How Teachers Value Pension Wealth: A Reexamination of the Illinois Experience

Shawn Ni, University of Missouri-Columbia

Michael Podgursky, University of Missouri - Columbia and Saint Louis University

Fangda Wang, University of Missouri-Columbia Columbia

October, 2020

Abstract

In a widely-cited study, Fitzpatrick (2015) found that more than a quarter of Illinois teachers were unwilling to pay 19 cents for pension enhancements worth one dollar in present value. We revisit this finding by tracking the same cohort of teachers to retirement, which permits exact measurement of the annuity received and service years. The vast majority of teachers purchased the upgrade. Among the teachers who did not, the benefit on average had a negative value given their retirement timing. Our analysis finds that Fitzpatrick's instrumental variables fail to capture the underlying heterogeneity of preferences driving this result.

JEL classification codes: H75, I21, J26, J45

Shawn Ni: Department of Economics, University of Missouri-Columbia, 118 Professional Building, Columbia, MO 65211. Email: Nix@missouri.edu.

Michael Podgursky: Department of Economics, University of Missouri-Columbia, 118 Professional Building, Columbia, MO 65211. Email: PodgurskyM@missouri.edu.

Fangda Wang: Department of Economics, University of Missouri-Columbia, 118 Professional Building, Columbia, MO 65211. Email: fw48n@mail.missouri.edu.

The authors wish to acknowledge research support from the Center for Analysis of Longitudinal Data in Education Research (CALDER) and the Charles Koch Foundation, and helpful comments from Maria Fitzpatrick and Dan Goldhaber. The views expressed and findings reported in this paper are those of the authors and not necessarily those of the supporting organizations or individuals mention above.

1 Introduction

A growing literature finds that defined benefit (DB) teacher pensions induce early retirement of teachers and have created large fiscal liabilities in many states.[1] This has motivated a reexamination of the efficacy of these traditional plans versus alternatives more congruent with private sector practice. Many states are struggling with the challenge of paying for the liabilities that have already been accrued as well as those that will accrue in coming years. One important issue in these policy debates is how much teachers actually value benefits in their current plans. For example, if the accrued liabilities of these plans on the books are significantly less than the value attached to them by teachers, there may be buy-out options that can leave states and teachers better off.

A recent study of how Illinois public school teachers responded to a pension upgrade option offered in 1998 suggests that incumbent teachers may not place a very high value on a marginal dollar of pension benefits. Fitzpatrick (2015) analyzes the response of these teachers to an option allowing them to upgrade pension benefits at a bargain rate. Nonetheless, she finds that more than a quarter of the teachers declined to pay 19 cents for a dollar increase in the present value of pension wealth. Specifically, she reports that by 2009, 74% of teachers who had 22-28 years of experience in 1998 (and were most likely to benefit from the upgrade) had purchased the upgrade. She estimates that the "non-takers" seemingly passed up a substantial potential gain. "The average price of the upgrade offered to employees with 25 years of experience in 1998 was $15,245 while the expected benefit would have been $94,166." In a larger policy context, this suggests that if offered the option, many teachers would voluntarily cash out of the current, more expensive DB plan for an

[1]See, e.g., Novy-Marx and Rauh 2009, Costrell and Podgursky (2009), Burnette and Will (2018), Friedberg and Turner (2010); Costrell and McGee (2010); Brown (2013); Fitzpatrick and Lovenheim (2014); Knapp et al. (2016); Ni and Podgursky (2016).

alternative cheaper plan and an upfront cash payment that could on net reduce liabilities of the pension plan. Clearly this would offer one promising tool to help address pension funding challenges.[2]

In this study we reexamine Illinois teacher responses to the optional pension upgrade by tracking the 1998 cohort further to 2019 for upgrade purchases. This effectively takes the entire cohort to retirement. We find a much higher rate of purchases (87% versus Fitzpatrick's 74%). Importantly, by tracking these teachers all the way to retirement and observing their actual as opposed to predicted retirement annuities as well as their actual retirement timing, we find that of the 13% who did not purchase the upgrade 12% would not have benefited by doing so given Illinois pension rules, and conventional discount rates and life tables.

Tracking teachers to retirement highlights the important role of unobserved heterogeneity in work versus retirement preferences as well as the importance of taking into account the policy induced changes in retirement behavior. Illinois retirement benefits are capped at 75% of final average salary, with or without the upgrade. Thus, teachers who planned to work longer than 38 years would receive no benefit from the upgrade. For such job-committed teachers there is nothing irrational about declining the purchase. The ex ante pricing of the benefit is highly dependent on patterns of exit from work. Fitzpatrick used historical retirement patterns to price the benefit, but as we will show, these retirement patterns changed dramatically in response to the policy itself. Both of these factors make estimating a linear demand curve for the upgrade, and using average, ex ante historical values for price and cost, a problematic exercise. We also show that Fitzpatrick's use of IV methods based on ex ante observables does not address these problems or produce reliable estimates of a

[2]Not surprisingly, Fitzpatrick's surprising findings have been widely cited in the media has well. For references see Goldhaber and Holden (2020), a study of Washington teachers, which provides an alternative critique of Fitzpatrick's Illinois findings.

teacher's response to the benefits and price of the upgrade.

In the next section we briefly describe Illinois rules and the upgrade option and then present our revised estimates of the take-up rate and benefit of the upgrade. We also show that the takers and non-takers of the upgrade appear similar in 1998 but differ considerably at retirement. Next we show that Fitzpatrick's IV estimation produces biased estimates of the willingness to pay for the marginal teacher. In addition, we show that a linear probability model (LPM) is not well-suited to measure the way a marginal teacher values these highly non-linear benefits.

2 The Pension Upgrade and Teacher Responses

2.1 The "2.2 upgrade"

The annuity a teacher collects at retirement is based on the formula: $Annual\_Benefit = S \times FAS \times R$, where S is service years, FAS final average salary, and R the replacement factor. Here, service years are the sum of all years of service credit earned in the system.[3] The final average salary is the average of the four highest consecutive annual salaries within the last ten years. The replacement factor is 2.2% for each service year earned after July 1, 1998. For service years earned prior to July 1, 1998 the replacement rate varies by experience: i) 1.67 percent for each of the first 10 years, ii) 1.9 percent for each of the second 10 years, iii) 2.1 percent for each of the third 10 years, iv) 2.3 percent for each year over 30 years.[4]

[3]Some additional credits are also allowed such as up to two years of unused sick leave. Optional service credit can also be purchased if approved. The service years credit can be fractional.

[4]There is another wrinkle in studying retirement behavior for this cohort. The annual retirement benefit is the larger of the two amounts based on either an "actuarial annuity calculation" or the formula discussed above. The latter is usually the larger number and typically treated as the default. The actuarial annuity calculation was terminated for teachers who became members on or after July 1, 2005. It is based on interest and mortality rate factors that are subject to change. According to TRS (2018), the actuarial calculation may give a higher annuity than average salary/service credit calculation in two cases: i) members with exceptionally long career, or ii) members with long periods of inactive status. The exact formula for the actuarial calculation is not publicly available, but we observe a small percentage of teachers (6%) who

After 1998, teachers were given the option to upgrade their replacement rate for service years prior 1998 to 2.2 percent by paying a price equal to:

$$P_{it} = \min\left(\frac{Exp_{1998}}{100}, \frac{20}{100}\right) \times Salary_{it}$$

where Exp_{1998} is the service year a teacher have earned before 1998, $Salary_{it}$ the highest salary rate during the four school years before the teacher applies to make the upgrade contribution. Typically, this is the salary of the teacher at the time of payment. This means that the cost of upgrade is approximately 1% for each year of service before 1998, capped at 20%.

It is important to note that with or without the upgrade, the retirement annuity is capped at 75 percent of final average salary. Under the old formula, teachers hit this cap at 38 years of service. Under the new formula they hit it at 34 years. Thus the benefit of the upgrade declines after 34 years and hits zero at 38 years. This means is that teachers who planned on longer careers and retirement at a later age receive smaller or zero benefits from the upgrade.

This is seen in Figure 1 where we plot pension wealth accrual with and without the upgrade for a representative teacher with 22 years of experience. The solid line plots pension wealth without the upgrade and the dashed line with it. The bottom panel plots the present value of the upgrade. Note that the value of the upgrade rises sharply until 32 years experience. It starts falling sharply when the teacher hits the 75 percent cap with the upgrade (34 years experience) and falls to zero at the pension cap without the upgrade (38 years experience). Since the average cost of the upgrade is roughly $14,000 the net value of the upgrade turns negative at roughly 37 years experience.

Figure 1 also shows that the value of the upgrade not only depends on rules surrounding

retired with pensions based on the actuarial calculation. These annuities are included in our reported findings. Further details are given in the appendix.

the pricing of the upgrade, but also how teachers respond to the new pension incentives. Consider a teacher who would have retired at 36 years under the old rules, with pension wealth indicated by point A. Let us suppose that with the upgrade option the teacher chooses to retire at 32 years (a representative response as we will see in the next section). Thus the gain in pension wealth for this teacher is not B-A, but rather C-A. Pricing the benefit of the upgrade at either 32 years (C-D) or 36 years (B-A) does not correctly measure the gain for this teacher.

Figure 1 shows that the value of the upgrade is strongly related to teacher preferences regarding career length. A teacher with strong preferences for teaching relative to retirement is more likely to retire in the low to negative return zone for the upgrade, hence is less likely to purchase it. Unfortunately, ex ante we do not know how strong these work versus retirement preferences are for any given teacher.

2.2 Updated Panel Data

Figure 1 illustrates the interaction between the return to the upgrade and the length of the teaching spell for a representative teacher. In order to explore this empirically, we constructed a data file with additional payroll information from Teacher Retirement System (TRS), and extended the sample period from 2008-09 to 2019, by which time nearly all the teachers have retired. Details about data construction are given in the Appendix Section B.

Tracking the teachers to retirement is important for two reasons. First, a teacher has the option of purchasing the upgrade all the way up to the date of retirement. Once she/he has retired, however, the option expires. Second, the potential costs and benefits of an upgrade can be calculated with much greater accuracy because for both takers and non-takers the retirement annuity and timing of retirement is actually observed rather than estimated. As we will see below, estimating the benefit of the annuity ex ante is problematic because the existence of the 2.2 option itself likely changes behavior for many teachers (or at least leads them to sort based on work-retirement preferences), thus making historical data a poor foundation for pricing benefits for any individual teacher.

Our estimates of the benefits are reported in Table 1 which reports the breakdown of the entire sample of teachers into the sum of eventual takers (87%) and non-takers (13%). The first thing to note is that tracking this group of teachers all the way to retirement produces a considerably higher take up rate than what Fitzpatrick finds (87% versus 74%). This discrepancy in the take-up rate stems largely from her merging of TRS data with Teacher Service Records (TSR) data that do not have matching service credits (see Appendix Section B.7.) Table 1 shows that the take-up patterns are fairly similar for males (89%) versus females (86%).

Did the 13% non-takers pass up a substantial net benefit from the upgrade? Columns 3, 4 and 5 of Table 1 report the cost, benefit and net benefit of the upgrade for both takers and non-takers. As noted, unlike Fitzpatrick's tabulations, which are based on estimates of the final pensions, for the vast majority of these teachers, these estimates of cost and benefits are based on actual plan data (e.g., the actual annuity collected at retirement versus a forecast). A surprising finding from Fitzpatrick's study is that teachers, either due to irrationality or very high discount rates, were leaving "money on the table" in rejecting the upgrade. However, what we see from the simple calculations in Table 1, based on actual service years and the final annuity paid, and using the same 2% real discount rate as Fitzpatrick, is that teachers seem to be making perfectly rational decisions based on conventional discount rates. The average net benefits for takers was $106,297, while the average for non-takers was -$9,596. Again, these patterns are fairly similar for males and females.[5]

Column 6 of Table 1 reports the net estimated benefit following Fitzpatrick (2015). Instead of using the actual benefit, here we weight the expected benefit by retirement probability in each future year. The retirement probability is based on historical retirement data (pre-enhancement) given teachers' age and service experience. In sharp contrast to Column 5, the average estimated benefit (net of cost) of non-takers for all teachers is $93,839, similar to that of takers, replicating the finding of Fitzpatrick (2015). This clearly shows the problem of using average retirement behavior to estimate the net benefit of the upgrade.

Figure 2 plots the realized net benefits from the upgrade (pooling males and females) for takers and non-takers. The non-taker distribution is highly concentrated just below zero, with a small right tail of positive values. The taker distribution is less concentrated, with a small left tail of values below zero.[6]

Table 2 provides insight into why the net benefits for takers and non-takers are so different. Here we report statistics on the age and experience distributions of the eventual takers and non-takers in 1998 (on the left panel) and at the time of retirement claim (on the right panel). It shows that in 1998 the takers were slightly older and had slightly longer service experience than the non-takers. At the time of retirement, however, the non-takers are much older and more experienced. The median experience for takers and non-takers in 1998 is 25, while at retirement, the median experience of takers is 33, while for non-takers it is 37. The similarity in observable features between the takers and non-takers in 1998 leads to similar estimated benefits, and the large differences at the time of retirement results in

[5]Median values, reported in the appendix, show a similar pattern of large positive benefits for takers and negative values for non-takers.

[6]The benefit of the upgrade depends on the discount rate. Fitzpatrick (2015) uses nominal annual discount rate of 5.1%. With the assumed inflation rate of 3% this implies a real discount rate of 2.1%, which is lower than the values commonly used. In the appendix Table B12 reports the benefit and net benefit with nominal annual discount rate of 7.1% and 9.1% (or real discount rate of 4.1% and 6.1%). The general pattern is qualitatively the same as Table 1.

large differences in actual benefits. This explains the sharp contrast between Columns 5 and 6 of Table 1.

The upper panel of Figure 3 plots the age and experience distributions of the eventual takers and non-takers in 1998, while the lower panel plots these variables at the time of retirement. This shows clear evidence that the two groups sorted to values of service years associated with the 75 percent salary cap, with the takers clustered around 32 and the non-takers around 36. A reasonable interpretation of the evidence thus far is that teachers who liked teaching and preferred more years in the classroom turned down the upgrade and pursued longer careers. Teachers less enamored with the classroom (relative to retirement) chose the upgrade and exited earlier. The decision as to whether to purchase the upgrade provides information about these heterogeneous unobserved preferences. Importantly, Figure 3 illustrates why Fitzpatrick's use of historical data on retirement behavior (based on prior pension rules) to estimate the benefit of the upgrade, and applying these estimates of benefits to price the benefit of the upgrade likely yields biased estimates for both takers and non-takers.[7]

2.3 IV Estimates Reconsidered

Simple descriptive statistics using data on actual service years at retirement and the actual annuity received suggests that most takers and non-takers are making rational decisions based on conventional discount rates. There is no paradox to be explained. In this section, we explore further why Fitzpatrick finds such surprising results regarding teacher valuation

[7] We also find a sharp contrast in the age/experience distributions at retirement between cohorts with 22-28 years of experience in 1988 and 1998. The cohort of 1998 retired after the introduction of the upgrade option, while a substantial portion of the 1988 cohort retired before the introduction of the upgrade. The retirement experience of the 98 cohort is much more concentrated around 32 to 34 years. This suggests that because the upgrade elevates the "pull" and "push" of the pension incentives, many teachers alter their retirement timing for more pension wealth gain from the rule change. The shift in retirement behavior in the post-upgrade era suggests estimation of upgrade benefits in 1998 based on the pre-upgrade era retirement data likely results in bias.

of the upgrade. She estimates a linear probability model of individual demand for the pension upgrade:

$$D_i = \beta_0 + \beta_1 P_i + \beta_2 B_i + \epsilon_i \quad (1)$$

where D_i (= 1 or 0) is individual i's decision of whether to upgrade, P_i is the price of the upgrade, and B_i is the benefit of the upgrade based on expected retirement timing, the annuity collected and longevity. Teachers will have different values of P_i and B_i. For empirical analysis (1) also includes control variables which we omit here for simplicity.

