

Redistributive pensions in the developing world

Achim Kemmerling* and Michael Neugart[†]

December 14, 2018

Abstract

Redistributive so-called social pension schemes have seen a remarkable surge in developing countries. These schemes often target the rural elderly and correlate with urbanization rates, urban rural-wage differentials, and family norms. We use this stylized evidence to motivate a political economy model for a Beveridgean pension system with trade-offs between four groups: the (poorer) rural old and young, and the (richer) urban old and young. We show under which conditions governments will install a pension system and increase its generosity as the share of the urban population rises, productivity differentials between urban and rural workers widen, or the social norm erodes. Our conclusion is that the role of the rural-urban divide in shaping redistribution merits more scholarly attention, as in many developing countries the gap between cities and the countryside widens.

Keywords: pensions, developing countries, political economy, family transfers, crowding out, electoral support

JEL-classification: H55, D72, O18

*University of Erfurt, Willy Brandt School of Public Policy, Nordhäuser Straße 74, D-99105 Erfurt, Germany, E-mail: achim.kemmerling@uni-erfurt.de

[†]Corresponding author: Technische Universität Darmstadt, Department of Law and Economics, Hochschulstraße 1, D-64289 Darmstadt, Germany, E-mail: neugart@vwl.tu-darmstadt.de

This is the peer reviewed version of the following article: Kemmerling A, Neugart M. Redistributive pensions in the developing world. *Rev Dev Econ.* 2019; 23: 702–726, which has been published in final form at <https://doi.org/10.1111/rode.12582>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions. This article may not be enhanced, enriched or otherwise transformed into a derivative work, without express permission from Wiley or by statutory rights under applicable legislation. Copyright notices must not be removed, obscured or modified. The article must be linked to Wiley's version of record on Wiley Online Library and any embedding, framing or otherwise making available the article or pages thereof by third parties from platforms, services and websites other than Wiley Online Library must be prohibited.

1 Introduction

Redistributive pension schemes have seen a remarkable surge in developing countries (Pallares-Miralles et al., 2012; ILO, 2014). To a large extent, this surge is due to the rise of so-called social pension schemes (Leisering and Barrientos, 2013; Holzmann et al., 2009). Social pensions are non-contributory pension schemes, financed by taxation, and (usually) targeted to poor people above the retirement age. In the last decades the number of developing countries which have introduced such pension schemes has more than tripled (see Figure 1). It is noteworthy that these pension schemes, by and large, not only redistribute between generations but also within generations. This makes them different from many social security systems of developed countries and also different from individual pension accounts in developing countries (Lindert, 1994; Perotti and Schwiendbacher, 2009)

In this article we offer an explanation for the emergence of redistributive pension systems and the increase in their generosity in the less developed world. We build a political economy model of a social security system of the pay-as-you-go type. We show under which conditions a government will be willing to install a pension system and increase its generosity as the share of the urban population increases, urban to rural productivity differentials rise, or the social norm driving family transfers to the elderly erodes.

Governments are more inclined to introduce a redistributive public pension scheme and increase its size if such a policy reform finds electoral support. Increasing taxes on the income of the working population in order to provide for old-age security may affect societal groups differently. In our approach, we focus on four groups in society that seem to be important for these kind of policy reforms in developing countries. Generally speaking, there is a divide between urban and rural citizens with significantly higher wages being earned in urban areas. Furthermore, this divide of a (young) working population is mirrored in the well-being of the older population which in developing countries still

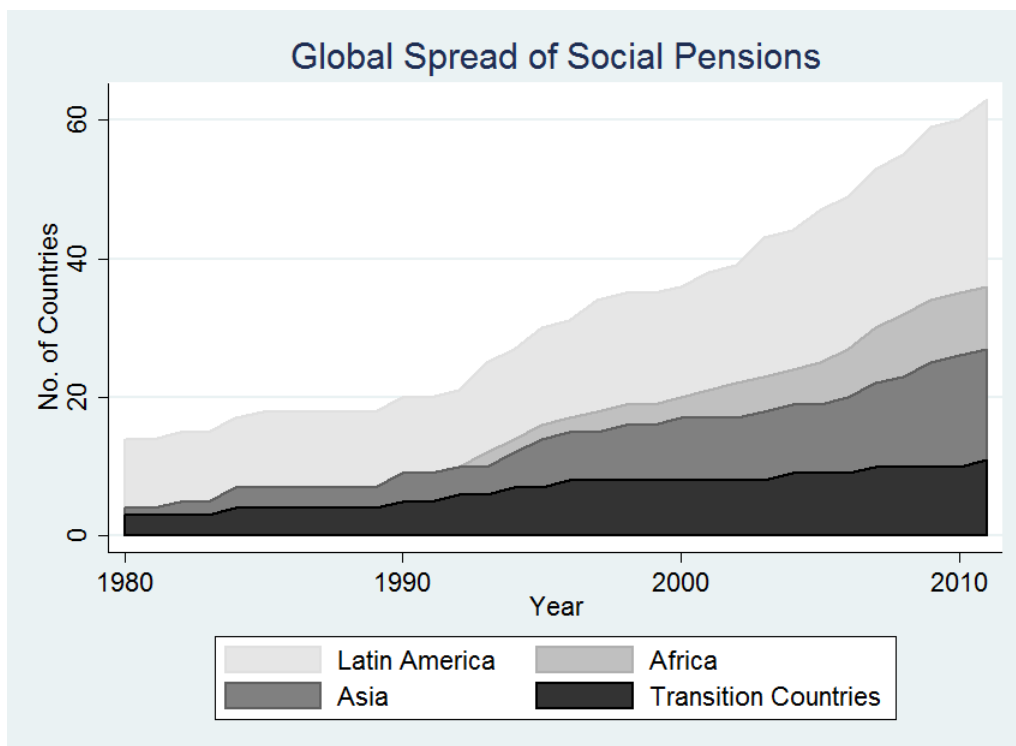


Figure 1: The Diffusion of Social Pensions

Sources: Holzmann et al. (2009) and own compilation.

depends to a large degree on transfers paid by younger family members. As the public pension schemes that have been introduced or enlarged more recently do not only transfer between generations but also redistribute within the younger generation, a conflict of interest on the size of such a public program arises between those living in high paying urban areas and the rural population. Moreover, if public pension schemes crowd out private transfers to the elderly the interest of the old population will be affected. We show that a process of urbanization reflected not only in an increase in the urban population share but also in higher urban-rural wage differentials, and an erosion of a social norm driving the transfers to the old population may explain the introduction of a redistributive public pension scheme and a rise in its generosity.

For our analysis we opted for a political economy model of the social security system that builds on a pay-as-you-go system and a government support function. The choice to model the pensions as a redistributive pay-as-you-go system, as opposed to, say, private individual accounts, follows the empirical evidence on how pension systems are organized in the developing world (Holzmann et al., 2009; Leisering and Barrientos, 2013).

Moreover, we believe that collective decisions are more appropriately modelled with a government support function, as opposed to majority vote, in the developing world which is hallmarked by various degrees of ‘defective’ democracies or hybrid systems. Government support functions capture a larger variety of political settings than the median voter approach, and have been used for analysing the political economy of pension system previously (Verhoeven and Verbon, 1991; Verbon and Verhoeven, 1992; Kemmerling and Neugart, 2009). The government support function can be thought of as taking into account access of competing interest groups, voter turnout, or societal groups’ political power that even a non-democratic leader seeking office may have to consider. Besides a non-funded pension system and a political support function, our model also features social norms for supporting the old and a rural urban divide in terms of per-capita incomes.

All of these features are related to the scholarly literature in Section 2 in which we will briefly give our empirical motivation for our modelling choices, summarizing recent trends in developing countries, and in which we will then survey the existing literature. We will introduce our model in Section 3 which we will analyze in Section 4. Section 5 will summarize our results and will discuss potential effects of alternative modeling assumptions.

2 Empirical motivation and literature review

Redistributive public pension systems of developing countries grow in importance (Pallares-Miralles et al., 2012; ILO, 2014). As an example, we showed the rise of social pension schemes in the introduction. One implication of these changes are rising coverage rates, defined as the share of those above the statutory retirement age who receive a public pension. Let us now connect this observation with three stylized empirical facts (see Table 1, for its sources see Table 4 in the appendix). First, newly introduced social pension schemes have become particularly popular in middle income countries. In general, coverage rates tend to go up with rising income levels in these countries. Similar things apply if we look at total spending on pension system rises and the generosity of these systems. There are also clear direct indications that these systems are redistributive: replacement

rates for wealthier pensioners tend to be smaller than for poorer and Gini coefficients show the dampening impact of pension systems on income inequality (World Bank, 2011; Pallares-Miralles et al., 2012). All in all, the findings imply that these pension systems correlate with average income in these societies and do not only redistribute between but also within generations. This is our first stylized fact.

The second stylized empirical fact we would like to highlight is related to the first. Not only do these pension systems take off with higher income levels, but they very often address a growing divide between rural, poor areas and zones of urbanization. The International Labor Organization (ILO, 2014, p. 89) shows a huge gap between coverage rates in rural and urban areas. On an anecdotal level, it is well known that earlier pension systems in these countries had an urban bias to the middle and working classes. What seems to be new nowadays is that (social) pensions seek to cover, first and foremost, the rural sector. Only later they are expanded to those parts of the urban sector not hitherto covered by earlier systems. Let us make some important examples. In Mexico, the social pension program *70 y mas* of 2007 was initially designed to cover the rural sector and only later on got expanded to the cities (Flores-Castillo, 2013; Willmore, 2014). The Brazilian Previdencia Rural, originally created in 1971, got substantially expanded after democratization in 1988 (Schwarzer and Querino, 2002) and is a program specifically targeting the rural old. Perhaps most impressively, the Chinese rural pension scheme of 2009 increased coverage of the rural population by more than 89 million people in its first three years, and pushed coverage rates from 30 to 55 percent (Vilela, 2013; Yang et al., 2010).

