Global pension systems and their reform: Worldwide drivers, trends and challenges

Robert Holzmann

University of Malaya, Kuala Lumpur, Malaysia

Abstract Across the world, pension systems and their reforms are in a constant state of flux driven by a shifting focus, moving reform needs, and a changing enabling environment that reflect objective events but also changes in views and perception. The ongoing worldwide financial crisis and the adjustment to an uncertain “new normal” will make future pension systems different from past ones. The objectives of this article are: i) to briefly review recent and ongoing key changes that are triggering reforms; ii) to outline the main reform trends across pension pillars over the last two decades; and iii) to present key policy areas on which the pension reform community will need to focus to make a difference.

Keywords pension scheme, social security reform, social security planning, demographic aspect, old age risk, economic conditions, international

Introduction

The outlook on global pension systems and their reforms since the early 1990s has changed markedly; the most recent reassessment is triggered by the ongoing global financial crisis and its implications for funded and unfunded pensions.

Address for correspondence: Robert Holzmann, Social Security Research Institute, Faculty of Economics and Administration, University of Malaya, 50603 Kuala Lumpur, Malaysia; E-mail: Robert.Holzmann@um.edu.my.

The article has profited from presentations of earlier drafts in Washington DC, Vienna, Kuala Lumpur and Chengdu, from comments and suggestions received, in particular from Mark Dorfman, David Robalino, Peter Rosner and two anonymous referees, and from editing support by Amy Gautam. The usual caveat applies.
After the fall of the Iron Curtain and the move in Central and Eastern Europe from central planning to market economies, the future for pension systems for some experts and policy-makers appeared bright and fairly certain once the initial crisis was overcome: transfer main parts of retirement income provisions from the public sector to the private sector i) to address fiscal unsustainability and projected further population ageing and ii) to accelerate financial market development that was expected to trigger higher economic growth to co-finance some of the transition costs. This policy vision emerged from various sources: the successful Chilean pension reform and similar reform attempts in Latin America; the seminal 1994 World Bank publication that proposed a multi-pillar pension system with a significant shift from publicly managed, unfunded defined benefit (DB) schemes to privately managed, fully funded defined contribution (DC) schemes (World Bank, 1994); and general enthusiasm and optimism for more market and financial intermediation instead of public intervention. This policy vision caught on in many countries: between 1988 and 2008, 29 countries introduced systemic reforms involving the establishment of a main funded pension pillar, but with variations in design, implementation and outcome (Figure 1). Before the financial crisis hit,

**Figure 1. Evolution of number of countries with (mandated and funded) second pillars as of 2008**

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of Countries</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>30</td>
</tr>
<tr>
<td>2006</td>
<td>28</td>
</tr>
<tr>
<td>2005</td>
<td>27</td>
</tr>
<tr>
<td>2004</td>
<td>24</td>
</tr>
<tr>
<td>2003</td>
<td>23</td>
</tr>
<tr>
<td>2002</td>
<td>21</td>
</tr>
<tr>
<td>2001</td>
<td>17</td>
</tr>
<tr>
<td>2000</td>
<td>15</td>
</tr>
<tr>
<td>1999</td>
<td>14</td>
</tr>
<tr>
<td>1998</td>
<td>12</td>
</tr>
<tr>
<td>1997</td>
<td>8</td>
</tr>
<tr>
<td>1996</td>
<td>7</td>
</tr>
<tr>
<td>1994</td>
<td>6</td>
</tr>
<tr>
<td>1993</td>
<td>3</td>
</tr>
<tr>
<td>1988</td>
<td>2</td>
</tr>
<tr>
<td>1980</td>
<td>1</td>
</tr>
</tbody>
</table>

*Note:* “Second pillar” is defined here as a pension system that has “mandatory personal retirement accounts”. For exceptions and other detailed information, see the source.

*Source:* World Bank Pension Database.
even more countries were poised for reform (e.g. Ukraine) and some may still do (e.g. Czech Republic).

The worldwide reassessment of the policy approach to pension system reform is broadly the result of three interrelated changes: a refocusing as regards the key objectives of pension systems (such as the refocus on basic protection for the vulnerable elderly); moving reform needs (such as recognizing the urgency of addressing the effects of population ageing and encouraging deferred retirement ages); and perceived and actual changes in enabling environments (such as more realistic views about the capacity of funded schemes to manage risks, the achievable rates of return, and the fiscal restrictions to finance transition deficits). This reassessment has led to reform reversals in a few countries (e.g. Argentina, Hungary and Slovakia) and to temporarily or permanently reduced funded pillar contributions in others (e.g. Estonia, Latvia and Poland), but not to a global rejection of the funding or the DC approach. The reassessment has strengthened the push for alternative or complementary reform approaches, such as Non-financial (or Notional) Defined Contribution (NDC) and Matching Defined Contribution (MDC) schemes. While these new approaches should help move pension systems towards greater coverage and financial sustainability, there are a number of issues that still await solutions, such as addressing uncertainty as regards increases to longevity.

Against this background, the remainder of this article is structured as follows: the next section highlights and reviews the key pension reform drivers that have triggered a reassessment of pension systems and reform approaches. How this reassessment translates into changes and main pension reform trends across pension pillars is then outlined. While these reform trends may help address some of the key policy concerns, knowledge gaps as regards how best to proceed remain. The article suggests a few knowledge areas on which the pension reform community will need to focus to make a difference. The latter include preparing solutions for implementation after the marginalization or, perhaps, even demise of Bismarckian schemes in countries with high levels of labour market informality; how to keep older workers in the labour market; and how pension products can satisfactorily address the challenges presented by lower rates of return, high volatility and increased longevity. Concluding comments are then offered.
while the basic objectives of pension systems – poverty alleviation and consumption smoothing – have remained unchanged, the focus of attention between the two has re-shifted or at least broadened. Furthermore, the intermediate objectives (goals) of old-age pensions have become the focus of greater attention and precision, i.e. issues such as coverage, adequacy, affordability, sustainability, robustness, and redistribution are nowadays more explicitly addressed, possible trade-offs highlighted, and the measurement of goals and the evaluation of the achievements of initiated pension reforms undertaken. Last but not least, the key instruments to achieve these intermediate and absolute objectives – in particular benefit structure and funding mechanisms – and their evaluated capability to deliver on these, are being reassessed in view of their recent and expected future performance.

In this section, four key events that are claimed to have changed the reform agenda are highlighted: the refocus on basic income protection for the elderly; the realization and implications of population ageing; the acknowledgment of lessons from the global financial crisis; and the reassessment of achievable rates of return on pension fund assets.

**Refocusing on basic income protection for the elderly**

The refocus on basic income provisions for the vulnerable elderly across countries has three distinct but interrelated origins:

- Disappointment among pension reformers with the levels of coverage extension achieved after systemic pension reforms is quite likely the first main reason for attention being directed towards the vulnerable elderly and their income protection. For emerging economies, there were strong expectations that systemic pension reforms (at that time directed towards funded individual account systems) would contribute to a major increase in coverage/reduced informality as the contribution-benefit link tightened and credibility of the scheme increased, financial sectors developed and per capita income grew. The coverage/per capita income link is strongly suggested by cross-country data (Figure 2). Yet, eight of 18 countries in Latin America continue to have a pension coverage rate of the labour force below 30 per cent, with only moderate improvements in some countries, and unsystematic fluctuations in others, over the last two decades. For the mean of the 18 countries, coverage decreased across the period 1990 to 2000, a reduction that was only partly recovered by 2010 (see Figure 3). As a result, when reform and coverage expectations were not met, countries were forced to consider other approaches to extend coverage (as discussed in the next section).

- The move towards a tighter contribution-benefit link that has characterized many recent reforms of earnings-related schemes has limited the capability to
Figure 2. The active coverage (contributor/working-age population)-per capita income connection (mid to late 2000s)

\[ y = -0.0004x^2 + 0.0325x + 0.0162 \]

\[ R^2 = 0.8151 \]

Source: World Bank Pension Database.

Figure 3. Percentage of the economically active population that contributes to pension systems in Latin America (1990-2010)

Source: Rofman and Oliveri (2012).
redistribute income towards low-income groups within the schemes. Furthermore, in many countries, effective coverage rates in contribution-based systems were being reduced due to lower contribution density, in part because of the difficult transition into the labour market for youth, and because of the increasing mobility of workers between formal and informal wage employment and into self-employment, also in Member countries of the Organisation for Economic Co-operation and Development (Holzmann, 2003).

- The International Labour Organization (ILO), which had been sidelined in the discussion in many countries about multi-pillar pension reforms and funded pensions, returned with force to the international social policy arena, advocating national “social protection floors”; i.e. access to essential health care for all, and income protection for the elderly, the unemployed and children, as part of a two-dimensional approach to building and maintaining comprehensive social security systems (see Bachelet, 2011). While implementation of basic old-age income protection in many low- and middle-income countries is still awaiting realization, the political push has caught on in development circles.

**Realizing the implications of population ageing**

Individuals and politicians in high-income countries are finally waking up to the challenge of population ageing. While the message of increased life expectancy, reduced fertility rates, and resultant deteriorating demographic (and system dependency) ratios has been around for some time, it has been largely ignored by politicians and the broader public until recently. The call for funding was at times motivated by incorrect arguments that funding by itself would be able to address population ageing and the incorrect assumption that the then high funded rates of return would continue in the future. Both arguments served also as a vehicle to motivate reforms that addressed short-term fiscal issues, but without a deeper understanding of what population ageing might imply for the pension system, other social programmes, the labour market and the society at large. Only gradually, it has been accepted that the effects of population ageing on pension systems can only be addressed in three ways: higher contributions, lower benefits, or later retirement; and this applies to both unfunded and funded systems. Individuals and policy-makers are also gradually acknowledging that the problem cannot be passed on to future generations, that higher contributions or lower benefits may not be the best approach, and that later retirement/working longer looks like the most natural and best solution. Yet, as we discuss more fully it has also become better understood that simply legislating an increase in the legal retirement age may not be sufficient. Reforms of the pension system to provide incentives for later retirement and policies to keep older workers in the
labour market are required to raise the effective retirement age (see European Commission, 2012; OECD, 2012a).

Effects of the global financial crisis

The financial, then economic, and now sovereign debt crisis that started in 2008 provides some sobering lessons for reformed pension systems, which are only gradually being understood and translated into policy actions. Three lessons stand out in particular:

• The fall, both, in GDP below the pre-crisis trajectory and in pension fund asset prices (not yet fully recovered in many countries) made a major dent in the financial situation of mandated pension schemes and individual benefit levels, whether unfunded or funded (Figures 4 and 5). The effects of alternative recovery scenarios for GDP compared to the pre-crisis benchmark and their impact on the deficit paths are presented in Figure 4. Under a severe crisis and low recovery scenario, the accumulated system deficits – in excess of the deficit path under the benchmark development for the former transition economies in Central and Eastern Europe – are projected to amount to 17 per cent of GDP. Yet the crisis impacts are still dwarfed by the cost impacts associated with future population ageing and the movement of the baby-boom “bulge” through the population structure. This population effect is particularly strong in the former transition economies and amounts for the years 2015 to 2075 to a simulated accumulated deficit of well over 50 per cent of GDP. The clear message is that greater efforts are needed, and sooner rather than later, to address population ageing and its effects on retirement schemes and public budgets if a future meltdown of pension systems is to be prevented (see IMF, 2011).

• The budgetary consequences of the financial crisis render the financing of transition costs for a newly introduced funded pillar more difficult. Cash flow problems, already substantial on their own, are for member States of the European Union (EU) aggravated by the debt accounting method applied under the Maastricht Treaty, which takes insufficient account of the fact that, with the reform, part of the increased explicit debt merely reflects a reduced implicit pension debt. These different financing issues have been used by some countries as an excuse to legally (e.g. Argentina) or virtually (e.g. Hungary and Slovakia) end the funded pillar and to divert the pension fund assets for public debt reduction purposes. Other countries have implemented temporary (e.g. Estonia and Latvia) or permanent (e.g. Poland) reductions in the contribution rate to the funded pillar, which should be to the benefit of the unfunded pillar by helping to reduce levels of public deficit and debt (see World Bank, 2009). While all reform countries were informed about the fiscal implications of a systemic pension reform, very few, if any, had a well thought-out plan for normal economic situations, let alone one for bad times.
**Figure 4. Simulations of the fiscal impact of the financial crisis on a synthetic country in Central and Eastern Europe**

Panel A. GDP growth scenarios

Panel B. Financial balance of first pillar

Source: Hinz et al. (2009).
Figure 5. Real Rates of Return (RRoR) of pension funds in CEE before and during the crisis

Panel A. RRoR until 2007

Panel B. RRoR 2008–2010

Source: Based on Rudolph (2012), with yearly-data partly interpolated as data requested but not received.
The temporary fall in asset prices and portfolio composition gave opponents of the systemic reform approach further ammunition (see Orenstein, 2011). But it also led reform supporters to review some of the design components and to propose improvements, such as life-cycle portfolios (i.e. a mandated move from an aggressive (higher-risk) to a more conservative (lower-risk) portfolio as an individual approaches retirement, as is done in Chile), and more flexibility around mandated annuitization to avoid a locking-in of losses (see World Bank, 2008).

**Rate of returns on assets**

The 2008 financial crisis and its ensuing impacts reinforced the already sober expectations for the rates of return of funded (and the growth rate of wages for unfunded) schemes and increased uncertainty regarding regulatory reforms of pension funds (see IPE, 2011). The high return expectations of the 1990s were first dampened by the bursting of the dot.com bubble in the early 2000s and have been under general pressure since 2008. The longer-term decline of real rates of return is clearly visible for key OECD countries with major financial markets (Germany, Japan, United Kingdom and the United States, see Figure 6). The current plight of pension fund investment returns in OECD and non-OECD countries can be appreciated in Figure 7 (Panels A and B).

The more recent and ongoing stark fluctuations in asset prices, the possible non-existence of an asset with zero risk (i.e. government bonds), and the likely “new normal” future of low real asset return for a protracted period of time create major uncertainties for individuals, policy-makers and pension fund professionals. More critically, all this begs the question of the future of the size of funded pension pillars (compared to the unfunded pillars) and possibly even about their very existence. This question is additional to the still unanswered ones about the (international) performance of pension funds on a comparable basis and about how to usefully define such a basis (see Hinz et al., 2010). There is some recent evidence that even in countries such as Chile the expected rates of return on financial assets may not necessarily surpass the growth rate of wages, which is the rate unfunded systems are able to pay (see Fajnzylber and Robalino, 2012).

To sum up, this section has put forward that the focus, reform needs and the enabling environment have changed since the early 1990s and has offered examples as concerns these changes as well as of underlying events, experiences and perceptions. The impact of these changes is visible in changing reform trends in countries across the world, reviewed next, but also present in the adjustment of

---

1. Albeit largely temporary, the situation had not fully recovered by the end of 2010 in OECD countries (OECD, 2011b), while pension fund asset to GDP ratios in countries of Asia and the Pacific by the end of 2010 were well above their 2007 levels (Hu, 2012). For the latter, this reflects both recoveries in asset prices as well as coverage levels.
the World Bank’s thinking about pension systems and reform (see Holzmann and Hinz, 2005; Holzmann, Hinz and Dorfman, 2008; World Bank, 2008, and 2009; Dorfman and Palacios, 2012).

**Main reform trends across pension pillars**

The changes in focus, reform needs and enabling environments outlined in the prior section have a bearing on the reform trends across the world that can be highlighted through the design and implementation innovations in key pension pillars witnessed since 1990. To highlight this trend as it applies to distinct pillars, we use the definition and structure of the World Bank’s 2005 “five pillars” model (Holzmann and Hinz, 2005). This approach separates a “zero pillar” from a “first public pillar” to better differentiate between the poverty alleviation (zero pillar) and consumption smoothing (first pillar) objectives of public and unfunded schemes. The second pillar refers to mandated funded schemes (DB or DC), the third pillar refers to occupational or personal voluntary funded schemes (defined originally in World Bank, 1994). The fourth pillar offers informal (family), market-based (voluntary), and public support (e.g. health care) to the elderly that has a strong bearing on the scope and design of the other pension pillars; this pillar is not discussed further herein.
The use of the pillar structure to highlight reform trends is motivated by the very broad and increasing support for the multi-pillar pension concept. The structure can be viewed: as an ordering principle for the analysis of existing systems; as a means of risk diversification (with unfunded pillars allocating savings to the pay-as-you-go asset and funded pillars allocating savings to financial assets); and as recognition that different pillars have varying degrees of importance for the key target groups in a population (e.g. formal-sector workers, those working in the informal sector, and the lifetime poor). For low- and many middle-income countries, those working in the informal sector represent by far the largest group.

Table 1 highlights the basic system architecture for the mandated pillars by World Bank regions for 2011 (and 1990). As can be seen, in 2011, the large majority of countries rely on first pillar schemes (that can be Notional Defined Benefit (NDB), Notional Defined Contribution (NDC), and public DC/provident fund arrangements) and almost half have a zero pillar (100 per cent of the 24 traditional OECD countries have a zero pillar). Only 32 countries have mandated and funded...
pillars, of which two have a DB structure (Iceland and the Netherlands); the rest have DC structures. The most significant changes since 1990 are the more than doubling of countries with a zero pillar, the transformation of NDB to NDCs schemes in eight countries, and the introduction of Funded Defined Contributions (FDC) schemes in 29 countries that mostly complemented and only rarely replaced NDB schemes.  

2. Data in Pallares-Miralles, Romero and Whitehouse (2012) reveal that in 2011 the majority of countries (60) were operating a separate scheme for civil servants, albeit some progress has been made in recent years to integrate these workers into the general scheme.
The main objective of the zero pillar is poverty alleviation, and as this has become the focus of development policy, its importance has increased over the past two decades as is visible in Table 1. In its simplest form, it is part of the social “safety net” that protects poor and vulnerable individuals of all ages through universal, means-tested, or conditional cash transfers; these instruments have been the revolution in low- and middle-income countries over the last decade (see Grosh et al., 2008). In the more old-age focused form of social pensions, zero pillar schemes provide income transfers to the elderly, typically via means-testing for the younger elderly, and at times without means-testing for the older elderly. These schemes are now ubiquitous in traditional OECD member countries, and are increasingly but slowly gaining traction in low- and middle-income countries.

An important step to integrate the new zero pillar with the earnings-related (funded) pillar took place in Chile in 2008; this reform is seen as a benchmark (see Rofman, Fajnzylber and Herrara, 2008). There are also recent initiatives to implement ex ante transfers in the form of matching contributions for informal wage-employment workers and the self-employed (discussed further below). Figure 8

Table 1. Basic system architecture by region, 2011 (and 1990)

<table>
<thead>
<tr>
<th>Region</th>
<th>Pillar 0</th>
<th>Pillar 1</th>
<th>Pillar 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Targeted</td>
<td>Basic</td>
<td>NDB</td>
</tr>
<tr>
<td>East Asia &amp; Pacific</td>
<td>4</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>Europe &amp; Central Asia</td>
<td>11</td>
<td>4</td>
<td>28</td>
</tr>
<tr>
<td>High income: OECD</td>
<td>8</td>
<td>9</td>
<td>16</td>
</tr>
<tr>
<td>Latin America &amp; Caribbean</td>
<td>16</td>
<td>2</td>
<td>29</td>
</tr>
<tr>
<td>Middle East &amp; North Africa</td>
<td>1</td>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>South Asia</td>
<td>3</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Sub-Saharan Africa</td>
<td>3</td>
<td>2</td>
<td>30</td>
</tr>
<tr>
<td>2011 Total</td>
<td>46</td>
<td>21</td>
<td>130</td>
</tr>
<tr>
<td>Grand total</td>
<td>67</td>
<td>155</td>
<td>34</td>
</tr>
<tr>
<td>1990 Total</td>
<td>20</td>
<td>10</td>
<td>140</td>
</tr>
<tr>
<td>Grand total</td>
<td>30</td>
<td>157</td>
<td>5</td>
</tr>
</tbody>
</table>

Note: NDB/NDC: Notional Defined Benefit or Defined Contribution scheme; FDB/FDC: Financial DB or DC scheme; PF: Provident Fund.

Source: Author’s compilation based on Pallares-Miralles, Romero and Whitehouse (2012).

Zero pillar

The main objective of the zero pillar is poverty alleviation, and as this has become the focus of development policy, its importance has increased over the past two decades as is visible in Table 1. In its simplest form, it is part of the social “safety net” that protects poor and vulnerable individuals of all ages through universal, means-tested, or conditional cash transfers; these instruments have been the revolution in low- and middle-income countries over the last decade (see Grosh et al., 2008). In the more old-age focused form of social pensions, zero pillar schemes provide income transfers to the elderly, typically via means-testing for the younger elderly, and at times without means-testing for the older elderly. These schemes are now ubiquitous in traditional OECD member countries, and are increasingly but slowly gaining traction in low- and middle-income countries.

An important step to integrate the new zero pillar with the earnings-related (funded) pillar took place in Chile in 2008; this reform is seen as a benchmark (see Rofman, Fajnzylber and Herrara, 2008). There are also recent initiatives to implement ex ante transfers in the form of matching contributions for informal wage-employment workers and the self-employed (discussed further below). Figure 8
offers a taxonomy of retirement income transfers, including the guarantee of minimum pensions within the mandated (first or second) pillar.

