

# Teacher Pension Enhancements and Staffing in an Urban School District

Shawn Ni

University of Missouri

Michael  
Podgursky

Saint Louis University  
and University of  
Missouri

Xiqan  
Wang

Beijing University  
of Technology

## Abstract

Many states enhanced benefits in teacher retirement plans during the 1990s. This paper examines the school staffing effects of one such enhancement in a major urban school district with mostly high poverty schools. Pension rule changes in 1999 for St. Louis public school teachers resulted in large increases in pension wealth for active teachers, as well as a powerful increase in “push” incentives for earlier retirement. Simple descriptive statistics on retirement patterns before and after the enhancements suggest much earlier retirement resulted. Shorter teaching spells imply a steady state with more teacher turnover and a larger share of novice teachers in classrooms. To better understand the long run effects of these changes and alternatives policies, the authors estimate a structural model of teacher retirement. Simulations of retirement behavior for representative senior teachers point to shorter completed teaching spells and earlier retirement age as a result of the enhancements. By contrast, moving from the post-1999 to a DC-type plan would extend the teaching career of a representative senior teacher by roughly two years.

JEL codes: J32, J26, H72

Key words: Teacher pensions, teacher compensation, state and local pension finance

Working Paper 21-01  
May, 2021

The researchers acknowledge support from the Charles Koch Foundation. Analysis and options are those of the authors.

# 1 Introduction

Most public school teachers are covered by traditional, or more precisely, final average salary defined-benefit (FAS-DB) pension plans that base a teacher's retirement annuity on the average of the highest (typically final) several years of earnings. These types of plans are commonplace for state and local government employees, but DB plans of any sort have largely disappeared in the private sector, in favor of defined contribution (DC) plans such as 401k, IRA or 403b plans (Butrica, Iams, Smith, and Toder, 2009). The costs of teacher pension plans have risen sharply over time and account for a growing share of per pupil expenditures. In current dollars, employer teacher pension costs (excluding Social Security and teacher contributions) amounted to \$547 per student or 4.8 percent of current per student expenditures in March, 2004. By June, 2020, these costs had risen to \$1510 per student or 11.1 percent of per student expenditures<sup>1</sup>.

The fiscal stress for schools and districts resulting from these rising pension costs has been widely noted (e.g., Moody and Randazzo (2020)). There are a variety of factors that have contributed to these rising costs, most notably a failure of state or local governments to make actuarially necessary contributions and less than forecast returns on plan investments. However, one often overlooked factor is the cost of pension enhancements that occurred during the 1990s. In response to the stock market boom during the 1990's, many pension plans enhanced benefits, and educator pensions were among the most actively enhanced. Between 1999 and 2001 alone, for example, the National Conference of State Legislators reported that educator pensions were enhanced in more than half the states (Koedel, Ni, and Podgursky, 2014). Teachers retiring from public schools today and for many years to come are doing so under these more remunerative formulas.

Our focus in this paper is on one consequence of these enhancements – the effect on the labor supply of senior teachers. In particular, we examine the effect of a substantial pension

---

<sup>1</sup><http://www.uaedreform.org/downloads/2017/12/employer-contributions-per-pupil-12-20-17.pdf>

enhancement on teacher labor supply and school staffing in St. Louis, an urban district that has difficulty attracting and retaining qualified teachers. In 1999, the pension formula for St. Louis teachers was changed substantially, with a large increase in the replacement rate accompanied by a cap on total benefits as percentage of final salary. These major changes in the pension rules provide an opportunity to study how pension rules influence teacher retirement decisions and school staffing.

A well-known finding is that teachers retire earlier than other comparable professionals (Harris and Adams, 2007). A growing literature connects teacher retirement decisions to DB plan incentives. Studies of senior teachers consistently show a high degree of responsiveness in retirement timing to pension system incentives (Brown, 2013; Costrell and McGee, 2010; Furgeson, Strauss, and Vogt, 2006; Knapp, Brown, Hosek, Mattock, and Asch, 2016; Ni and Podgursky, 2016). In general, pension enhancements, by lowering the price of leisure (retirement) relative to work, tend to encourage more of the former. Given popular concern with shortages of experienced teachers, especially in high need schools, the consequences of expensive compensation decisions that have the effect of reducing teacher labor supply deserve careful scrutiny. In this paper we highlight the strong “push” and “pull” incentives in the St. Louis plan and show descriptively how changes in pension wealth accrual are associated with changes in retirement behavior in pre- and post-panels.

To explore the longer term effects and investigate alternative policies, we estimate a structural model of retirement behavior, with parameters describing underlying teacher preferences that are invariant to pension rules. There are important advantages to estimating a structural model. First, it allows us to isolate and unbundle the effects of more complex changes in plan rules. Second, it allows us to predict the effect of the rule changes in retirement beyond the sample period, since it takes many years for the full effects of pension rule changes on teacher retirement decisions to materialize. Finally, and perhaps most importantly, structural estimates allow us to examine the effect of alternative plans. In the

context of St. Louis, it allows us to explore the labor supply and fiscal effects of partial or total conversion from the current FAS-DB to defined contribution (DC) alternatives.

This is not the first study of the labor supply effects of the Saint Louis teacher pension enhancements. Koedel and Xiang (2017) examine the teacher retention effects of the 1999 enhancements in a difference-in-difference study before and after the enhancements occurred. They found no significant effects on retention for younger teachers who were not retirement eligible. Retirement eligible teachers, aware of forthcoming enhancements, postponed retirement until the enhancements were in place. Our focus is on how the enhancements affected the long run timing of retirement, an issue not explored in their paper.

Specifically, this paper makes several contributions to the research literature on public employee and teacher pensions. First, it provides a descriptive analysis of a major shift in retirement behavior after an enhancement of teacher pensions in a high poverty urban district. As noted, these types of enhancements were commonplace during the 1990's, but little analysis has been undertaken regarding their staffing effects in high poverty schools. Second, it shows that this structural model can fit more challenging data than it has to date in the existing retirement literature on teachers (e.g., Knapp et al. (2016); Ni and Podgursky (2016); Ni, Podgursky, and Wang (2020)). We show that these structural estimates of teacher preferences are similar across different groups of teachers in retirement plans with different rules – including Social Security coverage. Moreover, the sharp changes in retirement behavior of St. Louis teachers after the 1999 enhancements provide a useful empirical challenge, i.e., to fit very different retirement data with one set of parameters. We find that the behavioral parameters of the structural model estimated on the *post-enhancement* sample accurately fit *pre-enhancement* data and behavior. The good out-of-sample fit under very different pension rules provides external validation for the structural model and highlights the importance of pension rules on retirement behavior. In addition, we use the structural estimates to unbundle the separate effects of a complex enhancement to analyze their effects on female and male

teachers at different levels of age and service years. Finally, we use the structural estimates to simulate the effects of mandatory and voluntary DC conversion plans for teachers.

In the next section we describe the operation of the St. Louis pension plan and the manner in which pension wealth accrues over a teacher’s work life, with the resulting “pull” and “push” incentives. We show how the 1999 enhancements dramatically increased the “push” incentives for retirement and report descriptive data for retirement age and service years of teachers before and after the enhancements. We then specify and estimate our structural retirement model. The model is found to exhibit good fit within and out-of-sample. Using the structural model estimates we then simulate the long run effect of the enhancements on a representative senior teachers. These simulations imply a significant reduction in the length of teaching careers as a result of the enhancements. We then consider some DC alternatives, including voluntary conversions. In a concluding section we discuss ways in which the research methods and findings for this particular district can be extended to other urban school districts.

## **2 Pension Wealth Accrual Under Plan Changes**

Teachers in Missouri are in three separate FAS-DB systems. Teachers in the St. Louis and Kansas City districts are covered by Social Security and are in two separate district plans. The rest of the public school teachers in the state are in the Public School Retirement System (PSRS), and are not covered by the Social Security system. There is no reciprocity between the St. Louis and other plans, meaning that the years of service in the formula below apply only to service in the Saint Louis plan.

Table 1 summarizes the St. Louis plan rules before and after 1999. Benefits at retirement

are determined by a standard FAS-DB formula:

$$\text{Annual Benefit} = rf \times S \times FAS$$

where  $rf$  stands for replacement factor.  $S$  denotes service years in the system.  $FAS$  is the average of the highest consecutive three years of salary, which are typically the three years prior to retirement. There is also a provision for early retirement at age 60 with a reduced annuity for teachers who do not qualify for regular retirement. St. Louis teachers become eligible for a regular pension when they reach age 65 with at least five years of service or when the sum of their years of service and age equals at least 85 (known as “Rule of 85”).

During the 1990’s there were some small pension rule changes. However, the big change occurred in 1999. Prior to 1999, the annual replacement factor was 1.25 percent. After 1999, the replacement factor increased to 2 percent retroactively. For example, before 1999, a retired teacher with 30 years of service and a final average salary of \$50,000 received an \$18,750 annuity ( $.0125 \times 30 \times \$50,000$ ). After 1999, this same teacher would receive an annual annuity of \$30,000 ( $.02 \times 30 \times \$50,000$ ). This is a very large enhancement awarded to all active teachers, including those who were about to retire.

While the replacement rate increased sharply, the new benefit was capped at 60 percent of final average salary. Thus a teacher with 31 or more years of service can still only receive 60 percent of her final average salary, in spite of the fact that she continues to contribute five percent of salary to the plan. Table 1 summarizes the pension rules before and after enhancement.<sup>2</sup>

(Table 1)

---

<sup>2</sup>There were some subsequent rule changes regarding pension COLAs. Before 2000, “COLA adjustments were equal to any increase of 1% or greater in the CPI, subject to an annual maximum of 3% and a cumulative maximum increase of 10%”. Effective July 1, 2000, retired members received a catch-up COLA of at least 65% of CPI. After 2002, COLAs are paid to retired members when such COLAs are approved by both the Board of Education and the Board of Trustees.

