How Much Do Public School Teachers Value Their Retirement Benefits?

Maria D. Fitzpatrick

Cornell University

Department of Policy Analysis and Management, Cornell University. Mail: 103 Martha Van Rensselaer Hall, Ithaca, NY 14853. Email: maria.d.fitzpatrick@cornell.edu Phone: (607) 255-1272. Webpage: http://www.human.cornell.edu/pam/people/maria_fitzpatrick.cfm

Acknowledgements:
Special thanks to Cynthia Fain, Ed Maybe, Sally Sherman and Sandra Cochran for their help with the Teacher Retirement Service data. This paper has greatly benefited from the comments of Liran Einav, Gopi Shah Goda, Caroline Hoxby, William Johnson, Damon Jones, Ed Olsen, Greg Rosston, John Shoven and Sarah Turner. All errors are my own.
Abstract

Public sector employees receive large fractions of their lifetime income in the form of deferred compensation. It is posited this attracts high-quality public employees. The introduction of the opportunity provided to Illinois public school employees to purchase additional pension benefits allows me to estimate employees' willingness-to-pay for benefits relative to the cost of providing them. The results show employees are willing to pay 19 cents on average for a dollar increase in the present value of expected retirement benefits. The findings suggest substantial inefficiency in compensation and cast doubt on the ability of deferred compensation schemes to attract employees.
1. Introduction

Public sector employees receive relatively large fractions of their lifetime income in the form of deferred compensation, like defined benefit pensions, as compared to their counterparts in the private sector.\(^1\) At the same time, estimates suggest that pension funds for public employees are $3 trillion underfunded in present value terms (Novy-Marx and Rauh 2009). Given the potential burden on taxpayers and the stark contrast to private sector compensation mechanisms, in this paper I investigate whether public sector employees' demand for deferred compensation can explain its use.

Economists have long been interested in the motives for the use of deferred compensation, however, neither the broad analysis of public sector employment nor the specific analysis of teacher labor markets provides theoretical and empirical justification for such a large emphasis on pension compensation. Most of the literature, especially that related to teacher pensions, has focused on the incentive effects of existing pension structures on teacher labor supply decisions.\(^2\) Rather than focus on the effects of the system in place, the question addressed in this research is whether teachers' valuation of deferred compensation equals the cost of providing it.

The most straightforward explanation for generous deferred compensation packages in the public sector is that these packages attract high quality employees because these employees have low discount rates or for some other reason prefer to trade lower current wages for

\(^1\) To illustrate, on average, schools and other public sector employers contribute nearly three times as much per hour worked to the pension benefits of their employees as their counterparts in the private sector. (Based on a comparison of compensation per hour worked of public school teachers versus professionals in the private sector using the Bureau of Labor Statistics' National Compensation Survey.)

\(^2\) Defined benefit pension plans in the public sector have non-linear pension structures that may affect teacher behavior (Costrell and Podgursky 2009, Costrell and Podgursky Forthcoming, Brown 2010). It has been posited that this structure allows employers to manipulate worker labor supply in order to maximize productivity (Lazear 1989).
guaranteed future compensation (Gustman et al. 1994).\(^3\) Investigations have generally shown that employees are willing to trade current wages for future pension benefits, but have not been conclusive about the size of the compensating differential (Ehrenberg 1980, Schiller and Weiss 1980, Smith 1981, Montgomery and Shaw 1997). Whether teachers are willing to forgo current wages for future compensation is difficult to test (Ehrenberg and Smith 1983), in part because wages and pension benefits generally both move cyclically, making it difficult for the econometrician to disentangle differences in teachers' valuations of each. Comparing the wages and pension benefits of public school teachers to those of private school teachers would also be unproductive because apparently similar teachers get less of both forms of compensation in private schools.

In this study, I offer an important test of employee valuation of retirement benefits. To sidestep the identification problems inherent in cross-sectional or panel comparisons of the pension-wage tradeoff, I utilize an opportunity provided in 1998 to Illinois Public School (IPS) teachers to purchase extra retirement benefits. By allowing teachers to choose between current dollars and increased pension benefits, the introduction of this product allows me to estimate employees' demand for retirement benefits and compare it to the expected present value of the benefits.

The price of the contract to upgrade pension benefits is directly proportional to a teacher's annual compensation at the time of purchase. However, a teacher's salary is likely correlated with her health and other factors affecting demand. If so, ordinary least squares estimates of the relationship between price and quantity demanded will be positively biased, possibly producing

---

\(^3\) Risk aversion is one possibility, though risk aversion does not necessarily imply that these types of workers will be attracted to deferred compensation. For example, risk averse younger workers may be attracted to jobs with higher levels of current compensation relative to future compensation because they are uncertain about living or remaining with the employer long enough to receive the deferred compensation.
an upward-sloping demand curve. I create two alternative simulated instruments for price using exogenous variation in program rules. The instruments also use variation in salary schedules at the district level, purging unwanted individual-level variation in prices that might be correlated with employee take-up. In one strategy, identification follows from district differences in teacher salary schedules prior to the product’s introduction. In the other strategy, identification stems from differences in salary schedules across teachers within the same district who have different amounts of experience at the time of the offer. The instrumental variables results vary little between the two strategies. Conditioning on observed demographic factors and market characteristics does not statistically change the coefficients, instilling confidence that the estimates of employee willingness-to-pay are identified from exogenous variation in price.

The results show that IPS employees value pension benefits at the margin much less than the cost of providing them. On average, IPS employees’ marginal willingness-to-pay for the increased pension benefits is over $70,000 less per teacher than the expected present value of the benefits. Said differently, on average, these employees are willing to trade just 19 cents of current compensation for each expected dollar of future compensation. The employees’ low valuation rules out worker preferences as the main justification for large pensions. The implications of these results for other theories justifying the use of generous pensions in the public sector are also discussed.

Today, there is great concern about the inability of governments to pay their promised pension benefits at a time when budgets are already under pressure (Powell 2010). Recent studies suggest projected pension obligations outweigh available pension funds by $3 trillion in present value terms (Novy-Marx and Rauh 2009). The policy options offered to deal with the underfunding problem generally fall into two categories: cut benefits for public sector employees
or increase taxes. The findings of this study imply an alternative Pareto-improving and politically-feasible solution: governments could offer to buy back pension benefits from teachers and other public sector employees. The results here suggest that at least parts of promised pension benefits may be recouped for as little as twenty cents on the dollar. As discussed in the conclusion, a successful buy-back program could go a long way to improving the balance sheets of state and local pension systems.

The next section explains the pension system for employees of IPS, including the benefit formula change in 1998 and the associated upgrade opportunity. Section 3 details the administrative data, which I obtained from the Illinois State Board of Education (ISBE) and the Illinois Teacher Retirement System (TRS). A theoretical model and the simulated instruments framework are described in Section 4. In Section 5, I present and discuss the results. Section 6 concludes with a discussion of alternative models and their importance for future research and policy.

2. Illinois Teacher Retirement Benefits and the Upgrade

A Brief History of Illinois Teacher Retirement Benefits

The first system of teacher pensions in Illinois was introduced in 1915. At that time, Social Security did not exist and teaching was an occupation reserved for unmarried women. Those who remained in teaching for many years were those who never married and therefore could not rely on a spouse's income for retirement support. As such, the system was seen as a mechanism for providing late-in-life support to women who might not otherwise have it. Over time, the system became more structured and some of the parameters changed, but since its

---

4 Reports of women being fired from teaching upon marriage were common and were not isolated to the state of Illinois (TRS 2009).
conception the program has remained a traditional defined benefit pension plan where workers pay in a set fraction of their salary in return for a set benefit upon retirement.

1998 Legislative Change

In the late 1990s, the booming economy and stock markets led to unanticipated surpluses in state budgets and pension funds (relative to their mandated levels). Led by a few legislators from Springfield, a city with many government employees, and encouraged by the unions, the state legislature used the 'extra' funds to grant teachers and other public sector employees more generous pensions.5 The option to purchase additional retirement benefits that I study was the byproduct of this pension 'sweetener'.

The 1998 legislation increased the rate at which pension benefits accrued for teachers' future years of service. As described in more detail below, benefits for service accrued prior to 1998 are subject to a less generous rate of accrual unless a teacher chooses to pay a fee to 'upgrade.' This offer to upgrade retirement benefits is the setting I use to estimate teacher's valuation of pension compensation. Put simply, teachers pay a one-time fee in order to receive an increased stream of payments in the future, much like a single-premium deferred annuity.

This opportunity offered to the employees in IPS allows me to estimate the rate at which teachers are willing to trade-off future compensation for current dollars.

To be more specific, for service accrued prior to 1998, the statutory formula for calculating the annual retirement benefit is nonlinear and depends on the accumulation of

---

5 The decision to spend budget surplus funds and stock market earnings on increased public sector pension benefits was not unique to Illinois. In the 1990s, several other states, including New York and California introduced similar increased.
creditable service in the system. The contribution proportions are 1.67, 1.9, and 2.1 percent of end-of-career salary per year for the first, second and third decades of service, respectively, and 2.3 percent per year for any service beyond 30 years. After 1998, each year of service accrued by any teacher contributes 2.2 percent of a teacher's end-of-career salary to the annual benefit amount. However, the new formula does not apply to years of service already served, which is why the legislature also allowed the option to upgrade, which I study here. The relevant end-of-career salary is the average of the teacher's creditable earnings in the four consecutive highest salary years of the previous 10 years of creditable service. The maximum allowable benefit is 75 percent of a teacher's end-of-career salary.