The key questions about the zero pillar have changed little, but a few new ones have been added. For low- and middle-income countries, discussions on fiscal affordability, disincentive effects, and administrative issues of universal benefits compared to means-tested approaches remain evergreen. New to the discussion is the potential role of ex ante interventions to address poverty and adequacy issues upstream, and the impact of social pensions on informality and thus coverage under formal earnings-related schemes (discussed below).

**First pillar**

The typical mandated, unfunded, and DB-type first public pillar has undergone reforms to various degrees across rich and poor countries. In OECD countries, reforms have been mostly parametric, and have included: a reduction in generosity (such as a lower annual accrual rate); lengthening of the assessment period, at times to all contribution periods; the introduction of decrements for earlier and increments for later retirement; and, in a number of countries, an increase in the standard retirement age (see OECD, 2011a, 2012a). Although all of these measures should ensure that the first pillar is sustainable, this seems to have been achieved in some, but not many, countries. In most countries, further parametric reforms are needed to address population ageing-related fiscal as well as labour market- and social policy-related incentive concerns. The challenge here is to deal with the political discretion that surrounds this type of reform which risks compromising the long-term solvency of pension schemes.
Against this background, the NDC scheme, a systemic reform innovation that maintains the unfunded character of the public first pillar, is attracting increasing attention in Europe and elsewhere. The NDC scheme operates as a DC scheme in the accumulation phase and for the annuity calculation at retirement, but remains unfunded (except for, perhaps, a reserve fund to address short-term liquidity issues). To achieve solvency, the NDC scheme offers only the notional rate of return that keeps the system solvent and only the annuity amount that is consistent with the remaining (projected cohort) life expectancy at retirement.

This pension reform innovation was introduced in the second half of the 1990s in Italy, Latvia, Poland and Sweden and has weathered the financial crisis fairly well (see Chlon-Domińczak, Franco and Palmer, 2012). In 2009, Norway legislated a reform now under implementation that mimics many but not all of the NDC features. In 2010, Egypt legislated an NDC reform for which implementation is envisaged for 2013, but is quite likely to be again delayed. This reform approach is also under discussion in many EU countries, as well as in countries as diverse as Belarus, the People’s Republic of China, Lebanon and Uruguay.

The attractions of the NDC scheme are: the promise of solvency, even during adverse economic times and under severe population ageing; the DC-type incentive structure to address labour market concerns and broader social changes (such as increasing life expectancy and rising divorce rates); and the openness to future partial or full shifts towards FDC schemes once the enabling environment has been created. While promising, the NDC approach is not foolproof, i.e. immune against policy mistakes, and there are still a few conceptual and operational issues that have not yet been satisfactorily solved, such as the design of an effective balancing mechanism, including the measurement of assets and liabilities; the interactions of NDC schemes with other pillars and benefits (e.g. disability and survivors); reliable methods to project cohort life expectancy and equitable approaches for sharing longevity risks; and defining and establishing the enabling environment for NDC implementation in low- and middle-income countries (see the recent anthology on NDCs in two volumes by Holzmann, Palmer and Robalino, 2012 and 2013; and the first anthology exploring the early period of NDC developments by Holzmann and Palmer, 2006).

Second pillar

The mandated and funded DC pillar has been the main innovation in pension reform design since it was introduced in Chile in 1981. While the Chilean reform is considered the most successful benchmark worldwide, very few of the other 29 countries that had introduced second pillars as of 2008 have closely copied its design and implementation; in Europe and Central Asia, only Poland has done so. This may explain some of the variance in outcomes. Other than Chile, no country
has conducted such a rigorous analytical evaluation of its second pillar scheme, or introduced reforms to improve existing structures (e.g. introduction of life-cycle funds or an integrated “solidarity pillar”, or measures to reduce costs and fees) based on thorough analysis.

Consequently, many of the smaller and larger changes in countries with systemic reforms represent experimental corrections meant to address issues as they emerged. Three are highlighted:

- The high costs and fees of funded pensions have long been a concern, as their size presents a major reduction of the future benefit level. Fees amounting to 100 or more basis points lead to a reduction in ultimate benefits of 20 per cent or more. Attempts to control costs and fees include: limits and caps; constraints on marketing efforts; innovative differentiated fee limits to create a competitive market; and the creation of clearing houses linked with blind accounts to reduce administrative and marketing costs, and to limit pension funds to an asset management function. While broadly moderately effective, these attempts have not been successful in limiting costs and fees to basis points in the low double digits. And there are conceptual considerations that the current approaches may never be able to produce sufficiently low costs and that attempts to increase competition risks actually deteriorating the outcome. Such potential market failures would result from the specific production technology of financial institutions (working with increasing returns to scale) as well as power asymmetry (between the financial institution and individual), both of which call for the application of an industrial organization model.

- At their initiation, pension funds were typically subjected to a “Draconian regime” to avoid early mishaps and a discrediting of the reform approach. Tight regimes included quantitative restrictions on asset classes in which the pension fund could invest. Over time, the restrictions were relaxed, and in advanced countries, even abandoned. Concurrently, the financial significance of pension funds increased, making them a critical component of the financial market, on par with banks and insurance companies. This led to the extension of risk-based

3. For the ease of use, the returns on investment and the related fees charged by pension funds are often expressed in basis points, i.e. percentage points times 100. Hence, a gross real rate of return of, say, 4.5 per cent equals 450 basis points; the overall fees charged by the pension fund may reduce the net rate of return to 3.0 per cent, i.e. by 1.5 percentage points or 150 basis points. To translate all charges (on flows such as contributions, on stocks such as managed capital, or lump-sum charges for exit and entry) into a single number requires modelling in a life-cycle context. The same model (and assumptions) can be used to translate fees per year into a reduction of the benefit level. As a rough but useful guide on the latter: fees amounting to 1 per cent (or 100 basis point) on accumulated capital each year across the whole life cycle are broadly equivalent to a 20 per cent tax on contributions or a reduction of 20 per cent of the pension benefit (see Whitehouse, 2000).

4. For a more detailed discussion of these issues, see Impavido, Lasagabaster and García-Huitrón (2010).
supervision methods, developed originally for bank supervision, to pension funds in a number of developed economies (such as Australia, Denmark and the Netherlands) but also in emerging economies (such as Mexico). The different approaches applied provide a rich set of information for professionals involved in pension fund issues in both emerging and developed countries (see Brunner, Hinz and Rocha, 2008).

- Much of the focus on the development of the second pillar in emerging economies in Latin America and Central and Eastern Europe was on the accumulation phase, as the payout phase was not to happen for many years. Yet the payout phase has almost arrived and the reform countries that introduced a second pillar now face the challenge of organizing the payout for retiring workers. This effort entails introducing a well-organized and well-supervised market for retirement products, including marketing activities and intermediaries. Alternatively, governments could provide the annuity in exchange for handing over the accumulated resources (as is done in Sweden). While some advanced economies (and Chile) provide useful lessons for the structure and operations of annuity markets (see Rocha, Vittas and Rudolph, 2011), the features and trade-offs for key products are still little known (Vittas, 2011), and financial assets to address inflation and longevity risks are not available (as discussed in the next section).

Third pillar

Establishment of this pillar of occupational or personal voluntary funded pensions has often preceded the creation of the mandated second pillar, but how best to regulate and supervise these schemes remains a challenge across the world. However, this pillar is receiving increasing levels of interest from policy-makers in all regions as a means to offer some coverage to those employed in the informal sector in low- and middle-income countries, and to offer individuals an opportunity to compensate for reduced public generosity with individual saving efforts in high-income countries. The links between public generosity and voluntary pension coverage clearly exist for OECD economies (Figure 9). To motivate voluntary participation, countries have traditionally used favourable tax regimes with limited success and questionable distributive effects. Increasingly countries across all income levels are using ex ante subsidies in the form of matching contributions (hence, Matching Defined Contribution (MDC) schemes) and/or other nudging devices and advocacy, particularly for informal workers (see Palacios and Robalino, 2009; Hinz et al., 2013). The latter include implementing massive information campaigns and dramatically reducing transactions costs for enrolling and contributing through the use of mobile phones, such as in the Mbao programme in Kenya (Retirement Benefits Authority, 2011).
MDC schemes are well known in developed economies such as the United States (401(k) plans), Germany (Riester pension), and New Zealand (KiwiSaver) for supplementing public benefits. However, MDCs are also gaining traction in emerging economies such as India and China for offering basic benefit coverage. For example, China started pilot schemes for the rural population in 2009, an experiment that was extended to the urban population in 2011. Whether the expectations of these schemes can and will be met is under discussion and review (Hinz et al., 2013). This adds to the better known issues of third pillar coverage, such as high administrative costs, lack of good annuity products, and the role and scope of regulation.

Overall, the move toward funded pensions through mandated (second pillar) or voluntary (third pillar) arrangements is visible in the size and growth of assets over the last decade for 13 major pension markets, from USD 14.8 trillion to an estimated aggregate of USD 26.5 trillion by end-2011; a record high if measured in absolute terms (Table 2). Pension assets in per cent of GDP reached 72.3 per cent, still below the 2007 level of 78.9 per cent but also below the 2010 ratio of 75.5 per cent. The crisis after 2007 left a dent in most but not all countries and regions, which has not yet been recovered by end-2011 (Towers Watson, 2012). Data for Asia and the Pacific signal higher ratios of pension assets to GDP in 2010 compared to 2007 in all reviewed economies (Hu, 2012).
Centralized public pre-funding

The trend to create and expand public pension reserve funds to support unfunded pillars, to more generally allow for intergenerational consumption smoothing, or to offer societal cushioning against adverse future events, expanded over the last decade to some 18 OECD countries and a number of other major economies, such as Argentina and China. By the end of 2010, for OECD countries for which funds data are available, public pension reserve funds held USD 4.8 trillion (compared to USD 4.6 trillion in 2009; Table 3). Given the budgetary crises in many of these countries, it is doubtful that these funds will receive additional resources or even survive. The situation is likely to be different in resource-rich countries with their earmarked reserve funds (such as in Australia and Norway) or in countries with

<table>
<thead>
<tr>
<th>Country</th>
<th>Total assets 2011 In USD (billions)</th>
<th>In percentage of GDP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>1,301</td>
<td>96</td>
</tr>
<tr>
<td>Brazil</td>
<td>321</td>
<td>15</td>
</tr>
<tr>
<td>Canada</td>
<td>1,303</td>
<td>78</td>
</tr>
<tr>
<td>France</td>
<td>129</td>
<td>5</td>
</tr>
<tr>
<td>Germany</td>
<td>468</td>
<td>14</td>
</tr>
<tr>
<td>Hong Kong (China)</td>
<td>84</td>
<td>34</td>
</tr>
<tr>
<td>Ireland</td>
<td>101</td>
<td>50</td>
</tr>
<tr>
<td>Japan</td>
<td>3,383</td>
<td>55</td>
</tr>
<tr>
<td>Netherlands</td>
<td>1,046</td>
<td>133</td>
</tr>
<tr>
<td>South Africa</td>
<td>227</td>
<td>62</td>
</tr>
<tr>
<td>Switzerland</td>
<td>693</td>
<td>115</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>2,394</td>
<td>101</td>
</tr>
<tr>
<td>United States</td>
<td>16,080</td>
<td>107</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>27,510</strong></td>
<td><strong>72</strong></td>
</tr>
</tbody>
</table>

Notes:
1. Assets include only those from closed entities.
2. Excludes RRSP.
3. Pension assets from company schemes only.
4. Only includes total of autonomous pension funds.
5. Excludes personal stakeholders DC assets.
6. Includes IRAs.

Table 3. Size of public pension reserve fund markets in selected OECD countries and other major economies, 2010

<table>
<thead>
<tr>
<th>Country</th>
<th>Names of the fund or institution</th>
<th>Founded in</th>
<th>Assets (USD billions)</th>
<th>% of GDP</th>
<th>% increase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selected OECD countries</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>United States</td>
<td>Social Security Trust Fund</td>
<td>1940</td>
<td>2,609.0</td>
<td>17.9</td>
<td>2.7</td>
</tr>
<tr>
<td>Japan</td>
<td>Government Pension Investment Fund</td>
<td>2006</td>
<td>1,312.8</td>
<td>25.9</td>
<td>nd</td>
</tr>
<tr>
<td>Korea (Rep. of)</td>
<td>National Pension Fund</td>
<td>1988</td>
<td>280.4</td>
<td>27.6</td>
<td>16.7</td>
</tr>
<tr>
<td>Canada</td>
<td>Canadian Pension Fund</td>
<td>1997</td>
<td>136.0</td>
<td>8.6</td>
<td>13.0</td>
</tr>
<tr>
<td>Sweden</td>
<td>National Pension Funds (AP1-AP4 and AP6)</td>
<td>2000</td>
<td>124.7</td>
<td>27.2</td>
<td>8.1</td>
</tr>
<tr>
<td>Spain</td>
<td>Social Security Reserve Fund</td>
<td>1997</td>
<td>85.3</td>
<td>6.1</td>
<td>7.3</td>
</tr>
<tr>
<td>France</td>
<td>AGIRC-ARRRCO</td>
<td>nd</td>
<td>71.7</td>
<td>2.7</td>
<td>nd</td>
</tr>
<tr>
<td>Australia</td>
<td>Future Fund</td>
<td>2006</td>
<td>65.8</td>
<td>5.5</td>
<td>8.4</td>
</tr>
<tr>
<td>France</td>
<td>Pension Reserve Fund</td>
<td>1999</td>
<td>49.0</td>
<td>1.9</td>
<td>11.1</td>
</tr>
<tr>
<td>Ireland</td>
<td>National Pensions Reserve Fund</td>
<td>2000</td>
<td>32.3</td>
<td>15.9</td>
<td>9.3</td>
</tr>
<tr>
<td>Belgium</td>
<td>Zilverfonds</td>
<td>2001</td>
<td>23.3</td>
<td>5.0</td>
<td>4.3</td>
</tr>
<tr>
<td>Norway</td>
<td>Government Pension Fund – Norway</td>
<td>2006</td>
<td>23.1</td>
<td>5.6</td>
<td>16.9</td>
</tr>
<tr>
<td>Portugal</td>
<td>Social Security Financial Stabilisation Fund</td>
<td>1989</td>
<td>12.8</td>
<td>5.6</td>
<td>2.5</td>
</tr>
<tr>
<td>New Zealand</td>
<td>New Zealand Superannuation Fund</td>
<td>2001</td>
<td>11.2</td>
<td>7.9</td>
<td>17.1</td>
</tr>
<tr>
<td>Chile</td>
<td>Pension Reserve Fund</td>
<td>2006</td>
<td>3.8</td>
<td>1.9</td>
<td>12.2</td>
</tr>
<tr>
<td>Mexico</td>
<td>IMSS Reserve</td>
<td>nd</td>
<td>3.6</td>
<td>0.3</td>
<td>–6.7</td>
</tr>
<tr>
<td>Poland</td>
<td>Demographic Reserve Fund</td>
<td>2002</td>
<td>3.4</td>
<td>0.7</td>
<td>39.1</td>
</tr>
<tr>
<td>Total selected OECD countries</td>
<td></td>
<td></td>
<td>4,848.1</td>
<td>19.6</td>
<td>5.0</td>
</tr>
</tbody>
</table>

Other major economies

<table>
<thead>
<tr>
<th>Country</th>
<th>Names of the fund or institution</th>
<th>Founded in</th>
<th>Assets (USD billions)</th>
<th>% of GDP</th>
<th>% increase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Saudi Arabia</td>
<td>General Organization for Social Insurance</td>
<td>1969</td>
<td>400.0</td>
<td>106.4</td>
<td>nd</td>
</tr>
<tr>
<td>China</td>
<td>National Social Security Fund</td>
<td>2001</td>
<td>126.5</td>
<td>2.2</td>
<td>10.3</td>
</tr>
<tr>
<td>Argentina</td>
<td>Sustainability Guarantee Fund</td>
<td>2007</td>
<td>45.7</td>
<td>12.3</td>
<td>26.4</td>
</tr>
<tr>
<td>Total other major economies</td>
<td></td>
<td></td>
<td>572.2</td>
<td>75.9</td>
<td>14.6</td>
</tr>
</tbody>
</table>

Memo item: Sovereign Wealth Funds with a pension focus

<table>
<thead>
<tr>
<th>Country</th>
<th>Names of the fund or institution</th>
<th>Founded in</th>
<th>Assets (USD billions)</th>
<th>% of GDP</th>
<th>% increase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Norway</td>
<td>Government Pension Fund – Global</td>
<td>1990</td>
<td>509.1</td>
<td>122.8</td>
<td>16.6</td>
</tr>
<tr>
<td>Russian Federation</td>
<td>National Wealth Fund</td>
<td>2008</td>
<td>88.4</td>
<td>5.9</td>
<td>–2.7</td>
</tr>
</tbody>
</table>

Source: OECD (2011b).
sovereign wealth funds with a pension focus (such as in the Russian Federation). And among a number new resource-rich developing countries in Asia (such as Kazakhstan and Mongolia), Latin America (such as Brazil, following Chile’s copper fund) and perhaps soon also some East and West African countries, there is interest in creating wealth funds to address current policy challenges (such as an appreciating exchange rate) and expected future shocks, including population ageing. Such funds have a tradition in the oil-rich countries of the Gulf Cooperation Council.

Key challenges ahead

The key trends in pension reform discussed above are very likely to stay and strengthen, in particular the provision of basic pension benefits (zero pillar), the move toward a closer linking of contributions to benefits through NDC-type structures or proxies (pillar 1), and the expansion of supplementary funded pensions of a voluntary nature to compensate reduced public generosity (pillar 2). Mandated funded pensions of the DC type (pillar 3) are likely to stay but their expansion as a complement to reformed NDB/new NDC (pillar 1) schemes may have to await fiscal consolidation across the world (except, perhaps, in old and new resource-rich countries).

These international trends in pension reform are helpful in support of the key objectives of pension schemes – poverty alleviation and consumption smoothing – but not sufficient for delivering on the performance goals (i.e. coverage, adequacy, affordability, sustainability and robustness). There are a few minor challenges around the redesign of disability and survivors pensions to respond to socio-economic changes and to make these benefits consistent with DC-type designs for old-age protection. There are already larger challenges as regards how to establish full portability of acquired rights across professions, sectors and countries (for pensions and other social benefits, in particular health care). And there are the perennial political economy challenges of making a pension reform happen (i.e. start it and make it successful), even when a convincing reform approach is on the table.

This section highlights three crucial challenges for which the relevant questions have often not even been asked and by consequence good answers are largely missing: i) will Bismarckian systems with their mandated and often high contribution rates survive in countries with high levels of labour market informality, and what happens after their marginalization or, perhaps, demise? ii) how can older workers be kept in the labour market?; and iii) how to address the likely lower rates of return in the accumulation phase and unknown longevity increases in the payout phase?
Preparing for Bismarck’s demise and succession

As discussed above, to address low pension coverage, many low- and middle-income countries introduced basic provisions in the form of social pensions, assuming that over time workers would move towards formal-sector employment and participation in a mandated and earnings-related (funded or unfunded) scheme. Yet, these very provisions (essentially subsidies) risk becoming a tax on formal work and may create an incentive for individuals to take informal jobs or move into self-employment while they build up their own retirement provisions (e.g. businesses, houses, financial assets, etc.), knowing that the “safety net” or “floor” will be there for them if everything else fails. Such a tendency seems to be particularly pronounced in Latin America, as suggested by recent analytical work (see for example, Levy, 2008; Aterido, Hallward-Driemeier and Pagés, 2011; Ribe, Robalino and Walker, 2012). If confirmed, this may risk sounding the death knell of Bismarckian systems in many low- and middle-income countries. But what would a future retirement income scheme look like: only basic mandatory provisions plus unstructured voluntary retirement provisions? Or innovative new schemes in which basic provisions are also based on individual accounts funded across the life cycle by government and augmented by individual savings supplements of unknown design? In the absence of clear answers, this question calls for much greater attention by researchers and policy-makers.

Keeping older workers in the labour market

There is a growing understanding and emerging consensus among many (but not yet all) policy-makers in OECD economies that part of the solution to the ageing problem is to be found in longer labour market participation and hence the later retirement of individuals. Thus, there is a willingness to raise the legal retirement age as witnessed in many OECD countries (OECD, 2011a and 2012a). However, there is not yet full recognition that a legal decision alone will do little to raise the effective retirement age unless major reforms in the labour market take place to allow older workers to remain productively employed and to provide employers with incentives to keep them employed/offer them jobs. While the basic ingredients are simple to pronounce (i.e. keep workers healthy, skilled, and motivated), the policies to do so are less known or implementable across varying cultural contexts. And implementing successful reforms requires reviewing many or all of a country’s institutions and regulations, including the incentive structure of the

5. Much of this thinking is still vivid in the ILO’s two-dimensional social protection agenda (see, ILO, 2010).
pension system. But unless workers are convinced that they will have a suitable job when they are older, they are likely to oppose a legal increase in the retirement age. While OECD countries have started introducing promising reforms at the firm level to improve labour market opportunities for older workers (see for example European Centre for the Development of Vocational Training, 2008), any scaling-up or transfer to other countries is, for the time being, limited by the lack of sound analytical preparation and rigorous monitoring and evaluation of these innovations. Such issues, however, are not yet even on the horizon of most middle-income, not to mention low-income, countries.