With $\beta_1 < 0$ and $\beta_2 > 0$, the ratio β_2/β_1 determines the combinations of cost and benefit that offset each other so that the propensity for the upgrade is unchanged. Suppose teachers who face a very high (low) benefit relative to cost will choose to upgrade (not to upgrade). If teachers are willing to pay $-\beta_2/\beta_1$ dollars of cost for each dollar of benefit, then some of them along the demand curve with the ratio of coefficients β_2/β_1 choose to upgrade and some choose not to upgrade. In principle, the fact that the fraction of upgrade is between 0 and 1 provides an opportunity to estimate this willingness to pay for a marginal teacher.

In estimating her demand model, Fitzpatrick notes that P_i and B_i may be correlated with the error term ϵ_i, thus producing biased estimates of the price effects. She uses the district beginning salary of BA holders and the expected upgrade benefit generated by district maximum salary of MA holders as two instruments for P_i and B_i. When she does this the paradox emerges. With these IVs, her 2SLS estimate of the absolute value of the ratio β_2/β_1 is generally 0.2 or less, which means that a teacher is willing to pay no more than twenty cents for a dollar of pension benefits. This finding is central to her policy conclusion that the marginal teacher attaches a very low value to an additional dollar of retirement benefits relative to price.

We reproduce in Table 3 her IV estimates of the demand for the upgrade. The coefficient on price, β_1 is -0.082 and the coefficient on the benefits β_2 is 0.014. The ratio $-\beta_2/\beta_1$ is 0.17. In her data set with estimated benefits, among non-takers the average value of B is $\$90,864$ and the average of P is $\$14,273$, which clearly suggests that non-takers attached a low value to the benefit upgrade.

In order to explore the credibility of this IV, we conducted a series of placebo tests on Fitzpatrick's sample. The key paradox in Fitzpatrick's findings is that the non-takers are turning down benefits that are well in excess of the cost (price). In the first panel of Table 3 we reproduce Fitzpatrick's estimates using data on estimated benefits. We then report a series of placebo tests in which we consistently held the value of benefits (B) below price (P) for all non-takers. We conducted three tests in which we change the value of benefits (B) for non-takers (estimates of B for takers are unchanged): (i) B = 0 , (ii) B=P the price of the upgrade, and (iii) a random B that is always below P. While the coefficient estimates on P and B do change with these placebos, we find in all three cases that the estimated ratio of $-\beta_2/\beta_1$ remains roughly 0.2. At a minimum this suggests that the estimated cross-sectional regression seems to be driven by the benefit/price ratio of the takers, who are the vast majority of the sample.

We next reestimate her model on our updated sample with observed benefits. The results are shown in the upper panel of In Table 4. As in her case, the coefficient on P is negative and significant and the coefficient on B is positive and significant, although both are smaller in absolute value. What is noteworthy for this discussion is that the ratio β_2/β_1 from the IV estimates is nearly identical, in the upper panel of Column 2 (based on estimated benefit from Fitzpatrick's data), -0.17, and in the lower panel of Column 4 (based on the actual benefit from our data), -0.14. However, in the updated sample the average value of B for the non-takers is $\$4,406$ and the average value of P is $\$14,002$. Thus, even with our updated sample based on actual benefits, the paradox persists.

In order to further explore this paradox, we generated data from a model in which each teacher solves for the optimal retirement decision jointly with the upgrade decision, given her salary and the pension rules in place in 1998. The teacher's utility depends on the present value of flows of salary before retirement and the pension benefit after retirement. We employ a plausible set of structural parameters regarding preferences (based on Ni and Podgursky (2016)), and with the real discount factor set at 2% to be consistent with the calculation of pension benefits thus far. Teachers differ in unobserved preferences for teaching. This unobserved heterogeneity results in different benefits from the upgrade, and different decisions regarding retirement and upgrade purchase by teachers with the same age and experience. In the appendix we provide more details on the theoretical model that generates data similar to the observed data in many dimensions, including the observed take-up rate for the upgrade.

A key feature of the optimal choice model is that teachers are discounting the benefits at exactly the same 2% rate that Fitzpatrick and we use in estimating B. What this means is that by construction teachers value a dollar of B at the same rate as P, hence if in the IV model the ratio β_2/β_1 captures the willingness to pay by the marginal teacher it should be roughly equal to unity. We conducted two sets of simulations based on two decision models. In the first, the decisions are entirely independent (conditional on salary) of the district salary schedule parameters. In the second simulation we added an unobserved effort that is multiplicative in the salary schedule in the district d and teacher i's year of service at period t. To reproduce the endogeneity that those exerting more effort tend to have higher demand for the upgrade, we then make the effort correlated with unobserved heterogeneity. That is, those exerting high effort would have higher salaries and tend to retire earlier, thus have greater demand for the upgrade as assumed with Fitzpatrick's IVs.

These simulated data both reproduce similar results concerning the ratio $-\beta_2/\beta_1$. In both cases the ratio is roughly 0.2. Note that in the first case, by construction teacher effort is uncorrelated with the district salary schedule which is the underlying theoretical justification for Fitzpatrick's cross district IV. However, the key point is that even in these simulated data sets in which a dollar of benefits is valued precisely at cost, the paradoxical finding is observed.[8]

The analysis above suggests that Fitzpatrick's IVs are problematic. However, an additional problem arises with her use of a linear probability model (1) to identify the preferences of a marginal teacher in these data. In the following we show that the ratio $-\beta_2/\beta_1$ from the estimates of the model (1) does not capture the "willingness to pay" for a marginal teacher if the true probability of an upgrade purchase is highly nonlinear in benefit and price. This problem is present even when the price and benefit of the upgrade are correctly measured.

Consider the case of a continuum of teachers, each is assigned a price (P) and a benefit (B) for the pension upgrade. Teacher i upgrades ($D_i = 1$) if the benefit $B_i > P_i$ and chooses $D_i = 0$ otherwise. So the cost/benefit ratio for the marginal teacher is 1. We choose the set of parameters $(\beta_0, \beta_1, \beta_2)$ that minimizes the average of squares of residuals $E_{B>P}[1 - (\beta_0 + \beta_1 P + \beta_2 B)]^2 + E_{B \leq P}(\beta_0 + \beta_1 P + \beta_2 B)^2$. To simplify discussion we assume B and P are independent and uniformly distributed on $(0, B_{max})$ and $(0, 1)$ respectively, with $B_{max} \geq 1$. Among the teachers who have $B > 1$ the probability of upgrade is 1, a

[8]Goldhaber and Holden (2020) provide an alternative critique of Fitzpatrick's use of district salary schedule parameters as valid IVs. A fundamental problem in estimating a price effect independent of benefits in the demand equation that Fitzpatrick attempts to estimate is that both the price and benefit of the upgrade are a function of teacher salaries. Thus, price effects are inherently mixed with income effects. Goldhaber and Holden carefully consider the conditions under which district salary schedule parameters used as IVs can address this problem. They show that teacher base salaries, set by salary schedules, would have to be uncorrelated with supplemental salaries (e.g., extra duty pay, bonuses). Given the facts of teacher salary determination in practice, and evidence from Washington state teacher salaries, they find this assumption implausible.

constant, instead of a linear function of price and benefit. In general the ratio $-\beta_2/\beta_1$ for the least squares estimate of (β_1, β_2) depends on the joint distributions of price and benefit across teachers. It can be shown that for independent and uniformly-distributed prices and benefits, $\beta_1 = -1/B_{max}, \beta_2 = (3B_{max} - 2)/(B_{max})^3$, and $-\beta_2/\beta_1 = (3B_{max} - 2)/(B_{max})^2$. If $B_{max} = 10$ then $-\beta_2/\beta_1 = 0.28$.[9] As B_{max} goes to infinity the ratio $-\beta_2/\beta_1$ shrinks to 0, despite the fact that for a marginal teacher the price to benefit ratio is 1. This is because the estimate of (β_1, β_2) that fits the probability of upgrade for the whole population is heavily driven by the distribution of benefit and price of the takers, who have much higher benefit than cost. This problem is caused by the use of a misspecified model given the nature of the data and is not remedied by the use of instrumental variables.

3 Conclusion

Fitzpatrick (2015) finds that only 74% of Illinois teachers were willing to purchase a pension upgrade that in present value terms was seemingly worth six or seven times the cost. On the face of it, then, this suggests that the compensation of Illinois teachers is inefficient, with too much compensation loaded into retirement benefits, and too little into salary. Indeed, Fitzpatrick's estimates suggests that the marginal senior teacher only valued an additional dollar of retirement benefits at roughly 20 cents.

Our reexamination of the Illinois experience, with data on the actual retirement timing and annuity collected, as opposed to estimates, found that a higher share of teachers actually opted to purchase the upgrade (87%). Importantly, for teachers who declined to purchase the upgrade, the discounted net benefit was actually negative. The fundamental problem is that the value of the upgrade for any teacher depends on unobserved preferences for

[9]In the data, the distribution of benefits is not uniform so this number is not meant to be compared to those in the table.

work versus retirement. Given the 75 % cap on benefits, teachers who worked longer spells and approached the cap derived no net benefits from purchasing the upgrade. The ex ante administrative teacher data used by Fitzpatrick provides little information about these unobserved preferences for work versus retirement. The IV methods employed in her analysis do not adequately address this issue.

Subsequent studies have attempted to assess how much teachers value retirement benefits. Biasi (2019) analyzes retirement behavior of Wisconsin teachers in response to legislation restricting collective bargaining rights for teachers (Act10). Goldhaber and Holden (2020) analyze decisions of teachers in a Washington state hybrid plan providing teachers with a range of choices regarding DC contributions. Johnston (2020) and Fuchsman et al. (2020) analyze survey data presenting teachers with various options of pay and benefits to back out preferences. These studies generally find that a typical teacher values an incremental dollar of retirement benefits less than an additional dollar of salary, but at values above Fitzpatrick's 20 cent estimate. Such findings have obvious implications for efficient compensation design for teachers, as well as for policies to "buy down" pension plan liabilities by offering teachers less than a dollar for a dollar of liabilities. Our reexamination of the Illinois 2.2 upgrade experience, however, leads us to conclude that key features of this program, particularly its binding benefit cap combined with subsequent teacher choices regarding retirement timing, made it largely uninformative regarding these important policy questions.

References

Biasi, B. (2019). Higher salaries or higher pensions? Inferring preferences from teachers' retirement behavior.

Brown, K. M. (2013). The link between pensions and retirement timing: Lessons from California teachers. *Journal of Public Economics 98*, 1–14.

Burnette, D. and M. Will (2018). Costly pension plans are fanning the flames of teacher unrest. *Education Week 37*(28), 1.

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.

Fitzpatrick, M. D. (2015). How much are public school teachers willing to pay for their retirement benefits? *American Economic Journal: Economic Policy 7*(4), 165–88.

Fitzpatrick, M. D. and M. F. Lovenheim (2014). Early retirement incentives and student achievement. *American Economic Journal: Economic Policy 6*(3), 120–54.

Friedberg, L. and S. Turner (2010). Labor market effects of pensions and implications for teachers. *Education Finance and Policy 5*(4), 463–491.

Fuchsman, D., J. McGee, and G. Zamarro (2020). Teacher's willingness to pay for retirement benefits: A national stated preferences experiment.

Goldhaber, D. and K. Holden (2020). How much do teachers value deferred compensation? Evidence from defined contribution rate choices.

Johnston, A. (2020). Teacher preferences, working conditions, and compensation structure.

Knapp, D., K. Brown, J. Hosek, M. G. Mattock, and B. J. Asch (2016). *Retirement Benefits and Teacher Retention*. Rand Corporation.

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.

TRS (2018). Evolution of the TRS benefit structure.

Table 1: Summary Statistics on Takers and Non-Takers

	number	percent	average cost ($)	average benefit ($)	average net benefit ($)	average estimated net benefit ($)
	(1)	(2)	(3)	(4)	(5)	(6)
All Teachers	19,126	100.0%	15,093	106,411	91,318	100,771
–taker	16,654	87.1%	15,255	121,552	106,297	101,800
–non-taker	2,472	12.9%	14,002	4,406	-9,596	93,839
Female	12,506	100.0%	14,135	104,441	90,305	98,860
–taker	10,794	86.3%	14,277	120,286	106,009	99,893
–non-taker	1,712	13.7%	13,245	4,540	-8,704	92,345
Male	6,620	100.0%	16,902	110,133	93,230	104,382
–taker	5,860	88.5%	17,057	123,884	106,827	105,313
–non-taker	760	11.5%	15,708	4,103	-11,605	97,206

Note: The sample is teachers with 22-28 years experience in 1998. The "takers" are teachers who purchased the 2.2 upgrade by 2019. The "non-takers" are those who did not purchase by 2019. Teachers with at least 22 years experience in 1998 and still working in 2019 are considered non-takers. We use the same nominal rate of 5.1% as Fitzpatrick (2015). This along with the cost of living adjustment of 3% implies the real discount factor of 2.1%. The benefit in Column (4) is based on the actual annuity at retirement. The estimated benefit used for Column (6) is computed from forecast of retirement timing for each teacher based on historical data observed before 1998, and assuming teachers paid for the upgrade cost in 1998.

Table 2: Mean and Median Age and Experience in 1998 and at Retirement

	1998				at claim			
	mean age	median age	mean exp	median exp	mean age	median age	mean exp	median exp
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
All Teachers	49.25	49.0	25.09	25.0	57.11	56.0	32.95	33.0
–taker	49.40	49.0	25.14	25.0	56.82	56.0	32.53	33.0
–non-taker	48.23	48.0	24.76	25.0	59.38	60.0	35.75	37.0
Female	49.30	49.0	24.97	25.0	57.26	56.0	32.93	33.0
–taker	49.47	49.0	25.01	25.0	56.98	56.0	32.50	33.0
–non-taker	48.21	48.0	24.74	25.0	59.25	60.0	35.63	37.0
Male	49.17	49.0	25.32	25.0	56.84	56.0	32.98	33.0
–taker	49.28	49.0	25.39	26.0	56.52	56.0	32.59	33.0
–non-taker	48.26	48.0	24.80	25.0	59.66	60.0	36.02	37.0

Note: Teachers with 22-28 years experience in 1998. Experience is the cumulative total service credits based on data provided by TRS.

