DISCUSSION PAPER SERIES

The Political (In)Stability of Funded Social Security

Roel Beetsma
Oliwia Komada
Krzysztof Makarski
Joanna Tyrowicz

I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

The Political (In)Stability of Funded Social Security

Roel Beetsma
University of Amsterdam, Copenhagen Business School, European Fiscal Board, CEPR, CESifo and Tinbergen Institute

Oliwia Komada
FAME|GRAPE and Warsaw School of Economics

Krzysztof Makarski
FAME|GRAPE and Warsaw School of Economics

Joanna Tyrowicz
FAME|GRAPE, University of Warsaw and IZA

ABSTRACT

The Political (In)Stability of Funded Social Security*

We analyze the political stability of funded social security. Using a stylized theoretical framework we study the mechanisms behind governments capturing social security assets in order to lower current taxes. The results and the driving mechanisms carry over to a fully-fledged and carefully calibrated overlapping generations model with an aging population. Funding is efficient in a Kaldor-Hicks sense. We demonstrate that, even though we can rationalize the actual introduction of a two-pillar defined-contribution scheme with funding through a majority vote, a new vote to curtail the funded pillar through asset capture or permanent diversion of contributions to the pay-as-you-go pillar always receives majority support. For those alive and thus allowed to vote, the temporary reduction in taxes outweighs the reduction in retirement benefits. This result is robust to substantial intra-cohort heterogeneity and other extensions, and only overturned with a sufficient degree of altruism. Our analysis rationalizes the experience of Central and Eastern European countries, who rolled back their funded pension pillars soon after setting them up.

JEL Classification:	H55, D72, E17, E27
Keywords:	social security, funding, pay-as-you-go, asset capture, majority vote, welfare

Corresponding author:
Joanna Tyrowicz
Group for Research in APplied Economics (GRAPE)
Foundation of Admirers and Mavens of Economics
ul. Mazowiecka 11/14 00-052
Warszawa
Poland

E-mail: j.tyrowicz@grape.org.pl

1 Introduction

Political support for social security has been extensively studied in the literature with key questions encompassing the very existence of inter-generational transfers (e.g. Samuelson 1958, Aaron 1966, Breyer 1989, Boll et al. 1994, Krieger and Ruhose 2013), the size of these transfers (e.g. Browning 1975, Boldrin and Rustichini 2000, Casamatta et al. 2001), and the political economy of arriving at a social contract (e.g. Sjoblom 1985, Boadway and Wildasin 1989a,b, Cooley and Soares 1996, 1999a,b, Tabellini 2000, Conde-Ruiz and Galasso 2005, Kelley 2014, Parlevliet 2017). This literature typically refers to social security as a pay-as-you-go (PAYG) transfer from current workers to the retired. The findings in this literature have been synthesized in different review studies (Galasso and Profeta 2002, de Walque 2005, Mulligan and Sala-i Martin 2015) that conclude that PAYG social security can be a politically stable arrangement. The mechanism typically relies on the assumption that if a working generation refuses to pay its transfer to the currently retired, it will not receive a transfer from the next working generation either, once it is old itself, causing the arrangement to be permanently abolished.

However, social security does not need to be financed on a PAYG basis. The alternative is a funded system in which the participants save for their own retirement. The political stability of such a system cannot rely on the retaliation mechanism described above. The contribution of this paper is the analysis of the political stability of funded social security in a stylized theoretical overlapping-generations model and subsequently study the same problem in a computational model, which we calibrate to the Polish economy. Such an analysis is of policy relevance because a number of Central and Eastern European (CEE) countries, among them Poland, introduced (partial) funding between the mid-1990s and early 2000s (see Holzmann and Stiglitz 2001, Bonoli and Shinkawa 2006, Gruber and Wise 2009), but abolished it again a number of years later. The intuition is that the change in the state of the economy, in particular accumulated pension savings, changes the trade-offs for the voters.

