

Olivia S. Mitchell
The Wharton School

April 15, 2008

New Directions for Pension System Performance Measurement ¹

Good morning, ladies and gentlemen. I am honored to be speaking to the delegates and participants of this very important International Seminar on pensions, and I offer special thanks to the International Federation of Pension Fund Administrators, FIAP, for the kind invitation to open this seminar.

It is a pleasure to be once again in Peru, a country in which I lived and studied in the early 1970's. Many things have changed in the more than 30 years since I left, but perhaps one of the key accomplishments, from my perspective, has been the funded pension system implemented in 2004. I look forward to finding out more from the Peruvian experts, as well as those from the other countries represented here, about how things are going.

My goal today is to provide comments about how I think we should measure retirement system performance going forward.² My remarks should not be interpreted as downplaying the importance of recent financial market developments which, no doubt, pose serious challenges for those charged with managing retirement systems for the future. Rather, today I would like to focus on recent research regarding how one might judge a retirement system's performance, particularly in terms of meeting *consumption-smoothing* and *adequacy* goals. My recommendation is that pension systems must begin to more clearly define and measure payout benchmarks if we are to understand when and where a retirement system will be successful or may fall short.

¹ This presentation was prepared for the May 2008 International Seminar on "Pensions for the Future: Developing Individually Funded Programs," hosted by the International Federation of Pension Fund Administrators (FIAP) and the Peruvian Association of Private Pension Fund Administrators. The research reported here was supported in part by the Pension Research Council and the Boettner Center for Pensions and Retirement Security, at the Wharton School of the University of Pennsylvania. © 2008 Mitchell. All Rights Reserved.

² My remarks draw on my joint research with several authors listed in the Bibliography.

Retirement System Objectives

Economists posit that the goal of a retirement system is to help workers build a more secure retirement by inducing and incentivizing retirement saving, generating strong investment results, pooling longevity risk; and protecting against old-age poverty.

In the past, policymakers and researchers working on funded pension systems have devoted substantial – and in my view, very valuable – attention to constructing an environment favorable to pension funding and investment. The focus, quite naturally, has been on investment safety, portfolio diversification, and investment return and volatility. With this has come much debate and discussion what types of pension fund schemes are most cost-effective in getting contributions deposited, tracked, invested, and paid out – with due attention to fees and commissions, including annuitization costs. That is, there has been much analysis and work on how to structure the pension fund investment marketplace including regulations over investments, fees, and commissions. And the results have been impressive: global pension assets in funded systems now amount to more than \$23 trillion globally.³

As a result of the past several decades of intensive work, much of it by people in this room, funded pension systems now offer better investment choices, more low-cost investment options, and diversification options that did not exist two decades years ago in the funded pension arena. Additionally, the research on behavioral economics has found its way into pension system design in both developed and developing countries, helping to make contributions more automatic, enhancing the role of target date maturity funds, and emphasizing annuities as a sensible way to manage longevity risk.⁴ Along many dimensions, then, retirement systems are becoming more resilient to political and economic shocks that were more prevalent three decades ago (Mitchell 2000, 2004).

Developing New Standards for Retirement System Performance

Many countries have reported real success stories in their funded pensions over the last several years. Nevertheless, critics of these systems are beginning to make waves, and in some cases

³ Watson Wyatt Worldwide (2007).

⁴ Mitchell and Utkus (2004).

there is concern that they might even “tip the pension system boat” and bring about reforms which could even sink the important progress made by funded pensions over the last generation.

Such critiques arise from several sources. Some analysts, including myself, express concerns about financial illiteracy and lack of pension system knowledge. For instance, my recent work using the US Health and Retirement Study, shows that most mature persons – in their mid 50’s – cannot respond correctly to simple questions about compound interest calculations, do not understand the erosive power of inflation, and are underinformed about investment risk. Similarly, in the Chilean case using the Encuesta de Proteccion Social, we find that workers report pension contributions and accumulations inaccurately, do not know how much they pay in commissions, are unaware of the rules for minimum pensions, and cannot say how their funds are invested. Such a profound degree of financial illiteracy is troubling, and of further concern is that lack of knowledge appears to be concentrated among low-educated and lower-paid workers, women, and minority groups (Lusardi and Mitchell 2007; Arenas de Mesa et al. 2008). Such findings require that plan sponsors and regulators must take on a much broader role in educating retirement system participants, if the pension schemes are to be successful at attracting and incentivizing pension savings.

