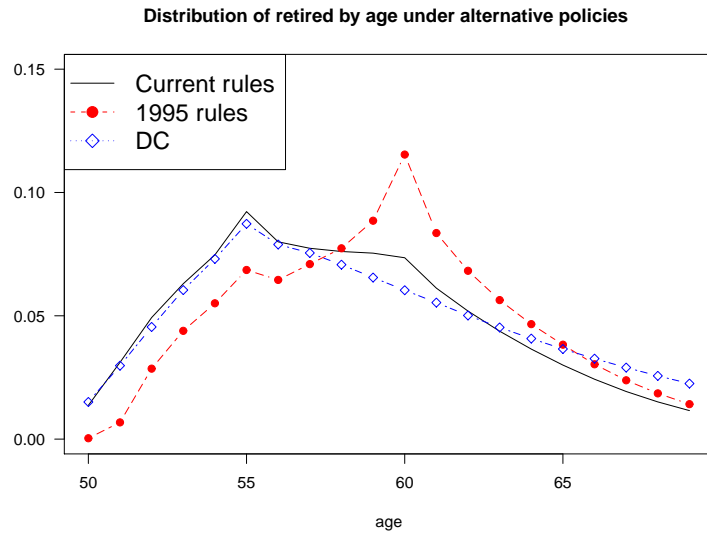
Note for Figure 5: The simulation is based on the estimates in the first column of Table 2.

Figure 6: Survival Rate Under Alternative Policies



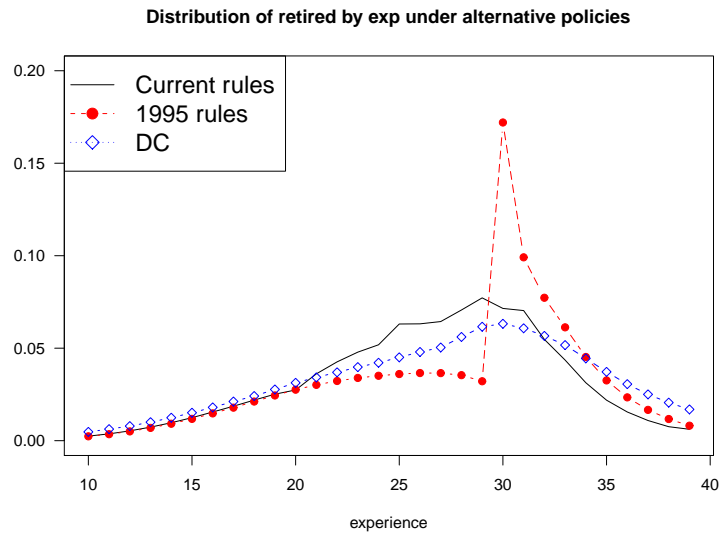
Note for Figure 6: The simulated survival rates are based on the estimates in the first column of Table 2, under alternative pension rules.

Figure 7: Retirement Age Distribution Under Alternative Policies



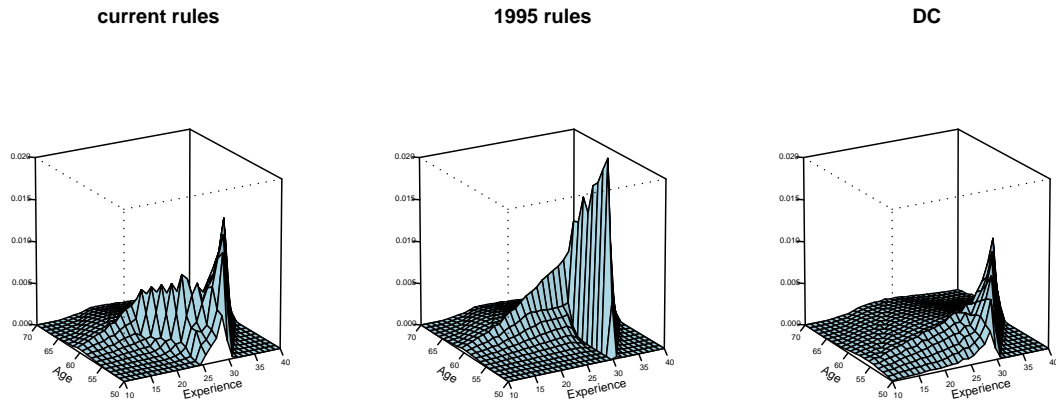
Note for Figure 7: The simulation is based on the estimates in the first column of Table 2.

