Volume Title: Public Sector Payrolls

Volume Author/Editor: David A. Wise, ed.

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-90291-9

Volume URL: http://www.nber.org/books/wise87-1

Publication Date: 1987

Chapter Title: Promise Them Anything: The Incentive Structures of Local Public Pension Plans

Chapter Author: Howard L. Frant, Herman B. Leonard

Chapter URL: http://www.nber.org/chapters/c7155

Chapter pages in book: (p. 215 - 242)

9 Promise Them Anything: The Incentive Structures of Local Public Pension Plans

Howard L. Frant and Herman B. Leonard

Public pension systems have been much criticized, but their details have been studied relatively little. Studies of federal pension plans have revealed substantial accumulations of unfunded liabilities facing future taxpayers, and both government and private studies of state and local pension plans have indicated that these problems are common, though not universal, in lower-level jurisdictions as well. But while there have been some studies of the aggregate impacts of these plans, little attention has been paid to the level and form of the incentives they create. The differences across jurisdictions are frequently quite dramatic. The level and timing of pension benefits and of the accrual of pension rights by employees—and the work incentives thereby created—are strikingly variable across plans. Our primary purpose in what follows is to describe that variation and give some insight into its sources. We will not explicitly concern ourselves with developing a theory to account for the observed facts, but neither will we wholly resist the tendency of some of the more remarkable facts to speak for themselves about theory.

We examine 94 local employee public pension plans from thirty-three states. Of these, 67 cover general employees or teachers, and 27 cover police or fire employees. Some plans are state-administered; most are locally administered. The plans we describe are among those investigated in Arnold (1983); they represent a subset for which there were adequate data to conduct our examination. These systems cover more than 2.9 million employees.[1] The plans do not represent a random sample, so the statistics we will cite should be taken as roughly indicative rather than precisely descriptive.

Herman B. Leonard is associate professor of public policy at the John F. Kennedy School of Government, Harvard University.

Howard L. Frant is a doctoral candidate in public policy at the John F. Kennedy School of Government, Harvard University.

This chapter describes the character and variety of public pension plans, examines the roles played by certain features of these plans, and assesses their relative importance. We focus on the time profile of pension wealth and wealth accruals. Pension wealth accrual is the increment to a worker's wealth in a given year as a result of increases in pension rights granted in that year, just as conventionally measured labor income is the increase in a worker's wealth resulting from wages and salaries. Pension wealth accruals are thus an element of total worker compensation; to understand the time profile and consequent incentive effects of public compensation, we need to understand the time profile of pension accruals.

Our work parallels research of Kotlikoff and Wise (1984) describing private sector plans. Aside from the fact that public sector plans cover large numbers of employees, there are two (possibly contradictory) reasons why we might be interested in looking at these plans. First, they may have different labor market properties or be determined by different factors than private sector plans. Second, because these plans are not covered by federal pension law, they represent a less constrained and therefore richer universe of possible features.

9.1 Some Features of the Plans

Form. All of the plans we are examining are defined benefit plans—pensions are determined by formula, typically related to years of service and to salary in the last year or last few years before retirement. Nearly all of our plans have formulas of the form

$$\text{Pension} = BAR \times YOS \times SALAVG,$$

where *BAR* is the benefit accrual rate; *YOS*, the years of service; and *SALAVG*, the average salary received in a specified number of years prior to retirement. Three- and five-year final salary averaging are the most common, though pensions based only on salary in the last year are not uncommon in our plans. A few plans have two- or four-year final salary averaging; one plan averages salaries in the final ten years.

Benefit accrual rates. In general, these plans appear to be more generous than private sector plans. While Kotlikoff and Wise (1984) describe a typical private plan as having a benefit accrual rate (the percentage of average final earnings that the worker receives per year of service) of 1 percent, rates in public plans with a single rate ranged from 1 percent to 3.33 percent, with a mean of 1.9 percent and a mode and median of 2 percent. About three-fifths of the plans had some ceiling on accrual of benefits.

Cost-of-living increases. Nearly half of the plans have explicit provision for a cost-of-living (COL) increase to pensioners. The provisions

are generally far less generous than the full CPI increase of federal retirement systems and Social Security; four-fifths of these plans cap COL increases at 3 percent or less per year, and a few also have caps on the total COL adjustment a retiree may receive over the length of the pension. Only a half-dozen plans are explicitly integrated with Social Security.

Vesting. Vesting in some public plans contrasts sharply with that in private plans covered by ERISA. Nine of our plans have no vesting at all—workers become entitled to the pension at the same time they become eligible to begin drawing it. Eight of these are police or fire plans. Seven others have vesting of twenty years or more; five of these cover police or fire employees. Thus 13 of the 27 police/fire plans in our group have no vesting or very long vesting, while only 3 of 67 general plans do. Among the remaining plans, vesting ranges from one to fifteen years, with ten years being typical. All but three plans have "cliff" vesting—that is, workers receive full entitlement to a pension in a single year.

Early retirement. The contrast between police/fire and general plans is also striking with respect to early retirement. Only a third of the police/fire plans have a provision for a reduced pension before normal retirement age, while more than three-quarters of the general plans have such a provision. The difference is no doubt related to the generally earlier normal retirement age in police/fire plans: the mean age for *unreduced* retirement for someone entering one of these plans at age 25 is 51, while the mean age for *first* retirement (reduced or unreduced) in general plans is over 54, and for unreduced retirement almost 59.

Eligibility for benefits. Only twenty-two of the plans have age-only requirements for full retirement (or age-only plus vesting), and only four, all police/fire plans, have service-only requirements. The remainder have various age and service combinations.

9.2 Methodology

Our approach to analyzing these plans was to calculate wealth and accruals for a single hypothetical worker. We chose a worker who enters the system at age twenty-five. Using a single worker rather than some composite of various ages gives a clearer picture of incentive patterns. As we will illustrate later, however, the time profile can change markedly when different assumptions are made about entry age. The profiles that we present, therefore, should not be considered as complete characterizations of the plans in question, but rather as illustrative of the ways in which varying plan provisions can produce different effects on similar individuals.

In order to make these calculations, we must make assumptions about the real interest rate, the inflation rate, the real rate of salary growth associated with increased experience, and the real rate of general wage increase in the economy. We used 3 percent as the real interest rate and 5 percent as the inflation rate. To put all plans on a comparable basis, we used the same assumed salary growth trajectory for every plan. Experience growth rates were assumed to be the same as in the federal civil service, as reported by the Office of Personnel Management (1980). These rates range from 5.5 percent at age 25, to 2.2 percent at 45, to 1.1 percent at 65. In addition, we assumed a real annual growth rate of 0.6 percent in general salary levels over time; this is consistent with assumptions used for federal workers by OPM.

The pension is an annuity whose expected duration equals the pensioner's expected remaining life from the date he or she begins receiving benefits.[2] The value of that annuity will differ across plans, depending on their provisions for cost-of-living increases.[3] We take the value of pension wealth in any given year to be the value of pension rights acquired up to that point—in effect, the value of the rights a worker would have if he left his job in that year. Thus, a worker who is not vested has pension wealth of zero. A worker who leaves after becoming vested, but before she qualifies to begin collecting a pension, has a future right whose value must be discounted to the present. The appropriate discount factor is (almost always) the nominal discount rate, since the vested pension right is granted (almost always) in nominal terms. Given that a worker has departed (call it either resignation or retirement) but is not yet eligible to begin collecting a pension, from what year should we discount pension rights to arrive at a present value? One answer would be to discount the pension from the year in which one first becomes eligible to begin receiving it. In some plans, though, age-based early retirement penalties are large enough to make it worthwhile, in present value terms, for a retiree to wait one or more years after initial eligibility before starting to receive a pension. A fully rational retiree will wait to begin receiving payments until the optimal year, that is, the year that the pension annuity has its highest present value.[4] (Note that taking account of this makes the accrual profile smoother than it would appear in a naïve model that assumes someone leaving work will take a pension as soon as it becomes available.) The optimal year is, of course, sensitive to starting age and discount rate assumptions.