As an illustration we show that pension coverage and spending are clearly related to the share of the urban population across countries. On the horizontal axis of Figure 2 we plot the share of the urban population for the year 2010. On the vertical axis we plot either coverage rates around 2010 for all pension systems (left panel) or spending on old-age security as percentage of GDP (right panel). In both cases we see a strong and positive bi-variate relationship. Note that a growing divide between rural and urban areas doesn't only imply more people moving to the urban sectors but also a growing economic

Table 1: Characteristics of Pension Systems in Developing Countries in the late 2000s

	Per capita income		
	Low	Lower middle	Upper middle
Countries with a pension system			
with any pension scheme (% of all countries in respective income group)	97	88	86
with social pension (% all countries in respective income group)	21	46	53
Coverage rates (% of elderly)			
of contributory systems	13	27	41
of social pensions	15	46	31
Spending (in % of GDP)			
on all pension systems	1.15	2.51	4.88
on social pensions	0.14	0.72	0.58
Replacement rates (in %)			
Net replacement rate at 100% of average wage	75	79	86
Percentage difference in net replacement rates at 150% and 50% of average wage	0.7	-14	-15
Redistribution (reduction in post-transfer Gini coefficient in %)			
by all old-age security	3.8	8.8	13.7
by social pensions	1.5	0.8	1.1

Notes: The table summarizes some major features of public pension schemes using the World Bank income groups as categories (columns). The exact year of observation changes from country to country, but the modal value is for 2010.

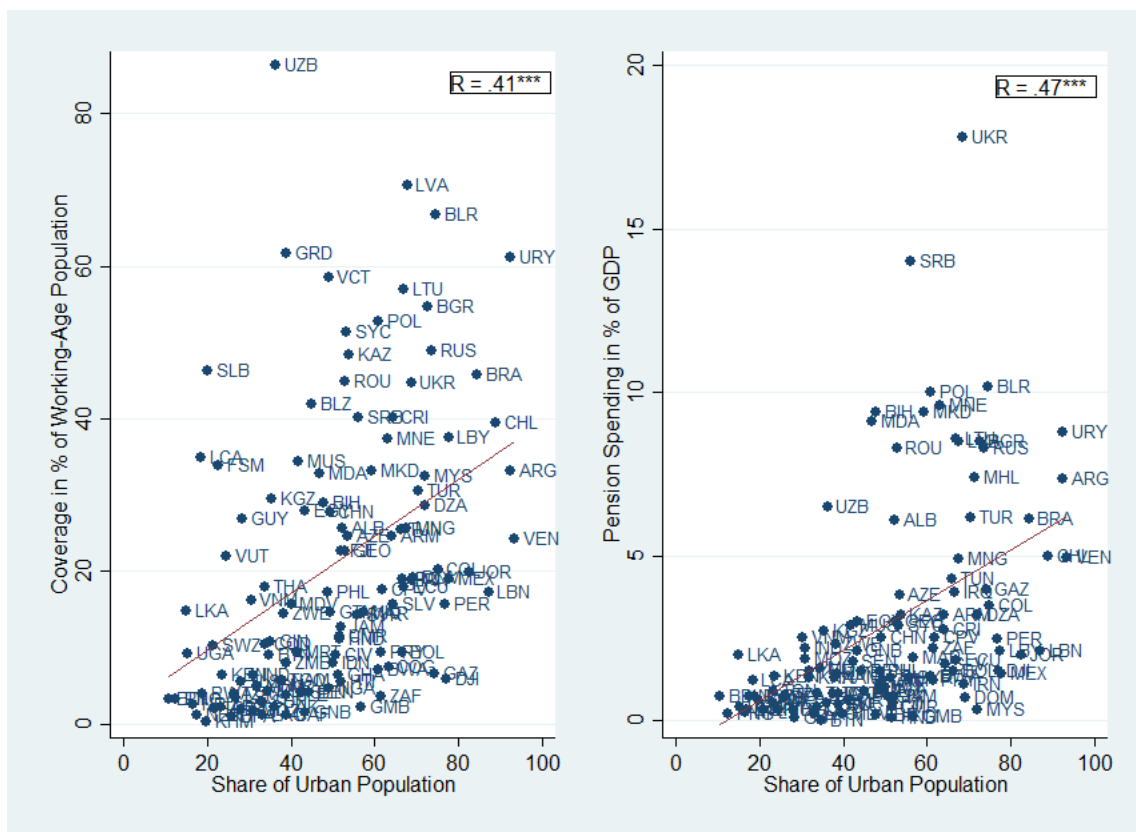


Figure 2: Size of Pension System and Urban Population; see Appendix for data sources.

inequality between the two sectors. Income disparities between rural and urban areas are especially large in middle income countries. Income in typical metropolitan regions is more than five times, in extreme cases even more than ten times above the levels of rural income (Florida, 2014).¹ We believe that this anecdotal evidence gives a strong prior for our theoretical expectation: the rise of recent redistributive pension schemes, and the tremendous boost which they gave to coverage rates, is related to growing inequalities between rural and urban population.

Our third and final stylized empirical fact illustrates the importance of family norms, intra-household transfers and the existence of public safety systems. There is a rich empirical literature illustrating these complex relationships (see below). In some African countries, for instance, more than 90 percent of the elderly population receives direct

¹It is difficult to get reliable cross-country data for the differences between income and wage levels of the urban relative to the rural population in a country. Yet, some proxies such as the productivity gap between agriculture and industry (Bourguignon and Morrisson, 1998; Gollin et al., 2014), or inequality in night-time luminosity (Henderson et al., 2012; Chen and Nordhaus, 2011) point in similar directions across countries and time.

help from their children (World Bank, 1994). Many scholars argue that the importance of traditional family norms, as opposed to more individualized value systems, is still one of the most distinctive differences that sets poorer, developing countries apart from richer countries (e.g. Arrondel and Masson (2006)). As an illustration we show the importance of family norms in developing countries (Figure 3).² On the horizontal axis we see a response item from the World Value Survey (Wave 5). The question was whether respondents think that it should be a major goal in life to ‘make your parents proud’. We admit that this is an imperfect proxy for the size of intra-family transfers from the young to the old, but the variation in the data is consistent with findings in the literature on actual transfers. The figure shows those (strongly) agreeing with the statement as a share of all respondents. The vertical axes are identical to the previous graph. There are strong negative bi-variate relationships, as is to be expected if the social norm is related to the relative importance of informal nets of social security. Based on this it is safe to argue that family norms play a fundamental role in shaping the political economy of pension systems in developing countries.

The three empirical facts show that a political economy model of pension systems in developing countries should reflect the redistributive nature of the system, the role of the (rural-urban) disparities and family norms. The three facts relate to three strands in the literature which are particularly relevant for our purposes: first, the political economy of pension systems; second, the literature on social norms and intra-family transfers; and third, the literature on crowding out between public and private systems.

Starting with the first, the political economy literature on pension systems, we note the overwhelming majority talks about old-age security systems of developed countries (for an exception see Jung and Tran (2012)). This very large literature addresses the question why there is electoral support for a redistributive system in which the contributors outnumber the recipients and where contributors possibly have superior savings devices at hand (for excellent surveys see Breyer (1994), Galasso and Profeta (2002), or De Walque (2005)).

²The figure only shows the values for developing countries, but it has to be noted that richer countries would exhibit much lower values for the social norm. In fact, the difference between the average for rich and developing countries is more than 20 percentage points. In other words, on average slightly less than 70 percent of the population in rich countries agrees with the statement.



Figure 3: Size of Pension System and Social Norms; see Appendix for data sources.

Two more recent contributions address the rise of social security systems in less developed countries but, compared to us, on somehow different grounds. Caucutt et al. (2013) also draw on an urban-rural divide to explain the emergence of the social security system in the U.S. in the early last century. But, rather than pointing towards the role of norms interacting with the urbanization process as we do, they focus on the changing role of land as an old-age savings device to explain the rise of a social security system. In a nutshell, in their contribution it is the devaluation of land driven by the urbanization process which shifted people's preferences from using land to a social security system for securing old-age consumption. Another closely related contribution is by Leroux and Pestieau (2014). Here, agents' investment into children in terms of either more or better educated offsprings partly secures old-age consumption. Their model predicts that as family solidarity crumbles, people prefer pension systems. In contrast to us they model a funded pension system and employ a different political mechanism to derive the collective decision on the pension system. Moreover, we focus on a different rationale that explains

the emergence of pension systems in the developing world: the interplay of an urbanization process that comes with rural-urban productivity differentials and that crowds out the intra-family transfers of the young to the old differently depending on whether they are part of the urban or rural constituency. It is through this channel that we get yet another conflict of interest among the old workers on top of the intra-generational conflict of the young workers, and the inter-generational conflict which usually drives the decision to introduce an pension system.

