How Much Do Public School Teachers Value Their Retirement Benefits?

Maria D. Fitzpatrick
Stanford Institute for Economic Policy Research
Stanford University

This draft: November 15, 2010
Abstract

Public employers contribute three times more to retirement benefits per hour worked than their counterparts in the private sector. Given the potential burden on taxpayers of underfunded pensions, efficient compensation mechanisms for teachers and other public employees are an imperative. The question addressed in this research is whether teachers value deferred compensation at the cost of providing it.

One explanation offered for the heavy use of deferred compensation in the public sector is that these packages attract high quality employees. These employees may have low discount rates or, for some other reason, prefer to trade off lower current wages for guaranteed future compensation. At what rate teachers are willing to forgo current wages for future compensation is difficult to measure because wages and pension benefits both move cyclically, making it hard to disentangle differences in valuations of each. The introduction of the opportunity provided in 1998 to Illinois public school employees to purchase additional pension benefits allows me to estimate employees' demand for retirement benefits relative to the cost of providing them.

The price charged for this product is directly proportional to a teacher's current salary. Because a teacher's salary is likely correlated with her health and other factors affecting demand, I create a simulated instrument for prices using exogenous variation in program rules. The instrument uses variation in salary schedules at the district level, purging unwanted individual-level variation in salaries that might be correlated with employee take-up. Identification follows from district differences in teacher salary schedules prior to the product's introduction. The fact that conditioning on observed demographic factors and market characteristics does not change the estimate instills confidence that the relationship between price and demand is identified from exogenous variation in price.

The results show that the majority of Illinois public school teachers are willing to pay just 17 cents for a dollar increase in the present value of expected retirement benefits. The findings therefore suggest substantial inefficiency in compensation as the public cost of deferred compensation exceeds its value to employees. The results of this study suggest a Pareto-improving policy solution: governments can offer to buy back promised pension benefits for just a fraction of their expected present value.
1. Introduction

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.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). Economists have long been interested in the motives for the use of these types 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. By shedding light on employee valuation, I can differentiate between the models used to explain the use of generous deferred compensation packages in the public sector.

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 guaranteed future compensation.³ Whether teachers are willing to forgo current wages for future compensation is difficult to test because wages and pension benefits generally both move

---

¹ 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).
³ Risk aversion does not necessarily imply that these types of workers will be attracted by 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.
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 cost of providing those 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 an upward-sloping demand curve. I create a simulated instrument for price using exogenous variation in program rules. The instrument also uses variation in salary schedules at the district level, purging unwanted individual-level variation in prices that might be correlated with employee take-up. Identification follows from district differences in teacher salary schedules prior to the product's introduction. The instrumental variables results vary little with a host of market-level controls and county fixed effects. The fact that conditioning on observed demographic factors and market characteristics does not statistically change the coefficients instills 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 $50,000 less per teacher than the cost of providing the benefits. Said differently, these employees are willing to trade just 17 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.

Another set of prominent theories explaining the heavy use of defined benefit pension plans in the public sector relies on the fact that wages are set through collective bargaining, as is the case for many teachers (Freeman 1986). The objective function of union leadership, who does the bargaining, is less obvious than that of the individual worker (who maximizes only her own lifetime utility). Because of their senior status or for other reasons, union leadership may weigh member utility in a way that results in stronger preferences for deferred compensation relative to current compensation than the marginal worker’s preference.

To test whether the political representation role of union leadership can be driving generous public sector pension benefits, I look for differences in the valuation employees place on current compensation relative to future benefits.\(^4\) 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.

---

\(^4\) This is akin to using new techniques adapted from the literature on adverse selection in insurance markets (Einav, Finkelstein and Cullen 2010).
Other mechanisms are reflected in models that incorporate the role of politicians, who are motivated by the desire to get elected (or reelected), in public sector wage determination. If voters are myopic or if for some reason the costs of pension benefits are less salient to them (perhaps because they are less publicized), then the perceived costs of pensions at the time they are granted may be much lower than the actual costs of pensions. For example, the true costs of pension benefits 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.

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 promised employee pension benefits, or at least parts of these promised benefits, may be recouped for as little as twenty cents on the dollar. As discussed in the conclusion, a successful buyback 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). The simulated instruments framework and results are described in Section 4. Section 5 concludes with a discussion of alternative models and their importance for future research and policy.

2. Illinois Teacher Retirement Benefits and the Annuity

1998 Legislative Change

In the late 1990s, the booming economy and stock markets led to surpluses in state budgets and pension funds (relative to their mandated levels). At the time, Illinois had the least generous pension system of any state that did not allow its teachers to participate in Social Security. 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. 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 prior to 1998 were 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. The one-time fee depends on the teacher's service at the initial offering of the upgrade, future service and salary-at-time-of-purchase; the annual increase in retirement benefits depends on the teacher's service at the time of the initial offering, future service and her end-of-career salary. 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.