The product of the benefit accrual rate, years of service, early retirement reduction factor (if any), final salary averaging factor, and annuity factor is equal to pension wealth as a fraction (or multiple) of current salary. This number times cumulative real salary growth gives pension wealth as a fraction (multiple) of age-25 salary. Although we

will give some results in terms of current salary, most of our discussion will be in terms of age-25 salary. We prefer to avoid using current salary as a metric because it does not capture one of the sources of pension wealth increases: increases in the salary base from which pensions are calculated. Using a reference point that represents a *fixed* number of real dollars, such as salary at age twenty-five (or any other age), thus gives a truer picture of a pension plan's incentive profile.

Accruals are calculated directly from pension wealth. We are interested in accruals as a component of labor income. What does this imply about the relationships between wealth and accruals? Think of the analogous situation for a defined contribution plan—that is, for a plan that consists of an actual account for each employee. In a defined contribution plan, the accrual would simply be whatever amount was deposited in the account that year. But the account balance would also increase as a result of the interest earnings on the funds already invested. Thus,

$$PW_t = PW_{t-1} \times (1 + r) + ACC_t,$$

where r is the real rate of return in the economy. The appropriate definition of the accrual in a defined benefit plan should be just the same. If at the end of the fifteenth year an employee has accumulated pension wealth of $100,000 and the real rate of return is 3 percent, then by the end of the sixteenth year she will have pension wealth of $103,000; any difference (positive or negative) is that year's accrual. The correct baseline from which to assess the annual accrual is thus the preceding year's pension wealth adjusted upward by the real rate of interest. We therefore define accruals as the increase in wealth from one year to the next *above* the increase due to interest on existing wealth.

For the six plans with Social Security integration, replacement rates were approximated using data for technical and clerical workers in service industries.[5] Because of computational complications, the optimal year to begin collecting a pension in these plans was simply assumed to be the first available year. This assumption appears to have little effect on any of the results.

9.3 Accrual Profiles: What Creates Them?

A striking fact about pension accrual profiles is that they often include "spikes" or discontinuities. In a particular year, the accrual may increase sharply over the previous year, then decline as sharply the following year. These features are costly and have potentially large incentive effects, and it seems unlikely that the time profile of wages exhibits similar features in either the same or the offsetting direction. As Kotlikoff and Wise (1984) noted, these facts are difficult to reconcile

with a spot-market view of labor, in which workers are paid their marginal product at each point in time.

What plan features create these spikes? Briefly, spikes are created by discontinuous or discrete assignment of pension rights. The simplest example is initial vesting, which we refer to as "primary" vesting. The sudden assignment of a right to a deferred pension, where no such right existed before, creates a one-year jump in accruals. The size of this spike depends both on the size of the deferred pension being awarded and on how far in the future benefits will be collected. The latter point implies larger vesting spikes not only for those plans with relatively early retirement, but also for those with relatively late vesting. As we will see, the effects can be dramatic.

There are other pension rights, however, that may be vested later than the primary vesting of basic entitlements. We refer to vesting of such additional entitlements as "secondary vesting." One example is the right to begin collecting a pension early at a reduced rate.[6] Whether pension accrual at the reduced retirement date is discontinuous depends on whether the right is assigned discretely. An "early retirement spike" is not created by the mere existence of an option for reduced retirement at some age. What creates a spike is that in the previous year, the only vested right that existed was to retire at some *later* age.

For instance, the first part of figure 9.1 shows the Denver police and fire plan, in which at age 49 the worker has a vested right to retire at 65. The following year he is awarded the right to retire immediately (although at a reduced pension). This creates an accrual at age 50 which is dramatically higher than that at 49 or 51. Note that the presence of an early retirement penalty keeps accruals substantially positive after the reduced retirement age of 50, even though in this example the final salary percentage reaches its ceiling at age 50. In contrast, the second part of figure 9.1 shows the Danbury, Connecticut, plan in which the worker in the year before reduced retirement has a vested right to retire the following year (age 55). (In this case the optimal year to begin collecting the pension is actually age 58, but that is not the essential feature here.) There is no discontinuity between age 54 and age 55. There may, of course, still be a discontinuity on the other side, if the worker gets most of the value of the pension in the year of reduced retirement. But early retirement penalties can go far toward smoothing out this discontinuity, as in this example. Among the plans we studied that have a reduced retirement feature, those permitting deferral to the reduced retirement date, and therefore not having a spike at that date, outnumber those with a spike by about 3 to 2.

Secondary vesting features have in common what we call *acceleration:* they result in some vested right moving nearer to the present. The early retirement spike discussed above is one example. In this

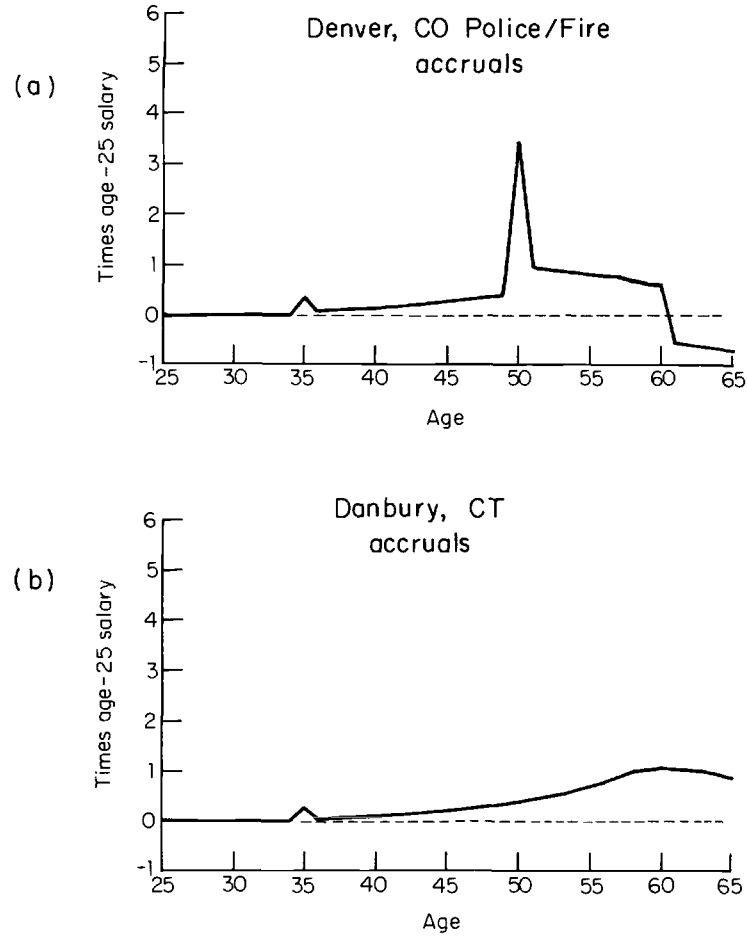

Fig. 9.1

case, we have *full* acceleration—that is, moving of a vested right all the way to the present. In the public plans we examined, *partial* acceleration was also an important cause of spikes in accrual patterns. This occurs in some plans with several age-service combinations for retirement. For example, a plan might permit retirement at age 60 with 10 years of service, or at 55 with 25 years of service. In some plans this means that a person with 25 years of service can leave and take with him the right to begin collecting a pension at age 55.[7] Since in the previous year he had only the less valuable right to collect a pension at 60, we observe a spike in the accrual at 25 years of service, representing the difference in value of those two rights. Another example would be a person eligible at 55 for a pension reduced, say, 5 percent

for each year before 65, who at 60 becomes eligible for a full pension under a different age-service combination. One can think of this situation either as a sudden increase in the benefit amount, or as a sudden acceleration in the date of full retirement. The two dimensions of full or reduced retirement and full or partial acceleration give four possibilities, any of which may create an accrual spike—and each of which is represented somewhere in the public plans we examined.

