Issues in Social Security Reform with a Focus on Spain
Peter Diamond
Massachusetts Institute of Technology

Like many other countries, Spain is contemplating further changes in the provision of retirement pensions, having approved guidelines for reform (Pacto de Toledo) and made a start on reform in 1997 legislation. There are three good reasons for making further changes. One is that the cost of retirement pensions is expected to grow much faster than the economy under current rules. Second, the current system does not have good labor market incentives. And third, the provision of benefits to older survivors (primarily widows) does not provide a good pattern of replacement rates for what was received when both husband and wife were alive.

I begin by discussing labor market incentives and the shortcomings of current law. Then I consider the financial picture, considering alternative ways of cutting benefits and the possible role of funding. In considering funding I consider the issue of portfolio choice and the contrast between centralized investment and individual accounts. Then, I briefly consider survivor benefits.

I. Labor market incentives

A. Incentives for young workers

1 Prepared for Lliçó d'Economia, Universitat Pompeu Fabra, Barcelona, October 29, 2001. Some of the material has appeared in lectures on other European social security systems. The author is grateful to Juan Jimeno for extensive comments, to Michele Boldrin, Joan Gil, and Guillem López for comments, to Ivan Fernandez-Val for research assistance, and to the National Science Foundation for research support under grant SBR-9618698.

2 For projections of the Spanish pension system, see Jimeno, 2000.

3 For more detailed analysis of labor market incentives, see Boldrin, Jimenez-Martin and Peracchi, 1999, 2000a and b.
Social security taxes affect the decisions of workers and employers. Anticipation of larger future benefits as a consequence of current earnings also affects their decisions. This underlies the importance of designing benefit rules that give good labor market incentives. Consider how benefits are determined for a worker covered by the general regime and retiring at age 65 in 2002 – with a changing law, the rules vary a bit from year to year.\(^4\) Provided the worker has been in the system for at least 15 years, benefits are determined by a two-step process. The first step is to average earnings subject to a minimum and maximum (bases de cotización) for the last 15 years (the first 13 of these are indexed up by the CPI, while the last 2 are not). This is called the "base reguladora." The pension is the result of multiplying the base reguladora by a coefficient that depends on the number of years of contribution. The coefficient is 0.50 for 15 years of contributions and increases by .03 for each additional year up to 10 additional years and .02 for each additional year up to 10 further years, until it reaches 1.0 for 35 years of contributions. The coefficient does not get larger for additional years beyond 35.

Consider a worker who expects to work at least 15 more years before retiring. While the taxes currently paid add to the number of years used in determining the coefficient (assuming that will be less than 35) the exact amount of taxes does not add anything to the level of benefits. Thus, from the perspective of distorting taxes in the labor market, all of the tax is a distortion on any marginal decisions affecting how much to earn during a year that already has enough earnings to count as a year toward benefit calculations. Earnings decisions reflect not just the number of hours to work, if there is control over that variable, but also choices across different jobs that might have different levels of pay and different disutilities of work.

Next, consider a worker in the last 15 years of work (and earning less than the taxable maximum). This worker is paying the tax. But this worker also sees an increase in benefits as a consequence of work. Ignoring issues of indexing and discounting, we see that the marginal euro earned (for a worker receiving more than the guaranteed minimum pension and ignoring the

\(^4\) For someone retiring in 2001, 13 years are used, as the law phases in a longer averaging period. The government has indicated an intention to take further steps in increasing the number of years affecting benefits. There are separate systems for some workers.
presence of a maximum pension) adds 1/15 euro to the base reguladora. Depending on the number of years worked, this adds between 1/30 and 1/15 euro to the benefit. This larger benefit will be collected for the rest of the worker’s life (and will affect the survivor benefit of a surviving spouse). With life expectancy at 65 being substantial and growing, a worker with average life expectancy is subsidized at the margin, not taxed. Adjusting for inflation and interest rates would not change this conclusion.

Economic theory tells us that it is inefficient to have such different levels of distorting taxes at different ages. Not only is the use of a small number of years poor for incentives, but it also introduces an extremely capricious element in the redistribution that is done through social security. Those who happen to have good earnings years at the end of their careers are winners at the expense of others. Who are the winners from this limited basis for benefits is complex, reflecting both earnings patterns and the impact of both the cap on earnings used for calculating benefits (máximo de cotización) and the presence of the minimum and maximum pensions. Disproportionately, high earners have steeper age-earnings profiles, but this may not be so important for benefits at present. Reliance on a longer period of earnings would change the pattern of benefits across workers to be more fair. This need not result in lower benefits on average – that depends on how benefit calculations are changed along with any change in the length of the averaging period used in benefit calculations. That is, fairness and efficiency can be improved without necessarily lowering average benefits. Although lowering average benefits for the long run is an important consideration.

In the past, in many countries it was common to base pensions on earnings in a small number of years and on the total number of years of service. The recent reform moving the basis of benefits from an average of 8 years of wages to an average of 15 years of wages was a move in the right direction. But, it does not go nearly far enough. Indeed the system in Sweden was based on a 15 years earnings average. The Swedes found this unsatisfactory and have moved to a system that reflects the level of earnings in all years. In particular they have adopted what is called a notional defined contribution (NDC) system for the major portion of their system.
(having small funded individual DC accounts as well). Italy has adopted such a system for currently young workers, although their current system will continue to hold for older workers. I will describe this approach to a benefit formula below.

Economists widely agree that a system with a short averaging period is a poor design for a national system. Harmful to the efficiency of the economy are the distortions of labor supply incentives and the creation of incentives to manipulate the formula by concentrating earnings in the small number of years with earnings that count for benefits. As an extreme example from my home city of Boston, the subway system bases pensions on earnings (not base pay) of workers at the end of their careers. As a result older workers do a great deal of the overtime work in the system. This has caused accidents when older workers, having put in too many hours, fall asleep at the control of trains. One need not go so far as endangering lives to see that such systems are harmful.