Table 3: Placebo Tests

	OLS	Cross-district IV		
	(1)	(2)	(3)	(4)
(i) $B_i^* = Estimated\ Benefit$ **(Fitzpatrick)**				
Price	0.010***	-0.082***	-0.071***	-0.070***
	(0.001)	(0.007)	(0.008)	(0.008)
Benefit	0.001***	0.014***	0.012***	0.012***
	(0.000)	(0.001)	(0.001)	(0.001)
Ratio of coefficients	0.142	-0.174	-0.174	-0.174
(ii) $B_i^* = 0$ **for non-takers**				
Price	-0.029***	-0.036***	-0.033***	-0.034***
	(0.001)	(0.005)	(0.006)	(0.006)
Benefit	0.007***	0.007***	0.006***	0.007***
	(0.000)	(0.001)	(0.001)	(0.001)
Ratio of coefficients	-0.240	-0.192	-0.195	-0.193
(iii) $B_i^* = P_i$ **for non-takers**				
Price	-0.032***	-0.043***	-0.038***	-0.040***
	(0.001)	(0.006)	(0.007)	(0.007)
Benefit	0.007***	0.008***	0.007***	0.007***
	(0.000)	(0.001)	(0.001)	(0.001)
Ratio of coefficients	-0.231	-0.185	-0.188	-0.186
(iv) $B_i^* \sim Unif(0, P_i)$ **for non-takers**				
Price	-0.030***	-0.040***	-0.035***	-0.037***
	(0.001)	(0.005)	(0.006)	(0.006)
Benefit	0.007***	0.007***	0.007***	0.007***
	(0.000)	(0.001)	(0.001)	(0.001)
Ratio of coefficients	-0.237	-0.189	-0.191	-0.189
County fixed effect	X	X	X	X
District characteristics	X		X	X
Experience × Age fixed effects				X
Experience × Age fixed effects × District characteristics				X

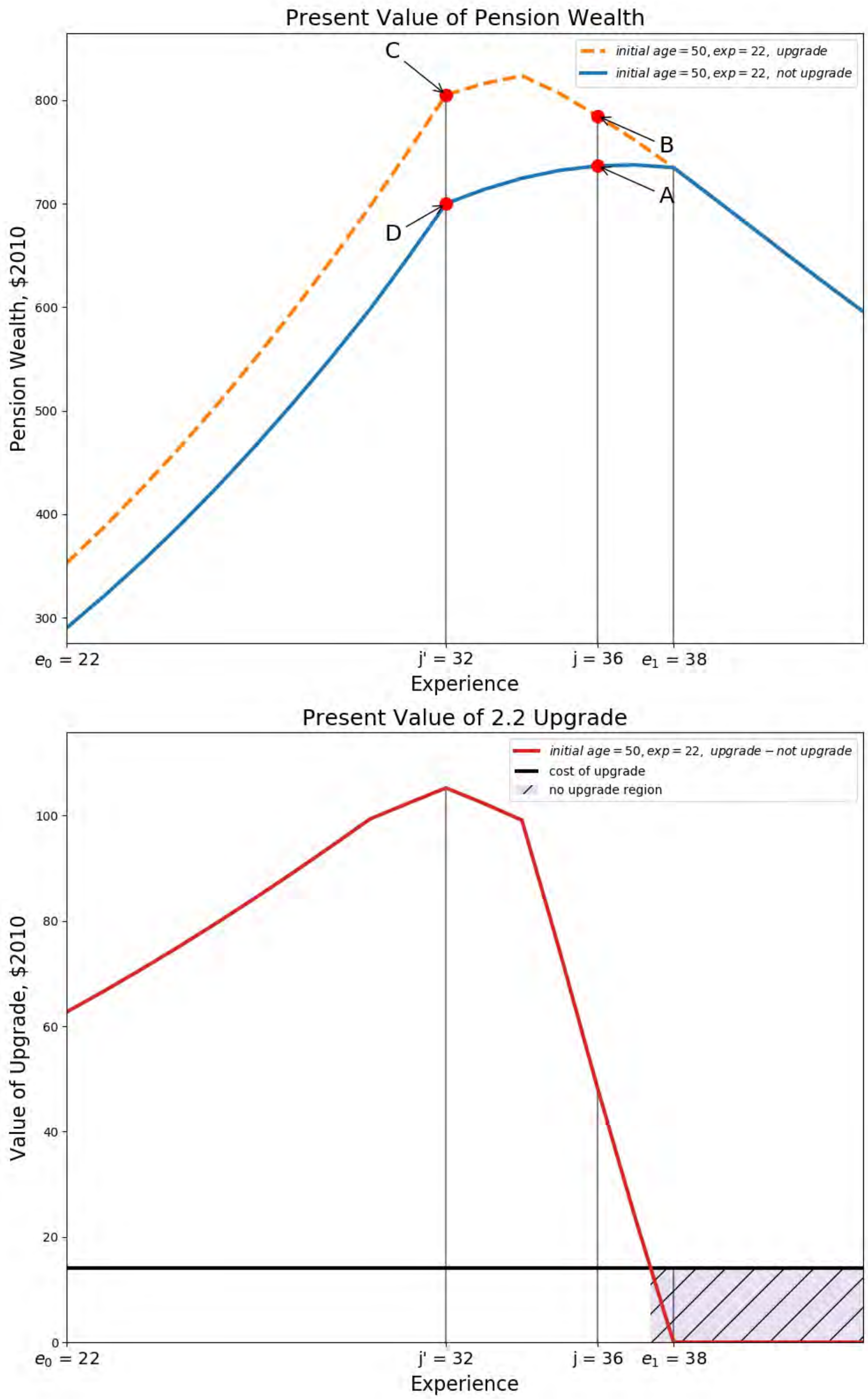
Note: "Ratio of coefficients"= coefficient of benefit /coefficient of price, i.e., β_2/β_1 in model (1). Only the results in panel A of Fitzpatrick (2015)'s Table 2, i.e., using beginning salary for BA holders as IV, are reported here. Additional results can be found in the appendix.

Table 4: Estimates of Demand for Upgrade

		Fitzpatrick (2015)		Our data		Simulated, w/o sample selection		Simulated, w/ sample selection	
	coef	OLS	IV	OLS	IV	OLS	IV	OLS	IV
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimated Annuity	price	0.010***	-0.082***	0.006	-0.100***	-0.061***	-0.109***	0.097***	-0.107***
		(0.001)	(0.007)	(0.005)	(0.032)	(0.004)	(0.009)	(0.006)	(0.020)
	benefit	0.001***	0.014***	0.0010	0.015***	0.016***	0.016***	0.013***	0.015***
		(0.000)	(0.000)	(0.001)	(0.004)	(0.001)	(0.001)	(0.001)	(0.001)
	ratio of coef	0.142	-0.174	0.128	-0.152	-0.256	-0.146	0.131	-0.137
Actual Annuity	price			-0.009***	-0.055***	-0.067***	-0.079***	-0.167***	-0.084***
				(0.002)	(0.020)	(0.001)	(0.004)	(0.002)	(0.008)
	benefit			0.003***	0.008***	0.012***	0.010***	0.015***	0.010***
				(0.000)	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)
	ratio of coef			-0.381	-0.138	-0.182	-0.126	-0.088	-0.117

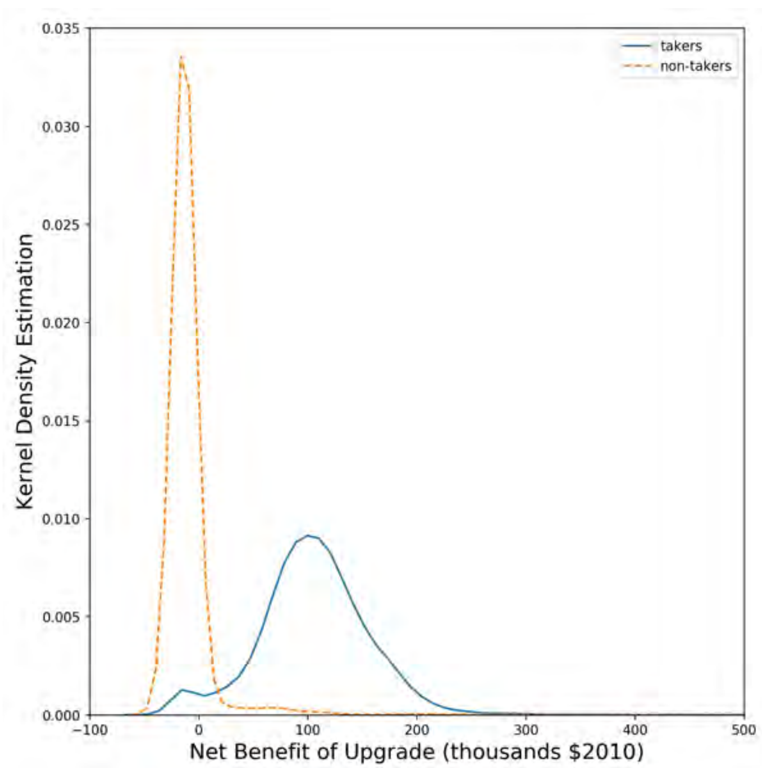
Note: Each pair of columns corresponds to the first and second column of panel A in Fitzpatrick (2015)'s Table 2. Columns (1) and (2) are based on Fitzpatrick (2015)'s data obtained from the AEJ website. Columns (3) and (4) are based on our data summarized in Table 1. In the panel of Estimated (Actual) Annuity, the benefit of the pension upgrade is measured through predicted (observed) timing of retirement for each teacher. Columns (5) to (8) are based on data simulated from a model of optimal decision on retirement and upgrade under the TRS pension rules. The instruments for the IV regressions are the district beginning salary for BA holders and expected upgrade benefit generated with district maximum salary for MA holders. Besides controlling for teacher characteristics such age, experience, highest degree, position, we also control for the county fixed effects and district characteristics in the OLS regressions and the county fixed effects in the IV regressions. In regressions using data simulated from the model of optimal retirement and upgrade, we control for age and experience. In the optimal decision model teachers differ in unobserved preference error and effort, and correlation between the preference error and effort creates endogeneity. In simulated data without (with) sample selection, the preference error and effort are uncorrelated (correlated) by construction.
"ratio of coef" = coefficient of benefit/coefficient of price, β_2/β_1.

Figure 1: Present Value of Pension Wealth With and Without Upgrade



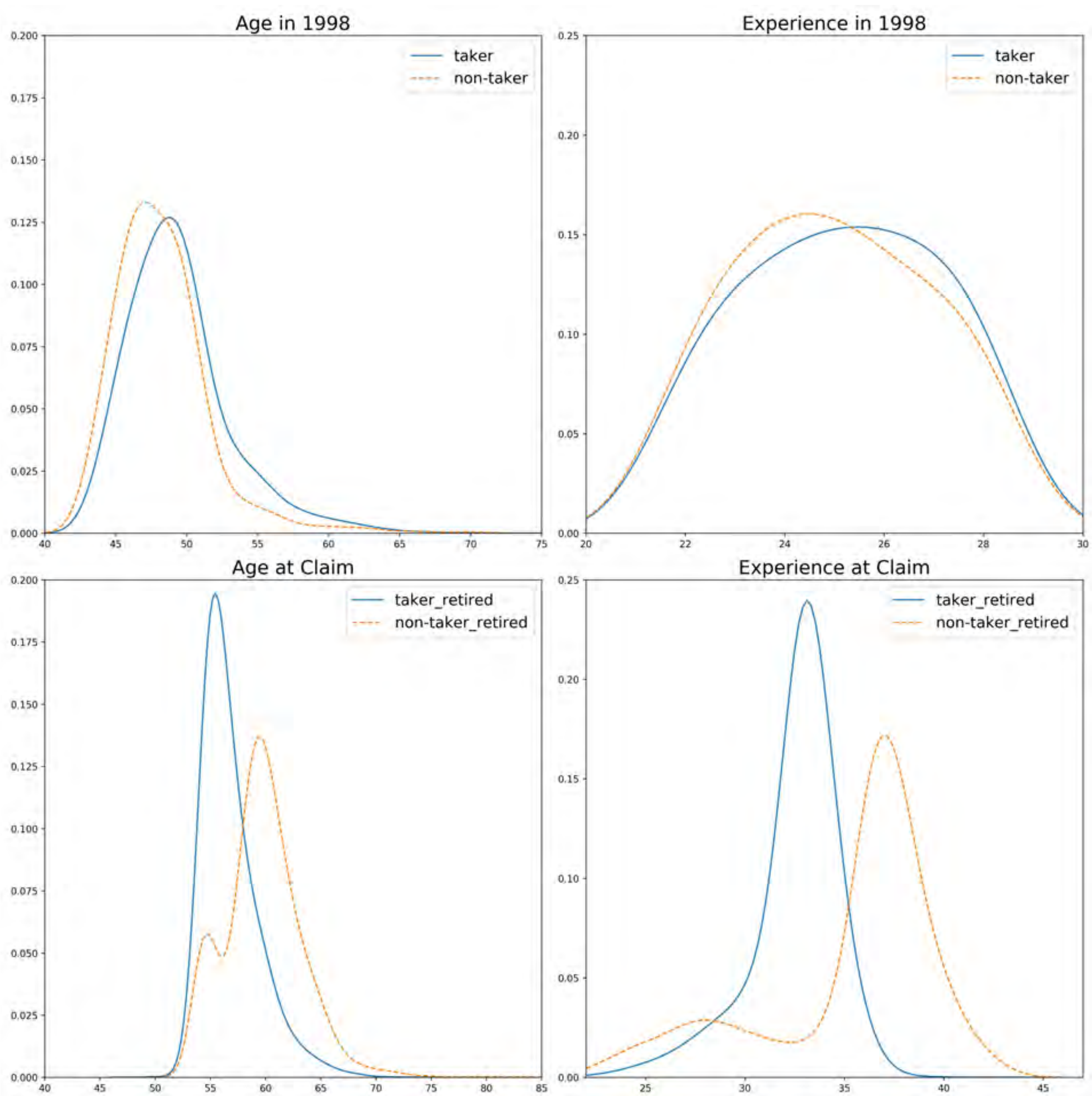
Note: In the upper panel, the solid line plots pension wealth without the upgrade and the dashed line with it. The bottom panel plots the present value of the gain in pension wealth from the upgrade. The horizontal line in the bottom panel is the cost of the upgrade. Pension wealth and retirement decision are based on the TRS pension rules. The graph plots the scenario that a teacher who would have retired at 36 years under the old rules (with pension wealth indicated by point A) chooses to retire at 32 years and take the upgrade (with pension wealth indicated by point C). Benefits and costs are in thousands of 2010 dollar.

Figure 2: Distribution of Net Benefit from Upgrade for Takers and Non-Takers



Note: Kernel density smoothing with bandwidth = 5.

Figure 3: Age and Experience Distribution of Takers and Non-Takers in 1998 and at Retirement Claim Date



Note: Kernel density smoothing with bandwidth = 1.

Online Appendix to How Teachers Value Pension Wealth: A Reexamination of the Illinois Experience

The appendix includes Illinois pension rules and the upgrade option, data construction and supplemental tables for robustness checks, replication and comparison with data used in Fitzpatrick (2015), more details in estimation of the linear probability model and simulation from a model of optimal choice of pension upgrade and retirement.

Contents

A Illinois TRS Pension Rules and the 2.2 Pension Upgrade **2**

B Data Description and Sample Construction **6**

C Tables for Regressions **22**

D Simulating from a Model of Optimal Choice of Retirement and Pension Upgrade **29**

A Illinois TRS Pension Rules and the 2.2 Pension Upgrade

Before discussing the pension upgrade in 1998, we first summarize the defined benefit pension rules of public school teachers in Illinois. Illinois public school teachers are enrolled in the Teacher Retirement System (TRS) (except teachers in Chicago, who are enrolled in Chicago Teachers' Pension Fund, CTPF). The TRS teachers are classified into several tiers of pension plans. Before September 2018, there were two tiers: Tier 1 (2) are those first contributed to TRS before (after) Jan. 1, 2011 or (and) have (no) pre-existing creditable service with a reciprocal pension system prior to Jan. 1, 2011. Tier 1 and Tier 2 members are covered by defined benefit plans of different parameters. In September 2018, a new group, Tier 3, was created. Tier 3 is a hybrid of defined benefit and defined contribution plans.

Teachers can raise the replacement factor for their service years prior to July 1, 1998 to 2.2% by paying a cost. We call it the "2.2 upgrade".

The 2.2 upgrade in 1998 only concerned Tier 1 members. Pensions of Tier 1 teachers experienced several waves of pension enhancements. The 2.2 upgrade was one of them.

A.1 Eligibility for Pension

During the time of employment, teachers contribute a portion of their salary to the pension system, matched by the employer. Eligibility for pension is based on combinations of age-experience (accumulated service credit) requirement:

Table A1: Retirement Eligibility

Year of Service	Age	Description
5	62	Normal Retirement
10	60	Normal Retirement
20	55	At reduced rate: 6% for each year under 60; or under ERO*
35	55**	Normal Retirement

Note: * ERO stands for Early Retirement Options, which will be discussed later in the Appendix.
**If the retirement annuity is at least 74.6 percent of the final average salary and the teacher will reach age 55 between July 1 and Dec. 31, TRS considers him/her to have attained age 55 on the preceding June 1. Moreover, if a teacher meets some criteria of the state of Illinois, he/she can apply for rule of 85.
Source: TRS (2018a), TRS (2018b)

Vested members who do not satisfy the above requirements may retire under actuarial annuity calculation. The overall retirement qualification requirements are plotted in Figure A1.