**Payout from Purchasing the Upgrade**

Figure 1 depicts the formulas in place for service accrual before and after the formula change. The flat rate formula, shown by the solid line, applies to service after July 1, 1998. Benefits for service prior to that point are based on the formula shown by the dashed line, unless members chose to "upgrade". Upgrading increases retirement benefits by an amount equal to the difference between the two formulas times a teacher's end-of-career salary. In other words, the annual benefit of purchasing the upgrade equals the vertical difference between the two lines at a given level of service multiplied by the teacher’s end of career salary.

To be clear, the 1998 pension reform itself made the marginal benefit accrual rate of one additional year of service more generous, but that change is not the focus of this paper. Rather, I

---

6 Creditable service includes years employed in IPS as well as some years spent on medical or military leave and, for a fee, years of service spent in private schools. Annual benefits are available to members of the TRS when they terminate active service with IPS and meet the following age and service requirements: age 55 with 35 years of service, age 60 with 10 years of service, or age 62 with 5 years of service.

7 Should an employee with at least 30 years of experience choose not to upgrade, she will still receive the 2.3 rate of accrual for her remaining experience.

8 With the flat rate for benefits, a teacher receives the maximum benefit with 35 years of service. Under the preexisting formula, the maximum benefit was reached when a teacher accumulates 38 years of experience.
focus on the upgrade opportunity, which was the chance to pay a one-time fee in exchange for an increased retirement benefit in the future. In order to illustrate the value of the upgrade, Table 1 recreates a table of information presented to the employees themselves by the TRS. Almost all of the parameters defining the value of the upgrade are fixed as of 1998, except the teacher's end-of-career salary. For this reason, it is convenient to represent the upgrade in terms of its value relative to end-of-career salaries, as is done in the table.

Teachers who retire with 10 years of experience (all of which was accrued under the pre-1998 formula) receive an annual retirement benefit equal to 16.7 percent of their end-of-career salaries. Purchase of the upgrade would increase the value of their annual retirement benefit to 22 percent of their end-of-career salaries. Therefore, as a result of purchasing the upgrade, these teachers would receive an additional annual benefit of 5.3 percent of their end-of-career salaries, or a 31.7 percent increase in their annual retirement benefit. Similar comparisons can be made for teachers with other levels of experience in 1998 by looking at the other rows of the table.

**Price of the Upgrade**

To receive the upgrade for service years prior to 1998, a teacher had to pay a one-time fee equal to one percent of her salary for each year of service accrued by 1998, $Exp_{1998}$. The relevant salary for the upgrade fee, what I will call the salary-at-time-of-purchase or Salary, is one's highest annual salary in the four years prior to the decision to upgrade. Wage growth, either because a teacher is moving along the salary schedule or because salary schedules increase, causes the price of the upgrade to rise as a teacher delays purchase. The maximum upgrade fee
is 20 percent of the salary-at-time-of-purchase. The price of the upgrade to teacher \( i \) who purchases in year \( t \), \( Price_{it} \), is therefore \( Price_{it} = \min\left(\frac{Exp_{1998}}{100}, \frac{20}{100}\right) \times Salary_{it}.^{9} \)

The optimal time to purchase therefore varies across employees depending on wage growth, interest rates, discount rates and future employment expectations. Under reasonable assumptions about each of the first three of these factors and regardless of future years of work, the present discounted value (PDV) of the price was lowest for most employees if they purchased in 1998.\(^{10}\)

**Expected Financial Return on the Upgrade**

To understand the purchase decisions of employees, it is helpful to have a sense of how valuable the upgrade is. First, I show a back-of-the-envelope calculation, as this may be how many people evaluate investment opportunities like this annuity. Consider a teacher who has 20 years of experience in 1998 and purchases the upgrade immediately. The price she pays for the annuity is 1 percent of her salary-at-the-time-of-purchase per year of experience, or 20 percent of her salary in 1998. The annual benefit to her from this purchase will be 8.2 percent of her end-of-career salary in each year of retirement. In the simplest case, if she retires right away and there is no discounting, she will earn her money back on the investment (and then some) within 3 years. This simple case is illustrated in Figure 2, which shows the price of the upgrade, the increase in annual benefits associated with the upgrade and the lifetime value of the upgrade.

---

9 There are a few other details about the upgrade price worth mentioning. For every three years of service the teacher accrues after July 1998, the number of years used in the upgrade calculation is reduced by one year. Finally, once the teacher works enough to reach the maximum retirement benefit, the annuity no longer increases her pension benefit amount. To compensate teachers who work so many years the upgrade will have no value, once a teacher is actively employed in a creditable position for more than 35 years, the purchase amount of the upgrade is refunded (with interest) at a rate of 25 percent per year of service.

10 Teachers with little experience in 1998 who face greater than 5 percent wage growth experience lower PDV prices if they wait to purchase. This is because the salary growth does not compound as quickly to put upward pressure on the price. This is another reason I focus on teachers with over 22 years of experience in 1998, since I am most likely to observe their purchase decisions in my censored data (see Section 3).
(without discounting) for a 60 year old teacher who retires immediately in 1998 and expects to live until age 76.

Of course, some teachers will not retire for many years after 1998. In this case, each annual payment will be discounted at a higher rate to account for the greater delay in payment. However, discount rates would have to be extremely high, or life expectancy low, to make the upgrade a poor investment. Since most teachers can expect to collect many years of payments (the average in my censored data is 14 years) and experience real wage growth, for most employees under reasonable assumptions, the upgrade is a good investment.

More generally, one can calculate an expected rate of return on the product and compare that to other assets the teachers could have purchased (Friedman and Warshawsky 1990, Mitchell et al. 1999). For each eligible employee in IPS, I calculate the expected internal rate of return, or the discount rate at which the price of the upgrade equals the expected PDV of the extra retirement benefits. In expectation, the annuity was a rather good investment for most employees. For 99 percent of eligible teachers, the expected rate of return is above 7 percent.

Facilitated Enrollment: Information and Payment Options

In this study, I examine the upgrade purchase decisions of employees. Since the upgrade was generally a good deal for most employees, take-up should be high. One thing that could prevent high take-up of this type of product is a lack of information. If IPS employees did not know about the upgrade or the extra retirement benefits it would provide. This is because the 2.2 formula change and the

---

11 Because I use average mortality and retirement probabilities this is not necessarily the employee's own expected rate of return. Instead, it represents an average across employees in a cohort with a given amount of experience. See the Appendix for more information about how I calculate this rate of return.
upgrade were widely publicized. Both were the subject of many news reports and newspaper articles in 1998 (Finke 1998, Erikson 1998, Thomson and Mask 1998). Also, after the formula change, detailed mailings were sent out to IPS employees documenting how the price and value of the upgrade could be calculated and providing assistance to employees in determining the price and benefits of the upgrade for themselves. Examples of worksheets mailed to employees in the weeks following the formula change are provided in Appendix A.

Moreover, teachers and other public school employees are generally a communicative and informed group. First, they are unionized; one role for unionization is information provision. In fact, the teachers unions in Illinois also sent information to their members about the upgrade and its value to them. Second, benefits are fairly standardized and have few parameters, making it easier for employees to share information with each other. It is therefore unlikely that a lack of information prevents take-up for many employees. Existing survey evidence has shown that teachers in traditional defined benefit pension plans were quite knowledgeable about their pension plans, more so than workers in the private sector more generally (DeArmond and Goldhaber 2010).

Alternatively, teachers may be credit constrained and unable to purchase the upgrade. However, credit constraints likely play much less of a role in this instance than they would in a typical private annuity market because employees could spread payments over as many as five years. The ability to have the payment for the upgrade deducted from one's paycheck, deducted from one's retirement benefit check (if retiring immediately after the new policy came into place) or even transferred from other retirement accounts lessens concerns about binding credit

---

12 Consider, for example, the relative simplicity of a defined benefit pension program as compared to a defined contribution pension program. While the former has essentially two parameters - the contribution rate and the benefit accrual rate - that are constant across employees, the latter has many pieces - including contribution rates, matching rates, asset distribution decisions, interest rate-risk tradeoffs, etc. - that may vary across employees.
constraints. Moreover, in addition to extending actual payments over several years, teachers could delay purchase to allow themselves time to save.

3. Data Description

To conduct the analysis, I use data on employees of IPS amalgamated from two sources, the Teacher Service Record (TSR) and the Teacher Retirement System. Because the data are from an administrative record of all employed service in IPS over the entire period from 1987 to 2009, I can completely characterize the employment, retirement contribution and benefit receipt experiences of every employee of IPS over the period. Next, I describe each data set in turn and then the process by which I combine them. The section concludes with a description of procedures for creating the information used in the analysis, including the prices and costs of the upgrade.

*Illinois Teacher Service Record*

The first resource for my data on IPS teachers is the Teacher Service Record (TSR) covering the years from 1987 to 2009. The TSR is a database compiled by the Illinois State Board of Education (ISBE) from school district administrators to track employment and salaries of teachers and administrators in public schools throughout the state. Each observation in the TSR is an employee record for a given school year. The TSR includes the following information about employees in IPS: name of the teacher or administrative employee, a unique identifier for the employee,\(^\text{13}\) the school and district in which the teacher is employed, total compensation (as reported to the relevant retirement system), number of months employed at the position, full-time equivalent percentage of the position and the percent of time that is administrative. The data also

\(^{13}\) This identifier has been provided for the years prior to 2005. For later years, I have used employees' names and employment information to match records to the years prior to 2006.
contain information on the number of years of experience (within the district, within Illinois and out-of-state) and the highest degree held by the employee.

The reported compensation includes, among other things, extra-duty pay (coaching, clubs, etc.), vacation and sick day buyouts, bonuses, school-board-paid retirement contributions, and other compensation that the Teachers Retirement System (TRS) includes in total creditable earnings. This measure of compensation data does not include the cost of employer-paid health insurance or other benefits provided by the school-board to the employee. Importantly for the current work, the compensation measure recorded in the TSR is a precise measure of creditable earnings toward the retirement system. I will use the terms salary and compensation interchangeably to refer to this measure of creditable earnings.