Most of the CEE countries countries originally featured a defined-benefit (DB) Beveridge type of system financed on a PAYG basis. Reform implied the replacement of these systems with an individual defined-contribution (DC) Bismarckian system combining mandatory publicly-managed PAYG and funded pillars with contributions set by the law and collected by the state social security institution.[1] These reforms were supposed to deliver long-run welfare gains, even when accounting for the transition costs, i.e. they would improve welfare in the Kaldor-Hicks sense.[2]

However, a decade or so after funding was introduced, it was partially or completely rolled back with governments appropriating the assets accumulated in the funded pillar, while raising the share of contributions going to the PAYG pillar (Schwarz 2011, 2014).[3] Although these policy reversals are projected to reduce future retirement benefits (by roughly 10-20%, see Jarrett 2011, Hagemejer et al. 2015, for Hungary and Slovakia, respectively Poland), they turned out to be politically acceptable.

---

[1] In addition, voluntary funded pillars were established with tax incentives to encourage private savings for retirement.

[2] A study for Poland shows that these welfare gains can be sizable, on the order of 2-3% of lifetime consumption (Makarski et al. 2017). Without exception these reforms honored the existing social security obligations to the transition cohorts, i.e. those already retired or too close to retirement to meaningfully adjust to the new situation. They also kept overall contribution rates unchanged. The result was a fiscal gap, as current social security benefits were paid in full, while at the same time part of the collected contributions were channeled to the funded pillar. The benefit of the reforms was enhanced efficiency by establishing a link between labor supply and future retirement benefits and by promoting capital accumulation.

[3] None of these countries abandoned the DC feature of their social security.

Our analysis can explain these reversals, because a majority of those alive on net benefit from the fact that the tax reductions financed by these reversals dominate the implied reduction in future social security benefits. This result is robust to substantial intra-cohort heterogeneity and other extensions, and only overturned with a sufficiently high degree of altruism.

We construct an overlapping-generations general-equilibrium framework able to rationalize the rolling back of a funded social-security pillar, despite overall welfare gains from a two-pillar system. The rolling back can take place through the diversion of contributions to the PAYG pillar or through the capture of accumulated assets for retirement.[4] We refer to any such roll-back measures as "unfunding". As the return on the funded pillar exceeds that of the PAYG pillar, individuals in effect face a trade off between lower current taxes and lower future retirement benefits. The overall benefit is in addition affected by the general equilibrium effects associated with adjustments in the labor supply and private voluntary savings.[5] The distribution of the benefits and costs of unfunding differs across the cohorts alive at the time of the vote. In particular, younger cohorts forego more of the higher return on funded retirement contributions. However, for both a permanent reduction of the funded pillar and a temporary asset capture, we demonstrate that the tax reduction channel dominates the reduction in retirement benefits for a majority of the voters at any point in time. Hence, in line with the experience of a number of CEE countries, our overall conclusion is that a funded social security pillar is politically unsustainable if the alternative of an enlarged PAYG pillar is available. We find that this result is robust to the presence of substantial intra-cohort heterogeneity, demonstrating that potential coalition building among subcohorts does not affect the results, and voting at alternative moments during the demographic transition.[6]

Our paper relates to the literature on "privatizing social security", which tends to emphasize (Hicksian) welfare improvement.[7] There, the key policy challenge is the appropriate timing of the introduction of funding, because the working cohorts have to finance the retirement benefits of both their predecessors and themselves (Huang et al. 1997). This double burden may be ameliorated by deploying public debt to smooth the costs of the reform across the cohorts (Belan and Pestieau 1999, Makarski et al. 2017).[8] This also matters from a political perspective, because a transfer of future welfare gains to the present helps to compensate the current losses of the transition cohorts that need to support the reform (Conesa and Garriga 2008, compare various ways to introduce a funded pillar with sufficient political support). While this literature analyzes the political economy of introducing funded social security, to the best of our knowledge there exists little or no work on whether a funded pillar, once introduced, is politically stable.

---

[4]We exclude the possibility of capture of assets from voluntary private savings.

