

Wages, Productivity and Retirement

FIRST DRAFT

April 14, 2009

Edward P. Lazear

Hoover Institution
and
Graduate School of Business

Stanford University

for Netspar Conference
The Netherlands

Do workers have the right incentives to retire? Do some stay too long and others leave too early? How should we define “too long” and “too early?” These questions have well-formulated and unambiguous answers that are the outgrowth of a significant literature in personnel economics.

“Efficient retirement,” defined as the time at which an individual should retire, involves a comparison between the value of work and the value of leisure. If a worker’s time is more valuable at work than it is in leisure in retirement, it is socially appropriate for that individual to continue working. Surplus is generated when an individual works in this case because the value of the output he produces more than offsets the value of the forgone leisure. Put differently, the firm at which the worker is employed could pay him what he produces and the worker would voluntarily choose to continue working because that amount exceeds the value of his leisure. Conversely, if a worker’s time is more valuable in leisure than the value of what he produces at work, then it is efficient for him to retire. There is no way that the firm could pay him enough (without taking losses) and induce him to give up his leisure voluntarily.

Self-employed workers always have appropriate incentives to retire efficiently, that is, only when the value of leisure exceeds the value of work. Because self-employed workers receive the full value of their output as wage or profit, they compare this amount to the value of their leisure. If the value of work falls short of the value of leisure, then the worker retires. If not, the worker continues to work. The decision of self-employed workers is an appropriate benchmark against which to judge efficient retirement.

The same is not necessarily true of workers who are employed by others. Such workers

will only retire when the wage they receive falls below the value of their leisure in retirement. If the wage that workers were paid always equaled the value of their output, there would be no difference between the incentives of self-employed workers and those employed by others. But it may well be the case that workers are not always paid exactly the value of their output. Some theories in labor and personnel economics have proposed reasons why wages might differ systematically from the value of output. Those theories have received some empirical support over the years. If they are correct, then workers may have the wrong incentives to retire and a country's pattern of retirement, like that of the Netherlands, may deviate from that which might be expected to prevail were the workers to behave efficiently.

Two Theories of Wage Determination

There are two primary theories of wage determination where wages deviated from a worker's marginal product. The first, which dates back to Becker (1962), is the theory of firm-specific human capital. Firms supply workers with two kinds of training in the context of working on the job. One type of training, which Becker refers to general training, makes a worker more productive not only in his current firm, but in other firms as well.¹ As a result, the firm cannot pay the worker less than his output after he is trained. Were it to attempt to do so, other firms, which value the worker's services to the same extent as the firm that provided the training, would steal the worker away by offering a wage greater than his current wage, but less than the worker's output value. It is for this reason that the theory of general human capital, or

¹It is not necessary that it affect productivity in all firms. Sufficient is that the training is sufficiently useful in a large enough number of firms to produce competition for the worker's services.

more exactly, general on-the-job training, does not provide a motivation for wages which deviate from the worker's output. Competition from other firms that value the worker's services to the same extent force the wage up to the worker's marginal product.

Firm specific on-the-job training is somewhat different in its implications. Firm-specific on-the-job training is defined as training that affects the worker's productivity in the current firm, but does not increase it elsewhere. Examples traditionally include learning who does what at a particular firm and how a given firm implements particular procedures.² Because the training is only valuable in the firm that provides it, there is no direct competition by other firms. As a result, it is not necessary for firms that provide firm-specific training to pay wages that are equal to a worker's marginal product. The wage that the worker is paid after the training is the outcome of bargaining. The worker is worth more to the current firm than to any other firm so the wage that is paid is bounded by the worker's output and his alternative wage. The details are provided in appendix A, but the basic point is that the theory of firm-specific on-the-job training has a very concrete implication: The trained worker is paid more than his wage in an alternative firm, but less than the value of his output.

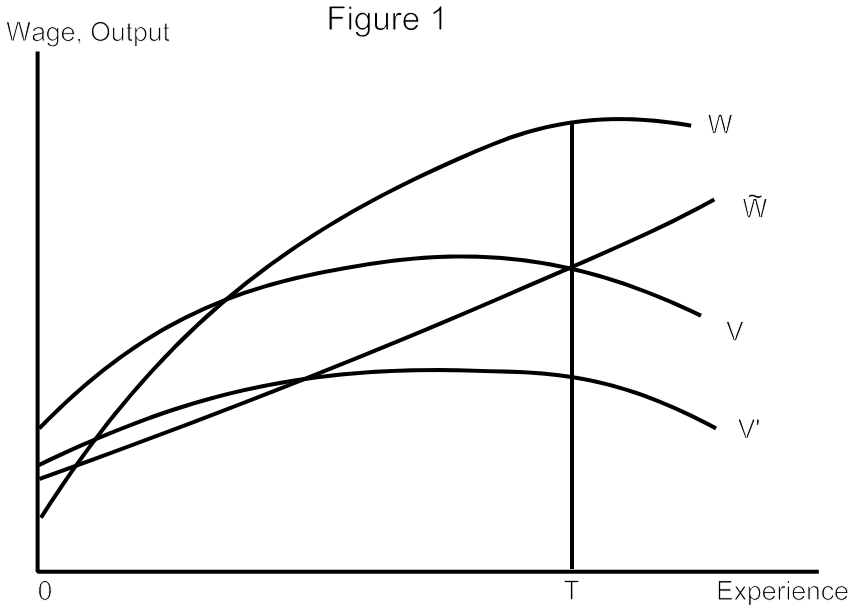
This relationship is important because it implies that the trained worker is too anxious to leave the firm. Because he only takes into account the part of output that he captures, the worker leaves when the social value of work exceeds the social value of his leisure. Workers abandon the firm too often.