Addressing lower rates of return in the accumulation phase and unknown longevity increases in the payout phase

There are many indications that the rates of return of funded and unfunded pensions may be lower in the future than in the past. For funded pensions, lower (real) rates of return may be temporary, spanning the current period of fiscal consolidation and slow growth in developed economies – a period that might, however, last a couple of decades. For unfunded pension in developed economies a lower (implicit) rate of return will be the result of lower labour productivity (perhaps related to an ageing population) plus a low and at times negative labour force growth rate (due to a projected shrinking of the population in some OECD countries). Thus changing the funding mode of pension schemes may do little to improve the benefit outlook; more likely will be higher contributions, lower benefits and an increase in the retirement age beyond that dictated already by higher life expectancy and a lower fertility rate. The question is to what extent such a choice between higher savings and later retirement (to avoid inadequate benefit levels) should be mandated by government, for example by an indexed increase in the minimum and standard retirement age? Or whether such a decision about the choice between higher voluntary contributions or later retirement should be left to individuals?

Perhaps the main open design issue for NDC and FDC (and indirectly also for FDB and NDB) schemes concerns the payout phase and how best to address shocks in longevity. The financial, social, and political sustainability of a DC scheme requires translating accumulated savings at retirement into an annuity that takes account of the projected remaining cohort life expectancy. Required also are adjustment features that are fair and transparent if the longevity projections turn out to be wrong. At the moment, DC systems do not have reliable projection methods for mortality rates and remaining life expectancies (see for example Alho, Bravo and Palmer, 2013). Neither NDC nor FDC (or FDB and NDB) schemes have
robust methods for distributing the risks if projection errors take place. In FDC schemes that typically use the private-sector annuity market, the financial sector provider officially takes the risk, while individuals and/or the government bear the final risk if insurers go bankrupt; yet, how the bearing of the risk ultimately would be distributed remains untested and is little addressed in the writings of the insurance industry (see for example Liedtke and Schanz, 2012). In NDC schemes run according to the rule book, projection errors can be corrected by adjusting the notional interest rate and the annual indexation of pensions – effectively distributing risks between active and retired plan members in specified proportions. For FDC and NDC schemes, there are several proposals to address the longevity issue through longevity bonds that create an effective hedge against longevity risk (see Blake, Boardman and Cairns, 2010; Palmer, 2013). These proposals need further analysis and piloting prior to full implementation, however. Across the world, financial market instruments to hedge longevity risk are still virtually non-existent (Roy, 2012). Whatever progress can be made in this area, paying attention to population ageing and longevity issues must form part of reforms to increase confidence among investors and the population at large in public- and private-sector balance sheets (IMF, 2012).

Concluding remarks

Across the world, pension systems and their reforms are in a constant state of flux driven by a shifting focus, moving reform needs, and a changing enabling environment. These reform drivers reflect objective events but also changes in views on objectives and feasible instruments by an international pension community of varying influence and strength. Over recent decades, this has led to a number of redirections and innovations throughout the world, including: the introduction or strengthening of basic protection for the vulnerable elderly; the move towards funded and unfunded mandated DC schemes; and increased nudging by governments to encourage benefit coverage and “top-ups” under voluntary and funded provisions.

The move towards pre-funded old-age income provisions is now itself under review as a consequence of the fallout of the financial crisis/recession/borderline depression. Areas of concern include the fall in asset prices, high fluctuations in rates of return, and the possibility of lower real risk-adjusted rates of return as the “new normal”. Some of the lessons from the crisis are straightforward and easy to implement, such as the move towards life-cycle funds. Others may be more difficult to deal with, such as the outlook to lower rates of return plus higher return-rate volatility.

However, such a review is unlikely to ring the death knell of pre-funded old-age pensions, if only for the simple reason that the fiscal conditions after the crisis and
the fiscal implications of the expected further ageing of populations limit both the capacity and the willingness of governments to take complete responsibility for ensuring retirement income. However, providers of funded provisions will need to work hard to re-establish confidence and to deliver what is promised to keep their share in the retirement income market.

Both funded and unfunded pension schemes will need to address a number of challenges for which the relevant questions are only slowly emerging and by consequence good answers are largely missing. This includes design issues for linking basic income protection with earnings-related schemes in countries with a high level of informality, effective policies to keep older workers healthy, skilled and motivated to stay in the labour market and to support an ever-increasing effective retirement age, and approaches to share the risks of unknown longevity increases in an efficient and equitable manner.

Bibliography


Global pension systems and their reform: Worldwide drivers, trends and challenges


Global pension systems and their reform: Worldwide drivers, trends and challenges


Global pension systems and their reform: Worldwide drivers, trends and challenges


Actuarial balance sheets as a tool to assess the sustainability of social security pension systems

Assia Billig and Jean-Claude Ménard

Office of the Chief Actuary, Ottawa, Canada

Abstract  The choice of the methodology used to produce a social security pension system’s balance sheet is mainly determined by the system’s financing approach. In this article, it is shown using the example of the Canada Pension Plan that if the assessment of the financial sustainability of a pay-as-you-go or partially funded system is done through the means of an actuarial balance sheet, then the methodology used should take into account future contributions of current and future participants. The balance sheets produced using the open group approach, as well as methodologies used in United States and Sweden, are discussed.

Keywords  social security scheme, pension scheme, method of financing, actuarial valuation, Canada, Sweden, United States

Addresses for correspondence: Assia Billig, Office of the Chief Actuary (OCA), Office of the Superintendent of Financial Institutions, 255 Albert Street, 12th Floor, Ottawa, Ontario K1A 0H2, Canada; Email: assia.billig@osfi-bsif.gc.ca. Jean-Claude Ménard, Office of the Chief Actuary, Office of the Superintendent of Financial Institutions, 255 Albert Street, 12th Floor, Ottawa, Ontario K1A 0H2, Canada; Email: jean-claude.menard@osfi-bsif.gc.ca.

This article is derived from a recent Actuarial Study published by the Office of the Chief Actuary (OSFI, 2012), and both authors were part of the working team for this Actuarial Study. The authors would like to thank the OCA staff who assisted in the preparation of the original Actuarial Study as well as this article. The authors would also like to express their gratitude to Ole Settergren for answering questions regarding the Inkomstpension system and providing valuable feedback, and to Alice Wade for answering questions regarding the United States OASDI programme.
Introduction

This article addresses the methodologies used to produce actuarial balance sheets for social security pension systems. The balance sheet is often a useful measure of the financial position of the pension system. The balance sheet results are relatively easy to explain, and this concept is widely used for pension plans other than social security systems. At the same time, the balance sheet can produce very misleading results if the proper methodology is not used.

The choice of the methodology used to produce a social security pension system’s balance sheet is mainly determined by the system’s financing approach. For fully funded systems, the accrued liabilities are assumed to be funded in advance; therefore, balance sheets under closed groups (i.e. that take into account current contributors only) with or without future accruals are appropriate for such plans. In contrast, pay-as-you-go (PAYG) and partially funded systems represent social contracts where, in any given year, current contributors allow the use of their contributions to pay current beneficiaries’ benefits. As a result, such social contracts create a claim for current and past contributors to contributions of future contributors. The proper assessment of the financial sustainability of a social security PAYG or partially funded system by means of its balance sheet should take these claims into account. The traditional closed group methodologies do not reflect these claims since only current participants are considered. In comparison, the open group approach accounts explicitly for these claims by considering the benefits and contributions of both current and future plan participants (OSFI, 2012).

Another example of the balance sheet approach that takes into account the partially funded nature of the pension system is the balance sheet of the Swedish Inkomstpension system. While this balance sheet is based on the closed group approach, its assets side includes the so-called contribution asset that accounts for future contributions as a major source of the system’s financing.

It should be noted that the recent International Actuarial Association (IAA) exposure draft, the International Standard of Actuarial Practice – Valuation of Social Security Programs, promotes the concept of choosing the balance sheet methodology based on the programme’s financing approach (IAA, 2012).

In this article, the different balance sheet methodologies are applied to the Canada Pension Plan (CPP or the “Plan”), the partially funded second pillar of the Canadian retirement income system. Before doing so, an overview of the Canadian retirement income system is presented and the financing of the Canada Pension Plan discussed. In turn, we will look at the closed and open group methodologies with the emphasis on the open group balance sheet. Finally, we will conclude by discussing the United States methodology (the actuarial balance) and the Swedish methodology (the balance sheet including the contribution asset).
Canadian retirement income system

Canada’s retirement income system is based on a diversified approach to savings in terms of the sources (private and public), coverage (mandatory and voluntary), and the financing methodology (PAYG, partial funding, or full funding). The Canadian system rests upon three pillars. Pillar one is the Old Age Security (OAS) programme that provides a universal basic benefit payable at age 65 based on citizenship and years of residence. The OAS programme is financed on a pure PAYG basis from Government tax revenues. The second mandatory pillar consists of the earnings-related Canada and Québec Pension Plans (CPP and QPP). The CPP and QPP are sister plans that both came into effect on 1 January 1966, and are deemed to be equivalent. Both plans provide a retirement pension, as well as disability, death, survivor, and children’s benefits. The CPP covers virtually all Canadians between the ages of 18 and 70 working outside of the province of Québec, while those working in Québec are covered by the QPP. Both plans are financed by employer and employees contributions, and also by investment earnings. Funds not immediately required to pay benefits are invested in the markets by the respective plans’ fund managers – the CPP Investment Board (CPPIB) and the Caisse de Dépôt et Placement du Québec. Therefore, these plans are partially funded.

In Canada, the first two pillars replace about 40 per cent of pre-retirement earnings for individuals with earnings at the average level. Such a replacement rate is consistent with the International Labour Organization’s Social Security (Minimum Standards) Convention, 1952 (No. 102).

The third pillar of the Canadian retirement income systems is a voluntary one, and consists of all private savings, which are expected to be fully funded. It includes tax-deferred employer-sponsored pension plans, as well as individual savings plans such as Registered Retirement Savings Plans and Tax-Free Savings Accounts.

The diversification of the Canadian system through its mix of public and private pensions and different levels of funding mitigates the multitude of risks to which the system and individuals’ retirement incomes are exposed. As stated by the Organisation for Economic Co-operation and Development (OECD): “Taking the long view, a diversified pension system – mixing public and private provision, and pay-as-you-go and pre-funding as sources of finances – is not only the most realistic prospect but the best policy” (OECD, 2011).

Canada Pension Plan

The CPP was initially established as a PAYG plan with a small reserve and a combined employer-employee contribution rate of 3.6 per cent of covered earnings. A
set of major changes introduced in 1998 modified, in particular, the financing approach from a PAYG basis to a hybrid of PAYG financing and full funding, called “steady-state funding”. This partial funding approach has resulted in a greater reserve of assets accumulating over time. Steady-state funding involves a steady-state contribution rate that is the lowest rate sufficient to ensure both the stabilization of the ratio of assets to the following year’s expenditures over time and the long-term financial sustainability of the Plan without recourse to further rate increases. Another change introduced in 1998 was the requirement that any benefit improvements or new benefits for the Plan be fully funded. The sum of the steady-state rate and the incremental full funding rate is called the minimum contribution rate. The current CPP legislated combined employer-employee contribution rate is 9.9 per cent of covered earnings and it has been the same over the past decade.

Actuarial valuation reports on the CPP are prepared every three years. These reports determine a minimum contribution rate and show projections of the Plan’s contributions, expenditures and assets for the next 75 years. Each actuarial valuation report is based on a number of best-estimate assumptions that reflect the best judgment of the Chief Actuary of the CPP as to future demographic and economic conditions. The latest report is the Twenty-Fifth Actuarial Report on the CPP as at 31 December 2009 (the 25th CPP Actuarial Report) (OSFI, 2010b). A set of the best-estimate assumptions of this report is presented in Table 1. The main findings in this article are based on these assumptions.

The 25th CPP Actuarial Report determined the minimum contribution rate to be 9.86 per cent, i.e. below the legislated contribution rate. Table 2 shows the projected financial status of the CPP on the basis of this report using the legislated contribution rate of 9.9 per cent. It can be seen that, although net cash flows are projected to turn negative, investment earnings will be more than sufficient to cover the shortfall. The PAYG rate, which equals the contribution rate required to fully meet expenditures in a given year (that is, the ratio of expenditures to contributory earnings) reaches a level of 10.9 per cent in 2050. The excess resulting from the legislated rate of the Plan (9.9 per cent) over the minimum contribution rate (9.86 per cent) allows for the Plan’s assets to grow more quickly (OSFI, 2010b).

Canada Pension Plan actuarial balance sheet

The key financial measure for evaluating the Canada Pension Plan is the minimum contribution rate, specifically, its adequacy and stability over time. However, a number of indicators may be used to assess the CPP’s financial sustainability. One of these measures is the actuarial balance sheet, i.e. a summary of the scheme’s
assets and liabilities. It is especially favoured by some experts involved in assessing the financial sustainability of a pension scheme (for instance, accountants and economists) due to its prevalence in fully funded pension systems. This section examines and compares the assets and obligations (liabilities) of the CPP derived using different closed and open group methodologies. Two indicators of the Plan’s financial status are presented for each methodology: the difference between the Plan’s assets and its liabilities (this difference is termed “asset excess”, if positive, or “asset shortfall”, if negative) and the Plan’s total assets as a percentage of its liabilities. Both measures provide an indication of the extent to which the Plan’s obligations are covered by its assets.

A closed group includes only current participants of a plan, with no new entrants permitted, while, an open group includes all current and future participants of a plan. Two types of closed groups are discussed in this section: a closed group without future benefit accruals for the group’s members, and a closed group with future accruals for its members.

In the past, the CPP actuarial reports presented the CPP actuarial balance sheets; however, the methodology used to produce these balance sheets evolved over time. The most significant changes took place after the Twenty-Third Actuarial Report on the Canada Pension Plan as at 31 December 2006 (the 23rd CPP

<table>
<thead>
<tr>
<th>Table 1. Best-estimate demographic and economic assumptions of the 25th CPP Actuarial Report</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Canada</strong></td>
</tr>
<tr>
<td>Total fertility rate</td>
</tr>
<tr>
<td>Mortality</td>
</tr>
<tr>
<td>Canadian life expectancy:</td>
</tr>
<tr>
<td>at birth in 2010</td>
</tr>
<tr>
<td>at age 65 in 2010</td>
</tr>
<tr>
<td>Net migration rate</td>
</tr>
<tr>
<td>Participation rate (ages 15-69)</td>
</tr>
<tr>
<td>Employment rate (ages 15-69)</td>
</tr>
<tr>
<td>Unemployment rate (ages 15+)</td>
</tr>
<tr>
<td>Rate of increase in prices</td>
</tr>
<tr>
<td>Real-wage differential</td>
</tr>
<tr>
<td>Real rate of return</td>
</tr>
</tbody>
</table>

Source: Adapted from OSFI (2010b).
<table>
<thead>
<tr>
<th>Year</th>
<th>PAYG rate (%)</th>
<th>Contribution rate (%)</th>
<th>Contributions (CAD)</th>
<th>Expenditures (CAD)</th>
<th>Net cash flow (CAD)</th>
<th>Investment income$^2$ (CAD)</th>
<th>Assets at 31 Dec.$^3$ (CAD)</th>
<th>Yield (%)</th>
<th>Asset/expenditure ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>8.65</td>
<td>9.90</td>
<td>36,862</td>
<td>32,192</td>
<td>4,670</td>
<td>2,391</td>
<td>133,897</td>
<td>1.80</td>
<td>3.94</td>
</tr>
<tr>
<td>2020</td>
<td>9.83</td>
<td>9.90</td>
<td>55,983</td>
<td>55,608</td>
<td>375</td>
<td>16,646</td>
<td>275,099</td>
<td>6.31</td>
<td>4.68</td>
</tr>
<tr>
<td>2030</td>
<td>10.78</td>
<td>9.90</td>
<td>85,193</td>
<td>92,803</td>
<td>-7,610</td>
<td>28,444</td>
<td>464,687</td>
<td>6.34</td>
<td>4.79</td>
</tr>
<tr>
<td>2040</td>
<td>10.73</td>
<td>9.90</td>
<td>130,283</td>
<td>141,263</td>
<td>-10,980</td>
<td>44,686</td>
<td>733,329</td>
<td>6.31</td>
<td>4.98</td>
</tr>
<tr>
<td>2050</td>
<td>10.94</td>
<td>9.90</td>
<td>195,398</td>
<td>215,909</td>
<td>-20,511</td>
<td>71,427</td>
<td>1,169,230</td>
<td>6.33</td>
<td>5.18</td>
</tr>
</tbody>
</table>

Notes:
1. CAD 1 is approximately equal to USD 1.
2. Investment income includes both realized and unrealized gains and losses.
3. As at 30 June 2012, the investment portfolio of the CPP totalled CAD 165.8 billion, and 60 per cent of the portfolio was invested outside of Canada.

Source: Adapted from OSFI (2010b).
Actuarial Report). The 23rd CPP Actuarial Report showed the actuarial balance sheet on a closed group without future accrual basis, and, not surprisingly, the liabilities significantly exceeded the available assets (OSFI, 2007). The independent peer reviewers of this report commented on the misleading nature of the CPP closed group. They noted that the closed group balance sheet is “appropriate for systems that have full funding as an objective. However, the CPP’s primary funding objective is steady-state funding, with full funding only for benefit increases. [...] The inclusion of items such as a balance sheet suggest to the casual reader that the CPP’s funding is hopelessly inadequate, which is simply not true. The meaningful measure is the adequacy and sustainability of the 9.9 per cent contribution rate, not the funded percentage” (Campbell, Flanagan and Levy, 2008).

As a result, the Office of the Chief Actuary prepared Actuarial Study No. 8, where the application of the open group balance sheet to the CPP was explored (OSFI, 2010a). Following the publication of this study, the Plan’s balance sheet based on an open group methodology was included in the 25th CPP Actuarial Report (OSFI, 2010b), as well as the Annual Report of the Canada Pension Plan 2010-2011 (HRSDC, 2011) and the Public Accounts of Canada 2010-2011 (Government of Canada, 2011).

Notes on the methodology

For all balance sheets discussed in this article, it is assumed that future contributions are determined using the current legislated contribution rate of 9.9 per cent of covered earnings. For the purpose of determining actuarial liabilities and assets, the future inflows and outflows are discounted using the assumed nominal rate of return on the invested CPP assets, unless otherwise specified. This nominal rate is determined to be 6.3 per cent ultimately (4.0 per cent real) under the assumption that the assets of the Plan are invested in the best-estimate portfolio of the 25th CPP Actuarial Report (ultimately 42 per cent equity, 40 per cent fixed income securities, and 18 per cent inflation-sensitive assets, such as real estate and infrastructure) (OSFI, 2010b). It could be argued that it is not appropriate to determine the Plan’s liabilities and assets using the expected rate of return on invested assets, since the CPP is a partially funded plan largely financed by future contributions. The use of alternative discount rates based on the growth of the contributory base is addressed later in this article.

Another important element of the methodology is the length of the projection period. The Canada Pension Plan legislation specifies that the CPP actuarial report should present financial information for at least a 75-year period following the valuation date. For a closed group with or without future accruals, the balance sheet based on the projection of future contributions and expenditures for a
A 75-year period presents an accurate picture since it includes all future contributions and expenditures associated with the group’s participants. Alternatively, the use of a 75-year projection period for the open group balance sheet could be viewed as insufficient. Limiting the projection period to 75 years excludes from the liabilities part of the future expenditures for cohorts that will enter the labour force during the projection period; however, most of the contributions for these cohorts are included in the assets. This imbalance could lead to a perceived better financial position. The United States 2010 Annual Report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Federal Disability Insurance Trust Funds (the Trustees’ Report) states that “the overemphasis on summary measures (such as the actuarial balance and open group unfunded obligation)\(^1\) for the 75-year period can lead to incorrect perceptions and policies that fail to address financial sustainability for the more distant future” (OASDI, 2010). Thus, in this article the cash flows are projected over an extended time period of 150 years. It should be noted however that, while enhancing the assessment of the financial sustainability, increasing the length of the projection period also increases the uncertainty of projections.

For all balance sheet approaches examined in this article it is assumed that for beneficiaries, benefits received are increased annually in line with inflation only, as provided under the current Plan. In addition, projected expenditures include administrative expenses.

### Closed group without future accruals

“Closed group without future accruals” methodology assumes that no new entrants to the Plan are permitted. Furthermore, current plan participants who are not receiving benefits at the valuation date are assumed to make no further contributions beyond that date, and hence accrue no future benefits.

Contributors’ projected benefits are calculated by increasing pensionable earnings prior to the valuation date in line with wages and by assuming zero future pensionable earnings. The resulting projected expenditures are then discounted to determine their present value. This is the actuarial liability of the Plan under the closed group without future benefit accruals approach. The assets under this approach consist of the Plan’s current assets.

Under this methodology, the Plan’s asset shortfall is CAD 748 billion with the total assets representing 14.5 per cent of the actuarial liabilities as at 31 December 2009, as shown in Table 3.