Finally, we should take note of other features that affect accrual profiles. As one might expect, the *benefit accrual rate*—the number that is multiplied by years of service to give the pension as a fraction of final average salary—affects the level but not the shape of the accrual profile. Figure 9.2 shows two Pennsylvania counties with plans that

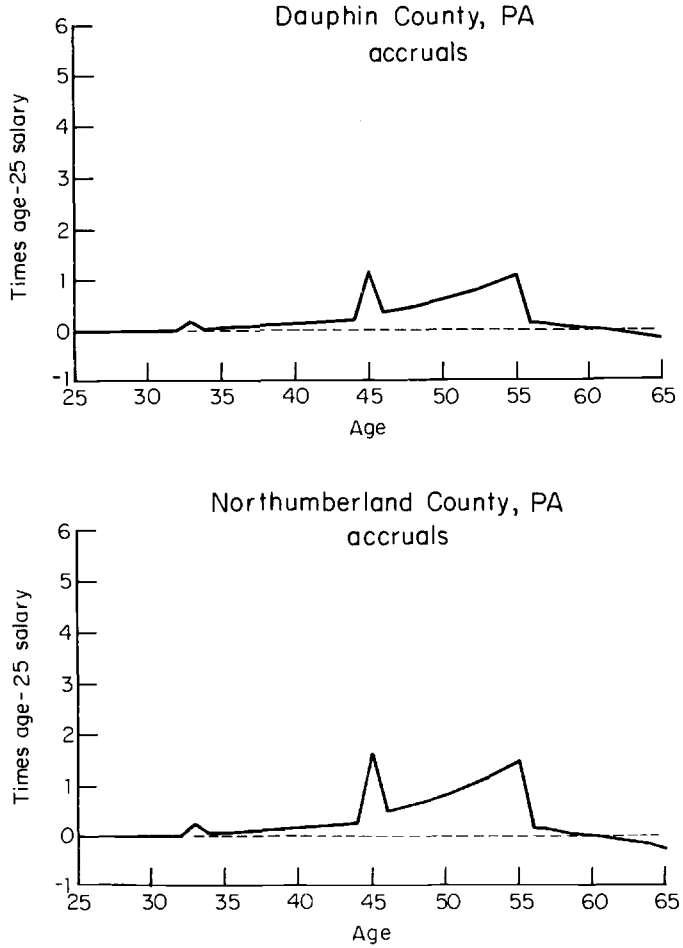

Fig. 9.2

are identical in structure but have different rates. A change in the benefit accrual rate in midcareer will create a kink in the accrual profile, but this effect is usually small. *Early retirement penalties* will increase the difference in pension wealth between one year and the next. Thus, they will make accruals more positive. *Benefit ceilings* will have the opposite effect: by reducing the gain from staying another year they will make accruals more negative.

9.4 Types of Accrual Profiles

The ninety-four plans we have studied display a broad range of accrual profiles. We found it convenient to group them into four broad categories.

The "simple" type displays a primary vesting spike and then relatively smooth accruals up to the date of full retirement, followed by a drop-off if the age of full retirement is before sixty-five. (This includes some plans in which there are reduced retirement provisions, but eligibility for reduced retirement occurs at primary vesting, so that there is no further discontinuity.) There are twenty-two such plans in our sample. Plans with this classic pattern can still exhibit tremendous variation in timing and levels of accruals, however, and can thus look strikingly different. Figure 9.3 shows the accrual patterns for Oregon;

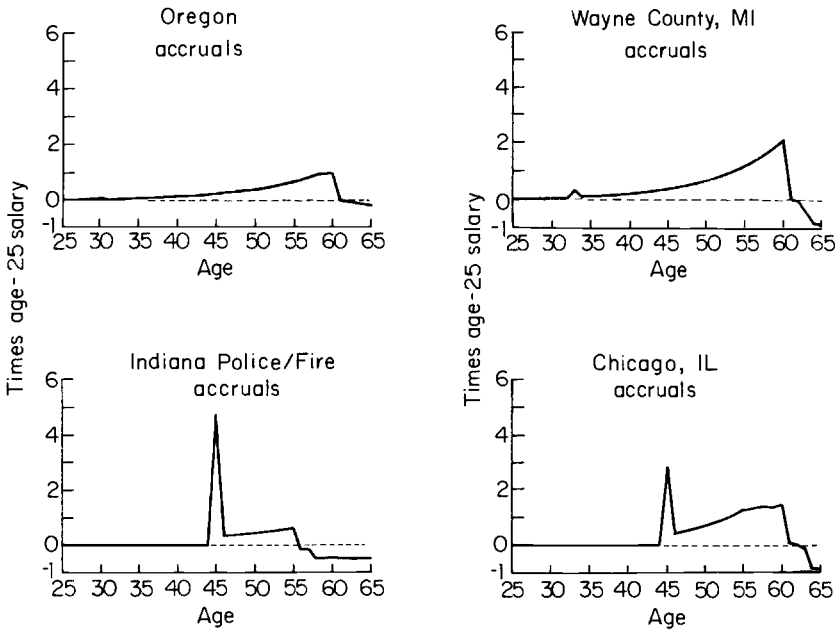

Fig. 9.3

Wayne County, Michigan; Indiana police and fire; and Chicago. The Oregon and Wayne County plans show small primary vesting spikes followed by smoothly increasing accruals over a long period, with those in Wayne County considerably larger. Both the Indiana police/fire and Chicago plans have late primary vesting and as a consequence have dramatically larger initial spikes. The Chicago plan then has a longer period with higher continuing accruals. Oregon and Chicago both show a flattening of the slope of the accrual profile after reduced retirement; the other two have no provision for reduced retirement. Even within the simple accrual pattern, then, the plans we examined showed enormous variation.

A more common form of accrual pattern is a primary vesting spike with one secondary vesting spike. There are forty-eight plans in this category. Again, plans with these essential features may look very different. Figure 9.4 shows patterns for Washington State and Lansing, Michigan, police and fire. The Washington plan has almost negligible spikes, while the Lansing police and fire plan has dramatic spikes for both primary and secondary vesting. Conversely, plans may look similar in their accrual profiles as a result of quite different provisions. For example, figure 9.5 shows that New York State, Grand Rapids police and fire, and Dauphin County, Pennsylvania, appear similar, yet the secondary spike is produced in the first case by a one-time retroactive increase in the benefit accrual rate (an unusual mechanism), in the second case by a conventional "early retirement" mechanism, and in the third by a partial acceleration of full retirement from sixty to fifty-five. We might also include in this category some plans such as those in figure 9.6 (Fresno, California, and Phoenix, Arizona), where a noticeable discontinuity is produced, in the first case, by a drop in the benefit accrual rate or, in the second case, by a ceiling on it. This is a close call, though, because such provisions do not produce spikes in the sense of a discontinuity on both sides of the year in question.

A third class is the "pot of gold at the end of the rainbow" group. In these plans, essentially all of the pension wealth is awarded in a single year, producing spikes that go far off the scale we are using here. Two examples, from San Antonio, Texas, and Birmingham, Alabama, are in figure 9.7. We found six such plans, five of them for police and/or fire fighters.

Finally, we have eighteen plans that exhibit various sorts of multiple spikes or other marked discontinuities. Examples are shown in figure 9.8, which shows the plans of the Lansing, Michigan, Board of Water and Light; Minnesota; Mobile, Alabama, police and fire; and Memphis, Tennessee. These spikes are most often produced by interaction of various age-service requirements, but as the examples show, they may also result from benefit accrual ceilings, discontinuous early retirement

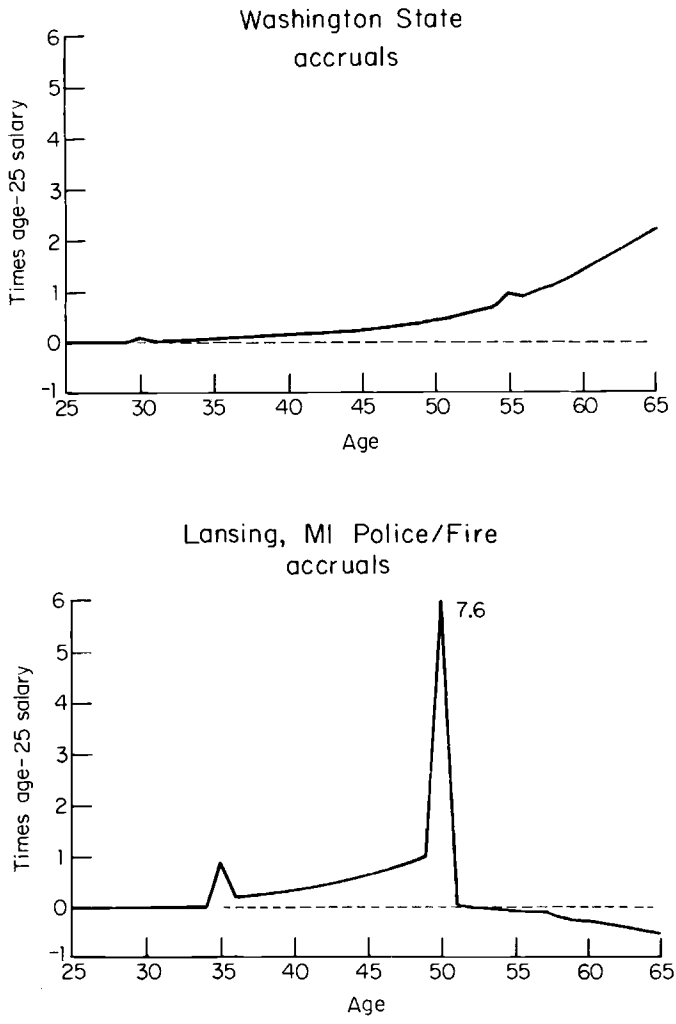

Fig. 9.4

reductions, discrete accruals, and other features. Examining each of these plans in turn will help show how the factors operate.