**Teacher Retirement System**

I also use data collected by the TRS, the pension fund for teachers and administrators in IPS, but not Chicago Public Schools (CPS). The TRS has provided data on the retirement benefits paid to its members. This includes information about the name of the benefit recipient, the size of the annual benefit payment, the timing (beginning and ending) of benefit receipt, the creditable years of service and age of the member (employee). Additionally, the data contain information about upgrade purchases, including their timing and cost.

**Combining the Data**

The two administrative sources (the ISBE and the TRS) do not use a common identifier for teachers. Therefore, I use fuzzy matching techniques to combine the data. Both sources

---

14 Employees of Chicago Public Schools are covered by the Chicago Teachers Pension Fund. For more information, see the Data Appendix.
provided me with the names of teachers as recorded in their systems. I also have information about recorded service accrued in both systems and employer information. Using this information, I find matches for 97 percent of teachers in the TRS data who were eligible for the upgrade. I include all teachers that match between the two datasets in the sample.\(^\text{15}\)

**Defining the Prices and Costs**

Two crucial measures for the analysis are the price of the upgrade and how much an employee can expect to collect in benefits, which is equal to the expected costs to the TRS of providing the extra retirement benefits purchased. (In what follows I will refer to these as either price and either expected benefit or costs.) To repeat, the basic formula for the price charged to employee \(i\) who purchases in period \(t\) is

\[
\text{Price}_{it} = \min\left(\frac{\text{Exp}_{1998}}{100}, \frac{20}{100}\right) \times \text{Salary}_{it}.
\]

In practice, I define the price variable as 1 percent of the highest salary earned in the last four consecutive years of earnings prior to purchase times the number of years of service the employee had accrued by 1998, up to a maximum of 20 percent. For teachers that did not purchase the upgrade, the price variable is defined using their highest salary earned as of 1998.\(^\text{16}\)

Additionally, I inflate prices to 2010 dollars as described in Appendix B.

The cost to the pension fund of a person purchasing the upgrade is equal to the PDV of the extra amount that is paid out annually for each year that a person receives retirement benefits. (In calculating these costs, I exclude any survivor benefits. Since many spouses and survivors collect benefits after an employee is deceased, this results in my understating the actual costs of

---

\(^{15}\) This choice of matching algorithm does not affect the results. More information about the exact matching techniques and results with different samples is available from the author upon request.

\(^{16}\) Note that neither of these definitions of price takes future years of service into account. Doing so makes little difference to the results.
IPS employees make their purchase decisions without observing their actual payout of benefits. I therefore define the cost to the pension fund of the upgrade as the expected PDV of the extra retirement benefits based on an employee's expected mortality, expected retirement behavior and her highest salary as of 2009.

These extra benefit costs accrue over time, so it is necessary to discount them to their present value in 1998 to facilitate the comparison between costs and prices. I use the true interest cost rate for Illinois state bonds sold in 1998, 5.10 percent per year, as the discount factor for measuring the costs of the increased retirement benefits in 1998 terms. As a proxy for an employee's salary at retirement, I use a teacher's highest observed salary to calculate her additional retirement benefit from the upgrade purchase, $A$. TRS increases benefits by 3 percent per year, to keep pace with inflation, so I include the increases in my calculation of the cost. The present value of the total cost of providing the upgrade to an employee is therefore

$$\text{Cost}_i = \sum_{j=\text{Age}_{1998}}^{119} \frac{(1.03)^{A_i+M_j+R_j}}{(1.051)^j}.$$  

The annual additional benefit to teacher $i, A_i$, is essentially weighted by the probability the employee collects a payment in that particular year, which in turn depends on her cohort's probability of surviving until year $j, M_j$, and her cohort's probability of retiring before year $j, R_j$. Once I have this measure of the expected PDV of costs as of 1998, I inflate it to 2010 dollars to make it comparable with the price measure and place its value in a more current context.

Table 2 presents information on IPS employees' take-up, price, cost and retirement decisions. I focus on employees with 22 to 28 years of experience in 1998. Each row represents employees with the amount of creditable service accrued in 1998 indicated by the row

---

18 More senior cohorts are a selected sample because members of their cohort had already begun retiring when the upgrade was introduced. On the other hand, because teachers have until retirement to purchase an upgrade, I may not observe the true purchase decisions of cohorts with less experience (because my data are censored in 2009).
header. The second column reports the fraction of retirees who have retired by 2009, the endpoint of my sample. Here, retirement is defined as the collection of benefits, though it may not be synonymous with leaving the classroom. Teachers with less experience are less likely to have retired by 2009 than teachers with more experience. For example, 83 percent of teachers with 25 years of experience in 1998 had begun collecting pension benefits by 2009, while 91 percent of the teachers who had 28 years of experience in 1998 had retired by 2009. This difference illustrates why I focus on those employees with the highest experience and the largest eligible populations. Censoring of the data may keep me from observing purchase behavior of employees who have not yet retired, so I focus on those whose retirement history I am most likely to observe.

The third column of the table reports the fraction of teachers who have purchased the annuity by 2009. The take-up rate ranges from 70 percent for teachers with 22 years of experience in 1998 to 78 percent for teachers with 25 years of experience in 1998. Recall that if a teacher accumulates at least 38 years of experience, there is no reason to purchase the upgrade. Teachers with more experience, i.e. those with 28 years or more in 1998, may be less likely to purchase the upgrade because they anticipated staying employed long enough to reach the maximum benefit.

Columns 4 and 5 of Table 2 illustrate the prices offered to and the costs to the pension fund of providing the upgrade. The average price of the upgrade offered to employees with 25 years of experience in 1998 was $15,245 while the expected costs of providing them with the extra retirement benefits if they all purchased would have been $94,166. That more than 20

---

19 The average costs to purchasers are slightly higher than the average costs to the whole population. This could be indicative of selection into who purchases the upgrade. If the people who expect to benefit the most from the upgrade are the most likely to purchase it, the average costs of providing the upgrade to purchasers will be higher than the average costs to non-purchasers.
percent of these employees do not purchase more retirement benefits even when offered them at just 16 cents on the dollar, on average, indicates that at least some teachers prefer current compensation to increases in deferred compensation.

4. Demand for and Costs of Upgraded Pension Benefits: Theory and Estimation

Demand for the Upgrade

Underlying the estimation of the demand and the expected benefit cost curves described below is an environment where employees with rational expectations about the future are choosing between two pensions to maximize their utility subject to a lifetime budget constraint. The first pension available to employees is an annual retirement benefit that is provided free of charge, indexed by \( l \). The choice employees make is of whether to pay a price, \( p \), to purchase a pension with additional retirement benefits, \( h \). The price of the pension with additional retirement benefits is proportional to the employee's salary at the time of purchase, \( s_1 \), while the value of both \( l \) and \( h \) are proportional to the employee's end-of-career salary, \( s_2 \).

I assume employees have well behaved utility functions, that depend on the price of the extra benefits and the characteristics of the employees, \( X \) (e.g. risk preferences, health status, etc.). As such, the utility to a consumer of type \( i \), with characteristics \( X_i \), of purchasing the additional retirement benefits, or the upgrade, is \( U^N(X_i, p(s_1), s_2) \) and the lifetime utility to consumer of type \( i \) of not purchasing the upgrade is \( U^I(X_i, s_2) \). An employee will purchase the extra retirement benefits if the increase in lifetime utility she gets from doing so is larger than the decreased utility from the loss in current income, i.e. if \( U^N(X_i, p(s_1), s_2) > U^I(X_i, s_2) \). For each

---

20 This model and exposition are similar in spirit to a model for health insurance described in Einav, Finkelstein and Cullen (2010). It differs from the actual upgrade setting in that the price and benefits of the extra pension are delineated in dollars rather than in fractions of one's salary. The estimation strategy incorporates this feature of the upgrade more directly.
employee, there is a reservation price, which is the highest price she is willing to pay to purchase the retirement benefits.

In order to determine employees' willingness to pay, I estimate the aggregate demand curve for the upgrade. The aggregate demand curve traces out the distribution of reservation prices in a population of employees. In other words, employees who purchase at price $P$ but do not purchase at price $P + \varepsilon$ are those whose reservation price is $P$. Estimating the aggregate demand function for the upgraded retirement benefits involves specifying the underlying relationships between take-up and the prices at which consumers are offered the upgrade.

I estimate demand, $D_i$ (a dummy variable equal to one if a teacher purchases the upgrade), as a function of the price at which she is offered the upgrade, $P_i$, and the value to her of purchasing the upgrade, $B_i$:

$$D_i = \beta_0 + \beta_1 P_i + \beta_2 B_i + \varepsilon_i.$$  

Recall that the price of the upgrade is one percent of a teacher's salary-at-time-of-purchase (Salary) per year of service accrued in 1998 ($1998\text{Experience}$). In equation (1), the price of the upgrade to teacher $i$ in district $d$ of county $c$ at time of purchase $t$ is therefore:

$$P_{idct} = 0.01 \times \text{Salary}_{idct} \times 1998\text{Experience}_{idc}.^{21}$$

Demand for the upgrade also depends on the value of the upgrade to the teacher. First, at any given price, the upgrade will be more attractive to the employee if she expects to collect more in retirement benefits as the result of purchasing it. Second, both the price and the additional amount of retirement benefits collected by the employee are related to her salary at the end of her career.\textsuperscript{22} In this sense, teachers who face higher prices (because of higher salaries at

\textsuperscript{21} Note that for simplicity, I have dropped the subscripts $d$, $c$, and $t$ in equation 1.