The old and new formula after the 2.2 upgrade is depicted in Fitzpatrick (2015), which we redraw here for convenience in Figure A2.

A.1.1 Details on service credit

According to Chapter 5 of TRS (2018a), credible service years include regular service, sabbatical leave, sick leave, optional service and reciprocal service, among others. Although regular service makes up the majority of total service credit, sick leave and optional service play big role at the margin, especially when teachers make 2.2 upgrade decisions (see the table in Figure A3, extracted from TRS (2007), for the magnitude of these two items.) For example, a teacher with 36 years of regular service, 1 year of unused sick leave and 1 year of optional service, (i.e., 38 years of total service) does not need the upgrade since the old formula already gives the teacher max benefit. Neglecting the 2 additional service years of sick leave and optional service would result in a substantially higher estimate of the upgrade benefit, given the rules of the upgrade discussed later.

Figure A1: Retirement Eligibility by age and service years

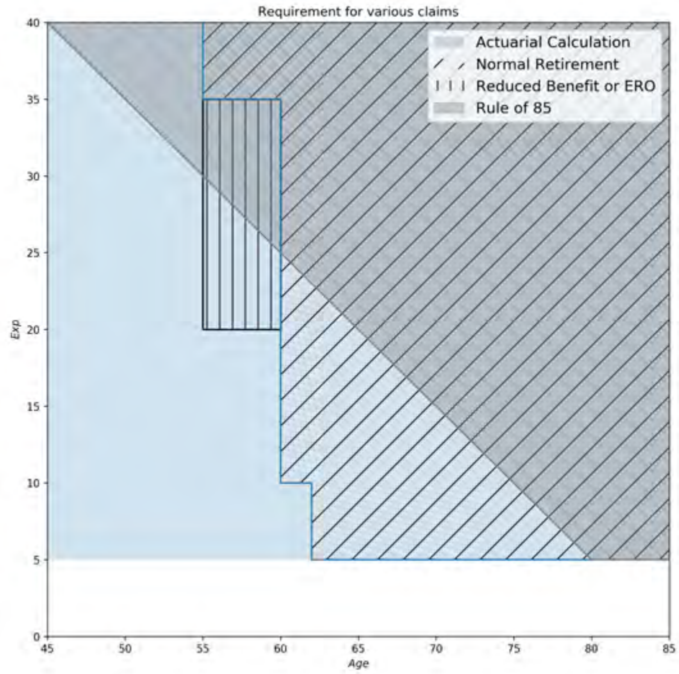
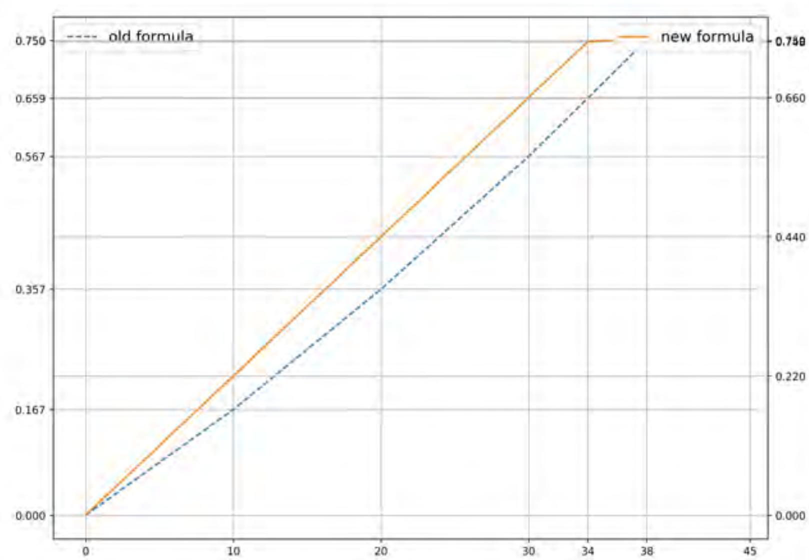


Figure A2: Old and New formula: pension benefit as the fraction of the final average salary by service years at retirement



Note: The dashed line is the old formula. The solid line is the new formula after the 2.2 upgrade.

Figure A3: Credible service years during fiscal year 2002-2006 in TRS (2007)

Page 64

TABLE 9

SUMMARY OF OPTIONAL AND SICK LEAVE SERVICE CREDIT AMONG ACTIVE SERVICE RETIREES WHO RETIRED WITH AN ANNUITY

MALES AND FEMALES
FY 2002 – FY 2006

SERVICE CREDIT AT RETIREMENT

Type of Service	Number of Retirees	Years of Service	Average Over All Retirees	As a Percent of Regular
Regular	21,985	652,738	29.690	100.00%
Optional Service	7,953	20,350	0.926	3.12%
Sick Leave	20,989	24,018	1.092	3.68%
Total	21,985	697,105	31.708	106.80%

AVERAGE AMOUNTS OF DIFFERENT TYPES OF SERVICE

Type of Service	Total Years of Service at Retirement					
	Under 20	20-24.999	25-29.999	30-33.999	34 or more	Total
Regular	13.309	20.085	24.882	29.626	33.447	29.690
Optional Service	0.616	1.440	1.720	1.359	0.613	0.926
Sick Leave	0.643	0.881	1.012	1.094	1.177	1.092
Total	14.569	22.406	27.614	32.078	35.237	31.708

AVERAGES BY FISCAL YEAR OF RETIREMENT

Fiscal Year	Optional Service	Sick Leave	Total
2002	1.028	0.840	1.868
2003	0.880	1.030	1.910
2004	0.929	1.106	2.035
2005	0.913	1.212	2.125
2006	0.866	1.269	2.135
Average	0.926	1.092	2.018

TEACHERS' RETIREMENT SYSTEM OF THE STATE OF ILLINOIS

BUCK CONSULTANTS, LLC

A.1.2 Restrictions on retirement benefit

TRS has several restrictions in the benefit formula:

1) The retirement benefit is capped at 75% of the final average salary. However, if the factor is greater than 74.6%, TRS will round it to 75%. [1]

2) There is a 20% cap on salary increases in calculation of final average salary: annual salary rate after June 30, 1979 cannot exceed the previous year's full-time rate by more than 20 percent. However, this applies only to salaries earned from the same employer.

Finally, for Tier 1 members, the retirement benefit increases at an annual compound rate of 3%.

A.1.3 Early Retirement Options (ERO) and Early Retirement Initiatives (ERI)

There are two TRS programs encouraging early retirement. Additional contributions were required from the member and employer to participate in these programs.

[1] For the actuarial calculation this cap does not apply. Thus for those with exceptionally long careers, actuarial calculation may yield a higher annuity than one based on service credit.

Early Retirement Options (ERO) was enacted on July 1, 1979 and was available until 2016. Under ERO a member between the ages of 55 and 60 with 20 or more years of service credit could retire without an early retirement reduction. (see TRS (2018a).)

Early Retirement Initiatives (ERI) allowed employees to buy an additional five years of age and service credit (see Fitzpatrick and Lovenheim (2014).) ERI was only available in 1992-1994. Because the 2.2 upgrade took place in 1998, ERI has no direct impact on teachers' decision on the upgrade.

A.1.4 Impact of pension rules on the calculation of expected benefit of the upgrade

In Fitzpatrick (2015) the benefit of the 2.2 upgrade is estimated by using retirement probability. In the main text we argue that the estimated benefit does not capture idiosyncratic factors that influence decisions on retirement and upgrade.

In addition to the omission of the unobserved heterogeneity, there are several problems in calculating the estimated benefit. The retirement probabilities are estimated from the historical data for 1988-1997 and may be a biased estimate of the retirement probabilities after 1998 due to pension rule changes:

i) The introduction of 2.2 upgrade policy itself results in changes in the timing of retirement.

ii) The new ERO policy since 2005 could also influence teacher's retirement decision (note that teachers retiring under ERO typically have less experience and hence the 2.2 upgrade is more valuable to them).

iii) The new ERO policy allows employers to limit the number of ERO retirements to 10 percent. Those prefer to retire early but cannot retire under ERO may choose the actuarial calculation instead of reduced benefit. The actual benefits of upgrading may be zero for those teachers.

iv) The ERI policy only existed in 1992-1994. Retirement probability calculated that covered this period may not be the same as that after 1998.

A.2 Details on the 2.2 upgrade

A.2.1 Cost and benefit of the upgrade

Teachers can upgrade their service year prior to 1998 by paying a price:

$$Price_{it} = \min\left(\frac{Exp_{1998}}{100}, \frac{20}{100}\right) \times Salary_{it}$$

where Exp_{1998} is the service year earned before 1998, $Salary_{it}$ is the highest salary rate during the four school years before one opt for the upgrade. Typically, this is the salary of the teacher at the time upgrade purchase. It means that the teacher pays 1% for each year of service before 1998, capped by 20%. [2]

The main effect of the 2.2 upgrade is to help teachers reach the 75% cap sooner. Under the old formula, it took 38 years of service to reach the cap. Under the new formula and the upgrade, it only takes 34 years (see Figure A2). On the other hand, the 2.2 upgrade has no value for at least three groups of teachers:

i) those are not eligible for normal, reduced normal or ERO retirement plans;

ii) those who already reached the 75% cap under the old formula;

iii) those who can collect more annuity under the actuarial calculation.

A.2.2 Reduction, refund, and restrictions on the upgrade

TRS imposed some restrictions on the upgrade (see, e.g., TRS (2019)).

i) On the amount of upgrade: a member who chooses to upgrade must upgrade all the service years prior to 1998 under the four-step formula, but the maximum charge is 20 percent, or 20 years, times the 1 percent contribution rate.

The upgrade charge can be refunded or reduced under some conditions:

ia) One may receive a refund with interest or a reduction of the 2.2 upgrade cost: For every three years with earned regular service credit after July 1, 1998, the equivalent of one year of the 2.2 upgrade cost will

[2] There is a special term for members with at least 24 years of service credit as of July 1, 1998: a member who did not take the upgrade and had at least 24 years of service prior to July 1998 would receive 2.2 percent for each year of service earned after June 1998 up to 30 years and then 2.3 percent for each year of service over 30 years.

be reduced or refunded with interest. For our sample of senior teachers, it is 1% of salary, or 1/20 of the total upgrade price.

ib) Teachers who retire with more than 34 years of service credit may receive a 25 percent reduction in upgrade costs for each year of creditable service beyond 34 years, up to a maximum of 100 percent. Partial years are prorated.

Table A2: Reduction in upgrade cost

Years of service after 98	Reduction	Years of service at retirement	Reduction
3	5%	34	0%
6	10%	35	25%
9	15%	36	50%
12	20%	37	75%
15	25%	38	100%

Note: Years of service are examples. For senior teachers, the cost is 0.2 times salary, and the reduction is as percent of cost instead of salary.

ii) The timing of upgrade: one can only upgrade before collecting any retirement benefit. In another words, "The 2.2 benefit formula does not apply to retired TRS members because they are receiving a retirement benefit."

iii) The frequency of upgrade: "You may make an election to upgrade only once during a five-year period. The election will remain in effect for five years from August 15 following your election to upgrade."

iv) The payment of upgrade can be spread out over at most 5 years according to Fitzpatrick (2015) (page 184). The maximum payment period is 24 months if paid by a reduction in retirement benefits, according to TRS (2019).

B Data Description and Sample Construction

Our main results on take-ups of the upgrade and the cost-benefit analysis are based solely on the TRS data, rather than the merged data of TSR and TRS as in Fitzpatrick (2015). Since TRS is responsible for recording creditable years of service and annual earnings for determining pension benefits as well as the 2.2 upgrade, analysis based on TRS records is appropriate. The availability of teacher payroll data by TRS makes it possible to calculate the years of service at any time, and in 1998 in particular. We discuss merging of TRS and TSR data in Section B.7 for replication of the sample statistics reported by Fitzpatrick (2015); and in Section C.1 for regression analysis, where we need variables such as highest degrees, teaching positions for controls, and district code to construct cross-district IVs.

B.1 Data sources

The data used for this study consist of five data sets from three sources:

a) the Teachers' Retirement System (TRS) data for teachers active in 1980 to 2014

We obtain the data for teachers who are on record at least once from 1980 to 2014. We have the TRS member's name, gender, age (at separation if retired, or current age otherwise), choice of the 2.2 upgrade, first day of work, last day of work, final cumulative service credit, and (if retired) retirement claim date and claim type. Moreover, for each job the member took, we have the payroll data recording the fiscal year, employer name, contract days, paid days, service credit, salary and earnings. Because the Illinois public system allows fractional service credit, one can earn

$$\min\{\frac{\sum \text{paid days}}{\text{contract days}}, 1\}.$$

The payroll data contain some records before 1980.

b) the Teacher Service Record (TSR) data for teachers active in 1980 to 2014

Table B1: Data sources

Data	Description	Source	Time	Note
a)	teacher information in TRS	TRS	1980 to 2014	cross section
	teacher payrolls	TRS	up to 2014	panel
b)	teacher service records (TSR)	ISBE	up to 2014	panel
c)	teacher pension	TRS	2015 to 2019	panel
	2.2 upgrade information	TRS	1998 to 2019	cross section
d)	Fitzpatrick (2015) data	Fitzpatrick (2015)	as of 2009	cross section
e)	district salary table	ISBE	2010-2019	panel

Note: Teacher information in TRS includes all teachers on record between 1980-2014. Teacher payrolls and service records start their first year of service (which can be earlier than 1980). Some records were created before the electronic system was built.

The TSR is obtained from Illinois State Board of Education (ISBE). It also contains name, gender, salary, district experience, state experience and first date of work for public school employees. Moreover, it has the position code, district code, and details of teaching assignment (e.g., subject and grade information) as well as highest degrees, etc. However, there is no unique identifiers between TRS and TSR, hence we need to match the records in the two systems. Fitzpatrick (2015) uses fuzzy matching based on member name, employer name, and "recorded service accrued" in both systems in 2009, and reports a 97% matching rate. We use member name and matched 99% of TSR members in the 2014 data. Our primary reason for matching the TRS and TSR data is to use the teacher and district data in TSR for regression analysis. Data a) and b) were requested in Feb, 2014. However, they do not include the pension benefits (although retirement claim types are recorded).

c) the TRS data on annual pension benefits and 2.2 upgrade cost up to 2019

We requested the amount of pension and the cost of 2.2 upgrade for teachers active in 1980 to 2019. We requested the data in Feb 2019, and the requested data came in several batches until June 2019. In this newly requested TRS data, the purchase date of 2.2 upgrade is "final date of purchase" which may be updated when one took refund at the time of retirement. Because there is no unique identifier of teachers in the two waves of our requested TRS data and we match them based on personal characteristics including name, first year of work, last year of work, etc.

d) the data used by Fitzpatrick (2015)

We also obtained the data and program used by Fitzpatrick (2015) from the AEJ website. Her sample period ends in 2009. We run her code on her data and obtain the same results as in her paper (with a few exceptions detailed in Section B.6). However, her data do not contain teacher name or sufficient information of other variables to allow an adequate match of her data with ours.

e) district salary tables

To reconstruct the cross-district IV in Fitzpatrick (2015), we also obtain the district salary table from the ISBE website. However, spreadsheets of salary data now went back only as far as 2010-2011, which we use instead of the 1998 salary table used in Fitzpatrick (2015).