[5]Note that this problem is analogous, *au rebours*, to the introduction of a funded pillar in an economy with PAYG social security. Introducing a funded pillar yields delayed gains in the form of higher retirement benefits and lower future taxes, but at the expense of a contemporaneous increase in taxation. Here, the gains from unfunding are immediate, whereas the costs in terms of lower benefits are born later.

[6]We do not consider repeated voting as in, for example, Cooley and Soares (1999a) and Krusell and Rios-Rull (1999). When given the chance to unfund, the currently-alive prefer to do it now, rather than to wait for some future moment. Moreover, the resulting shift away from funding is assumed to be permanent: with an experience of unfunding in the memory of voters, it will be politically difficult to reintroduce or expand funding.

[7]The general conclusion of welfare gains is a robust finding, see a discussion by Conesa and Krueger (1999) and Fehr (2009).

[8]Song et al. (2015) consider the case in which wage growth exceeds the financial market return, a situation not uncommon in fast-growing economies with underdeveloped financial markets. In this case PAYG-DB may be more efficient than funding.

The remainder of our paper is structured as follows. In Section 2 we demonstrate our main results and their driving mechanisms using a stylized three overlapping generations model. Section 3 presents our fully-fledged quantitative overlapping generations model, while Section 4 describes the calibration of the economy and social security for the case of Poland, which is to a large extent exemplary for the CEE countries.[9] Section 5 discusses the results generated with our quantitative overlapping generations model. Finally, Section 6 concludes the paper. The appendices are not for publication, but are available online.

2 A stylized illustrative model

This section presents a stylized theoretical model to demonstrate that social security with a funded pillar is not politically stable. We assume an open economy with endogenous savings and a single asset earning an exogenous rate of return ("interest rate") r. Initially, there is a social security system in place with a PAYG and a funded pillar. Voters decide whether to retain the social security system in its entirety, whether to abolish only the funded part or whether to abolish the system in its entirety. The advantage of our simple framework is that it allows us to clearly lay out the mechanisms behind the lack of political stability of funded social security. This model provides the groundwork for the later analysis based on a fully-fledged quantitative model, which allows for a realistic demographic structure, general equilibrium effects and intra-cohort heterogeneity.

2.1 Environment

Our stylized economy is populated by overlapping generations of individuals who live for $J = 3$ periods. Cohorts of newborns are of equal size, normalized to 1, in each period and, hence, the population and its demographic structure are constant over time. We allow for an individual life span of three periods in order to have meaningful majority votes over alternative policies. Within each cohort individuals are identical.

The preferences of an individual born in period t are

$$\mathcal{U}_t = \sum_{j=1}^{J} \delta^{j-1} u(c_{j,t+j-1}) \tag{1}$$

where $c_{j,t+j-1}$ denotes the period $t+j-1$ consumption of an individual in period j of her life. Function $u(\cdot)$ is strictly increasing and strictly concave. The real wage rate grows at a constant rate g from period to period. Individuals inelastically supply one unit of labor in each of the first two periods of their life, and are retired in the third period of their life. Their net income $\Psi_{j,t}$ at age j originates either from labor earnings ($j < J$) or a (potential) social security benefit $b_{j,t}$ ($j = J$):

$$\Psi_{j,t} = \begin{cases} (1-\tau_t)w_t, & \text{for } j < J, \\ b_{j,t}, & \text{for } j = J, \end{cases}$$

---

[9]In 1999 Poland reformed its PAYG-DB system into a DC scheme, with a PAYG and a funded pillar. In 2011, it started shifting contributions from the funded to the PAYG pillar, while in 2013 part of the assets in the funded pillar were captured by the government.

where w_t is the real wage rate in period t and τ_t is the sum of the contribution rates to the different social security pillars. The social security benefit is

$$b_{J,t} = f_{J,t}^P + f_{J,t}^F, \tag{2}$$

where $f_{J,t}^P$ are the entitlements in the PAYG pillar and $f_{J,t}^F$ are the assets in the funded pillar.