The second theory of wage determination that is relevant in the retirement context is

²A somewhat different motivation and structure of firm-specific training is the "skill-weights" approach (see Lazear (2006)).

incentive theory. This theory (see Lazear (1979, 1981)) posits that workers are paid less than they are worth when they are young and more than they are worth when old in order to create appropriate incentives. Appendix B contains a formal model of the structure, but the basic point can be stated verbally. Workers make two choices regarding work. The first is how much effort to put into the job. The second is whether to work at all. When effort is imperfectly observed, an incentive structure that provides motivation to put forth the necessary level of effort induces too much work. In particular, it creates an incentive for workers to want to work too long, that is, beyond the efficient date of retirement when the value of leisure just equals the value of their output on the job.

Figure 1 illustrates the point.



A worker can choose to work at a high level of effort or he may shirk, putting forth a low level of effort. Suppose that a worker's output if he works at a high level of effort is given by the V profile in figure 1. If he shirks, his output is V' . Paying the worker V will not provide much of an incentive for him to perform at the high level of output. The \tilde{W} curve in figure 1 is the worker's alternative use of time. In this context most easily thought of as the value of his leisure. Time T is the date of voluntary and efficient retirement. If workers receive compensation V , they would choose to retire voluntarily at time T , because that is the point at which the alternative use of time just equals the worker's marginal product or payment. Any scheme that duplicates the first-best outcome must have, as one of its features, separation of the worker at time T .

Consider two alternative payment schemes. One pays the worker path V throughout his career and the other pays the worker W , which starts below V , but rises above V . W is constructed such that the present value of the W path, from 0 to T , equals the present value of the V path, from 0 to T . Workers who have access to capital markets are indifferent to paths W and V , if all else were equal. But all else is not equal.

Suppose the worker is being paid according a wage equal to V . As he approaches time T , he must reconsider his decision of whether to put forth the high level of effort, or to shirk. As he nears time T , the incentives to shirk become overwhelming. Imagine the worker's decision one day before time T . On that day, he may either work at the high level of effort, or he may shirk. If he shirks, the worst thing that can possibly happen is the worker gets fired. If he gets fired, he does not get to receive wage V during the next period, but he does get to enjoy the value of his leisure. At time T , V and the value of leisure are equal, so the worker loses nothing by shirking,

even if he gets fired. Thus, there will come a point at which all workers will opt to shirk if they are simply paid their marginal product.

Alternatively, the worker can be paid according to W . Now consider the worker's decision one period before time T . If the worker shirks and is dismissed, he gets to enjoy the value of his leisure \tilde{W} next period, but he forgoes wage W . Since W is set such that it is well above V at time T , a shirking strategy causes the him to forfeit utility. Thus, if the W profile is sufficiently steep, it will induce workers to perform at a higher level of effort than will the V profile. In fact, it would be impossible to pay workers along the V profile, because the worker's output, if V were paid, would fall to V' at some point in the worker's career.

The steeper age-earnings profile W induces workers to put forth high levels of effort, for which they receive additional compensation. If workers were paid according to V , they would be induced to put forth lower levels of compensation and the entire experience-earnings profile would have to be adjusted downward to take into account this lower productivity.

One byproduct of this particular compensation scheme is the need for mandatory retirement. Since the wage path W lies above the wage path V , workers will not choose to retire voluntarily at time T . Instead, they would prefer to continue working as long as the firm were willing to offer wage W . But doing so would not only bankrupt the firm, it would induce workers to remain with the firm inefficiently long. Mandatory retirement is a consequence of using a profile that pays workers more than their worth at their retirement age.

While this theory was first put forth as an explanation of retirement, the principle is much more general. The argument here relates more to incentives than it does to mandatory retirement,

and even in the absence of mandatory retirement, this kind of scheme can be used. In fact, there are alternative institutions which can actually act as well or better than mandatory retirement to insure that efficient separation occurs in the firm. Pensions, the value of which varies with age of retirement, can be used as a device to buy workers out of their current contracts. Efficient pension schemes accomplish the same goal as mandatory retirement, and they sometimes do it better.