One way to assess the retirement incentives arising from these pension plan rules is to compute pension wealth at different points in a teacher’s work life. Pension wealth is calculated as the expected present value of pension benefits at retirement age  $r$  as follows:

$$PW(r) = \sum_{s \geq r} (1 + b)^{r-s} \pi(s|r) P(s|r),$$

where  $b > 0$  is the annual discount rate,  $\pi(s|r)$  is the conditional probability of survival given by age-dependent mortality rates, and  $P(s|r)$  is the pension benefit if the teacher retires at age  $r$ . Pension wealth captures the market value of obtaining this stream of annuity payments over time.

Following Costrell and Podgursky (2009) and others, pension wealth accumulation can be calculated as follows:

$$pw(r) = \frac{PW(r) - (1 + inf) \times PW(r - 1)}{Salary(r)},$$

where  $PW(r)$  is pension wealth if retiring at age  $r$ ,  $inf$  is inflation rate,  $Salary(r)$  is salary at age  $r$ .

Figure 1 plots total pension wealth before and after pension enhancement for a representative teacher in St. Louis who enters the system at age 25 and works continuously until separation. Figure 2 plots annual pension wealth accrual before and after pension enhancement.

(*Figures 1-2*)

Figures 1 and 2 are consistent with the prior literature analyzing the “pull” and “push” incentives implied by different pension rules, e.g., Costrell and Podgursky (2009). The steeply rising accrual up to peak value tends to lock in teachers, while the decline in pension wealth past the peak tends to push teachers in to retirement. What is notable in both figures is how

much more powerful the “push” incentives became after 1999. This is a direct result of the 60 percent of FAS annuity cap. Pension wealth accrued faster because of the sharp increase in the formula factor from 1.25 percent to 2.0 percent, and then declines sharply because of the 60 percent cap. The result is much more powerful incentives pushing our representative teacher into retirement around age 55.

### **3 Retirement Behavior Before and After Pension Enhancements**

In this section we examine empirical retirement behavior before and after the 1999 pension enhancements. We use administrative data for St. Louis public school teachers and track two cohorts of teachers aged 50-62 at the beginning of the period forward for five years. The pre-enhancement cohort begins in Fall, 1992 and is followed to Fall, 1996, well before enactment or discussion of the 1999 changes. The post-enhancement cohort begins in Fall, 2008 and is tracked to Fall 2012. Again, this is well beyond the initial implementation of the enhancement. We have a cohort of 696 female and 184 male teachers in the 1993 cohort. By the end of the panel 186 females (27%) and 69 males (38%) had retired. The 2009 cohort had 692 female teachers and 198 male teachers of whom 396 female teachers (57%) and 116 male teachers (59%) had retired by the end of the panel. Table 2 and Table 3 report the basic descriptive statistics in the base year and subsequent five years for the two cohorts.

*(Tables 2-3)*

Table 3 suggests that teacher retirement behavior differs substantially under pre- and post-enhancements rules in two important ways. The first difference is that in the post enhancement period teachers retire earlier than in the pre-enhancement period. This is



seen clearly in Figure 3 which reports the employment survival rates over five years. The post enhancement survival curve is well below the earlier cohort for both female and male teachers. Figures 4 and 5 report age and service years distributions of retired teachers in the two cohorts. Here it is important to note we only observe completed age and experience spells for teachers who have retired. Nonetheless, among the retirees, the cumulative age distribution is clearly displaced left for females. The results for males are somewhat more ambiguous. For both groups, however, there is a pronounced leftward displacement in the distribution of service years.

*(Figures 3-5)*

It is also noteworthy that the average baseline service years of the senior teachers in the post-enhancement sample period is considerably lower than in the pre-enhancement sample, despite the fact that the average age is nearly identical. In 2009, at the beginning of the post-enhancement sample the female teachers are about 2.8 years less experienced than their same age-group counterpart in 1993, and male teachers in 2009 are about 4.8 years less experienced. This substantial difference in service years suggests in the post-enhancement sample teachers older than 50 tended to retire earlier (with fewer service years) due to the change in pension incentives. There are two implications of this observation. First, conditional on age, one would expect less experienced senior teachers to retire later, hence the earlier retirement by the post-enhancement cohort shown in Figures 4 and 5 is even more striking. Second, some teachers of age 50-62 are retirement-eligible prior to the beginning of the sample period. Their decision to stay on the job reflects preferences unobserved to the researchers. The distributions of these preferences are likely differ at the beginning of the pre- and post-enhancement periods because the pension rules themselves have reshaped the pool of senior teachers. This illustrates why it is important to take account of baseline sample selection in estimating a statistical retirement model. We will take up this issue of baseline sample selection when we discuss structural model estimation in the next section.

Before turning to modeling, we round out our empirical discussion by presenting simple regression-adjusted retirement rates for the two cohorts. Table 4 reports the results of fitting a simple linear probability model that predicts probability of retirement as a function of baseline age, service years, and whether a teacher is in the post-enhancement cohort. The coefficient on the post-enhancement cohort dummy implies that the probability of retirement, averaged over all age and service year cells was 37 percentage points higher for females, and 30 percentage points higher for males over the five year period.

*(Table 4)*

The evidence in Tables 3 and 4 and Figures 3-5 suggests that the 1999 enhancements significantly shortened teaching careers and that the “push” effects were very potent. To explore further the long-run effect of these changes on senior teachers, and consider how teachers would respond to alternate retirement plans, in the next section we estimate a structural model of retirement behavior and use it to examine the long run effect of these rule changes on representative senior teachers.

## **4 Estimating a Dynamic Retirement Model**

Stock and Wise (1990) developed a structural model of retirement which was extended by Ni and Podgursky (2016) to an analysis of public school teachers. We apply this model, with several innovations, to an analysis of St. Louis teachers.

Consider an employed teacher who has not retired at the beginning of year  $t$ . The present

value of expected lifetime utility for the teacher if she retires in year  $r$  ( $r \geq t$ ) is:

$$\mathbb{E}_t V_t(r) = \max_{c_1, c_2} \mathbb{E}_t \left\{ \sum_{s=t}^{r-1} \beta^{s-t} \pi(s|t) U_w(Y_s, B_s(c_1)) + \sum_{s=r}^T \beta^{s-t} \pi(s|t) U_r(P_s(c_2), B_s(c_1)) \right\} \quad (4.1)$$

where

$$U_w(Y_s, B_s(c_1)) = (k_s((1-c)Y_s + I_s^1 B_s(c_1)))^\gamma + \omega_s, \quad (4.2)$$

$$U_r(P_s(c_2), B_s(c_1)) = (I_s^2 P_s(c_2) + I_s^1 B_s(c_1))^\gamma + \xi_s. \quad (4.3)$$

$$I_s^1 = \begin{cases} 1 & \text{if } s \geq c_1 \\ 0 & \text{otherwise} \end{cases} \quad (4.4)$$

$$I_s^2 = \begin{cases} 1 & \text{if } s \geq c_2 \\ 0 & \text{otherwise} \end{cases} \quad (4.5)$$

The value function  $V$  depends on future annual earnings  $Y_s$  before retirement, pension benefits  $P_s(c_2)$  if the teacher starts collecting a pension benefit in year  $c_2$ , and Social Security benefits  $B_s(c_1)$  if she starts collecting Social Security in year  $c_1$ . The collection years  $c_1$  and  $c_2$  are optimally chosen to maximize the expected utility given the retirement year  $r$ . The parameter  $c$  is the contribution rate for teachers before retirement, and  $k_s$  represents the teacher's preference for teaching versus retirement, which, in turn, depends on the teacher's age in year  $s$ . We assume  $k_s = k(\frac{60}{age})^{k_1}$ . The errors  $\omega_s$  and  $\xi_s$  measure preferences for teaching and retiring respectively. The conditional survival rate  $\pi(s|t)$  is based on U.S. mortality data.

The expected gain from postponing retirement to year  $r$  is:

$$G_t(r) = \mathbb{E}_t V_t(r) - \mathbb{E}_t V_t(t). \quad (4.6)$$

Each year, an employed teacher has two choices: continue teaching or retire. The expected gain from postponing retirement can be seen as the “option value” of continuing to work, which is a key feature of this model. Retirement occurs if the value of continuing teaching is less than the value of retiring, i.e., the option value of continued teaching is negative.

The teacher’s future salary, pension benefits, and Social Security benefits are assumed to be predictable<sup>3</sup>. We know the salary of teachers in the first period and then we predict the future salary with forecasts based on historical data.

The expected gain from retiring in year  $r$  (later than  $t$ ) is

$$G_t(r) = \mathbb{E}_t V_t(r) - \mathbb{E}_t V_t(t) = g_t(r) + K_t(r)\nu_t, \quad (4.7)$$

where the first term,  $g_t(r) = \sum_{s=t}^{r-1} \beta^{s-t} \pi(s|t) ((k_s((1-c)Y_s + I_s^1 B_s(c_1)))^\gamma) +$

$\sum_{s=r}^T \beta^{s-t} \pi(s|t) ((I_s^2 P_s(c_2) + I_s^1 B_s(c_1))^\gamma - [\sum_{s=t}^T \beta^{s-t} \pi(s|t) (I_s^2 P_s(c_2) + I_s^1 B_s(c_1))^\gamma])$ , depends on pension rules. The second term,  $K_t(r) = \sum_{s=t}^{r-1} (\beta\rho)^{s-t} \pi(s|t)$ , depends on the mortality rates. Both terms depend on the parameters we need to estimate. The preference error  $\nu_t = \omega_t - \xi_t$ . We assume  $\nu_t$  follows an AR(1) process:  $\nu_t = \rho\nu_{t-1} + \epsilon_t$ , where  $\epsilon_t$  is iid  $N(0, \sigma^2)$ .