B.2 Construct a matched sample between two TRS datasets

To calculate the cost and benefit of the 2.2 upgrade, we only need two data sets from TRS: a) and c). First, we construct the sample of senior teachers (those with 22-28 years of service in 1998) from TRS, similar to Fitzpatrick (2015). Our sample size differs slightly from hers: she has a sample of 19,429 teachers, we have 19,126. The sample size depends on the selection criteria in Table B2. Moreover, a stricter standard yields higher quality of the merged data set, which is crucial for the precision of our calculation of upgrade cost and benefit. Hence we apply the most restrictive Selection Criterion 3 and focus on the sample with 19,126 observations.

Table B2: Sample selection and merging two TRS datasets

Selection Criterion	1	2	3
experience of service 22-28 in 1998	Yes	Yes	Yes
work full time in 1998	Yes	Yes	Yes
match on name	Yes	Yes	Yes
—— first year of work	No	Yes	Yes
—— last year of work	No	No	Yes
—— total service credit at retirement	No	No	Yes
retire by 2014	No	Yes	No
observation	21,261	19,450	19,126

Note: Experience in 1998 is calculated as the cumulative sum of service credit by 1998, not including sick leaves, optional credit, etc. We require exact match for name and allow fuzzy match on other variables. Work full time in 1998 means the total service credit earned in 1998 should be 1.

B.3 Detailed procedures

step 1 selection based on data set a) TRS data up to 2014

With the TRS payroll data, we first obtain those who had worked in the 1997-1998 school year (by selecting those who have earned at least 0.006 service credit, i.e., one work day in 1997-1998), and the service credits in every fiscal year. TRS provides the final years of service but not the cumulative years of service by 1998. We calculate the latter by adding the annual service credits. One may earn fractional service credit from multiple employers in the same year, we first add all the service credit for one year (if it is greater than 1, we record it as 1), keep those with 1 service credit for 1997-1998, and then calculate the cumulative sum up to that time.

There are two potential issues: first, we may miss those working part time in 1997-1998. By the policy rules, they are eligible for the upgrade as long as they have not collected the retirement benefit. Second, since the service credit can be fractional, we round them to integers: so 22 years of experience may include teachers with in fact 21.50-22.49 years of experience. For robustness check against the first issue we include teachers who work part time in 1998. To address the second issue, we first select those with 21-29 years of service at the end of 1997-1998 school year and round their service credit to integers. Then we select those with 22, 23,..., 28 years of service credit at 1998 to create an alternative sample.

Accuracy of our service credit calculation

To verify that our calculation of the service credit is consistent with that of TRS, we compare the final service credit calculated using our approach with the total service credit reported by TRS. We find that the mean error is -0.003, indicating our service credit calculation is quite accurate. This suggests our calculation of the service credit in 1998 is likely accurate.

step 2 merge with data set c) TRS data on pension and upgrade up to 2019

We follow Selection Criterion 3 and match on name, first year of work, last year of work, total service credit and select those retired by 2014.

step 3 descriptive statistical analysis

With this sample, we can track teachers to 2019 and record their retirement claim types, retirement date, upgrade decisions, and the date of upgrade purchase. We divide the sample into takers and non-takers and examine the difference in age and experience across the subgroups. Moreover, we further decompose the non-taker group by their retirement claim types and repeat the comparison in age and experience. The results are reported in Tables 1-2 and Figures 2-3 in the text and various tables in the Appendix.

B.4 Robustness check of sample construction

Before describing how we calculate the upgrade cost and benefit in the next section, we discuss some alternative criteria for data matching and procedures for robustness. Here, we only report the results for replication of Table 1 in the text.

B.4.1 observations with exact years of service of 22-28

As noted earlier, Illinois allows fractional years of service and we round the years of service in 1997-1998 to integers, which may cause measurement errors. We build another restrictive sample with years of service exactly equal to 22, 23, ..., 28 in 1997-1998. That sample has 11,114 observations and we replicate the analysis in the main context on this data set. The result for this robustness check is reported in Table B3.

B.4.2 teachers work part time in 1997-98

As we noted in step 1 of Section B.3, some teachers may have worked part time in school year 1997-1998 and we excluded them from our sample. Here we include them in the sample and repeat the analysis. The result for this robustness check is reported in Table B4.

B.4.3 alternative selection criteria

We use several alternative selection criteria and repeat our quantitative analysis for samples obtained from those criteria:

i) Instead of using the cumulative service credit, we assume that one year equals one service credit, no matter how many service credit one actually earns. The result is in Table B5.

ii) We include all teachers with 20-29 years of service at the end of 1997-98 school year and report the result in Table B6.

Table B3: Replicating Table 1 for teachers with exact years of service in 22, 23, ..., 28 at 1998

type	number	percent	average cost	average benefit	average net benefit	median net benefit
taker	9530	85.75%	15.81	122.46	106.65	105.37
non-taker	1584	14.25%	14.34	4.13	-10.21	-12.96
non-taker-retired	1351	12.16%	14.28	4.84	-9.44	-12.88
whole sample	11114	100.0%	15.60	105.59	89.99	95.30

Table B4: Replicating Table 1 by including for teachers working part time in 1998

type	number	percent	average cost	average benefit	average net benefit	median net benefit
taker	16733	86.78%	15.23	121.57	106.33	104.74
non-taker	2550	13.22%	13.80	4.38	-9.43	-12.67
non-taker-retired	2175	11.28%	13.76	5.13	-8.63	-12.56
whole sample	19283	100.0%	15.04	106.07	91.03	96.59

Table B5: Replicating Table 1 by restricting one year = one service credit

type	number	percent	average cost	average benefit	average net benefit	median net benefit
taker	17235	86.57%	15.00	121.04	106.03	104.13
non-taker	2673	13.43%	13.32	4.40	-8.92	-12.44
non-taker-retired	2258	11.34%	13.25	5.21	-8.04	-12.33
whole sample	19908	100.0%	14.78	105.37	90.60	95.90

Table B6: Replicating Table 1 for teachers with 20-29 years of service at 1998

type	number	percent	average cost	average benefit	average net benefit	median net benefit
taker	23364	86.24%	15.11	120.15	105.04	103.03
non-taker	3727	13.76%	13.66	4.49	-9.18	-12.61
non-taker-retired	2941	10.86%	13.67	5.68	-7.99	-12.48
whole sample	27091	100.0%	14.91	104.24	89.33	94.48

The robustness checks confirm two patterns reported in Table 1 in the text. First, the overall take-up rate is around 85%. Second, the average net benefit of upgrade is negative for non-takers and positive and large for takers.

B.5 Calculating cost and benefit of 2.2 upgrade

B.5.1 Benefit of upgrade

We contrast our calculation of the benefit of the upgrade with Fitzpatrick (2015). We use the actual benefit after observing the retirement decision, Fitzpatrick (2015) uses an estimated benefit based on retirement probability.

B.5.2 Calculation of the "actual" benefit of the upgrade

$$B = \sum_{i=0}^{T-a-i} \frac{1.03^i \pi(a+i|a)\Delta B_i}{1.051^i} \qquad (1)$$

step 1 obtain ΔB_i

We use the observed actual annual retirement benefit for 2018 to recover the actual benefit at the time of retirement claim. For takers, we calculate the benefit they would receive at the time of retirement claim without the upgrade and subtract it from the actual benefit, to obtain the benefit of upgrade, ΔB. For non-takers, we calculate the benefit they would receive at the time of retirement claim with the upgrade and subtract it with the actual benefit. Note that at retirement claim date, the final average salary as well as the total service credit are determined, hence our calculation of the potential retirement benefit without the upgrade is straightforward.

step 2 compute the present value weighted by mortality and the real discount rate

We use gender-specific mortality rate and the real discount rate of 1.03/1.051 to calculate the discounted expected value of the upgrade. T is the maximum lifespan (set at 101 years), a is age at the year of retirement $\pi(a+i|a)$ is the probability of surviving to age $a+i$ for a teacher of age a.

step 3 convert to 2010 dollars

Finally, we convert the value at the time of retirement claim to 2010 dollars by CPI inflation.

B.5.3 calculation of "estimated" benefit of the upgrade

Following Fitzpatrick (2015) we calculate estimated benefit weighted by retirement probability at year j based on the following formula:

$$B^* = \sum_{j=0}^{T-a-j} \frac{1.03^j \pi(a+j|a)(\Delta B_j)R_j}{1.051^j} \qquad (2)$$

step 0 estimate the wage schedule and retirement probability

The pension benefit depends on the salary at the time of retirement. Thus the benefit of the upgrade for a teacher who will retire in the future depends on future salary. Teacher salary is highly predictable by experience. Here we use a regression to estimate salary schedule. We obtain a cross section of teacher salary in different years by experience and regress the logarithm of salary on a cubic function of experience. The salary at year j is fitted value for $e + j$ multiplied 1.03^j to adjust for inflation.

The retirement probability R_j across time for a teacher with initial age-exp (a, e) in 1997-98 is approximated by teachers with the same age and experience at an earlier period (for example, late 1980's to early 90's). We track the teachers in the earlier period over time and record their retirement years. We use the frequency of retiring in year j as the estimate of retirement probability in year j.

step 1 obtain ΔB_j for each $j = 0, ..., T - a$

Now we can calculate the pension benefits (with or without upgrade) if one retires at year j, and the additional pension benefit from the upgrade, with the estimated salary schedule.

steps 2 and 3 are similar to those in subsection B.5.2

step 4 weight by retirement probability

Finally, we weight the expected discounted benefit by retirement probability for all feasible years j to obtain the one used in Fitzpatrick (2015).

B.5.4 Cost of the upgrade

The date of purchase in our data set from TRS is the date of final transaction. We were informed by TRS that the record for the first date of purchase is unavailable. We estimate the first date of purchase using data on the final payment of 2.2 upgrade and the number of years of service purchased. Dividing the former by the latter gives us the salary in the year of purchase. We then compare that salary with the payroll data for each teacher to find the closest match, which corresponds to the year of first purchase.

The estimated first purchase is likely earlier than the reported date of the final transaction by TRS, which may be for a later payment or rebate for the upgrade. But the final transaction occurs only when the upgrade was purchased earlier. Hence for a given cohort the cumulative take-up rate for a given year based on the estimated first upgrade purchase is higher than that based on the final transaction date.

The estimated upgrade purchase date may contain errors but this is a minor issue for two reasons. First, there is no ambiguity in "whether" a teacher took the upgrade. Second, our robustness checks show that "when" the upgrade took place is empirically unimportant in the calculation of the cost and benefit of the upgrade.

For robustness check to the purchase date, we calculate the upgrade cost with two extreme timings of the purchase of upgrade. And we convert both costs to 2010 dollars.

cost paid at retirement

We assume the upgrade the cost is paid at the time of retirement. For our sample of senior teachers, it is 20% of the annual salary at the time of retirement minus potential reductions for service credit after 1998 as well as total service credit exceeding 34.

cost paid in 1998

This is 20% of salary in 1997-1998.

Finally, note that the actual cost may be between the two cases and it is possible to spread the payment over a period up to 24 months and deduct it from the retirement benefit. But the cost difference at the different timing of upgrade does not materially change net benefit of the upgrade.

B.5.5 four cases

We have two methods for calculating upgrade cost and and two for benefit, hence four cases in total:

Table B7: Four measures of upgrade cost and benefit calculation

benefit / cost	at 1998	at retirement claim
actual benefit based on formula (1)	Table B8	Table B9
estimated benefit based on formula (2)	Table B10	Table B11

Note: Table B10 and Table B11 use the retirement probability and salary growth based on the 1991 data.

Table B8 presents a finer decomposition of Table 1 in the text. Table B9 uses the cost at retirement claim date for the upgrade cost.

Table B8: Summary Statistics by Types of Teachers, alternative decomposition

type	number	percent	average cost	average benefit	average net benefit	median net benefit
taker	16654	87.08%	15.26	121.55	106.30	104.77
non-taker	2472	12.92%	14.00	4.41	-9.60	-12.76
non-taker-retired	2106	11.01%	13.97	5.17	-8.80	-12.65
non-takers-with-benefit	120	0.63%	13.83	84.60	70.77	49.32
whole sample	19126	100.0%	15.09	106.41	91.32	96.83

Note: The "takers" are those taking the 2.2 upgrade eventually in our sample, up to 2019. The "non-takers" are those not taking it by 2019; "non-takers retired" are those retired by 2014; "non-takers with benefit" is a subgroup of "non takers retired" who were eligible for the upgrade with positive net benefit but failed to upgrade. All values are in $1000 2010 dollars.
Our detailed retirement-related data (including final average salaries and total years of service, which are essential in determining the potential gain of upgrade) are up to 2014. We label active teachers who have not purchased the upgrade by 2014 as non-takers, because for a teacher with at least 22 years of experience in 1998 and worked continuously to 2014 would have a service year of 38 would not benefit from the upgrade.
Average price (cost) is calculated assuming all teachers upgrade in 1998, following Fitzpatrick (2015); the results with price of upgrade at retirement claim date are in Table B9.
The average benefit for takers is the present value of the actual benefit minus the potential benefit without the upgrade; for non-takers, it is potential benefit with upgrade minus the actual benefit. Note that for non-takers, the potential benefit of not upgrading is the maximum of two options: one for old formula and another for actuarial calculation.
We use the same nominal rate of 5.1% as Fitzpatrick (2015). This along with the cost of living adjustment of 3% implies the real discount factor: $\beta = 1.03/1.051 = 0.980$ or $r = 1/\beta - 1 = 0.020$.

Table B9: Summary Statistics by Types of Teachers, with upgrade cost at retirement claim date

type	number	percent	average cost at claim	average benefit	average net benefit	median net benefit
taker	16,654	87.08%	16.55	121.55	105.01	102.73
non-taker	2,472	12.92%	2.84	4.41	1.57	0.00
non-taker-retired	2,106	11.01%	3.33	5.17	1.84	0.00
non-takers-with-benefit	431	2.25%	3.38	25.03	21.65	2.22
whole sample	19,126	100.0%	14.77	106.41	91.64	95.74

Note: The cost and benefit are in $1,000 2010 dollars. Average cost at claim is the payment one has to make for upgrade at the time of retirement claim. It depends on the salary one earned for the last year of work, and takes into account potential reduction and refunds. These reduction and refunds include discount for those with total service credit greater than 34 years at retirement, as well as for years of service after 1998.

Tables B10 and B11 show that using the weighted average calculation of the upgrade benefit, the gap in the average benefit between takers and non-takers shrinks to around $10k, and most of those non-takers have positive net benefit, with the cost at paid in 1998 or at retirement claim. Moreover, the upgrade benefit seems to distributed uniformly across teachers with different years of service at 1998.

Table B12 compares the average actual benefit at different discount rate and the weighted average benefit with retirement probability and salary based on the 1991 data.

Table B10: Summary Statistics by Types of Teachers, estimated benefit, and upgrade cost in 1998

type	number	percent	average cost	average benefit	average net benefit	median net benefit
taker	16,654	87.08%	15.26	117.06	101.80	97.00
non-taker	2,472	12.92%	14.00	107.84	93.84	89.41
non-taker-retired	2,106	11.01%	13.97	106.98	93.01	88.32
non-takers-with-benefit	2,469	12.91%	14.00	107.97	93.97	89.41
whole sample	19,126	100.0%	15.09	115.87	100.77	95.77

Note: All measurements are the same as Table 1 in the text, with two exceptions. Upgrade cost is measured by 20% of the salary in 1998; and upgrade benefit is measured by the weighted average with retirement probabilities and salary growth of the 1991 cohort.

Table B11: Summary Statistics by Types of Teachers, estimated benefit, and upgrade cost at retirement

type	number	percent	average cost at claim	average benefit	average net benefit	median net benefit
taker	16,654	87.08%	16.55	117.06	100.51	95.44
non-taker	2,472	12.92%	2.84	107.84	105.01	99.92
non-taker-retired	2,106	11.01%	3.33	106.98	103.65	98.61
non-takers-with-benefit	2,469	12.91%	2.83	107.97	105.14	99.94
whole sample	19,126	100.0%	14.77	115.87	101.09	96.06

Note: Upgrade cost is based on payment at the time of retirement claim, net of possible refund/reductions. Benefit is the weighted average with retirement probabilities and salary growth both from the 1991 cohort.