A different source of criticism has noted that countries often impose restrictions on pension investments, for instance, limiting them to domestic holdings or capping holdings in particular assets. This practice, in turn, exposes pension participants to political and economic risk ranging from inflation to expropriation. A related point is that high fees and commissions can and all too frequently do erode the small savings which low-wage workers manage to accumulate when they pay into their systems. These and related topics have been the subject of previous FIAP conferences and publications, and I commend to anyone interested the many excellent publications on these topics.⁵

Retirement System Performance Measures: Consumption-smoothing and Adequacy

A different issue, one that has received far less attention, but is the crucial challenge for the future, has to do with how well a pension system is at paying benefit promises. In particular, I

⁵ See for instance the selection listed at <http://www.fiap.cl/>

believe we must start judging retirement systems in terms of their success at two key performance outcomes: (1) helping provide insurance, so that participants can smooth their consumption over their life cycles, and (2) providing welfare or income redistribution to the elderly poor. In order to assess whether a pension system is doing this, of course, it is necessary to specify these objectives in measurable performance terms. Further, it is critical to have good enough measures of actual or forecasted outcomes so that analysts can assess what a system will produce and whether it is satisfactory. While both efforts have been plagued with difficulties in practice, recent developments in data and modeling can inform future efforts to conduct pension system performance evaluation.

Consumption Smoothing. The insurance criterion I refer to requires the pension system to help households consumption smooth over their life cycles. That is, if the pension insures workers against shocks to consumption in old age, it would be expected to pay more to those experiencing higher shocks than lower ones. A key obstacle to conducting this evaluation is that consumption data are difficult to obtain in practice – that is, it has been quite difficult to measure household consumption patterns to see if they are doing about as well after retiring, as before. In practice, analysts tend to make assume that consumption and labor earnings are positively correlated, so then analysts can ask to what extent retirement income payments are higher for those who had higher pre-retirement earnings. Presumably a well functioning pension insurance system would pay higher-earning households more in retirement, as these might be anticipated to experience the greatest loss of income in old age. This rather naturally leads to the proposition that retirement benefit payments must be compared to workers' own past labor earnings, to evaluate how good a job the pension is doing in terms of *earnings replacement*.

Techniques to implement this approach have evolved over time. In the US for instance, the 1935 committee designing Social Security introduced the idea of a *target replacement rate* under President Franklin Delano Roosevelt. That group proposed that “payment of benefits at a rate . . . approximating 50 percent of previous average earnings is socially desirable,” though it did not explain why it selected that particular threshold nor did it indicate whether each worker's own lifetime earnings or some economy average would be the proper reference point. Three decades later, McGill (1966) opined that a worker's retirement income net of taxes should be

“approximately equivalent to the worker's spendable wages” which meant he was targeting 100% of the employee's own pre-retirement pay after tax. This was consistent with the presumption that a well-run pension system has as an “implicit or explicit goal ... the maintenance of preretirement standards of living” (Meier et al. 1980).

In the real world, those of us who compute and use replacement rates as pension performance indicators have rarely had data sufficiently rich to permit comparison of a given retiree's own benefits to his own lifetime pay, much less to derive this for a sample of actual workers followed over their lifetimes. Instead, researchers generally have been forced to rely on *aggregate* data and actuarial models to compute benefit replacement rates for hypothetical full-time steady workers employed over their entire careers. For instance the stylized earnings paths used in OECD (2007) and Whitehouse (2007) are derived from average economy-wide earnings paths, on the assumption that workers never change their relative position in the pay distribution and assuming no volatility in the returns to human capital. The work by Bernal et al. (2008) employs aggregate coverage data in an actuarial cell-based model to predict eventual replacement rates from the Peruvian system, and related simulations for Chile are conducted by Berstein et al. (2005) and Favre et al. (2006).