**Payout from Purchasing the Upgrade**

For service accrued prior to 1998, the statutory formula for calculating annual retirement benefit is nonlinear. 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 contributes 2.2 percent of a teacher's end-of-career salary to the annual benefit amount. 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.

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 (the vertical difference between the two lines at a given level of service) times a teacher's end-of-career salary.

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. The relevant salary for the upgrade fee, what I will call the salary-at-time-of-purchase, is one's highest annual salary

---

5 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.
6 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.
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.

Price of the Upgrade

The maximum upgrade fee is 20 percent of the salary-at-time-of-purchase. 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.

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 1998.7

Value of the Investment

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

---

7 Teachers with little experience in 1998 who face 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 20 years of experience in 1998, since I am most likely to observe their purchase decisions in my censored data (see Section 3).
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. If she retires right away, she will earn her money back on the investment (and then some) within 3 years (even with a reasonable discount rate to take account of the fact that the retirement benefits are not paid until the future).

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, it is also the case that most teachers can expect to collect many years of payments (the average in my censored data is 14 years) and experience real wage growth. These factors make it likely that, 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.\(^8\) 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. More than half of eligible teachers had expected IRR above the 15 percent Real S&P Index Return in the 1990s (Fama and French 2001).

\(^8\) 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.
Information Provided to Teachers

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 opportunity, they could not have purchased it, regardless of its value. However, in this setting it is unlikely that teachers were unaware of the upgrade or the extra retirement benefits it would provide. This is because the 2.2 formula change and the upgrade were widely publicized. Both were the subject of many news reports and newspaper articles in 1998. Also, after the formula change, detailed mailings were sent out to IPS employees documenting how the upgrade could be calculated and providing assistance to employees in determining the price and benefits of the upgrade for themselves. It is therefore unlikely that a lack of information prevents take-up for many employees.

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 had several possible purchase mechanisms through which payments could have been spread over as many as five years. Also, the ability to have the payment for the upgrade deducted from one's paycheck, deducted from one's retirement benefit check (if retiring immediately) or even transferred from other retirement accounts lessens concerns about binding credit 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, 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 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.

---

9 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.
The reported compensation reported 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 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

---

10 Employees of Chicago Public Schools are covered by the Chicago Teachers Pension Fund. For more information, see the Data Appendix.
information, I find matches for 97 percent of teachers in the TRS data who were eligible for the upgrade and were vested. I include all teachers that match between the two datasets in the sample.\(^{11}\)

**Defining the Prices and Costs**

The terms of the annuity contract depend on the amount of service a teacher had in 1998, when the upgrade option was first offered. Recall that the additional retirement benefits paid to an upgrade purchaser are directly proportional to her service at the time of the initial offering and her end-of-career salary. Similarly, the price of the extra retirement benefits depends on the employee's service as of 1998, salary at time of purchase and, in the end, her service accrued after 1998. It is therefore the teacher's service accrued in 1998 that defines the terms of the upgrade contract. Regardless of when the teacher purchases the upgrade, how long the teacher works or what her end-of-career salary is, both the price and the payout of the upgrade will be directly proportional to her years of service in 1998.\(^ {12}\)

Two crucial measures for the analysis are the price of the upgrade and how much it costs to provide the extra retirement benefits purchased. The price variable is defined 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.

---

\(^{11}\) 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.

\(^{12}\) As an example, consider the difference between two identical teachers, one has 10 years of experience in IPS in 1998 and the other has 11. Assume their salaries in 1998 were such that, initially, the price of the annuity offered to them is identical. If we further assume that the two teachers have the same end-of-career salary, amount of experience at retirement and length of life, the teacher with 11 years of experience would receive more total payout from the upgrade offered to her than the teacher with 10 years of experience would from the upgrade offered her.\(^ {12}\) The two teachers were offered different products. Each set of teachers with the same experience is, in essence, its own market. Therefore, I perform the estimation separately for each level of experience in 1998.
For teachers that did not purchase the upgrade, the price is based on their highest salary earned as of 1998. Additionally, I inflate prices to 2010 dollars as described in the Appendix.

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 the upgrade.) IPS employees make their purchase decisions without observing their actual payouts 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 her 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. The present value of the total cost of providing the upgrade to an employee is therefore

\[ \text{Cost}_i = \sum_{j=\text{Age}_{1998}}^{19} \left[ \frac{(1.03)^j A_i M_j R_j}{(1.05)^j} \right] \]

The annual additional benefit, A, is essentially weighted by the probability the employee collects a payment in that particular year, which in turn depends on the probability of surviving, M, and.