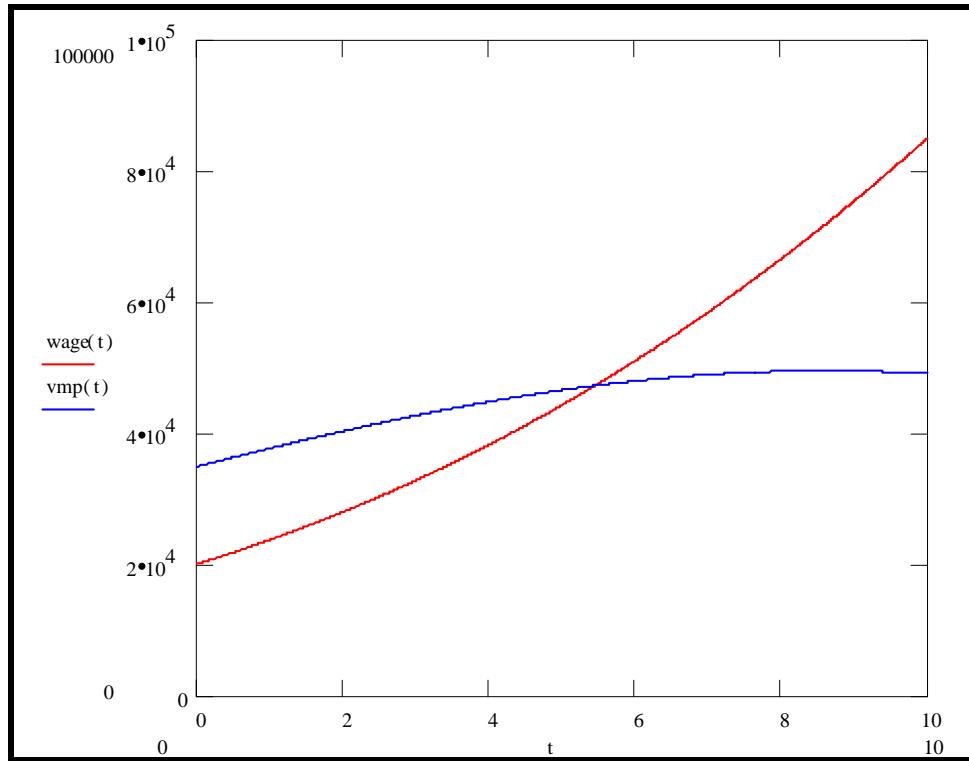
I believe that this view of life cycle earnings corresponds quite closely to that held by personnel managers. Senior workers are paid high salaries in the firm not so much for their current productivity, but as a reward for past productivity and as a motivator for current productivity of their more junior counterparts.

There has been empirical work to support the contentions of this theory. A number of papers by Hutchens (1987, 1989) support the view that upward sloping age earnings profiles are used as a motivator when workers are engaged in tasks that are difficult to observe or measure. Goldin (1986) finds that during the Industrial Revolution, women were more likely than men to be paid a piece rate. She argues that women had less permanent attachments to the work force so that work-life incentive schemes could not be used as effectively for women. Firms paid women piece rates as an alternative, despite the higher measurement costs.

Using Safelite data, I find that productivity rises more slowly than wages over the lifetime. The actual estimated relationship is shown in figure 2.³ Note that output (labeled VMP) starts above and ends below the wage series.

³From Lazear (1998).

Figure 2



There is other evidence on this point, but the importance of the idea for the purpose of retirement is that incentive pay creates a deviation between wages and output. As a result, the worker wants to work too long with the firm. Mandatory retirement was a response to that. In the current environment. The elimination of mandatory retirement in the US through the Age Discrimination in Employment Act and amendments thereto, removed a tool that firms used to adjust retirement to what would have been desirable in the absence of distorted labor supply decisions, either because of incentives or human capital.

Other Theories of Wages

In addition to the theory of specific human capital and incentives, there are a variety of other economic concepts that have implications for wage setting. Most of these deal not so much with the level of wages (and most fail to discuss the structure of wages over the life cycle), but deal primarily with the dispersion of wages within the firm. For example, insurance models suggest that firms insure their workers both over business cycles and across individuals by compressing wages, which makes wage variation less sensitive to output than would be the case in the absence of an insurance motive. The classic papers on insurance include Baily (19??), Azariadis(19??), and Holmstrom (1979). Insurance models have had a long history of theoretical development in the economics literature, but empirical evidence that they are used in practice is scarce and much of it suggests the opposite is true for a variety of different reasons, e.g., see Prendergast (??).

Other theories include those that emphasize politics in the industrial environment. For example, median voter theories that suggest that in the union context or other formalized wage setting processes, policies are dictated by the median rather than the extremes of the distribution. As such, pay tends to be compressed. There is a good bit of evidence that this is the case, in particular union wages tend to be more compressed than non-union wages, and wages in the public sector tend to be more compressed than wages in the private sector.

Monitoring costs provide another rationale for pay compression. Because it is sometimes difficult to measure the output of individual workers within the firm, it becomes a practical impossibility to pay workers on the basis of their individual output or of their output as it varies over time. As a result pay is compressed because differentiation is difficult and so there is no basis on which to base pay, other than subjective evaluation, which is imperfect and not free.