This replacement ratio approach is, however, suspect along several dimensions, making it an imperfect tool for judging pension system performance. Of particular relevance for developing economies is the fact that workers do not always pay into the pension system consistently and steadily for their entire worklives. Indeed recent analysis with administrative data linked to nationally representative household surveys such as the Encuesta de Proteccion Social (EPS) in Chile show that people tend to have quite uneven patterns of attachment to the pension system – partly because they are often out of the labor force, and partly because they are working as self-employed and they have not been required to contribute while in that state (Arenas de Mesa et al. 2008). The same is true in the US, as the Health and Retirement Study (HRS) shows (Mitchell and Phillips 2006; Mitchell et al. 2000). Accordingly, projected replacement rates that assume steady earnings and a full career of contribution will likely dramatically overstate expected retirement system benefits, and the bias is likely to be greater for subgroups such as women and the less educated. In the Chilean case, for instance, Arenas de Mesa et al. (2008) show that those

with the highest density of contributions over their worklives are the highly paid, the better educated, and men..

A related point is that even when people do work in the covered labor force, their earnings are subject to rather large shocks which produce remarkably different retirement wealth accumulations (Mitchell et al. 2007; Mitchell et al. 2000). And volatility of labor income shocks need not be constant over the lifetime; for instance Mazumder (2001) uses US administrative records tracking workers through time and reports that (a) the variance in the transitory component of earnings varies over the lifecycle, (b) the transitory earnings variance follows a U-shaped pattern over the lifecycle, and (c) certain groups, particularly blacks and the least educated, have the highest lifetime earnings variance. What this means is that projected replacement ratios which assume steady worker profiles are likely to overstate eventual retirement benefits and the effect is more biased for some than others, as indicated for the US case by Mitchell and Phillips (2007).

A far deeper criticism of the replacement rate concept is that it does not directly measure consumption. Because of this, this measure cannot be used to directly assess whether a pension system is successfully helping retirees smooth consumption at older ages. A recent and extremely influential critique along these lines is found in some very new research by Scholz et al. (2007), who use a life cycle model to estimate that 80 % of U.S. older households have saved optimally for retirement, after taking into account the impact of children on consumption patterns. Specifically, that research shows that an older couple could reduce its consumption by one-third after the children leave home, and by another 30 percent on the death of one spouse. In other words, some drop in consumption during retirement makes sense when we consider the measures used to judge retirement system adequacy. The problem is that most replacement rate measures do not do this, leading to an overstatement of the benefit amount required in retirement relative to pre-retirement pay. This concern has been exacerbated by recent analysis of retired households' food expenditures by Aguiar and Hurst (2005), who report that retirees tend to spend less money on food, but they make up for this by undertaking more "home production" – comparison shopping and home preparation – so as to maintain their total level of wellbeing. Thus Skinner (2007) concludes that "planning for consumption smoothing doesn't mean one has

to maintain consumption spending through retirement,” a point that tends to seriously undermine the conventional approach to retirement replacement ratio notions.

Adequacy. Since replacement rates can be said to be poor performance criteria for pension schemes, would something else be better? We turn next to the standard of *benefit adequacy*, which has some important advantages over the replacement ratio notion.⁶

Discussions over adequacy standards take on meaning mainly in the context of a discussion on redistribution. That is, a retirement scheme that meets adequacy goals would presumably be one that provides relatively higher benefits to retirees with low income and assets in old age. Yet here too, implementing an adequacy concept in practice can be controversial. For instance, an adequacy threshold might be set in terms of a poverty line defined in terms of income sufficient to purchase a minimum consumption level. Naturally the adequate pension would then be one where retirees received sufficient money to purchase a subsistence calorie diet for households of a given composition. This standard has the virtue of being relatively constant and measurable consistently through time. Of course even here, it can prove difficult to capture accurate caloric intake and there is debate over the quality of calories and possibly differential consumption needs of the elderly. Further, there is often a tendency to convert a so-called “fixed measuring rod” of caloric intake into a “rubber yardstick” so that poverty standards are set differently across countries and may change over time.