The important lesson is to base benefits on many years. As I said, Sweden is moving to an NDC system which uses all years of earnings (actually taxes paid). To take two more examples: The US defined benefit (DB) system uses the 35 best years (relative to the wages generally in the economy). Chile’s defined contribution (DC) system bases benefits on the payroll taxes paid in all years (compounded by the rate of return until retirement). Both of these approaches do a good job of labor market incentives. Evaluation of the exact details of design is complex because of the interaction with other labor market interventions and issues of insurance against earnings fluctuations. Avoidance of national systems using a small number of years is important. It seems to me important to review the basic structure of benefits here, not merely to increase the number of years used to determine the base reguladora. If many years (including possibly years of zero earnings) are being used to determine the base for benefits, also counting the number of years of contributions may not be a useful additional element. None of the systems I have just mentioned count years of contributions in determining benefits (although a minimum of 10 years in order to receive retirement benefits is present in the US. I have not seen anyone consider this issue, which is ripe for analysis.
B. Incentives for retirement

The impact of social security on retirement depends on two separate rules. One rule is the determination of what workers receive when they can first retire and claim retirement benefits. The second rule determines the financial consequences of continuing to work and so delaying the start of benefits beyond the point of first eligibility. Economists recognize that it is difficult to form a precise judgment on exactly when workers should first be eligible to claim benefits - there are many different effects coming from a change in early eligibility and some of them are extremely difficult to measure. Here in Spain, while the normal retirement age is 65, benefits are generally available after age 61 (although some qualify after age 60) and a sizable majority retire early.\textsuperscript{5} It needs to be recognized that an early retirement at a reduced benefit may well make sense for some workers. For example, in the US those retiring earlier have higher mortality rates (Waldron, 2001). Workers vary in many ways – in life expectancy, in job opportunities, in enjoyment or dislike of work, in degree of difficulty in continuing work, and in wealth and benefits for financing retirement. Thus, different workers should be retiring at different ages. A good system needs to have flexibility in retirement ages with sensible incentives.

Flexible retirement options are important. The age for first eligibility for benefits needs to be set low enough so that once there are significant numbers of workers for whom retirement makes good sense, they have access to benefits. But setting this age in recognition of a significant minority of workers implies that most workers should continue working beyond the age at which they are first eligible for benefits. And they will not continue working if the financial penalty for delaying benefits is too large. To avoid such a large financial penalty, benefits that are delayed need to be significantly increased. So approximating what is called actuarial fairness is needed to make sense of a good policy of first eligibility. That is, as a first

\textsuperscript{5} The rules described in the text refer to workers who were affiliated before January 1, 1967. The government plans to modify the rules for those entering the system after this date.
approximation, the social security budget should break even when a worker decides to delay retirement for a year, which is the definition of actuarially fair.

In considering the incentive to continue working once eligible for benefits, we have to consider the effect of continued work on benefits. First, continued work generally changes the base reguladora. The base could go up or down, depending on the current level of earnings compared with that of 15 years earlier. This particularly discourages part-time work or a drop to an easier, lower-paid “bridge job,” as is common in the US. Also the minimum and maximum pensions affect the increase in benefits, making this calculation irrelevant for many workers. I will ignore the minimum (and maximum) pension in this discussion of retirement, returning to this issue briefly when discussing income distribution below.

Second, continued work adds another year to the number of years of contributions. This raises benefits by 2.0% or 3.0% unless there are already 35 years of work. Thus this portion of the determination of benefits has a smaller incentive to work for long career workers than short career workers. It is not clear that there is any logic to that.

Third is a change in the adjustment of benefits because of retirement before age 65. The adjustment factors are .60 or .65 for 60 year-olds. These factors rise by .08 or .07 for each year, until they reach 1 at age 65. (The second coefficient applies in case of displaced workers). This common pattern of a linear adjustment for early retirement does not match with the increasing rate of mortality as people age. Working from 60 to 61 increases the factor from .60 to .68, a 13.3% (8/60) increase, while working from 64 to 65 only increases the factor by 8.7% (8/92).

---

6 The government plans to modify this structure of 8% penalties by allowing early retirement only for those with at least 30 years of work and by varying the penalty for early retirement between 6% and 8% depending on years of work, with larger penalties for those with shorter careers. Larger penalties imply a greater incentive for additional work, so the incentive to continued work from this portion of benefit determination will be larger for short career workers. A further element is that an additional year of work may change the rate of penalty for early retirement, adding to the incentive to work.
The adjustments should be increasing not decreasing in percentage terms to avoid large implicit taxes.\(^7\)

This analysis is for a worker whose benefits are determined by these rules. Since there are both minimum and maximum pensions, workers who are at these limits do not get additional returns from continued work. These represent much larger distortions than those for workers who are between the minimum and maximum. This suggests that there may be other methods of improving income redistribution that are not so distortive (following the principles of optimal taxation applied to social security).\(^8\)

At present there are no further increases in benefits for work beyond age 65, although the government is considering a change in the rules. Either benefits should be paid independent of retirement after that age or benefits should increase. One rule or the other should be followed to avoid having large financial penalties for continued work past age 65. Indeed, the Toledo pact has called for allowing people to work beyond age 65, with an increase in benefits, although that has not been enacted.\(^9\) (Ending payroll taxes at that age would be a help.)

---

\(^7\) Economists are agreed that a delay in the start of benefits should trigger a significant increase in eventual benefits. As a rough approximation, the increase should be what is called actuarially fair. That is, the increase in benefits should cost social security about as much as is saved by not paying benefits for a year for the typical worker. I say about as much since there is nothing magical about exact fairness, and some arguments why some deviation from exact fairness may be a good thing. But approximately fair is important. As mortality rates increase with age, the size of a benefit increase needed to match a year without benefits gets larger.

\(^8\) On this subject in general see Diamond, forthcoming.