Other Incentives: The State

So far everything that has been discussed would be true in an environment where there is no state intervention, no taxes, no subsidies, nor other distortions to labor market decisions. But

the real world presents a variety of other distortions, many of which are created by state pension programs or by pension programs that are administered, not by individuals or individual firms, but by larger collective organizations.

Taxation is the most commonly treated factor that distorts labor supply. Because taxes reduce the compensation received by workers or increase the amount paid by firms for labor services, taxes reduce incentives to work.

In the retirement context perhaps the more important factor is the subsidy to leisure that occurs by placing earnings tests of one form or another on the receipt of pension benefits. To the extent that pension payments are not independent of work activity or decreases in expected value with retirement age, leisure is subsidized. These subsidies to leisure are very significant and historically have been a key determinant of the amount of labor that is supplied by elderly workers.

The Net Effect

Among the direct labor market effects that are independent of the government, firm specific human capital and life cycle incentive effects go in opposite directions. Specific human capital makes elderly workers too anxious to retire whereas life cycle incentive pay makes elderly workers too anxious to stay on the job.

As mentioned earlier, there is both direct and indirect evidence to this point. The direct evidence comes from examining actual measures of output and actual measures of wages. Unfortunately, such evidence is rare. There are also indirect methods of drawing inference on the direction of the effect. For example, the existence of mandatory retirement is prima facie evidence that at least in firms that feature this particular institutional design, elderly workers are overpaid relative to their productivity, not underpaid, and have incentives to stay too long. An additional indirect method is to analyze the wage and pension structure. For example, in Lazear (Pensions and Severance Pay), one can infer the implicit difference between wages and output that exist among the elderly by examining the price at which firms are willing to buy out elderly workers. Those values then can be examined directly to see whether they are sensible and

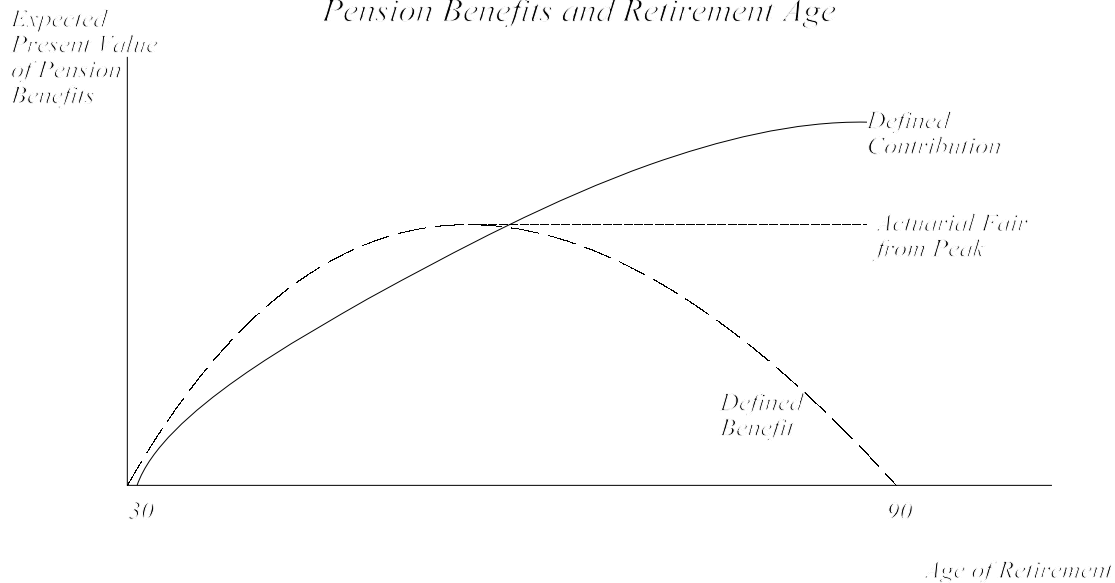
whether they conform with other observable characteristics as theory would predict. Most of the evidence on this suggests that elderly workers are overpaid relative to productivity at least in the fields examined.

Taxes and subsidies to leisure are forces that work in the opposite direction. Both of these factors induce the elderly to supply less labor than would otherwise be optimal in a competitive economic environment.

Determining which effect dominates is obviously an empirical question and is country specific. However, in the context of the Netherlands, certainly at least until the most recent years, it seems clear that the subsidy to leisure implicit in pension formulas was the driving force in inducing workers to retire early, where early is defined both in a conceptual sense (i.e., earlier than a self-employed worker absent a state would retire) and early relative to other countries.

First, pensions are very large in the Netherlands, with levels that approach 80% ratios of pensions to the average wage. But that is not sufficient. If the pension amount is granted independent of work then there is an income effect inducing people to take more leisure, but no substitution effect away from work. Further, the income effect depends in large part on an individual's position in the income distribution because the pension payment is offset by the taxes imposed on individuals that are necessary to cover the payment. Furthermore, the income effect on leisure would not necessarily show up most prominently during the retirement years. For example, it could take the form of longer vacations during the work years. As a result, the actual structure of the pension is a key consideration. Figure 3 illustrates the point.