The second type of literature our article relies upon deals with social norms and income transfers between generations (Arrondel and Masson, 2006; Laferrere and Wolff, 2006). It is well documented that old age security is an important motive for having (large) families (see, e.g., Nugent, 1985) and that there is a strong obligation, especially in less developed countries, for children to provide assistance to elderly parents. The actual motives for intra-family transfers are often a mix of altruism and indirect exchange (Altonji et al., 1996), but they clearly also constitute some type of a social norm binding younger generations. One important stylized fact in this literature is that the nexus between old age security and intra-family transfers is much weaker in advanced economies. Many authors explain this by some sort of modernization, and especially the transformation from rural to urban societies with a high degree of mobility (World Bank, 1994; Folbre and Wolf, 2013). In the formal literature, the role of families norms has found particular resonance in explaining the politics of long-term care arrangements. For instance, Canta and Pestieau (2013) build a model with individuals varying to the degree of whether or not they adhere to the family norm and analyze the implications for the existence of intra-family and market-based forms of care. On basis of this they draw normative conclusions about government interventions. Donder and Leroux (2017) model the individuals' choice between intra-family transfers (in both time and money) and their demand for a long-term care system run by the government. They find that a political majority for such a system is hard to identify, especially if parents prefer family help and the probability of becoming dependent is not too high. These models are very helpful in our context, despite the fact that we model a pension system as a flat rate benefit to all elderly and that people differ

in their wage and savings incomes in our case.

The third strand of literature we integrate into our analysis deals with the question whether formal systems of old-age security crowd out intra-family transfers (see, for a survey, Fafchamps and Cox, 2008). In one of the first studies Cox and Jimenez (1992) find a substantial reduction in intra-family transfers for those families covered by formal pension plans in Peru. More recent evidence on various developing countries reveals partly striking crowding out effects (see Dercon and Krishnan, 2003; Heemskerk et al., 2004; Jensen, 2004; Juarez, 2009; Gerardi and Tsai, 2014). Canta et al. (2016) show in a theoretical model that effects of crowding out are even more complex, if one allows for interactions between family help, private insurance and different types of public insurance systems. In particular, they show that if family is very important, public and private schemes can actually become complements rather than substitutes. This is especially true for rapidly eroding social norms. Our model relates to this literature as the choice of the young generation on the transfers for the elderly will be a function of the scope of the public pension system and their own income.

3 The Model

3.1 General outline of model

We consider an overlapping generations model. People live for two periods t and $t + 1$. They work and save when young. The interest rate on savings is r .³ Labor supply is exogenous. There are two types of young workers: urban with share β and rural with share $(1 - \beta)$, with $0 < \beta < 1$. Urban workers earn a fixed wage w^u and rural workers earn a fixed wage w^r , with $w^u > w^r$. There is population growth denoted with n which we assume to be smaller than the interest rate on savings ($n < r$).⁴ Old workers receive a

³We allow for savings although our context is one of developing countries where financial markets are typically underdeveloped and households' choice set is restricted. One may, however, interpret savings more broadly and include, as is typical for developing countries, contributions to savings associations or investments in livestock (e.g., Morduch et al., 2009; Zeller and Sharma, 2000; Swanepoel et al., 2010). If we set-up a model without a possibility for households to save, we arrive at qualitatively similar results compared to what we show in the following.

⁴We rely on the assumption of a dynamically efficient economy (Aaron, 1966) which has been discussed on theoretical and empirical grounds by Abel et al. (1989), Homburg (1991), or Koethenbueger and

transfer T from their kids. The transfer is driven by a social norm. Non-compliance to the norm is costly. Finally, the political decision is on whether to introduce a pay-as-you-go pension scheme and if so on its generosity.

3.2 Voters

3.2.1 Preferences

Utility of a young (y) representative household member depends on consumption of a private good (c) in periods t and $t + 1$, and on compliance with a social norm:

$$U_t^y = \ln(c_t^y) - \alpha \ln(c_t^y/T) + \gamma \ln(c_{t+1}^y) \quad (1)$$

with current consumption c_t^y and consumption when being old of the currently young c_{t+1}^y . Future payoffs are discounted with $\gamma > 0$, and costs of the social norm are weighted with $0 < \alpha < 1$. The decisions of the young generation are on savings s , the transfer T , and the size of the social security system, i.e. τ .

It is assumed that utility for the young is lower the smaller their transfers to the old are relative to their own consumption when being young. Such a specification implies a social norm according to which one should give to the elderly in relation to one's own well-being. Thus, a young agent's decision to support his parents and by how much is driven by his assessment of what is an appropriate amount given his own level of consumption. What he considers appropriate is a function of the social norm parameterized by α . In this sense our formulation relates to the literature on social comparisons although the reference income or consumption level is not explicitly incorporated in our specification of the utility function.⁵ An advantage of our specification lies in the fact that it is easily interpretable and consistent with the empirical literature (Altonji et al., 1996). Using the properties of the \ln function the first two terms of the utility function can be rewritten as

Poutvaara (2006), among others, and is mostly seen as the more appealing assumption for models on the political economy of social security systems (see Galasso and Profeta, 2002, p. 7).

⁵In the literature on social comparisons one often finds disutility terms by comparing an individual's payoff with the payoff of a reference group either in terms of a difference (see, e.g., Ljungqvist and Uhlig, 2000; Choudhary and Levine, 2006; Pérez-Asenjo, 2011) or in ratio form (see, e.g., Persson, 1995; Corneo, 2002; Goerke and Hillesheim, 2013).

the sum of a young agent's weighted valuations of his own consumption and his transfers to the older generation: $(1 - \alpha)\ln(c_t^y) + \alpha\ln(T)$. The stronger the social norm in society, the larger the weight of transfers and the lower its own consumption. In general, social norms guaranteeing mutual insurance have been given a prominent role for countries in the developing world (see, e.g., the survey by Platteau, 2006). More specifically, it has been put forward as a major enforcement mechanism for inter-generational transfers already by Nugent (1990).

Utility of a currently old (*o*) household member is

$$U_t^o = \ln(c_t^o). \quad (2)$$

The decision of an old household member is only about the size of the social security system (τ) as these voters already made their savings decisions in the prior period.

In (1) and (2) we opted for a specific functional form (the \ln over consumption) as we could otherwise only model the choice for a pension system implicitly depending on the relative size of the social norm (α), urbanization (β), wages (w^i), the interest rate (r), and the population growth rate (n). Being able to explicitly solve for the choice variables, however, allows us to get crispier results.

3.2.2 Budget constraints

Private consumption c_t in the first period of a voter's life satisfies (without superscripts indicating age for convenience)

$$c_t = w^i(1 - \tau) - s - T, \quad (3)$$

with $i = u, r$. Wages net of the contribution τ to the social security system minus private savings s and the transfers T have to equal consumption c_t .

Private consumption in the second period when people are old is

$$c_{t+1} = P + s(1 + r) + (1 + n)T \quad (4)$$

i.e. the sum of the pension payment P , the returns on the private savings, and the transfers T corrected by the rate of population growth as in Buiter (1979) or Burbidge (1983).

3.3 Government

The government imposes a tax τ on labor income out of which pensions P are financed. In any period t we have N_t^o old members of society and N_t^y young members of society. Therefore, a balanced government budget for the pension scheme of the pay-as-you-go type requires

$$\beta N_t^y \tau w^u + (1 - \beta) N_t^y \tau w^r = N_t^o P_t \quad (5)$$

which can be re-written as

$$(n + 1) \tau \bar{w} = P_t \quad (6)$$

with $\bar{w} \equiv \beta w^u + (1 - \beta) w^r$ and $n \equiv \frac{N_t^y - N_t^o}{N_t^o}$. Note that this is a Beveridgean pension system. All members receive the same pension but contribute in relation to their wage income. Consequently, the pension system redistributes within a generation from urban to rural workers. We opt for a model with a Beveridgean pension scheme due to the empirical evidence of the previous section that shows that pensions in the developing world are strongly redistributive. As long as there is sufficient redistribution of resources between urban and rural workers our results should also hold, however, for a scheme in which pensions are partly linked to previous wage income.

4 Choice of pension system

4.1 Voters' policy preferences

Citizens have diverse interests with respect to the social security system. In particular, we have young urban and young rural workers, and voters in the old generation receiving transfers from young family members either earning high urban or relatively low rural wages. In order to derive the preferred pension system for each of these four societal

groups, first, one needs to determine the optimal decisions of the young generation on savings s , and transfers T as functions of the contribution rate τ . In the next step, we determine the optimal choice on τ for each of the four groups.

4.1.1 The young generation

Inserting the consumption budget constraints (3) and (4), and the government's budget constraint (6) in (1), we get for the utility function of the currently young

$$U^y = \ln(w^i(1 - \tau) - s - T) - \alpha \ln\left(\frac{w^i(1 - \tau) - s - T}{T}\right) + \quad (7)$$

$$+ \gamma \ln((n + 1)\tau\bar{w} + s(1 + r) + (1 + n)T).$$

From the first order conditions on savings and transfers ($\frac{\partial U^y}{\partial s} = 0$ and $\frac{\partial U^y}{\partial T} = 0$) we derive optimal savings \hat{s} and transfers \hat{T} as a function of the tax rate τ :

$$\hat{s} = (1 - \tau)w^i \frac{(r - n)\gamma - \alpha(1 + n)}{(1 + \gamma)(r - n)} - (1 + n)\tau\bar{w} \frac{\alpha(1 + n) + (r - n)}{(1 + \gamma)(r - n)(1 + r)} \quad (8)$$

$$\hat{T} = \alpha \frac{(1 + n)\tau\bar{w} + (1 + r)(1 - \tau)w^i}{(1 + \gamma)(r - n)}. \quad (9)$$

Calculating the partial derivatives for the optimal decisions on savings and transfers as voters would face different contribution rates to the pension scheme we get

$$\frac{\partial \hat{s}}{\partial \tau} = -w^i \frac{(r - n)\gamma - \alpha(1 + n)}{(1 + \gamma)(r - n)} - (1 + n)\bar{w} \frac{\alpha(1 + n) + (r - n)}{(1 + \gamma)(r - n)(1 + r)} < 0, \quad (10)$$

$$\frac{\partial \hat{T}}{\partial \tau} = \alpha \frac{(1 + n)\bar{w} - (1 + r)w^i}{(1 + \gamma)(r - n)} \geq 0. \quad (11)$$

Higher tax rates (and thus higher pensions) crowd out private savings of the young generation. As can be seen from (8) for sufficiently high social security contributions, savings may become negative. Essentially, what happens is that the provision of a social security system crowds out the private incentives to provide for old-age security by saving. The effect may become so strong that households do not want to save at all. In order to avoid

negative savings we impose an upper restriction on the social security system $\bar{\tau}$ which we formally derive in the Appendix. We assume that it holds throughout our analysis.