Suppose  $r^\dagger$  solves  $\max_{r \in \{t+1, t+2, \dots, T\}} \mathbb{E}_t V_t(r)$ . Thus, the teacher will continue working at  $t$  if  $G_t(r^\dagger) = \mathbb{E}_t V_t(r^\dagger) - \mathbb{E}_t V_t(t) > 0$ . The probability of retirement at time  $t$  is  $P[R = t] = P[G_t(r) \leq 0, \forall r \geq t + 1] = P[G_t(r^\dagger) \leq 0]$ , which can be rewritten as

$$P[-g_t(r^\dagger)/K_t(r^\dagger) \geq \nu_t]. \quad (4.8)$$

---

<sup>3</sup>Teachers’ salaries are largely determined by the district’s salary schedule and can be predicted using the following formula:  $Salary_{t+i} = Salary_t * \exp(b_1 * i + b_2 * [(t+i)^2 - t^2] + b_3 * [(t+i)^3 - t^3])$ , where  $b_1=0.03569$ ,  $b_2=-0.00053652$ ,  $b_3=-0.00000372$ .

In this model there are six unknown structural parameters to be estimated, which are listed in Table 5. The threshold  $g_t(r_t^\dagger)/K_t(r_t^\dagger)$  depends on teacher’s age, service years, and the six structural parameters. Details concerning the computation of the retirement probabilities are given in the Appendix A.1.

(Table 5)

The structural model assumes that retirement decisions by teachers are governed by the six parameters. Teachers with the same age and service years differ in timing of retirement due to differences in preference errors  $\nu_t$  that measure the unobserved willingness to teach (relative to retirement). The preference errors are partly endogenous because they depend on teachers’ past retirement decisions. Condition (4.8) implies that the preference errors  $\nu_t$  for teachers who choose to continue teaching are greater than  $-g_t(r_t^\dagger)/K_t(r_t^\dagger)$ . Teachers with above average age and service years and who repeatedly choose to remain working imply higher values of  $\nu_t$ . In the model estimation we factor in the dependence of the thresholds on preference errors by computing the sequential probabilities of the teachers’ decisions.

The dependence of the preference errors pertains to the initial sample as well as the panel and results in the baseline sample selection problem noted in the previous section. The estimation data set includes teachers who were eligible for retirement before 2008 (the beginning of sample period) but who chose to wait, but it excludes those who chose to retire prior to 2008. Failing to take this sample selection into account would result in biased estimates. Hence the sample selection in this setting is an example of the familiar “initial condition problem” in dynamic panel data models (e.g., Heckman (1981); Wooldridge (2005).)

We solve the “missing early leavers” problem by estimating the model conditioning on the probability that retirement-eligible teachers appear in the initial sample. The option-value model depicts how the sample selection depends on the preference errors prior to the initial sample period.

The likelihood of the sample can then be computed by integrating out these baseline preference errors along with the preference errors in the sample period. Our approach to correcting this sample selection problem is discussed in greater detail in Appendix A.2.

The data we use for estimation includes all teachers aged 50-62 in the 2008-09 school year, which is the cohort of teachers under the post-enhancement pension rules described in Table 3. We track this cohort of teachers for the next five years. Using Maximum Likelihood estimation (MLE) methods described in the Appendix, we obtain the estimates of six structural parameters reported in Table 6.

*(Table 6)*

All of these parameter estimates are statistically significant, plausible, and estimated with reasonable to high precision. Compared with the estimates in Ni and Podgursky (2016), these “deep” parameters, i.e., capturing basic preferences, are quite similar even though the St. Louis teachers are in a separate and substantially different FAS-DB pension system. In addition, St. Louis teachers are covered by Social Security, whereas teachers in the state-wide plan are not.

The parameter  $\beta$  measures the teacher’s discount rate. The parameter  $\gamma$  is significantly less than unity for both female and males, which implies risk-aversion.  $\sigma$  and  $\rho$  are two parameters related to unobserved retirement preferences. These unobserved preferences (“preference shocks”) are substantial and persistent (i.e.,  $\rho$  is positive and significant). Since the option value is the key variable in this model that predicts retirement behavior, the large standard deviation estimates of 4001.108 and 3770.432, for females and males respectively, are reasonable since they capture all other omitted variables that influence retirement decisions. The parameters  $k$  and  $k_1$  together measure the disutility of teaching relative to retirement. These estimates imply that the attractiveness of retirement (leisure) relative to

teaching rises with age (Ni and Podgursky, 2016).

The structural model provides a good fit to these data both in and out of sample. Figure 6 reports employment survival rates in-sample for the 2009 senior teacher cohort and out-of-sample for the 1993 cohort. Note that the 1993 out-of-sample cohort is a non-overlapping group of teachers who face a very different set of pension rules. Nonetheless, the fit is good for both groups. More extensive plots of age and service years are provided in the Appendix B. Overall, the structural model exhibits good fit, both in and out-of-sample.

*(Figure 6)*

## 5 Simulating the Effects of Pension Plan Changes

As noted in the introduction, the 1999 enhancement of the replacement factor was accompanied by a cap on pension benefits of 60% of final average salary. Using the estimated structural model we conduct counter-factual simulations on the effect of each aspect of the pension rule changes.

In order to remove the influence of the composition of teachers in the post-enhancement sample and highlight the policy effect, we choose a representative female and male teacher who is 50 years old and with 25 years of service. Such a teacher is entering the retirement window. We analyze the effects of the pension enhancements and three counterfactual pension alternatives over a 30 year time horizon, by which time this teacher will have retired.

Table 7 reports the effects of the different pension rules on the timing of retirement. We treat the after-enhancement pension rules as our baseline case (Column 2) and subtract enhancement effects. Column 5 reports the effect of eliminating all of the enhancements. The elimination of pension cap of 60%, a reduction of 0.75% in the replacement factor, and a COLA cap leads to a higher level of average retirement age and extends the teaching career

by 2.6 years for females and 1.9 years for males. This shows that the less remunerative and uncapped pre-enhancement pensions exerted a weaker “push” effect, and hence yielded longer teaching careers.

(Table 7)

Table 7 also decomposes the effects of the pension enhancements into the components due to increasing the replacement factor and the effect of a benefit cap. Eliminating the cap on pension benefits increases the average service years at retirement by 0.7 years for female teachers and 0.5 years for male teachers. Imposing the COLA cap increases service years of retirement from 0.3-0.6 years. Reducing the replacement factor from 2% to 1.25% increases the average service years of retirement by 1.3 years for female and 1.1 years for male teachers.

In short, the somewhat surprising finding from Table 7 is that while the 60% cap, which has a clear “push” effect, shortened teaching careers, the larger effect came from the increase in the replacement factor from 1.25% to 2%. The COLA cap also had a potent effect.

Over the last several decades many private sector employers have frozen or closed their DB plans and put workers into defined contribution (DC) plans (Butrica et al., 2009). Several papers in the retirement literature have used reduced form and structural estimates of retirement behavior of employees in DB plans to simulate what the effect of retirement behavior under an alternative DC plan (e.g., Friedberg and Webb (2005), Costrell and McGee (2010), Ni and Podgursky (2016), Ni et al. (2020)). In part this is necessitated by the fact that an involuntary conversion of state and local employees (including teachers) to a DC is illegal under state constitutions (Monahan (2010)). Thus, the only such conversions to date have been for newly hired public sector employees. whose retirement behavior will not be observed for many years to come. A DC plan, while clearly a major change from the DB plan on which we base our estimates, may not be too far “out of sample” to justify simulating its workforce effects. In the context of our model, the major difference between DB and DC



is that the latter removes all “push” incentives.<sup>4</sup> In this respect it mirrors in reverse the powerful push incentives introduced in 1999 – i.e., an increase in the replacement rate from 1.25 % to 2.0 % combined with a 60 % benefit cap.

With these caveats in mind, the final column of Table 6 simulates the effect of a DC conversion for St. Louis teachers. This plan would freeze accrual of pension wealth under the current DB plan and convert all future contributions toward a less expensive (but still generous by private sector standards) DC plan. Specifically, the teacher would receive the full value of her accrued pension wealth. We assume a 4% annual nominal return on this balance. Going forward, the teacher would contribute 5% of her salary to the DC plan (which matches her contribution rate to the DB plan) and the district would contribute 10%. The district contribution rate is less than the current cost of the DB plan.

The last column of Table 7 shows the behavior of our representative senior teacher under such a hypothetical DC plan. The mandatory conversion to a DC plan produces a substantial increase in additional teaching years and retirement age. As compared to the current plan, completed service years at retirement rises by 2.5 and 1.7 years for females and males, respectively. Average age at retirement rises from 56.4 to 58.9 for females and 56.5 to 58.2 for males. These substantial increases from a DC conversion are robust to modest changes in underlying assumptions regarding returns or contribution. This exercise highlights the potent effect of the “push” incentives built into both the current and pre-enhancement DB plans. A DC conversion neutralizes the “push” and thus substantially extends a teaching career.<sup>5</sup>

---

<sup>4</sup>The return in a DC plan depends on investment risk taken by teachers. In our analysis we assume a 4% nominal return and 3% inflation, which implies a 1% real return. Our behavioral model does not incorporate investment risk as a factor affecting retirement timing.

<sup>5</sup>Unpublished tabulations available from the authors show that the DC effect on career lengths are relatively robust around our age 50, 25 service years teacher. For a given age, the DC impact rises with experience. For a given experience, the impact declines with age. This pattern reflects the impact of the 60%/30 years service annuity cap. For a given age, the “push” effect is larger as service years increase. For a given experience, the push effect falls with age because older teachers would have been more likely to retire as they passed 30 years of service anyway.

Of course, a mandatory conversion to DC may reduce the welfare of late-career teachers who are far from reaching the pension wealth peak. As a consequence such a mandatory conversion would almost certainly be unconstitutional in most states (Monahan (2010)). However, some states that have implemented hybrid plans (combined DB and DC) for new hires have permitted more senior to teachers to voluntarily opt into the new plan (e.g., Goldhaber and Grout (2016)). In unpublished simulations, we find that if faced with a voluntary DC conversion option (even at rates of conversion between 80-90 percent of full pension wealth), many male and female teachers at or near peak value opt for a conversion. Teachers below the peak value in service years may also opt for the conversion, but their probability of accepting the conversion is much more sensitive to the terms of the offers than that of teachers near peak value. At a minimum, dynamic structural models can provide rough empirical estimates of the take up rates for buyout or conversion options.