Figure 3
Pension Benefits and Retirement Age



The typical defined benefit plan has the pattern of value shown by the inverted “U” shaped curve in the figure. The easiest way to see this is to consider someone who begins work at a firm at age 30 and who will live with certainty until age 90. Were he to “retire” on the day that he entered the firm, he would have accrued no benefits and so the value of the pension would be zero at that point. Conversely, were the individual to work until the instant at which he died, he would receive no benefits as well. Even though the pension payment per year might be extremely high he would have no years of life left over which to receive the benefits. Because the pension value must be positive at some year between 30 and 90, a connect-the-dots approach ensures that the function must be an inverted “U” shape. Of course, it need not be smooth. It can have spikes, nonlinearities, discontinuities, zigs and zags, but the basic structure must be one of initially rising and eventually declining with age. As a result, there is a disincentive to stay beyond a certain age. This does not mean that everybody retires at the peak of the defined benefits function. Other factors enter including the wage that the individual receives for continued work as well as the non-monetary value of remaining on the job. But the point is that individuals give up retirement benefits in order to continue to work and that is a cost of delayed retirement.

By contrast the upward sloping curve, labeled “defined contribution” shows that, at least in an expected sense, the value of a defined contribution plan necessarily rises monotonically with

age of retirement. Because additional contributions are made each year and because the pension fund is actually owned by the worker, the expected amount received in pension benefits, which can even be bequeathed, always rises with retirement age. This provides incentives to remain on the job. Once again, this does not mean that workers who have defined contribution pension plans work forever, because other factors enter. The most obvious factor is increases in the value of leisure associated with advancing age. Once a worker gets beyond a certain age the value of leisure may be extremely high. For example, leisure may take the form of repairing one's body after a major health event. But once again, the point is that the defined contribution pension plan provides incentives to remain on the job rather than to retire early.

A final variant, and one consistent with some changes that have been made to the Dutch pension structure in recent years, is to make the pension payments actuarial at some point. In the diagram, pension payments become actuarially fair at the peak of the defined benefit structure. Workers accrue pension benefits up to age 55 in this diagram and from then on the age of retirement has no effect on pension benefits. This implies that the benefit-per-year must rise as retirement is deferred. In this stylized example, were the worker to work to the instant before death, the pension flow would approach infinity to ensure that the total amount received equal the value of the pension shown by the green line. This is both impractical and unrealistic. Most pension structures, even actuarially fair ones, impose caps at a maximum age at which retirement benefits per year continue to escalate or have lump-sum payments to accommodate later retirement.

Even the actuarially fair pension structure does have implicit in it incentives to stay on the job until at least to the peak. Wages are higher, other things equal, for ages before the peak than they are after because the true wage is the observed wage plus the increment to pension value that accrues from working one additional year. Since the change in the pension value is positive for ages up to the peak and zero thereafter, this hybrid structure creates positive incentives to stay up to the benefit value age. It is neutral incentives thereafter. Again, while these are forces that work in the direction of favoring the peak, they are by no means dispositive in ensuring that workers retire at or beyond the peak as shown in the diagram.

In addition to incentives that alter the age of retirement, pay compression creates other, often adverse, incentives. Buyout programs are the most obvious manifestation of what is sometimes adverse selection problem induced by the interaction between a pension buyout and a compressed wage structure. In the United States, as a result of changes in the Age Discrimination in Employment Act, mandatory retirement was prohibited for most jobs. This was a particular problem in academia. A number of universities decided to buy their workers out by offering a sweetened pension plan to workers who would agree to retire early. These plans often backfired, resulting in the best professors leaving and the worst professors staying on. This classic example of adverse selection occurred primarily because the better professors were able to find alternative jobs at other institutions whereas the poorer professors had no alternatives and had to remain at their current employer simply because the pension buyouts were insufficient to cover their remaining years. This is a result not so much of the buyout per se, but of the interaction between the buyout and the wage. Because the poorer quality workers were overpaid relative to the high quality workers, they were unable to find jobs at wages close to those they received at their current institutions. Because the high quality workers were underpaid relative to their productivity, the buyout was a catalyst that induced them to move elsewhere. Were there no pay compression, this kind of adverse selection would not have been a necessary consequence. Even the poor quality workers would be able to find alternative jobs that would pay them an amount close or equal to what they were receiving on their current job because their low productivity would be matched by their low wage at the current firm. Similarly highly productive workers would get higher wages, but their outside alternatives would also be in line with those that they received inside and there would be no bias toward retirement of a particular kind of worker.

Pay compression is an important incentive feature of most wage structures and tends to result in adverse selection particularly among retiring workers. But it is neither a necessary consequence of the pension structure nor of the desire to buy workers out. Indeed, a well-tailored pension buyout structure could reverse the pattern and could keep the better workers on the job. An extreme form of this would simply offer higher buyout payments to workers who are currently overpaid and lower buyout payments to workers who are relatively underpaid. This is not without its problems, however, because it creates moral hazard when workers know that they can increase

their buyout offers by performing poorly in their current assignments.