In the case of Lansing Water and Light, the spike at 35 is 10-year vesting with pension benefits starting at age 60. At 40, this worker has 15 years and becomes eligible for reduced retirement at age 55. At 50, the worker has 25 years and reduced retirement is accelerated from 55 to 50—that is, there is full acceleration of reduced retirement benefits. Finally, at 55 the worker has 30 years of service, and full retirement is accelerated from 60 to 55—there is full acceleration of full retirement benefits.

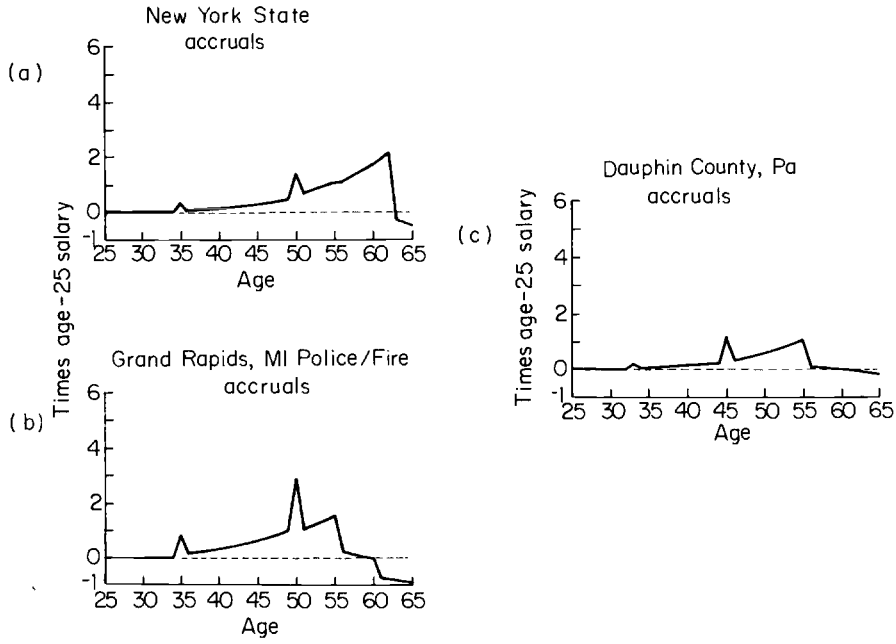

Fig. 9.5

The Minnesota plan has both vesting of the right to retire at age 55 and an increase in the benefit accrual rate (from 2 percent to 2.5 percent) after 10 years (age 35 for our illustrative worker). At age 55 there is eligibility for reduced retirement. This eligibility does not produce a discontinuity, however, because at vesting (and at age 54) the worker is eligible to leave and collect a reduced pension at 55. But our illustrative worker reaches 30 years of service at age 55, so the reduction is calculated not based on number of years before age 65 but on number of years before 62. This jump from one reduction schedule to another creates a spike. Finally, at age 58, the worker's age plus service equals 90, so he is able to jump from the higher early retirement schedule to full retirement.

The Mobile police and fire plan grants 50 percent of final salary after 20 years, 52.5 percent after 25 years, 55 percent after 30 years, and 60 percent after 35 years. Full retirement is possible at age 55. Note that accruals become dramatically negative after age 55, except in year 35 (age 60).

Memphis, Tennessee, has vesting at 10 years. At 25 years (age 50) there is an acceleration of the full-retirement age from 65 to 62. At age 55 there is a reduced-retirement spike, caused by acceleration from full retirement at 62 to reduced retirement at 55. The accruals are slightly irregular from age 55 to 60 because of varying penalties. At age 60 the

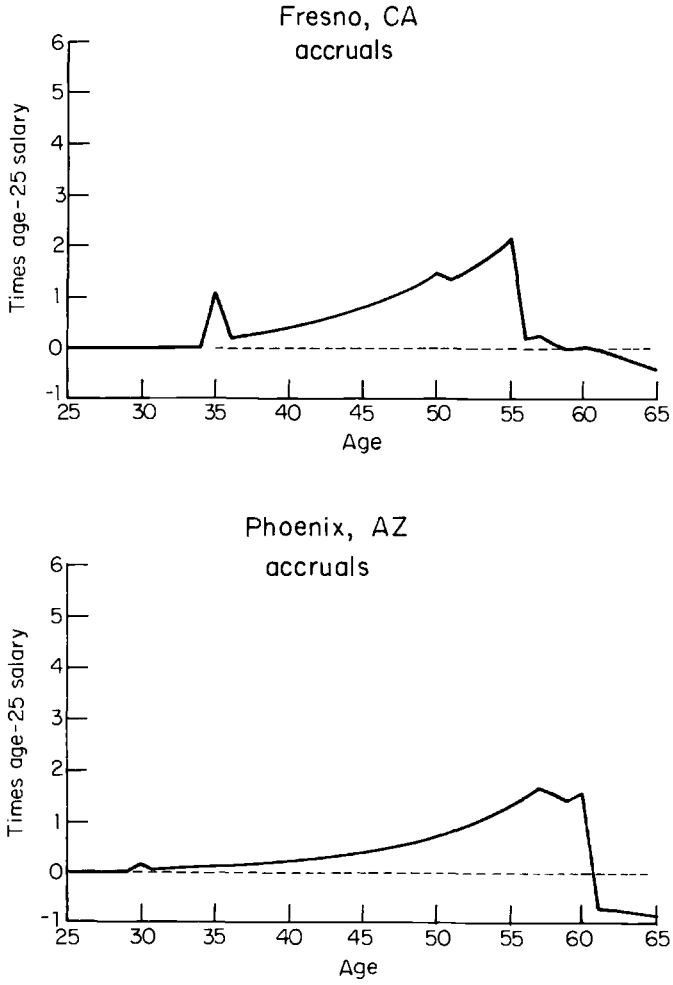

Fig. 9.6

worker receives full retirement (age 60 with 30 years) and simultaneously reaches the benefit accrual ceiling (35 years). The result is a dramatic drop-off in accruals to significantly negative numbers.

The plans we have examined, then, create an extremely wide range of accrual profiles and use a broad range of instruments to form them. Even those that are quite similar in kind may differ dramatically in degree.

It is worth noting that, despite the wide variations in accrual profiles described above, there are certain types of profiles we never observed. For example, none of the plans we examined had accrual profiles that were downward sloping, or even level, in real dollar terms.