## 6 Conclusion

Many states enhanced benefits in teacher retirement plans during the 1990s. This paper examines the effect of a major enhancement for teachers in St. Louis public schools in 1999. Pension rule changes in 1999 resulted in very large increases in pension wealth for active teachers, as well as a powerful increase in “push” incentives for earlier retirement. Simple descriptive statistics on retirement patterns before and after the enhancements suggest earlier retirement resulted. However, to better understand the long run effects of these changes, we estimated a structural retirement model using panel data for the St. Louis teachers. Simulations based on the structural estimates imply shorter completed teaching spells and earlier retirement ages for a representative senior teacher as a result of the enhancements. These shorter teaching spells imply a steady state with more teacher turnover and a larger share of novice teachers in classrooms. A large teacher value-added literature finds that

novice teachers (e.g., teachers with less than three years teaching experience) are less effective in raising student achievement (e.g., Clotfelter, Ladd, and Vigdor (2006); Wiswall (2013)). The latter would imply lower overall workforce quality. When combined with the Koedel and Xiang (2017) finding that the large Saint Louis teacher enhancements had little effect on retaining younger teachers, this would seem to call into question the efficacy of such a back-loaded compensation policy.<sup>6</sup>

By contrast, conversion to a less expensive DC type plan, which eliminates the “push” incentives into retirement, would substantially lengthen expected teaching careers for senior St. Louis teachers. A mandatory conversion to a DC plan would substantially extend expected teaching careers among senior teachers. A study of Tennessee teachers (Ni et al., 2020), who are covered by a similar DB retirement plan, finds substantial improvement in the quality of the senior workforce, along with a lengthening of teaching careers in simulated conversions to a DC type plan. Our simulations also suggest that many senior teachers would opt into a voluntary DC plan as well.

This paper has focused on the retirement behavior of senior teachers. We have not attempted to model the effect of pension plan changes on younger teachers or recruitment. As noted, Koedel and Xiang (2017) find little effect of the large pension enhancements on the short term turnover behavior of younger teachers. However, it is possible that fundamental changes to the plan itself, such as partial or total conversion to DC, would affect recruitment into the profession and early career turnover. While young teachers display little knowledge of the specifics of their retirement plans, there is some evidence that they would prefer a DC type plan (e.g., (Goldhaber and Grout, 2016), (DeArmond and Goldhaber, 2010)). However, recruitment and early career retention effects of fundamental plan redesign are beyond the

---

<sup>6</sup>In a study of the effect of a one-time early retirement incentive in Illinois in the 1992-93, Fitzpatrick and Lovenheim (2014) find null and some positive effects on student achievement in the first two years following the program. They conclude that the test scores increased as a result of an overall teacher quality, possibly due to less effective teachers being more likely to take the early retirement option, a result consistent with the finding of (Ni et al., 2020) from Tennessee data. We do not have direct measures of teacher quality to test this hypothesis for St. Louis teachers.

scope of this study.

As a post-script to this study, rising pension plan costs and growing unfunded liabilities have led plan administrators and school leaders to cut retirement benefits for new St. Louis teachers (hired in 2018 or later) and raise contribution rates for new teachers from 5 to 9 percent of salary. However, it will be many years before these newly hired teachers are eligible for retirement. In the meantime, the pension incentives described in this paper will be shaping retirement incentives, and shortening teaching careers of incumbent teachers for many years to come. The structural methods used in this paper can help policy makers better understand the short and long run staffing and fiscal consequences of changes in pension plan design.

**Table 1: Summary of Pension Rules**

	Before 1999	After 1999
Normal Pension	$Age + ServiceYears \geq 85$ (Rule of 85) or $Age \geq 65$ with $ServiceYears \geq 5$	$Age + ServiceYears \geq 85$ (Rule of 85) or $Age \geq 65$ with $ServiceYears \geq 5$
Early Pension	$Age \geq 60$	$Age \geq 60$
Replacement Factor	1.25%	2%
Maximum Benefit	No	60%

**Table 2: Sample Statistics in Base Year: Before and After 1999 Enhancements**

<b>Female Teachers</b>	# of Teachers	Age	Service Years (SY)	SY $\leq$ 20 (%)	20<SY $\leq$ 30 (%)	SY>30 (%)
Teachers in 1993	696	55.29	24.36	29.2%	44.4%	26.4%
Teachers in 2009	692	55.88	21.55	43.5%	36.6%	19.9%
<b>Male Teachers</b>	# of Teachers	Age	Service Years (SY)	SY $\leq$ 20 (%)	20<SY $\leq$ 30 (%)	SY>30 (%)
Teachers in 1993	184	55.00	24.54	27.7%	44.6%	27.7%
Teachers in 2009	198	56.19	19.78	54.5%	31.8%	13.6%

Note: The table compares two cohorts of teachers aged 50-62 in the 1992-93 and 2008-09 school years in the base year. The age and service years columns report the averages for teachers. The last three columns report intervals of service years

**Table 3: Sample Statistics: Before and After 1999 Enhancements**

<b>Female Teachers</b>	<b># of Teachers</b>		<b>Age</b>		<b>Service Years</b>	
	93-97	09-13	93-97	09-13	93-97	09-13
Female Teachers in 1993(2009)	696	692	55.29	55.88	24.36	21.55
Retired in 1993(2009)	37	64	58.70	56.70	27.19	25.75
Retired in 1994(2010)	34	131	59.97	58.71	30.88	25.31
Retired in 1995(2011)	32	48	60.91	58.92	31.06	24.79
Retired in 1996(2012)	39	56	60.69	59.09	31.36	23.77
Retired in 1997(2013)	46	97	60.59	60.29	30.48	25.68
Not Retired	508	296	58.27	58.54	26.99	23.33

<b>Male Teachers</b>	<b># of Teachers</b>		<b>Age</b>		<b>Service Years</b>	
	93-97	09-13	93-97	09-13	93-97	09-13
Male Teachers in 1993(2009)	184	198	55.00	56.19	24.54	19.78
Retired in 1993(2009)	16	17	57.38	56.41	28.38	23.24
Retired in 1994(2010)	6	35	59.67	58.86	31.67	22.74
Retired in 1995(2011)	18	7	56.83	58.71	27.44	19.57
Retired in 1996(2012)	12	22	60.42	60.14	32.92	22.55
Retired in 1997(2013)	17	35	60.71	60.26	31.53	23.46
Not Retired	115	82	58.00	59.11	26.54	22.62

Note: The table compares two cohorts of teachers aged 50-62 in the 1992-93 and 2008-09 school years. We track the two cohorts forward for five years. The age and service years columns report the averages for teachers in the base year, those who retired in each subsequent year, and who were not retired at the end of the sample period.

**Table 4: Regression Estimates: Linear Probability Models**

	<i>Dependent variable: Retirement</i>	
	(1) Female	(2) Male
With Pension Enhancement	0.366*** (0.039)	0.300*** (0.077)
Service Years	0.011*** (0.001)	0.010*** (0.003)
Age	0.048*** (0.003)	0.040*** (0.007)
Salary(\$000)	-0.003** (0.002)	-0.006* (0.003)
Constant	-2.524*** (0.184)	-1.872*** (0.376)
Observations	1,388	382
R <sup>2</sup>	0.264	0.161
Adjusted R <sup>2</sup>	0.262	0.152
Residual Std. Error	0.424 (df = 1383)	0.461 (df = 377)
F Statistic	124.116*** (df = 4; 1383)	18.090*** (df = 4; 377)

Note: The table shows the results of linear probability model estimated on the two cohorts of teachers in Table 3 in the base year.

Model:  $Retire_i = \beta_0 + \beta_1 Enhancement_i + \beta_2 ServiceYears_i + \beta_3 Age_i + \beta_4 Salary_i + \epsilon_i$ ;

Dummy variable  $Retire = 1$  if the teacher retired during the panel;

Dummy variable  $With Pension Enhancement = 1$  if the teacher is in the 09-13 cohort;

\*p<0.1; \*\*p< 0.05; \*\*\*p<0.01. The standard errors are in parentheses.

**Table 5: Structural Model Parameters**

Parameters	Economic Interpretation
$\beta \in (0, 1)$	discount factor
$k \in (0, 1) \ \& \ k_1 > 0$	$k_s = k(\frac{60}{age})^{k_1}$ : preference of teaching versus retiring
$\gamma \in (0, 1]$	curvature in the utility function ( $\gamma < 1$ indicates concavity)
$\sigma > 0$	magnitude of unobserved preference shocks
$\rho \in (-1, 1)$	persistence in unobserved preference shocks

**Table 6: Estimates of Structural Parameters**

	Female	Male
$\beta$	0.964 ( 0.014)	0.963 (0.026)
$k$	0.682 (0.061)	0.688 (0.108)
$k_1$	0.900 (0.100)	0.751 (0.189)
$\sigma$	4001.108 (388.878)	3770.432 (853.972)
$\gamma$	0.684 (0.015)	0.693 (0.027)
$\rho$	0.637 (0.030)	0.682 (0.053)
Log-Likelihood	-1053.697	-314.897
Number of Teachers	692	198

Note: Parameters are estimated on post-enhancement (09-13) cohort.

The standard errors are in parentheses.