The general point here is that while adverse selection is an important issue it does not result in a bias toward early or deferred retirement per se, but rather a change in the nature of the individuals who retire early.

Most of the discussion so far has focused on worker incentives, but firms also have incentives as a result of the pension structure and the wage productivity relationship. In most cases, worker and firm desires go in opposite directions. For example, in the case of life cycle incentives, where workers are underpaid when young but overpaid when old, workers have a desire to stay on the job whereas firms have a desire to rid themselves of these older workers. But with respect to state created distortions and incentives, either through the tax structure or through subsidies to leisure in retirement, incentives between workers and firms tend to be aligned. For example, taxation of wages, particularly among elderly workers, reduce incentives for workers to remain on the job and for firms to employ them. This sometimes is put in the context of “the incidence of a tax.” Even when taxes are levied on employers for hiring workers or retaining their services, it has been shown that a significant fraction of the tax is borne by workers and this is true in the context of taxes imposed on elderly workers as well. Thus, the cost of employing workers goes up. Either the firm bears the tax or workers bear it implicitly or explicitly. Taxes levied on worker income reduce incentives to work directly. Those levied on firms that reduce worker wages reduce labor supply incentives as well.

As mentioned earlier, the subsidy to leisure implicit in many state pension plans is likely to be the most important factor determining age of retirement among older workers. As a result, these distortions must be offset by changes in the pension structure or other methods if a country is worried about retaining workers beyond a certain age.

Supporting Pensioners

As a result of changes in demographic patterns that have and will continue to occur in most developed countries, dependency ratios are expected to rise in the future. This means that a smaller group of workers will have to support a larger group of elderly recipients of benefits which will place a burden on the fisc and force governments to seek solutions.

Adaptation to this changing demographic structure is likely to take one of three forms. Either the government will have to levy higher taxes on workers; it will have to mandate or create incentives to increase in the length of the work life; or it will have to reduce the state component of pension benefits. If it chooses the latter it is likely to increase the requirements for pensions done privately.

The extent of the challenge facing pension structures depends on the nature and depth of changing demographics. Some countries will be hit harder than others. Countries that have large flows of immigrants, for example, are less susceptible to this problem than others because immigrants tend to be at the younger part of the age distribution. That said, a number of studies have shown that immigrants take awhile to become significant net contributors to the budget. Other factors that affect dependency ratios include especially birthrates among the natives and Japan and Italy are among the leaders in facing the most significant problems going forward because of their low birthrates.

In order to choose between higher taxes on workers and increased work life or reduce state pensions it is important to think more concretely about the goal of a state pension system and how it would be affected by the various measures. Most state pension plans incorporate implicitly or explicitly three features: forced saving, insurance, and redistribution.

Forced saving is an appropriate goal of state pension plans when any of the three following conditions hold.

Moral hazard: individuals know that the state is unwilling to allow its elderly to starve. As a result an individual might have insufficient incentive to save on his own because he knows that the state will take care of him in old age. Forced saving remedies that difficulty.

Imperfect foresight: People make mistakes. Some are unwilling to save simply because they do not believe that they will have a need for income in old age. This may be based on expectations about length of life that turn out to be inaccurate or because they expect to have an income source in their latter years that does not materialize. The cause is unimportant but imperfect foresight makes forced saving another remedy for a problem that may occur and impose costs on the state.

Hyperbolic discounting: Some have emphasized the view that people's behavior is time inconsistent. There is experimental evidence to support this viewpoint. Whether hyperbolic discounting is important in the real world or not remains to be shown, but its potential existence is another reason why forced saving might be required.

Another current goal of pension plans is to provide insurance. Here, to the extent that the pension plan is used at all for insurance, it is to insure unanticipated events that occur late in life like worse-than-expected health outcomes. Pensions are not the right tool for insuring all adverse health outcomes. Dealing with poor health that occur earlier would presumably be incorporated into a general system of redistribution and health care and income insurance that is incorporated into the social welfare structure. Note that making a pension plan actuarial actually harms the insurance function because those who become ill early receive a lower per-year stream of income despite the fact that their needs for that income may increase. This, of course, depends in part on how the health care system is financed and how health benefits are allocated. It also depends crucially on how other parts of the state transfer structure treats those who become ill.

Caps, floors and other quirks in the pension formula create sometimes implicit redistribution from one income class to another. These are complicated and country specific. They also change over time and affect different groups in a variety of ways. For example, groups with long life expectancies tend to be favored relative to those with shorter life expectancies. Sometimes this works in the direction of redistributing from high income to low income as is case in male-female life expectancy differences. Women, by virtue of living longer, receive more benefits for a given lifetime or final income than men. Sometimes redistribution works in the opposite direction. For example, because the poor have shorter life expectancies than the rich,

pension systems that assume all have the same life expectancy redistribute away from the poor, at least with respect to this component. Additionally, in a number of systems the very poor are not taxed to fund pensions. And in some systems like that in the United States, there is a cap on income that is subject to taxation, which works to the benefit of the rich.