\(^9\) Spain has moved toward universal coverage and a more uniform system. A unitary public system does not mean that all workers in all jobs should make the same preparation for retirement. Obviously not. Some careers are naturally longer than others. Some workers and their families expect to live longer after retirement than others. The way to respond to this need is through a combination of individual savings decisions and provisions associated with jobs. Job-related pensions should be part of negotiations between employers and unions or nonunionized workers. In order to fit with a mobile labor force, these pensions need to be defined contribution systems. This approach, using supplementary DC pensions, also makes sense for civil servants. And the desirability of leaving room for this flexibility should be kept in mind when deciding how large a public system to have. That is, taxes to support mandatory retirement pensions should not be so large as to prevent supplementary pensions. It is also the case that labor market
II. Finances

Under current law, expenditures for retirement benefits will rise much faster than GDP. Thus it seems appropriate to consider cutting benefits as well as increasing taxes. I will not consider issues in choosing a particular mix between benefit cuts and tax increases (i.e., the size of the program). I will discuss the basis (and method) for making different percentage cuts in benefits for different cohorts, in reflection of the increasing life expectancies we expect for future cohorts. And I will consider the timing of any increases in taxes, recognizing that collecting taxes prior to expenditures amounts to creating a fund for future benefits. I will also consider alternative approaches to managing such a fund.

A. Increased longevity of future cohorts

To set up the issue of increased longevity, let us consider what a worker without social security would do upon learning that life expectancy would be longer, either for the worker or the worker’s spouse. Such information is likely to come with changing expectations of being able and willing to work at advanced ages, but I will ignore that. Upon learning that you are likely to live longer, the previous combination of consumption before retirement, consumption after retirement and retirement age is no longer financially feasible. That is, the previous length of career is no longer sufficient to finance the previously planned lifetime consumption package. A sensible response would use all three available margins. Working somewhat longer is sensibly part of the package. So too is consuming somewhat less both before and after retirement. Consuming less before retirement is an increase in savings or contributions to a pension plan. Consuming less after retirement is having a lower monthly benefit for a worker relying on a pension for retirement income.

---

incentses depend on how large the system is. If tax rates are very high, that also is harmful to the efficiency of the labor market.

10 On this subject, see Diamond, 1995.
We need to examine how a pension system should adapt to the same arrival of information for a population relying primarily on the pension system for its retirement income. For a decrease in consumption before retirement there should be an increase in contributions. For a decrease in consumption after retirement there should be a decrease in monthly benefits for any given retirement age. To have workers choose longer careers in order to have higher benefits, a system needs good incentives to work longer. Only then will workers partially offset the decline in retirement benefits by working longer. Thus a good financial incentive to work more, including an increase in future benefits as a consequence of working longer, is a critical part of enabling workers to adapt to changing circumstances. Any improved ability of workers to earn at advanced ages would further strengthen the importance of significant incentives for continued work. I have already discussed the returns from later retirement and will not repeat that here.

I have described this for a given cohort of workers who receive new information. We can use this as the basis for considering how retirement income systems should vary across cohorts when later cohorts are expected to live longer than earlier ones. In addition to living longer, later cohorts are likely to be wealthier than earlier ones. That is, we expect technological progress to continue raising wages and we expect technological progress and rising incomes to contribute to continuing declines in mortality rates. Since we expect a population receiving more income to take more leisure, this increased wealth will tend to offset the impact on retirement ages of an increased ability to earn. Historically the wealth effect has dominated the improvements in mortality and health and there has been a long-term trend to earlier retirements. This has been further encouraged at some times by social security systems, although the trend to earlier retirements is not restricted to times of increasing generosity of social security. I do not know which element, wealth or health, will dominate the trend in the future. While it is useful to predict the trend to have a sense of system costs, this prediction becomes less important when systems accommodate different retirement choices without large financial impacts. This will be the case when adjustments of benefits for the age of retirement are close to actuarially fair.
Moreover, good labor market incentives are important since there will be different changes in wealth and ability to work for different workers.

Increased longevity puts pressure on retirement systems that try to maintain monthly benefits, whether they are funded or unfunded. Indeed, in present discounted value terms, the increased cost of financing monthly benefits for longer retirements is the same in both types of systems. What may be different is the timing of responses, as funded systems that try to stay fully funded will be pressed to respond more in the early years after a realization of improved mortality. Earlier responses allow later responses to be smaller when solving a given financial problem.

I want to review how the retirement income systems in three different countries are legislated to adapt to increasing longevity (assuming no further changes in legislation and no trend in interest rates) – Chile, the US and Sweden.

Chile has a mandatory defined contribution, or DC, pension system. That is, Chilean workers are mandated to make deposits in accounts held by mutual funds, with their retirement benefits financed by the accumulations in their own accounts, although there is also a guaranteed minimum benefit. Neither the mandatory savings rate nor the age at which benefits can be claimed is legislated to change. That is, none of the response to improved mortality rates is legislated to happen in mandatory contributions. Because it is a DC system, longer life expectancy for a worker who is purchasing an annuity would result in lower monthly benefits for any given age at purchase. A worker choosing to work longer can claim benefits nonetheless. Or the worker can wait to get higher benefits from both additional savings and later annuitization, since annuities are cheaper for older workers. So the response of the Chilean system depends not only on the work decision but also on the separate decision of whether to start benefits. Thus, individuals can work longer and start benefits later to offset some of the decline in monthly benefits from longer life expectancy. For a worker who does this, all of the
additional contributions and all of the savings from delayed benefits accrue to the worker. In a DC system such as in Chile, there will be little or no pressure on the legislature to increase the mandatory savings rate. Thus, this part of the response to improved mortality rates is likely to be absent. For many workers there will not be a voluntary savings response either. Without a savings response, the Chilean system is likely to see a drop in replacement rates as we look across cohorts towards ones who live longer. This does not seem to me to be optimal.