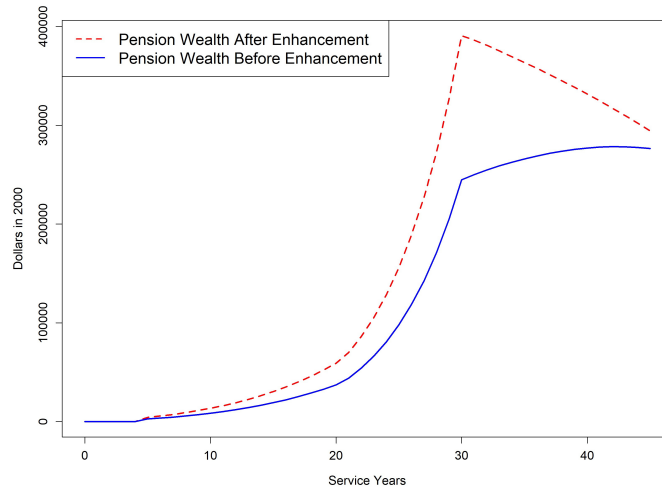


**Table 7: Effects of Alternative Pension Rules on a Representative Senior Teacher**

<b>Representative Female Teacher</b>		With Enhancements	Policy A	Policy B	No Enhancements	DC Policy
	Replacement Factor	2%	2%	2%	1.25%	n.a.
Pension Rules	Cap	60%	No	No	No	n.a.
	COLA Cumulative Max.	No Cap	No Cap	1.1	1.1	n.a.
<hr/>						
	Average Retired Age	56.4	57.1	57.7	59.1	58.9
	Average Retired Service Years	31.4	32.1	32.7	34.1	33.9
	Additional Years of Teaching	6.3	7.0	7.7	9.0	8.8
	Difference in Additional Years	-	0.7	1.3	2.6	2.5
<hr/>						
<b>Representative Male Teacher</b>		With Enhancements	Policy A	Policy B	No Enhancements	DC Policy
	Replacement Factor	2%	2%	2%	1.25%	n.a.
Pension Rules	Cap	60%	No	No	No	n.a.
	COLA Cumulative Max.	No Cap	No Cap	1.1	1.1	n.a.
<hr/>						
	Average Retired Age	56.5	57.0	57.3	58.4	58.2
	Average Retired Service Years	31.5	32.0	32.3	33.4	33.2
	Additional Years of Teaching	6.4	6.9	7.2	8.3	8.1
	Difference in Additional Years	-	0.5	0.8	1.9	1.7

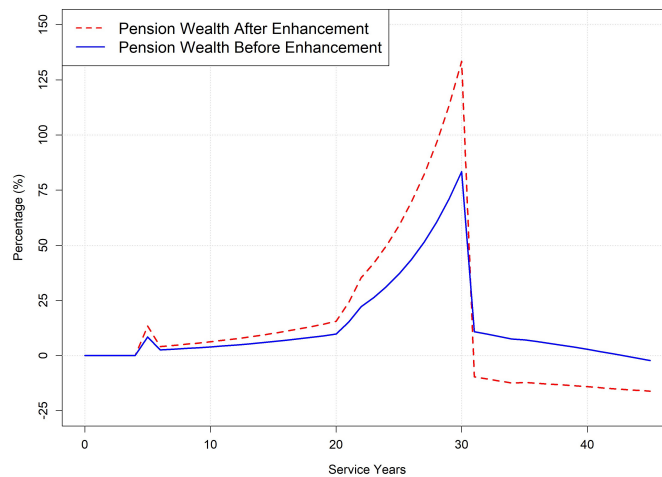
Note: The simulation tracks a representative female and male teacher aged 50 with 25 years of service for the next 30 years under four cases: a baseline with enhancements; Policy A eliminates the pension cap of 60%; Policy B further imposes the maximum increase of 10% on the retirement COLA adjustment; before-enhancement policy; Defined Contribution (DC) Policy with investment returns of 4%. Additional teaching years is the number of years the teacher is expected to teach before retirement. Difference in additional teaching years is the additional teaching years under an alternative policy minus that under the baseline DB policy (with the enhancements in place.) The DC plan assumes that teachers contribute 5% and the district contributes 10% of salary.

**Figure 1: Accumulated Pension Wealth by Service Years for a Teacher Entering at Age 25**



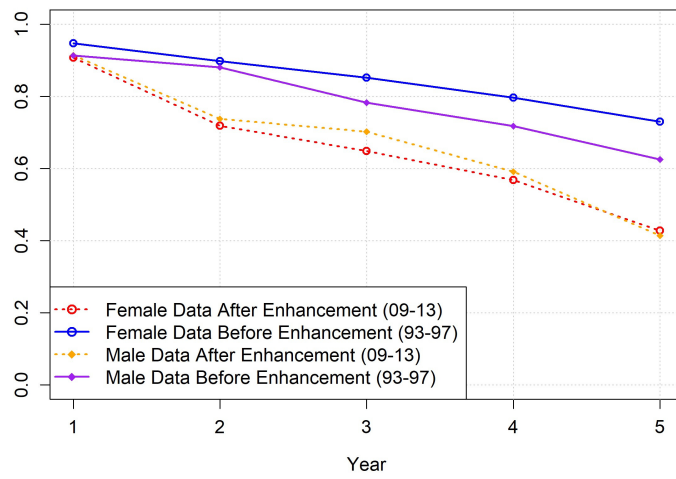
Note: The figure compares pension wealth accrual before and after the 1999 pension enhancements for a representative teacher in St. Louis who enters the system at age 25. For prediction of salary we use a cubic function of service years.

**Figure 2: Annual Changes in Pension Wealth for a Teacher Entering at Age 25**



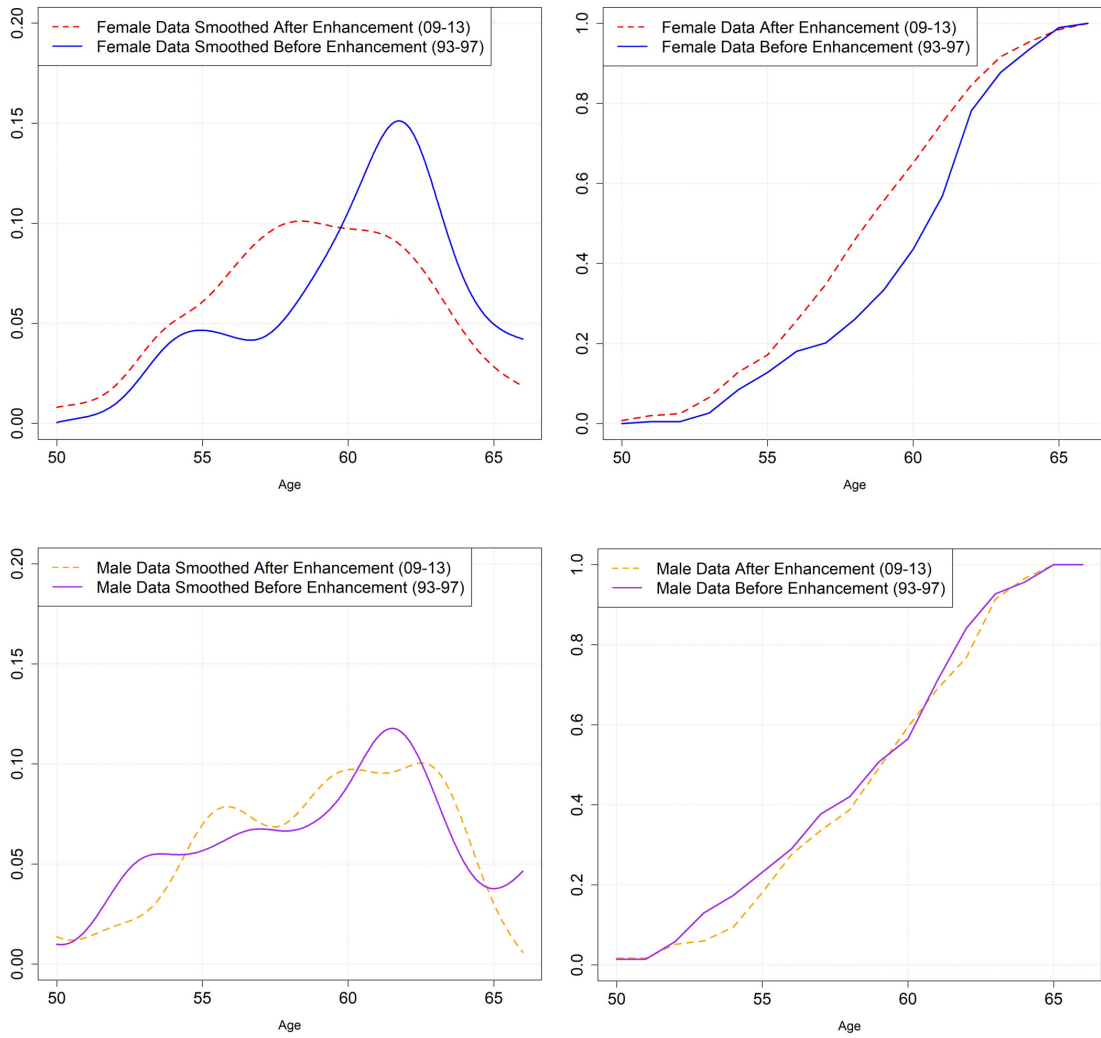
Note: The figure compares the pension wealth accumulation under the pension rules before and after the 1999 pension enhancements.