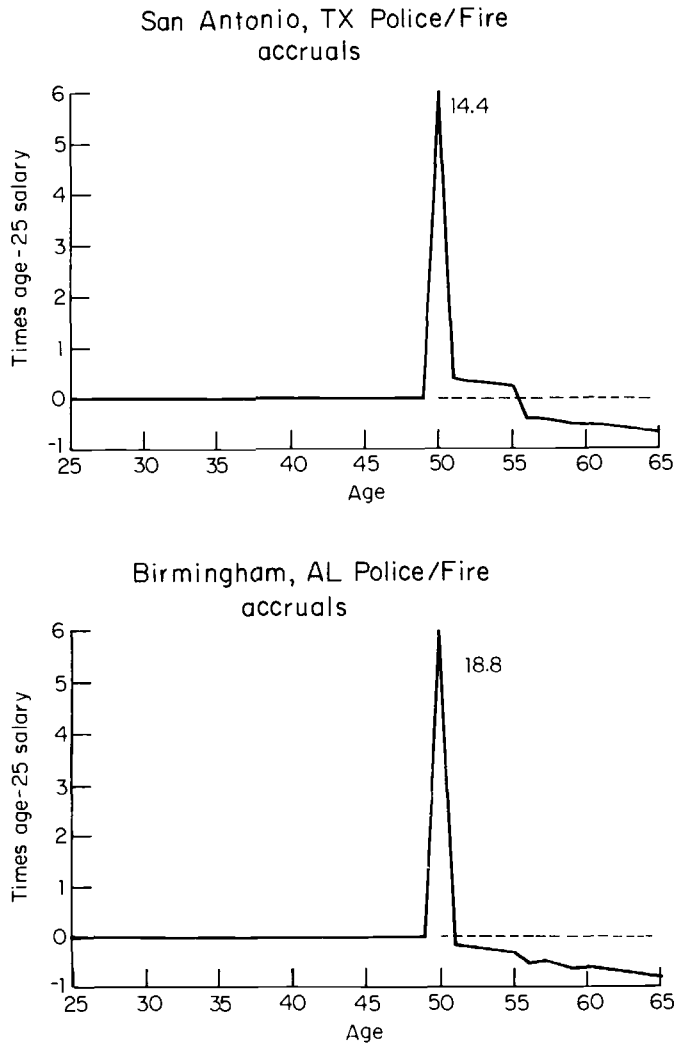

Fig. 9.7

9.5 What Is Interesting about These Profiles?

9.5.1 Wealth: Large, Convergent

Figure 9.9 shows the distribution across plans of pension wealth at five-year intervals, compounded forward to age sixty-five for comparability. Three things are striking about the wealth results.

First, the numbers are large. As noted earlier, these plans have large benefits relative to private plans. The mean value of pension wealth at age sixty-five in our plans is 24 times age-25 salary, with a standard

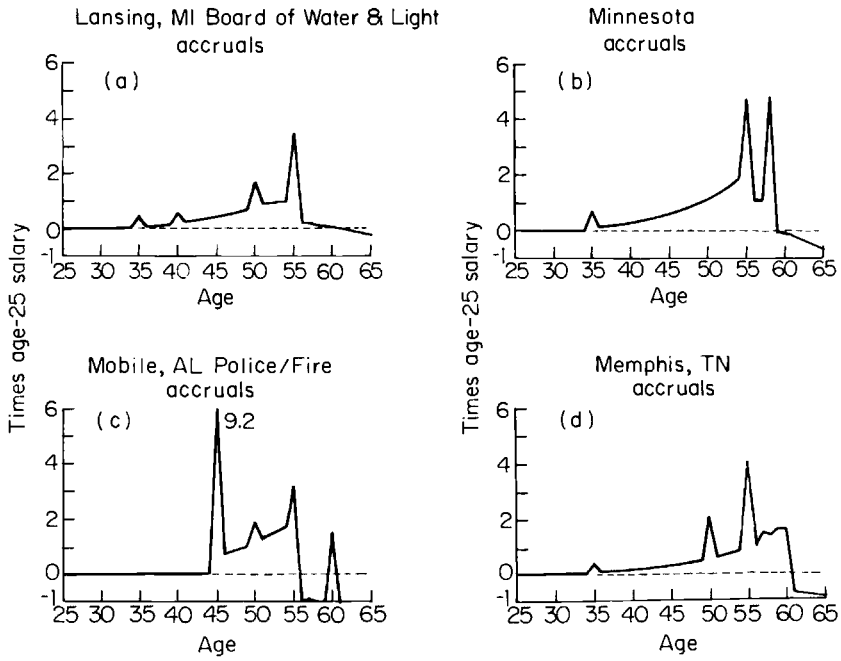

Fig. 9.8

deviation of about 6 times age-25 salary. Converted to normal costs (that is, divided by the cumulated compounded value of salary), this translates to a mean of 15.4 percent and a one-standard-deviation range of 11.7 percent to 19.2 percent. Kotlikoff and Wise (1984), using slightly different actuarial assumptions, calculate that a typical *private* pension, by contrast, represents 2.6 percent to 7.2 percent of discounted salary, depending on retirement date. It should be recalled, of course, that workers in many of our plans are not covered by Social Security.

Second, wealth tends to peak before age sixty-five. While accruals to wealth are generally positive up until the age of full retirement, they quickly drop and become negative thereafter.

Finally, the wealth associated with different plans tends to converge with increasing age—plans differ less in where they end up than in how they get there. At age forty-five, for instance, the standard deviation of pension wealth is 76 percent of the mean; by age sixty-five it is only 24 percent.

9.5.2 Big Spikes

In our sample, 25 of 27 police/fire plans and 28 of 67 non-police/fire plans had at least one spike in excess of *100% of current salary*. Thirteen

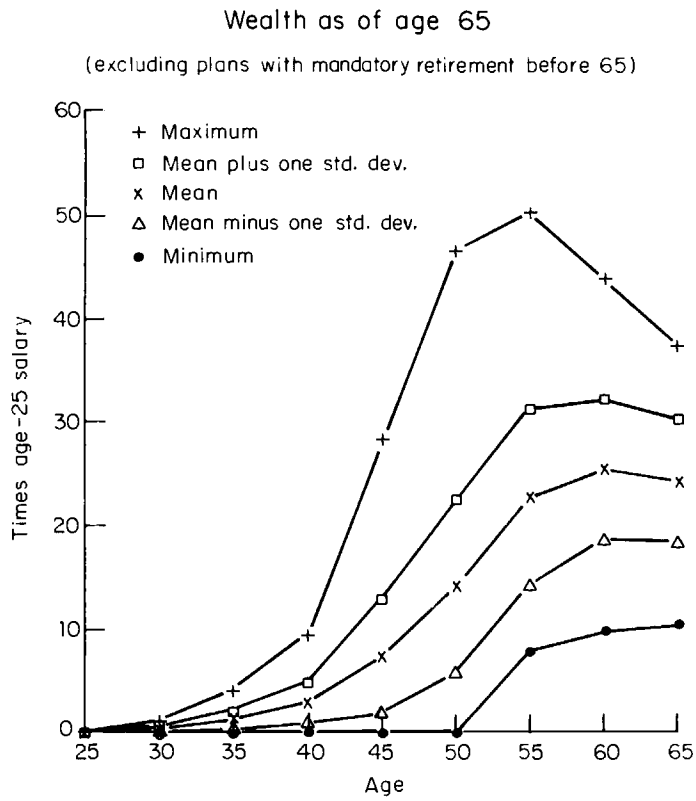

Fig. 9.9

police/fire plans and nine non-police/fire plans had spikes in excess of *three times* current salary. We know of no jurisdiction in which salaries are adjusted downward sharply in years in which these spikes occur. Thus, the profile of total compensation in many of these plans is highly irregular, and its variance is driven mainly by variation in pension accruals. It seems clear that even a tortuous story could not support the claim that these employees receive their marginal product each year.

The thirteen police/fire plans with very large spikes are, not coincidentally, the thirteen plans with no vesting or with vesting of more than twenty years. These plans are not markedly more generous than average in terms of benefit accrual rates. The size of these spikes demonstrates the sensitivity of vesting accruals to the time of receipt of the pension.

Among the non-police/fire plans with very large spikes, however, only one arises because the plan has no vesting. The remainder are

caused by some form of secondary vesting, generally involving an interaction of age and service requirements. In figure 9.10, which shows the accrual profile for Kent County, Michigan, the story is typical. An employee at age 49 has a vested right to receive a full pension at age 60. The following year, experience of 25 years qualifies the worker for immediate full retirement. The spike represents a complete acceleration of the right to full retirement. Acceleration need not be complete or dramatic, however, to produce large spikes. For instance, one plan with deferral to age 58, or to age 55 with 25 years, produces a spike at year 25 (for our illustrative worker, at age 50) in excess of 100 percent of current salary.