The US social security system is a partially-funded defined benefit (or DB) system, which currently has roughly $1 trillion in assets. The partial funding of an otherwise pay-as-you-go system serves two purposes. One is to smooth the financing of benefits, increasing net revenues while the funding is built up in order to have less financial need once there is a fund that is earning a return. Secondly, by having a fund, there is a decrease in the tendency to lower benefits in light of short-run government revenue shortfalls. Given the difficulty of the elderly to adapt to real income cuts, pension systems should be adjusted slowly and with considerable lead time and separation from short-term budget needs is a plus. In 1983 legislation, the US Congress lowered future benefits in anticipation of longer life expectancies. It did this by increasing what is called the Normal Retirement Age (referred to as the NRA) from age 65 to eventually be 67. Congress did not change the age at which benefits could first be claimed, which remains 62. But, by raising the NRA, Congress lowered the benefits that would be paid for any given age at retirement. That is, those retiring earlier than the NRA receive larger reductions for earlier retirement, and those retiring later than the NRA receive smaller increments for late retirement. This does not have the same structure as would follow from actuarial calculations, since a larger percentage cut in benefits happens for workers retiring earlier. As a DB system, increases in life expectancy, beyond what is accommodated by the legislated increases in the NRA, result in financial pressures on the system. Thus, unlike the DC system, this DB system has pressure to increase taxes in order to preserve funding for longer lives. Like the other system, additional work does raise benefits significantly, at least for some ages.

11 A similar analysis holds for workers who do not annuitize, but use monthly withdrawals, thereby
Sweden has recently legislated a Notional Defined Contribution systems, which I will refer to by the acronym NDC. This is part of the Swedish system and also a very distant replacement for the Italian system. Since the concept is new and very interesting, I want to briefly describe the workings of the NDC system in Sweden. It is partially funded (a buffer stock of assets) and uses automatic adjustments to preserve financial stability, rather than relying on future legislative actions. An NDC system is similar to a DC system in that benefits are based on the entire history of contributions and it is contributions, not earnings subject to tax that affect benefits. The contributions are accumulated until retirement. The NDC system differs from a DC system in that these contributions are accumulated using a legislated notional return rather than actual returns on actual investments. The growth rate of average or total wages or GDP is the kind of notional return that makes sense for an NDC system, since it relates benefit claims to the central source of revenue – wages. Sweden uses average wages (although it also has a fallback with slower accumulation in case of financial need). As with a DC system, an NDC system needs to convert its notional accumulation into a monthly benefit. Sweden follows the DC approach of basing benefits on life expectancy and interest rates. However it deviates from what would be private insurance company practice. In addition to adjusting benefits for cohort mortality at the first age of benefit eligibility, an NDC system, like a DC system, adjusts benefits for the age at which a worker begins receiving benefits. Thus, as in Chile, Sweden is likely to have little pressure on the political system to raise taxes in response to longer lives, and individual workers can increase benefits significantly by working longer and delaying the start of benefits.

\[12\] In Italy, the adjustments for life expectancy need to be approved by parliament, rather than being automatic.

\[13\] Sweden adjusts benefits for life expectancy at the time a cohort reaches the age when benefits can be claimed (without disability), using the actual (period) mortality table, not a projected (cohort) table. A standard DC system would use projected mortality when determining benefits. By using a period mortality table, the NDC system sets benefits higher than if a projected cohort table were used in the same formula. Nevertheless the relative decline in benefits with life expectancy is roughly parallel to that if a cohort table had been used. Indeed, in a simple example, benefits, though higher, fall more rapidly in percentage terms than if a cohort table were used.
The idea of lowering benefits for longer-lived cohorts as a way of limiting cost does seem appropriate. We have seen that automatic adjustments for increased life expectancy and legislated benefit cuts in anticipation of longer lives are two ways to adapt to the trend of improving mortality. While this issue could be left for later parliaments, there are asymmetries in these processes – it is easier to raise benefits than to lower them, and easier to lower benefits in the distant future than in the near future. This suggests that it might be useful to legislate now some future decreases in benefits to partially but not fully offset the expected cost of longer lives, whether done directly or as part of a more basic reform. It is also possible to legislate a future increase in tax rates as part of the response. In the US future tax increases were on the books from 1935 to 1990, but not currently.

**B. Funding**

I turn now to the issue of funding of social security. Populations are aging. Without large changes in retirement ages, there must be a trend of rising taxes and/or declining annual benefits. Adding some funding (or more funding) to a basically pay-as-you-go (PAYG) system can reduce this trend, although funding at levels that are politically plausible do not change the direction of the trend over the next few decades, just the magnitude. A funded system can finance part of benefits out of the excess of the rate of return over the rate of growth times the level of funds. This helps to lower future taxes for any level of benefits or to allow larger benefits for any level of taxes.

But funds must come from somewhere. For an advanced country, funds come from currently raising taxes, cutting benefits or finding some other source of revenue (which would have an alternative use). Raising taxes now in order to have lower taxes in the future reduces the trend toward higher taxes. So the purpose of funding is to increase the burden on current generations in order to lower the burden on future generations. This would be similar to a budgetary decision to raise taxes or cut spending in order to decrease the public debt. (Provided, of course, that the cut in spending is not a cut in public investment.) Economists see advantages
in smoothing tax rates and foresee large tax increases or benefit cuts (or both) in the future and so they generally favor more funding, although there is considerable disagreement on how much.

Economists recognize that there are two aspects to increased funding, both of which matter. One is the growth of national capital and the other is strengthening the fiscal position of social security. The growth of national capital increases resources available in the future; strengthening the fiscal position of social security affects the political process, which will determine how costs and benefits will be allocated across different cohorts in the future. So funding that is associated with increased national savings is what economists tend to favor, since that accomplishes both goals - unlike funding that is merely relabeling or shuffling of liabilities, for example by issuing more national debt in order to fund accounts. To this end, increased funding within social security should not be offset by larger government deficits outside social security.