---

\(^1\) The term “unfunded obligation” refers to the difference between plan’s obligations and assets, i.e. this is the opposite of the asset excess or shortfall.
As a result of the 1998 CPP amendments, the CPP has been moving away from pure PAYG financing (with a small contingency reserve) towards partial funding. At the same time, full funding of the Plan was never intended by the stakeholders. The balance sheet under the closed group methodology does not reflect the nature of the partial financing approach where future contributions represent a major source of financing of future expenditures. Therefore, it is inappropriate to reach a conclusion regarding the Plan’s long-term financial sustainability considering the asset shortfall or excess under the closed group balance sheet.

**Closed group with future accruals**

For this analysis, it is assumed that the CPP remains open to current participants, but with no new entrants. Thus, current Plan contributors continue to contribute and to accrue benefits with future salary increases in line with wage increases.

As a result, the obligations side of the balance sheet includes the present value of future expenditures for current Plan participants, while the assets side includes the present value of their future contributions in addition to the CPP’s current

### Table 3. Balance sheet as at 31 December 2009 for the CPP: Groups with and without future benefit accruals – Comparison of methodologies (9.9 per cent contribution rate, nominal discount rate 6.3 per cent)

<table>
<thead>
<tr>
<th>As at 31 December (in CAD billion)</th>
<th>Methodology</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Excluding future benefit accruals</td>
</tr>
<tr>
<td></td>
<td>Closed group</td>
</tr>
<tr>
<td><strong>Assets</strong></td>
<td></td>
</tr>
<tr>
<td>Current assets</td>
<td>127</td>
</tr>
<tr>
<td>Present value of future contributions</td>
<td>–</td>
</tr>
<tr>
<td>Total assets (a)</td>
<td>127</td>
</tr>
<tr>
<td><strong>Liabilities (present value)</strong></td>
<td></td>
</tr>
<tr>
<td>Current benefits</td>
<td>308</td>
</tr>
<tr>
<td>Future benefits</td>
<td>567</td>
</tr>
<tr>
<td>Total liabilities (b)</td>
<td>875</td>
</tr>
<tr>
<td>Asset excess (Shortfall) (a) – (b)</td>
<td>–748</td>
</tr>
<tr>
<td>Total assets as a percentage of total liabilities (%) (a)/(b)</td>
<td>14.5%</td>
</tr>
</tbody>
</table>

*Source: Adapted from OSFI (2012).*
assets. The asset shortfall under this methodology as at 31 December 2009 is CAD 509 billion. This balance sheet is presented in Table 3 together, for comparison, with the closed group without future accruals methodology and the open group methodology, discussed next.

Once again, this methodology is not consistent with the partial funding financing approach since it only takes into account assets and contributions of current contributors as a source of financing for future expenditures.

Open group

An open group includes all current and future participants of a plan, where the plan is considered to be ongoing into the future, that is, over an extended time horizon. This means that future contributions of current and new participants and their associated benefits are included in order to determine whether current assets and future contributions will be sufficient to pay for all future expenditures.

The actuarial liability of the Plan under the open group approach is equal to the present value of future projected expenditures with respect to the current and future Plan participants. The assets of the Plan under this approach are the sum of the present value of future projected contributions of current and future contributors and the Plan’s current assets. The asset shortfall under the open group methodology as at 31 December 2009 is CAD 7 billion and the total assets covers 99.7 per cent of the actuarial liabilities.

In our opinion, if the Plan’s financial sustainability is to be measured based on its asset excess or shortfall, it should be done on an open group basis. The independent peer review panel concurred with this statement in the 25th CPP Actuarial Report: “The Plan is intended to be long-term and enduring in nature, a fact that is reinforced by the federal, provincial and territorial governments’ stewardship through the established strong governance and accountability framework of the Plan. Thus, an open group valuation that emphasizes the long-term nature of the Plan could be deemed to be the most appropriate” (Andrews, Brown, and McGillivray, 2011). The inclusion of future contributions and benefits with respect to both current and future participants in the assessment of the CPP’s financial status confirms that the Plan is able to meet its financial obligations and is sustainable over the long term.

Table 3 shows that the closed group without future accrual approach results in the largest asset shortfall. This is because there are no future contributions to the Plan as well as no future accruals. Under the best-estimate assumptions, the present value of future contributions, either for only current or for both current and future participants, exceeds the present value of the associated future benefits. As a result, the asset shortfall decreases when future accruals are taken into account. At the same time, a substantial shortfall still exists under the closed group
with accrual, since this approach does not fully account for future contributions as a major source of financing of the Plan.

Further discussion of the open group approach – modified open group balance sheet

The partial funding of the CPP may be examined in greater detail by viewing the open group balance sheet in a modified form. The idea is to show the PAYG and funded components of the Plan, with their respective assets and liabilities, separately. The resulting modified balance sheet emphasizes the hybrid nature of partial funding and allows for a better understanding of how future expenditures are financed. Under the partial funding, part of a current year’s expenditures is financed from the same year’s contributions, forming the PAYG component of the Plan. The remaining expenditures, if any, are covered using the underlying pension fund from the funded component of the Plan. Although there is a funded component to steady-state funding, its goal is not to fully fund the CPP. Rather, by stabilizing the asset/expenditure ratio, steady-state funding ensures that the Plan’s contributions remain the primary source for covering the Plan’s expenditures.

The allocation of the projected contributions and expenditures to the PAYG and the funded components of the CPP is shown in Table 4. By definition, the PAYG component’s contributions and expenditures are exactly the same every year. Contributions for the funded component exist as long as the current year’s contributions exceed the same year’s expenditures. These excess contributions are added to

<table>
<thead>
<tr>
<th>Year</th>
<th>PAYG component</th>
<th>Funded component</th>
<th>Total*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Contributions</td>
<td>Expenditures</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(a)</td>
<td>(b)</td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td>32.2</td>
<td>32.2</td>
<td>36.9</td>
</tr>
<tr>
<td></td>
<td>4.7</td>
<td>0.0</td>
<td>56.0</td>
</tr>
<tr>
<td>2015</td>
<td>42.8</td>
<td>42.8</td>
<td>45.1</td>
</tr>
<tr>
<td></td>
<td>2.3</td>
<td>0.0</td>
<td>92.8</td>
</tr>
<tr>
<td>2020</td>
<td>55.6</td>
<td>55.6</td>
<td>56.0</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>0.0</td>
<td>101.3</td>
</tr>
<tr>
<td>2030</td>
<td>85.2</td>
<td>85.2</td>
<td>85.2</td>
</tr>
<tr>
<td></td>
<td>0.0</td>
<td>7.6</td>
<td>152.8</td>
</tr>
<tr>
<td>2040</td>
<td>130.3</td>
<td>130.3</td>
<td>130.3</td>
</tr>
<tr>
<td></td>
<td>0.0</td>
<td>11.0</td>
<td>241.3</td>
</tr>
<tr>
<td>2050</td>
<td>195.4</td>
<td>195.4</td>
<td>195.4</td>
</tr>
<tr>
<td></td>
<td>0.0</td>
<td>20.5</td>
<td>215.9</td>
</tr>
</tbody>
</table>

Note: * As shown in Table 2.
Source: Adapted from OSFI (2012).
the Plan’s assets, which are invested by the CPPIB. Since the 25th CPP Actuarial Report projects that contributions will exceed expenditures up to and including the year 2020, the funded component’s contributions exist during this period. Starting in 2021, the expenditures are then projected to be higher than contributions (OSFI, 2010b). These excess expenditures are in turn allocated to the funded component of the Plan, and are financed by the invested assets.

The open group balance sheet then can be modified as follows:

- On the assets side, the present value of future contributions is broken down into two parts: the present value of future contributions that cover future expenditures (PAYG component future contributions shown in column (a) of Table 4) and the present value of future contributions in excess of future expenditures, which are invested (funded component future contributions shown in column (c) of Table 4);
- On the liabilities side of the balance sheet, the present value of future expenditures is similarly broken down into the present value of future expenditures covered by future contributions (PAYG component future expenditures shown in column (b) of Table 4) and the present value of future expenditures not covered by future contributions and therefore financed by the invested assets (funded component future expenditures shown in column (d) of Table 4).

As the second step, the PAYG and funded components are separated in the open group balance sheet. Figure 1 illustrates the structure of the open group modified balance sheet.

It can be seen clearly that no asset excess or shortfall ever exists with respect to the PAYG component. Under PAYG financing, while both the present values of the contributions and expenditures could vary depending on past experience and future actuarial assumptions, they will always remain equal. Under the funded component, an asset shortfall results when the funded component’s total assets, consisting of the invested assets and the present value of future contributions in excess of future expenditures, are not sufficient to pay the future expenditures not covered by future contributions. In the case of the CPP, the asset shortfall is CAD 7 billion (under the best-estimate scenario).

Breaking down the open group balance sheet into PAYG and funded components raises questions regarding which discount rate should be used to determine present values. So far in this article, all present values were determined using a discount rate equal to the expected nominal rate of return on the invested CPP assets. A similar methodology is used for the Trustees’ Reports on the OASDI programme in the United States, where the open group balance sheet entries are determined by discounting the programme’s future contributions and expenditures using the effective yield on the trust fund assets (OASDI, 2010). However, in a partially funded plan, the cash flows of the PAYG component are not invested and depend on demographic and economic factors other than market returns. Therefore, it
could be argued that the use of the expected return on the invested CPP assets as a discount rate for the PAYG component is not necessarily appropriate.

It is desirable for the discount rate used to determine the present values of future cash flows of a pension system to be connected to the growth in the system’s financing base. Mikula and Settergren (2005) suggest that for a pure PAYG system financed by contributions only, such a discount rate is equal to the growth in the contributory base. The financing base of a partially funded system, such as the CPP, has two components: future contributions (contributory base) and invested assets of the system. As such, discounting the cash flows of the PAYG component using the growth in the contributory base and discounting the cash flows of the funded component using the expected return on the invested CPP assets represent a logical choice.

The nominal growth in the contributory base is a combination of three components: growth in the real-wage differential, inflation and growth in the number of contributors. Under the best-estimate assumptions of the 25th CPP Actuarial Report, the ultimate nominal growth in the contributory base is 4.1 per cent, or 1.8 per cent in real terms (OSFI, 2010b). Table 5 presents a modified open group balance sheet as at 31 December 2009 with the PAYG component’s cash flows discounted either with the expected rate of return on the CPP assets (totals as presented in Table 3) or with the growth in the contributory base. The funded
### Table 5. Modified open group balance sheet (9.9 per cent contribution rate)

<table>
<thead>
<tr>
<th>Present value (PV) as at 31 December 2009 (in CAD billion)</th>
<th>Discount rate for PAYG component</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rate of return on the CPP assets (6.3%)</td>
</tr>
<tr>
<td><strong>PAYG component</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Assets = Liabilities</strong></td>
<td></td>
</tr>
<tr>
<td>PV of future contributions that cover future expenditures</td>
<td>1,836</td>
</tr>
<tr>
<td>PV of future expenditures covered by future contributions (a)</td>
<td></td>
</tr>
<tr>
<td><strong>Funded component</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Assets</strong></td>
<td></td>
</tr>
<tr>
<td>PV of future contributions in excess of future expenditures</td>
<td>25</td>
</tr>
<tr>
<td>Current assets</td>
<td>127</td>
</tr>
<tr>
<td>Total assets for funded component (b)</td>
<td>152</td>
</tr>
<tr>
<td><strong>Liabilities</strong></td>
<td></td>
</tr>
<tr>
<td>PV of future expenditures not covered by future contributions (c)</td>
<td>159</td>
</tr>
<tr>
<td>Asset excess (shortfall) with respect to funded component (d) = (b) − (c)</td>
<td>−7</td>
</tr>
<tr>
<td><strong>Total for Plan</strong></td>
<td></td>
</tr>
<tr>
<td>Total assets (e) = (a) + (b)</td>
<td>1,988</td>
</tr>
<tr>
<td>Total liabilities (f) = (a) + (c)</td>
<td>1,995</td>
</tr>
<tr>
<td>Total asset excess (shortfall) (g) = (e) − (f)</td>
<td>−7</td>
</tr>
<tr>
<td>Total assets as a percentage of total liabilities (h) = (e)/(f)</td>
<td>99.7%</td>
</tr>
</tbody>
</table>

Component liabilities as a percentage of total liabilities:

- PAYG (a)/(f) 92% 97%
- Funded (c)/(f) 8% 3%

**Note:** 1. Discounted at the rate of return on the CPP assets. Discount rates for PAYG component: Growth in contributory base and expected rate of return on CPP assets. Discount rate for funded component: Expected rate of return on CPP assets.

**Source:** Adapted from OSFI (2012).
component’s cash flows are always discounted using the rate of return on the invested CPP assets (ultimate real rate of return of 4.0 per cent or nominal rate of return of 6.3 per cent).

Since the rate of growth in the contributory base is lower than the assumed rate of return on the CPP assets, the PAYG component assets and liabilities are much higher than if the rate of growth in the contributory base were used as the discount rate. However, since the Plan’s asset excess or shortfall is generated only by the funded component, it remains the same regardless of which discount rate is used for the PAYG component. Therefore, although the total Plan’s liabilities increase significantly from CAD 2 trillion to CAD 5.7 trillion, a similar increase is observed on the assets side of the balance sheet. Thus, the asset shortfall remains at CAD 7 billion.

The decomposition of the CPP into the PAYG and funded components provides insights into the nature of the risks to which the CPP is exposed. Both the PAYG and funded components are subject to demographic and economic risks. The PAYG component, however, is not exposed to financial market risk since the associated cash flows are not invested. This is in contrast to the funded component, which is subject to financial market risk since its assets are invested. It is important for this additional risk to the funded component, and hence to the Plan as a whole, to be taken into account in considering both the short-term and long-term financial sustainability of the Plan.

While the funded component may appear to be small (its share of the total obligations is shown in Table 5 to be between 3 per cent and 8 per cent, depending on the discount rate scenario), the impact of the Plan’s exposure to the financial market risk on the stability of the CPP contribution rate should not be underestimated. Both short-term and long-term negative market experiences could result in an increase in the minimum contribution rate above the legislated rate of 9.9 per cent, as illustrated in the “Uncertainty of Results” section of the 25th CPP Actuarial Report (OSFI, 2010b).

Balance sheet measures used in the United States and Sweden

This section discusses balance sheet measures used in the assessment of the long-term financial sustainability of the Old-Age, Survivors and Disability Insurance

2. The PAYG component’s contributions and benefits may become affected if CPP’s benefits and/or contributions are altered as a result of the financial market performance. It is also true for systems that possess automatic balancing mechanisms, the application of which is contingent on the value of system’s assets.
programme in United States and the Inkomstspension system in Sweden. These two national programmes have a different design; however they have a similar financing principle – both schemes are PAYG programmes with a fund. While the balance sheet measures discussed below are very different, both measures reflect the nature of the financing of these programmes. For illustration purposes, these measures are applied to the CPP. Although it is instructive to examine the financial position of the CPP using these alternative measures, the main indicator of the CPP’s financial sustainability remains the steady-state contribution rate, in particular, its adequacy and stability over time.

**United States Old-Age, Survivors and Disability Insurance (OASDI) programme**

The United States OASDI programme is an earnings-related defined benefit social security programme that provides old-age, survivor and disability benefits to insured participants. Similar to the CPP, the benefits are based on adjusted career-average earnings. The programme is financed mainly on a PAYG basis with trust funds aimed at covering short-term fluctuations.

The Annual Reports of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Federal Disability Insurance Trust Funds in the United States analyze the financial sustainability of the OASDI programme using several balance sheet measures based on the open group approach; that is, projections of incomes, expenditures and contributory earnings take into account current and future OASDI programme participants. A primary measure used in the Trustees’ Reports is called the actuarial balance. In addition, these reports present the programme’s open group unfunded obligation using the 75-year and infinite projection periods.

The actuarial balance is the difference between present values of annual incomes and expenditures over a selected projection period expressed as a percentage of the present value of contributory earnings over the same period. An actuarial balance of zero for any period indicates that the estimated cost for the period is met, on average, with assets remaining at the end of the period equal to the following year’s expenditures. On the contrary, a negative actuarial balance indicates that, over the period, the income to the programme combined with existing assets is not sufficient to cover the programme expenditures.

The long-term actuarial balance in the Trustees’ Reports is determined over a 75-year period, and the present values used to determine the components of the actuarial balance are based on the effective yield on the trust funds assets. The ultimate assumed nominal effective yield as of 31 December 2009 is 5.7 per cent (2.9 per cent real) (OASDI, 2010). This discount rate is based on the assumptions that the trust funds assets are invested in United States Treasury securities.
Table 6. Components of 75-year actuarial balance (2010-2084) best-estimate assumptions (9.9% contribution rate for the CPP)

<table>
<thead>
<tr>
<th>Item</th>
<th>OASDI (in $ billions)</th>
<th>CPP (in $ billions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contribution revenue (income) (a)</td>
<td>40,118</td>
<td>1,533</td>
</tr>
<tr>
<td>Plan expenditures (b)</td>
<td>48,065</td>
<td>1,609</td>
</tr>
<tr>
<td>Income minus plan expenditures (c) = (a) – (b)</td>
<td>-7,947</td>
<td>-76</td>
</tr>
<tr>
<td>Fund assets at start of period (d)</td>
<td>2,540</td>
<td>127</td>
</tr>
<tr>
<td>Open group asset excess (shortfall) (e) = (d) + (c)</td>
<td>-5,407</td>
<td>51</td>
</tr>
<tr>
<td>Ending fund target (f)</td>
<td>441</td>
<td>10</td>
</tr>
<tr>
<td>Open group asset excess (shortfall) minus ending fund target (g) = (e) – (f)</td>
<td>-5,848</td>
<td>41</td>
</tr>
<tr>
<td>Contributory earnings (h)</td>
<td>304,530</td>
<td>14,978</td>
</tr>
</tbody>
</table>

Percent of contributory earnings

| Actuarial balance (100 x g ÷ h)                        | -1.92%³                | 0.27%               |

Notes:
1. OASDI (2010), p. 63, Table IV.B5.
2. CAD for the CPP and USD for the OASDI.
3. In the most recent 2012 Trustees’ Report, the 75-year actuarial balance for the period 2012-2086 under the intermediate assumptions is determined to be -2.67 (OASDI, 2012). 

Source: Adapted from OSFI (2012).

Table 6 presents the actuarial balance of the OASDI programme over a 75-year horizon as at 31 December 2009 and compares it with the actuarial balance of the CPP. The CPP figures are calculated using the expected rate of return on the CPP assets invested in the diversified portfolio.

The negative actuarial balance of the OASDI programme is an indicator that the programme in its current form is not financially sustainable over the long term. At the same time, applying the actuarial balance measure to the CPP confirms its financial sustainability under the current legislation and best-estimate assumptions.

**Inkomstpension system in Sweden**

The Swedish Inkomstpension system is an earnings-related notional defined contribution (NDC) scheme that maintains individual notional accounts for the scheme’s participants. These accounts increase every year with new contributions and are indexed using the income index determined on the basis of the growth in average income. At retirement, the accumulated individual notional account is converted into an annuity.
The Inkomstpension is a PAYG social security scheme with a buffer fund where pension contributions received are used to pay pension disbursements for the same year. Differences between contributions received and pensions paid are transferred to the buffer fund.

The financial sustainability of the Inkomstpension system is assessed through the means of a modified closed group without future accruals balance sheet. Unlike the traditional closed group without future accruals balance sheet discussed earlier, the modified one includes a “contribution asset” on the assets side in addition to the assets held in the buffer fund. As discussed, pure PAYG as well as partially funded systems by design do not possess a fund sufficient to cover the system’s liabilities. The concept of the contribution asset developed in Sweden recognizes that a PAYG or partially funded system does not have any legal requirement to hold assets to fully guarantee its liabilities. Since such a system relies on contributions as a major source of its financing, it implies that the flow of future contributions represents an asset for the system (OSFI, 2012).

The contribution asset is determined as the product of the most recent year’s contributions to the system and the average time during which the obligations generated by these contributions are expected to remain in the system. This time is called the “turnover duration”. Figure 2 illustrates this concept. The white area on the left-hand side of the figure corresponds to the distribution of contributions for a given year by age of contributors. The black area on the right-hand side of the figure corresponds to the distribution of benefits for the same year by age of beneficiaries. The turnover duration is the difference between the benefit-weighted average age of beneficiaries (\(A_r\)) and the contributions-weighted average age of contributors (\(A_c\)). It is based on the profile of the system’s participants at the time

![Illustration of turnover duration (TD) concept](image-url)
of the valuation. It could also be interpreted as the sum of the expected pay-in and pay-out durations.

The ratio of the assets to the liabilities is called the balance ratio, and this balance ratio serves as a trigger for the application of an automatic balance mechanism. If this ratio is less than one, the automatic balance mechanism is activated, whereby the indexation of accounts and pensions in pay is reduced in order to restore the financial sustainability of the system. When the balance ratio reverses to values higher than one, the automatic balance mechanism allows for indexation at a rate higher than the change in the income index. This period of higher indexation lasts until the cumulative indexation reaches the level where it would have been if the automatic balancing had never been applied.