The effect of higher tax rates on the within family transfers from the young to the old is ambiguous. For the young urban generation, which earns $w^u > \bar{w}$, higher contributions to the pension system reduce within family transfers \hat{T} . For the young rural generation, however, the effect of higher tax rates τ is ambiguous. This is the case because the pension scheme can make them better off when being old so that they are willing to give up some of their savings when being young to transfer them to their old family members – but only so if the pension scheme is sufficiently redistributive. The partial effects on the transfers are, of course, stronger the stronger is the social norm α that the young should support the old.

Inserting optimal savings \hat{s} and transfers \hat{T} in (7) and differentiation with respect to the policy rate τ yields after simplifying

$$\frac{\partial U^y}{\partial \tau} = -(1 + \gamma) \frac{w^i(1 + r) - (n + 1)\bar{w}}{w^i(1 - \tau)(1 + r) + (n + 1)\tau\bar{w}}. \quad (12)$$

The sign of the partial derivative is determined by a threshold

$$\omega^y \equiv \frac{1 + n}{1 + r} < 1 \quad (13)$$

such that if $\frac{w^i}{\bar{w}} > \omega^y$ we have $\frac{\partial U^y}{\partial \tau} < 0$, and $\frac{\partial U^y}{\partial \tau} > 0$ if $\frac{w^i}{\bar{w}} < \omega^y$.

The definition of the threshold (13) suggests that the young urban workers do not want to have a pension system at all ($\frac{\partial U^y}{\partial \tau} < 0$) because $\frac{w^u}{\bar{w}} > 1$ and $r > n$ by assumption. The young rural workers, however, want to tax away as much of the income as possible, i.e. $\bar{\tau}$ and put it into the social security system if their own wage w^r relative to the average wage in the economy is sufficiently low. Note, that this may happen even under the assumption of $r > n$ which we apply throughout our analysis. The divided interests of the young rural and urban workers is essentially due to the intra-generational distribution of the pension system which is financed by a tax proportional to the wage w^i and pays the same pension for all old no matter whether they earned the higher urban or the lower

rural wage during their working life.⁶ While these channels are well established in the literature, in our analysis we pin down their consequences for the electoral support of a pension system by the old generation in the presence of transfers driven by a social norm which may be crowded out by a public pension system.

4.1.2 The old generation

Using the budget constraint (4) and pension payments (6) utility of the currently old can be written as

$$U^o = \ln((n+1)\tau\bar{w} + s_{t-1}^o(1+r) + (1+n)\hat{T}). \quad (14)$$

When evaluating their best choice on τ they will take into account the effect on pensions P and the transfer of the young \hat{T} , but not on $s_{t-1}^o \equiv s^o$ as this decision was already taken in the past period. Differentiation with respect to τ gives

$$\frac{\partial U^o}{\partial \tau} = \frac{(n+1)\bar{w} + (1+n)\frac{\partial \hat{T}}{\partial \tau}}{(n+1)\tau\bar{w} + s^o(1+r) + (1+n)\hat{T}}. \quad (15)$$

The comparison of the marginal effect on pension payments and the marginal effect on family transfers as given in the numerator of (15) yields whether the old want to increase pensions. Inserting the marginal effect on optimal transfers $\partial \hat{T} / \partial \tau$ using (11) shows that the sign of (15) is determined by a threshold

$$\omega^o \equiv \frac{(1+\gamma)(r-n) + \alpha(1+n)}{\alpha(1+r)}, \quad (16)$$

such that if $\frac{w^i}{\bar{w}} < \omega^o$ one has $\frac{\partial U^o}{\partial \tau} > 0$, and $\frac{\partial U^o}{\partial \tau} < 0$ for $\frac{w^i}{\bar{w}} > \omega^o$. The definition of the threshold (16) reveals how the existence of a social norm may twist the support of the old generation for a pension system. A typical finding of the political economy literature (see references above) is that the currently old will always support the introduction of a pension system because they will not have to contribute but benefit from pension payments. This result is a special case of equation (16) if we abstract from the social norm and calculate

⁶See Casamatta et al. (2000) and Tabellini (2000) for early contributions recurring to intra-generational effects stemming from different incomes within a generation, and, e.g., Ignacio Conde-Ruiz and Profeta (2007); Koethenbueger et al. (2008) for more recent explorations.

$\lim_{\alpha \rightarrow 0} \omega^o = \infty$. In this case, $\frac{w^i}{w} < \omega^o$ is always fulfilled and we have $\frac{\partial U^o}{\partial \tau} > 0$, i.e. the old want a pension scheme of scope $\bar{\tau}$. There are no transfers paid by the young to the old and, consequently, as no crowding out can occur the old will always support the pension scheme. As the social norm becomes more prevalent and transfers rise, the threshold ω^o declines. Eventually, the threshold will fulfill $\frac{w^u}{w} > \omega^o$, and if it falls even further as α continues to increase the threshold will fulfill $\frac{w^u}{w} > \frac{w^r}{w} > \omega^o$. Essentially, only those old will support the pension system for whom pensions will more than compensate the decline in transfers. First, the old who are supported by the young urban workers will withdraw support as α increases, and for a yet higher social norm also the old who are supported by the young rural workers do not want a social security system.

Summing up, we have four distinct groups in society with potentially different interests with respect to the introduction of a social security system. There are two groups among the elderly whose interests are not necessarily aligned. It may be that all of them want to have a social security system. But it also may be the case that only those who are supported by young rural workers have an interest in a social security system, or that none of the elderly want a social security system. It depends on the extent to which pensions can compensate for the change in private transfers which is driven by the prevalence of the social norm and the distinct wage levels of the young contributors. Furthermore, we found that young urban workers will never want a social security system and young rural workers want one if their wage is sufficiently low.

Given these preferences we may have a first look into a voting outcome if a simple majority rule was applied. A majority vote for the introduction of a social security scheme may arise for two coalitions.⁷ First, the young rural workers may support a pension scheme together with all old workers. Second, a pension scheme may arise based on the support of the rural population only. In both cases, the pension system would have scope $\bar{\tau}$. Having these results in mind helps to better grasp the intuition of our findings based on a more elaborate collective decision rule that takes into account that a large share of political systems in developing countries show elements of defective democracies or hybrid

⁷A formal derivation of the result can be found in the Appendix.

systems which can hardly be described with a majority rule . We turn to a possibly more appropriate collective decision rule now.

4.2 Collective decision

4.2.1 Fixed weights on electoral groups

We assume a government choosing once and for all the size of the social security system τ by maximizing a political support function

$$\max_{\tau} \Lambda_{y,u}U^{y,u} + \Lambda_{y,r}U^{y,r} + \Lambda_{o,u}U^{o,u} + \Lambda_{o,r}U^{o,r} \quad (17)$$

over the payoffs of the four distinct societal groups: young urban (y, u) and rural workers (y, r), the old who receive a transfer from a young urban (o, u) worker, and the old who receive a transfer from a young rural worker (o, r), with $\Lambda > 0$.

Such a government objective function can be rationalized with a probabilistic voting model (see Coughlin and Nitzan (1981), Lindbeck and Weibull (1987) or, for a textbook exposition, Mueller (2003)). Political support functions have been previously used by Verhoeven and Verbon (1991), Verbon and Verhoeven (1992), Grossman and Helpman (1996, 1998), Neugart (2009) or Kemmerling and Neugart (2009). We choose such an approach because it captures a larger variety of political settings than a simple median voter approach. It takes into account that some societal groups may be favored by the government. A government may cater more likely to a group of society the more it can expect that it can provide for the government's survival. Thus, the support of a group of society may be a function of the political system that controls access of competing interest groups, the likelihood that a voter in that group turns out at the ballot, or the group's economic power.

The first order condition for (17) using the partial derivatives (12) and (15) along with the optimal savings and transfers (8), (9) becomes:

$$H = \Lambda_{y,u} \frac{\partial U^{y,u}}{\partial \tau} + \Lambda_{y,r} \frac{\partial U^{y,r}}{\partial \tau} + \Lambda_{o,u} \frac{\partial U^{o,u}}{\partial \tau} + \Lambda_{o,r} \frac{\partial U^{o,r}}{\partial \tau} = 0 \quad (18)$$

with

$$\frac{\partial U^{y,i}}{\partial \tau} = -(1 + \gamma) \frac{w^i(1 + r) - (n + 1)\bar{w}}{w^i(1 - \tau)(1 + r) + (n + 1)\tau\bar{w}},$$

$$\frac{\partial U^{o,i}}{\partial \tau} = \frac{(1 + n)\bar{w}((1 + \gamma)(r - n) + \alpha(1 + n)) - \alpha(1 + n)(1 + r)w^i}{(1 + r)s^o(1 + \gamma)(r - n) + (1 + n)\tau\bar{w}((1 + \gamma)(r - n) + \alpha(1 + n)) + \alpha(1 + n)(1 + r)(1 - \tau)w^i}$$

and $i = u, r$. The first order condition implicitly defines the tax rate $0 \leq \tau^* \leq \bar{\tau}$ which maximizes the support for the government.⁸

We summarize our main results with respect to the political economy of the emergence of social security systems in the following Proposition.