Figure 8: Retirement Experience Distribution Under Alternative Policies



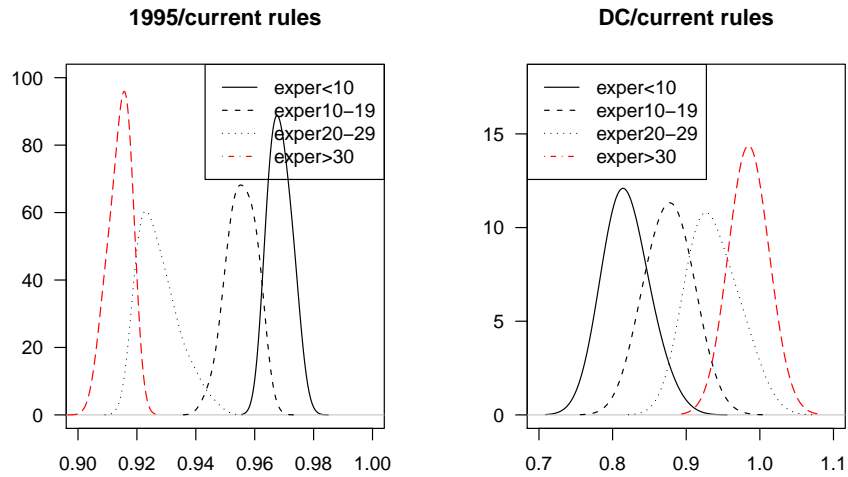
Note for Figure 8: The simulation is based on the estimates in the first column of Table 2.

Figure 9: Retirement Age-Experience Joint Distribution Under Alternative Policies



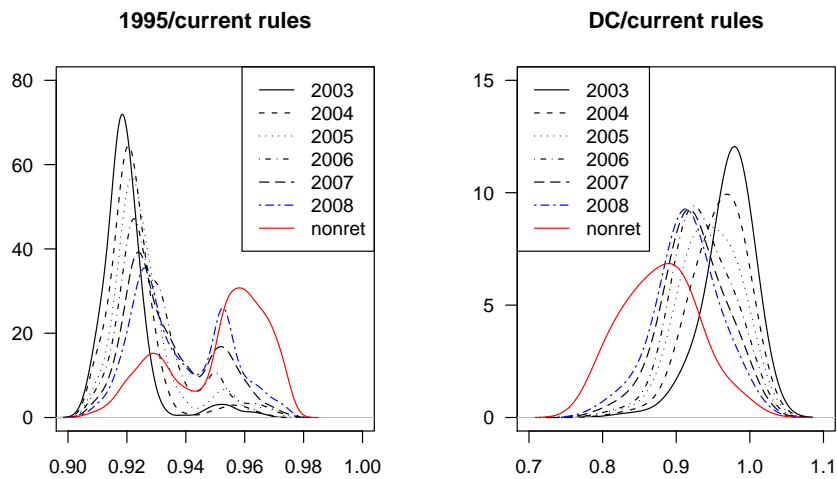
Note for Figure 9: The simulation is based on the estimates in the first column of Table 2.

Figure 10: Salary-Equivalent Utility Ratio Relative to the Current Rules by Experience



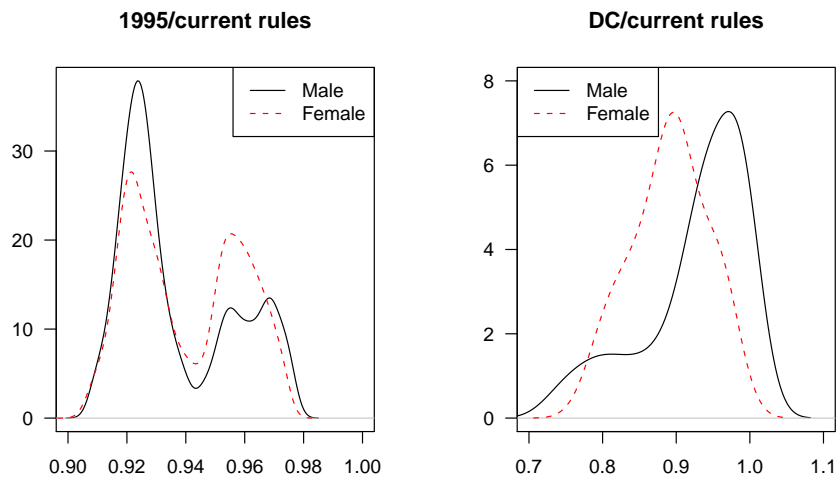
Note for Figure 10: The simulation is based on the estimates in the first column of Table 2. A ratio less than unity means the teacher is better off under the current rules relative to the alternative rules.

Figure 11: Salary-Equivalent Utility Ratio Relative to the Current Rules by Retirement Year



Note for Figure 11: The simulation is based on the estimates in the first column of Table 2. A ratio lower than unity means the teacher is better off under the current rules relative to the alternative rules.

Figure 12: Salary-Equivalent Utility Ratio Relative to the Current Rules by Teacher Gender



Note for Figure 12: The simulation is based on the estimates in the third column of Table 2 (for females) and fifth column of Table 2 (for males). A ratio lower than unity means the teacher is better off under the current rules relative to the alternative rules.