The total social security contribution rate is $\tau_t = \tau_t^P + \tau_t^F$, where τ_t^P is the contribution rate to the PAYG pillar and τ_t^F the contribution rate to the funded pillar. PAYG entitlements are indexed at the growth rate g of the aggregate wage bill, which is the same as the growth rate of the individual wage rate, because the working-age population is constant. Contributions to the funded pillar are invested against the exogenous and constant interest rate r. We assume that the return on the funded social security pillar exceeds that on the PAYG pillar, i.e. $r > g$. With this assumption we stack the odds against PAYG and thus raise the bar for terminating the funded pillar. We will nevertheless show that the funded pillar is politically not stable.

The social security budget constraint The government collects all the contributions and pays out the benefits related to the PAYG social security. The period-t budget constraint of the PAYG pillar is:

$$b_t^P = \sum_{j=1}^{J-1} \tau_t^P w_t + Subs_t. \tag{3}$$

The government budget constraint The government uses lump-sum taxes Υ_t to finance exogenous government purchases $G_t = (1+g)G_{t-1}$, which grow at the same rate as the economy's wage bill, and a potential subsidy $Subs_t$ to the social security system. The period-t government budget constraint reads:

$$G_t + Subs_t + (1+r)D_t = \sum_{j=1}^{J} \Upsilon_t + \tilde{F}_t^F + D_{t+1}, \tag{4}$$

where D_t denotes the outstanding public debt at the end of period $t - 1$. The first term on the right-hand side is the sum of the lump-sum taxes paid by all cohorts currently alive. The term $\tilde{F}^F$ denotes social security assets from the funded pillar captured by the government. Details are presented later.

If retirement assets are captured in period t, a government surplus emerges in this period. This is followed by a government deficit in period $t + 1$, because $r > g$ and the original promise on the return on the funded pillar contributions made in period $t - 1$ is maintained, but the entitlement associated with this contribution is now recorded in the PAYG pillar. In contrast to the funded pillar contribution in period $t - 1$, all contributions in period t get uprated at the growth rate of the wage bill g.

The individual budget constraint Individuals can accumulate assets $a_{j,t}$ in order to smooth consumption over their lifetime. The budget constraint of an individual of age j in period t is given by
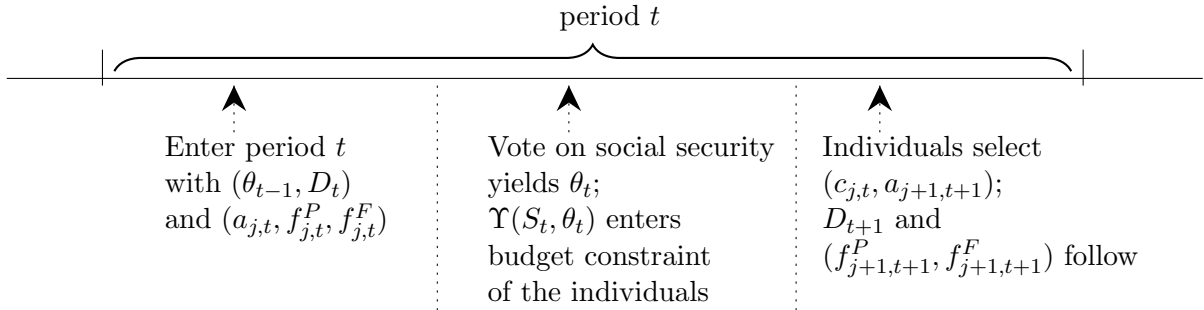
$$a_{j+1,t+1} + c_{j,t} + \Upsilon_t = (1+r)a_{j,t} + \Psi_{j,t}, \tag{5}$$

where Υ_t is a lump-sum tax.

Notation For convenience, we introduce following notation. We denote by $\theta_t \in \mathcal{P}(\theta_{t-1})$ the set of possible social security arrangements in period t, conditional on arrangement θ_{t-1} in period $t-1$. Here, we define $\theta_t \in \{0, 1, 2\}$, where 0 indicates no social security system at all, 1 a system with only a PAYG pillar, and 2 a system with both a PAYG and a funded pillar.