---

13 Note that neither of these definitions of price takes future years of service into account. Doing so makes little difference to the results.
the probability of retiring, \( R \). 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 1 presents information on IPS employees' take-up, price, cost and retirement decisions. In Table 1, I focus on employees with 25 to 28 years of experience in 1998, but the results are generalizable to all cohorts.\(^{15}\) Each row represents employees with the amount of creditable service accrued in 1998 indicated by the row 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 censoring of the data leads me to focus on those employees with the highest experience and the largest eligible populations.

The third column of the table reports the fraction of teachers who have purchased the annuity by 2009. The take-up rate ranges from 78 percent for teachers with 25 years of experience in 1998 to 72 percent for teachers with 28 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

\(^{15}\) I focus on the cohorts with 25 to 28 years of experience for brevity's sake. Older cohorts are a selected sample because members of their cohort had already begun retiring. Also, 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).
purchase the upgrade because they anticipated staying employed long enough to reach the maximum benefit.

Columns 4 and 5 of Table 1 illustrate the prices offered to and paid by employees, respectively, while columns 6 and 7 present information about the costs to the pension fund of providing the upgrade to the whole set of employees and to the purchasers, respectively. 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 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. Estimating Demand for and Costs of Upgraded Pension Benefits

Underlying Theoretical Framework

Underlying the estimation of the demand and cost curves described below is an environment where employees with rational expectations about the future are choosing between two pension amounts to maximize their utility subject to a lifetime budget constraint.\(^{16}\) The first pension level available to employees is a regular 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 additional retirement benefits, \(h\).

\(^{16}\) 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.
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 \( i \) of purchasing the upgrade is \( U^h(X_i, p) \) and the utility to consumer of type \( i \) of not purchasing the upgrade is \( U^l(X_i) \).

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 \( U^h(X_i, p) > U^l(X_i) \). For each employee, there is a reservation price, which is the highest price she is willing to pay to purchase the retirement benefits. 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 \( P \).

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.

**Instrumental Variables Estimation of Demand**

In order to determine employees' willingness to pay, I estimate the aggregate demand curve for the upgrade. 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 of the upgrade, $P_i$:

$$D_i = \beta_0 + \beta_1 P_i + \epsilon_i.$$  \hspace{1cm} (2)

Recall that the price of the upgrade is one percent of a teacher's salary-at-time-of-purchase per year of service accrued in 1998. Since I estimate the demand equation separately by level of experience in 1998, the price of the upgrade to teacher $i$ in district $d$ of county $c$ at time of purchase $t$ is defined by one's salary-at-time-of-purchase:

$$P_{idct} = 0.01 \times \text{Salary}_{idct}.$$

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. by components of the error term in equations (2). 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. 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 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. For this reason, 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:

\[ P_{lart} = a_0 + a_1 \overline{SalaryStatistics}_{d1998} + \omega_i. \]

My instrument for the price of the upgrade is a set of order statistics from the 1998 salary schedule offered to teachers with a Bachelor's Degree in the same district as the employee. 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 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 be sure the instrument I have created is exogenous to the error term in equation (2), I also include measures of employee characteristics as of 1998, e.g. Master's Degree attainment, a polynomial in age, controls for her position in the school, etc. Additionally, in some specifications, I include market-level characteristics 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.\(^{17}\)

---

\(^{17}\) These district-level characteristics come from tabulations done by the U.S. Census and provided for the National Center for Education Statistics.
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.

Finally, to the extent that teachers with certain characteristics may be drawn to reside and be employed in different areas within Illinois, estimates of equations (2) and (3) may produce biased results. For example, if teachers with the highest valuation of retirement benefits are all drawn to the highest salary areas, this would produce a more positive relationship between price and demand than the true relationship. Therefore, in some specifications, I also include county fixed effects. When these are included, identification of the demand curves stems from within-county differences in take-up and salary schedules. The assumption in these specifications is that teachers with similar characteristics and family circumstances are likely to be assigned at random to different 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.

*Estimation Strategy: Costs*

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, \( D_i \):

---

18 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.
\[ C_i = \gamma_0 + \gamma_1 D_i + \nu_i. \] (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 this 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 would be 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.\(^{19}\) This method would allow identification of 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.