Table B12: Actual Benefit and Estimated Benefit

exp at 98	actual benefit	actual benefit 7.1%	actual benefit 9.1%	prob 91 slry 91	number of teachers
22	93.55	74.96	61.98	117.52	2385
23	96.49	77.27	63.85	117.80	2430
24	104.52	83.61	69.04	116.16	2897
25	109.80	87.83	72.53	112.11	3006
26	112.71	90.10	74.36	113.91	2972
27	113.59	90.86	75.03	117.61	2599
28	110.89	88.70	73.24	116.94	2837
whole sample	105.94	84.76	70.00	116.01	19126

Note: Actual benefit based on formula (1) with different discount rates: 5.1%, 7.1%, and 9.1%.
"prob91 slry 91": benefit uses the weighted average (formula (2)) with retirement probabilities and salary growth both from the 1991 cohort.

B.6 Working with the data and code in Fitzpatrick (2015)

Using the data and code provided at the AEJ website we are able to replicate the results of Fitzpatrick (2015) for the most part, with some discrepancies in the calculation of take-up rate. Because her data set does not contain teacher names, we are not able to extend the records of the same cohort by merging her data with our data.

B.6.1 Replicating summary statistics in Table 1 of Fitzpatrick (2015)

The original program from the AEJ website does not reproduce Table 1 in Fitzpatrick (2015). Table B13 is produced with the data and code from the AEJ website. The sample size in Table B13 is 20,350, instead of 19,429 reported in Table 1 in Fitzpatrick (2015).

We managed to replicate her Table 1 by using one criterion for sample selection for the upgrade and retirement probability and another for the cost and benefit of the upgrade. This is reported in Table B14. Table B14 matches Table 1 in Fitzpatrick (2015) very closely with one exception that the retirement ratio for those with 22 years of service at 1998 is 0.51, instead of 0.6 reported in her paper. In Table B13 the retirement ratio for those with 22 years of service at 1998 is also 0.51.

Figure B1: Original Table 1 of Fitzpatrick (2015)

TABLE 1— TAKE-UP RATES, RETIREMENT RATES, PRICES, AND COSTS BY LEVEL OF EXPERIENCE IN 1998

Years of experience in 1998	Fraction who retire by 2009	Fraction who purchase upgrade by 2009	Mean price [$1,000s] (SD)	Mean benefit [$1,000s] (SD)	Mean ratio of benefits to price (SD)	Observations
22	0.6	0.7	14.02	86.06	6.29	2,409
			(4.08)	(28.41)	(2.87)	
23	0.67	0.71	14.61	90.41	6.27	2,589
			(4.35)	(30.62)	(1.99)	
24	0.78	0.76	14.90	93.01	6.27	2,991
			(4.40)	(31.55)	(1.20)	
25	0.83	0.78	15.26	96.94	6.37	3,121
			(4.51)	(33.62)	(1.26)	
26	0.88	0.77	15.40	100.40	6.54	2,909
			(4.47)	(35.39)	(1.43)	
27	0.89	0.74	15.75	102.73	6.53	2,674
			(4.73)	(37.56)	(1.37)	
28	0.91	0.73	15.90	102.17	6.42	2,736
			(4.63)	(36.42)	(1.27)	
Whole sample						19,429

Notes: Based on the author's calculations using data from Illinois TRS and TSR. Years of service in 1998 is the number of creditable years of service the teacher has accrued by 1998. The fraction who retire is the fraction of the teachers with the indicated number of years of experience in 1998 who have begun collecting retirement benefits as of 2009. The fraction who purchased the upgrade is the fraction of teachers with the recorded amount of service who have purchased the upgrade by 2009. The average price of the upgrade is based on the teacher's salary and experience at the time of purchase (and is in thousands of 2010 US dollars). The cost of the upgrade (in thousands of 2010 US dollars) is the present value in 1998 of the extra retirement benefits paid out as of 2009 as explained in the text.

Table B13: Replication of Fitzpatrick (2015) Table 1 with the code from the AEJ website

Years of experience in 1998	Fraction who retire by 2009	Fraction who purchase upgrade by 2009	Mean price [$1,000s] (SD)	Mean benefit [$1,000s] (SD)	Mean ratio of benefits to price (SD)	Observations
22	0.51	0.69	13.79	85.22	6.39	2542
			(4.19)	(28.31)	(2.96)	
23	0.66	0.71	14.42	89.63	6.35	2711
			(4.44)	(30.51)	(2.15)	
24	0.77	0.75	14.73	92.24	6.35	3141
			(4.48)	(31.44)	(1.51)	
25	0.82	0.78	15.08	96.10	6.45	3267
			(4.57)	(33.53)	(1.53)	
26	0.87	0.77	15.21	99.37	6.61	3042
			(4.55)	(35.32)	(1.70)	
27	0.89	0.74	15.58	101.72	6.58	2795
			(4.78)	(37.30)	(1.59)	
28	0.90	0.73	15.73	101.23	6.47	2852
			(4.69)	(36.28)	(1.48)	
Whole sample						20,350

Note: The table is produced with the code and data labeled for Table 1 from the AEJ website.

Table B14: Replication of Fitzpatrick (2015) Table 1 with revised code

Years of experience in 1998	Fraction who retire by 2009	Fraction who purchase upgrade by 2009	Mean price [$1,000s] (SD)	Mean benefit [$1,000s] (SD)	Mean ratio of benefits to price (SD)	Observations
22	0.51	0.70	14.02	86.06	6.29	2409
			(4.08)	(28.41)	(2.87)	
23	0.67	0.72	14.61	90.41	6.27	2589
			(4.35)	(30.62)	(1.99)	
24	0.77	0.76	14.90	93.01	6.27	2991
			(4.40)	(31.55)	(1.20)	
25	0.83	0.79	15.26	96.94	6.37	3121
			(4.51)	(33.62)	(1.26)	
26	0.88	0.77	15.40	100.40	6.54	2909
			(4.47)	(35.39)	(1.43)	
27	0.89	0.75	15.75	102.73	6.53	2674
			(4.73)	(37.56)	(1.37)	
28	0.90	0.73	15.90	102.17	6.42	2736
			(4.63)	(36.42)	(1.27)	
Whole sample						19,429

Note: The table is produced with a revised code that uses a different sample-selection criterion (which is used to generate the sample for the regression analysis in her Table 2) from the AEJ website.

B.6.2 Two measures of takers of upgrade

More importantly, in her programs there are two variables recording the upgrade decisions. She used "takeupgrade" for the summary statistics in her Table 1 and "takeUPmiss" for the regression analysis in her Table 2. There are missing values of "takeupgrade", and for those not missing, "takeupgrade" coincides with "takeUPmiss". There are no missing values in "takeUPmiss".

In Table B15, for "takeupgrade" and "takeUPmiss", "1" means a teacher took the upgrade, and "0" means a teacher did not take the upgrade, "n.a." means missing value for "takeupgrade". Table B15 shows that 1,735 out of the 5,535 (or 31%) non-takers categorized by "takeUPmiss" have missing values in "takeupgrade", and 2400 out 13,894 (about 17%) takers have missing values in "takeupgrade".

Table B16 shows that the missing value of "takeupgrade" is perfectly correlated with the observed retirement behavior. All the teachers with missing "takeupgrade" are not retired.

Table B15: "takeupgrade" and "takeUPmiss"

takeUPmiss	takeupgrade n.a.	0	1	Total
0	1735	3800	0	5535
1	2400	0	11454	13894
Total	4175	3800	11454	19429

Note: The first (second) row decomposes those with "takeUPmiss" being 0 (1) into different values of "takeupgrade". The columns decomposes those with a given value of "takeupgrade" into different values of "takeUpmiss".

Table B16: "takeupgrade" and "retire"

retire	takeupgrade n.a.	0	1	Total
0	4175	0	0	4175
1	0	3800	11454	15254
Total	4175	3800	11454	19429

Note: The first (second) row decomposes those with "retire" being 0 (1) into different values of "takeupgrade". The columns decomposes those with a given value of "takeupgrade" into different values of "retire".

In Table B14, we use the indicator upgrade "takeupgrade" that contains missing values. We redo her Table 1 using "takeUPmiss" as indicator for the upgrade decision. The result is reported in Table B17. The fraction of takers is lower than that in Table B14, as is expected.

Finally, we revisit the question of take-up rate. The take-up rate in Table 1 of her paper (here in Figure B1) is 74.3%. Table B18 reports sample composition based on the count of "takeUPmiss". The take-up rate of 71.5% is lower than that based on "takeupgrade".

B.7 Supplemental tables of cumulative take-ups by year

The data we have for tracking the upgrade decision is up to 2019, however, the retirement claim data we have is up to 2014. Table B19 tracks upgrades for the subgroup of teachers retired by 2014 and for the whole sample.

There are classification errors for non-takers. We remove teachers with five types of retirement claims in the subgroup of "non-takers by 2019" to build the subgroup of "true non-takers by 2019": "Regular 2.2", "2.2 ERO Employer Pay", "2.2 ERO Member Pay", "Rule of 85 - 2.2", and "Regular 2.2 - Disability". These claims are only available for teachers who took the upgrade. There are 322 such teachers. We call the remaining non-takers true non-takers. The eventual take-up rate (by 2019) based on the TRS record given in Table B19 is 1 minus the eventual true non-takers (12.92%), which results in a take-up rate of about 87%.

Table B17: Replication of Fitzpatrick (2015) Table 1 with revised code and measure the upgrade decision by "takeUPmiss"

Years of experience in 1998	Fraction who retire by 2009	Fraction who purchase upgrade by 2009	Mean price [$1,000s] (SD)	Mean benefit [$1,000s] (SD)	Mean ratio of benefits to price (SD)	Observations
22	0.51	0.72	14.02	86.06	6.29	2409
			(4.08)	(28.41)	(2.87)	
23	0.67	0.70	14.61	90.41	6.27	2589
			(4.35)	(30.62)	(1.99)	
24	0.77	0.73	14.90	93.01	6.27	2991
			(4.40)	(31.55)	(1.20)	
25	0.83	0.74	15.26	96.94	6.37	3121
			(4.51)	(33.62)	(1.26)	
26	0.88	0.72	15.40	100.40	6.54	2909
			(4.47)	(35.39)	(1.43)	
27	0.89	0.70	15.75	102.73	6.53	2674
			(4.73)	(37.56)	(1.37)	
28	0.90	0.69	15.90	102.17	6.42	2736
			(4.63)	(36.42)	(1.27)	
Whole sample						19,429

Note: The table is produced with a revised code that uses an alternative sample-selection criterion (which is used to generate the sample for the regression analysis in her Table 2) from the AEJ website. In addition, the indicator for the upgrade decision is "takeUPmiss" instead of "takeupgrade".

Table B18: Recalculated Summary Statistics by Types of Teachers in Fitzpatrick (2015)

type	number	percent of total	average cost	average benefit	average net benefit
taker	13894	71.51%	15.49	101.20	85.71
non-taker	5535	28.49%	14.27	90.86	76.59
non-taker-retired	3800	19.56%	14.41	98.12	83.71
non-takers-with-benefit	5532	28.47%	14.27	90.91	76.63
whole sample	19429	100.00%	15.14	98.26	83.11

Note: The table is produced with the data from the AEJ website but with "takeUPmiss" as the take-up indicator.

Table B20 decomposes the true non-takers by the retirement claim types. Table B21 compares the retirement and upgrade data in Fitzpatrick (2015) and our sample. Table B21 is based on the estimated first upgrade purchase. Fitzpatrick (2015) indicated that her sample period ends in 2009 without giving the exact date. We consider two extreme cases for the end of sample period: one is at the end of 2008 and another at the end of 2009. The estimate of the cumulative take-up rate based on the estimated first date of purchase given in Table B21 is 85%, the same for 2008 and 2009. The cumulative take-up rates based on the final transaction dates supplied by TRS are in Table B22. As noted in Section B.5.4, some of purchases occurred before the final transaction dates. Hence the take-up rates in Table B22 are lower than the cumulative actual rates by the year.

As noted earlier, our calculation of the take-up rate of final transaction date is based on the TRS records and is robust to the end of sample period. The cumulative take-ups by our estimated first purchase date plateaued by 2008. One may ask why the take-up rate calculated by Fitzpatrick (2015) is about 11% lower than ours. [3]

We discussed above the missing value of take-up indicators in the data used by Fitzpatrick (2015) that can affect the take-up rate. But a more likely explanation is that merging TRS with TSR creates a discrepancy in the samples defined by service credits of 22-28 years in 1998. As mentioned earlier, TRS provides the final cumulative service for retired members and current cumulative service for active members. We calculated the cumulative service credit in 1998 using a supplemental data set provided by TRS. The note in Table 1 of Fitzpatrick (2015) states that it was "Based on the author's calculations using data from Illinois TRS and TSR." It appears that Fitzpatrick (2015) used the TSR payroll data for service years (her footnote 18 notes that TSR gives "precise measure of creditable earnings toward the retirement system".) The data on the AEJ website includes a variable called "trssvc" which we believe measures the estimated years of service of TRS. For some teachers labeled as non-takers this measure is well below 22. If the sample of Fitzpatrick (2015) was selected based on 22-28 years of TSR service, then some teachers whose TRS service years were much lower than the TSR service years would not take the upgrade because the benefit of upgrade before 2009 was too low given how far they were from retirement.

To replicate the calculation of take-up rate by Fitzpatrick (2015), we merged 2014 TSR payroll data and the 2014 TRS teacher information file. There are 122,659 records for the fiscal year 1997-1998 in TSR, and a total of 572,528 records in the TRS teacher information file. Exact match based on name (first name, last name, middle initial) results in 121,855 records, or a matching rate over 99% of the TSR records.

We then match these 121,855 teachers with those in the TRS payroll data. Based on our calculation of TRS service from the TRS payroll data we identify a cohort of 18,875 who had 22-28 TRS service years in 1998. Among these 18,875 teachers 15,951 checked "Yes" for the question "Has 2.2 Upgrade", which implied a 84.51% take-up rate by 2014.

We also construct an alternative sample from the matched TRS and TSR data by selecting those with service experience in TSR of 22-28 years in 1998. This sample includes 22,121 teachers, among them 14,756 took upgrade by 2014 (based on the "Has 2.2 Upgrade" indicator provided by the 2014 TRS file). So the take-up rate of based on the TSR experience data is a much lower 66.7%.

This exercise replicates the pattern we reported that the sample with 22-28 TSR service years in 1998 has a lower average take-up rate of upgrade than one with 22-28 TRS service years. Because TRS is responsible for administrating pension benefit, we maintain that the TRS service years are more relevant than TSR service years in calculation of benefit for upgrade.

We find evidence that casts doubt on using TSR service years for retirement related calculations. Table 1 of Fitzpatrick (2015) (Figure B1 in this appendix) shows that among 14,431 teachers with 24-28 years of service experience in 1998, 2,079 were not retired by 2009. They would have 35-39 years of service in 2009, assuming teachers earn one service credit per year. (This assumption is supported by the observations that the average annual service credit earned is about 0.97 for full time/part-time teachers in the age group 55-64, based on TRS (2007).) According to the TRS annual actuarial report, the number of active members who have 35-39 years of service as of June 30, 2009 is 1,097 (see page 56 of TRS (2010)). This is consistent with the scenario that the service years of 1998 used in Fitzpatrick (2015) are lower than those of TRS. In comparison, based on our data in Table B21 the non-retired teachers have 35-39 years of service in 2009 is 1,102, very close to the number (1,097) given in the TRS annual actuarial report.