Figure 3: Employment Survival Rate Before and After Pension Enhancements



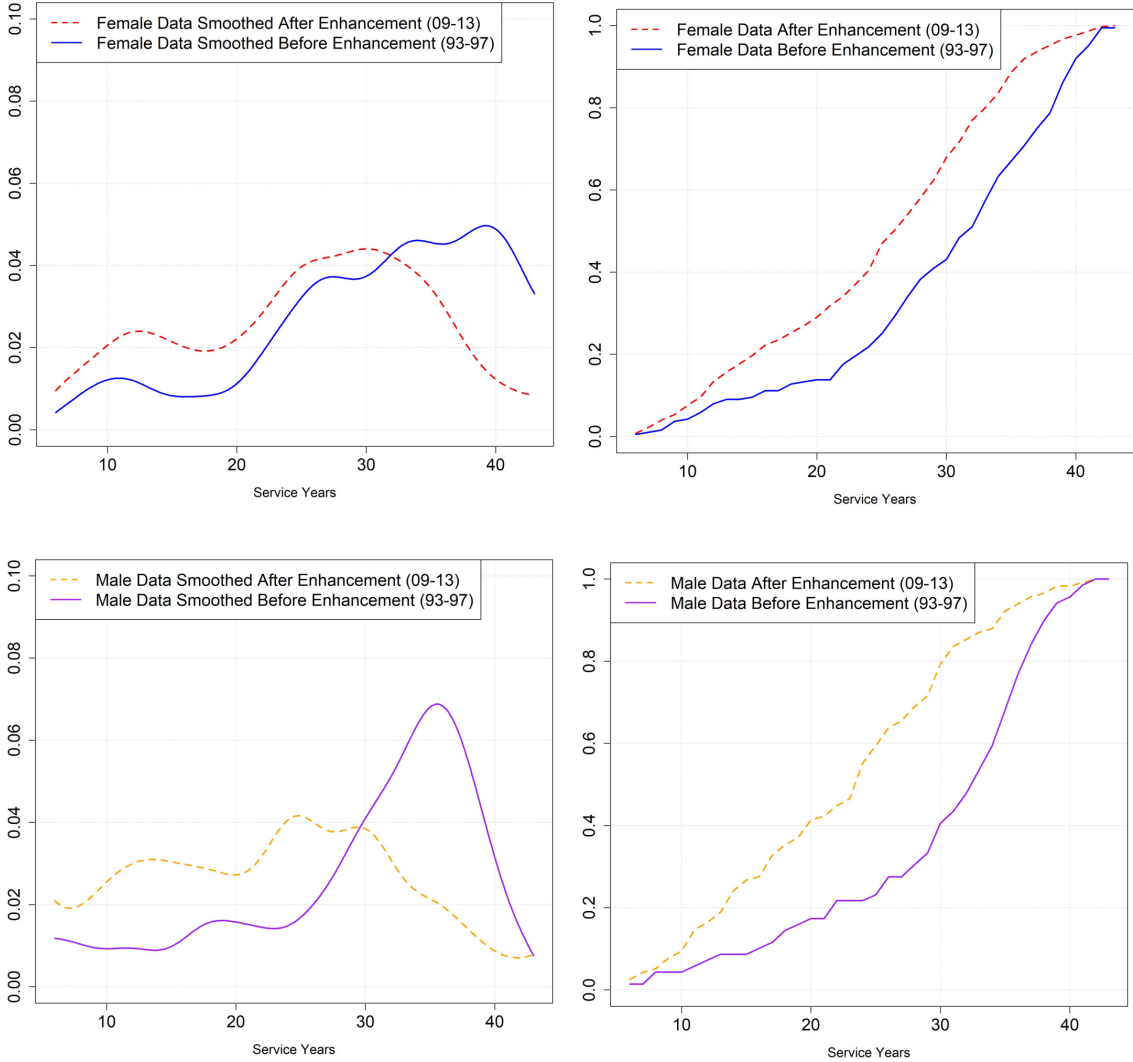
Note: The figure compares the employment survival rates for two cohorts of teachers (for female and male teachers) aged 50-62 at the base year and then tracks them for the next five years (93-97 and 09-13).

Figure 4: Age Distribution Before and After Pension Enhancements



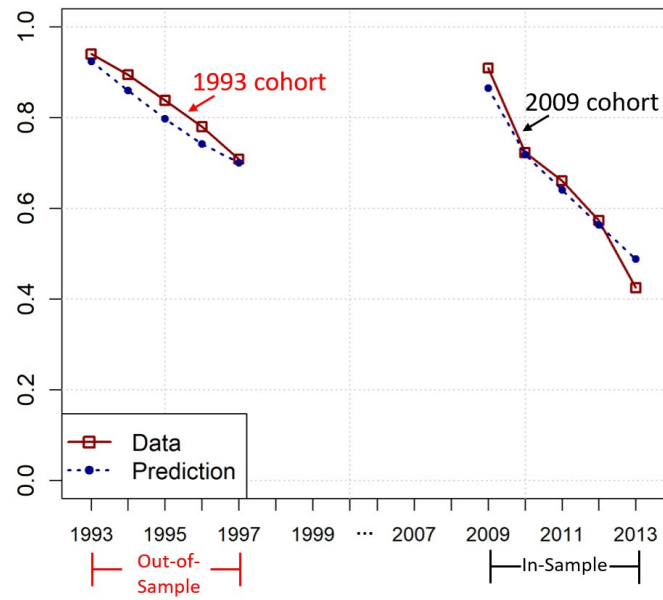
Note: The figure compares the age distributions for retired teachers (female and male teachers) before and after the 1999 enhancements. The teachers are aged 50-62 at the base year of two periods. The left graph compares the probability density, while the right graph compares the cumulative distribution. Please note that the figures are kernel-smoothed density estimates of data.

Figure 5: Service Years Distribution Before and After Pension Enhancements



Note: The figure compares the service years distributions for retired teachers (female and male teachers) before and after the 1999 enhancements. The teachers are aged 50-62 in the base year of the two periods. The left graph compares the probability density, while the right graph compares the cumulative distribution. The figures are kernel-smoothed density estimates.

Figure 6: Observed and Predicted Survival Rate for All Teachers (In Sample and Out of Sample)



Note: The figure compares the observed and predicted survival rate of two cohorts of teachers: in-sample (2009 cohort) and out-of-sample (1993 cohort) teachers. Male and female teachers are pooled.

## **Acknowledgments**

This research was supported by the National Center for the Analysis of Longitudinal Data in Education Research (CALDER), which is funded by a consortium of foundations. For more information about CALDER funders, see [www.caldercenter.org/about-calder](http://www.caldercenter.org/about-calder). All opinions expressed in this paper are those of the authors and do not necessarily reflect the views of our funders or the institutions to which the author(s) are affiliated. CALDER working papers have not undergone final formal review and should not be cited or distributed without permission from the authors. They are intended to encourage discussion and suggestions for revision before final publication. We also thank the editor and two referees for insightful comments and suggestions.

## References

- Börsch-Supan, A. and V. A. Hajivassiliou (1993). Smooth unbiased multivariate probability simulators for maximum likelihood estimation of limited dependent variable models. *Journal of Econometrics* 58(3), 347–368.
- Brown, K. M. (2013). The link between pensions and retirement timing: Lessons from California teachers. *Journal of Public Economics* 98, 1–14.
- Butrica, B., H. Iams, K. Smith, and E. Toder (2009). The disappearing defined benefit pension and its potential impact on the retirement incomes of baby boomers. *Social Security Bulletin* 69(3), 1–28.
- Clotfelter, C., H. Ladd, and J. Vigdor (2006). Teacher-student matching and the assessment of teacher effectiveness. *Journal of Human Resources* 41(4), 778–820.
- Costrell, R. M. and J. B. McGee (2010). Teacher pension incentives, retirement behavior, and potential for reform in Arkansas. *Education Finance and Policy* 5(4), 492–518.
- Costrell, R. M. and M. Podgursky (2009). Peaks, cliffs, and valleys: The peculiar incentives in teacher retirement systems and their consequences for school staffing. *Education Finance and Policy* 4(2), 175–211.
- DeArmond, M. and D. Goldhaber (2010). Scrambling the nest egg: How well do teachers understand their pensions and what do they think of alternatives? *Education Finance and Policy* 5(4), 558–586.
- Fitzpatrick, M. and M. Lovenheim (2014). Early retirement incentives and student achievement. *American Economic Journal: Economic Policy* 6(3), 129–154.
- Friedberg, L. and A. Webb (2005). Retirement and the evolution of pension structure. *Journal of Human Resources* 40(2), 281–308.
- Furgeson, J., R. P. Strauss, and W. B. Vogt (2006). The effects of defined benefit pension incentives and working conditions on teacher retirement decisions. *Education Finance and Policy* 1(3), 316–348.



- Goldhaber, D. and C. Grout (2016). Which plan to choose? the determinants of pension system choice for public school teachers. *Journal of Pension Economics and Finance* 15(1), 30–54.
- Harris, D. N. and S. J. Adams (2007). Understanding the level and causes of teacher turnover: A comparison with other professions. *Economics of Education Review* 26(3), 325–337.
- Heckman, J. J. (1981). The incidental parameters problem and the problem of initial conditions in estimating a discrete time process. In Charles F. Manski and Daniel McFadden, eds. *The Structural Analysis of Discrete Data*, MIT Press, Cambridge, MA, pages 179–195.
- Knapp, D., K. Brown, J. Hosek, M. G. Mattock, and B. J. Asch (2016). *Retirement Benefits and Teacher Retention*. Rand Corporation.
- Koedel, C., S. Ni, and M. Podgursky (2014). Who benefits from pension enhancements? *Education Finance and Policy* 9(2), 165–192.
- Koedel, C. and B. Xiang (2017). Pension enhancements and the retention of public employees. *Industrial and Labor Relations Review* 70(6), 519–551.
- Monahan, A. (2010). Public pension plan reform: The legal framework. *Education Finance and Policy* 8(4), 617–647.
- Moody, J. and A. Randazzo (2020). Hidden education funding cuts: How growing teacher pension debt is eating into k-12 education budgets. Technical report, Equable Institute.
- Ni, S. and M. Podgursky (2016). How teachers respond to pension system incentives: New estimates and policy applications. *Journal of Labor Economics* 34(4), 1075–1104.
- Ni, S., M. Podgursky, and X. Wang (2020). Teacher pension plan incentives, retirement decisions, and workforce quality. *Journal of Human Resources*, published ahead of print, doi:10.3368/jhr.57.1.1218–9912R2.
- Stock, J. H. and D. A. Wise (1990). Pensions, the option value of work, and retirement. *Econometrica* 58(5), 1151–1180.
- Wiswall, M. (2013). The dynamics of teacher quality. *Journal of Public Economics* 100(1), 61–78.

Wooldridge, J. M. (2005). Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of Applied Econometrics* 20(1), 39–54.