For more details on the Swedish system accounting and the operation of the automatic balance mechanism, the reader is referred to Settergren (2001), Mikula and Settergren (2005), Settergren (2009b), and the Swedish Pension Agency (2010).3

Table 7 presents the balance sheet for the Inkomstpension as at 31 December 2009, as well the CPP modified closed group without future accruals balance sheet that takes into account the contribution asset. At the end of the 2009, the financial position of the Inkomstpension system was close to 96 per cent.

It should be noted that the Swedish methodology of assessing the financial sustainability of the Inkomstpension system cannot be applied automatically to the CPP. The main reason is that there are major differences in the designs of the two plans, the Inkomstpension being an NDC system, and the CPP being a defined benefit plan. The nature of an NDC system, the indexation provisions of the Inkomstpension system, as well as the automatic balance mechanism, minimize the need for actuarial projections and, therefore, minimize the number of assumptions that should be made for the future. This is not the case for defined benefit plans, including the CPP, where the assessment of the Plan’s financial sustainability is based on extensive actuarial projections. At the same time, it is interesting to examine the Swedish balance sheet applied to the CPP, since the introduction of the contribution asset on the assets side of the closed group without future accruals balance sheet takes into account the partially funded nature of the CPP.

The calculations for the CPP are performed using the best-estimate assumptions of the 25th CPP Actuarial Report with the exception of the discount rate. Since the calculation of the contribution asset is based on future contributions, a discount rate equal to the growth in the contributory base appears to be the most appropriate choice. In attempting to apply the notion of a contribution asset to a

---

3. The "Orange Report" of the Swedish Pension Agency is published annually and provides detailed information on the financial status of the Inkomstpension.
partially funded defined benefit plan’s balance sheet, several simplifications were made. For more details on the underlying assumptions the reader is referred to Actuarial Study No. 10 (OSFI, 2012).

Table 7 shows that the current CPP assets together with the contribution asset are sufficient to finance the accrued liabilities of the CPP, confirming that the CPP is financially sustainable.

Conclusion

PAYG and partially funded social security retirement income systems represent social contracts where, in any given year, current contributors allow the use of their contributions to pay current beneficiaries’ benefits. Such social contracts create claims for current and past contributors to contributions of future contributors. The proper assessment of the financial sustainability of a social security PAYG or partially funded system by means of its balance sheet should take into...
account these claims. The traditional closed group methodologies do not reflect these claims since only current participants are considered. On the contrary, the open group approach accounts explicitly for these claims by considering the benefits and contributions of both the current and future plan’s participants (OSFI, 2012).

Countries, such as the United States and Sweden, assess the financial status of their old-age social security programmes using balance sheet measures that reflect the largely PAYG nature of these systems.

In the case of Canada, major amendments introduced in 1998 led to the change in financing of the Canada Pension Plan from a PAYG basis to a form of partial funding called steady-state funding. These changes, and particularly steady-state funding, restored the Plan’s financial sustainability for current and future generations. The purpose of the steady-state financing methodology is to produce an asset/expenditure ratio that is relatively stable over time. Although a number of approaches may be used to assess the Plan’s financial status, the key financial measure for evaluating the Plan is the steady-state contribution rate, in particular, its adequacy and stability over time.

From its inception, the CPP was never intended to be a fully funded plan. Therefore, it is inappropriate to assess the Plan’s long-term financial sustainability considering the asset shortfall or excess under the closed group balance sheet that does not reflect the nature of the partial financing approach.

The Canada Pension Plan is a nationwide social security programme under the stewardship of the federal, provincial and territorial governments, and with a strong governance and accountability framework. As such it is unlikely that the Plan would become insolvent. Therefore, if the Plan’s financial sustainability is to be measured based on its asset excess or shortfall, it should be done on an open group basis that reflects the long-term nature of the Plan as well as the partially funded approach to the Plan’s financing – that is, its reliance on both future contributions and invested assets as means of financing future expenditures. The inclusion of future contributions and benefits with respect to both current and future contributors in the assessment of the CPP’s financial status shows that the Plan is able to meet its financial obligations and is sustainable over the long-term.

Bibliography


Actuarial balance sheets as a tool to assess the sustainability of social security pension systems


Benefit dependency: The pros and cons of using “caseload” data for national and international comparisons

Johan De Deken and Jochen Clasen

University of Amsterdam, Netherlands; University of Edinburgh, United Kingdom

Abstract  Policy-makers in advanced welfare states have increasingly expressed concerns over large numbers of working-age people claiming social security support. Accordingly, policies aimed at reducing the level of “benefit dependency” have gained prominence. However, such policies rest on shaky empirical evidence. Systematic collections of national “caseload” data are rare, social security programmes overlap and administrative categories vary over time. The internationally inconsistent treatment of national transfer programmes provides a further challenge for cross-national comparisons. This article first identifies and discusses several of these problems, and ways in which they may be addressed. It then employs administrative claimant data from six European countries as a way of illustrating trends over time and across countries. The underlying aim is to explore the scientific potential of benefit recipient numbers as an indicator for welfare state change over time and across countries.

Keywords  social policy, social security planning, benefit administration, beneficiary, longitudinal analysis, comparison, international

Addresses for correspondence: Johan De Deken, Lecturer at the Department of Sociology, University of Amsterdam, OZ-Achterburgwal 185, 2012 DK Amsterdam, Netherlands; Email: J.J.deDeken@uva.nl. Jochen Clasen, Professor of Comparative Social Policy, School of Social and Political Science, University of Edinburgh, 15a George Square, Edinburgh EH8 9LD, Scotland, United Kingdom; Email: jochen.clasen@ed.ac.uk.

The authors thank Daniel Clegg, Irene Dingeldey, Marcel Hoogenboom and Jørgen Goul Andersen for valuable discussions on caseload data, and for the use of initial data collected for the countries featured in this article.
Introduction

One of the common characteristics of the development of modern welfare states over the past 20 years has been governments’ concerns over the number of people in receipt of benefit. Whereas in the 1980s the focus had been on containing or reducing registered unemployment, by the mid-1990s policy-makers had become increasingly focused on boosting labour force participation as a key strategy for coping with an ageing population. More recently debates have gone further and policies have shifted from “ex post” remedies that address the consequences of being out of work towards “ex ante” prevention of people from becoming dependent on a transfer income. Most current governments claim that they are pursuing policies aimed at a reduction of benefit dependency and, by implication, social security spending.

The onset of the financial and economic crisis has exacerbated such concerns. Rising unemployment and tight public budgets have reinforced perceptions of economically unsustainable levels of transfer spending. Moreover, if in the past benefits were seen as at least in part tackling material deprivation, nowadays they have come to be perceived as creating poverty traps, reinforcing social exclusion or undermining the moral foundations of the social contract. Activating the working-age population and extending the active phase in the life course have thus become central tenets of social policy reform. Against this context it comes somewhat as a surprise to find that countries rarely collect data on the development of the number of benefit claimants (or “caseloads”) in any systematic fashion. Obviously, this makes it difficult to assess any policies aimed at reducing “benefit dependency”. Equally, it makes systematic research based on claimant data challenging both within and, even more so, across countries.

There are a number of reasons for the difficulty of producing comprehensive claimant data across programmes and over time, some of which we will discuss below. Such difficulties may also be the principal reason for the absence of theoretically-guided empirical social research based on benefit recipient numbers. However, the recourse to the more conventional “dependent variables” such as social rights or social spending is faced with serious conceptual and methodological problems too (see Clasen and Siegel, 2007). Thus, it seems to us worthwhile to at least explore the options which existing and potential claimant data may offer comparative social security research. Concretely, building on preliminary considerations (De Deken and Clasen, 2011) we assess to which extent caseload trends may function as informative indicators of social policy change, possibly complementing social spending or social rights.

Prior to addressing this key concern, in the next sections we discuss methodological challenges faced by research based on national data and international comparisons of recipient numbers, before proposing a categorization of working-
age benefits. Making use of administrative claimant data for a collaborative project (Clasen and Clegg, 2011), we then explore caseload developments in six countries (Belgium, the Netherlands, Denmark, Germany, France and the United Kingdom) that have followed different pathways, thereby allowing us to identify further problems of categorization and measurement. We juxtapose caseload trends and the development of unemployment and social spending respectively. The aim is to illustrate functional equivalence and “communicating vessels” between different working-age benefit programmes, and particularly the persistence of work incapacity and early retirement schemes as reflecting labour market developments. Finally we discuss a series of policy innovations which illustrate recent reconfigurations of benefit systems, including leave schemes, the integration of unemployment assistance and social assistance, and the combination of social transfers and earnings. As we show, all the latter pose considerable challenges for cross-national research based on caseload data, while at the same time highlighting the political and administrative nature of the construction of benefit dependency.

**National claimant data – some methodological challenges**

One reason for the paucity of comprehensive national data on claimant numbers is certainly the complexity of national social security arrangements, with distinctive benefit programmes at times catering for the same needs and similar social groups. Some governments have begun to address this problem by combining what are functionally similar systems, as in Germany for example, where the so-called *Arbeitslosengeld II* (unemployment benefit – UB II) has been created as a basic income security scheme for all employable persons outside of unemployment insurance (Clasen and Goerne, 2011). The United Kingdom is in the process of merging several programmes within the forthcoming Universal Credit, which is heralded as a “single working-age benefit system”, even though this is to some extent a misnomer since parallel programmes, such as contributory unemployment and disability transfers, will remain in place. Thus, even in these two countries persons out of work will continue to receive income transfers from a range of schemes that differ significantly in terms of generosity and conditionality, but often overlap in terms of risks covered.

National programmes also differ in the ways in which respective claimant numbers are documented. For example, for most countries it is relatively straightforward to collect claimant data on unemployment insurance benefits in a systematic fashion. For other schemes, particularly incapacity but also early retirement, the availability of administrative data is more problematic, partly due to the existence of parallel systems (e.g. early retirement options within unemployment as well as pension programmes; incapacity benefits within social assistance schemes;
short-term incapacity benefits in the form of statutory sickness pay) and the termination of some and introduction of what are often merely slightly different programmes.

Systematically collecting comparable social assistance claimant data is particularly challenging for reasons such as divided administrative and financial responsibilities. In Spain, for example, social assistance (*Renta mínima*) is administered by the autonomous regions and recipient numbers cannot be found in national statistics. Moreover, the “target population” of social assistance can be very broad (“general” social assistance) or narrow (e.g. particular social assistance schemes for older people, lone parents, immigrants, etc.). In Germany, for example, social assistance used to be an encompassing scheme for persons in need irrespective of citizenship, age or labour market status. Since the early 1990s separate social assistance (or basic security) programmes have emerged for asylum seekers, people of retirement age and, most recently, persons who are deemed not to be employable. A further complication is the nature of social assistance as a poverty alleviating measure, making it impossible in some countries to distinguish between beneficiaries who receive only social assistance and others who claim additional social security transfers, or to distinguish between social assistance and other means-tested (e.g. housing or unemployment) assistance transfers, some of which may supplement earned income.

National data collection is also challenged by a blurred distinction between statutory and non-statutory benefits. Early retirement schemes in the Netherlands are a case in point. Formally considered as voluntary agreements, these schemes are initiated and run by social partners at the level of industrial sectors. Such agreements are quasi mandatory and thus encompassing via the practice of administrative extension. Yet it is hard to obtain comprehensive data because these schemes are administered by numerous organizations and lack the formal status embodied in public early retirement programmes.

**International comparisons**

The first and potentially most comprehensive attempt to assemble and standardize national administrative data on caseloads across different countries dates back to the late 1990s, when the Dutch Ministry of Labour (NEI) commissioned a pioneering study that covered nine European countries, Japan and the United States for the period 1980–1997 (*Arents, Cluitmans and van der Ende, 2002*). Subsequently, the Organisation for Economic Co-operation and Development (OECD) has sought to improve the comparability of this database, and updated the data by a few years (*OECD, 2003*). For illustrative purposes in Figure 1, Panels A and B, we reproduce the OECD estimates of working-age benefit dependency in six European countries. Figure 1, Panel A, shows the evolution of total working-age benefit dependency...
**Figure 1. Working-age benefit recipients in six European countries**

Panel A. As a percentage of the population aged 15 to 64, between 1980 and 2004

Panel B. Broken down according to functional category, 2004

dependency between 1980 and 2004. Figure 1, Panel B, breaks down the total case-
load into four main categories: unemployment, work incapacity (which includes
sickness and disability benefits); old age (early retirement), and social assistance
(which includes lone parent benefits).1

Figure 1, Panel A, suggests remarkably similar trends of increasing claimant
numbers across the six countries between 1980 and the mid-1990s, followed by
modest declines in some countries. The composition of the total caseload indicates
considerable differences in the prevalence of different benefit categories. Building
upon earlier work (De Deken and Clasen, 2011) we have rearranged and extended
the time series of the NEI–OECD database for a selected group of countries up to
the year 2011. It should be pointed out that our aims are less ambitious than those
of the original NEI study, Arents, Cluitmans and van der Ende (2002) sought to
estimate a total “benefit dependency ratio”, i.e. the percentage of the population
that is dependent upon some kind of benefit, which is an endeavour confronted
with a number of problems (see below). By contrast in this article we are merely
interested in tracing caseload trends, i.e. changes in the volume and composition
of those working-age benefit schemes which can be considered as an alternative to
or substitute for income from paid work. Concretely this means that we have not
collected caseload data on survivor benefits or maternity benefits, and only exam-
inied old-age benefits in so far as they allow people to leave the labour market prior
to the statutory retirement age.

The OECD too focused primarily on the working-age population, but, as did
the NEI study, aimed to estimate a total dependency rate. Such an endeavour
invites several methodological challenges, which, in contrast to what has some-
times been suggested (see for example CESifo, 2003), have at best only partially
been resolved. Even though we are not interested in the total benefit dependency
rate, and thus avoid some of these methodological problems, in what follows we
have aimed to apply as far as possible the same principles which were adopted in
the original NEI study. Accordingly, for calculating claimant trends over time and
across countries we have identified four major challenges.

**Partial benefit receipt**

First, partial benefit receipt refers to instances when claimants are employed on a
part-time basis while concurrently receiving benefits. Partial benefits are often
prevalent in disability schemes and also exist in early retirement and in unemploy-
ment benefit schemes in some countries. In order to create greater comparability, it

1. The authors thank David Grubb of the OECD for providing the original data files and the update for
2004.
may be possible to transform partial benefits into full benefit years. In some countries, including the Netherlands, partial unemployment benefits are already expressed in full-time equivalents. For other countries where such schemes exist, including Belgium, Denmark and France, we applied a factor of 0.5 to beneficiaries of part-time benefits. Similar procedures were used for calculating the caseload of part-time early retirement schemes.

Periodic benefits

Second, periodic benefits should be, wherever possible and appropriate, expressed in benefit years. For recipients of unemployment, early retirement, disability and social assistance benefits this was done by calculating the caseload of any particular calendar month. For sickness benefits, we calculated benefit years by dividing the number of days during which a benefit was received by the maximum number of days per year for which a benefit can be paid out. This differs from country to country, i.e. 260 days in France, 312 days in Belgium and 365 days in Germany. Not limited to sickness benefits, this problem applies to all schemes in which claimants receive benefits for less than a year or where benefits are not paid for every day of the week. One problem is that the information of the number of days benefits are paid in particular countries is either missing or not available, another is that only data for a particular benchmark month might be collected or accessible. In such instances, the NEI and the OECD used either the number of beneficiaries in December of the year or “the figures of from whatever month ... available” (Arents, Cluitmans and van der Ende, 2002, p. 11). Evidently it is impossible to know whether persons who received a benefit during the benchmark month claimed benefits also for the rest of the year. Moreover, persons who received a benefit during months other than the benchmark month are not included, and the assumption might be that these two effects might somehow cancel each other out. It should be noted that this is a rather big assumption however, neglecting, for example, seasonal effects. In what follows, we chose whenever possible September rather than December figures, as seasonal effects are less likely to manifest themselves. Nevertheless, other problems of taking monthly figures as an approximation of average annual caseloads remain. For example, a reduction in average spells of unemployment periods in one country but not another might explain divergent trends in benefit expenditure and caseloads.

Double counting

Third, there is a potential problem of double counting. One person should count for no more than one (full-time equivalent) benefit claim. In practice however,
fragmented national social security administrations make it often impossible to implement this principle in countries where the simultaneous receipt of different benefits is permissible. As part of their efforts to combat benefit fraud, some governments have started to set up integrated databases for separate social security benefits in a single registry. In the future, databases such as the “datawarehouse labour market and social protection” of the Belgian KSZ–BCSS (Crossroads Bank for Social Security) system might form a valuable tool in eliminating double counts, but the recent introduction of such initiatives prevents the development of historical time series. In the NEI study (Arents, Cluitmans and van der Ende, 2002) this problem manifested itself particularly in the case of survivor benefits which are often combined with an old-age pension, and for old-age pensions which are supplemented by social assistance. As we are only interested in working-age benefits, the problem of double counts is less relevant, except for the caseloads of active labour market policies (ALMP), for example in Denmark. At some point in the 1990s participants of those programmes started to be reported separately, i.e. in addition to, rather than included in, the caseloads of unemployment and other working-age benefits (Goul Andersen, 2011).

Payments to couples

Finally, payments to couples should ideally be individualized. Primarily this problem manifests itself in the case of old-age pensions and social assistance which is typically paid on a household basis. Within the countries we discuss in this article, a breakdown of benefit receipt between singles and married couples was not available.2 As a consequence, in all our six countries the total number of persons who are dependent upon social assistance is underestimated.

Problems of categorization

As indicated, despite the increased political salience of “benefit dependency”, it is difficult to obtain reliable comprehensive time series of recipient numbers. Moreover, the quality of readily available data varies not only between countries but within the same country depending upon the branch of the social security system. At times this seems to be a consequence of the fragmented and complex nature of those schemes. For example in the NEI and OECD studies, the category of “early retirement” ignored one of the many early retirement schemes in

2. In the larger sample of 12 countries analysed in De Deken and Clasen (2011), only Sweden was seen to systematically report a breakdown of benefit receipt between singles and married couples.
Belgium, and thus significantly underestimated the caseload in this particular category.\(^3\)

In the remainder of this article we aim to illustrate shifts in the mix of caseloads of different working-age benefit programmes over time. For that purpose we distinguish between five basic types of programmes providing benefits which allow temporary or permanent exits from the labour market. The distinction is based on differences in employment-related behavioural requirements:
- unemployment: beneficiaries of unemployment insurance and unemployment assistance are expected to re-enter the labour market;
- work incapacity: beneficiaries of sickness and disability benefits are exempt from labour market participation on medical grounds;
- early retirement: beneficiaries younger than the statutory retirement age who are permanently exempt from labour market participation;
- sabbatical and leave schemes: claimants are temporarily exempt from labour market participation, allowing periods of non-remunerated activity;
- social assistance: a residual category typically including persons in need and facing problems other than, or in addition to, lack of employment.

It should be noted that these analytically relatively clear distinctions have become less explicit in recent years. For example, in countries such as Germany, social assistance programmes (for some claimants) have gradually been transformed into quasi unemployment assistance schemes. In other countries, such as the Netherlands and the United Kingdom, eligibility to disability transfers has become more employment oriented and subjected to regular “work tests”.

As discussed in the introduction, one problem for an international comparison of caseload data is the decision to include or exclude benefit programmes based on criteria which are not always easy to apply cross-nationally. For example, the NEI study and the OECD claim that they “only included social security benefits that are regulated by law . . . regardless of the way they are administrated and financed” (Arents, Cluitmans and van der Ende, 2002, p. 8). Based on industry-wide collective agreements, which are regarded as “private”, the application of this principle led to the exclusion of Dutch early retirement schemes. By contrast, in what follows below, we have included Dutch early retirement schemes since we did not consider those as “private voluntary” contracts. After all, such schemes are concluded in the shadow of the Dutch neo-corporatist system which ascribes collective agreements a status which is akin to legislation and backs such schemes up with a procedure of administrative extension (De Deken, 2013). Moreover, it could

\(^3\) In the 1980s four main early retirement options existed in Belgium, two of which were administered by the old-age pension system: the so-called “exceptional bridging pension” and the “early retirement pension”. The two others were essentially run within the unemployment insurance administration: the “conventional bridging pension” and the “statutory bridging pension” (De Deken, 2011). The OECD failed to take into account the former.
be argued that the NEI justification for the exclusion of “private” schemes is inconsistent since it does not rule out non-statutory programmes such as the Danish unemployment insurance system, for example, which in essence is voluntary. According to the NEI “not including [Danish unemployment insurance schemes] would render international comparison difficult because one important benefit category . . . would not be included. . . ” (Arents, Cluitmans and van der Ende, 2002, p. 8). We would agree with this statement, but see no reason why it should not be extended to Dutch early retirement schemes.