Proposition:

1. There exist population weights $\Lambda_{y,u}$, $\Lambda_{y,r}$, $\Lambda_{o,u}$, and $\Lambda_{o,r}$ for which $0 < \tau^* \leq \bar{\tau}$ holds.
2. Given $0 < \tau^* \leq \bar{\tau}$, the generosity of the pension scheme is related to the wage differential and the social norm in the following way:
 - (a) As urban wages (w^u) rise relative to rural wages the pension scheme becomes more generous ($\frac{d\tau^*}{dw^u} > 0$) if the government weighs the rural population sufficiently strongly relative to the urban population.
 - (b) An eroding social norm increases the generosity of the pension scheme.

Proof: see Appendix

We already showed that the rural and urban workers as well as the elderly receiving transfers by an urban or rural worker, respectively, have diverging interests on the size of the pension system. The pension system redistributes income away from the young urban workers. Under no circumstances they support a pension scheme. If the pension system is sufficiently redistributive, young rural workers gain from the social security system because their contribution to the pension system is relatively small due to their lower

⁸In the Appendix we show that the second order condition for a maximum is fulfilled.

income. In contrast, the payment they will receive is based on all workers' contributions including those of the high income workers. Whether or not the old generation wants a pension system depends on the distribution of wages, the prevalence of the social norm, the population growth rate, the discount rate, and the return on savings as described by the condition $\frac{w^i}{w} \leq \omega^o$ (see (16)). Essentially, the old trade off additional income from the pension system with a reduction of the family transfers. The other source of old-age income, informal transfers, will decrease with a more generous pension scheme since the young will reduce their payments when being subject to a higher contribution rate to the public pension system. The government balances out the effects on the well-being of the four electoral groups when it decides on the size of a pension system.

The Proposition also answers the question on how the government would change the generosity of pensions to secure maximum electoral support as the following changes occur: urban wages w^u increase or the social norm α becomes less important. How the government responds is basically a question of the effect of these parameter changes on the marginal utility of a tax change on any of these four groups of society and the corresponding weights with which the payoffs are credited by the government.

For the rural population the marginal utility of a larger pension scheme increases as the wages of the young urban workers rise. Contrarily, the marginal utility of a larger pension scheme may decrease for the urban population as urban wages rise. Thus, if the government puts a high enough weight on the electoral support of the rural population relative to the urban population, an increase in w^u will always increase the size of the pension scheme. Now, the pension scheme redistributes even more within the young generation which makes the young rural workers more in favor of increasing its scope. Moreover, the currently old receiving transfers from the young rural workers profit from more generous pensions without being too much punished by decreasing family transfers. For the elderly who receive transfers from young urban workers, however, the crowding out of transfers may be so strong that the effect of a larger pension on their marginal utility is weakened.

As we look into changes of the social norm, we find that the weights on the young are

irrelevant for the government's choice. At the margin, a change in the social norm does not alter the marginal effect of a change in the contribution rate on utility because the young internalize the implied change this will have on the old age transfers they would get. Moreover, the marginal effects on young- and old-age consumption, and costs arising from the social norm outweigh each other. As the strength of the social norm decreases the only factor influencing the government's choice on the size of the pension system is the effect it has on the fate of the old voters. It turns out that the interests of the old rural and old urban voters are aligned. As a result, the erosion of the social norm leads to a larger pension scheme.

4.2.2 Population weights on electoral groups

In the analytical discussion of our results we assumed that the weights of the electoral support function are arbitrarily chosen positive numbers. One may ask how results look like if the weights in the electoral support function reflect the relative size of the population. The underlying assumption for the evaluation of such a case would then be that electoral groups influence the government only according to their population size. Groups of a given size are not able in this special case to organize their interests more powerful than other groups of the same size. A government support function arising from such assumptions would then look like

$$G = \beta(1+n)U^{y,u}(\tau, \hat{s}, \hat{T}) + (1-\beta)(1+n)U^{y,r}(\tau, \hat{s}, \hat{T}) \quad (19)$$

$$+\beta U^{o,u}(\tau, \hat{s}, \hat{T}) + (1-\beta)U^{o,r}(\tau, \hat{s}, \hat{T}).$$

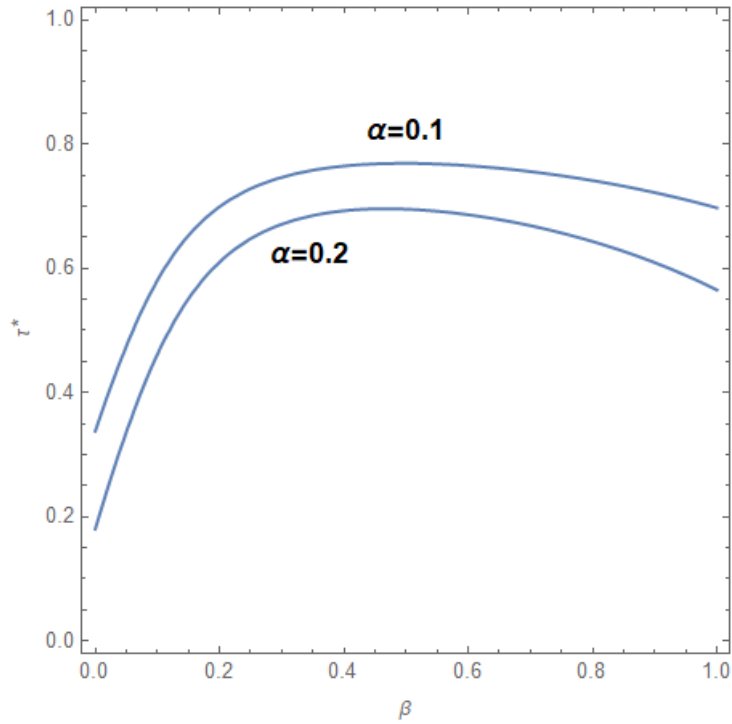
The share of the urban population β and the growth rate n determine the weights. One can straightforwardly apply the reasoning of the proof of our main Proposition to analytically show that also in this specification a social security system with $0 \leq \tau^* \leq \bar{\tau}$ exists. No clear-cut analytical results, however, can be derived for a partial effect of an increase in the urban population on the size of the pension scheme. There are various countervailing effects. Although the young urban generation is not in favor of a public pension scheme

having more urban workers lessens the burden of a more generous pension scheme as, *ceteris paribus*, more workers contribute who earn high wages. Moreover, as the tax base increases, a marginal increase in the contribution rate will also raise the utility of the young rural workers because the intra-generational distribution of the pension scheme is magnified. Finally, a larger tax base also strengthens the inter-generational redistribution of the pension scheme favoring the elderly so that the government can increase its political support by proposing a larger public pension scheme. However, given that a larger urban population share increases the weight on the urban population while it decreases it for the rural population, the sign of $\frac{d\tau^*}{d\beta}$ is ambiguous.

Therefore, we turn to a numerical evaluation of how the optimal policy rate τ^* of a government changes with respect to the various parameters of interest. In Figure 4 we plot the size τ^* of the pension system that maximizes the electoral support to the government as a function of the share of the urban population and two different parameter values for the social norm ($\alpha = 0.1$ and $\alpha = 0.2$).

Holding all else constant, Figures 4 illustrates that a government increases the size of the pension system as the share of urban population and those older citizens supported by the urban young increases. This is, at least true for relatively low urban population shares irrespective of the prevalence of the social norm. Given that for the high wage urban population public pensions are an inferior savings device, support for the government flattens out as the share of the urban population increases. Moreover, the decline in support for the government is slightly stronger when the social norm to support the elderly is larger because there is more crowding out of transfers to the old population supported by the urban young which would make those old worse off if pensions increased even further. What is more, for any given share of the urban population, a lower prevalence of the social norm lets governments increase the size of the pension scheme. All in all, the numerical examples underpin the analytical results derived from a model with fixed weights in the government support function.

Figure 4: Numerical example



Notes: Size of pension system τ^* over share of urban population; parameter values are $w^u = 3$, $\gamma = 0.61$, $n = 0$, $r = 0.64$, $w^r = 1$, $s^o = 0.5$. The choices of γ and r reflect the fact that pension payments accrue upon retirement. For the numerical example, we assume that this occurs after 25 years with a yearly interest rate of 2%. Thus, a unit saved today will be worth 64% more upon retirement. Similarly, the (compound) discount rate is $\gamma = 0.61$. The wage differential is chosen based on the evidence presented in Section 2.

5 Discussion and conclusions

In the developing world redistributive public pension systems grow in size, which prompts the question why this happens. Empirically we observe that these pension systems are mainly of the Beveridgean type, i.e. they do not only redistribute between generations but also within generations. We also see that many of these programs explicitly target the rural old. The phenomenon that these pension schemes are redistributive is different from what nowadays developed countries installed when they introduced social security systems for the first time, typically more than a hundred years ago, and also different from often short-lived attempts to introduce (non-redistributive) private account systems in developing countries (see e.g. authors 2017).