Economists also recognize that this intergenerational redistribution is the real gain from increased funding; that a widely-made argument for funded individual accounts is not right. Let me present and correct that argument. Some analysts and politicians compare the long-run return on assets with the long-run return in a PAYG system, which, as is well-known, is the rate of growth. Since long-run rates of return exceed rates of growth this is sometimes presented as a pure gain. But it is wrong to analyze policy by considering only the long run, not including the short-run costs and benefits associated with going to a different long run. It would be wrong to say that having the rate of interest exceed the rate of growth implies that a funded system is better. Once the analysis is done fully, then it is seen that there is no gain for everyone that is available from funding per se, just an intergenerational redistribution.

This correct argument can be seen by considering the infinite horizon present-discounted-value (PDV) budget constraint for social security. Defining benefits in terms of individual accounts does not change this constraint per se, although it may be bundled with benefit cuts. The constraint is changed by raising revenues or lowering benefits. So, taking some social
security revenues and moving them into individual accounts leaves behind a revenue gap that must be filled. Similarly, issuing more national debt in order to fund individual accounts also leaves a revenue gap that is needed to finance this debt. Combining the need to fill this revenue gap with the other effects of creating the accounts leaves the PDV constraint roughly unchanged. This simple arithmetic is altered to the extent that the rate of return on assets can be changed. But whether individual accounts raise or lower the rate of return on assets is a complex economic and political question.\textsuperscript{14}

If more funding is wanted, there are multiple alternatives. Funding can be in government bonds or in a diversified portfolio. Funding can be in a central fund controlled by a government agency or in individual accounts controlled by individual workers. The central gain from funding, that of accumulating resources to benefit future cohorts, is similar whichever route one takes.\textsuperscript{15}

C. Portfolio diversification

If there are assets, there is a choice as to what assets to hold. In Spain, any assets that social security might have are to be fully invested in government bonds. If a fund is built up, by and large, portfolio diversification is a plus. And portfolio diversification can be done with or without individual DC accounts.

The presence of many available routes to building a diversified fund is not just a theoretical option, since we have seen different countries consider and go down different routes.

\textsuperscript{14} Some people have argued that a gain for everyone is available by bundling analysis of static efficiency improvements in social security with funding. But those short run gains are available without funding, so this is misattribution, unless it really is the case that the short run gains are not available politically without such bundling. In politics bundling does matter, but such bundling at this stage of debate, at least in the US, is not necessary for reform.

\textsuperscript{15} While portfolio diversification can be done with or without individual DC accounts, putting the choice of portfolio in the hands of workers makes little sense unless workers bear a large fraction of the consequences of such choices.
Funding a central diversified portfolio within a DB system has been done in Sweden for years, has started in Canada and Switzerland and was proposed by the finance minister in Ireland and the Clinton administration in the US. Funding with a central fund within a DC system has been done by the provident funds of Malaysia and Singapore. And funding can be done with worker choice over portfolios made available by private providers within a DC system, as was pioneered in Chile and followed in other Latin American countries, implemented as a voluntary system in the UK, and as a mandatory small portion of the system in Sweden.

One word on the choice of assets. Long before reform, the Chilean social security reserves were invested in government bonds that were not indexed for inflation. The social security reserves were wiped out by inflation. While the world is more sensitive now than in the past to the costs of inflation, this does not mean that inflation is banished forever. Designing systems to react well in adverse circumstances is central for insurance and government bonds indexed for inflation can help.

It seems to me to be sensible to allow some investment in stocks by a centralized trust fund that is earmarked for retirement benefits. In this way the fund would expect a higher return while taking on some more risk. This could be done with a restriction to index funds and with fiduciary responsibilities for the trustees overseeing such investment in order to keep investment decisions as far from political interference as possible. A portfolio completely in government bonds does not optimize the risk-return tradeoff. That is, given the use of the fund to permit benefits to be higher in the future than without such a fund, a fund with a higher expected return but more variation in return can result in varying levels of benefits that makes the risk-return tradeoff favorable. The risk can be spread widely over cohorts.

In considering the gain from portfolio diversification, there are two elements. One is the ability of a DB system to spread risks across cohorts in a way that the market can not match since future cohorts are not alive to trade in today’s markets (Bohn, , Gale, ). Second is the fact that many workers have no personal investment in stocks and little private pension exposure to
stocks and would therefore not be taking on too much risk correlated with existing pension risks (Diamond and Geanakoplos, 1999).

**D. Central trust fund vs. individual account**

While a central trust fund can invest in a diverse portfolio, some people have argued for the advantages of using individual accounts as the method of investing. This can be done as a small supplementary account (as in Sweden and has been proposed by President Bush) or as a full replacement for the existing DB system (as in Chile). Argentina represents another route with both the DB and the DC systems being large.

There are both economic and political arguments for and against individual accounts as opposed to a central trust fund planned to hold a similar level of funds and a similar mix of assets.

On the economic side, individual accounts allow different individuals to hold different portfolios, allowing different choices on the risk-return frontier for those with different risk aversion. Countering this argument is questioning how well people will choose their portfolios, given the difficulty of good choice. Also countering it is the considerably higher

---

16 We know some, but not as much as I would like, about individual choice. There have been studies of the choices in employer-provided DC plans in the US, called 401(k) plans. The simple summary is that broad averages tend to move in the directions that finance theory suggests is sensible, but considerable numbers invest nothing in stocks and some invest completely in stocks, and it is likely that a significant fraction are not making really good choices. For example, many people invest heavily in the stock of their own employer - suggesting limited understanding of the advantages of diversification - both across stocks and across the combination of earnings and asset returns. Others invest fully in money market funds – a poor choice for a long-term investment. In addition, many people who choose individual stocks appear to trade too much - lowering the point on the risk-return frontier after trades, on average, as well as incurring trading costs. There is suggestive evidence that individuals do less well than the mutual funds in which they invest by trying to time the market, moving between classes of assets in a way that increases risk relative to expected return and indeed seems to lower expected return on average as well. As in consumer markets generally, people choose products at higher costs than seemingly identical products also available in the market. So there is lots of reason to be skeptical about the gains from individual choice per se in mandatory accounts. Will learning-by-doing take care of that? The evidence suggests not. US experience
administrative costs associated with individual accounts (and with the annuitization of the funds) and the fact that market risk is concentrated on asset holders, not spread through adjustments over cohorts as can be done (well or badly) in a DB system.