9.5.3 Effect of Early Retirement Penalties and Accrual Ceilings

Our intuition was that early retirement penalties would not be of much importance. By working for another year rather than taking an immediate pension, after all, one gains both a salary increase (real and nominal) and an increase in the final average salary percentage. One loses a year's pension, but many years in the future. We thought that accruals would be significantly positive after reduced retirement age, more or less irrespective of the penalty. Even without large penalties for early retirement, it would seem that the standard increases in pen-

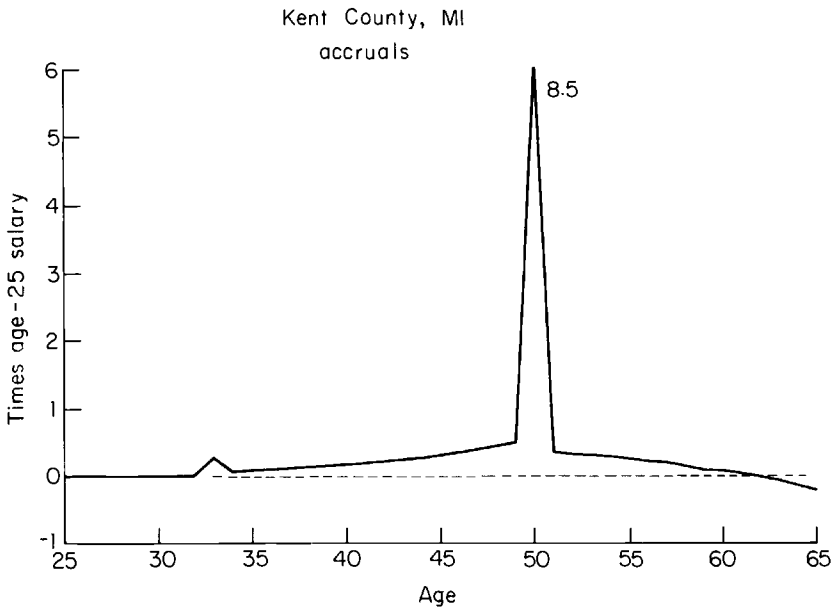

Fig. 9.10

sion benefit entitlements should provide a substantial incentive for employees to work up to the year of normal retirement.

The reduced retirement penalty generally has a much more important role than we anticipated in making accruals positive between the reduced retirement date and the full retirement date. As soon as the penalties stop (at the full retirement date), accruals tend to fall to near zero and then gradually drift downward. The growth of pension wealth caused by increases in final average salary percentage and in current salary tend to be about as much as the interest that would be due on the wealth to date, which means that the net accrual is about zero. Reductions in the penalty account for most of any positive accrual from year to year. And if pension wealth by this date is large (15 or 20 times age-25 salary at age fifty-five is not unusual), then staying another year to reduce the penalty by even a small percentage can yield a large accrual. Of course, the larger the penalty, the larger the accruals over this period.

To illustrate the effect of early retirement penalties in maintaining the pension-induced incentive to work in the final years before full retirement, figure 9.11 shows the accrual profile for Chicago both with the early retirement penalty it imposes (actual) and without (hypothetical). Without the early retirement penalty, accruals fall essentially to zero after the retirement date (in the absence of a penalty, this

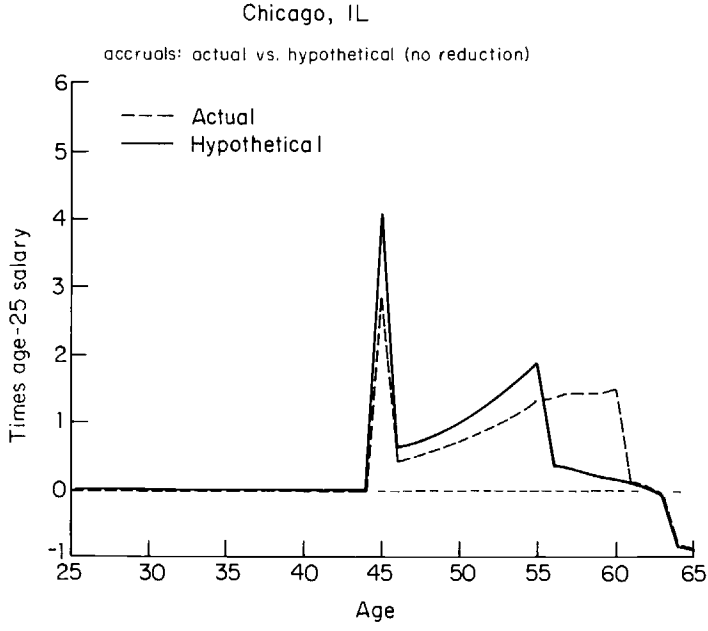

Fig. 9.11

would be age fifty-five for our illustrative worker). With no reduction, the increase in pension wealth from salary increases and additional benefit formula accruals (offset by the one-year reduction in the pension's expected length) is just enough to provide the interest that should be due on the existing balance—that is, the accrual for the year is essentially zero. With the penalty, the accruals remain positive up to the normal retirement date. Of course the penalty reduces the size of the primary vesting spike and the early accruals that follow it. Thus, the system with a retirement penalty spreads the positive pension accruals—the pension-based inducements to keep working—more smoothly and over a longer time period.

Ceilings on the benefit accrual rate have a significant effect in the other direction. If they occur after the full retirement date, they typically push the accruals as a multiple of age-25 salary from near zero to about -1, as in figure 9.11.

9.6 Implications

The patterns of pension wealth accruals that these plans display are puzzling. Since we have not attempted to develop a theoretical framework in which to evaluate efficiency, or to specify what employers' goals might be with respect to retention incentives, we cannot say definitively that these plans are inefficient. The data, however, strongly suggest this.

Lazear (1983) has noted that non-immediate vesting gives rise to an inefficiency. He suggests that the need to sort workers may provide an explanation, but one that is less than fully satisfying. Can one find a plausible explanation for plans, such as those in figure 9.8, that have several dramatic primary and secondary vesting spikes, or for those, as in figure 9.7, that are essentially nothing but a vesting spike?[8]

To argue that these are optimal contracts we must also explain the extraordinary sensitivity of some of these profiles to entry age. Figure 9.12, which shows the accruals for Minnesota, for example, shows the same plan with entry ages of 25 and 30. The two versions show peaks at similar points, to be sure, but in markedly different ways. In the case of age-25 entry, there are dramatic spikes of about five times age-25 salary at ages 55 and 58. In the case of age-30 entry, we find a gradual buildup of accruals to age 55, followed by a drop, followed by another gradual increase through age 60. The highest point is barely three times age-25 salary. These profiles present radically different incentives to the two workers. At age 54, is the difference between age-25 entry and age-30 entry really significant enough to justify this radical difference?

Exploration of the theoretical implications of these data is certainly in order. It seems likely, though, that some of these features arise from factors that are difficult to model: the political economy of the work-

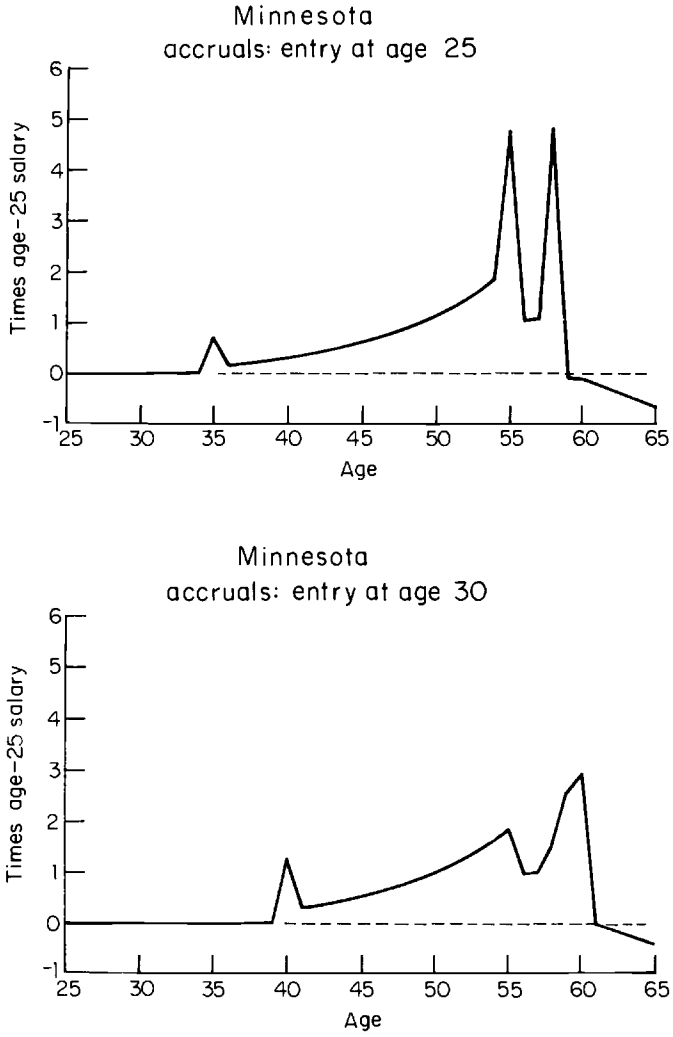

Fig. 9.12

place, institutional rigidities, or simple accident. This in turn suggests that these plans as presently constituted may not be an efficient expenditure of public money.