Using stylized empirical facts, we observe that developing countries with a larger share of the population residing in urban areas have larger pension schemes. As, in addition, per-capita income in urban areas is manifold of what can be earned in rural areas, political support for an increase in a pension system may come about as the share of urban citizens increases, urban to rural income gaps rise, or the social norm driving transfers to the elderly erodes. In our theoretical analysis support for a more generous pension system comes from young and low income rural workers to whom the pension system redistributes, and the old generation up to the point where an increase in the public pensions is not overcompensated by a decrease in the norm-driven family transfers. In societies characterized by an urbanization process, these social norms may, however, erode and curb the family transfers making it more likely that eventually the old will support a pension scheme. We show under which conditions governments may introduce pension schemes and increase its generosity.

As usual, for deriving our results, we had to agree on a range of potentially critical assumptions. We would like to take up some of these and discuss to which extent alternative ones might have changed our findings, or could be taken up in future research. First, we abstracted from fertility issues. Our conjecture is that different fertility rates between the rural and the urban population may change the analysis of the support of a pension scheme in a sense that the redistributive feature between the urban and the rural popu-

lation could be reinforced. Similar consequences may arise if one would look deeper into differing life expectancies for the rural and urban sector. Second, our analysis grounds on the assumption that the elderly can be sure that there is a transfer from the young generation. While we varied the strength of the social norm that ultimately determines the size of the transfer, another interesting extension could be to allow for uncertainty with respect to whether the young will actually support the elderly. Third, our model could also be applied to the formal vs. informal divide (Carnes and Mares, 2014). A richer model would then include the probability of the young in the cities finding a job in the formal sector and hence paying payroll taxes. However, we find clearer evidence for differences between the rural and urban sector than for the formal vs. the informal sector: Wage differentials between informal and formal voters are not as straightforward, as the informal sector is very heterogeneous. Fourth, we chose a pay-as-you-go type pension system in light of recent empirical trends. Our model would be less suitable to explain other types of pension reforms such as the introduction of private pension accounts which often exhibit features of special-interest politics (e.g. Kemmerling and Neugart (2009)). Finally, the social norm itself may be responsive to the pension system. Once the new system is in place, the norm to support the elderly may become less prevalent. We think that this would rather accelerate the dynamics we have described.

As mentioned above, testing our model faces serious empirical challenges. Whereas data on urbanization and agglomeration has improved over the last decades even for poor countries, data on productivity and income differentials between rural and urban areas are subject to fundamental problems of availability and quality (Jerven, 2013). However, we are optimistic that such data will be available in the future (Florida, 2014). Better data would allow us to tease out the role of redistributive conflicts between different groups of the population in developing countries. This data would also allow us then, to compare the predictive power of our model compared to alternative explanations as, for example, those based on the growing importance of the informal sector or those based on the relevance of the demand for insurance. Finally, we fully acknowledge that the social pension systems are not the only systems in developed countries. This opens a new and

barely touched research agenda: When and why do governments opt for different forms of providing social security for the elderly in less developed countries?

Acknowledgments

We would like to thank the participants at the research seminar at Philipps-Universität Marburg, the European Public Choice conference in Groningen, the conference of the Verein für Socialpolitik in Münster, the WINIR conference in Rio de Janeiro, and the European Political Science Association conference in Vienna for their helpful suggestions. We are also extremely grateful for the comments of the editor and two anonymous reviewers.

Appendix

Condition on positive savings and consumption of the young

Savings are positive if

$$\hat{s} = (1 - \tau)w^i \frac{(r - n)\gamma - \alpha(1 + n)}{(1 + \gamma)(r - n)} - (1 + n)\tau\bar{w} \frac{\alpha(1 + n) + (r - n)}{(1 + \gamma)(r - n)(1 + r)} > 0 \quad (20)$$

which can be re-arranged as

$$\frac{1}{\frac{\bar{w}(1+n)(\alpha(1+n)+(r-n))}{w^i((r-n)\gamma-\alpha(1+n))(1+r)} + 1} > \tau. \quad (21)$$

Inequality (21) defines an upper bound on the pension scheme $\bar{\tau}$, for $w^i = w^r$, under which savings are always positive. We assume that this holds throughout our analysis.

Furthermore, consumption of the young is positive if

$$w^i(1 - \tau) - \hat{s} - \hat{T} = \frac{1 - \alpha}{1 + \gamma} \left(w^i(1 - \tau) + \tau\bar{w} \frac{1 + n}{1 + r} \right) > 0, \quad (22)$$

see (3), (8), and (9). For $\alpha < 1$ consumption is always positive. Intuitively this can be explained by rearranging the utility function for the consumption of the young when being

young: $\ln(c_t^y) - \alpha \ln(c_t^y/T) = (1 - \alpha) \ln(c_t^y) + \alpha \ln(T)$. To guarantee a positive weight on c_t^y one needs that they value it at all, i.e. $\alpha < 1$.

Majority vote

For the determination of the majority vote we proceed in two steps. First we derive possibly supporting coalitions for the introduction of a pension scheme. Second we check whether a majority may arise.

Table 2 summarizes all combinations of coalitions arising from the different electoral groups, i.e. young rural workers (y, r) , old rural voters (o, r) , and old urban voters (o, u) for the support of a pension scheme. (Remember, that the young urban workers never want a pension scheme.) In principle, eight combinations exist. In each case we denote the support of a group with 1, and use 0 if the group does not support a pension scheme.

Not all of the combinations are, however, feasible given the preferences of the electoral groups. We may exclude several of them using our productivity thresholds that were defined as

$$\frac{w^r}{\bar{w}} < \omega^y = \frac{1+n}{1+r} \quad (23)$$

and

$$\frac{w^i}{\bar{w}} < \omega^o = \frac{(1+\gamma)(r-n) + \alpha(1+n)}{\alpha(1+r)} \quad (24)$$

with $\bar{w} = \beta w^u + (1 - \beta)w^r$. Remember that $\frac{w^r}{\bar{w}} < \omega^y$ implies that the young rural workers want a pension scheme, and that $\frac{w^r}{\bar{w}} < \omega^o$ and $\frac{w^u}{\bar{w}} < \omega^o$ imply that the old rural and old urban voters want a pension scheme, respectively. Note also that $\omega^o > \omega^y$. If we have $\frac{w^r}{\bar{w}} < \omega^y$, i.e. the young rural workers support a pension scheme, it must also hold that $\frac{w^r}{\bar{w}} < \omega^o$, i.e. the old rural voters support it. It follows that combinations 5 and 6 are not feasible. If we have $\frac{w^u}{\bar{w}} < \omega^o$, i.e. the old urban voters support a pension scheme, it is also true that $\frac{w^r}{\bar{w}} < \omega^o$, i.e. the old rural voters support a pension scheme. It follows that combination 2 is also impossible.

We proceed by deriving for each of the remaining combinations the conditions under

Table 2: Coalitions supporting a pension scheme

	y, r	o, r	o, u	
1	0	0	0	
2	0	0	1	not possible
3	0	1	0	
4	0	1	1	
5	1	0	0	not possible
6	1	0	1	not possible
7	1	1	0	
8	1	1	1	

which a majority for the introduction of a pension scheme would arise. Note, that the total population becomes for $N^o = 1$

$$\beta N^y + (1 - \beta)N^y + \beta N^o + (1 - \beta)N^o = 2 + n. \quad (25)$$

Considering the remaining cases listed in Table 2, we get:

- 3: We would have to have

$$(1 - \beta) > \frac{1}{2}(2 + n) \quad (26)$$

$$-\beta > \frac{1}{2}n \quad (27)$$

which is never fulfilled.

- 4: We would have to have

$$\beta + (1 - \beta) > \frac{1}{2}(2 + n) \quad (28)$$

$$0 > \frac{1}{2}n \quad (29)$$

which is never fulfilled

- 7: We would have to have

$$(1 - \beta)(1 + n) + (1 - \beta) > \frac{1}{2}(2 + n) \quad (30)$$

$$(1 - \beta)(2 + n) > \frac{1}{2}(2 + n) \quad (31)$$

$$(1 - \beta) > \frac{1}{2} \quad (32)$$

$$\frac{1}{2} > \beta \quad (33)$$

which can be fulfilled.

- 8: We would have to have

$$(1 - \beta)(1 + n) + \beta + (1 - \beta) > \frac{1}{2}(2 + n) \quad (34)$$

$$\frac{1 + \frac{1}{2}n}{1 + n} > \beta \quad (35)$$

which can be fulfilled.

Thus, a majority vote for the introduction of a pension scheme may arise for two reasons: the rural population wants it and is sufficiently large ($\frac{1}{2} > \beta$), or the rural population together with the old urban population want it and the coalition is sufficiently large ($\frac{1 + \frac{1}{2}n}{1 + n} > \beta$). Empirically speaking, the latter is the more realistic.