Estimates of the Demand and Cost Schedules

Standard economic theory predicts demand should be negatively related to price. Table 2 presents both OLS and two-stage least squares (2SLS) estimates separately for each level of experience in 1998. For each experience level, the OLS results are in the first row and the 2SLS estimates are in the second row. Moving across the columns for each experience level, I include additional controls as indicated at the bottom of the table. Because of space constraints, I discuss

\(^{19}\) 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. The answers are similar to those using individual data.
the results in the context of teachers with 28 years of experience in 1998. The results are similar for teachers with other amounts of experience in 1998.

The OLS estimates of the relationship between price and demand are positive, suggesting an upward sloping demand curve. For example, for teachers with 28 years of experience, the OLS results in the first column suggest that employees are 1.4 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, suggesting the positive relationship is not driven by older employees with higher salaries or the administrative employees who also have higher salaries.

Even once all of the market- and district-level controls are included, the OLS estimates imply that take-up of employees with 28 years of experience increases by 2.1 percentage points with a $1,000 increase in price (column 4). That the estimated relationship between price and take-up changes little with these controls suggests public school employees are not sorting across markets based on their propensity for take-up. The estimates change little when either county fixed effects (column 3) or district fixed effects are added (not shown, but available upon request). Based on these OLS results, teachers do not appear to be sorting across districts or counties in ways that are correlated with their demand for increased retirement benefits.

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. However, it would
not be true that the teachers who face higher prices were more likely to purchase the upgrade because they face higher prices. 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 2SLS strategy described in the previous section. In Table 2, the exogenous instrument excluded from the second stage is the maximum salary paid to BA holding teachers in the district where a teacher was employed in 1998. The 2SLS estimates accord with economic theory: demand decreases as price increases. For example, teachers with 28 years of experience are 1.7 percentage points less likely to purchase the extra retirement benefits when the cost increases by $1,000. When county fixed effects and other market-level controls are added, the results are slightly more negative. From the last column of results for 28 year employees, we see they are 2.9 percentage points less likely to purchase the extra retirement benefits when the cost increases by $1,000. However, the results are not statistically different when market-level controls are included. This suggests sorting of teachers across counties is not systematically related to their demand for deferred compensation.

I combine the estimates in Table 2 about demand's sensitivity to price \((dQ/dP)\) with the average take-up \((Q)\) and price information \((P)\) for each experience level group from Table 1 to obtain an estimate of the elasticity of demand. These estimates of the price elasticity of demand for the upgrade at the average take-up and price level generally range from -0.24 and -0.59. Demand for the upgrade is fairly inelastic and does not vary systematically with experience.

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). The regression is estimated at the individual-level, but can
be translated to the aggregate level by interpreting the coefficient on take-up as the fraction that purchase the upgrade. 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, the extra retirement benefits paid to purchasers with 28 years of experience in 1998 is $8,327 more than the expected cost to non-purchasers would be. The estimates range from cost increases of $7,591 to $10,502 with purchase, depending on teachers' experience in 1998 and the set of controls included in the regression. The relationship between the demand curve and the marginal cost curve allows me to say something about the valuation of teachers for current income relative to future income. Before describing this relationship, I detail several specification checks to be sure my choice of instrument and functional form do not drive the results.

Controlling for End-of-Career Salary

The price a teacher is charged for the upgrade is proportional to her current salary while the benefit she receives from purchasing the upgrade is proportional to her end-of-career salary. 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 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. In this sense, teachers who face higher prices may be more likely to purchase the upgrade because they expect to collect more from the upgrade. 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.\textsuperscript{20} The maximum scheduled salary for BA teachers in the same district as a teacher is a good proxy for expected future benefits because most teachers reach the maximum scheduled salary before retiring. Since I used this as an instrument in previous specifications and I am now including it in both stages of the estimation, I use a new instrument for price: the scheduled salary in the teacher's district for someone with her level of experience. Conditional on the maximum schedule in a district, the salary paid to teachers with a particular level of experience varies because salary schedules have a different number of steps and/or have different sizes of steps, i.e. increases in salary with additional years of experience. This estimation framework allows me to be sure I compare the take-up rates at different prices for teachers who have the same potential gain from the benefit.

Columns (1) and (2) of Table 4 present results of the 2SLS estimation using this new instrument. Comparing the estimates in the first column to the results in Tables 2 and 3, the estimated relationships between price and demand are similar with this new instrument to what they were with the maximum salary as the instrument. The second column of Table 4 shows the results of specifications controlling for the maximum salary paid to a teacher with a BA in the teacher's own district in both stages of the estimation. For most experience levels, this variable is positively related to take-up (though it is not statistically significant). This is what I expected given that teachers with higher end-of-career salaries can expect to benefit more from upgrade purchase. Also, while the inclusion of this control for the potential benefit pushes the estimated

\textsuperscript{20} This is akin to controlling for the product's characteristics in a hedonic framework. Note that I have already included controls like age and experience of the teacher, which should proxy for the fact that, \textit{ceteris paribus}, an employee can expect to collect more from the upgrade if she expects to live longer or retire earlier.
relationship between price and demand to be more negative, the differences in the estimates are not statistically different from those seen previously. I therefore return to using specifications without this control.