# Appendices

## Appendix A Statistical Appendix

### A.1 Numerical Computation of Retirement Probability

Continuing with the notation from Section 4, denote  $f_t^+ = g_t(r^+)/K_t(r^+)$ . Suppose year 1 is the first year of our sample period, and year  $N$  is the last year of the sample period. Denote  $V_{t_1,t_2} = (\nu_{t_1}, \dots, \nu_{t_2})$ ,  $F_{t_1,t_2}^+ = (f_{t_1}^+, \dots, f_{t_2}^+)$ . The unconditional likelihood (unconditional probability) that teacher  $i$  chooses to retire at year  $n$  ( $n < N$ ) is

$$\text{prob}(V_{1,n} \in A_i) = \text{prob}(((V_{1,n-1} > -F_{1,n-1}^+) \cap (v_n < -f_n^+))) \quad (\text{A.1})$$

The unconditional probability that teacher  $i$  chooses to stay until year  $N$  is

$$\text{prob}(V_{1,N} \in A_i) = \text{prob}(((V_{1,N} > -F_{1,N}^+))) \quad (\text{A.2})$$

These probabilities depend on pension rules, teachers' age and service years, and structural parameters  $(\gamma, k, k_1, \beta, \sigma, \rho)$ . Under the normality assumption of preference errors, computing the probabilities for the retirement decision made by teacher  $i$  involves a high dimensional numerical integration of the correlated preference errors  $V$  over the region  $A_i$ , for which we use the Geweke–Hajivassiliou–Keane (GHK) algorithm. The GHK algorithm transforms the problem of high dimensional integration to a problem of sequential one-dimensional integrations. For more applications of the GHK algorithm see for example Börsch-Supan and Hajivassiliou (1993).

## A.2 Censoring in the Baseline Sample

The sample excludes those teachers whose age and service years would have put them in our sample but who chose to retire before our sample period. This results in sample selection bias. Sample selection bias of this sort will lead to overprediction of retirement for relatively young teachers (Ni and Podgursky, 2016). We correct for sample selection bias in our baseline sample by using conditional probability to construct the likelihood function. The probability of teachers retiring in our sample period is based on the condition that teachers are not retired at the beginning of sample period.

Suppose in the first year of our sample period, teacher  $i$  was eligible for retirement  $J$  years ago. The probability of teacher  $i$  retiring at time  $n < N$  conditional on being in the sample is:

$$\begin{aligned}
 & \text{prob}(V_{1,n} \in A_i \mid \text{appearing in sample}) \\
 &= \text{Prob}(\text{retiring at } n \mid \text{appearing in sample}) \\
 &= \frac{\text{Prob}(\text{retiring at } n \ \& \ \text{appearing in sample})}{\text{Prob}(\text{appearing in sample})} \\
 &= \text{Prob}[\left((V_{1,n-1} > -F_{1,n-1}^+) \cap (v_n < -f_n^+)\right) \mid V_{-J,0} > -F_{-J,0}^+] \\
 &= \frac{\text{Prob}[(V_{-J,n-1} > -F_{-J,n-1}^+) \cap (v_n < -f_n^+)]}{\text{Prob}(V_{-J,0} > -F_{-J,0}^+)}
 \end{aligned}$$

The conditional probability of teacher  $i$  chooses to stay until the end of sample period  $N$  is:

$$\begin{aligned}
 & \text{prob}(V_{1,N} \in A_i \mid \text{appearing in sample}) \\
 &= \text{Prob}(\text{staying at } N \mid \text{appearing in sample}) \\
 &= \frac{\text{Prob}(\text{staying at } N \ \& \ \text{appearing in sample})}{\text{Prob}(\text{appearing in sample})} \\
 &= \text{Prob}[\left((V_{1,N} > -F_{1,N}^+)\right) \mid V_{-J,0} > -F_{-J,0}^+] \\
 &= \frac{\text{Prob}[(V_{-J,N} > -F_{-J,N}^+)]}{\text{Prob}(V_{-J,0} > -F_{-J,0}^+)}
 \end{aligned}$$

Given the observations of teachers  $i = 1, \dots, I$  who retired during the sample period (at  $n_i < N$ ) or not retired until the end of the sample period ( $n_i = N$ ), the conditional likelihood

for the whole sample is

$$L(\gamma, k, k_1, \beta, \sigma, \rho) = \prod_{i=1}^I \text{prob}(V_{1,n_i} \in A_i \mid \text{appearing in sample}). \quad (\text{A.3})$$

We use the likelihood in equation (A.3) to estimate our structural model.

## Appendix B Goodness of Fit

### B.1 In-Sample Goodness of Fit

Figure B1 reports the observed and predicted distributions of age, service years, and the sum of age and service years for retired teachers. There is a sharp spike at 85 and 86 reflecting the “Rule of 85” (i.e., age + service years = 85) option for regular retirement.

While the observed age and service years distributions are somewhat choppy because of the relatively small sample sizes (particularly for male teachers,) overall they are in line with the predicted distributions. All the above figures show that the model nicely mimics the distribution of both age and service years, in particular the “Rule of 85”.

*(Figure B1)*

### B.2 Out-of-Sample Goodness of Fit

The out-of-sample simulation provides a robust check for the validity of our model. The data we use for out-of-sample simulation includes all teachers aged 50-62 in the 1993-97 school year. The basic statistics are shown in Table 3. The table shows that under different rules the patterns of retirement rates differ before and after the enhancement as well as the distributions of teacher service years at the time of retirement. Using the estimated

parameters in Table 6, we simulate the teachers retirement behaviors in 1993-97 school years. During this period, the teachers are under pre-enhancement pension rules. The out-of-sample simulation also take into account of the fact that teachers in 1993 are significantly different from those in 2009 in age and service years. Figure B2 reports the observed and predicted distribution of age, service years, and the sum of age and service years for retired teachers for the out-of-sample simulation.