Second order condition

For having a maximum, we need to show that

$$\frac{dH}{d\tau} < 0. \quad (36)$$

Calculating the derivatives for (18) term by term gives

$$\frac{d}{d\tau} \left(-(1 + \gamma) \frac{w^i(1 + r) - (n + 1)\bar{w}}{w^i(1 - \tau)(1 + r) + (n + 1)\tau\bar{w}} \right) = \quad (37)$$

$$= -(1 + \gamma) \frac{(w^i(1 + r) - (n + 1)\bar{w})^2}{(w^i(1 - \tau)(1 + r) + (n + 1)\tau\bar{w})^2} < 0 \quad (38)$$

Now, let us look into the derivative of

$$\begin{aligned} \frac{d}{d\tau} \left(\frac{(1+n)\bar{w}((1+\gamma)(r-n) + \alpha(1+n)) - \alpha(1+n)(1+r)w^i}{(1+r)s^o(1+\gamma)(r-n) + (1+n)\tau\bar{w}((1+\gamma)(r-n) + \alpha(1+n)) + \alpha(1+n)(1+r)(1-\tau)w^i} \right) &= \\ &= - \frac{\left((1+n)\bar{w} + \frac{\alpha(1+n)((1+r)w^i - (1+n)\bar{w})}{(1+\gamma)(n-r)} \right)^2}{\left((1+r)s^o(1+n)\tau\bar{w} + \frac{\alpha(1+n)(-(1+r)(1-\tau)w^i - (1+n)\tau\bar{w})}{(1+\gamma)(n-r)} \right)^2} < 0 \end{aligned} \quad (39)$$

$$\quad (40)$$

Therefore, we have

$$\frac{dH}{d\tau} < 0. \quad (41)$$

Proof of main results

Government installing a pension system

A government will introduce a pension system, i.e. $0 < \tau^* \leq \bar{\tau}$, if $H(\tau^*) = 0$ exists:

$$\begin{aligned} H(\tau^*) &= \Lambda_{y,u} \frac{\partial U^{y,u}}{\partial \tau} \Big|_{\tau^*} + \Lambda_{y,r} \frac{\partial U^{y,r}}{\partial \tau} \Big|_{\tau^*} + \\ &\quad \Lambda_{o,u} \frac{\partial U^{o,u}}{\partial \tau} \Big|_{\tau^*} + \Lambda_{o,r} \frac{\partial U^{o,r}}{\partial \tau} \Big|_{\tau^*} = 0 \end{aligned} \quad (42)$$

From the discussion of the partial effects of τ on the marginal utilities of the four groups – see (12) and (15) in the main text – we already know that the first term is always negative while the sign of the remaining three terms is ambiguous. Table ?? shows that there are four possibilities for which there is at least one societal group that supports the introduction of a pension scheme. As there is no upper restriction on the weights $\Lambda_{h,i}$ with $h = y, o$ and $i = u, r$. one can always increase the weights on the supporting electoral groups up to a level where (42) holds, given any weights for the non-supporting electoral groups.

Effect of w^u

An increase in w^u raises τ^* if

$$\frac{d\tau^*}{dw^u} = -\frac{\frac{\partial H}{\partial w^u}}{\frac{\partial H}{\partial \tau^*}} > 0. \quad (43)$$

Given that $\frac{\partial H}{\partial \tau^*} < 0$ we need to check if $\frac{\partial H}{\partial w^u} > 0$. Again, doing this term by term for each of the societal groups gives:

$$\frac{\partial^2 U^{y,u}(\tau, \hat{s}, \hat{T})}{\partial \tau \partial w^u} < 0, \quad (44)$$

$$\frac{\partial^2 U^{y,r}(\tau, \hat{s}, \hat{T})}{\partial \tau \partial w^u} > 0, \quad (45)$$

$$\frac{\partial^2 U^{o,u}(\tau, \hat{s}, \hat{T})}{\partial \tau \partial w^u} \leq 0, \quad (46)$$

and

$$\frac{\partial^2 U^{o,r}(\tau, \hat{s}, \hat{T})}{\partial \tau \partial w^u} > 0. \quad (47)$$

The effect of an increase in urban wages on the marginal effect on utility of a larger pension scheme differs between societal groups. For the young rural workers and the old population receiving transfers from rural workers the marginal effect of a larger pension scheme on utility is amplified with higher urban wages. For the young urban workers the effect is negative, and for the old receiving transfers from the young urban workers it is ambiguous.

Effect of α

Let us turn to the effect of a change in the social norm α on τ^* and ask if

$$\frac{d\tau^*}{d\alpha} = -\frac{\frac{\partial H}{\partial \alpha}}{\frac{\partial H}{\partial \tau^*}} < 0. \quad (48)$$

Given that $\frac{\partial H}{\partial \tau^*} < 0$ we need to check if $\frac{\partial H}{\partial \alpha} < 0$. Again, looking into the effects on each of the societal groups separately, reveals that the cross-derivatives for the young workers are zero because the marginal effects on the utility of the young workers arising from an increase in the scope of the public pension scheme are independent from the social norm. Doing it term by term for the old population gives:

$$\frac{\partial^2 U^{o,u}(\tau, \hat{s}, \hat{T})}{d\tau \partial \alpha} < 0 \quad (49)$$

and

$$\frac{\partial^2 U^{o,r}(\tau, \hat{s}, \hat{T})}{d\tau \partial \alpha} < 0. \quad (50)$$

A decrease in the social norm amplifies the positive effect of a marginal tax increase on the utility of the old population.

Partial effects on scope of pension system

Collecting all the marginal effects on the four groups of society as shown in Table 3 it is straightforward to derive the partial effects on the scope of the pension system.⁹ If the weights on the rural population ($\Lambda_{y,r}$ and $\Lambda_{o,r}$) are sufficiently larger than the weights on the urban population ($\Lambda_{y,u}$ and $\Lambda_{o,u}$) the partial effect is $\partial H / \partial w^u > 0$ which implies $d\tau^* / dw^u > 0$. Furthermore, one gets $d\tau^* / d\alpha < 0$.

Table 3: Effects of changes in w^u and α on the marginal effect of a tax increase on the utility of the four societal groups

	Societal group			
	y,u	y,r	o,u	o,r
w^u	-	+	+/-	+
α	0	0	-	-

Selected Indicators on Pension Systems

⁹A detailed derivation of the partial effects is available upon request.

Table 4: Summary Statistics and Sources

Variable	Num. obs.	Mean	St. dev.	Source
Year of introducing social pension	80	1981	31	ILO (2014)
Coverage of social pensions (in % elderly population)	85	33	35	ILO (2014)
Coverage of contributory systems (in % of elderly population) in 2010	154	52	38	ILO (2014)
Coverage of contributory systems (in % of elderly population) in 2000	154	42	39	ILO (2014)
Net replacement rate at 100% of average wage	61	70	26	Pallares-Miralles et al. (2012)
Net replacement rate at 50% of average wage	61	101	174	Pallares-Miralles et al. (2012)
Net replacement rate at 150% of average wage	61	66	26	Pallares-Miralles et al. (2012)
Spending on all Pension Systems in % of GDP	145	3.94	3.88	Pallares-Miralles et al. (2012)
Spending on social pensions in % of GDP	53	0.55	0.71	ILO (2014)
Redistribution all old age (reduction in Gini in %)	36	10	13	World Bank (2011)
Redistribution social pensions (reduction in Gini in %)	16	1	1	World Bank (2011)
Urban population (in % of total population)	201	56	24	United Nations (2014)
'Make Parents Proud?'	55	81	16	World Value Surveys (2009)

References

- AARON, H. (1966): “The social insurance paradox,” *Canadian Journal of Economics and Political Science/Revue canadienne de economiques et science politique*, 32, 371–374.
- ABEL, A. B., N. G. MANKIW, L. H. SUMMERS, AND R. J. ZECKHAUSER (1989): “Assessing dynamic efficiency: theory and evidence,” *The Review of Economic Studies*, 56, 1–19.
- ALTONJI, J. G., F. HAYASI, AND L. KOTLIKOFF (1996): “The Effects of Income and Wealth on Time and Money Transfers Between Parents and Children,” *NBER Working Paper Series*, No. 5522.
- ARRONDEL, L. AND L. MASSON (2006): “Altruism, exchange or indirect reciprocity: What do the data on family transfers show?” in *Handbook of Economics of Giving, Altruism, and Reciprocity, Volume 2*, ed. by S.-C. Kolm and J. M. Ythier, North-Holland: Elsevier, 971–1052.
- BOURGUIGNON, F. AND C. MORRISSON (1998): “Inequality and development: the role of dualism,” *Journal of Development Economics*, 57, 233–257.
- BREYER, F. (1994): “The political economy of intergenerational redistribution,” *European Journal of Political Economy*, 10, 61–84.
- BUTER, W. H. (1979): “Government finance in an overlapping generations model with gifts and bequests,” in *Social security versus private saving*, ed. by G. M. von Furstenberg, Cambridge, MA: Ballinger, 395–429.
- BURBIDGE, J. B. (1983): “Government debt in an overlapping-generations model with bequests and gifts,” *American Economic Review*, 73, 222–227.
- CANTA, C. AND P. PESTIEAU (2013): “Long-term care insurance and family norms,” *The B.E. Journal of Economic Analysis & Policy*, 14, 401–428.

- CANTA, C., P. PESTIEAU, AND E. THIBAUT (2016): “Long-term care and capital accumulation: the impact of the state, the market and the family,” *Economic Theory*, 61, 755–785.
- CARNES, M. AND I. MARES (2014): “Coalitional realignment and the adoption of non-contributory social insurance programs in Latin America,” *Socio-Economic Review*, 12, 695–722.
- CASAMATTA, G., H. CREMER, AND P. PESTIEAU (2000): “The political economy of social security,” *The Scandinavian Journal of Economics*, 102, 503–522.
- CAUCUTT, E. M., T. F. COOLEY, AND N. GUNER (2013): “The farm, the city, and the emergence of social security,” *Journal of Economic Growth*, 18, 1–32.
- CHEN, X. AND W. D. NORDHAUS (2011): “Using luminosity data as a proxy for economic statistics,” *Proceedings of the National Academy of Sciences of the United States of America*, 108, 8589–8594.
- CHOUHARY, M. A. AND P. LEVINE (2006): “Idle worship,” *Economics Letters*, 90, 77–83.
- CORNEO, G. (2002): “The efficient side of progressive income taxation,” *European Economic Review*, 46, 1359–1368.
- COUGHLIN, P. AND S. NITZAN (1981): “Electoral outcomes with probabilistic voting and Nash social welfare maxima,” *Journal of Public Economics*, 15, 113–121.
- COX, D. AND E. JIMENEZ (1992): “Social security and private transfers in developing countries: The case of Peru,” *World Bank Economic Review*, 6, 155–169.
- DE WALQUE, G. (2005): “Voting on pensions: A survey,” *Journal of Economic Surveys*, 19, 181–209.
- DERCON, S. AND P. KRISHNAN (2003): “Risk sharing and public transfers,” *The Economic Journal*, 113, C86–C94.