Some of these ceilings and floors are appropriate from a straight economic point of view. For example, conditioning on low income is a way to exempt from taxation individuals who have suffered negative transitory shocks to their income. It is probably not a good idea to tax those individuals during those phases of their worklives, in part because liquidity constraints make it difficult to finance consumption when their income is transitorily low. But most redistribution is better handled through the general tax system and broader version of social transfers that is not tied to age or retirement per se. There is no obvious reason why a poor 63-year old should receive benefits from one system whereas a poor 66-year old should receive benefits from another system.

Minimizing Labor Supply Distortions

Unfortunately most state pension plans have associated with them incentives that distort labor supply, especially later years. An ideal state pension plan would eliminate these distortionary incentives. The goal would be to mimic the labor supply behavior that our stylized self-employed individual absent a state would exhibit. Again, the goal is not later retirement but efficient retirement, where efficient simply means that individuals make the right choice on labor supply based on the value of work versus the value of leisure. I focus on state pension plans rather than private ones simply to avoid the discussion of whether the state or private sector handles pensions more effectively. That is an interesting and important discussion but outside the scope of this paper.

A state defined contribution plan such as that used in Sweden (notional defined contribution and financial defined contribution) eliminates most of the distortionary incentives associated with defined benefit pension plans that I have described earlier. Individuals would be

required to set aside a fixed amount of their income each year and that would then be contributed to a pension fund in their own names. There would be no redistribution nor insurance through this system. Redistribution and insurance would be handled through general revenues that would be separate from the pension structure and would likely treat all individuals similarly, whether they were in retirement years or not.

This structure solves the problem of rising dependency ratios by putting each individual on his own footing, but there are issues that require mention.

First, recall that incentives based wage structures over the life cycle create a wedge between wages and output that induce senior workers to want to stay too long. This distortion still exists even with a defined contribution pension plan, but to the extent that taxes (for purposes other than funding his own pension) lower take home pay there is an offset to this effect.

Second, there is still a need for insurance. For example, individuals who suffer an unexpected blow to their health in their later years may need to retire earlier than others. One would like to have the flow of pension benefits higher to those individuals and not adjusted in an actuarially fair way. Again, however, this need not be done through the pension system per se. There is no reason why general welfare structures could not take into account pension earnings in the same way they take into account any other wage earnings.

Third, there is a problem with transition because most countries currently have a pay-as-you-go system where the current young generation cares for the current old generation. Moving to a stand-on-your-own basis kind of pension plan, even one run through the state, creates a one-generation transition problem. There are a variety of ways that public finance economists have thought about dealing with this problem, one of which includes deficit financing which would be repaid through higher taxes later. Repaying the deficit is offset by the elimination of a liability inherent in pay-as-you-go systems that require future generations to take care of their parents.

Other Policy Issues

A frequent suggestion for dealing with demographic shifts and budgetary shortfalls of state pension structures is to tie benefits to income or to lifetime wealth. The rich elderly, it is

argued, do not need the state pension payments and a significant amount of resources could be saved by reserving benefits for those who need them. There are adverse incentives associated with this suggestion. To the extent that pension payments are tied to current income, such a system would distort labor supply incentives. To the extent that it is tied to lifetime wealth or some measure of accumulated assets, the structure would distort incentives to save. Finally, at least in the United States, this suggestion has been championed by conservatives and opposed by liberals. The conservatives have bought into the notion that the wealthy do not need the benefits and it is more important to enhance the solvency of the Social Security system. Liberals have opposed this in large part based on the philosophy first espoused by Franklin D. Roosevelt, who worried that a system for the poor will be a poor system. A concern is that to the extent that benefits are limited to a sub-group of the population, the system will lose political support that it needs to be maintained. Whether this is true or not is an empirical question. Moving to a defined contribution structure that is coupled with redistribution through general revenues would make tying benefits to income redundant and unnecessary.

Another suggestion that has been made is that taxes on the elderly could be reduced to reduce early retirement incentives that exist in the current system. Again, a better approach would be to remove the distortions directly by eliminating the subsidy to leisure, say through a state structured defined contribution plan. Short of that, reducing taxes on the elderly would move in the right direction. It comes at a cost, however. Those tax revenues must be made up elsewhere and that would imply an increase in taxation of the young. The general rule in public finance is that broadly-based taxes are less distortional than narrowly-based taxes, and that taxes should be levied on goods and services with the least elastic demand or supply. Empirical evidence suggests that reducing taxes on the elderly and spreading the tax burden more broadly on the rest of the population would probably move in the right direction, but again, this is a second best approach.