Results with Alternative Specifications

In columns 3 and 4 of Table 4, I report the results of analyses where I use alternative definitions of the simulated instrument. In column 3, the instrument is defined using the beginning and maximum salaries paid to a teacher with a BA. In column 4, the relevant salary measure is just the beginning salary for a teacher with a BA degree. Across the specifications using these different order statistics from the salary measure, the coefficient estimates for the effect of price on take-up are fairly stable, though they are no longer as likely to be statistically significant.

Up to this point, I have assumed that both the demand and cost curves are linear in prices. That is not necessarily the case. With the results in Table 5, I explore the sensitivity of my conclusions to this functional form assumption for the demand curve as well as to some other constraints suggested by theory.\textsuperscript{21} The first row of Table 5 presents estimates from Table 2 with limited controls. Each subsequent row in the table presents estimates of both the constant term and the coefficient on price under different functional form assumptions. The second row of the table presents results using IV probit estimation rather than 2SLS. In this specification, IPS employees are slightly more price sensitive, but this does not change the main conclusion of the prior results.

\textsuperscript{21} Other specifications for the functional form of the average cost curve fit the data poorly. For example, chi-squared tests cannot reject the null hypothesis that the coefficient on a squared term in price is zero (at the 40 percent level).
Standard price theory provides guidance for how much consumers should value a product. Assuming there is free disposal, when the price of the upgrade is zero, everyone should purchase it. Therefore, one constraint I model is that when price equals zero, demand should equal one. Similarly, the price a consumer is willing to pay should never exceed the most she could expect to receive from purchasing the upgrade. The 99th percentile of expected costs to the retirement fund of an employee is at $162,000. This implies that when the price is $162,000, demand should be zero. I test the effect of using these constraints individually and combined in the remaining rows of Table 6. The restrictions have little effect on the estimated demand curve.

**Employee Valuation of Current Income Relative to Future Income**

The coefficient estimates from Tables 2 and 3 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 (2) defines the demand curve, while equation (3) can be used to draw the average cost 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 (2) and (3) in the following formula:

\[
MC(D) = \frac{\partial TC(D)}{\partial D} = \frac{\partial (AC(D) \cdot D)}{\partial D} = \frac{\partial AC(D)}{\partial D} D + AC(D).
\]

Figure 2 presents an example of these curves for IPS employees with 28 years of experience in 1998. The solid line is the demand curve as predicted using the estimates from Table 2. The dashed and dotted lines represent the average and marginal cost curves, respectively, as calculated from the estimates in Tables 2.
The first and foremost 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 demand curve in Figure 2 implies 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 $61,000. Yet, this high valuation employee can expect to collect over $90,000 in benefits. So, my estimates suggest that even the highest value employee is willing to trade just 66 cents of current compensation for a future dollar of benefits. Averaging along the entire demand curve, employees in IPS are willing to trade just 30 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 just over the data points themselves show that over half of the employees had a willingness to pay for the upgrade that was over $50,000 less than the expected benefits they would later collect. Comparing the demand and marginal cost curves over the region of observed data suggests that most teachers are unwilling to trade just 17 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. 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.

The preference for current funds relative to annuitized payments is therefore 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 this 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, including teachers, 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.

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 decision points necessary to make decisions about defined contribution pensions, which includes
decisions about the portfolio choice. Third, recent literature suggests that people are likely to opt-out of the default investment option when the returns to doing so are larger (Beshears et al. 2010). As described earlier in the paper, the returns to purchasing the upgrade are rather large. This is probably why the large majority of public school employees purchase it. Finally, the purchase of the upgrade is not 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.

Similarly, one might make the case that the complexity of discounting or financial calculations may lead to behavior that differs from the rational expectations model described earlier. If the reason for differences in take-up were differences in ability to assess the information and make a complex financial calculation, it might be the case that some IPS employees, e.g. math teachers or administrators, would be more likely to determine the value of the upgrade and, hence, purchase it. However, take-up rates and price sensitivity do not vary with observable characteristics of teachers that are unrelated to the value of the upgrade (results not shown, but available upon request). This makes it unlikely that employee upgrade purchase decisions were driven by the costliness of gathering information or computing the price and expected payout of the upgrade.

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 frequently close to 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.

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.\textsuperscript{22} 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

\textsuperscript{22} 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.
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. It is therefore unlikely that models relying on the representative nature of union leadership and its members can explain current levels of pension funding in the public sector.