\textsuperscript{22} In theory, a teacher can move from one district to another and/or take on additional responsibilities to increase her end-of-career salary relative to the price she paid for the upgrade. In practice, however, because teachers are
the time of purchase) may be more likely to purchase the upgrade because they expect to collect more from the upgrade (because of higher end of career salaries). Said differently, teachers who face different current salaries may be presented with a different product because their potential to collect benefits is correlated with the price they are charged. To disentangle the effect of price from the effect of the expected benefit, I add a control for future benefits, $B_i$. The maximum scheduled salary for MA teachers in the same district as a teacher is a good proxy for expected future benefits because most teachers i) have MA degrees and ii) reach the maximum scheduled salary before retiring.

*Instrumental Variables Frameworks for Demand Estimation*

The price charged to a teacher depends directly on her salary at the time she purchases the upgrade. A concern is that this salary is not randomly assigned to teachers, but that a teacher’s salary is determined by factors that also affect demand, i.e. $\text{Cov}(\text{Salary}_{it}, \epsilon_i) \neq 0$. There are three dimensions along which teachers' choices determine their compensation at time of purchase: purchase timing, effort and choice of employer. Each may be related to unobservable teacher characteristics driving demand for deferred compensation. For example, consider the case of teachers who take on extra duties to earn extra pay, i.e. those who induce more 'effort.' Those who do so may be those without spouses or other family members to rely on for current or pension income, making them both want to take on extra duties for increased current income and to purchase the upgrade for its extra income upon retirement. If so, ordinary least squares estimates of the relationship between price and quantity demanded will be relatively immobile and salary schedules are rather rigid and compressed, the salary used to calculate the price of the upgrade will be correlated with a teacher's end-of-career salary.

---

23 This is akin to controlling for the product's characteristics in a hedonic framework. Note that I also include controls like age and experience of the teacher, which should proxy for the fact that, ceteris paribus, an employee can expect to collect more from the upgrade if she expects to live longer or retire earlier.
positively biased (in part because I am unable to control for marital status), possibly producing an upward sloping demand curve.

Given the potential endogeneity of realized teacher salary in the demand for deferred compensation, an appropriate instrumental variables strategy makes use of the variation in teacher salaries among teachers in Illinois that is most logically exogenous to a teacher's compensation and career trajectory after the upgrade is offered. I use two distinct instrumental variables techniques to estimate demand. Each has different underlying assumptions about how teachers' salaries and their demand for additional retirement benefits are related to one another, but, as I detail in the next section, they produce quite similar results.

First, I create a simulated instrument based on the prevailing salary schedule in a teacher's district of employment before the upgrade is offered. Specifically, I estimate the following first-stage equation:

\[ P_{idct} = a_0 + a_1 \text{BeginningSalary}_{d1998} + \omega_i. \]

My instrument for the price of the upgrade is the beginning salary paid to a teacher with a Bachelor's Degree according to the salary schedule in the employee's district of employment in 1998.\(^{24}\)

The use of this simulated instrument purges the price of many sources of variation correlated with or driven by the choices teachers make. First, by eliminating the timing dimension of the decision, I remove the risk that teachers may be updating beliefs about their future health and then factoring those updated beliefs into their purchase decisions. Second, exerting effort, through attainment of a Master's Degree or by taking on duties like coaching, offers teachers higher salaries than if they did not exert these types of effort. If people who are

\(^{24}\) As will be seen in the discussion of the results, I also use alternative measures based on the scheduled salary for employees with other educational attainment levels.
motivated to exert effort are those who are likely to collect more in retirement benefits (because they either live longer or retire earlier), then, *ceteris paribus*, teachers with higher salaries (and therefore higher prices), may be more likely to purchase the upgrade. Because the instrument is the same for all teachers within a district, this instrument is invariant to teacher effort.

To instill confidence the instrument I have created is uncorrelated with the error term in equation (2), I also include measures of employee characteristics as of 1998, $X_t$, e.g. Master’s Degree attainment, year of experience fixed effects, a polynomial in age, controls for her position in the school, subject matter taught, etc. Additionally, in some specifications, I include market-level characteristics, $X_d$, to control for differences in school and area-level characteristics that might lead teachers to sort across districts. When possible these characteristics of the market in which a teacher is employed are measured at the district-level.\(^{25}\) When district-level characteristics are not available, I use county-level or Public-Use Microdata Area (PUMA)-level data. These market-level variables include measures of the distribution of household income, unemployment rates, marriage rates, educational attainment of residents, etc.\(^{26}\)

Finally, to the extent that teachers with certain characteristics may be drawn to reside and be employed in areas within Illinois that are different from each other in ways that are unobservable, estimates of equations (2) may produce biased results. For example, if teachers with the highest valuation of retirement benefits are all drawn to the areas with the highest beginning salaries (conditional on maximum salaries), this would produce a more positive relationship between price and demand than the true relationship. Therefore, I also include

\(^{25}\) These district-level characteristics come from tabulations done by the U.S. Census and provided for the National Center for Education Statistics.

\(^{26}\) The set of market-level variables is broad and defined by what is available from the U.S. Decennial Census, Common Core of Data and the Bureau of Labor Statistics. Here, when using county or PUMA level information from the Census, it is from the Decennial 2000 Census. Similar conclusions are drawn if more aggregated data from the 1990 Census are used instead.
county fixed effects, $\varphi_c$; identification of the demand curves therefore stems from within-county differences in take-up and the slope of salary schedules.\footnote{27}

The resulting set of estimation equations is therefore:

\begin{align*}
D_i &= \beta_0 + \beta_1 p_i + \beta_2 B_i + \theta X_i + \pi X_d + \varphi_c + \varepsilon_i. \quad (2) \\
P_{idct} &= a_0 + a_1 \text{BeginningSalary}_{d1998} + \mu X_i + \kappa X_d + \vartheta_c + \omega_i. \quad (3)
\end{align*}

The assumption in these specifications is that teachers with similar characteristics and family circumstances are likely to be assigned at random to different, but observationally similar, districts within a county. In other words, variation in teacher salary structures within local labor markets determines the price that a teacher pays for additional pension coverage and is unrelated to a teacher’s choices about effort that determine pension demand.

In the second instrumental variables strategy, I rely on within-district variation in salaries paid to teachers with different amounts of experience at the time of the upgrade. Specifically, rather than using county-level fixed effects, I include district-level fixed effects, $\vartheta_d$. Since the original instrument for price, the beginning salary paid to a BA teacher in one’s district of employment in 1998, is constant across all teachers in the same district, I use an alternate order statistic from the salary schedule: the scheduled salary paid to a teacher with a BA degree and the same amount of experience as teacher $i$ in her district of employment in 1998, $\text{BAExperience}_{i1998}$. The set of estimation equations in this instance is:

\begin{align*}
D_i &= \beta_0 + \beta_1 p_i + \theta X_i + \varphi_d + \varepsilon_i. \quad (4) \\
P_{idct} &= a_0 + a_1 \text{BAExperience}_{i1998} + \mu X_i + \vartheta_d + \omega_i. \quad (5)
\end{align*}

Note that equation (4) does not include $B_i$, the end-of-career salary of MA degree-holding teachers, because it will be the same for all teachers within a district.

\footnote{27} Identification stems from the slope of salary schedules rather than the levels, because the level of the salary schedule is essentially held fixed when I control for $B_i$, the maximum salary paid to an MA degree holding teacher.
Much like the other instrument, the use of $BAExperience_{id1998}$ purges unwanted variation across teachers in labor supply responses to the upgrade opportunity because the instrument is measured before the upgrade was offered. Also, this instrument is once again unrelated to any choice about effort, since it relies only on the scheduled salary order statistics. With this specification, the underlying assumption is that there are no differences between teachers in the same district who have similar characteristics but had different levels of experience in 1998. In other words, there are no differences in cohorts of entering teachers or selection into who leaves IPS after between 22 and 28 years of experience that are not adequately captured by the included controls.

**Expected Costs of the Upgrade**

Employees’ willingness to pay for the benefits alone do not allow me determine whether employees’ valuation of the upgrade because in the population of employees the value of the upgrade will depend on an employee’s characteristics, like her expected length of life and retirement date. In this way, the upgrade product can be thought of like an insurance product where the expected present value to the employee is the same as the expected cost to the employer of providing her with the long-life insurance (Eniav, Finkelstien and Cullen forthcoming). In this setting, the supplier is the government, who offers to sell the higher level of retirement benefits to any employee who wishes to purchase at a set price. The expected PDV of the cost of providing the extra retirement benefits varies over consumers depending on their expected mortality, labor supply and wage trajectories. These costs of providing the extra retirement benefits can be averaged over purchasing employees to obtain the aggregate average
cost curve and, in turn, the marginal cost curve. In doing so, I implicitly assume there are no administrative or other additional costs involved in providing the extra retirement benefits.

I estimate the aggregate average cost curve by using information about the relationship between the expected costs to the pension fund of providing the extra benefits to a purchasing employee, $C_i$, and the take-up of eligible employees, $D_i$:

$$C_i = \gamma_0 + \gamma_1 D_i + u_i. \tag{3}$$

This expected cost to the pension fund, $C$, is a proxy for the extra benefit collected by the employee who purchases the upgrade. It is not precisely the expected benefit to the employee because the employee may have information about her own life expectancy or retirement date that is unobservable to me, the econometrician.

I am interested in the average cost curve for eligible employees. In (3), I estimate the average cost curve using individual-level data on expected costs and take-up. As such, it is identified off of the average costs among purchasers and the average costs among non-purchasers. An alternative strategy is to group the data by the exogenous price faced by the employees to estimate average cost among purchasers as a function of the proportion of the group that purchases at a particular price.\textsuperscript{28} This method identifies the average cost curve from the changes in average costs across populations with different take-up rates (because they faced exogenously different prices). The estimates using grouped data are similar to those reported using individual level data.