Timing The timing of events in period t is summarized in Figure 1. First, at the start of the period, the outcome of the previous period vote and public debt D_t are given, i.e. the aggregate state is $S_t = (\theta_{t-1}, D_t)$. Also, the private state $s_{j,t} = (a_{j,t}, f_{j,t}^F, f_{j,t}^P)$ is given. Next, individuals vote on the social security outcome, which yields θ_t. This outcome is determined by a majority vote. Then, the lump-sum tax $\Upsilon_t = \Upsilon(S_t, \theta_t)$ is determined via the government budget constraint, described below. Finally, given their private state $s_{j,t}$, the voting outcome θ_t and resulting lump-sum tax, individuals choose period-t consumption and the amount of assets with which to enter the next period.

Figure 1: Timing

period t

| Enter period t with (θ_{t-1}, D_t) and $(a_{j,t}, f_{j,t}^P, f_{j,t}^F)$ | Vote on social security yields θ_t; $\Upsilon(S_t, \theta_t)$ enters budget constraint of the individuals | Individuals select $(c_{j,t}, a_{j+1,t+1})$; D_{t+1} and $(f_{j+1,t+1}^P, f_{j+1,t+1}^F)$ follow |

Voting on the evolution of the social security system We study Markov perfect equilibria. If a social security system is in place, it consists of a PAYG pillar only or the combination of a PAYG and a funded pillar. It never consists of a funded pillar alone. This assumption is based on the observation that in practice we do not observe funded pillars without also some basic PAYG scheme in place. Moreover, we will show that, starting from the presence of a social security system, the voting outcome will always be that the PAYG pillar is maintained. In our analysis of the voting, and in line with the related literature, we assume that once a social security pillar is abolished, it will not be resurrected. Formally, we denote $\mathcal{P}(\theta_{t-1} = 2) = \{0, 1, 2\}$, $\mathcal{P}(\theta_{t-1} = 1) = \{0, 1\}$, and $\mathcal{P}(\theta_{t-1} = 0) = \{0\}$. This assumption seems reasonable in view of the fact that voters cannot expect a newly erected pillar to be more sustainable than one that has been abolished. Finally, as already mentioned and also in line with reality, we assume that the benefit payments to the retirees in the period in which the abolishment of funding takes place are respected.

We assume that as long as a social security system is in place, the total contribution rate is constant, i.e. $\tau_t = \tau$, and that as long as both a PAYG and a funded pillar are in place, the contribution rates to these pillars are also constant, i.e. $\tau_t^P = \tau^P$ and $\tau_t^F = \tau^F$. Hence, rolling back the funded pillar implies that the contribution to this pillar will be diverted to the PAYG pillar, as long as the latter remains in place.

6

The evolution of pensions and contributions   With the above assumptions, we can now describe the possible ways in which the social security arrangements, their contributions and their entitlements and assets evolve. Trivially, if there is no social security left in place after the vote in period $t-1$ (hence, $\theta_{t-1} = 0$), then there will be no ensuing vote in period t and onward.

With only PAYG social security in period $t-1$ (hence, $\theta_{t-1} = 1$), the possible voting choices in period t are to maintain PAYG social security ($\theta_t = 1$) or abandon it ($\theta_t = 0$).[10]   Hence, the contributions in period t and the implied PAYG entitlements in period $t+1$ of someone of age $j+1$ are:

$$\tau_t^P(\theta_t) = \begin{cases} \tau, & \text{if } \theta_t = 1 \\ 0, & \text{if } \theta_t = 0 \end{cases} \qquad \text{and} \qquad f_{j+1,t+1}^P(\theta_t) = \begin{cases} (1+g)(f_{j,t}^P + \tau_t^P w_t), & \text{if } \theta_t = 1 \\ 0, & \text{if } \theta_t = 0 \end{cases} \tag{6}$$

Hence, from period to period entitlements grow through contributions and indexation of existing entitlements to wage growth, as is the case in a textbook model of a notional DC system.