5. 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 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 the 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 are have a willingness-to-pay for the upgrade that is $50,000 less than what the upgrade is worth to them in expectation. Said differently, teachers would prefer $1.80 in current wages to $10 in 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.

In this context, the main finding of this paper, that the majority of IPS employees value their pension benefits at about 17 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 how many states could eliminate their pension funding deficit with such a program. I use data on the assets and liabilities, specifically the liabilities discounted at the risk-free rate of return, from Novy-Marx and Rauh (2009). To determine

---

whether a state could use a buyback program to remove its funding deficit, I compare the liabilities of the pension fund in the state to the worth of the assets at a particular buyback rate, assuming all pensioners took up the buyback opportunity. At a buyback rate of 40 cents on the dollar, 16 states would eliminate their pension fund deficit. The number of states that would be fully funded increases to 35 and 50 at buyback rates of 30 and 20 cents on the dollar, respectively. All of these rates would provide teachers with more current income than necessary, given the valuations estimated in this paper, though care should be taken when extrapolating from the increase in benefits studied here to the buyback of all benefits.

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 suggest 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 66 cents on the dollar.

---

24 For example, Illinois had liabilities of $285 billion and assets of just $66 billion. At a buyback rate of 40 cents on the dollar, the assets are worth $164 billion, not enough to cover the liabilities. At a buyback rate of 20 cents on the dollar, the assets are worth $329 billion.
It does not appear the differences across teachers are able to explain the existence of generous pension benefits in the public sector.

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 the last set 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 voters are either myopic or if for some reason the costs of pension benefits are less salient to them (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*


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. Estimated Demand and Cost Curves for Employees with 28 Years of Experience in 1998.

Note: Based on the author's calculations using the TRS and TSR.
Table 1. Take-up Rates, Retirement Rates, Prices, and Costs by Level of Experience in 1998

<table>
<thead>
<tr>
<th>Years of Service in 1998</th>
<th>Fraction Who Retire by 2009</th>
<th>Fraction Who Purchase Upgrade by 2009</th>
<th>Average Price of Upgrade</th>
<th>Average Price of Upgrade to Purchasers</th>
<th>Average Cost of Providing Upgrade</th>
<th>Average Cost of Providing Upgrade to Purchasers</th>
<th>Number of Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>0.83</td>
<td>0.78</td>
<td>$15,245</td>
<td>$15,801</td>
<td>$94,166</td>
<td>$97,843</td>
<td>3,213</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(4,482)</td>
<td>(4,631)</td>
<td>(36,867)</td>
<td>(37,740)</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>0.88</td>
<td>0.77</td>
<td>$15,371</td>
<td>$15,758</td>
<td>$97,880</td>
<td>$101,341</td>
<td>2,984</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(4,460)</td>
<td>(4,550)</td>
<td>(38,314)</td>
<td>(38,029)</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>0.89</td>
<td>0.74</td>
<td>$15,722</td>
<td>$16,149</td>
<td>$99,423</td>
<td>$103,204</td>
<td>2,763</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(4,713)</td>
<td>(4,771)</td>
<td>(41,165)</td>
<td>(41,526)</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>0.91</td>
<td>0.73</td>
<td>$15,866</td>
<td>$16,385</td>
<td>$99,338</td>
<td>$103,241</td>
<td>2,814</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(4,632)</td>
<td>(4,817)</td>
<td>(39,635)</td>
<td>(40,169)</td>
<td></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 2. Instrumental Variables Estimates of the Relationship Between Annuity Demand (Purchase of the Upgrade) and Price