The major arguments, however, are political not economic. Proponents of individual accounts argue that there will be greater political drive to build and maintain a given size fund if the assets are individually owned rather than collectively owned. They also argue that the choice of assets and the use of the voting power of asset shares will be less politicized (and so better done in terms of risk and return) with individual accounts. Countering these arguments are the experience of some governments in successfully insulating portfolio decisions from political interference (both some countries and many state and local governments in the US (Munnell and Sundén)), as well as the issue of whether the difference in returns even if present would be as large as the loss from higher administrative costs. On the political side opponents argue that individual accounts will undercut the solidarity built up by a collective system and so tend to undercut protections for low earners and survivors that are built into current systems.17

The question of whether to have individual accounts is an ongoing issue in the US. While President Clinton proposed using part of the projected future surpluses to transfer resources to social security to be invested by the central trust fund, President Bush has taken a very different tack. The large tax cut (income and estate taxes) had used up the projected surplus even before the terrorist attack. Thus President Bush does not have a surplus to transfer to social security. Any transfer of general revenues would require a tax increase, an expenditure cut, or

---

17 For more on these arguments, as developed by a panel including both proponents and opponents of individual accounts, see Diamond, 1999, Question 3. The widely read World Bank book, Averting the Old Age Crisis, highlighted the poor performance of some governments in investing centralized funds. However, many of the countries cited had governments that didn’t function well on many dimensions and capital markets that offered poor choices to everyone. So one needs caution in deciding for which countries this message is important. Moreover, there has been progress in the ability of governments to invest well and an easier task because of index funds. However, the feasibility of heavy reliance on index funds depends on social security funds not being too large relative to the capital market.
running of deficits. Moreover, President Bush stated that the payroll tax rate should not increase. President Bush has called for individual accounts and has appointed a panel to propose a particular design for such accounts (but not to evaluate whether there should be such accounts – that debate will happen in Congress).

**E. Organizing individual account investments**

If there are individual accounts there is a major issue of how they are organized (Diamond, 1999). A major distinction is between government-organized DC accounts and individually-organized DC accounts. Government-organized accounts are those where the government makes available a small number of alternative investment vehicles. The government negotiates contracts with private firms to manage these investment vehicles. The pricing that occurs is the pricing that the market provides to large investors – referred to as institutional pricing in the US. These prices are considerably lower than the prices offered to individuals. Interestingly, Bolivia used competitive bidding for two halves of its market in setting up its new DC system. The costs are very low, but as with other bidding settings, Bolivia now needs a mechanism to ensure that adequate services are provided in return for the charges. And the US federal government uses this approach for its 2 million civil servants, with very low costs (and limited services).

In contrast, the Chilean approach has individuals free to deal with any approved investment vehicle provided in the market, and the approval mechanism is meant to be quite easy and open – easy entry to hold down market power. Thus, the mandatory markets are meant to mimic voluntary capital markets, although highly regulated. But voluntary capital markets that deal with individuals are expensive everywhere in the world. In the US, which has the best-developed capital markets in the world, the typical individual investing in equity mutual funds pays roughly 1.25 percent of the value of assets in annual charges (Rea, Reid and Lee, 1999). Over the course of a 40-year career such charges will cut the value of the accumulation by roughly 25 percent. That is, these annual charges on benefits compound over time since deposits
are subject to annual charges each year until retirement. But the annual charge falls on deposits roughly 20 times on average if deposits are made annually over a 40-year career. The relationship between annual charges on balances and total loss at retirement is close to proportional with a factor of roughly 20. These costs are large, suggesting that wide choice may not be worthwhile. Nor does this cost estimate reflect the fact that many US investors pay separately for investment advice, sometimes as much as 1 per cent of assets per year in what are called wrap accounts.

Chile has done well, having costs that reduce benefits by 15-20% (Diamond and Valdés-Prieto, 1994). In other Latin American countries, accounts are smaller and the cost is larger in percentage terms. With some costs being fixed costs per account, it is not surprising that costs are higher in percentage terms for smaller accounts. In contrast, in the UK, with less structure to the regulation of funds, costs of accumulation reduce benefits by roughly 35% (Murthi, Orszag, and Orszag, 1999). And these costs are just the cost of accumulation – annuitization adds to these costs by roughly 10% (Murthi, Orszag, and Orszag, 1999). ¹⁸

Sweden has taken another route. Swedish workers have wide access to the market (500 funds) but restrict individual contacts to being with the government, which then invests in bulk in individual funds. ¹⁹ Moreover, the Swedes regulate the prices that can be charged. It is hard to know, but I do not see anything in the Swedish system that will make demand more price sensitive and so competition more concentrated on low prices rather than heavy advertising. That the investment firms may not know who their customers are will change advertising, but may not lower its overall cost. Worldwide, the record of price controls is not terribly good. And even if successful, the costs will be considerably higher than could be achieved with government-organized individual choice. For example, if the government reduces charges to half those in the voluntary market in Sweden, the reduction in benefits from charges is still over 15

¹⁸ On the cost of annuities in the US, see Mitchell, Poterba, Warshawsky, and Brown, 1999.
¹⁹ That is, in Sweden the government keeps all records for the 2.5 percent of payroll devoted to individual accounts. Financial firms negotiate a price structure with the government to become eligible to accept deposits, and those deposits are made in bulk by the government, not separately by individual workers.
per cent. Thus, I see no reason to think that this form of organization itself will result in a particularly cheap system in the long run.