Conclusion

This essay has attempted to discuss a wide variety of issues that relate to retirement

behavior, primarily by focusing on compensation schemes and incentives implicit in them as well as in pension formulas, both state and private. The deviation of wages from output either because of specific human capital, lifecycle incentive-based compensation plans, insurance or other industrial political features, causes a distortion in labor supply incentives. This could be particularly pronounced for the elderly workers. Additionally, taxes and subsidies that are imposed by the state may also create major labor supply distortions and sometimes these work in directions opposite those from the private incentives. Finally, in considering the challenges that countries will face as the demographic structure changes, a move toward a defined contribution structure that makes benefits independent of labor supply behavior late in life is one approach that should be considered.

Appendix A

Wage Determination With Firm-Specific On-the-job Training

The theory of wage determination with firm-specific on-the-job training can be captured in a two period model. In period one, the training is provided while the worker works. In period two, the worker only works.

Let output in the current firm be given by V'_t , where t refers to the period. Output at all other firms is given by V_t . The wage that the worker receives in W_t . It is clear that

$$V_2 \leq W_2 \leq V'_2$$

Since period 2 is the last period, the wage can never exceed the worker's output at the current firm, V'_2 , because the firm will not pay it. Nor can the wage ever fall short of V_2 or the worker would choose to work elsewhere. The exact level of W_2 cannot be determined without a more specific bargaining model. For example, Nash bargaining would set

$$W_2 = (V_2 + V'_2) / 2$$

but all that matters is that W_2 does not exceed, and in general falls short of V'_2 .

The wage during the training period must exceed the net output, V'_1 , where net output is defined as output minus the cost of training.

Ex ante competition by firms to hire labor in period one implies that

$$W_1 + W_2 = V'_1 + V'_2$$

So

$$W_1 = V_1' + (V_2' - W_2)$$

But since $W_2 \leq V_2'$, $W_1 \geq V_1'$ and normally will take the form of strong inequalities.

Thus, wages exceed output in the current firm in period one and fall short of output in the current firm in period 2.

Appendix B

This two period model shows that it is necessary to have wages greater than output in period two to provide proper incentives in period one. It also shows that when the wage profile provides correct incentives, it distorts the work-leisure decision.

Key is that there are two margins. First, the worker must choose effort-per-hour. Second, the worker must decide whether to work or not.

Assume that the worker can choose to work at level low or high at a pain cost of C_L and C_H , respectively, where $C_L < C_H$ and where the output from working at the low level and high level is Q_L and Q_H .

Suppose further that

$$(B1) \quad Q_H - Q_L > C_H - C_L$$

which implies that it is efficient for the worker to work at the high level of effort. Wages are stated at the beginning of the period and paid at the end of the period. They cannot be made contingent on the level of output. Piece rate pay makes compensation directly contingent on output. The assumption here is that output cannot be observed without a sufficient lag and that the worker cannot finance his consumption in the absence of a wage payment, which precludes ex post contingent pay. Additionally, for now, pensions are not permitted.

First consider period two. Since wages are not contingent on output in period two, the worker necessarily choose to work at low effort, producing output Q_L . The firm knows this and would therefore pay wage

$$W_2 = Q_L$$

but for the implications that such a wage scheme would have for effort in period one.

Suppose the firm tells the worker that it will assume that the worker is going to work at the high level of effort, but if he is caught “shirking,” he will be fired and will not have the opportunity to work in period two.

If the worker works at the high level of effort in period one (and low in period two), the rent that he receives is

$$(B2) \quad W_1 - C_H + W_2 - C_L$$

If he works instead at the low level of effort in period one (and low in period two), the rent he receives is

$$(B3) \quad W_1 - C_L$$

Recall that he is then fired and does not have the opportunity to work in period two.

The worker chooses the high level of effort only if

$$W_1 - C_H + W_2 - C_L > W_1 - C_L$$

or if

$$(B4) \quad W_2 > C_H$$

In order for the firm to earn zero profit (the competitive equilibrium) , it is necessary that total wages equal total wages paid or that

$$(B5) \quad W_1 - Q_H + W_2 - Q_L = 0$$

To provide appropriate incentives for the worker to work or not in period two, it is necessary that

$$W_2 = Q_L$$

Then the worker will work in period two if and only if

$$Q_L > C_L$$

which is the efficiency condition.⁴

Consider the case where $Q_L < C_L$ so that the worker should not work. Under these circumstances, paying the appropriate period two wage that prevents him from working implies that

$$(B6) \quad W_2 < C_L$$

But if (B6) holds, then (B4) cannot. A contract that induces the appropriate work - no-work decision in period two results in too little effort in period one. Stated alternatively, if the firm sets up a contract that provides the proper effort incentives for effort in one period one, it will result in work too frequently in period two. If (B4) holds, then

$$W_2 > C_L$$

When $Q_L < C_L$, work will occur in period two, even though the worker should retire after period one from an efficiency point of view.

Pensions that are paid after the end of the work life and are contingent on performance can help to solve this problem, which is fundamentally that there are more margins than instruments.

⁴Because there are no externalities, the efficiency criterion is equivalent to maximizing worker utility subject to a zero profit condition, or conversely, to maximizing profit subject to wage sufficient to cover the costs of worker leisure.

References

TO BE COMPLETED LATER