- DONDER, P. D. AND M.-L. LEROUX (2017): “The political choice of social long term care transfers when family gives time and money,” *Social Choice and Welfare*, 49, 755–786.
- FAFCHAMPS, M. AND D. COX (2008): “Extended Family and Kinship Networks: Economic Insights and Evolutionary Directions,” in *Handbook of Development Economics*, Vol. 4, ed. by H. Chenery and T. Srinivasan, Amsterdam et al.: Elsevier, vol. 4, 3712–3784.
- FLORES-CASTILLO, A. (2013): “Transferencias no contributivas a personas mayores. Analisis comparativo de dos programas: 70 y mas y pension alimentaria de la Ciudad de Mexico,” *CEPAL Mexico - Naciones Unidas*, LC/MEX/L.1125.
- FLORIDA, R. (2014): “Why big cities matter in the developing world,” *The Atlantic Cities*, Jan 14.
- FOLBRE, N. AND D. WOLF (2013): “The intergenerational welfare state,” *Population and Development Review*, 38, 36–51.
- GALASSO, V. AND P. PROFETA (2002): “The political economy of social security: A survey,” *European Journal of Political Economy*, 18, 1–29.
- GERARDI, K. AND Y. TSAI (2014): “The effect of social entitlement programmes on private transfers: New evidence of crowding out,” *Economica*, 81, 721–746.
- GOERKE, L. AND I. HILLESHEIM (2013): “Relative consumption, working time, and trade unions,” *Labour Economics*, 24, 170–179.
- GOLLIN, D., D. LAGAKOS, AND M. E. WAUGH (2014): “The agricultural productivity gap,” *The Quarterly Journal of Economics*, 129, 939–993.
- GROSSMAN, G. M. AND E. HELPMAN (1996): “Electoral competition and special interest politics,” *The Review of Economic Studies*, 63, 265–286.
- (1998): “Intergenerational redistribution with short-lived governments,” *The Economic Journal*, 108, 1299–1329.

- HEEMSKERK, M., A. NORTON, AND L. D. DEHN (2004): “Does public welfare crowd out informal safety nets? Ethnographic evidence from rural Latin America,” *World Development*, 32, 941–955.
- HENDERSON, J. V., A. STOREYGARD, AND D. N. WEIL (2012): “Measuring economic growth from outer space,” *American Economic Review*, 102, 994–1028.
- HOLZMANN, R., D. A. ROBALINO, AND N. TAKAYAMA, eds. (2009): *Closing the Coverage Gap. The Role of Social Pensions and other Retirement Income Transfers*, Washington: World Bank.
- HOMBURG, S. (1991): “Interest and growth in an economy with land,” *Canadian Journal of Economics*, 24, 450–459.
- IGNACIO CONDE-RUIZ, J. AND P. PROFETA (2007): “The redistributive design of social security systems,” *The Economic Journal*, 117, 686–712.
- ILO (2014): *Social Protection Report 2014/5*, Geneva: International Labour Office.
- JENSEN, R. T. (2004): “Do private transfers ‘displace’ the benefits of public transfers? Evidence from South Africa,” *Journal of Public Economics*, 88, 89–112.
- JERVEN, M. (2013): *Poor Numbers: How We Are Misled By African Development Statistics And What To Do About It*, Ithaca London: Cornell University Press.
- JUAREZ, L. (2009): “Crowding out of private support to the elderly: Evidence from a demogrant in Mexico,” *Journal of Public Economics*, 93, 454–463.
- JUNG, J. AND C. TRAN (2012): “The extension of social security coverage in developing countries,” *Journal of Development Economics*, 99, 439–458.
- KEMMERLING, A. AND M. NEUGART (2009): “Financial market lobbies and pension reform,” *European Journal of Political Economy*, 25, 163–173.
- KOETHENBUERGER, M. AND P. POUTVAARA (2006): “Social security reform and investment in education: Is there scope for a Pareto improvement?” *Economica*, 73, 299–319.

- KOETHENBUERGER, M., P. POUTVAARA, AND P. PROFETA (2008): “Why are more redistributive social security systems smaller? A median voter approach,” *Oxford Economic Papers*, 60, 275–292.
- LAFERRERE, A. AND F.-C. WOLFF (2006): “Microeconomic models of family transfers,” in *Handbook of Economics of Giving, Altruism, and Reciprocity, Volume 2*, ed. by S.-C. Kolm and J. M. Ythier, North-Holland: Elsevier, 889–969.
- LEISERING, L. AND A. BARRIENTOS (2013): “Social citizenship for the global poor? The worldwide spread of social assistance,” *International Journal of Social Welfare*, 22, 50–67.
- LEROUX, M.-L. AND P. PESTIEAU (2014): “Social security and family support,” *Canadian Journal of Economics/Revue canadienne d'économique*, 47, 115–143.
- LINDBECK, A. AND J. W. WEIBULL (1987): “Balanced-budget redistribution as the outcome of political competition,” *Public Choice*, 52, 273–297.
- LINDERT, P. H. (1994): “The rise of social spending, 1880-1930,” *Explorations in Economic History*, 31, 1–37.
- LJUNGQVIST, L. AND H. UHLIG (2000): “Tax policy and aggregate demand management under catching up with the Joneses,” *The American Economic Review*, 356–366.
- MORDUCH, J., D. COLLINS, O. RUTHVEN, AND S. RUTHERFORD (2009): *Portfolios of the Poor. How the World's Poor Live on Two Dollars a Day*, Princeton University Press.
- MUELLER, D. C. (2003): *Public choice III*, New York et al: MIT Press.
- NEUGART, M. (2009): “Pensions with early retirement and without commitment,” *Applied Economics Letters*, 16, 257–260.
- NUGENT, J. B. (1985): “The old-age security motive for fertility,” *Population and Development Review*, 11, 75–97.

- (1990): “Old age security and the defense of social norms,” *Journal of Cross-Cultural Gerontology*, 5, 243–254.
- PALLARES-MIRALLES, M., C. ROMERO, AND E. WHITEHOUSE (2012): “International Patterns of Pension Provision II,” *The World Bank Social Protection And Labour Discussion Papers*, No. 1211.
- PÉREZ-ASENJO, E. (2011): “If happiness is relative, against whom do we compare ourselves? Implications for labour supply,” *Journal of Population Economics*, 24, 1411–1442.
- PEROTTI, E. AND A. SCHWIENBACHER (2009): “The political origin of pension funding,” *Journal of Financial Intermediation*, 18, 384–404.
- PERSSON, M. (1995): “Why are taxes so high in egalitarian societies?” *The Scandinavian Journal of Economics*, 569–580.
- PLATTEAU, J.-P. (2006): “Solidarity norms and institutions in village societies: Static and dynamic considerations,” in *Handbook of the Economics of Giving, Altruism and Reciprocity*, ed. by S.-C. Kolm and J. M. Ythier, Elsevier, vol. 1, 819–886.
- SCHWARZER, H. AND A. C. QUERINO (2002): “Non-contributory pensions in Brazil: Their impact on poverty reduction,” *ILO Extension of Social Security (ESS) Working Papers*, No. 11.
- SWANEPOEL, F., A. STROEBEL, AND S. MOYO (2010): “The Role of Livestock in Developing Communities: Enhancing Multifunctionality,” The Technical Centre for Agricultural and Rural Cooperation (CTA).
- TABELLINI, G. (2000): “A positive theory of social security,” *The Scandinavian Journal of Economics*, 102, 523–545.
- UNITED NATIONS (2014): *World Urbanization Prospects*, New York: United Nations.
- VERBON, H. A. AND M. J. VERHOEVEN (1992): “Decision making on pension schemes under rational expectations,” *Journal of Economics*, 56, 71–97.

- VERHOEVEN, M. J. AND H. A. VERBON (1991): "Expectations on pension schemes under non-stationary conditions," *Economics Letters*, 36, 99–103.
- VILELA, A. (2013): "Pension coverage in China and the expansion of the New Rural Social Pension," *Pension watch briefing*, *HelpAge International*, May.
- WILLMORE, L. (2014): "Towards universal pension coverage in Mexico," *Pension watch briefing*, *HelpAge International*, May.
- WORLD BANK (1994): *Averting the Old Age Crisis*, A World Bank Policy Research Report, Washington: World Bank Group.
- (2011): *Aspire: The Atlas of Social Protection Database*, Washington: World Bank Group.
- WORLD VALUE SURVEYS (2009): "World Value Survey 1981-2008 Official Aggregate v.20090901," <http://www.worldvaluessurvey.org/wvs.jsp>.
- YANG, Y., J. B. WILLIAMSON, AND C. SHEN (2010): "Social security for China's rural aged: a proposal based on a universal non-contributory pension," *International Journal of Social Welfare*, 19, 236–245.
- ZELLER, M. AND M. SHARMA (2000): "Many borrow, more save, and all insure: implications for food and micro-finance policy," *Food policy*, 25, 143–167.