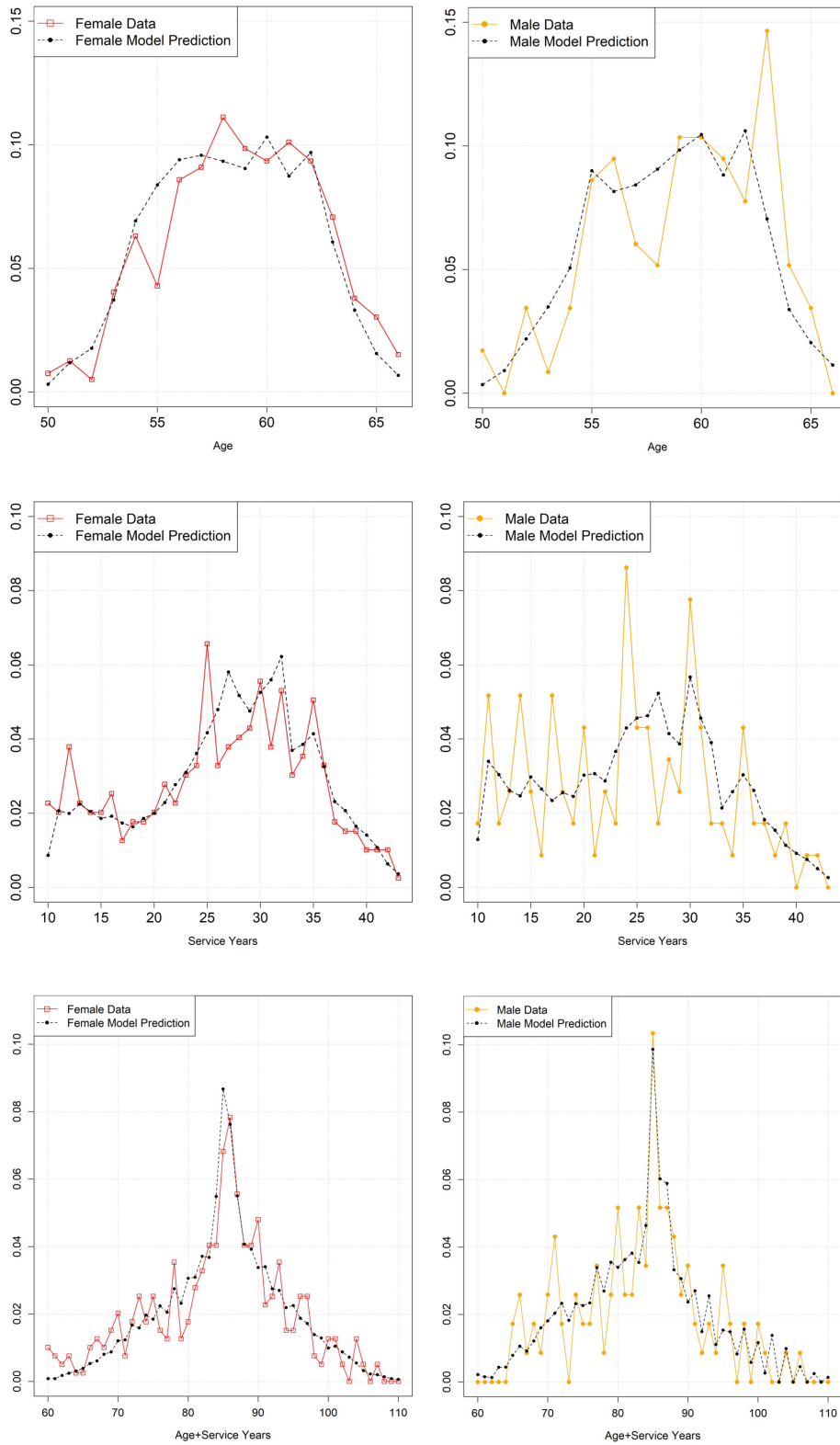
*(Figure B2)*

### **B.3 Overall Goodness of Fit**

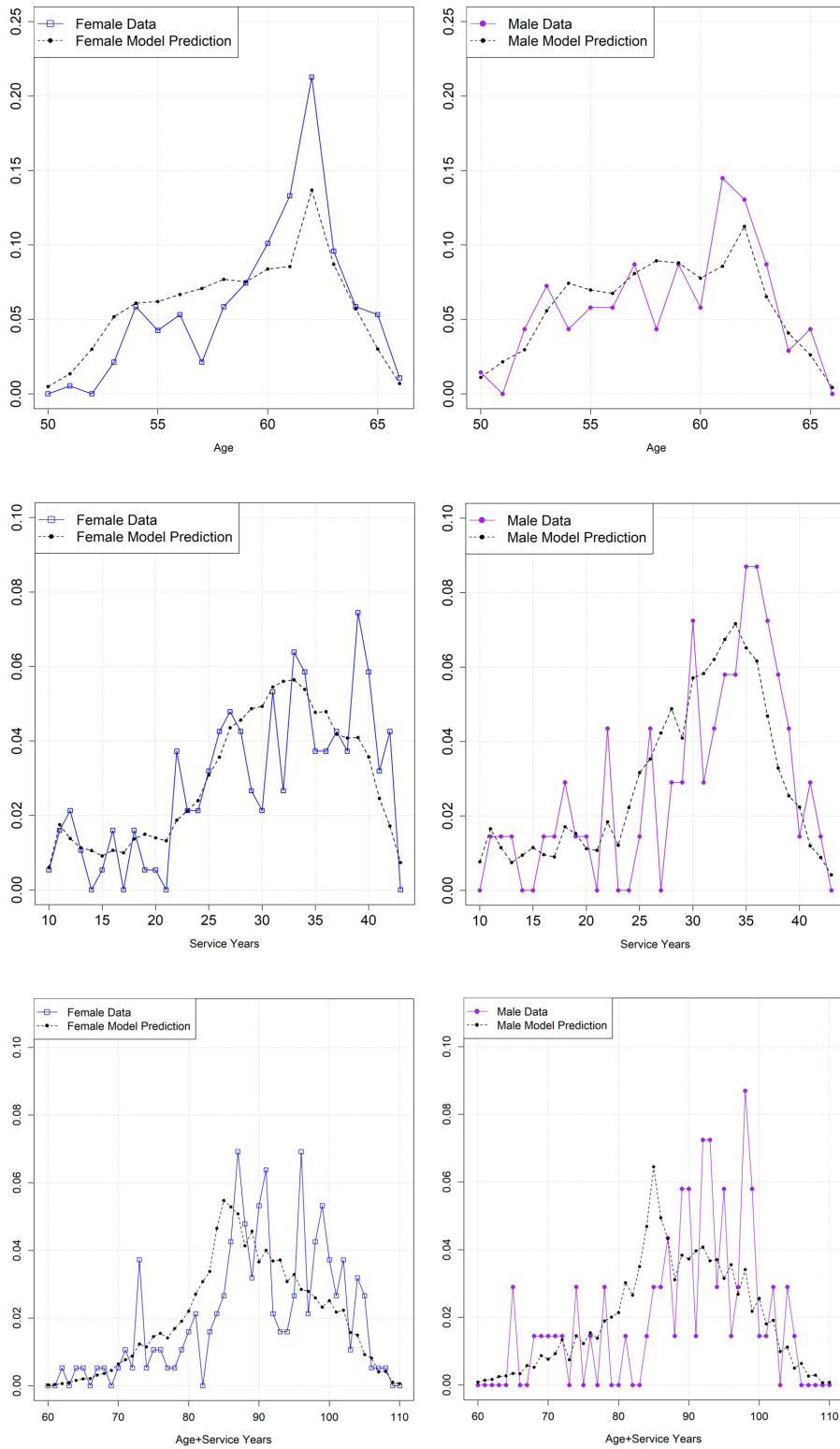
Figure B3 provides both the in-sample and out-of-sample goodness of fit of the survival rate. Figure B4 plots the observed and predicted in- and out-of-sample cumulative distributions of service years at the time of retirement. Overall, the model with the same parameters in teacher preference can capture the difference in retirement patterns for teachers under different pension rules.

*(Figure B3-B4)*

**Figure B1: Observed and Predicted Retired Age, Service years and Age+Service Years Distribution of Female and Male Teachers (In Sample)**

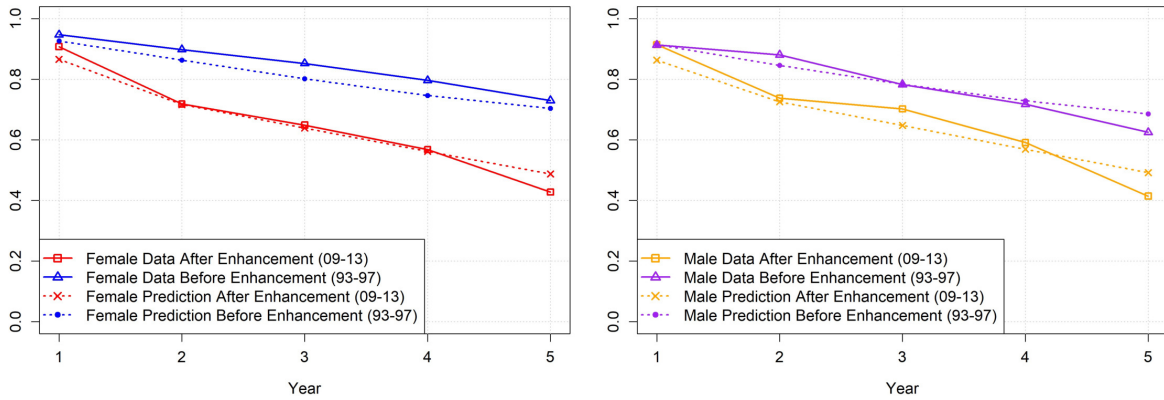


**Figure B2: Observed and Predicted Retired Age, Service Years and Age+Service Years Distribution of Female and Male Teachers (Out of Sample)**

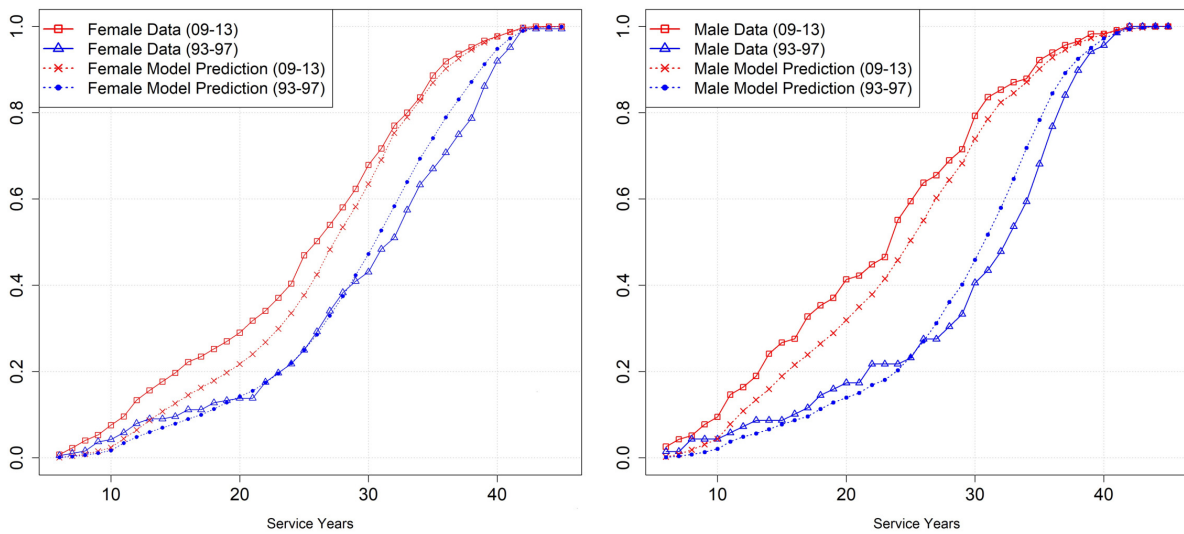




**Figure B3: Observed and Predicted Survival Rate (In Sample and Out of Sample)**



**Figure B4: Observed and Predicted Retired Service Years Cumulative Distribution of Female and Male Teachers (In Sample and Out of Sample)**



## Appendix C Defined Contribution Analytics

If in year  $t$  a teacher with age-experience  $(a, e)$  plans to retire in year  $r \geq t$ , her expected utility under the DB plan is denoted as  $\mathbb{E}_t V_t^{DB}(r)$  in (4.1). We use a similar notation for the expected utility under the following DC plan.

A teacher with (age,experience)  $(a, e)$  in the DB plan in a base year have cash balance  $W(a, e)$ . 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 of  $c_t$  and  $c_d$  fraction of salary for teachers and districts respectively, for a combined total of  $c_t + c_d$  times salary. A teacher's account accumulates with annual contributions and nominal investment returns of  $R - 1$  on the fund balance. The inflation rate is assumed to be  $i$ .

The teacher considers whether to retire or continue to work as in the SW model: her expected utility in period  $t$  is a function of expected retirement year  $r$  (with  $r = t, \dots, T$  and  $T = 101$  is an upper bound on age). For a teacher with a DC account value  $W_t = W(a, e)$ , the account nominal value in year  $r > t$  is the value of accumulation of contributions plus the compound return of the wealth in period  $t$ :  $W_r = W_t R^{r-t} + \sum_{k=t+1}^r (c_t + c_d) Y_k R^{r-k}$ ,  $Y_k$  is salary in year  $k$ .

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  $W_r$ , age  $a + (r - t)$  at the time of retirement (the teacher's age is  $a$  in year  $t$ ), let the expected nominal flow of the annuity be  $B_{r+n}$  in the  $n$ -th year of retirement. The retiree survives to  $r + n$  with conditional probability  $\pi(r + n|r)$ . The expected account value and the expected payment evolve from the collection  $r$  as

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

We set  $W_T = 0$ . It follows that given age  $a + (r - t)$  at the time of retirement has DC benefit

$$B_r = \frac{W_r(1+i)^{t-r}}{\sum_{n=1}^{T-a-(r-t)} \pi(r+n|r)\left(\frac{1+i}{R}\right)^n}. \quad (\text{C.1})$$

The benefit in time  $t$  dollar is  $B_r(DC) = B_r(1+i)^{r-t}$ .

In period  $t$ , the expected utility of retiring in period  $r$  under the DC plan,  $\mathbb{E}_t V_t^{DC}(r)$ , is the discounted sum of pre- and post retirement expected utility after replacing the DB annual benefit in (4.1) by the DC annuity  $B_r(DC)$ .