\textsuperscript{28} More precisely, I group eligible employees by the exogenous price at which they were offered the upgrade, as predicted by the first stage of my 2SLS estimation. I then calculate the average costs and take-up rates for each group and regress costs on take-up.
5. Results

Standard economic theory predicts demand should be negatively related to price. Table 3 presents both ordinary least squares (OLS) and two-stage least squares (2SLS) estimates of the effect of price on take-up. The OLS results in the first column of the table show a positive relationship between price and demand, suggesting an upward sloping demand curve. The OLS results suggest that employees are 1.9 percentage points more likely to purchase the upgrade when the price is $1,000 higher. This positive relationship exists even though all of the available controls for individual characteristics like age and employee position are in the regression, which suggests the positive relationship is not driven by older employees with higher salaries or the administrative employees who also have higher salaries. Also, all of the market- and district-level controls are included in Column (1), including county fixed effects. Their inclusion changes the OLS estimated relationship between price and take-up hardly at all, suggesting public school employees are not sorting across markets based on their propensity for take-up.

However, two observationally equivalent teachers who face different prices may be very different. For example, one may be more risk averse and therefore may take on extra duties so as to earn more income for saving. This teacher's salary would be higher. This risk averse teacher would also be more likely to purchase the extra long life insurance provided by the increased extra benefits of the upgrade. In this case, the OLS regressions would produce results implying that teachers who face higher prices are more likely to purchase the upgrade. But, higher prices do not generally induce demand. Rather, in this example, teachers' underlying risk aversion generates both the differences in prices and the difference in take-up. Because the prices used in the OLS estimation are not exogenous, they do not reveal employees' underlying demand for the upgrade.
To be sure it is exogenous variation in price that drives changes in the estimated take-up, I perform the estimation using the two 2SLS strategies described in the previous section. First, I discuss the results using the instrumental variables strategy that uses cross-district differences in salary schedules for identification. In Panel A of Table 3, the instrument is as specified in equation (3) -- the beginning salary paid to BA degree holding teachers in 1998 in teacher $i$'s district. Column (2) presents the estimated coefficients on $P_i$ and $B_i$ using a parsimonious specification that includes controls for individual characteristics, but not controls for district characteristics or county fixed effects. In columns (3) and (4), I add in county fixed effects and district level controls, respectively. To be sure the estimated effect is not an artifact of my choice to use the BA teachers’ salary schedule in estimation, in Panels B and C of Table 3, I repeat the estimation procedures just described using the beginning salaries paid to MA degree and PhD degree holding teachers, respectively.\(^{29}\)

The 2SLS estimates accord with economic theory: demand decreases as price increases. All of the estimates of the effect of price on take-up in Columns (2), (3) and (4) are statistically significant at the one percent level. The bulk of the estimates in the table suggest that a $1,000 increase in price leads to a decrease in take-up of between 8 and 11 percentage points. The two estimates using the beginning BA salary as the instrument, but no district level controls, suggest that teachers are even more sensitive to price, i.e. that a $1,000 increase in price decreases take-up by 14 or 15 percentage points. However, the inclusion of district level controls brings the estimate in line with the estimates using salary schedules for more educated teachers. That the inclusion of district controls changes the estimates using the BA salary schedule suggests there may be unobservable differences across teachers that are correlated with how the districts set

---

\(^{29}\) Some districts do not report having salary schedules specific to teachers with PhDs. For these districts I use the salary schedule for MA degree holders as the schedule for PhD holders.
their salary schedules for BA teachers. However, when using the other salary schedules, the estimates change little with the inclusion of district-level controls and are statistically indistinguishable from one another. The estimated relationship between price and take-up varies little using the preferred specification with the host of individual, district and market characteristics.

Before turning to the results using the alternate instrumental variables strategy, I first briefly discuss the coefficients on the proxy for the expected benefit of purchasing the upgrade, $B_i$. All else equal, an increase in the expected payout from the upgrade should increase take-up. Recall that $B_i$ is the maximum salary paid to a MA degree holding teacher in the district in 1998. The coefficient estimates in Table 2 suggest that a $1,000 increase in this maximum salary has the effect of increasing take-up by between 2.1 and 4.1 percentage points.

Estimates of the aggregate demand curve using the instrumental variable strategy that relies on the within-district variation in the scheduled salary across teachers with different amounts of experience are in column 5. Again the three panels present estimates using salaries for MA, MA and PhD teachers separately. The estimated effect of a thousand dollar increase in the price of the upgrade is to decrease take-up by 9 to 11 percentage points. These point estimates are quite similar to those in the previous table, but the inclusion of district fixed effects has caused the standard errors to increase and rendered the estimates statistically insignificant at the 5 percent level.

As just mentioned, the coefficients using the two instrumental variables strategies are remarkably similar to one another. Note that there is no reason to have expected this to be the case. The first instrumental variables strategy would produce biased estimates if teachers sorted across districts according to their preferences for additional retirement benefits. The second
would do so if cohorts of teachers within the same district were different from one another or if pre-1998 retirement behavior led to differences in preferences across teachers within a district who had different amount of experience in 1998.

I combine the estimates in Table 2 about demand's sensitivity to price (dQ/dP) with the take-up (Q) and price information (P) for teachers with different experience levels (e.g. from Table 1) and different prices to obtain an estimate of the elasticity of demand. These estimates of the price elasticity of demand for the upgrade at generally range from -0.01 and -0.66. Demand for the upgrade is fairly inelastic.

Meanwhile, the cost to the pension fund of providing the extra benefits is positively related to take-up. Table 3 presents the results of estimating the relationship between take-up and the cost shown in equation (3). In the table, I include increasingly more explanatory variables for market and district characteristics. The estimates in the table suggest that the average cost of providing the extra retirement benefits to an employee increases as the fraction of the population purchasing the upgrade increases. For example, controlling for individual and market-level characteristics including county fixed effects, the extra retirement benefits paid to purchasers is $7,377 more than the expected cost to non-purchasers would be. The estimates range from cost increases of $7,377 to $8,358 with purchase, depending on the set of controls included in the regression, however none of the estimates are statistically different from one another. As I explain in the next section, the higher expected costs for those who purchase the upgrade are overshadowed by the overwhelming dominance of the demand curve by the average and marginal cost curves in this particular setting.30

30 This upward sloping average cost curve leads to an upward sloping marginal cost curve. As detailed by Einav and Finkelstein (2010), the upward sloping cost curve is indicative of advantageous selection in this market. An interesting avenue for future research would be to investigate the drivers of the increased costs among purchasers, i.e. do they retire earlier or live longer than the teachers who do not purchase.
Employee Valuation of Current Income Relative to Future Income

The coefficient estimates from equation (1) and equation (6) can be used to trace out the upgrade's demand, average cost, and marginal cost curves. Comparing these curves gives a sense of the cost of providing the upgrade to IPS employees relative to the value the employees place on the benefits.

To be more precise, equation (1) defines the demand curve, while equation (6) can be used to draw the average cost curve. Note that equation (1) is slightly different than the actual equations underlying the estimates in Table 2, which were equations (2) and (4). The difference is that equation (1) only includes price and the expected benefits proxy as the explanatory variables, but no other covariates. I use equation (1) because tracing out the demand curve involves a constant vertical intercept. In equations (2) and (4) any choice of the intercept would be conditional on specified values of the characteristics. With equation (1), the only characteristic to fix when drawing a demand curve is the value of $B_i$, or the expected benefits proxy. I use the average level of $B_i$, the maximum MA Degree salary, across districts to assign the intercept of the demand curve. In order to determine the marginal cost curve of providing the upgrade to IPS employees, I use the coefficient estimates from the estimation of (1) and (6) in the following formula:

$$MC(D) = \frac{\partial TC(D)}{\partial D} = \frac{\partial (AC(D)D)}{\partial D} = \frac{\partial AC(D)}{\partial D} D + AC(D).$$

Figure 3 presents an example of these curves for IPS employees with 22 to 28 years of experience in 1998. The solid line is the demand curve. The dashed and dotted lines represent the average and marginal cost curves, respectively.

---

31 Importantly, the estimated coefficients on $P_i$ and $B_i$ from estimating either (1), (2) or (4) are similar when using the cross-district instrument based on the MA Degree holders' scheduled salary.
The first and most noteworthy conclusion from the figure is that the estimated demand curve is everywhere below the estimated marginal cost curve. (Actual data tell the same story.) Standard economic theory instructs that the demand curve traces out consumers’ willingness to pay for a good. The curves plotted in Figure 3 imply that the willingness of most teachers to pay for the upgrade is much less than the cost of providing them with the extra retirement benefits.

Taking the estimated demand curve as the true curve, the highest marginal willingness to pay for the upgrade of these employees – as indicated by the point of intercept between the demand curve and the vertical axis – is just $23,000. Yet, this high valuation employee can expect to collect over $85,000 in benefits. So, my estimates suggest that even the highest value employee is willing to trade just 27 cents of current compensation for a future dollar of benefits (23,000/85,000 = 27/100).\(^{32}\) Averaging along the entire aggregate demand curve, which represents the willingness to pay in the population of workers, employees in IPS are willing to trade just 19 cents for a dollar's worth of future benefits.