A market with a wide choice of mutual funds does not behave like an idealized competitive market. Individuals are somewhat responsive to differences in price and quality, but that responsiveness is limited in both size and speed. Thus, the pressure on pricing from consumer responses is present but limited. We would expect to find equilibrium with prices above marginal costs, with advertising to attract profitable customers, and with a wide range of prices for similar or identical products. Indeed these properties hold for all consumer markets, but seem particularly important here. A mandated market aiming at everyone will include a large number of inexperienced investors. It will include a large number of low-earners, with small accounts and so little incentive to monitor closely. Little monitoring will be done by some because of their limited attention to retirement issues, and by others because of the presence of income guarantees once they do retire. This is a setting where procrastination in reconsidering portfolio providers may be particularly rampant - there is little apparent gain from changing providers this month rather than next month, even if one had the ability to tell good providers from bad ones. And the great stochastic variability in returns makes that hard. This is further complicated if the pricing structure for administrative charges is complex. These high costs reflect limited price sensitivity by small savers and so an equilibrium with significant markups and heavy advertising costs.

Since retail markets generally are far from ideal markets, it is useful to review why the case for limited choice and a large role for the government does NOT extend to consumer goods generally. First we are considering a mandated market. This reflects the need for a mandate because workers do not do an adequate job of saving for retirement and creates an obligation for the government to see that the mandate does not lead to inferior outcomes. Second, saving for retirement during a working career is something people do just once. There isn’t the opportunity to learn to do it better the next time – once the full implications of investment choices are experienced directly, a worker has retired. Third, the principles of investment in the presence of
risk are not simple to master. Fourth, as long as the mandate is small enough that many people are saving outside the system, portfolios held outside the system can be adjusted to offset some of the limitations on choices for portfolios held inside the system. And fifth, the presence of the government is not likely to stifle innovation in financial markets since the mandated savings remain only a part of overall savings and investment.20

There are various ways to move portfolio choice away from the government and to individuals. Moreover, given the noise in returns, it is difficult for anyone to tell good portfolio managers from bad ones. Choices will be regulated and one needs to recognize that like the politics of direct government portfolio choice, the politics of portfolio regulation has its own potential shortcomings.

III. Protecting living standards after retirement

The purpose of the retirement portion of social security is to protect the living standards of workers and their families after retirement. This involves two issues – the level of benefits a worker receives at the start of retirement, and what happens afterwards. Adjusting benefits for inflation, or some combination of inflation, wage growth, and rates of return, makes sense and is not in dispute. The exact indexing rule, and so sharing of risks, can be analyzed. What is done

20 The government could select the portfolios for the individual accounts. Everyone could be given the same portfolio, as was done in Singapore and Malaysia. While this would minimize administrative costs, consideration of risk bearing suggests that different people should hold different portfolios. The importance of this issue varies with the size of the portfolio and the extent of savings outside the system, since many people could offset shortcomings in the balance between stocks and bonds in the social security portfolio by complementary changes on the balance of their portfolios. In theory, and I suspect in practice, one could do better by having the government vary portfolios based on age – for example, with a balance between stock and bond index funds shifting toward less risk as people age.
when a worker or the spouse of a worker dies should affect both initial and later benefits. With so much attention focused on intergenerational issues, on not leaving too large a burden on future generations, it is important not to lose sight of the importance of social security for distribution within each generation. This issue is also important and is also best addressed sooner rather than later.

A. Income Distribution

There are two perspectives on social security and income distribution. One perspective is the level of retirement benefits relative to needs. Both the protection from absolute poverty and limiting the extent of drop in living standards during retirement are important. The second perspective is that of redistribution – incorporating both benefits and taxes, and best approached on a lifetime basis. Both perspectives are relevant for ethical concerns. Both perspectives are relevant for the workings of the system in terms of outcomes and incentives. And both perspectives are relevant for the politics of social security reform. Income distribution concerns include both limiting redistributions to the better off because of poor system design and ensuring adequate redistributions to the less well off.

Countries approach these multiple issues through a variety of institutions. To deal with the most serious poverty problems, countries typically have a minimum income program for the elderly, one that is conditioned on both income and sometimes assets as well. Such a program discourages work, tax compliance and savings. While every redistribution has distortions, these programs tend to have very high implicit marginal taxes, making for large distortions. Such programs may provide a very low level of income guarantee. Such programs are obviously needed, but their disadvantages keep them small, leaving other needs unfilled. Guaranteed minima involve larger distortions on a smaller number of workers; broader redistribution involves smaller distortions on more workers. Thus the complexity in analysis.
Within social security, there are three parts to the distribution mechanism, which are generally used in different combinations. Some countries provide a guaranteed minimum social security benefit for workers who have had long enough careers. For example, this is done in Chile. Other countries have a process for determining benefits that contains a progressive element. This can come through a progressive benefit formula in a DB system, as in the US, or through a separate program providing, for example, a flat benefit (to eligible retirees) while having no explicit redistribution through the rest of the system. This is the approach taken in Argentina and the Netherlands. These methods have the advantage of not directly distorting savings decisions since benefits are not lost because of savings outside the system. And while some countries rely exclusively on a payroll tax that falls just on the earnings that count for determining benefits, other countries tap other sources of tax revenue that may be more progressive, such as some of earnings that do not count for benefits or general revenues that may be financed by a progressive income tax.

Male workers with higher earnings tend to live longer than low earners. Therefore, in the absence of some corrective mechanism, social security will redistribute from poor men to rich ones. So explicit attention to distribution is called for from a lifetime perspective. That is, progressivity is needed somewhere in the overall system since high earners benefit at the expense of low earners in a linear system. And replacement rate needs are thought to fall with income levels, making a further case for progressivity in benefit determination in some form. This issue would be worth reviewing as part of reform here.