<table>
<thead>
<tr>
<th>Experience in 1998</th>
<th>25 Years</th>
<th>25 Years</th>
<th>25 Years</th>
<th>25 Years</th>
<th>26 Years</th>
<th>26 Years</th>
<th>26 Years</th>
<th>26 Years</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Price</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ordinary Least Squares: Teacher's Own Salary</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.014***</td>
<td>0.023***</td>
<td>0.023***</td>
<td>0.021***</td>
<td>0.011***</td>
<td>0.020***</td>
<td>0.020***</td>
<td>0.017***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
<tr>
<td><strong>IV: District Level BA Maximum Salary</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Price</strong></td>
<td>-0.016***</td>
<td>-0.015***</td>
<td>-0.024**</td>
<td>-0.019*</td>
<td>-0.018***</td>
<td>-0.017***</td>
<td>-0.026**</td>
<td>-0.028</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.009)</td>
<td>(0.011)</td>
<td>(0.004)</td>
<td>(0.005)</td>
<td>(0.011)</td>
<td>(0.016)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Experience in 1998</th>
<th>27 Years</th>
<th>27 Years</th>
<th>27 Years</th>
<th>27 Years</th>
<th>28 Years</th>
<th>28 Years</th>
<th>28 Years</th>
<th>28 Years</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Price</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ordinary Least Squares: Teacher's Own Salary</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.010***</td>
<td>0.011***</td>
<td>0.020***</td>
<td>0.020***</td>
<td>0.014***</td>
<td>0.023***</td>
<td>0.023***</td>
<td>0.020***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.003)</td>
</tr>
<tr>
<td><strong>IV: District Level BA Maximum Salary</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Price</strong></td>
<td>-0.019***</td>
<td>-0.018***</td>
<td>-0.025*</td>
<td>-0.025</td>
<td>-0.017***</td>
<td>-0.023*</td>
<td>-0.029**</td>
<td>-0.029**</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.013)</td>
<td>(0.014)</td>
<td>(0.005)</td>
<td>(0.013)</td>
<td>(0.014)</td>
<td>(0.014)</td>
</tr>
</tbody>
</table>

- **PUMA Characteristics:** X
- **County Fixed Effects:** X X
- **County Characteristics:** X
- **District Characteristics:** X X

Note: Based on the author's calculations using data from the TSR and TRS. Sample includes all teachers eligible for the upgrade. Each column row set represents estimates from a separate regression limited to teachers with the number of years of experience in 1998 indicated. 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 where the actual price paid for the annuity is instrumented with the maximum scheduled salary paid to teachers with a BA in one's district or region of employment in 1998 (both measured in thousands of $2010). 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 3. Instrumental Variables Estimates of the Relationship Between Annuity Total Cost and Take-up

<table>
<thead>
<tr>
<th>Experience in 1998</th>
<th>25 Years</th>
<th>25 Years</th>
<th>25 Years</th>
<th>25 Years</th>
<th>26 Years</th>
<th>26 Years</th>
<th>26 Years</th>
<th>26 Years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1.258)</td>
<td>(1.096)</td>
<td>(1.083)</td>
<td>(1.084)</td>
<td>(1.342)</td>
<td>(1.185)</td>
<td>(1.167)</td>
<td>(1.174)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Experience in 1998</th>
<th>27 Years</th>
<th>27 Years</th>
<th>27 Years</th>
<th>27 Years</th>
<th>28 Years</th>
<th>28 Years</th>
<th>28 Years</th>
<th>28 Years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1.445)</td>
<td>(1.274)</td>
<td>(1.260)</td>
<td>(1.255)</td>
<td>(1.368)</td>
<td>(1.195)</td>
<td>(1.197)</td>
<td>(1.199)</td>
</tr>
</tbody>
</table>

PUMA Characteristics X X X X
County Fixed Effects X X X X
County Characteristics X X X X
District Characteristics X

Note: Based on the author's calculations using data from the TSR and TRS. Sample includes teachers who purchase the upgrade. Each column row set represents estimates from a separate regression limited to teachers with the number of years of experience in 1998 indicated. The dependent variable is the expected cost (at time of purchase) to the TRS if the employee purchased the upgrade. 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 Variable Estimates of the Effects of Price on the Demand for and Average Cost of the Annuity, Controlling for End-of-Career Income

<table>
<thead>
<tr>
<th>Experience in 1998</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Scheduled Salary for Given Experience Level as Instrument</td>
<td>Multiple BA Salary Measures</td>
<td>Beginning BA Salary</td>
<td></td>
</tr>
<tr>
<td>25 Years</td>
<td>Price</td>
<td>-0.012**</td>
<td>-0.009</td>
<td>-0.007**</td>
</tr>
<tr>
<td></td>
<td>Max Salary</td>
<td></td>
<td>-0.002</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.006)</td>
<td></td>
</tr>
<tr>
<td>26 Years</td>
<td>Price</td>
<td>-0.026***</td>
<td>-0.037**</td>
<td>-0.006**</td>
</tr>
<tr>
<td></td>
<td>Max Salary</td>
<td></td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.008)</td>
<td></td>
</tr>
<tr>
<td>27 Years</td>
<td>Price</td>
<td>-0.032***</td>
<td>-0.055***</td>
<td>-0.011***</td>
</tr>
<tr>
<td></td>
<td>Max Salary</td>
<td></td>
<td>0.010*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.008)</td>
<td></td>
</tr>
<tr>
<td>28 Years</td>
<td>Price</td>
<td>-0.022***</td>
<td>-0.032*</td>
<td>-0.005</td>
</tr>
<tr>
<td></td>
<td>Max Salary</td>
<td></td>
<td>0.004</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.007)</td>
<td></td>
</tr>
</tbody>
</table>