Although I have quite a bit of data, little of its mass sits in the range of demand proportions below 40 percent of the population. This makes it difficult to be certain what the levels of valuation are among the high valuation teachers because I am forced to go out-of-sample to make predictions. Demand and cost curves traced only over where the majority of the data are located show that over half of the employees had a willingness to pay for the upgrade that was over $71,000 less than the expected benefits they would later collect. Comparing the

---

\(^{32}\) The estimated valuation of the highest valuation person depends on the value of the vertical intercept of the demand curve. In this setting, the demand curve was drawn conditional on an expected benefit from upgrade being proportional to the average of the maximum scheduled salary paid to MA Degree teachers. Using a higher expected benefit value results in a demand curve that lies vertically above the one in Figure 3. If the expected benefit is proportional to the highest scheduled MA Degree salary in the state, the highest valuation person is willing to pay just 31 cents on the dollar and the average valuation in the population is estimated to be 26 cents on the dollar. Alternatively, if I use the lowest scheduled salary for a MA Degree teacher in the state as the proxy for expected benefits, the highest valuation is 15 cents per dollar and the population average is 10 cents per dollar of expected benefits.
demand and marginal cost curves over the region of observed data suggests that most teachers are unwilling to trade just 16 cents of current income for a dollar of expected future compensation.

In a perfectly competitive equilibrium, the price charged would be that which equates the revenue generated with the costs of provision. In other words, the perfectly competitive price would be set where the demand and average cost curves intersect. As can be seen in the figure, in this market for extra retirement benefits, a perfectly competitive equilibrium does not exist. What is more, since average cost is always larger than the employees' willingness to pay (as represented by the demand curve), there would be no private provision of this product. The fact that employees' willingness to pay for future compensation lies everywhere below the costs of providing it also implies that these teachers are not paying for the costs of these pensions benefits with decreased wages. Thus, the worker preference theories cannot explain why we see generous defined benefits pensions in the public sector.

Placing The Estimates in Context

The estimates presented indicate public school employees would prefer $2 in current wages to $10 in PDV of annuitized wealth in their retirement package. This estimate places the discount rates of teachers at the high end of those for employees in other settings. For example, Brown, Casey and Mitchell (2008) find that many retirees report they would exchange half of their Social Security benefits for a lump sum payment even if that payment were only 75 percent of the PDV of their annuitized benefits. Similarly, Warner and Pleeter (2001) found that, when given the opportunity to choose, the vast majority of enlisted military took a lump sum payment instead of a retirement annuity worth twice as much.
As the previous studies suggest, the preference for current funds relative to annuitized payments is not uncommon in the literature. Indeed the annuity puzzle has existed for years and has been explained using people's desire for liquidity and/or diversification of assets, bequest motives, or reliance on family members for insurance (Kotlikoff and Spivak 1981). What is striking about the present study is the magnitude of money that some teachers are willing to leave on the table. Likely, teachers are willing to forgo pension income in exchange for current income in part because they are oversaturated with deferred compensation. While private employees earn $1 in compensation through employer provided retirement benefits for every hour they work, teachers earn $3. Additionally, previous research has shown that union members have 50 to 100 percent more pension wealth than their counterparts in non-unionized firms (Allen and Clark 1985). Taking spousal benefits and Social Security payments into account makes it even more likely that teachers would prefer current compensation to "over-annuitization" through additional retirement benefits. This is consistent with recent survey evidence in which teachers were more likely to report wanting additional pension dollars invested in a defined contribution type plan, rather than a defined benefit style plan (DeArmond and Goldhaber 2010).

I interpret differences in take-up relative to price as revealing information about public school employees willingness to pay for increased retirement benefits. This requires the assumption that the employees are using rational expectations to make a informed choice about those retirement benefits. Before concluding the paper, it is worthwhile considering whether this is the correct framework for evaluating the purchase decision of employees and whether a different model of employee purchase behavior would change the interpretation of the estimates.
There are features of the upgrade setting and pieces of evidence from the data consistent with interpreting these results as revealed preferences of rational informed agents. First, as discussed earlier, the employees in IPS were repeatedly presented with information about the upgrade. This negates arguments that they were unable to make an informed decision because of a lack of information. Second, the features of the upgrade purchase decision are much more simple than other investment decisions in which the literature finds people behaving in ways at odds with the rational expectations framework (e.g. Madrian and Shea 2001). For example, the price of the upgrade and the potential benefit from purchasing are more straightforward than the calculations necessary to make decisions about defined contribution pensions, which include complicated decisions about the portfolio choice. Third, recent literature suggests that people are more likely to opt-out of the default investment option when the returns to doing so are large (Beshears et al. 2010). As described earlier in the paper, the returns to purchasing the upgrade are extremely large on average. This is probably why the large majority of public school employees purchase it. Finally, the purchase of the upgrade is sensitive to the price of the upgrade. If individuals were uninformed and unable to make a rational purchase decision, there is no reason that the propensity to purchase should vary with price.

Finally, the default behavior in this setting is for teachers not to upgrade. In this context, it is possible to place an upper-bound on the fraction of the population subject to the inertia of the default. This can be done by determining the take-up rate among the groups of employees offered the upgrade at the lowest of prices. Since the instrumental variables strategy implies the prices employees face are exogenously determined, each group acts as a randomly drawn subset of the population. If there is a portion of the population driven by inertia or something other than a rational expectations informed decision making process, they will not purchase the upgrade.
even at the lowest prices. In the data, take-up rates at the lowest prices are usually around 95 percent. This means that at most 5 percent of the population would display the default behavior, regardless of the price. It is therefore safe to interpret the demand curve estimation as representing the willingness to pay for retirement benefits of the overwhelming majority of employees in the IPS.

Variation in Propensity to Purchase the Upgrade and in Price Sensitivity

A lifecycle model of labor supply under the set of rules of the TRS generates predictions about how characteristics of teachers should be related to their purchase behavior. For example, conditional on age, price and the annual value of the expected payout, teachers with more experience can expect to retire sooner, on average, than those that have less experience. (This is because the TRS allows employees to retire as long as they meet certain age and experience level requirements.) Earlier retirement means beginning to collect benefits sooner and therefore less discounting of the future stream of retirement benefit payments. Therefore, teachers with more experience in 1998 should be more likely to purchase the upgrade than their colleagues with less experience. Similarly, conditional on experience, price of the upgrade and annual benefit amount, teachers who are younger will have to delay their collection of retirement benefits longer than their older colleagues, but they may also expect to live to collect more payments.

In Column (1) of Table 5, I investigate these hypotheses about the relationship between the decision of whether to purchase the upgrade and employees' age and years of experience at the time of the offer. The regression that produces these estimates includes controls for all the possible individual characteristics like age, education, experience, price and the proxy for expected benefits. Looking first at the coefficients in the bottom panel on the set of dummy
variables for different experience levels (the omitted category is teachers with 28 years of experience, the highest level in the sample), it can be seen that take-up rates are monotonically increasing in the experience levels of the teachers up to 25 years of experience. For example, conditional on age, price and annual benefit payment, teachers with 22 years of experience in 1998 are 14.5 percentage points less likely than their counterparts with 28 years of experience to purchase the upgrade. This concords with the theoretical predictions just described.

The part of Column (1) in the top panel reports the coefficients on dummy variables representing age categories, where the omitted category is the youngest set of workers, those under age 55. The relationship between age and take-up is not monotonically increasing, as it is with experience. Rather, take-up rates are highest among teachers ages 56 to 60 and lowest among teachers less than 56 and older than 71. This parabolic relationship between age and take-up may be reflective of the competing forces of the positive relationship between age and mortality timing (which would decrease take-up for older workers) and the negative relationship between age and time until retirement (which would increase take-up for older workers).

Although the results are not reported in Column (1) of Table 4, I also included interaction terms between price and the age dummies and price and the experience dummies. The intent was to determine whether teachers of different age or experience levels were more or less sensitive to price when making decisions about upgrade purchase. All of the coefficient estimates on these interactions were quite small and statistically insignificant, leading me to conclude that teachers of different ages or experience levels are no more or less sensitive to price.\textsuperscript{33}

\textsuperscript{33} The coefficient on the price term for the omitted group - teachers under age 56 with 28 years of experience - is essentially unchanged from the earlier estimates from regressions without these interaction terms.
Conditional on age, experience, price and the annual retirement benefit amount, a teacher's level of education does not change the timing of benefit receipt, etc. Therefore in a simple lifecycle model, teachers with different levels of educational attainment would be expected to take-up the upgrade at similar rates. However, if there are information constraints or complexities, it may be the case that teachers with higher levels of education are better able to navigate the retirement system and determine the upgrade was a good deal in expectation. Column (2) of Table 4 presents coefficient estimates from the demand equation when I include dummies for MA Degree holders and PhD holders and their interactions with price. None of the coefficients are statistically different from zero, suggesting either that these types of information constraints do not play a role or that MA and PhD attainment do not proxy well for the skills needed to overcome any information constraints.

The Role of Union Leadership

Despite the finding that, on average, teachers do not value the increased pension benefits, it could still be the case that large defined benefit pension plans exist because of the preferences of employees. Consider, for example, a model in which the government must negotiate with union leadership about the compensation package for public employees. If the union leadership has a strong preference for pension benefits or gives the preferences of those who value pensions the most weight in its objective function, it may negotiate for larger pension benefits than is optimal from the perspective of the marginal worker. Positing that this type of difference between union leadership and members exists has been popular in the literature (Freeman 1986), but does not yet have much empirical support.
At the crux of this set of theories is the idea that some employees value future compensation more than current compensation. The insurance nature of the upgrade implies that the upgrade's value will vary, even across teachers who are offered it at the same price (Einav, Finkelstein and Cullen 2010). If some employees value future compensation more than current compensation, and those employees become the managers of the union, it is possible the observed compensation package would differ from the compensation package optimal for the average or marginal worker. However, the demand and cost curve analysis suggests that none of the employees of IPS value the extra benefits at their expected value. The employee with the highest valuation of future benefits places a value of just 66 cents of current compensation for each expected dollar of future compensation. It is therefore unlikely that any model of union representation, whether it be a median voter model or a model in which the union representative only takes the highest valuation employees into account, can explain the existence of such generous defined benefit pensions in the public sector. The test is only suggestive, however, because it relies on out-of-sample predictions and because I cannot rule out other intervening mechanisms.