Finally, the 2014 TRS data contain an indicator labeled "Has 2.2 Upgrade". The take-up rate of the 22-28 service year cohort in 1998 is 85% by 2014 based on this directly reported indicator.

[3] The take-up rate calculated from Table 1 of Fitzpatrick (2015) is 74%, the take-up rates inferred from discussion through the text of Fitzpatrick (2015) range from 71% to 78%.

Table B19: Tracking the upgrade decisions and retirement claims, first year of purchase

retirement claims*	whole sample	upgrade by 1998	upgrade by 2000	upgrade by 2006	upgrade by 2008	upgrade by 2014	upgrade by 2019	non-takers by 2019	true non-takers by 2019***
Regular 2.2	7030	6028	6530	6876	6899	6915	6916	114	
2.2 ERO Employer Pay	6593	4982	5628	6433	6445	6446	6446	147	
2.2 ERO Member Pay	2424	1848	2093	2365	2380	2381	2381	43	
Normal - Actuarial calculation	1148	14	19	26	29	31	33	1115	1115
1 2/3% Formula - Graduated	987	13	15	17	18	19	19	968	968
Rule of 85 - 2.2	477	407	435	461	461	461	461	16	
Regular 2.2 - Disability	41	27	32	38	38	39	39	2	
Rule of 85 - Actuarial	8							8	8
ERO Employer Pay	7	1	1	2	2	2	2	5	5
State ERI	7	6	7	7	7	7	7		
ERO Member Pay	5			1	1	1	1	4	4
Rule of 85 - Formula	5							5	5
Age Type Formula Calculation	2			1	1	1	1	1	1
upgrade claims	whole sample	upgrade by 1998	upgrade by 2000	upgrade by 2006	upgrade by 2008	upgrade by 2014	upgrade by 2019	non-takers by 2019	true non-takers by 2019***
summary statistics									
retired by 2014	18734	13326	14760	16227	16281	16303	16306	2428	2106
not retired by 2014	392	12	17	21	21	26	26	366	366
number of teachers with upgrade claim decisions by year	19126	13338	14777	16248	16302	16329	16332	2794	2472
proportion of these teachers over whole sample (i.e., 19126 as denominator)	100.0%	69.74%	77.26%	84.95%	85.23%	85.38%	85.39%	14.61%	12.92%

*Note: Retirement claims need not to be made in the same year of upgrade, but upgrade decisions need to be made before or on retirement claims.
**Note: We track the upgrade decision to 2019 and retirement claim to 2014.
***Note: We remove teachers with five types of retirement claims in the subgroup of "not upgrade by 2019" to build the subgroup of "truly not upgrade by 2019": "Regular 2.2", "2.2 ERO Employer Pay", "2.2 ERO Member Pay", "Rule of 85 - 2.2", and "Regular 2.2 - Disability". These claims are only available to teacher who upgrade.

Table B20: Decomposition of true non-takers by retirement claims*

retirement options	no. obs	age at retirement claim		exp at retirement claim		benefit of upgrade**	
		mean	median	mean	median	mean	median
true non-taker retired	2106	59.74	60.0	35.34	37.0	5.17	0.0
—-1 2/3 % Formula - Graduated	968	59.85	60.0	36.75	37.0	6.05	0.0
—-Normal - Actuarial calculation	1115	59.64	60.0	34.14	36.0	4.27	0.0
—-Others	23			37.84	38.0	0.69	0.0
true non-taker not retired	366					0.00	0.0
total	2472	59.38	60.0	35.75	37.0	4.41	0.0

*Note: See the last column in table B19.

**Note: This is the expected discounted benefit calculated with actual retirement claim in $1,000 2010 dollars.

Table B21: Comparing retirements and upgrades of 2008/2009 with Fitzpatrick(2015), first date of purchase

	Fitzpatrick			Our data				
Year of Experience in 1998	Fraction who retire by 2009	Fraction who upgrade by 2009	Observations	Fraction who retire by 2008	Fraction who upgrade by 2008	Fraction who retire by 2009	Fraction who upgrade by 2009	Observations
22	0.60	0.70	2409	0.33	0.82	0.55	0.82	2385
23	0.67	0.71	2589	0.56	0.82	0.76	0.82	2430
24	0.78	0.76	2991	0.78	0.83	0.86	0.84	2897
25	0.83	0.78	3121	0.87	0.86	0.90	0.86	3006
26	0.88	0.77	2909	0.92	0.87	0.94	0.87	2972
27	0.89	0.74	2674	0.95	0.88	0.96	0.88	2599
28	0.91	0.73	2736	0.96	0.87	0.96	0.87	2837
whole sample	0.73	0.74	19429	0.78	0.85	0.86	0.85	19126

Note: Upgrade and retirement years are recorded by calendar year. Upgrade year is the year of upgrade purchase, retirement year is the year when one makes retirement claim.

Table B22: Comparing retirements and upgrades of 2008/2009 with Fitzpatrick(2015), final date of transaction

	Fitzpatrick			Our data				
Year of Experience in 1998	Fraction who retire by 2009	Fraction who upgrade by 2009	Observations	Fraction who retire by 2008	Fraction who upgrade by 2008	Fraction who retire by 2009	Fraction who upgrade by 2009	Observations
22	0.60	0.70	2,409	0.33	0.29	0.55	0.49	2,385
23	0.67	0.71	2,589	0.56	0.50	0.76	0.68	2,430
24	0.78	0.76	2,991	0.78	0.72	0.86	0.80	2,897
25	0.83	0.78	3,121	0.87	0.82	0.90	0.84	3,006
26	0.88	0.77	2,909	0.92	0.86	0.94	0.87	2,972
27	0.89	0.74	2,674	0.95	0.87	0.96	0.87	2,599
28	0.91	0.73	2,736	0.96	0.87	0.96	0.87	2,837
whole sample	0.73	0.74	19,429	0.78	0.72	0.86	0.78	19,126

Note: The final date of transaction is the date directly available in TRS. The reported cumulative take-up rate by the year is likely lower than the true take-up rate.

C Tables for Regressions

This section reports OLS and IV estimates of the linear probability model of demand for the pension upgrade: $D_i = \beta_1 P_i + \beta_2 B_i + control\ variables + \epsilon_i$ where D_i (= 1 or 0) is individual i's decision of whether to upgrade, P_i is the price of the upgrade, and B_i is the benefit of upgrade. Fitzpatrick (2015) interprets -β_2/β_1 as the marginal teacher's willingness to pay for each dollar of additional benefit from the upgrade. Fitzpatrick (2015) notes that P_i and B_i are potentially correlated with the error, and constructs instruments using cross-district IV based on district average starting and max salaries. In the text we explained that the ratio $-\beta_2/\beta_1$ is largely driven by the distributions of the benefit and cost of the upgrade across teachers, instead of the willingness to pay by the marginal teacher. The supplemental tables in this section show that our claim is robust.

In Section C.1 we reconstruct the cross-district IVs using our data. In Section C.2. we show the estimates of the ratio $-\beta_2/\beta_1$ using the cross-district IV are similar for the estimated benefit of the upgrade and the actual benefit, despite the fact that the estimated benefit is a poor measure of the benefit of the upgrade.

In Section C.3., we replicate Table 2 of Fitzpatrick (2015) using the data and code obtained at the AEJ website.

In Section C.4. we report placebo tests on the hypothesis that the non-takers' willingness to pay for each dollar of additional benefit from the upgrade, is captured by the ratio $-\beta_2/\beta_1$. We replace the benefit of the observed non-takers by an amount lower than the cost. So the re-constructed cost/benefit ratios for the upgrade for the non-takers are above 1. The re-estimated ratios $-\beta_2/\beta_1$ from the IV regressions are still about 0.2, very similar to those reported in Sections C.2 (using our updated data) and Section C.3 (using the original Fitzpatrick (2015) data). This means the ratio $-\beta_2/\beta_1$ from the regression does not capture the willingness to pay by the non-takers.

The pattern of the estimates of $-\beta_2/\beta_1$ is robust to sample and choice of instruments. More details of the robustness check are available upon request.

C.1 Reconstructing the cross-district IVs

To run the IV regressions of Fitzpatrick (2015), we first merge the TRS and TSR data. We then construct the two cross-district IVs, based on district level starting salaries and max salaries. Both IVs can be implemented in multiple ways.

C.1.1 Matching between TRS and TSR data

Fitzpatrick (2015) uses teacher name, name of employers, and "recorded service accrued" for fuzzy matching between the TRS and TSR data.

Teacher and employer names are available in our data; however, the "recorded service accrued" in 1998 in both system may contain measurement errors, especially for those still working at 1997-1998. For TRS data, we observe terminated total service, optional service and private school service as well as sick leaves for retired teachers and the current values for these variables at the time of request (i.e., 2014). For TSR data, we have several columns on district experience, state experience and out-of-state experience for each job each year. Hence it seems reasonable to use the state experience in 1997-1998 as the total service accrued in TSR; however, in TRS data this value is not directly available and we need calculate it (see section B.3.) We now summarize several implementation strategies for the fuzzy match:

i) merge based on teacher name only;

ii) merge based on teacher name and employer name;

iii) merge based on teacher name, employer name, and "recorded service accrued";

In Section B.7 we discussed statistics of the matching by i). We use the ii) as the baseline and i) and iii) for robustness check.

C.1.2 Constructing the two IVs

Since the ISBE website only keeps the salary table back to 2010-2011 school year, we use that to form our cross-district IVs.

District level starting salaries

We find the beginning salaries for BA, MA, and PhD degree holders (in the salary table, they are recorded as "BB", "MB", and "MPB" respectively) in each district.

District level max salaries

We first find the max salary for MA degree holders (in the salary table, it is recorded as "MM") in each district. Then we calculate the estimated upgrade benefit using probability of retirement in all possible years.

C.2 Regressions with cross-district IVs

In the two stages of regression we first regress cost and benefit on IVs, set of controls and fixed effects; then we regress the binary upgrade decision on the fitted values of cost and benefit as well as the set of controls and fixed effects. Here, we use the same controls and fixed effects, and report in the same format as Columns 1 to 4 in Table 2 of Fitzpatrick (2015).

We report the IV regressions both for the estimated benefit of the upgrade and the actual benefit, in Tables C1 and C2 respectively. The ratio of the coefficients on price over that on benefit is similar for different measures of upgrade benefit.

Table C1: IV regression with our data, estimated benefit of the upgrade

(0)	(1)	(2)	(3)	(4)
Panel A. Salary for BA degree holder	used in creating the instruments			
price	0.006	-0.100***	-0.058*	-0.097**
	(0.005)	(0.032)	(0.031)	(0.043)
benefit	0.001	0.015***	0.010**	0.016***
	(0.001)	(0.004)	(0.004)	(0.006)
Panel B. Salary for MA degree holder	used in creating the instruments			
price		-0.101***	-0.061*	-0.101**
		(0.033)	(0.032)	(0.044)
benefit		0.015***	0.010**	0.016***
		(0.004)	(0.004)	(0.006)
Panel C. Salary for PhD degree holder	used in creating the instruments			
price		-0.107***	-0.074*	-0.115**
		(0.039)	(0.043)	(0.055)
benefit		0.015***	0.011**	0.017***
		(0.004)	(0.004)	(0.006)
County fixed effect	X	X	X	X
District characteristics	X		X	X
Experience × Age fixed effects				X
Experience × Age fixed effects	× District characteristics			X

Table C2: IV regression with our data, actual benefit of the upgrade

(0)	(1)	(2)	(3)	(4)
Panel A. Salary for BA degree holder	used in creating the instruments			
price	-0.009***	-0.055***	-0.038*	-0.040*
	(0.002)	(0.020)	(0.021)	(0.020)
benefit_act	0.003***	0.008***	0.006***	0.006***
	(0.000)	(0.002)	(0.002)	(0.002)
Panel B. Salary for MA degree holder	used in creating the instruments			
price		-0.054***	-0.040*	-0.041**
		(0.020)	(0.021)	(0.021)
benefit_act		0.008***	0.006***	0.006***
		(0.002)	(0.002)	(0.002)
Panel C. Salary for PhD degree holder	used in creating the instruments			
price		-0.060**	-0.048	-0.048
		(0.027)	(0.030)	(0.030)
benefit_act		0.008***	0.006***	0.006***
		(0.002)	(0.002)	(0.002)
County fixed effect	X	X	X	X
District characteristics	X		X	X
Experience × Age fixed effects				X
Experience × Age fixed effects	× District characteristics			X

C.3 Replicating Fitzpatrick Table 2 OLS and IV regression results

Figure C1: Original Table 2 of Fitzpatrick (2015)

TABLE 2— ESTIMATES OF THE RELATIONSHIP BETWEEN DEMAND (*purchase of the upgrade*) AND PRICE

	OLS	IV cross-district			IV within-district	
	1	2	3	4	5	6
Panel A. Salary for BA degree holder used in creating the instruments						
Price ($1,000)	0.009***	−0.082***	−0.071***	−0.070***	−0.167*	−0.171*
	0.001	(0.007)	(0.008)	(0.008)	(0.089)	(0.094)
Benefit (MA maximum salary)	0.001***	0.014***	0.012***	0.012***	0.002	0.003
	0.0001	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Panel B. Salary for MA degree holder used in creating the instruments						
Price ($1,000)		−0.074***	−0.066***	−0.065***	−0.115	−0.121
		(0.008)	(0.009)	(0.009)	(0.104)	(0.116)
Benefit (MA maximum salary)		0.013***	0.012***	0.011***	0.002	0.002
		(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Panel C. Salary for PhD degree holder used in creating the instruments						
Price ($1,000)		−0.092***	−0.090***	−0.090***	−0.129	−0.130
		(0.012)	(0.015)	(0.015)	(0.106)	(0.116)
Benefit (MA maximum salary)		0.016***	0.016***	0.015***	0.001	0.001
		(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
County fixed effects	X	X	X	X		
District characteristics	X		X	X		
Experience × Age fixed effects				X	X	X
Experience × Age fixed effects × District characteristics				X		
District fixed effects					X	X

Table C3: Replication of Table 2 Columns (1) to (4) in Fitzpatrick (2015)

(0)	(1)	(2)	(3)	(4)
Panel A. Salary for BA degree holder	used in creating the instruments			
Price	0.010***	-0.082***	-0.071***	-0.070***
	(0.001)	(0.007)	(0.008)	(0.008)
Benefit	0.001***	0.014***	0.012***	0.012***
	(0.000)	(0.001)	(0.001)	(0.001)
Panel B. Salary for MA degree holder	used in creating the instruments			
Price		-0.074***	-0.066***	-0.065***
		(0.008)	(0.009)	(0.009)
Benefit		0.013***	0.012***	0.011***
		(0.001)	(0.001)	(0.001)
Panel C. Salary for PhD degree holder	used in creating the instruments			
Price		-0.092***	-0.090***	-0.090***
		(0.012)	(0.015)	(0.015)
Benefit		0.016***	0.016***	0.015***
		(0.002)	(0.002)	(0.002)
County fixed effect	X	X	X	X
District characteristics	X		X	X
Experience × Age fixed effects				X
Experience × Age fixed effects	× District characteristics			X

Note: The data for Table 2 at the AEJ website miss a control variable for annual salary. We remove it from the control variable list in OLS and IV regression and replicated the first 4 columns of Table 2 in Fitzpatrick (2015) without controlling the annual salary. We focus on the estimates of cross-district IVs since they are more robust than the estimates based on within-district IVs.

C.4 Further analysis of Fitzpatrick (2015)

C.4.1 Focus on retired teachers

Some late takers purchased the upgrade after 2008-09. Since one cannot upgrade after making retirement claims, we can remove the potential measurement error by focusing on retired teachers only. Table C4 shows that the estimates are very similar to those based on the whole sample.