The creation of meaningful comparable caseload data is not only a challenge at the international level but is also hampered by figures reported by national administrations which appear to be unreliable at times. For example, published by the Dutch national statistical office (Statistics Netherlands – CBS), annual claimant numbers of early retirement benefit suggest some erratic fluctuations. These tend to be attributed to internal revisions and a break in the series in the year 2001 when the number of beneficiaries was revised from 299,000 to 371,000 claimants (as shown in Figure 2). Moreover, discrepancies between trends in claimant numbers and benefit expenditure for the respective programme seem difficult to explain. Figure 2 illustrates this by plotting total expenditure on early retirement
benefits in the Netherlands (right Y axis) to the total number of recipients of benefits (left Y axis), as reported by the CBS. The figure suggests that the benefit caseload almost tripled during a 5-year period, but that the respective benefit spending remained more or less stable.4

Of course, a discrepancy between trends in the number of recipients and total expenditure does not necessarily need to imply a measurement problem. It might be caused by other factors, most notably a change in benefit generosity (in terms of level and duration of transfers paid). While this is unlikely to have been the case here, we return to the relationship between policy change, caseload and expenditure trends later in this article.

As indicated earlier, of interest here are benefits for people of working age, and caseloads are thus expressed as a percentage of the population between 15 and 64 years of age. This can pose a problem for comparisons of claimant numbers in early retirement programmes. In most (but certainly not all) countries the statutory retirement age for men is still around age 65. Women, and in some countries also men, used to enjoy a lower statutory retirement age and some still do. In principle, whenever making cross-national comparisons it would be preferable to consider any pensioners younger than age 65 as in receipt of a form of early retirement benefit rather than the standard pension. The NEI–OECD project has made such an attempt for France (where until the 2010 reform the statutory pension system used a flexible retirement age between 60 and 65). We have been able to replicate the NEI–OECD estimates for two more recent points in time: 2008 and 2010.5 The number of French claimants younger than age 65 in receipt of a statutory pension turns out be much higher than the number of those in receipt of an early retirement pension, i.e. the préretraités and dispensés (older beneficiaries of unemployment benefits exempted from looking for employment). Moreover, whereas the latter programmes have effectively been phased out since the late 1990s, claims for statutory pensions between ages 60 and 65 have remained stable.

In other countries, however, national data make it hard to distinguish between pensioners who are aged 65 or older from those who are younger than age 65. Depicting Dutch, Belgian and French trends, Figure 3 illustrates this. It suggests that comparisons restricted to recipients of early retirement benefits only would be inappropriate as a measure of all people who have retired, are younger than age 65 and in receipt of a pension. However, a broader measure, e.g. all claimants of a

4. To some extent this discrepancy might be related to an increase in part-time early retirement (the CBS statistics do not allow the splitting of the caseload of early retirement benefits into full-time and part-time), but the scale of the deviation in trends is more likely to be related to either a measurement/reporting problem on the expenditure or the caseload side. In one of its publications, the CBS recognized that a substantial part of the early retirement plans is implemented by companies and these pension funds remained invisible (Gebraad and Pfaff, 2006, p. 2).
Benefit dependency: The pros and cons of using “caseload” data for national and international comparisons

Fig. 3. Retirement before age 65 in France compared to the Netherlands and Belgium 1980-2011


statutory pension younger than age 65, is not without problems either, since this may involve persons who have never been part of the workforce. Another challenge is the fact that in some countries, such as Belgium, the statutory retirement age for women was raised from 60 to 65 only fairly recently, i.e. in the decade after 1997. In other words, proceeding, as we do in Figure 3, runs the risk of exaggerating the relative extent of early retirement in France.

Table 1 summarizes various problems of comparing caseloads, the ways in which they might affect the magnitude of benefit recipiency, and how some measurement problems may be tackled.

**Exploring and illustrating caseload trends**

Employing updated data from a book project (Clasen and Clegg, 2011) the remainder of this article selectively illustrates caseload trends and developments. Of course, problems of comparability across countries remain (e.g. in terms of...
early retirement, as discussed above). This is not our concern here, however. Rather than aiming to produce a calculation of total caseload volumes across countries, our more modest objective is to discuss trends over time within single countries. As a first step, this serves as a platform for a reflection on conceptual and methodological aspects of comparative research based on caseload data.

Many national social security reforms over the past decade or so have been geared towards reversing earlier policies of “labour shedding” during the 1980s and early 1990s. These have introduced behavioural requirements for unemployed and other working-age benefit claimants, for example, which would encourage shifts from “welfare to work”. To some extent the success of these policies may be measured in terms of a decline of the caseload of working-age benefit schemes. However, the decrease in claimant numbers in one benefit programme (e.g. unemployment) may lead to a concomitant growth in others (e.g. disability or social assistance), particularly for some groups such as low-skilled men (Clasen et al., 2006). Such a “substitution effect” has been noted in several countries and can be gauged from some of the Figures below. In what follows we map the direction and scale of the change in benefit schemes between 1980 and 2008 in terms of number of beneficiaries.

Initially however we explore the relationship between trends in caseload and developments in social spending. In Figure 4 we plot caseload trends (beneficiaries as a percentage of the population aged 15 to 64 – left axis) and total expenditure (in USD at constant 2000 prices at the 2000 purchasing power parity – right axis) of unemployment benefit schemes in three countries that witnessed different national trajectories during that period.

<table>
<thead>
<tr>
<th>Problem</th>
<th>Possible effects</th>
<th>Possible solutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partial benefit receipt</td>
<td>Increases caseload</td>
<td>Convert part-time benefit into full-time benefit years, or apply weighting factor</td>
</tr>
<tr>
<td>Periodic benefits</td>
<td>Seasonal variations; short spells of recipiency</td>
<td>“Neutral” reference month; convert days of benefit recipiency to benefit years</td>
</tr>
<tr>
<td>Double counting</td>
<td>Increases caseload</td>
<td>Recode programmes</td>
</tr>
<tr>
<td>Household recipiency</td>
<td>Reduces caseload</td>
<td>Individualize benefits</td>
</tr>
<tr>
<td>Functional categorization</td>
<td>Underestimates recipiency in policy-relevant categories</td>
<td>Re-categorize benefit categories on the basis of behavioural requirements</td>
</tr>
<tr>
<td>Non-statutory programmes and mandating</td>
<td>Underestimates caseload in countries that use mandating instead of statutory programmes</td>
<td>Include non-statutory programmes</td>
</tr>
<tr>
<td>Boundaries of working age</td>
<td>Underestimates caseload in countries with a lower statutory retirement age</td>
<td>Matching early retirement with the category of beneficiaries of statutory pensions younger than age 65</td>
</tr>
</tbody>
</table>
Figure 4 helps to illustrate that, whatever their respective uses, expenditure and caseload data should be regarded as separate indicators since trends in one are not necessarily reflected in developments of the other. Looking only at Dutch data in Figure 4, this point does not seem to be immediately obvious. In the Netherlands the level of benefit caseload and expenditure on unemployment benefit developed in parallel and both are in line with the trend in unemployment rates. This is not the case for the other two countries, however. In Belgium the number of beneficiaries fluctuated broadly in line with the economic cycle and with changes in unemployment rates which remained well above Dutch and British rates during the 1990s and beyond. Cross-national differences in unemployment rates are also reflected in the fact that British caseload data continued to decline in the early 2000s in contrast to Dutch and Belgian claimant numbers, which increased for a few years before declining again. Indeed, during the period 2002-2005 the United Kingdom economy performed marginally better than the Dutch and even more so than the Belgian economy, but it is unlikely that it was completely shielded off from the economic cycle. Hence the lack of “responsiveness” of the caseload can be assumed to be related to changes in eligibility conditions.

For Belgium it is striking that spending continued to rise without showing any of the cyclical patterns exhibited by the development of caseloads. This suggests...
that either unemployment benefits became more generous (which is not the case, see OECD, 2011, p. 40), or that the OECD’s spending category is broader than our caseload category, i.e. that it includes beneficiaries who are officially not counted as unemployed. Indeed, there are two schemes which might have contributed to the spending boom: early retirement (which may account for the increases during the first half of the period) and the paid sabbatical schemes that started to take off during the second half of the period. The most important Belgian early retirement schemes are financed by the unemployment insurance system, and so are the career break benefits (De Deken, forthcoming). Had the caseloads of these three programmes been plotted next to the OECD aggregate spending data, the inconsistency would have largely disappeared. In other words, the discrepancy between the spending and caseload lines in Belgium is an indicator of what has been termed “risk reconfiguration” (Clasen and Clegg, 2011). In this particular case this implies a transformation and widening of unemployment benefit into a type of umbrella scheme which incorporates a range of out-of-work benefits for claimants with quite different behavioural requirements.

Turning to the United Kingdom it is noticeable that the gap between British spending and caseloads is much larger than in Belgium and the Netherlands, respectively. This suggests a considerably less generous unemployment benefit system. However, it does not seem immediately obvious why this gap narrowed so much during the second half of the period. Legislative change is an unlikely explanation. The introduction of Jobseekers Allowance (JSA) in 1996 halved the contributory benefit entitlement and thus certainly led to some decline of claimant numbers, but also of expenditure. A more plausible explanation might be the combination between low unemployment benefit rates and changes in the composition of benefit caseloads. It can be assumed that the job growth after the mid-1980s and again after 1993 benefited particularly claimants with shorter unemployment spells. Since caseloads are measured as the number of claimants in a particular month, this would mean a faster depletion of recipient numbers than total expenditure given that the latter is disproportionately determined by long-term benefit claimants. Potentially there are other reasons to do with national data reporting and consistency over time, which would have to be explored further. For the time being the British trends illustrate our point that it is not possible to simply “read off” expenditure from caseload date or vice versa.

Since countries differ substantially in the ways in which they administratively configure the risk of unemployment (see also Erlinghagen and Knuth, 2010), an exclusive focus on unemployment benefit dependency would be deceptive. In particular those unemployed who are hard to reintegrate into the labour market are often referred to a range of other out-of-work benefit programmes, which entail different behavioural requirements. The most often used alternative exit routes are
work incapacity and early retirement. In the absence of a long-term insurance type of benefit, hard-to-employ persons may also end up in social assistance schemes. What complicates things is that in some countries the separation between these three out-of-work statuses has become blurred. In Belgium early retirement and sabbatical leave systems are part of the unemployment insurance system (and hence form a considerable part of this expenditure category in the OECD’s Soxx database or Eurostat’s ESSPROS database). In Germany, unemployment assistance (UB II) can be claimed not only by those who are registered as unemployed but by those who are “employable”, even in a minor capacity, as well as those in low-paid work (Clasen and Goerne, 2011).

The category of work incapacity appears to be less marred by discrepancies between trends of caseloads and total benefit spending. In Figure 5 we have plotted those two trends in a similar way as in Figure 4 for unemployment. The two sets of lines are fairly parallel within each country, but differ cross-nationally. In the

Netherlands work incapacity beneficiary numbers followed a cyclical development reminiscent of the caseload of unemployment insurance. In the United Kingdom there has been a steady rise in caseloads, whereas in Belgium the caseload has remained relatively stable.

Following unemployment trends, the cyclical pattern in the Netherlands suggests that work incapacity benefits throughout the period might have functioned as an alternative form of (long-term) unemployment benefit. In the United Kingdom, unemployment declined steadily after 1993 but disability benefit receipt continued to rise. This might be due to a number of reasons, including work incapacity increasingly covering more than the classic types of physiological impairment (see below), as well as policy reforms which appear to have contributed to transfers from unemployment to disability support for some groups (Clasen et al., 2006). Another reason may be changes in average benefit durations, possibly helping to explain the divergence between caseload and expenditure trends in the 1990s.

In Figure 6 we provide an overview of the evolution of caseloads of different out-of-work benefits in six European countries between 1980 and 2010 at five-year intervals. We have used the categorization as explained earlier. In addition to our five basic groups (unemployment, work incapacity, early retirement, paid leave and social assistance) we have created further categories in order to depict programmes which are hard to classify, or which merit special attention. These special categories are beneficiaries of statutory retirement benefits in France who are younger than age 65, claimants of UB II in Germany and recipients of working tax credit (WTC) in the United Kingdom.

Most striking in Figure 6 is the common pattern of relatively low out-of-work benefit caseloads in 1980 and the subsequent growth. Towards the end of the period only Denmark and the United Kingdom managed to significantly reduce claimant numbers which had peaked during the mid-1990s. The Netherlands only recently succeeded in reversing a trend of rising claimant numbers; and in Belgium, Germany and France the number of working-age benefit claimants continued to rise or at best stagnated well into 2011. However, if for Germany one excludes those recipients of UB II who are working (the so-called Aufstocker), that country too would show a significant decline in claimant numbers for 2010. Alternatively, if for the United Kingdom one includes recipients of the working tax

8. Hence it is irrelevant out of which national social security branch early retirement, for example, is financed or whether incapacity benefit in a particular country is manifest in sickness, disability or social assistance schemes (as long as benefits are granted on a medical basis).

9. Ideally one should take account of the exact scale of the relevance of UB II benefit in the total income package for Aufstocker – just as we control for the work intensity of those on part-time unemployment benefits, part-time early retirement benefits or temporary unemployment schemes. Unfortunately the data do not allow this.
credit (which to some extent can be regarded as equivalent to \textit{Aufstocker} in Germany), the British benefit recipiency rate becomes much higher, following a steep recent upward trend. This example illustrates problems of comparability and functional equivalents. Almost a third of recipients of unemployment assistance (UB II) in Germany are actually in paid work, and consequently these claimants are part of the working-age benefit caseload. Elsewhere those with similar subsidized types of (generally low-paid and/or part-time) jobs might receive tax-financed wage subsidies (tax credits) and are thus, as in the United Kingdom, excluded from the caseload claimant count (see also below). However, for some countries our caseload claimant counts include part-time unemployment benefits and time credit schemes, which in effect can also be considered as a form of subsidizing part-time work and as such fulfil the functions of working tax credits.

A cursory analysis of the benefit mix in our six countries suggests some degree of substitution between (low) unemployment caseloads and (comparatively high) work incapacity caseloads: countries with lower caseloads in the unemployment category tend to be faced with high claimant rates in the work incapacity category.

\textbf{Figure 6.} Changes in the caseload mix in six European countries, between 1980 and 2010

\begin{figure}
\centering
\includegraphics[width=\textwidth]{caseload_mix}
\caption{Changes in the caseload mix in six European countries, between 1980 and 2010}
\end{figure}

\textit{Note:} During the first years that ALMP beneficiaries were reported as a separate category, there probably is a problem of partial double counts between the unemployment category and the ALMP category in Denmark, which is why for 2005 and 2010 ALMP beneficiaries in those countries are depicted as a separate category. For Belgium, work incapacity data refer to 2009 instead of 2010. For France, unemployment and early retirement data refer to 1981 instead of 1980; work incapacity refers to 2004 instead of 2005, and 2006 instead of 2010; and statutory pensions for those younger than age 65 refer to 2008 instead of 2005.
(Netherlands and Denmark), and the successes in reducing the caseload of unemployment benefits concur with increasing work incapacity caseloads (the United Kingdom).

This mechanism of “communicating vessels” also becomes evident from a country-by-country perspective. For example, Figure 6 shows that the caseload of unemployment benefits in the United Kingdom halved on a long term basis, while the number of claimants on incapacity benefits almost tripled. Such a dramatic increase is not (only) the consequence of the disappearance of certain types of industrial jobs, but suggests a broadening of entitlement criteria or an increase in types of work incapacity of a psychological rather than merely physiological nature (or both). In Denmark, a spectacular decline in the caseload of unemployment benefit between 1992 and 2008 can be observed, but the number of beneficiaries of work incapacity benefits, as in the Netherlands, remained exceptionally high during the same period.

By contrast, in Belgium the number of unemployment benefit claimants remained high throughout, while the caseload of incapacity benefits rose only moderately. In Figure 7 we plotted the absolute caseload developments of our four basic working-age categories as well as unemployment levels as harmonized by the OECD.

The transition towards a post-industrial service sector economy in Belgium seems to have been facilitated by the use of early retirement, as the number of
work incapacity beneficiaries started rising only when government began to close down this particular exit route. In addition, since 2000 there has been a steep increase in the caseload of the paid sabbatical leave schemes (from about 70,000 to around 170,000 cases by 2011). This particular labour supply reduction route seems to fit more with the “cost disease problems” of a service sector economy (Baumol, 2001) than with the labour-shedding needs of a de-industrializing economy. The benefits of the Belgian sabbatical scheme are financed by the unemployment insurance system, which largely accounts for the divergence between the caseload and expenditure trends reported above.

The Netherlands illustrate yet another variation on the principle of communicating vessels. Here we see not so much a decline of one type of benefit (unemployment) being compensated by an increase in another (work incapacity), but what could be described as a delayed shockwave. Trends in unemployment benefits caseload seem to follow the economic cycle, and work incapacity schemes follow suit, albeit with a time lag of a few years. It is too early to judge to what extent the rise in unemployment following the financial crisis of 2008 will produce a new shock wave in work incapacity, but a recent CBS report suggests a dramatic continuous increase of disability benefits for young people (the Wajong scheme): in recent years about one out of 12 young adults have been entering the scheme (van Vuuren, van Es and Roelofs, 2011).

In some respects, the Netherlands “pioneered” the use of work incapacity benefits as a way of accommodating redundant workers during the era of the so-called “Dutch disease”, when the country’s labour market was plagued by an exceptionally high level of inactivity. As Figure 8 demonstrates, it was only towards the end of the period under study when policy reforms seem to have facilitated a reduction of the caseload of incapacity benefit receipt. The steep rise in early retirement benefits during the 1990s is a reason to cast doubt on the alleged Dutch employment “miracle” (Visser and Hemerijck, 1997) of that decade. While the volume of unemployment benefit receipt declined steadily, structural unemployment over time seems to have been accommodated initially by the work incapacity scheme, and since the early 2000s by collectively bargained early exit from the labour market (bearing in mind the discrepancies we discussed earlier between the caseload of early retirement and the spending figures). In other words, these early retirement schemes seem to have taken over the shock-absorbing role that the work incapacity schemes played during the heyday of the “Dutch disease”. The steep increase in early retirement caseloads between 2000 and 2004 is remarkable. It seems conceivable that this was at least in part a reaction to the Dutch government’s plan to close down the early retirement pathway.

Finally, the German case illustrates a number of problems of creating comparable categorizations across countries and over time. As we see in Panel A of Figure 9, German unification in 1990 and the collapse of the East German
economy manifested itself in steep rises in unemployment, incapacity and early retirement caseloads. A change in the registration of persons claiming a statutory pension benefit before retirement age masked the scale of the latter for some time. The pension reform law of 1992 led to the re-categorization of persons drawing a standard pension before the statutory retirement age, treating early retirees as “regular” pensioners and thus reduced the early retirement caseload considerably (Deutsche Rentenversicherung, 2010, p. 169). However, this was masked by the subsequent steady increase of early retirement and persistence at a high level until about 2004, which indicates the important role this programme played at a time when the German economy was faced with a considerable decline of industrial employment, relatively low economic growth and mass unemployment.

Early retirement, but also incapacity benefits, helped to contain the increase in unemployment-related caseloads. In Figure 9, Panel B, we show the shifts between six types of benefits: unemployment insurance (Arbeitslosengeld, as of 2005, unemployment benefit – UB I), temporary unemployment insurance (Kurzarbeitergeld), unemployment assistance (Arbeitslosenhilfe), which ended in 2005, and unemployment benefit II (UB II). We also list separately recipients of social assistance or, as of 2005, of the UB II who are in paid work (the so-called Aufstocker), as this can be considered more as a wage subsidy than an out-of-work benefit. Disregarding the latter, the figure illustrates the steep decline of the number of people without paid work and in receipt of working-age benefit after 2010.
Figure 9. Changes in the caseload mix in Germany, 1980-2011

Panel A. Unemployment, work incapacity, early retirement and social assistance

Panel B. Changes in cumulative caseloads of unemployment benefits
A few remarks need to be made in order illustrate this trend, which in turn is a prime example of the need for reflection in the use of benefit claimant data for comparative purposes. First, before 2005 a certain percentage of social assistance claimants were registered as unemployed. However, this proportion varied over time as well as across regions. In the absence of systematically collected data we have thus not included unemployed social assistance claimants in the total unemployment caseload in Figure 9, Panel A, which is therefore underestimated. On the other hand, we have included persons in receipt of temporary unemployment insurance (TU) (Kurzarbeitergeld), which is a temporary benefit for persons in employment whose company reduced their working hours for economic reasons, thereby avoiding redundancies. It is thus a benefit which replaces lost earnings in part and claimants might be considered temporarily (and partially) unemployed. This instrument was heavily used in response to the economic downturn in 2008, but also in earlier periods (Möller, 2010; Hijzen and Venn, 2011).

Plotting trends separately Figure 9, Panel B, indicates that means-tested unemployment assistance became increasingly important during the 1990s to the extent that overall unemployment protection in Germany became increasingly “dualized” in terms of caseload numbers, or rather heterogeneous if unemployed social assistance claimants are included (see Clasen and Goerne, 2011). However, in 2005 this changed fundamentally when, in the context of the most important reform in German labour market policy since the 1960s, unemployment benefit provision was administratively and institutionally reformed. Of particular interest for us is the merger between the former unemployment assistance with social assistance (for claimants deemed to be employable) into the new unemployment benefit (UB II). UB II is more than a simple merger however, but conceptually new in the sense that it is not (only) a form of unemployment protection but wider by (also) covering claimants not registered as unemployed but able to work, as well as some groups of persons who are in minor or poorly-paid employment that cause those workers to fall below the official poverty line.