6. Conclusion

In this paper, I empirically test whether teachers and other public school employees value pension benefits at the margin as much as it costs to provide them. Almost all states use a form of defined benefit program similar to the one studied as their main pension system for public employees. By guaranteeing a large fraction of pre-retirement salary until death, these programs offer generous deferred compensation to teachers and other public employees. Despite the fact

---

34 Otherwise, the union leadership would be maximizing an objective function orthogonal to its members’ preferences. While possible, this does not seem rational or the basis for the popular argument of union leadership's involvement in generous public sector pensions.
that 89 percent of teachers and 84 percent of public employees are in these types of defined benefit pension plans, the motivation for using them remains unclear.

The most straightforward argument for the large deferred compensation packages offered through defined benefit pension programs is that public employees prefer the guaranteed stream of income they provide to an equivalent increase in current wages. An opportunity offered to public school employees allows me to estimate the willingness-to-pay for increases in benefits and the cost of supplying them. I show that teachers' valuation of the increased pension benefits was much less than their cost. The majority of IPS employees have a willingness-to-pay for the upgrade that is more than $70,000 less than what the upgrade is worth to them in expectation. Said differently, teachers would prefer $2.00 increase in current wages to a $10 increase in the PDV of annuitized wealth in their retirement package.

This research provides new information critical to public policy decision making. Forty-six states have maintained defined benefit pension plans as the main provider of retirement benefits to their state employees (GAO 2007). These plans are becoming quite costly. In 2010, Illinois was on track to spend $9.6 billion from 17 of its pension funds for state and local employees. Meanwhile, Illinois is no more generous than many other states. The 1998 pension 'sweetener' brought pension benefit levels to the median of what they were in other states where teachers do not receive Social Security. Nor was Illinois the only state that decided to provide more generous benefits. In the 1990s, as high stock market returns improved the balance sheets of pension funds, many states, including California, New York and Florida, increased the generosity of retirement benefits for their employees. The margin at which I estimate pension benefit valuation in Illinois is quite representative of the margin on which many other states are currently making decisions about increasing or decreasing their pension benefits.

In this context, the main finding of this paper, that the majority of IPS employees value their pension benefits at about 19 cents on the dollar, has two important implications. First, it suggests a possible Pareto-improving and politically feasible solution to the current inability of states to pay their promised pension benefits to public employees. Governments could offer to buy back pension benefits from teachers and other public sector employees. If the results here generalize, governments may be able to buy back promised employee pension benefits, or at least some of these promised benefits, for as little as twenty cents on the dollar. Doing so would draw down the pension obligations of governments both significantly and immediately, rather than waiting for a reduction in benefits to take effect years in the future.

To illustrate the potential of a buy-back program to resolve state pension funding problems, it is instructive to consider what fraction of the states’ total pension funding deficits could be eliminated with such a program. I use data on the assets and liabilities of state employee pension funds, specifically the liabilities discounted at the risk-free rate of return, from Novy-Marx and Rauh (2009). I assume the buy-back plan would involve states offering to buy-back 30 percent of their current pension liabilities for some set price. I choose 30 percent because that is approximately the level of pension benefit increases experienced by teachers in my sample due to the upgrade.

At the low buy-back rate of 20 cents for each dollar of pension benefits, this 30 percent buy-back plan would decrease the amount of underfunding by 40 percent, or nearly $1.3 trillion. Of course, just because employees were only willing to pay about 19 cents, on average, for an increase in pension benefits of $1, they may not be willing to sell the benefits back at such a low price.36 A buy-back plan offering public employees 50 cents for each dollar of the 30 percent of

---

36 In a different setting, Chalmers and Reuter (2011) find that only 15 percent of public employees opt for a lump sum payment in exchange for part of their promised retirement annuity. However, 90 percent of the retirees they
their total benefits would still decrease the underfunding of public employee pensions systems in the U.S. by 25 percent, or about $775 billion. Of course, there may be other details that need to be taken into account by policymakers implementing such a program (such as how to raise the capital to pay for the buy-backs), but with such large potential to reduce the burden on taxpayers while also making public sector employees better off the option certainly deserves more exploration.

Second, by showing that workers value the increased benefits at only a fraction of the cost of providing them, the results of this paper offer clear evidence against a worker preference rationale for such generous defined benefit pension packages. If teachers were able to choose between higher pension benefits and equivalent increases in current salary, as is often assumed, they would choose higher current salaries. This evidence suggests that workers are not paying for the generous pension benefits with decreases in current wages, as is also often assumed. This should therefore turn our attention to other explanations for the existence of these generous benefits.

One possible alternative explanation for the use of defined benefit pension plans is that there is a subset of employees who prefer such benefits to current compensation. If these employees become union leadership, they may negotiate more generous pension benefits than the marginal worker would prefer. I show that even employees who are willing to pay the highest prices for the future retirement benefits value these benefits at just 27 cents on the dollar. It does not appear the differences in valuation across teachers are able to explain the existence of generous pension benefits in the public sector.

---

consider are in some form of a defined contribution setting so the results may not generalize to more traditional defined benefit settings.
The inability of the first two classes of theories to explain the use of generous defined benefit pension plans for public employees leads me to turn to a last group of models. These models rely on the political nature of public pensions to drive a wedge between the actual costs of pension benefits and the perceived costs of those benefits used to make decisions about compensation for public employees. Since in these models politicians care mostly about being elected (or reelected), they are concerned more with voters' perceived costs of pension benefits relative to current salaries, rather than the actual relative costs of these two compensation mechanisms. If for some reason the costs of pension benefits are less salient to voters (perhaps because they are less publicized), the perceived costs of pensions may be much lower than the actual costs of pensions. The true costs of pension benefits also may not be relevant for politicians if their constituents are likely to move out of the jurisdiction before the payment comes due (Inman 1982). The difference between real and perceived costs is exacerbated in a system that allows underfunding of public pensions, as most states do today. Understanding the role of salience, myopia and/or residential transitions in the determination of the lifetime compensation mix for public employees is an important avenue for future research.
References


Illinois Economic and Fiscal Commission. 1998. *Illinois Bond Reference*


Table 1. Annual Benefit Information Before and After the 1998 Formula Change, as Reported to Employees of IPS

<table>
<thead>
<tr>
<th>Years of Service</th>
<th>4 Step Formula</th>
<th>2.2 Formula</th>
<th>Percent Increase</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>16.7%</td>
<td>22.0%</td>
<td>31.7%</td>
</tr>
<tr>
<td>20</td>
<td>35.7%</td>
<td>44.0%</td>
<td>23.2%</td>
</tr>
<tr>
<td>30</td>
<td>58.7%</td>
<td>66.0%</td>
<td>16.4%</td>
</tr>
<tr>
<td>34</td>
<td>65.9%</td>
<td>75%</td>
<td>13.8%</td>
</tr>
</tbody>
</table>

Note: Recreated by the author from materials provided by the TRS to IPS employees at the time of the formula change.
Table 2. Take-up Rates, Retirement Rates, Prices, and Costs by Level of Experience in 1998

<table>
<thead>
<tr>
<th>Years of Experience in 1998</th>
<th>Fraction Who Retire by 2009</th>
<th>Fraction Who Purchase Upgrade by 2009</th>
<th>Mean Price ($1,000s)</th>
<th>Mean Cost ($1,000s)</th>
<th>Number of Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>22</td>
<td>0.60</td>
<td>0.70</td>
<td>14.02</td>
<td>83.83</td>
<td>2,473</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4.08</td>
<td>31.19</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>0.67</td>
<td>0.71</td>
<td>14.59</td>
<td>87.43</td>
<td>2,677</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4.35</td>
<td>34.15</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>0.78</td>
<td>0.76</td>
<td>14.89</td>
<td>90.29</td>
<td>3,081</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4.38</td>
<td>34.81</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>0.83</td>
<td>0.78</td>
<td>15.24</td>
<td>94.17</td>
<td>3,213</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4.48</td>
<td>36.87</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>0.88</td>
<td>0.77</td>
<td>15.37</td>
<td>97.88</td>
<td>2,984</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4.46</td>
<td>38.31</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>0.89</td>
<td>0.74</td>
<td>15.72</td>
<td>99.42</td>
<td>2,763</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4.71</td>
<td>41.16</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>0.91</td>
<td>0.73</td>
<td>15.89</td>
<td>99.34</td>
<td>2,814</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4.63</td>
<td>39.63</td>
<td></td>
</tr>
<tr>
<td>Whole Sample</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>20,005</td>
</tr>
</tbody>
</table>

Note: 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). The cost of the upgrade (in thousands of $2010) is the present value in 1998 of the extra retirement benefits paid out as of 2009 as explained in the text.
Table 3. Estimates of the Relationship Between Demand (Purchase of the Upgrade) and Price