C.4.2 A placebo test

We conduct a placebo test on the assumption that rational teachers will purchase the upgrade if and only if the benefit of the upgrade is higher than the cost (price). Assume teachers are heterogeneous in the true value of the upgrade. Then teacher i will purchase the upgrade if and only if $B_i > P_i$ and non-takers must have $0 \leq B_i \leq P_i$.

We run placebo tests by resetting the benefit of non-takers to, for example, $\mathrm{B}_i^* \sim Uniform(0, P_i)$ and keeping the values of B_i unchanged for the takers. Then with these values of benefit the observed upgrade behavior should be perfectly rational: non-takers should not purchase the upgrade since the net benefit is negative. We also consider other extreme cases by setting $B_i^* = 0$, $\forall i$ and $B_i^* = P_i$ for comparison. The results are summarized in Tables C5 to C7. We use the retired teacher for these tables; but the results are similar for the original sample. In these three cases all teachers are rational given the values of benefit of upgrade and no one "left money on the table", but the ratio between the coefficients of benefit to that of price is still around -0.2, well below 1 in absolute value.

Table C4: Retired teachers only

(0)	(1)	(2)	(3)	(4)
Panel A. Salary for BA degree holder	used in creating the instruments			
Price	0.019***	-0.096***	-0.075***	-0.081***
	(0.001)	(0.025)	(0.024)	(0.026)
Benefit	-0.000**	0.016***	0.013***	0.014***
	(0.000)	(0.004)	(0.004)	(0.004)
Panel B. Salary for MA degree holder	used in creating the instruments			
Price		-0.071***	-0.061***	-0.065***
		(0.022)	(0.022)	(0.023)
Benefit		0.013***	0.011***	0.011***
		(0.003)	(0.003)	(0.003)
Panel C. Salary for PhD degree holder	used in creating the instruments			
Price		-0.137**	-0.137**	-0.155**
		(0.055)	(0.060)	(0.072)
Benefit		0.023***	0.023**	0.025**
		(0.008)	(0.009)	(0.011)
County fixed effect	X	X	X	X
District characteristics	X		X	X
Experience × Age fixed effects				X
Experience × Age fixed effects	× District characteristics			X

Note: we select teachers with "retire=1" in the sample of Fitzpatrick (2015).

Table C5: Placebo test with $B_i^* = 0, \ \forall i$

(0)	(1)	(2)	(3)	(4)
Panel A. Salary for BA degree holder	used in creating the instruments			
Price	-0.029***	-0.036***	-0.033***	-0.034***
	(0.001)	(0.005)	(0.006)	(0.006)
Benefit	0.007***	0.007***	0.006***	0.007***
	(0.000)	(0.001)	(0.001)	(0.001)
Panel B. Salary for MA degree holder	used in creating the instruments			
Price		-0.031***	-0.029***	-0.030***
		(0.005)	(0.006)	(0.006)
Benefit		0.006***	0.006***	0.006***
		(0.001)	(0.001)	(0.001)
Panel C. Salary for PhD degree holder	used in creating the instruments			
Price		-0.043***	-0.044***	-0.045***
		(0.007)	(0.008)	(0.008)
Benefit		0.008***	0.008***	0.008***
		(0.001)	(0.001)	(0.001)
County fixed effect	X	X	X	X
District characteristics	X		X	X
Experience × Age fixed effects				X
Experience × Age fixed effects	× District characteristics			X

Table C6: Placebo test with $B_i^* = P_i$

(0)	(1)	(2)	(3)	(4)
Panel A. Salary for BA degree holder	used in creating the instruments			
Price	-0.032***	-0.043***	-0.038***	-0.040***
	(0.001)	(0.006)	(0.007)	(0.007)
Benefit	0.007***	0.008***	0.007***	0.007***
	(0.000)	(0.001)	(0.001)	(0.001)
Panel B. Salary for MA degree holder	used in creating the instruments			
Price		-0.036***	-0.034***	-0.035***
		(0.006)	(0.007)	(0.007)
Benefit		0.007***	0.007***	0.007***
		(0.001)	(0.001)	(0.001)
Panel C. Salary for PhD degree holder	used in creating the instruments			
Price		-0.052***	-0.052***	-0.054***
		(0.009)	(0.010)	(0.010)
Benefit		0.009***	0.009***	0.009***
		(0.001)	(0.001)	(0.001)
County fixed effect	X	X	X	X
District characteristics	X		X	X
Experience × Age fixed effects				X
Experience × Age fixed effects	× District characteristics			X

Table C7: Placebo test with $B_i^* \sim Uniform(0, P_i)$

(0)	(1)	(2)	(3)	(4)
Panel A. Salary for BA degree holder	used in creating the instruments			
Price	-0.030***	-0.040***	-0.035***	-0.037***
	(0.001)	(0.005)	(0.006)	(0.006)
Benefit	0.007***	0.007***	0.007***	0.007***
	(0.000)	(0.001)	(0.001)	(0.001)
Panel B. Salary for MA degree holder	used in creating the instruments			
Price		-0.034***	-0.032***	-0.032***
		(0.006)	(0.006)	(0.007)
Benefit		0.007***	0.006***	0.006***
		(0.001)	(0.001)	(0.001)
Panel C. Salary for PhD degree holder	used in creating the instruments			
Price		-0.048***	-0.048***	-0.049***
		(0.008)	(0.009)	(0.009)
Benefit		0.008***	0.008***	0.009***
		(0.001)	(0.001)	(0.001)
County fixed effect	X	X	X	X
District characteristics	X		X	X
Experience × Age fixed effects				X
Experience × Age fixed effects	× District characteristics			X

Note: the coefficients vary slightly to different seeds for random number generation.

D Simulating from a Model of Optimal Choice of Retirement and Pension Upgrade

We solve a model of optimal choice of the joint decision to upgrade and retire based on TRS pension rules and teacher salary schedules. The model is not designed for fitting the data, but numerical simulations of the simplified model are useful for illustration of two points.

One is for quantifying the importance of the unobserved heterogeneous preference error that drives both retirement and upgrade decisions. The benefit of upgrade depends on the optimal timing of retirement in the absence of the upgrade, and taking the upgrade changes retirement timing. We use the joint decision of upgrade and retirement simulated from this model for Figure 1 in the text.

The second use of the data simulated from the structural model is to show that $-\beta_2/\beta_1$ in the regression $D_i = \beta_1 P_i + \beta_2 B_i + \textit{control variables} + \epsilon_i$ does not capture the willingness to pay for the upgrade by a marginal teacher. From the structural model the willingness to pay for a dollar of benefit by a marginal teacher is 1. Table 4 in the text reports regressions using benefits simulated from the structural model show $-\beta_2/\beta_1$ is far below 1 in all regressions. The structural model is also used to study how the measure of benefits and endogeneity affects the estimates. Table 4 in the text shows that when using simulated expected benefits the sign of OLS estimate of β_2/β_1 depends on the presence of artificially generated endogeneity. The artificial data with sample selection are constructed with the endogeneity and IV in the spirit of Fitzpatrick (2015). The OLS and IV produce β_2/β_1 similar to those reported in Fitzpatrick (2015).

D.1 A Model of Optimal Choice of Retirement and Pension Upgrade

We denote the values by the age and experience in the initial period, (a, e). For a teacher alive in year t we denote the probability of survival to period $s > t$ as $\pi(s|t)$. The year t value of pension wealth is $PW_{(a,e)}(t, U)$. The upgrade status $U = 1$ if the teacher took upgrade already and 0 otherwise. [4] The year-t value of pension wealth for a teacher with initial (a, e) retired in year $s > t$ is $\beta^{s-t} E_t PW_{(a,e)}(U, s) = \beta^{s-t}\pi(s|t)PW_{(a,e)}(U, s)$. The annual benefit $B_{(a,e)}(U, s, t)$ in year s is a function of experience, where $B_{(a,e)}(1, s, t) \geq B_{(a,e)}(0, s, t)$. If the teacher retires with experience $e + t \geq e^*$ then there is no benefit from the upgrade (i.e., $B_{(a,e)}(1, s, t) = B_{(a,e)}(0, s, t)$).

The utility function for period t from teaching is $\kappa_t y_{(a,e)}(t) + \nu$, where $\kappa_t = \kappa(\frac{60}{a+t})^{\kappa_1}$ is an age-dependent parameter of leisure, with $0 < \kappa \leq 1$ during working years and captures the disutility of working. The unobserved in preferences for teaching is time-invariant random effect ν that differs by teacher. A lower value of ν favors earlier retirement and raises the benefit of pension upgrade.

The net cost of upgrade (cost net of rebate) bought in period t, $c_{(a,e)}(t)$, is introduced as an additive term to the utility instead of entering the utility of income. We assume $c_{(a,e)}(t)$ is increasing over time. This implies the cost does not affect the marginal utility and the timing of upgrade only affects the present value of the cost. In practice, the cost may be spread over multiple years and may differ by districts. Our assumption frees us from modeling the implementation details that are not crucial in illustrating the economic incentives.

The initial period when the upgrade is available is labelled t. The value function of current teacher with time-invariant preference error ν who takes upgrade is

$$V^u_{(a,e)}(\nu) = max_{r \geq t}\{v^u_{(a,e)}(r, \nu)\}$$

where

$$v^u_{(a,e)}(r, \nu) = \sum_{s=t}^{r} \beta^{s-t}\pi(s|t)[\kappa_t y_{(a,e)}(t) + \nu_s] - c_{(a,e)}(t) + PW_{(a,e)}(r, 1). \qquad (3)$$

Suppose $V^u_{(a,e)}(\nu) = v^n_{(a,e)}(r^{\dagger\dagger}, \nu)$ then with upgrade she retires in year $r^{\dagger\dagger}(\nu)$.

If the teacher does not upgrade, the value function with the optimal timing of retirement is

$$V^n_{(a,e)}(\nu) = max_{r \geq t}\{v^n_{(a,e)}(r, \nu)\}$$

[4]Let $G(a, a+1)$= survival probability from age a to age $a+1$. Then $\pi(s|t) = G(a+t, a+s) = G(a+t, a+t+1)..G(a+s-1, a+s)$. The year t value of pension wealth is $PW_{(a,e)}(t, U) = \sum_{s=t}^{T} G(a+s, a+s+1)\beta^{s-t}(B_{(a,e)}(U, s, t))$, where $B_{(a,e)}(U, s, t)$ is the year-t value of annual retirement benefit received in year s by the teacher with initial (a, e) retired in year t.

with

$$v^n_{(a,e)}(r,\nu) = \sum_{s=t}^{r} \beta^{s-t}\pi(s|t)[\kappa_t y_{(a,e)}(t) + \nu_s] + PW_{(a,e)}(r,0) \quad (4)$$

Suppose $V^n_{(a,e)}(\nu) = v^n_{(a,e)}(r^\dagger,\nu)$ then in the absence of upgrade she retires in year $r^\dagger(\nu)$.
The teacher's decision on upgrade is given by

$$V_{(a,e)}(\nu) = max\{V^u_{(a,e)}(\nu), V^n_{(a,e)}(\nu)\}.$$

If $V_{(a,e)}(\nu) = v^n_{(a,e)}(r^\dagger,\nu)$ then the teacher does not take the upgrade and retires in year $r^\dagger(\nu)$. If $V_{(a,e)}(\nu) = v^u_{(a,e)}(r^{\dagger\dagger},\nu)$ then the teacher takes the upgrade and retires in year $r^{\dagger\dagger}(\nu)$.

The parameters we use for simulation are $\beta = 0.98$, which corresponds to a subjective discount rate of 2%; the parameters of dis-utility to work are $\kappa = 0.64$ and $\kappa_1 = 1.60$.

D.2 Generating Data from the Model of Optimal Choice of Retirement and Upgrade

To simulate from the optimal decision model, we select three districts in Illinois, obtain their district salary table, and simulate the retirement and upgrade decisions for teachers who differ in unobserved preference and the effort they exert.

First, we choose three districts where the salary schedule differs:

Table D1: district salary schedules

district	rcdt	ba_beg	ma_beg	ma_max	ma_year_to_max
1	01-001-0010-26	31843	34086	55910	28
2	01-001-0020-26	35127	36451	50028	24
3	01-075-0100-26	31725	33156	62118	34

note: Data are based on salary table of ISBE 2010-2011, ba_beg is the starting salary for a BA holder, ma_beg is the starting salary for an MA holder, ma_max is the max salary for an MA holder, and ma_year_to_max is the years of service for an MA holder to get max salary.

We conduct two sets of simulations based on the decision model. In both cases, a teacher's salary is based on the district salary schedule and the teacher's own effort. In the first simulation, the teacher's effort is independent from the unobserved heterogeneity in preference error. In the second simulation, to reproduce the endogeneity that those exerting more effort tend to have higher demand for the upgrade, we make the effort negatively correlated with the unobserved preference for teaching. With these simulated samples, we replicate the OLS and IV regression in Fitzpatrick (2015) reported in Table 4 in the text.

We use the average salary schedule in Illinois public school system for the salary function $y_{(a,e)}(t)$ and the TRS pension rules for the original and upgraded benefit, $B_{(a,e)}(0,s,t), B_{(a,e)}(1,s,t)$. We assume teacher salary depends on two parts: the district-level base salary which is common to all teacher in that district, and the individual-level effort-pay which is a multiplier over the district base salary. We denote the teacher effort by eff and assume it takes three equally probable values: $eff \in \{-.1, -0.02, .1\}$. The salary of teacher i in district d at period t is determined by

$$salary_{idt} = e^{eff} \times base\_salary_{idt}$$

where the base salary is calculated by the salary schedule in district d and teacher i's year of service at period t. Moreover, we assume the unobserved heterogeneity in preference for teaching follow a uniform distribution: $\nu \in [-50000, 19300]$. We draw from a 100 point discrete uniform distribution with $\nu \in \{-50000, -493000, ..., 19300]$. And for each ν, we consider 5*4 initial age-experience combinations: $a_0 \in [51, 53, 55, 57, 59]$ and $e_0 \in [22, 24, 26, 28]$. All considered, we have 3*3*100*5*4 = 18000 distinct values of observable features of teachers who may take one value in each of the categories, among 3 districts, 3 levels of effort, 100 values of unobserved heterogeneity, 5 possible initial ages, and 4 levels of possible initial experience. For each teacher, we solve the decision model of retirement and upgrade and record their upgrade decisions and retirement timing with and without the option for upgrade.

With the simulated teacher-level decision we calculate cost and benefit of upgrade, based on actual and estimated benefit. The cost is assumed to be at the time of 1998. The actual benefit is based on the simulated retirement decision. The estimated benefit of Fitzpatrick (2015) is benefit of the upgrade weighted by the distribution of retirement timing in the absence of the upgrade.

Next, following Fitzpatrick (2015) we create cross-district IVs: the beginning salary at that district, and the expected benefit of upgrade with maximum salary paid to an MA holders calculated similarly as the expected benefit of an individual teacher.

Finally, to reproduce the endogeneity suggested by Fitzpatrick (2015), that teachers exerting more effort have a stronger demand for the upgrade, we create a negative correlation between eff and ν by dropping observations with high or low values in eff and ν. These simulated data correspond to Columns 7 and 8 of Table 4 in the text.

References

Fitzpatrick, M. D. (2015). How much are public school teachers willing to pay for their retirement benefits? *American Economic Journal: Economic Policy 7*(4), 165–88.

Fitzpatrick, M. D. and M. F. Lovenheim (2014). Early retirement incentives and student achievement. *American Economic Journal: Economic Policy 6*(3), 120–54.

TRS (2007). Investigation of demorgraphic and economic experience 2002-2006.

TRS (2010). Actuarial valuation and review of pension benefits as of june 30, 2009.

TRS (2018a). Evolution of the TRS benefit structure.

TRS (2018b). Tier 1 member guide.

TRS (2019). Understanding the 2.2 formula.