Note: Based on the author's calculations using data from the TSR and TRS. Each column row set represents estimates from a separate regression limited to teachers with the number of years of experience in 1998 indicated by the row header. The dependent variable is a dummy variable equal to 1 if the person purchased the upgrade and 0 otherwise. Reported estimates are from 2SLS estimation where the actual price paid for the annuity is instrumented with instruments indicated by the column header. In the first two columns, the excluded instrument is the salary paid to a teacher with the given experience according to the salary schedule in one's district in 1998 (measured in thousands of $2010). In column (2), the maximum salary paid to a BA teacher is also included in the estimation. In the third column, the excluded instruments include the maximum and beginning salaries paid to teachers with a BA. In column (4), the instrument is the beginning salary paid to BA teachers. Individual level, PUMA-level, county-level and district-level controls are included in each regression. Each regression also includes county fixed effects. Standard errors are clustered at the district level. Standard errors are in parentheses; *** p<0.01, ** p<0.05, * p<0.1.
<table>
<thead>
<tr>
<th>Experience in 1998</th>
<th>25 Years</th>
<th>26 Years</th>
<th>27 Years</th>
<th>28 Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.977***</td>
<td>0.986***</td>
<td>0.995***</td>
<td>0.963***</td>
</tr>
<tr>
<td>Coefficient on Price</td>
<td>-0.015***</td>
<td>-0.018***</td>
<td>-0.018***</td>
<td>-0.017***</td>
</tr>
<tr>
<td>Coefficient on Price</td>
<td>-0.004 (0.062)</td>
<td>-0.004 (0.070)</td>
<td>-0.004 (0.075)</td>
<td>-0.005 (0.075)</td>
</tr>
<tr>
<td>Coefficient on Price</td>
<td>0.004 (0.004)</td>
<td>0.011 (0.171)</td>
<td>0.012 (0.174)</td>
<td>0.012 (0.175)</td>
</tr>
<tr>
<td>Coefficient on Price</td>
<td>-0.018*** (0.001)</td>
<td>-0.018*** (0.001)</td>
<td>-0.019*** (0.001)</td>
<td>-0.020*** (0.001)</td>
</tr>
<tr>
<td>Coefficient on Price</td>
<td>-0.019*** (0.001)</td>
<td>-0.018*** (0.001)</td>
<td>-0.017*** (0.001)</td>
<td>-0.016*** (0.001)</td>
</tr>
<tr>
<td>Coefficient on Price</td>
<td>0.000 (0.000)</td>
<td>0.000 (0.000)</td>
<td>0.000 (0.000)</td>
<td>0.000 (0.000)</td>
</tr>
<tr>
<td>Coefficient on Price</td>
<td>-0.004*** (0.000)</td>
<td>-0.004*** (0.000)</td>
<td>-0.004*** (0.000)</td>
<td>-0.004*** (0.000)</td>
</tr>
<tr>
<td>Coefficient on Price</td>
<td>0.000 (0.000)</td>
<td>0.000 (0.000)</td>
<td>0.000 (0.000)</td>
<td>0.000 (0.000)</td>
</tr>
<tr>
<td>Coefficient on Price</td>
<td>-0.004*** (0.000)</td>
<td>-0.004*** (0.000)</td>
<td>-0.004*** (0.000)</td>
<td>-0.004*** (0.000)</td>
</tr>
<tr>
<td>Coefficient on Price</td>
<td>0.000 (0.000)</td>
<td>0.000 (0.000)</td>
<td>0.000 (0.000)</td>
<td>0.000 (0.000)</td>
</tr>
<tr>
<td>Coefficient on Price</td>
<td>0.000 (0.000)</td>
<td>0.000 (0.000)</td>
<td>0.000 (0.000)</td>
<td>0.000 (0.000)</td>
</tr>
<tr>
<td>Coefficient on Price</td>
<td>0.000 (0.000)</td>
<td>0.000 (0.000)</td>
<td>0.000 (0.000)</td>
<td>0.000 (0.000)</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. Each column row set represents estimates from a separate regression limited to teachers with the number of years of experience in 1998 indicated. The dependent variable is a dummy variable equal to one if the employee purchased the upgrade. Estimates are from 2SLS or IV probit estimation where the actual price paid for the annuity is instrumented with the maximum scheduled salary paid to teachers with a BA in one's district of employment in 1998 (both measured in thousands of $2010). Theory-based restrictions placed on the coefficients as indicated in the row header. Standard errors (in parentheses) are clustered at the district level; *** p<0.01, ** p<0.05, * p<0.1.
Appendix A. 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.