<table>
<thead>
<tr>
<th>Panel A.</th>
<th>OLS</th>
<th>IV Cross-District</th>
<th>IV Within-District</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>Price ($1,000)</td>
<td>0.019***</td>
<td>-0.144***</td>
<td>-0.152***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.011)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>Benefit (MA Maximum Salary)</td>
<td>0.000</td>
<td>0.041***</td>
<td>0.040***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.003)</td>
<td>(0.003)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B.</th>
<th>MA Degree Scheduled Salary as Instrument</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS</td>
</tr>
<tr>
<td>Price ($1,000)</td>
<td>-0.087***</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
</tr>
<tr>
<td>Benefit (MA Maximum Salary)</td>
<td>0.026***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C.</th>
<th>PhD Degree Scheduled Salary as Instrument</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS</td>
</tr>
<tr>
<td>Price ($1,000)</td>
<td>-0.083***</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
</tr>
<tr>
<td>Benefit (MA Maximum Salary)</td>
<td>0.024***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>PUMA Characteristics</th>
<th>County Fixed Effects</th>
<th>County Characteristics</th>
<th>District Characteristics</th>
<th>District Fixed Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A.</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panel B.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panel C.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Based on the author's calculations using data from the TSR and TRS. Sample includes all teachers eligible for the upgrade with between 22 and 28 years of experience in 1998. Each column-panel set represents estimates from a separate regression. The dependent variable is a dummy variable equal to one if the employee purchased the upgrade. As indicated, estimates are from OLS or 2SLS estimation. In the columns labeled IV Cross-District, the actual price paid for the annuity is instrumented with the beginning scheduled salary paid to teachers with a BA in one's district of employment in 1998 (measured in thousands of $2010). In the columns labeled IV Within District, the instrument used is the scheduled salary paid to a teacher in the district with the same amount of experience as the teacher. Each panel presents estimates using the scheduled salary for a different group of teachers. All specifications include individual characteristics. Market-level characteristics are added in groups as described in the text. Standard errors (in parentheses) are clustered at the district level; *** p<0.01, ** p<0.05, * p<0.1.
Table 4. Instrumental Variables Estimates of the Relationship Between Annuity Total Cost and Take-up

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Take-up</td>
<td>7.675***</td>
<td>8.358***</td>
<td>7.377***</td>
</tr>
<tr>
<td></td>
<td>(0.491)</td>
<td>(0.421)</td>
<td>(0.420)</td>
</tr>
<tr>
<td>PUMA Characteristics</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>County Fixed Effects</td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>County Characteristics</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>District Characteristics</td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

Note: Based on the author's calculations using data from the TSR and TRS. Sample includes teachers with 22 to 28 years of experience in 1998. The dependent variable is the expected cost (at time of purchase) to the TRS if the employee purchased the upgrade, measured in $1,000. All specifications include individual characteristics. Market-level characteristics are added in groups as described in the text. Standard errors (in parentheses) are clustered at the district level; *** p<0.01, ** p<0.05, * p<0.1.
Table 5. Estimates of the Relationship Between Demand (Purchase of the Upgrade), Employee Characteristics and Price

<table>
<thead>
<tr>
<th>Age</th>
<th>Education</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between 56 and 60</td>
<td>MA</td>
<td>0.120***</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.013)</td>
<td>(0.047)</td>
</tr>
<tr>
<td>Between 61 and 65</td>
<td>MA X Price</td>
<td>0.070***</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.012)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Between 66 and 70</td>
<td>PhD</td>
<td>-0.046***</td>
<td>1.095</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.016)</td>
<td>(0.685)</td>
</tr>
<tr>
<td>Over 71</td>
<td>PhD X Price</td>
<td>-0.145***</td>
<td>-0.036</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.023)</td>
<td>(0.032)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Experience</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>22 Years</td>
<td></td>
<td>-0.145***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.021)</td>
<td></td>
</tr>
<tr>
<td>23 Years</td>
<td></td>
<td>-0.123***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.016)</td>
<td></td>
</tr>
<tr>
<td>24 Years</td>
<td></td>
<td>-0.063***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.017)</td>
<td></td>
</tr>
<tr>
<td>25 Years</td>
<td></td>
<td>-0.022</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.017)</td>
<td></td>
</tr>
<tr>
<td>26 Years</td>
<td></td>
<td>-0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.019)</td>
<td></td>
</tr>
<tr>
<td>27 Years</td>
<td></td>
<td>-0.005</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.020)</td>
<td></td>
</tr>
</tbody>
</table>

Note: Based on the author's calculations using data from the TSR and TRS. Sample includes all teachers eligible for the upgrade with between 22 and 28 years of experience in 1998. Each column represents estimates from a separate regression. The dependent variable is a dummy variable equal to one if the employee purchased the upgrade. Estimates are from 2SLS estimation where the actual price paid for the annuity is instrumented with the beginning scheduled salary paid to teachers with a MA in one’s district or region of employment in 1998 (measured in thousands of $2010). The omitted age category is less than 56 years of age. The omitted experience category is 28 years of experience. The omitted education category is all other levels of education (the majority are BA Degree holders). All specifications include individual and market-level characteristics described in the text. Standard errors (in parentheses) are clustered at the district level; *** p<0.01, ** p<0.05, * p<0.1.
Figure 1. Teacher Retirement Benefit Formulas in Illinois, Before and After 1998

Note: The solid line depicts the flat rate formula used for calculating pension benefits for years of service accrued after 1998. The dashed line illustrates the graduated pension formula used before 1998.
Figure 2. Upgrade Details: Back of the Envelope Price & Value Calculations as Percent of Salary

Note: Calculations in the figure assume no discounting. Expected lifetime retirement benefits presented for a 60 year old teacher who immediately retires and expects to collect payments for 15 years.
Figure 3. Estimated Demand and Cost Curves for Employees with 22 to 28 Years of Experience in 1998

Note: Based on the author's calculations using the TRS and TSR.
Appendix A. Worksheet Examples of Information Provided by TRS to IPS Employees for the Purpose of Assessing Upgrade Purchase Decisions

**Calculating Cost to Upgrade**

Enter total years under old four-step formula or 20, whichever is less

Minus years worked under new formula divided by 3 (drop any decimals)

Equals number of years to upgrade

Multiply by 1% (.01)

Multiply by highest salary earned during last four years (do not include year of application)

COST TO UPGRADE

( )

______

years

X

% 

X$

$

Option: You do not upgrade to 2.2 and have less than 30 years of service credit as of July 1, 1998.

From the retirement formula table, find the total accrual percentage for years of service prior to 7/1/98 (Example: 20 years = .3570) _________ (1)

Years of service after 7/1/98 x .022 _________ (2)

Total years of service factor (add lines 1 and 2) _________

Multiply by average salary (from page 5) _________

Annual annuity _________

Divide by 12 (months) _________

Monthly annuity _________

Option: You upgrade all service credit to 2.2.

Average salary (page 5) $_______

Multiply total years of service by .022 _________

Annual annuity _________

Divide by 12 (months) _________

Monthly annuity _________
Appendix B. Data Appendix

A.1. Expected Rate of Return Calculation

I use projected mortality probabilities calculated from the cohort life tables underlying the 2007 Social Security Administration 2007 Trustees Report to create a conservative estimate of the expected payout. Thanks to Gopi Shah Goda for sharing this data with me. The SSA provides mortality probabilities under three scenarios ranging from more to less conservative assumptions. I use the intermediate set of probabilities for women, since 80 percent of public school teachers are women. I also use retirement probabilities estimated from the data. The new formula changed the amount of experience required to reach the maximum pension benefit. The data show teachers’ retirement timing changed concurrently. Therefore, I think the retirement probabilities based on pre-program behavior are a poor approximation. Results using pre-program retirement probabilities show expected rates of return that are similar. Finally, I use an employee’s highest observed salary to calculate the stream of benefit payments and incorporate the 3 percent cost-of-living adjustment built into pension payments by the TRS.

A.2. Chicago Teachers Pension Fund

Public school teachers in Chicago are covered by a separate pension system from the TRS, called the Chicago Teacher’s Pension Fund (CTPF). The state legislature sets most of the rules concerning retirement contributions and benefits, so the CTPF and TRS offer identical options to employees and employees in both systems were offered the 2.2 upgrade. I have requested information from the CTPF that is identical in form to what I have received from the TRS. In the meantime, however, I proceed without this information. The implication of this is that I may be misclassifying the upgrade purchase decision for teachers who, because they retire with CTPF, purchased the upgrade with the CTPF. The TRS and CTPF suggest that a teacher collect pension benefits from the system in which they recorded the most service and apply for reciprocal service benefits from the other system. The more service a teacher had in Chicago, therefore, the more likely she is to purchase the upgrade from and retire with the CTPF. Because of this, I include only teachers who either never taught in Chicago or who have more than two years of service recorded in the TRS in my estimation sample. This excludes the most likely candidates for retirement in CTPF, for whom I may be missing information. In order to be sure the misclassification of the purchase decision is not driving the results, I repeat the estimation excluding teachers who spent any time teaching in CPS and the results are qualitatively similar.

A.3. A Note on Inflating Monetary Variables

I inflate the prices of the upgrade to 2010 dollars using the Consumer Price Index for the appropriate year, e.g. for the year in which the relevant salary was earned. Since my salary information is reported on a school-year basis, I use the inflation measure corresponding to the spring of the school year in which a salary was earned. The price of the upgrade is based on the highest salary earned in the last four consecutive years of earnings prior to purchase. For some teachers, this will be her salary in 1998, so that the price of the upgrade is inflated to 2010 dollars using the relevant CPI measure for 1998. Other teachers, however, had their last year of service much earlier. Because they had not yet begun collecting benefits in 1998, they were still eligible for the upgrade. The price charged to them was based on the highest nominal salary
earned in the last four years of service before purchase. For these teachers I calculate the price of the upgrade based on this nominal salary measure and inflate it using the relevant inflation measure for 1998. For example, if a teacher's last stint of teaching before retirement in 2000 was in 1980, I use her nominal 1980 salary to price the upgrade, but inflate the price using the CPI for 1998. Finally, approximately 38 percent of teachers who purchased the upgrade did not do so immediately. The price to these teachers is inflated based on the relevant CPI measure for the year in which they purchase.