The difficulty of agreeing on an absolute adequacy standard sometimes implies that *relative adequacy* is used instead as a retirement system benchmark. In principle, the idea is that consumption in retirement could be compared to some external consumption stream. Since consumption data are difficult to get, analysts tend to focus in practice on income, as it is measured more easily than direct consumption. Yet here too there is controversy, since some seek to assess pension plan performance given worker’s *own* salary history, while others compare retirement payouts to *average earnings of those who remain employed*. For instance, Whitehouse (2007) generates *replacement rates* that compare a worker’s retirement benefit payments to his own lifetime earnings level. By contrast, OECD (2007) develops a concept

⁶ This section draws on McGill et al. (2004) and Mitchell and Phillips (2006).

termed the *relative pension level*, specified as the ratio of a worker's retirement benefit divided by *average economy-wide earnings* that same year. Obviously these two approaches are very different in spirit, and they can generate markedly different results in terms of performance. In a rapidly growing economy, for example, an individual might retire with a pension benefit that would be high relative to his own lifetime pay, but which would be low and lag behind the average pay of the average worker. Which relative standard to use is thus a matter of contention, and there is no common agreement on the right performance benchmark.

Which Pillar Should Do What?

A related question that arises when judging the performance of retirement systems is which pillar should be responsible for what goal, and what the interactions are between them. The approach often associated with the World Bank posits that a nation's old-age adequacy needs can be well satisfied by a social safety net type plan, typically an unfunded pay-as-you-go tax-financed scheme. The so-called "social pensions" which are growing increasingly popular are of this type, where passing some threshold age triggers potential eligibility. After that, benefit payments tend to be means-tested so that they are targeted on those with least income and assets. The second pillar of this multipillar structure would consist of the funded earnings-insurance-type plan, intended to provide for consumption smoothing. (Additional voluntary pillars may also be envisioned).

This pure structure is, however, rarely found in practice: rather, countries often tend to conflate the two objectives and the tasks of the retirement pillars. As one instance, the government of Chile has proposed adding to the funded scheme a large new publicly-financed minimum benefit guarantee under the new *Pilar Solidario*, which is to be a means-tested pay-as-you-go benefit costing an estimated 3% of GDP. Interestingly, this first-pillar safety-net benefit is described as raising the benefit replacement rate from 44% to 70% of workers' salaries, but little has been said about its ability to ensure older persons a minimum income/consumption standard. A different tack has been taken in Singapore, where the government has set a minimum subsistence income level for retirees and announced that participants in the national funded defined contributions system must amass sufficient assets in their funded defined contribution accounts to finance this minimum consumption flow in retirement. What is interesting in this second case

is that the funded system is being held to the minimum income flow standard, a task more commonly assigned to a safety-net pay-as-you-go system.

Concluding Observations

Those of us who focus on pensions are keenly aware that global aging will pose important new risks for the retirement systems we have worked so hard to establish, nurture, and bring to maturity. Funded systems have a strong track record in terms of building up assets in many countries. Yet there is increasing concern about people who fail to contribute will demand benefits, as well as others who do contribute but will never amass sufficient assets to guarantee retirement subsistence. Put simply, funded pension systems over the next few decades must do better in terms of payout risk.

It is our job to now devote more attention to this very important objective, which will take us beyond attention to fees, charges, and investment returns. This will require more clarifying the desired roles for each element in the multipillar system, about what might be affordable in terms of a minimum subsistence benefit, and about the possible negative as well as positive effects of means-testing in the context of a funded pension regime. This will also require more debate about the relative emphasis a country wants to and can afford to place on retirement scheme adequacy and consumption smoothing, and it will require pension system designers to focus more clearly on the tradeoffs between these goals.

Last but surely not least, analysis and judgments on pension payout performance will require far better data and models than we have had in the past. What is needed is information that traces individual workers' salary/earnings histories, to evaluate when people contribute and if they do not, why not. Further we must also understand how consumption patterns track income patterns both before and after retirement. This task is a large one, but it is starting to be feasible to carryout with the help of household surveys linked to administrative records. If we are to move beyond mainly financial market issues, to a broader discussion of pension payouts, these new sorts of data will be essential.

In closing, I thank you for your attention for my remarks. I particularly congratulate the organizers of this seminar on their success in achieving such a fine audience, on the excellent level of the presentations, and for giving all of us the opportunity to learn from the analyses, presentations, and discussion of so many who play such a key role in the retirement security of millions all over the world. There is much to be done and I welcome collaborations with you as we move forward.

Thank you very much.