The introduction of UB II has led to a relative marginalization of unemployment insurance in caseload terms and a dominance of means-tested unemployment support in the modern German welfare state. From one perspective it illustrates a form of risk re-categorization (Clasen and Clegg, 2011) and new orientation within unemployment protection, potentially signalling a shift from unemployment to “employability” as the more appropriate caseload category in the future. Alternatively it could be argued that the social, political and administrative category of “unemployment”, which during the onset and development of the industrial era was increasingly restrictively applied (Whiteside, 2008) is currently undergoing a major revision. It is these types of processes which cross-national research based on benefit caseload data need to take account of. Failing to do so
would make the use of social transfer claimant numbers, even if systematically collected, misleading.

Conclusion

This article has explored the viability of caseload data for comparative analysis while making use of available administrative data to illustrate trends in working-age benefit receipt in six European countries. The discussion has shown that the relative paucity of internationally available and comparable time series is not only due to the absence or inconsistency of national data sources. There are substantial methodological challenges for research aimed at calculating total benefit dependency ratios and comparing those across countries. While certainly of interest, not least politically, considerable efforts would be required to overcome problems such as double counting and partial and periodic benefit receipt. Another set of challenges relates to identifying benefit categorizations and conceptual boundaries between public and private programmes, as well as for functional equivalents between, for example, persons of the same pre-retirement age drawing pensions from different pension programmes. Our discussion has shown that what might appear to be a fairly straightforward task of selecting, allocating and adding caseload numbers often requires concept stretching, omissions of some data or compromises.

This does not mean that caseloads should be ignored. Indeed, in many respects the use of other key indicators in comparative social policy research, such as social expenditure, face very similar problems. In recent years information on “disaggregated” social spending on different areas over time has become more readily available, which has allowed more detailed programme-based analyses of developments in comparative perspective (e.g. see Castles, 2008; De Deken, forthcoming). However, depending on research aims the use of these data also requires conceptual deliberation and, as with the use of caseload data, rests on contestable assumptions about categorization or functional equivalence. This is underlined by the, at times, considerable degree of discrepancy between what is supposed to be comparable data at the international level (see De Deken and Kittel, 2007).

Neither does this mean that caseloads do not offer “added value”. The discussion has shown that trends in spending and caseloads of the same programme do not always co-vary. This can be due to changes in employment or unemployment structures, or because of policy reforms which affect the two indicators differently. Provided that methodological problems can be adequately addressed both types of data are therefore distinctive and important indicators of welfare state change. In any case, given the political saliency and concern over “benefit dependency”, social researchers should be clear about the options for and limitations of investigating social security claimant numbers over time both within and across countries.
Bibliography


De Deken, J. Forthcoming. “Is it possible to measure the social investment state? On the problems of defining a concept and operationalising it on the basis of social expenditure data”, in F. Vandenbroucke and B. Cantillon (eds.), For better, for worse: For richer, for poorer, labour market participation, social redistribution and poverty in the EU. Oxford, Oxford University Press.


Extending pension and savings scheme coverage to the informal sector: Kenya’s Mbao Pension Plan

Rose Musonye Kwena and John A. Turner

Retirement Benefits Authority, Nairobi, Kenya; Pension Policy Center, Washington, DC, United States

Abstract   The Mbao Pension Plan is a voluntary individual account savings plan to which all workers in Kenya may contribute without regard to income or age. It is designed to provide a programme that is suitable for the unique nature of the informal sector and to encourage a savings culture for those workers. The key innovation is that low-income workers can easily make small contributions at relatively low cost, considering the small contributions and small account balances. Participants can conveniently make contributions anytime and anywhere using their cell phones. This savings innovation is made possible by technological innovations that have reduced the costs of cell phones and airtime, and by the entrepreneurial innovation of mobile money. The plan is provided through private-sector businesses.

Keywords   pension scheme, social protection, gaps in coverage, saving, informal sector, telecommunications, Kenya, Africa

Introduction

In Kenya, about 80 per cent of the labour force works in the informal economy (Raichura, 2008) and less than 15 per cent of the population is covered by social
security old-age benefit programmes (Okulo, 2011). At the end of June 2011, the Retirement Benefits Authority in Kenya along with the National Federation of Jua Kali Associations launched an innovative programme using private-sector providers of cell phone technology for extending savings plan coverage to workers in the informal sector: the Mbao Pension Plan. This is not a social assistance cash transfer programme, but a voluntary contributory savings plan wherein individual workers choose the amount of their contributions. Albeit still in its infancy, the programme may yet offer a promising innovation for extending social protection coverage to the informal sector in Kenya and in other countries. This article considers in particular issues relating to its sustainable development in Kenya and its possible future extension to other countries in the East African Community.

Key objectives of contributory social security programmes, such as old-age pension schemes, are to provide income smoothing across the life course and insurance against the risk of longevity for workers (Barr and Diamond, 2009). Most old-age social security programmes are financed by mandatory monthly contributions paid by insured workers, with matching contributions paid by their employer. This model presupposes not only a formal employment contract but that the insured worker and employer adhere fully to the rules as regards the collection and payment of contributions to the social security system.

In countries with high levels of informal employment, social security coverage rates often remain legally restricted to those, frequently a small minority, working in the formal labour market. This coverage problem occurs in most middle- and lower-income countries. Worldwide, only 26 per cent of the working population is effectively covered by social security old-age benefits programmes (ILO, 2010). In Africa as of 2010, about a third of workers were legally covered by social security old-age benefit programmes, but effective coverage was substantially lower. In sub-Saharan Africa, effective coverage by contributory programmes is about 5 per cent, while in North Africa it is about 20 per cent (ILO, 2010).

Extending social security programmes designed originally for the formal sector to the informal sector is challenging. Leaving important legal and practical

---

1. In 2011, GDP per capita in Kenya was USD 1,700. In 2007, an estimated 75 per cent of the population worked in agriculture, but agriculture only accounts for 24 per cent of GDP (CIA, 2012). An estimated 22 per cent of the population lives in urban areas. With a population of 43 million, life expectancy at birth in Kenya is 63 years (CIA, 2012). The percentage of the population aged 60 or older was 5 per cent in 2010, and is projected to stay at that rate through 2020 and rise slightly to 6 per cent in 2030 (ESA, 2011). The birth rate of 32 births (per 1,000 population) ranks 37th in the world. An estimated 6.3 per cent of the adult population has HIV/AIDS, which is the eleventh highest in the world. The literacy rate for the adult population is 87 per cent (CIA, 2012).
2. The informal sector is called Jua Kali in Swahili. The Jua Kali sector is not just agricultural workers, but includes artisans in carpentry and woodworking, soapstone, ceramics, textiles, leather, beekeeping, as well as small business in electronics and mechanical engineering.
3. Effective coverage refers to the number of people who are actively contributing to a social security programme.
administrative issues aside, extending coverage to the informal sector is difficult because generally these workers have low and variable cash incomes. For these reasons, they cannot make regular contributions and contribution rates that may be reasonable for the formal sector are often too high for workers in the informal sector. Also, low-income workers might not see the benefits provided by those programmes as being relevant to their needs and for addressing the multiple risks with which they are confronted across the life cycle.4

Alongside social security systems, financial services products are important tools to better manage risk across the life cycle. Low-income people, including many in informal employment, despite having similar needs as higher-income and formally-employed people for loans, savings instruments, and pensions, generally are not well served by banks and other financial institutions. In this sense, when it comes to accessing formal “social risk management” (Holzmann and Jørgensen, 2000) mechanisms, the poor and those working in the informal economy are very often doubly excluded. In turn, a further challenge is that traditional “informal” support systems in many countries are weakening due to changes in demographic patterns, cultural behaviour and the socio-economic context, thus leaving vulnerable populations further exposed to the impacts of downside risks.

As one response to this complex set of challenges, the development of microfinance initiatives in many countries seeks to help workers and families better manage risk. Early microfinance comprised almost exclusively microcredit for microenterprises (Rutherford, 2012), providing small loans, sometimes as small as USD 100, to low-income persons to help them start or improve a business (Women’s World Banking, undated). Microfinance is now a well-established practice to help low-income persons who do not have access to banks or financial markets. Some of the loans have led to improved income and savings for poor persons, and thus to improved income security in old age.

However, micro-savings accounts and micro-pensions, which would be the savings and pensions equivalents of micro-loans, are not well established. Micro-savings accounts and micro-pensions would make it possible for low-income persons to easily, safely, and at low-cost make small contributions to savings accounts or to individual account pensions.

A micro-pension is a voluntary, defined contribution, individual account plan. Shankar and Asher (2011) explore issues relating to micro-pensions in India, one of the few countries where they have been developed. The first micro-pension in India was developed in 2006 by a mutual fund company in connection with a

4. Some countries have separate programmes with lower contribution rates for the informal sector. Efforts to extend coverage of contributory programmes to the informal sector have relied primarily on a tax being levied based on some assessment of their earnings. See Gillion et al. (2000).
bank. All participants needed to have a bank account with that bank. Shankar and Asher conclude that micro-pensions have the potential to be one of the most useful components in India’s multi-tiered pension system. Micro-pensions and micro-savings plans typically are provided by microfinance institutions and are normally long-term savings products or hybrids between pension schemes and savings products. Shankar and Asher argue that micro-pensions need to offer participants the ability to make small contributions at convenient locations. For this reason, it is difficult to construct a nationwide network for the collection of contributions. They suggest that microfinance institutions could be a promising approach, with representatives meeting with clients in their home or neighborhood. In India, microfinance clients are either served by self-help groups with a bank linkage or by microfinance institutions, which provide micro-loans.

In addition to tax-financed programmes that are important components of “social protection floors”, micro-savings accounts and micro-pensions provide a way of compensating for gaps in contributory social security coverage and reducing the vulnerability of low-income persons and informal-economy workers and their families, to help them better manage risk and combat economic insecurity. In this light, the implementation and design of Kenya’s Mbao Pension Plan is worthy of attention and analysis.

The rest of this article is structured as follows. The article starts by discussing an important aspect of social security infrastructure – the payment mechanisms for social security contributions and benefits, briefly comparing cash, branch banking, and mobile money, particularly in the context of developing countries. The discussion then focuses on Kenya, first discussing the penetration of mobile phones and mobile money in Kenya as a platform for making low-cost transactions for a pension or savings plan. The article next outlines the Mbao Pension Plan and analyses its potential to make a significant contribution to enhancing social protection, considers possible enhancements to the plan, and considers the extension of this plan to other countries in the East African Community. Final conclusions are then offered.

**Payment mechanisms for pension programmes**

Social security programmes, pension schemes, and savings schemes rely on infrastructure for the payment of contributions by workers and employers and for the payment of benefits to beneficiaries. In many countries, employers make contribution payments, including the employees’ share, by check or by electronic transfers. Social security programmes often pay benefits directly into the bank account of the beneficiaries or pay the beneficiaries by check. Thus, for payments of contributions and of benefits, employers and social security programmes generally do so through the banking system, rather than making them in person and in cash, which is con-
siderably more expensive and risky. For example, China’s new pension plan for rural workers makes benefit payments into the bank account of the beneficiary (Chen and Turner, 2012).

More than half of the world’s adult population, however, lacks access to banking services (Koech, 2012). In the developing world, 64 per cent of adults do not have bank accounts, compared to 17 per cent in developed countries. In developing countries, banks tend to be concentrated in high-income urban areas, while in many countries the majority of the population lives in rural areas. The majority of adults in developing countries thus only make cash transactions. For this reason, social security programmes need to develop low-cost alternative payment mechanisms, other than through traditional “brick and mortar” banking systems.

Compared to the low coverage for banking, 75 per cent of the world’s adult population has access to cell phones. Worldwide cell phone subscriptions have increased from 719 million in 2000 to 6 billion in 2011. Much of the growth has occurred in the developing world, due to low-cost phones and low-cost prepaid airtime (Koech, 2012). In Africa, cell phone connectivity has gone from virtually zero in the 1990s to now more than 60 per cent of Africans having mobile phone coverage. In Africa, mobile phones are more than ten times as numerous as land line phones (Aker and Mbiti, 2010; Mbiti and Weil, 2011).

In developing countries, cell phone subscriptions often outnumber bank accounts. This is the case, in South Africa, Algeria, and Ghana, for example, where cell phone subscriptions are more than twice as numerous as bank accounts (Koech, 2012). The set up costs for establishing cell phone signals in a rural or sparsely populated region are considerably smaller than the costs of establishing a physical infrastructure of banks (Vincent and Cull, 2011).

Cell phone usage has led to the development of mobile banking. Mobile banking is offered through different types of business arrangements. Some companies offering it are banks, some are mobile phone companies, and some are partnerships between banks and mobile phone companies. Regulatory requirements in a country can determine the types of companies involved. As of 2012, 129 mobile banking programmes worldwide were operating, with many more planned for future operation.

With mobile banking, individuals can use their cell phones to make deposits and withdrawals, transfer funds and pay bills. In the developing world, mobile money provides the promise of extending banking services to many people who did not previously have easy access to those services. Mobile money is an inexpensive, easy, and safe substitute for cash transactions, transactions by check or credit card (which generally are not available options), or transactions through Western Union or post offices. It has an advantage over programmes that deal in cash in that programmes based on the physical delivery of cash are more vulnerable to
fraud, theft of the cash in transit, and corruption. The cost of a transaction in cash is typically considerably higher than the cost of a cell phone transaction (Vincent and Cull, 2011).

**Mobile phones in Kenya**

The Mbao Pension Plan, the focus of this article, makes use of cell phone technology for payment of contributions and benefits. This section provides background concerning the mobile phone system that forms the basis of the Mbao Pension Plan.

Mobile phones are relatively inexpensive in Kenya and usage is also relatively inexpensive, with workers purchasing minutes of air time for as little as 20 Kenya Shillings (KES) (the equivalent of USD 0.25). Owing to the low cost, many poor people in Kenya own and use cell phones. The main mobile phone provider in Kenya is Safaricom with 15.2 million accounts, up from 9.5 million in 2010 (The Economist, 2010). A second company, Airtel, also provides mobile phone services.

The airtime for mobile phone service is mostly prepaid in Kenya, as in other parts of Africa and the developing world (Koech, 2012). For example, in Tunisia, when mobile phones were first introduced, companies required subscribers to have a bank account as a form of assurance of payment for phone services. The invention of prepaid phone cards allowed mobile phone companies to provide services to people without bank accounts or regular income (Delarue, 2013).

**Mobile money: M-Pesa**

Mobile phone companies in Kenya have established mobile money systems. These provide an inexpensive alternative to banks, and provide the transactions infrastructure for the Mbao Pension Plan – only 24 per cent of Kenyans have bank accounts, but 75 per cent of the adult population has mobile phones (Koech, 2012).

With mobile money, a mobile phone user gives cash to an agent of the mobile phone company. The agent typically is a retail store. The use of existing retail stores reduces the cost. The agent serves as a branch bank office, collecting and disbursing cash. The member can use his or her cell phone to transfer money to another account to pay bills. The system functions much like a bank, though officially it is not considered to be a bank as it is mainly regulated by the Communications

5. With a population of 43 million in 2012 and a population over age 15 of 25 million (CIA, 2012), 60 per cent of those over age 15 have cell phones through Safaricom, a Vodafone affiliate.
Commission of Kenya, the regulator for mobile telephone companies. However, mobile money transactions are regulated by the Central Bank of Kenya.

The Safaricom mobile money system is called M-Pesa.\(^6\) Launched in March 2007, M-Pesa was originally designed in trials so that microfinance borrowers could more conveniently receive and repay loans. The cost savings compared to dealing in cash permitted microfinance lenders to offer loans at lower rates. However, it was quickly extended to serve as a general money transfer service. Within eight months of its launch, M-Pesa had registered 1.1 million Kenyans (Safaricom, 2007) and is considered by some to be the most successful mobile money service in the developing world (Jack and Suri, 2011).

Mobile money was first adopted primarily by higher income persons. The early advertising focused on the message, “send money home”, offering a means to send remittances to parents (Mas and Ng’weno, 2010). Its adoption has since become widespread among lower-income persons as well (Jack and Suri, 2011). By 2008, a survey found that 43 per cent of households used M-Pesa, and by 2009 it was nearly 70 per cent. That survey also found that early adopters tended to be male, but that a year later adopters were more evenly split between the genders.

In developing countries such as Kenya, the transfer of money from one location to another can be done in various ways. Before M-Pesa, some bus companies provided money transfer services, as well as the post office, banks, and Western Union. Informal money transfers were also done through friends travelling to other parts of the country. Within a two-year span, however, M-Pesa became the dominant form of formal money transfer in Kenya (Mbiti and Weil, 2011). One study found that the success of M-Pesa has been due to its speed, reliability, safety, large network of outlets, and its low price relative to its competitors (Morawczynski, 2009).

A key feature of mobile money for pensions, savings schemes and other transactions is its convenience. Safaricom had over 27,900 agents across Kenya in 2010, compared to 840 bank branch offices (The Economist, 2010). Many small grocery stores, gas stations, cooperatives and courier services have become agents who work exclusively with Safaricom. By the end of 2012, it had 70,000 agents and this number is likely to expand further. Saraficom through M-Pesa currently transfers money each year equivalent to 11 per cent of the GDP of Kenya (The

---

\(^6\) M stands for mobile and Pesa means money in Swahili. Similar mobile money successes in other countries include GCash and Smart Money in the Philippines, Wizit and M-Pesa in South Africa, and M-Paisa (a version of M-Pesa) in Afghanistan and India (Koech, 2012). Vodacom also offers M-Pesa in Tanzania. Nearly all mobile phone companies in Africa have replicated M-Pesa and are now offering mobile money services, making Africa a leader in the developing world in electronic money transfers (Mwiti, 2012).
Economist, 2010). One of the uses of M-Pesa is to purchase Safaricom airtime. Users of M-Pesa are assigned an electronic money account that is connected with their phone number. It is accessed through a subscriber identity module (SIM), which is a memory chip in the phone. New users of M-Pesa need to get a new SIM card, which Safaricom provides for free.

Users of M-Pesa exchange cash for “e-float,” which is the credit to their account. It can be used to transfer money to other people, make payments, or later be exchanged back for cash. The “e-float” does not pay interest. All e-float is backed 100 per cent by deposits at three commercial banks in Kenya.7

Each mobile money platform has a PAYBILL number into which contributions are made by members. The user goes to an M-Pesa agent and gives that person the cash that he or she wants to have added to their account and shows an identification card. The documentation provided by the national identification card satisfies the Know Your Client requirements. The agent records the person’s telephone number and cash deposit in a book. Then the agent enters the customer’s phone number and contribution amount into his or her cell phone. The customer waits to receive a text message confirming that the money has been transferred to their account. The wait is about a minute, but can be longer during periods of peak usage. The instant recording of deposits and confirmation of receipt provide a form of protection for users to assure that their money has been safely deposited. Deposits are made free of charge, but a fee is charged for withdrawals or transfers.

The potential for using M-Pesa to facilitate savings was recognized early in its development. At the end of 2008, before the start of the Mbao Pension Plan, 75 per cent of M-Pesa users reported saving in their M-Pesa account, with 21 per cent saying it was their most important form of savings, even though it pays no interest (Jack and Suri, 2011). To provide a means of savings that pays interest, Safaricom and Equity bank have joined together to permit use of M-Pesa for savings through an interest bearing savings account. The new product is called M-Kesho. M-Kesho is a micro-savings account. There are no fees for opening an M-Kesho account, no periodic fees, and no minimum balance. The interest rate varies from 0.5 per cent to 3.0 per cent per year, reaching 3.0 per cent when account balances reach KES 10,000 (approx. USD 125) (Mbiti and Weil, 2011). One study has found that use of M-Pesa for savings reduced the use of “secret hiding places” for savings (Mbiti and Weil, 2011). It should be noted, however, that the interest rate on these accounts is far lower than the rate of inflation.8

7. Safaricom donates the interest earned on those accounts to charity, thus avoiding being regulated as a bank.
8. The inflation rate in 2011 was 14 per cent.
The Mbao Pension Plan

A micro-pension plan or savings scheme designed for low-income participants

The combination of widespread cellular phone use and the ability to transfer money instantaneously, safely, and inexpensively are having widespread effects on the organization of economic activity and on risk management and mitigation (Mbiti and Weil, 2011). In particular, they may provide the basis for a new approach to extending pension systems to persons in rural areas and the informal sector.

Going a step beyond M-Kesho, at the end of June 2011, the Retirement Benefits Authority in Kenya, which is the retirement benefits sector regulator in Kenya, with the National Federation of Jua Kali Associations and the participation of private-sector providers launched an innovative programme for extending pension and savings scheme coverage to the informal sector: the Mbao Pension Plan. The Mbao Pension Plan is a voluntary individual account savings plan provided through private-sector businesses. The government bears a small cost through a loss of tax revenue on the savings and through the cost of the regulator, the Retirement Benefits Authority.

Mbao is Swahili slang for 20 shillings, and the name of the plan refers to the minimum contribution of 20 shillings (approx. USD 0.25). It is a micro-pension or savings scheme in that it provides a plan where low-income workers can make small contributions at flexible intervals of their choosing. Since it is designed and marketed as a pension plan, the goal is to encourage retirement savings, but it can also be used for savings for other purposes. Experience in India has suggested that small, convenient, and frequent pay-ins make it easier to save (Rutherford, 2012). It is designed primarily for low-income workers who are not participating in social security, and thus is a substitute for social security, but can also be used as a complement to social security by workers who are covered by such programmes. Table 1 presents a number of facts about the Mbao Pension Plan.