B. Widows and widowers

Workers die, sometimes before retirement age and sometimes after. Thus it is important to design protection for the spouses of workers. It is obviously essential to protect nonworking spouses. It is also relevant even if both spouses have had a long career. For simplicity, I will consider only a death after retirement, assuming that the surviving spouse is also above the retirement benefit eligibility age and ignoring the possible presence of small children. Two
questions arise. One is how large the benefits should be for a survivor compared with the benefits received when both husband and wife were alive. And second is how to finance survivor benefits, an issue connected with redistribution across workers since one needs to adjust for family size when deciding which workers are needier than others.

There are economies of scale in living as a couple rather than two separate people. This is particularly the case for elderly who live alone, rather than with their children. In the US, the evidence is clear that by and large, the elderly who can afford it prefer to live separately from their children. And the children who can afford it prefer that too. The basic approach of US Social Security is to give a survivor somewhere between ½ and 2/3 of what the couple had (with deviations based on the ages at which benefits were claimed). However, there is no good logic on how the exact fraction is determined in the US. Studies of living costs generally find that an individual needs roughly 70 percent of what the couple had to preserve the standard of living. So it is important to build in protection of spouses that will achieve something resembling this outcome. Moreover, in the US, analysis of longitudinal data shows that on average women have a significant drop in their standard of living when they become widows. It appears there is a similar problem here, indeed it appears that widows, on average, have low incomes (Namkee Ahn y Florentino Felgueroso, 2001). This raises two issues. One is the level of survivor benefits on average. Second is the distribution of survivor benefits across couples with different earnings histories. It seems to me that the level is too low and the distribution does not have a good logic, although the guaranteed minimum pension helps greatly in reducing poverty.

Economies of scale in living suggest that a survivor should have more than one-half of the couple’s pension income. In the case of a worker who dies after retiring, the surviving spouse receives 45% of the "base reguladora" of the deceased worker as a widow or widower benefit, with the government planning to increase this percentage to 52%. Since the worker might have been receiving anywhere between 50% and 100% of the base reguladora (ignoring adjustments for early retirement), it is hard to see the logic in a survivor replacement rate that varies so much.

Again, I am ignoring the role of minimum pensions in the pattern of replacement rates that actually occur. A more thorough analysis would certainly include their role. While there is logic in giving a larger fraction to a survivor in a couple with a lower benefit relative to past earnings, the range of variation seems too wide. Looking ahead, the range varies from 52% (52/100), which seems too small to 104% (52/50) which seems too large. And reduced benefits for early retirement make the upper bound of replacements even larger since there is not a comparable reduction for survivors.

Further complication comes from the fact that the survivor benefit is a supplement to whatever retirement benefit the survivor might have on his or her own account (subject to an overall maximum). Moreover all of this is provided with no decrease in a worker benefit to help finance the survivor benefit. While dual career marriages are less common in Spain than some countries, they do occur and may well occur more frequently in the future. There is a sharp lack of symmetry in the treatment of the two in the couple, depending on their history of earnings. As one example, assume that husband and wife have the same base reguladora, but he worked 35 years and she worked 15. Assuming the overall maximum does not come into effect, together they receive 150% of their common base reguladora. If she dies, he gets 152% of their common base reguladora, while if he dies, she gets 102%. While not all couples would want exact symmetry in benefits, this seems too far from symmetry. Similarly, with a nonworking spouse of a worker with 35 years of contributions, there is no loss in benefit if the nonworker dies, but a loss of 48% of benefit if the worker dies.

This issue is of further concern since there is diversity in incomes of spouses. Insofar as a lower earning spouse is the survivor, there is a further fall in the standard of living. This may be further exacerbated if husband and wife chose to retire at the same time even though the wife may have been younger, and so had a larger reduction for early retirement in determining her pension. An important part of financing such benefits can be by adjusting the benefits of a worker (or both husband and wife if both worked) to reflect the financial advantage of survivor benefits. In the case of a dual earner in a couple, the survivor benefit is often a supplement to retirement benefits that the survivor may have already received on her or his own account.

There are more benefits if there are other family members and little outside income.
benefits. At one extreme, the expected cost of worker and survivor benefits can be the same as what the cost would have been had the members of the couple not been married. However, this assumes that the redistribution between single and married should not pay attention to marital status. And that does not appeal as an absolute principle. At the other extreme is the US approach, which provides spouse and survivor benefits with no cost to a worker, an approach that is much criticized. Some compromises can be readily designed - good survivor benefits financed by lower worker benefits are widely supported by many economists.

In sum, it is important to study the economic position of widows. It is not good policy to have widows experience a large drop in living standard on average. Raising widow benefits and having a systematic relationship between the total benefits of a survivor and the total benefits that had been received as a couple makes for good policy. And equity in terms of lifetime net returns of the couple can be achieved by adjusting initial benefits for a couple for some fraction of anticipated benefits of survivors.

IV. Concluding remarks

Because of projections of rising costs, it is important to focus on benefit levels. However, there are other improvements in the system that can be legislated without necessarily changing average benefit levels. Changes in pension benefits should happen with considerable lead time, particularly when benefits may be lower than previously expected. If one is protecting expectations to some degree, then it is important to act early before even more of the expectations are strongly held. Moreover, it is best to consider all aspects of a system, including such details as the treatment of widows, early in the reform process, instead of fixing just some of the details and leaving the rest until the debate is more pressing and, most likely, more contentious. With the payroll tax rate already so high, improving the pattern of benefits is likely to entail increasing some pensions at the expense of lowering some others. This sort of improvement is best legislated well ahead of implementation.
Similarly, the sooner the remaining financial difficulties of the system are faced the better. Waiting until a crisis is on hand may be comfortable for today’s politicians but makes tomorrow’s problems more severe.

References

Namkee Ahn y Florentino Felgueroso, 2001, La situación económica y familiar de la viudas en España (Economic and family status of widows in Spain), unpublished, FEDEA y Universidad de Oviedo.


Boldrin, Michele, Sergi Jimenez-Martin y Franco Peracchi, 2000b, Sistema de Pensiones y Mercado de Trabajo en España, Fundación BBVA, Madrid.


Gale, Douglas,