Selected Bibliography

- Aguiar, Mark and Erik Hurst. 2005. "Consumption versus Expenditure," *Journal of Political Economy*, 113(5), October: 919-948.
- Arenas de Mesa, Alberto. 2000. "Proyecciones de Déficit Previsional Chileno: Gasto Público en Pensiones Asistenciales 1999-2010". CEPAL – Serie de seminarios y conferencias N°3.
- Arenas de Mesa, Alberto, David Bravo, Jere R. Behrman, Olivia S. Mitchell, and Petra E. Todd. 2008. "The Chilean Pension Reform Turns 25: Lessons from the Social Protection Survey." In *Lessons from Pension Reform in the Americas*, Eds. Stephen Kay and Tapen Sinha. Oxford, Oxford University Press: 23-58.
- Bernal, Noelia, Angel Munoz, Huga Perea, Johanna Tejada, and David Tuesta. 2008. *Una Mirada al Sistema Peruano de Pensiones*. Lima, BBVA.
- Berstein, S., Larraín, G. and Pino, F. 2005. "Cobertura, Densidad y Pensiones en Chile: Proyecciones a 20 años plazo". SAFP Documento de Trabajo N°12, Santiago.
- Favre, M., A. Melguizo, A. Muñoz and J. Vial. 2006. *A 25 años del la Reforma del Sistema Previsional Chileno: Evaluación y Propuestas de Ajuste*. Santiago: BBVA.
- Lusardi, Annamaria and Olivia S. Mitchell. 2007. "Baby Boomer Retirement Security: The Roles of Planning, Financial Literacy, and Housing Wealth." *Journal of Monetary Economics*. 54(1) January: 205-224.
- McGill, Dan 1966. *Fundamentals of Private Pensions*. New York: Richard D Irwin.
- McGill, Dan, Kyle N. Brown, John J. Haley and Sylvester J. Schieber. 2005. *Fundamentals of Private Pensions*. 8th ed. Oxford: Oxford University Press.
- Mazumder, Bhashkar. 2001. "The Mismeasurement of Permanent Earnings: New Evidence from Social Security Earnings Data." Federal Reserve Bank of Chicago Working Paper 2001-24.
- Meier, Elizabeth, Cynthia Dittmar, and Barbara Boyle Torrey. 1980. *Retirement Income Goals. Report to The President's Commission on Pension Policy*. Washington, D.C.: March.
- Mitchell, Olivia S. 2004. *Challenges for Retirement Financing*. American College, Bryn Mawr, PA.: 27th Engle Lecture.
- Mitchell, Olivia S. 2000. "Building an Environment for Pension Reform in Developing Countries". In *Foundations of Pension Finance*. Zvi Bodie & E Phillip Davis, eds. London: Edward Elgar: 480-503.
- Mitchell, Olivia S., James Moore, and John Phillips. 2000. "Explaining Retirement Saving Shortfalls". In *Forecasting Retirement Needs and Retirement Wealth*. Eds. O.S. Mitchell, B. Hammond, and A. Rappaport. Philadelphia, PA: University of Pennsylvania Press: 139-166.
- Mitchell, Olivia S. and John W. R. Phillips. 2006. "Social Security Replacement Rates for Alternative Earnings Benchmarks." *Benefits Quarterly*. 4th Q: 37-47.
- Mitchell, Olivia S., John W. R. Phillips, Andrew Au. 2007. "Lifetime Earnings Variability and Retirement Shortfalls." In *Retirement Provision in Scary Markets*. Ed. H. Bateman. Cheltenham, UK: Edward Elgar: 78-99.
- Mitchell, Olivia S. and Stephen P. Utkus, Eds. 2004. *Pension Design and Structure: New Lessons from Behavioral Finance*. Oxford: Oxford University Press.
- OECD. 2007. *Pensions at a Glance: Public Policies Across OECD Countries*. OECD, Paris.
- Scholz, John Karl, Ananth Seshadri and Surachai Khitatrakun. 2006. "Are Americans Saving 'Optimally' for Retirement?" *Journal of Political Economy*, 114(4), August: 607-643.
- Skinner, Jonathan. 2007. "Are You Sure You Are Saving Enough for Retirement? NBER Working Paper 12981.
- Watson Wyatt Worldwide. 2007. *Global Pension Assets Study 2007*. <http://www.watsonwyatt.com/research/deliverpdf.asp?catalog=GPAS>
- Whitehouse, Edward. 2007. *Pensions Panorama: Retirement Income Systems in 53 Countries*. Washington, D.C.: The World Bank.