Workers must pay KES 100 to register with the Mbao Pension Plan, and must fill out a registration form and present ID documentation. The forms are sent to the fund administrator’s office. The user needs a national photo identification card to establish an account. However, this registration procedure will change in the future.

9. Mbao is pronounced similarly to the English word “pound” but dropping the last two syllables. It derives from the Kenyan pound having been the equivalent of KES 20.
10. While it is popularly known as the Mbao Pension Plan, its official name is the Blue MSMEs Jua Kali Individual Retirement Benefit Scheme. Blue represents the colour of the KES 20 note, while MSMEs (medium and small micro enterprises) represent the informal sector.
near future as the administrator is updating the platform to cater for full electronic registration using the phone.

Though the Mbao Pension Plan is barely over a year old, as of November 2012 it had 38,000 members who have saved KES 37 million (USD 463,000). By the end of December 2012, that amount had increased to KES 41 million. While voluntary savings programmes often have limited impact, and many people may not participate, the early experience with the Mbao Pension Plan suggests that it is an appealing programme to some people.

Table 1. Facts about the Mbao Pension Plan, 2012 (end of year)

<table>
<thead>
<tr>
<th>Item</th>
<th>Fact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum allowable contribution</td>
<td>KES 20 a day</td>
</tr>
<tr>
<td>Target minimum weekly contribution</td>
<td>KES 500 a month</td>
</tr>
<tr>
<td>Maximum tax deductible contribution</td>
<td>KES 20,000 a month</td>
</tr>
<tr>
<td>Maximum allowable contribution</td>
<td>KES 140,000 a day</td>
</tr>
<tr>
<td>Average contribution per contribution transaction</td>
<td>KES 180</td>
</tr>
<tr>
<td>Number of members</td>
<td>38,000</td>
</tr>
<tr>
<td>Amount saved</td>
<td>KES 37 million</td>
</tr>
<tr>
<td>Registration fee for participation (one-time fee)</td>
<td>KES 100</td>
</tr>
<tr>
<td>Fees on contributions</td>
<td>KES 3 for a contribution up to KES 49</td>
</tr>
<tr>
<td>Fee on minimum contribution of 20 KES</td>
<td>KES 3, or 15%</td>
</tr>
<tr>
<td>Fees on assets</td>
<td>Waived for first three years of the plan; thereafter, 0.95% per year</td>
</tr>
</tbody>
</table>

Note: KES is Kenyan shillings. KES 100 = approx. USD 1.25 in 2012.

Sources: Developed by authors.

Contributions

The minimum contribution to the Mbao Pension Plan of KES 20 a day largely appeals to the lowest income earners, such as hawkers, who foresee for themselves a decent retirement. Scheme members commit to contributing at least KES 100 a week and KES 500 a month, but contributions are not mandatory, and no penalty is charged for not contributing. Contributions of KES 500 a month (KES 6,000 a year) would be roughly equal to USD 75 a year. The Mbao Pension Plan does not have a maximum contribution, but the mobile phone companies will not handle contributions of more than KES 140,000 a day, which is the daily contribution.
maximum remittance. For all pensions in Kenya, the maximum tax deductible contribution is KES 20,000 a month. Contributions are made only by members, with no matching contributions by employers or the government. Though the system is designed with its low minimum contribution to be a system that informal-sector workers can use, all workers may participate.

With Safaricom, the fee for a contribution of KES 20 is KES 3, which is the set fee for transfers between KES 10 to KES 49. This ensures that all low-income members are subsidized by those making larger contributions, who pay higher fees. In addition, Safaricom charges the Mbao Pension Plan a fee of KES 2. The fee rises in steps with higher contributions, but as a per cent of contributions it falls. Safaricom’s market competitor, Airtel, charges workers a 10 per cent fee for small contributions. If contributions are withdrawn in the first year, a penalty is levied, but after that no penalty or fee is charged for withdrawals.

The key innovation of the Mbao Pension Plan is that low-income workers can easily make small contributions at relatively low cost, considering the plan is dealing with small contributions. Workers can conveniently contribute anytime and anywhere using their cell phones. This innovation is made possible by the technological innovations that have reduced the costs of cell phones and have reduced the costs of airtime, and by the entrepreneurial innovations of pre-paid phone cards and mobile money.

The current average contribution is KES 180 (Kwena, 2012). A survey found that 42 per cent of the informal-sector workers participating in the Mbao Pension Plan earned less than KES 6,000 (approx. USD 75) a month (Anami, 2012). Thus, the system is serving the needs of substantial numbers of low-income persons. The fact that members can make such small contributions towards pension saving has helped to demystify the notion that saving for pensions is only for people with disposable income. It further confirms longstanding arguments made in the financial services literature concerning the needs of low-income persons for financial services.

**Investment and savings aspects for participants**

To prepare for the launch and to encourage participation in the Mbao Pension Plan and other pension schemes in Kenya, in June 2009, the Government amended the Retirement Benefits Act to allow retirement benefit scheme members to assign up to 60 per cent of their accumulated pension or savings accounts to access mortgage facilities. This enables the mortgage financier to lend up to 115 per cent of the property value, with the additional lending going to finance the initial fees which consist of the government tax in the form of stamp duty, valuation fees, and legal fees. Mortgage lending in Kenya without this backing generally requires a down payment of 10 per cent of the purchase price of the house. When members use
their Mbao Pension Plan for this purpose, they cannot make a withdrawal from the plan until the mortgage is paid off.

The pension-backed mortgage serves as a key incentive to saving through a pension or savings scheme rather than through other saving vehicles like bank accounts, cooperative societies, insurance products and investment groups. Members therefore understand the importance of accumulating substantive reserves in their pension or savings scheme in order to qualify for a mortgage. This programme provides an immediate, tangible benefit to workers for participating in a pension or savings scheme. The house is the first form of guarantee of the mortgage. If an individual loses his or her job and defaults on his or her mortgage, but the value of their house exceeds the amount remaining on the mortgage, the house is sold but the individual does not lose any of his or her pension or savings account. The worker would lose part of his or her pension or savings account only if the price of the house had fallen below the value of the mortgage.

Contributions to the Mbao Pension Plan are taxed under the same tax regime as contributions to other pension schemes in Kenya. That is the EET tax regime, where contributions are tax deductible, investment earnings are tax exempt, and benefits are taxable. However, lump-sum payments in Kenya of up to KES 600,000 per year are tax exempt, so most payments from the Mbao Pension Plan will be tax exempt. Most workers in the informal sector pay no income taxes, so they do not benefit from the preferential tax treatment.

Policy-makers in Kenya are considering whether low-income participants should receive a subsidy to encourage their participation and to increase the amount in small accounts. A complication in doing this may be that participation in the Mbao Pension Plan is not limited to low-income participants. A subsidy for participants would need to be structured so that it only went to low-income participants not benefiting from a tax subsidy because they did not pay taxes.

The Mbao Pension Plan as of the end of 2012 was invested entirely in interest bearing assets, with more than a third (37.8 per cent) in Kenyan government bonds, and another third (37.0 per cent) in fixed-term bank deposits. The remainder is in corporate bonds (15.6 per cent) and cash (9.6 per cent). As of this writing, the rate of return for 2012 had not been declared.

By way of international comparison, like in Kenya, rural workers and farmers in China do not pay income tax, so in both countries workers do not have a tax incentive for participating in the pension system. The lack of a tax incentive may explain in part the matching contribution in China. In China, subsidies are used to encourage participation in the National Rural Pension System, where participation also is voluntary. The subsidy, which is provided by the government, varies across regions, but the minimum subsidy is a flat 30 Chinese Yuan (CNY) a year, where the minimum contribution is CNY 100 a year (approx. USD 16), with nearly half of participants making the minimum contribution (Dorfman et al., 2013).
**Benefit payments**

Individual savings in the Mbao Pension Plan can only be drawn down as a lump-sum payment, and in this respect the Mbao Pension Plan is similar to a provident fund. A lump-sum withdrawal can be made at any age, after a year of participation in the plan. Upon death of the account holder, the plan makes a lump-sum payment to the designated beneficiary. Members can access their account balance using their mobile phone. Clearly, with the plan having only been in existence for little more than a year it is too early to assess the level of benefits that it will ultimately provide.

Kenya’s social security programme, the National Social Security Fund, is also a provident fund, providing benefits as a lump sum. However, the Kenyan government as of 2013 is considering converting it to a social security fund, providing periodic benefits. Kenya could also consider offering periodic benefits through the Mbao Pension Plan.

**Design features that are different to those of a traditional pension plan**

Rutherford (2012) argues that micro-pension plans will not look, at least at first, like miniature versions of developed country private pension plans, and work best if they are medium-term savings commitment devices. Since formal employment and formal retirement are rare in the informal sector, an expanded view is needed as to how a micro-pension best functions.

Besides its innovative payment mechanism, several aspects of the Mbao Pension Plan differ from a traditional pension plan. The feature that the total amount can be withdrawn after one year without penalty makes it more like a savings plan than a pension plan. However, the requirement that the worker cannot access the money so long as he or she is using it as security for a mortgage means that some workers in old age might not be able to access it. While it is possible that both these features might be modified in some way in the future to strengthen the pension aspects of the plan, the Mbao Pension Plan may be viewed as an innovative hybrid savings arrangement that is structured to meet the needs of informal-sector workers.

**Financial services infrastructure**

The Mbao Pension Plan is officially registered with the Kenya Revenue Authority (tax authority), because it receives preferential tax treatment as a pension plan, and with the Retirement Benefits Authority, which is the pension regulator.
The financial services infrastructure for the Mbao Pension Plan involves several private-sector companies. The trustee and custodian for the Mbao Pension Plan is KCB Group (Kenya Commercial Bank), which is a well-known and widely trusted bank. (KCB allows its customers to transfer money from their accounts to M-Pesa, which can be then transferred to the Mbao Pension Plan.) The custodian holds the funds while the fund manager makes the decisions as to how they are invested. Individual participants do not decide how their funds are invested. By law in Kenya, the sponsor of the plan, which is the Kenya National Jua Kali Co-operative Society, must be different from the trustee. All pension schemes must have a sponsor who establishes them before appointing trustees and other service providers to manage the scheme funds. The sponsor has to make the declaration of a “trust deed” for a retirement benefits scheme to be established. The sponsor also ensures efficient administration of the scheme by appointing service providers like fund manager, trustee and custodian.

The fund manager, Co-op Trust Investment Services, carries out the investment decisions of the trustee. The fund administrator, Eagle Africa Insurance Brokers, handles the registration of new members and also keeps records for the scheme. The fund administrator is mandated with the responsibility of all administrative affairs of the retirement benefits scheme; ensuring that the scheme is run in accordance with the trust deed and rules and that the scheme is run within the law. Safaricom M-Pesa and Airtel Money are the mobile money transfer platforms. Other mobile services are expected to participate in the future.

**Regulation**

The Retirement Benefits Authority (RBA) has regulatory responsibility over the Mbao Pension Plan. RBA licenses all service providers of pension schemes in Kenya. The division of responsibility among different private-sector service providers concerning pensions ensures that there is a Chinese wall between them, thus giving an element of insurance against financial fraud. Fraud is a criminal offence subject to prosecution. If fraud were to occur, RBA, through its risk profiling process, would identify the lapses in the scheme management. Once the lapses are ascertained, RBA will immediately remove the trustees from office and appoint an interim administrator to take over the scheme management. If the interim administrator can prove that the fraud was caused by service providers, RBA can withdraw their licenses, in addition to instituting a legal (prosecution) process. If it is established that service providers colluded with trustees, then the trustees can be made to pay back any losses, in addition to being prosecuted.
Fees

For the first three years of operation of the Mbao Pension Plan, all the service providers except the two mobile phone companies have agreed to waive their fees. After the first three years, it is anticipated that the total fees charged by those service providers will be 95 basis points, or 0.95 per cent of assets per year. Those fees are 0.35 per cent for custody, 0.3 per cent for administration, and 0.3 per cent for the fund manager (Kwena, 2012). Because the expenses of managing a large investment account are similar to those for a small account, the system in Kenya of charging all accounts the same percentage fee in effect causes large accounts to subsidize small accounts in terms of fees paid. In addition to these fees, fees are charged by the mobile phone companies. The fees charged on contributions imply that relatively short duration participation is less advantageous than longer-term participation. A fee of 1 per cent of assets per year charged over the history of a person’s participation in a pension fund will reduce the benefit at retirement by about 20 per cent, compared to the hypothetical benchmark of no fees (Whitehouse, 2000).

To place the level of fees in perspective, they can be compared with trends in fees in the United States. To understand trends in mutual fund fees, the Investment Company Institute (ICI) has combined major fund fees in a single measure. ICI created such a measure by adding a fund’s annual expense ratio to an estimate of the annualized cost that investors pay for one-time sales charges. By this measure, mutual fund fees that investors pay have trended downward since 1980. In 1980, investors in stock funds on average paid fees of 2.32 per cent of fund assets; by 2006, that figure had fallen to 1.07 per cent of fund assets. Fees paid on bond funds have declined by a similar amount (Investment Company Institute, 2007). By comparison, Shankar and Asher (2011) find asset management fees in India ranging from 1.75 per cent to 2.5 per cent for the relatively new micro-pensions.

Marketing

Marketing the Mbao Pension Plan includes both informing participants about its availability and explaining the advantages of using it. The Retirement Benefits Authority has taken an active role in marketing the Mbao Pension Plan with information featuring prominently on its website.11 The endorsement of the plan by this governmental authority is designed to encourage trust in the plan as a safe way to save.

The issue of trust by workers has been viewed as important to the success of the Mbao Pension Plan. This issue has been important in part because of an infamous Ponzi scheme in Kenya, where many people lost money.\footnote{A Ponzi scheme is a fraudulent investment scheme where investors are paid a return from their own investments or the investments of later investors.} To instil trust in the Mbao Pension Plan, the KCB bank was chosen as corporate trustee due to its solid roots in the country and the region.

Another benefit to participants in the Mbao Pension Plan is that they have the opportunity to learn about savings and retirement issues through training provided by the service providers. The government regulator sponsored training on the Mbao Pension Plan for two champions (opinion leaders) from each of the 47 counties across the country during the launch of the scheme. These champions were equipped with knowledge on the scheme and went back to their respective counties to share this knowledge with their members.

Why do people voluntarily contribute to the Mbao Pension Plan?

Much of the economic literature on participation in voluntary pension and savings plans is based on the assumption that people generally need an economic incentive to do so. For this reason, pension plans in most countries receive favourable tax treatment. Because favourable tax treatment is often not sufficient, there has been increasing interest in the use of matching contributions to encourage contributions (Hinz et al., 2013). Even matching contributions have proven to be a relatively weak incentive, but automatic enrolment in voluntary pension schemes has proved to be fairly effective (Madrian, 2013).

The puzzle raised by the participation in the Mbao Pension Plan is that for many participants, none of these incentives apply. They do not have a tax incentive, they do not have matching contributions, and they are not enrolled through automatic enrolment. The ability to use the account as security for a mortgage may be an incentive for some people.

Much of the research on reasons why people participate in voluntary savings plans is based on experience in high-income countries, where people already are participating in a social security programme (Madrian, 2013). One finding that may be relevant from that research is that changes that make it simpler or easier to participate may play a role in encouraging voluntary participation. Delarue (2013) argues that a lesson from behavioural economics (Thaler and Sunstein, 2009) is that one approach to encouraging informal-sector workers’ participation is to make it as easy as possible to do so. In Kenya, many of the participants in the Mbao
Pension Plan are not participating in social security and presumably do not have any other financial savings. The feature of the Mbao Pension Plan that workers can easily make small contributions may be the key feature, combined with their desire to have some financial savings.

Possible future changes to the Mbao Pension Plan

The Mbao Pension Plan is new, and it is likely to undergo changes as experience with it develops. The Women’s World Banking (2003) recommends that micro-pensions develop first as a hybrid between micro-savings plans and micro-pensions as a step toward building capacity for micro-pensions. This section discusses possible changes to the plan as it develops that would strengthen the plan and provide greater options for participants.

The financing options could be changed in various ways. The plan could allow employer contributions to worker’s accounts, offering a low-cost option to employers for providing pensions. The plan could be changed to allow workers to borrow from it, providing liquidity and an added benefit of participating, while maintaining the account (Rutherford, 2012).

A government subsidy through a matching contribution could be provided for workers who do not receive a subsidy through the tax system because they do not pay income taxes. Such government expenditure could be justified because the government would otherwise be incurring tax expenditure through lost tax revenue if these workers had benefited from a tax preference. Matching contributions are provided for voluntary government social security pension programmes in China (Dorfman et al., 2013), India (Palacios and Sane, 2013) and Thailand (Wiener, 2013). In India, one pension programme offered for a four-year period a matching contribution for workers who made contributions under a certain small amount (Shankar and Asher, 2011). While studies in the United States have found only a limited effect of matching contributions on increased participation (Madrian, 2013), the effect could be more substantial in Kenya where workers do not also have a social security pension and are not also receiving a financial incentive through a tax subsidy.

Also, the benefit options could be changed so as to allow other benefit payment options than lump-sum payments, including phased or periodic withdrawals or partial withdrawals, or the purchase of an annuity when the accumulated account balance had reached at least a threshold level. An alternative could be that an annuity would be offered for a fixed term, such as ten years (Rutherford, 2012). These options would provide participants greater flexibility in risk management.

The option of withdrawal of funds after one year could be modified so that funds could only be withdrawn after five or ten years, or before a set age only for
particular purposes, such as education or medical expenses. These changes would move the plan toward being a pension plan rather than a savings plan. One option would be to start on a trial basis offering a savings plan that could not be withdrawn for five years.

Finally, continued efforts at financial education may lead to increased participation as the target population better understands the need for accumulating financial savings for retirement, and better understands the advantages of participating in the Mbao Pension Plan.

**Mbao Pension Plan in the East African Community**

Kenya is a member of the East African Community (EAC), along with Burundi, Rwanda, Tanzania and Uganda. The combined mobile phone subscriptions in Kenya, Rwanda, Tanzania and Uganda are 42.3 million, which is three times the number of bank account holders. More than USD 1 billion is transferred monthly through M-Pesa in these countries (Mwiti, 2012). In addition, in 2009 another mobile phone company Zain launched a service called Zap that allows customers in Kenya, Tanzania and Uganda to transfer money (Vincent and Cull, 2011).

The widespread usage of mobile phones facilitates the extension of mobile money systems and the Mbao Pension Plan in East Africa. For example, the use of M-Pesa and the Mbao Pension Plan in Tanzania could extend banking services and pension and savings scheme coverage to many people in the informal sector. By comparison, one study found that 9 per cent of adults in Tanzania had a bank account (FinScope, 2008). To extend the mobile money and the Mbao Pension Plan, countries need to have supportive legislation that permits mobile phone companies to provide these services.

The East African Community is attempting to encourage greater economic cooperation among the member states, including fostering cross border movement of workers. To do so, it needs to develop pension arrangements that do not hinder job change. That goal could be reached by extending M-Pesa and the Mbao Pension Plan to the other four countries of the EAC. As well as providing a pension plan that is suited to the needs of the informal sector, the Mbao Pension Plan is suitable for the needs of mobile workers. Having a defined contribution plan with immediate vesting of contributions would facilitate cross border movement of workers. Workers could even continue contributing to the Mbao Pension Plan in their home country, so long as they maintained an M-Pesa account in their home country. Alternatively, they could establish a Mbao Pension Plan in their new country of residence. In either case, job-changing workers would not suffer a loss of pension rights, and they would continue to have the opportunity to contribute to a pension plan.
Conclusions

Low-income people around the world generally are not served by banks and other financial institutions, yet they often have the same needs as higher-income people for loans, savings instruments, and pensions. Microfinance is a well-established practice to help low-income persons who do not have access to banks or financial markets, providing small loans to low-income persons to help them start or improve a business. Micro-savings accounts and micro-pensions, which would be the savings and pensions equivalents of micro-loans, are not well developed. Micro-savings accounts and micro-pensions would make it possible for low-income persons to easily, safely, and at low-cost make small contributions to savings accounts or to individual account pensions.

The Mbao Pension Plan is an innovative plan that seeks to tailor a savings product particularly to marginal population groups and contribute to their improved social and economic security while also supporting the further development of the financial services and communication sectors. The Mbao Pension Plan might have a longer-term role in social protection strategies as a “stepping stone” to encourage a movement towards greater formality in working relationships/practices, and thus be supportive of greater inclusion in formal social security programmes. Given that the Mbao Pension Plan is barely over a year old, it is too early to be able to assess its long-term impact on savings for retirement and elderly well-being, and its extension to other countries remains to be seen. However, in spite of its short history, it may be viewed as an interesting experiment that merits attention as to its future development.

Bibliography


Extending pension and savings scheme coverage to the informal sector: Kenya’s Mbao Pension Plan


Extending pension and savings scheme coverage to the informal sector: Kenya’s Mbao Pension Plan