One approach to correcting this would be to simplify the plans. Complex interactions of early retirement penalties, entitlements to pension rights defined discontinuously in terms of age and service, and other features of some plans may have unintended consequences in terms of accrual profiles and resulting incentives. Simplifying plans that currently have such features might well result in more (or more

appropriate) incentives per dollar of required funding. Alternatively, one may note that incentives are much clearer if plans are on a defined contribution basis. Both workers and taxpayers could then see directly both the timing and the magnitude of the incentives provided, as well as their cost. In principle, any desired accrual profile could be achieved by varying contribution rates; in practice we would be very surprised if any defined contribution plan had a time profile similar to some of those shown here. Note that this in itself argues that some features of the accrual profiles of defined benefit plans are accidental.

9.7 What Remains to Be Done?

There are several areas in which further investigation of public sector plans is likely to be fruitful. First, one could simply expand the universe of plans examined to get a better statistical picture of public plans in general. We urge great care in doing so, however. Differences in plans are often more subtle than one would realize from a mere list of parameters—whether a plan permits vesting of accelerated retirement rights, for instance. Examining a large number of plans properly is a very tedious task.

Second, one could examine actuarial data for a number of plans to see how effective particular profiles actually are in achieving the incentive effects one would hypothesize from looking at them. It is possible—though, we think, unlikely—that workers do not really understand where plan spikes are. It is more likely that they have no way of accurately assessing the size of the spikes.

Finally, we noted that many systems have several coexistent plans, with older employees grandfathered under previous plans. Such an arrangement provides an opportunity to examine the direction of change of plans over time. Is it simply random or are there consistent trends? Understanding this question may help answer whether the process generating these patterns should be thought of as a market or a political one.

9.8 Conclusion

Pension payments are an important component of labor income in the state and local public sector. They differ dramatically across jurisdictions in form, in timing, in level, and in the incentives they provide workers. Some are so complex that their incentive patterns appear to have arisen more by accident than by design. They may also be too complex to be fully understood by workers. This in itself may be a reason to simplify some of the more complicated plans.

Notes

This research was conducted with support from the National Bureau of Economic Research Program on Public Sector Payrolls.

1. This total is for the *systems,* rather than the particular *plans* we are describing. Many of the systems have large numbers of employees grandfathered under previous rules. We have chosen to confine ourselves to the version covering new employees. The figure also includes some state employees.

2. We calculated the value of the pension annuity as the value of an annuity for the expected remaining life. Strictly speaking, the correct value is the expected value over one's lifetime. The difference is minor, however.

3. For those plans with no explicit provision, we followed Arnold 1983 in assuming that adjustments average half of the CPI.

4. It may be useful to emphasize that this "optimal year" is *not* the optimal year of *retirement,* simply the optimal year to begin receiving payments. Calculating the optimal year of retirement would be a daunting task indeed, especially since workers most likely differ dramatically in their valuations of leisure and perhaps in their other opportunities as well. We are not attempting here to provide a comprehensive account of local public employees' decisions about mobility, but only to suggest how pensions contribute to that picture. Thus, we do not, for example, discuss Social Security except insofar as it explicitly affects the size of pension rights.

5. Data were from a program developed by Douglas Phillips. We thank Gary Heaton for his assistance.

6. This right is commonly called *early retirement.* The term can be confusing, however, because it is sometimes used to refer to departure with a vested right to pension later, or to unreduced retirement at an earlier age due to some age-service combination. To avoid ambiguity, we will refer to *reduced* and *full* retirement.

7. Not in all, though. Some would require him to be still working at age 55 in order to exercise this option.

8. Becker and Stigler 1975 offers a model of law enforcement and corruption in which the optimal compensation schedule for an enforcer includes a large payment at retirement which one loses if one leaves before retirement age. While actual compensation plans do not mirror Becker and Stigler's proposal exactly (in particular, with respect to "entrance fees"), the resemblance is suggestive and agrees with the intuition of some of our readers that pensions of this form serve as an organizational control mechanism. The difficulty with this view is that we apparently do not find special pension plans in other corruption-prone local government jobs, such as building inspector or cashier, while we do find pensions of this form for fire fighters. We also find a strikingly similar form for U.S. military pensions (see Leonard, chap. 3). When a distinction is made, it seems to be not between enforcement and non-enforcement jobs, but between uniformed and non-uniformed.

References

Arnold, Frank S. 1983. State and local public employee pension funding: Theory, evidence, and implications. Ph.D. diss. Harvard University.

Becker, Gary S., and George J. Stigler. 1975. Law enforcement, malfeasance, and compensation of enforcers. In *Capitalism and freedom: Problems and prospects: Proceedings of a conference in honor of Milton Friedman*, ed. R. T. Selden. Charlottesville: University Press of Virginia.

Ehrenberg, Ronald G. 1980. Retirement system characteristics and compensating wage differentials in the public sector. *Industrial and Labor Relations Review*, July.

Ehrenberg, Ronald G., and Robert S. Smith. 1980. A framework for evaluating state and local pension reform. In *Urban public labor markets*, ed. P. Mieszkowski and G. Peterson. Washington, D.C.: Urban Institute.

Inman, Robert P. 1984. Public employee pensions and the local labor budget. *Journal of Public Economics*, October.

Kotlikoff, Laurence J., and David A. Wise. 1984. The incentive effects of private pension plans. NBER Working Paper No. 1510.

Lazear, Edward P. 1983. Incentive effects of pensions. NBER Working Paper No. 1126.

———. 1984. Pensions as severance pay. In *Financial aspects of the U.S. pension system*, ed. Z. Bodie and J. Shoven. Chicago: University of Chicago Press.

Leonard, Herman B. 1984. The federal civil service retirement system: An analysis of its financial condition and current reform proposals. NBER Working Paper No. 1258.

Office of Personnel Management. 1980. *Board of Actuaries of the Civil Service Retirement System Fifty-seventh annual report*. Washington: Government Printing Office.

Comment

Edward P. Lazear

Frant and Leonard do two things in their chapter on state and local pension plans. First and foremost, they provide a detailed description of the various pension accrual patterns that can be found in the ninety-four plans they examine. They do an admirable job of presenting these findings in a clear and careful way. Second, they attempt to draw some conclusions about the optimality of labor contracts. It is this second aspect that I find most troublesome. Most of my comments are directed there. For the most part, the next few pages will explore what can and cannot be learned from an examination of the differences in pension accrual patterns.

The authors make the point that it is difficult to square the various pension accrual patterns with a simple story of optimal contracts. There is too much diversity in pension plans to conform to a simple story. Even if all plans were alike, it would be difficult to present a straight-

Edward P. Lazear is Isidore and Gladys J. Brown Professor of Urban and Labor Economics at the Graduate School of Business, University of Chicago, and senior fellow, Hoover Institution.

forward theory that would reconcile the discontinuous nature of pension accrual. On the whole, I agree. It is difficult, if not impossible, to claim that the data they generate are absolutely consistent with an optimal contracts story. But I stop short of admitting that the evidence they present speaks strongly to the issue of optimal contracts. There are a number of reasons.

First, because Frant and Leonard do not discuss what pensions are supposed to be doing, it is difficult to determine whether the plans they consider are efficient. Efficiency can only be determined within the context of a model. A hypothesis must be presented before it can be refuted; absent a raison d'être, it is nonsensical to ask whether pensions accomplish their goals. The inability to make reason from the observed patterns is not sufficient evidence to reject any general statement about efficiency. An efficiency criterion must be presented.

Specifically, the discussion does not define the choice variables over which efficiency is to be considered. Some obvious possibilities are labor supply variables, for example, hours of work, age of retirement, and labor quality variables such as the level of effort and investment in human capital. Do the pensions they examine operate on all of these variables? For which ones are inefficient outcomes produced?

Most of these questions are logical rather than empirical. A certain amount of evidence is needed before one can even construct a theory, and this paper provides that evidence. But it does not tackle the second task of determining efficiency. For example, in "Incentive Effects of Pensions" (Lazear 1985), I consider the effect of various pension provisions on work effort, human capital, age of retirement and hours worked, and worker turnover. I conclude that many pension provisions, such as non-immediate vesting and pension plans that make the pension a function of final salary, are inefficient with respect to those labor supply and human capital variables. Frant and Leonard document convincingly that non-immediate vesting of various types is prevalent in the public sector. As such, I conclude that they do in fact cause some inefficiencies. But that cannot be gleaned from the authors' analysis. Does non-immediate vesting induce workers to work too many or too few years? How does it affect effort? I believe this chapter is complementary with my theoretical analysis and welcome it. But I also believe that as a matter of style, the authors' careful empirical work should not be coupled with loose statements about efficiency. This is particularly troublesome in the policy implications sections. Nothing in chapter 9 tells us whether defined contribution plans are superior to defined benefit plans, for example.

Another theoretical difficulty, which causes some minor empirical problems (discussed later), is that the pension is treated independent of wage compensation. Workers' decisions are affected by their total

compensation, not merely by the pension part. It is important to recognize the constraint on total compensation paid (even by a nonprofit organization) because the effects can be offset or exacerbated by the other component of compensation. This comes back to the issue of why there are pensions at all. That compensation is divided between pension and wage payments is unlikely to be a random event. Understanding the effects of pensions requires that we understand simultaneously the effects of wages. I am not arguing that the weird accrual patterns the authors find are likely to be offset by similar and opposite weird patterns in the wage profile. I merely claim that I do not know what to make of their evidence unless I examine it in context.

A third point is that hoping to explain discrete phenomena is asking too much. Few economic models are successful in this regard, even though discontinuities are common in the real world. This point is not specific to pensions, but occurs in other aspects of labor and product markets. For example, raises are discrete and sometimes quite large. It would be difficult to argue that marginal product takes discrete upward jumps at these points. Product prices, too, change discontinuously. Consider, for example, the price of newspapers. (There are two components to price: the price to the reader and the price to the advertiser. One may be smoother than the other. Similarly, the wage part of compensation may move more smoothly than the pension part.)

Additionally, the amount of discontinuity that one observes depends on the unit of analysis. It is rare in labor economics that researchers are able to analyze data at the level of the firm. Usually what is reported is some average across a large number of firms or individuals. Generally, regression coefficients are presented and these average out the discontinuities. I do not suggest that Frant and Leonard have not performed a valuable service by discussing the variation across firms in accrual patterns. I merely point out that most results in labor economics (or in empirical economics in general) would be more discrete if researchers did not report results that are derived from averaging across a large number of individual units. Thus, the benchmark is different here.

A final point on diversity is that there is no reason to presume that a market displaying a great deal of heterogeneity is inefficient. The fact that clothes come in many sizes and shapes does not imply that something is necessarily wrong in that market. Similarly, pension plans may differ because workers differ in their savings desires, labor force participation behavior, and other assets and wealth, or because firms differ in their credibility and ability to raise capital. I do not suggest that these factors can explain the diversity of plan accrual patterns. I merely point out that one cannot tell the players without a scorecard. Without some clearer statement of what pensions are doing, it is difficult to conclude that variance implies inefficiency.

Some Technical Points

Failure to integrate wages with pensions leads to a minor technical mistake. The authors must select some date of retirement on which to base accruals. Since the pension that a worker receives is a function of final salary and years of service, it is necessary to know that date to compute the amount accrued at each point. Frant and Leonard select the optimal date of retirement, defined as the date at which the expected present value of pension flows is maximized. The problem is that the optimal date of retirement is not the date when pensions reach a maximum. The date depends on the relation of the wage compensation to the alternative use of time as well. This is best seen by examining figure C9.1.

The expected present value of pension benefits is a function of age of retirement shown by the curve labeled EPV. The wage and value of leisure functions are labeled accordingly. They define T as the optimum date of retirement. If there were no pension at all, T' would be the optimal date of retirement. With the pension, the true optimum is at T^*. T^* falls short of T' because the worker must take into account that although his wage exceeds his alternative use of time, he loses pensions by continuing to work. If the value of leisure function were above the wage at T, then T^* would lie to the left of T since it is total compensation and not merely one component of it that affects the retirement decision. The necessary conditions and algebra are spelled out in detail in my paper "Pensions as Severance Pay" (Lazear 1983), but the point is clear: Accruals depend on date of retirement, and that is a function of more than just the pension plan.

The differences in leisure value and wage rates across individuals can help account for different selected dates of retirement by workers who face the same pension plan. In fact, some identifying assumption

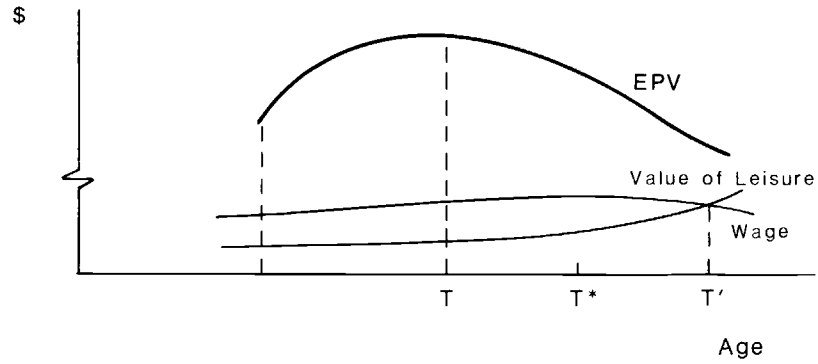

Fig. C9.1 Present value of pension benefits as a function of retirement age.

must be made in order to derive the effects of pension plans on retirement behavior. Most of the literature implicitly assumes that workers are randomly confronted with a pension plan and that all workers have the same tastes. That assumption will not work here. Since all workers within a firm or local government unit face the same plan, differences in retirement dates across individuals must be caused by something else. If it is variations in wage profiles and/or the value of leisure, then it will affect the estimates of the accrual patterns. I am not worried that this will have a major effect on Frant and Leonard's findings, but the implicit assumptions should be made explicit. Additionally, since they assume a wage profile for the typical worker anyway, they can obtain the sensitivity of the computed optimal retirement date to variations in the value of leisure function. My guess is that most of their results will be robust with respect to this kind of variation.

A related issue is that Social Security payments affect the choice of retirement date. For many plans, this is not a problem because their workers do not participate in the Social Security program. But for those plans that are integrated with Social Security, the effects of ignoring it are likely to be important. This is especially so because many plans have offset provisions, which create a deviation between the amount paid by the employer as pension and the amount of retirement income received by the worker. For some questions of efficiency, it is the amount paid that is important. For others it is the amount received. In any event, Social Security offsets that kick in and out at various ages are likely to affect the spikes. Additionally, an examination of these provisions might assist in understanding what pensions are actually doing and why the spikes are there in the first place.

A few minor empirical issues are worth noting. First, it is not clear whether the wage growth figures reported are actually earnings or stated annual salaries. Since older workers suffer health problems, there often is a large deviation between the two. The growth rate that should be used is the one corresponding to the definition of income on which the plan is based. Most plans in the private sector are based on actual earnings, often including overtime. If this is so, then the growth of actual earnings should be used. This definition should be spelled out in the chapter.

Second, it would be useful to perform the simulations with different salary levels. In my "Pensions as Severance Pay" (1984), I found some progressivity in the pension plans. Some have suggested an insurance interpretation of those plans. More evidence on the nature of the progressivity, especially from the public sector, would be welcome.

Finally, if the data are available, it would be informative to relate the various accrual patterns to the characteristics of the workers employed in those firms. The proportion female, black, the average salary

levels, and tenure on separation are obvious candidates. Although some endogeneity is clearly present, even simple correlations might shed some light.

In sum, this chapter does a fine job of presenting a considerable amount of information on pension accrual patterns. It is likely to stimulate more thinking on why pensions take these particular forms. It also may provide some clues to the causes of the growth of pensions during the past thirty years.

References

Lazear, Edward P. 1983. Pensions as Severance Pay. In *Financial Aspects of the U.S. Pension System*, ed. Zvi Bodie and John Shoven. Chicago: University of Chicago Press.

———. 1985. Incentive effects of pensions. In *Pensions, Labor, and Individual Choice*, ed. David Wise. Chicago: University of Chicago Press.