Finally, with both a PAYG and a funded pillar in place in period $t-1$ (hence, $\theta_{t-1} = 2$), all voting options are available: abandon social security completely ($\theta_t = 0$), continue with the PAYG pillar only ($\theta_t = 1$), or continue with both pillars ($\theta_t = 2$). Elimination of the funded pillar implies that the social security assets of the working individuals are transferred to the government's budget (see below). These individuals will receive a benefit according to the rules of the PAYG pillar. For the different possible cases, individual contributions in period t and the implied PAYG entitlements and funded pillar assets in period $t+1$ of someone of age $j+1$ are:

$$\tau_t^P(\theta) = \begin{cases} \tau^P, & \text{if } \theta_t = 2 \\ \tau, & \text{if } \theta_t = 1 \\ 0, & \text{if } \theta_t = 0 \end{cases} \qquad \text{and} \qquad f_{j+1,t+1}^P(\theta_t) = \begin{cases} (1+g)(f_{j,t}^P + \tau_t^P w_t) & \text{if } \theta_t = 2 \\ (1+g)(f_{j,t}^P + f_{j,t}^F + \tau_t^P w_t) & \text{if } \theta_t = 1 \\ 0, & \text{if } \theta_t = 0 \end{cases} \tag{7}$$

$$\tau_t^F(\theta) = \begin{cases} \tau^F, & \text{if } \theta_t = 2 \\ 0, & \text{otherwise} \end{cases} \qquad \text{and} \qquad f_{j+1,t+1}^F(\theta_t) = \begin{cases} (1+r)(f_{j,t}^F + \tau_t^F w_t) & \text{if } \theta_t = 2 \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

The second line in (7) is the result of the fact that accumulated pension assets in period t are captured by the government and the individual who is expropriated is compensated with equally valued entitlements in the PAYG pillar, which are then indexed against the wage growth rate g. Rolling back the funded pillar in period t does not affect social security benefit payments in that period, because period-t PAYG benefits are paid out of the current contributions by the working-age population, while the funded part of the benefits is paid out of the retirees' accumulated assets. Hence, as in the case in which the funded pillar is maintained in period t, the funded part of the benefits paid out in period t is still given by $b_{J,t}^F = (1+r)(f_{J-1,t-1}^F + \tau^F w_{t-1})$. The total asset capture when the funded pillar is abandoned is given by the accumulated funded retirement assets (including current contributions to this pillar) of those not yet retired in period t, i.e. by $\sum_{j=1}^{J-1}(f_{j,t}^F + \tau^F w_t)$, using that the size of each cohort is normalized to 1. Because the indexation rate in the PAYG pillar differs from

---

[10]The possibility to abolish the PAYG pillar in the absence of a funded pillar allows us to compare our results to those of Cooley and Soares (1999a), by exploring whether a PAYG pillar is a politically stable outcome in this situation.

the rate of return in the funded pillar, there are differences in the retirement benefits in period $t+1$ depending on the arrangement with which period t is entered:

$$\theta_{t-1} = 1 \quad \text{and} \quad \theta_t = 1 \quad : b_{J,t+1} = (1+g)((1+g)\tau w_{t-1} + \tau w_t)),$$

$$\theta_{t-1} = 2 \quad \text{and} \quad \theta_t = 1 \quad : b_{J,t+1} = (1+g)((1+r)\tau^F w_{t-1} + (1+g)\tau^P w_{t-1} + \tau w_t)). \qquad (9)$$

The first expression is the total benefit in period $t+1$ of someone who starts her working career when only PAYG social security is in place, while the second expression is the benefit in period $t+1$ of someone who starts her working career when the funded pillar is still in place, but this pillar is abolished in the second period of her working life.

Evolution of taxes, $\Upsilon(S_t, \theta_t)$ We assume that in case of no asset capture from the funded pillar $\theta_t \in \{0, 1\}$ in period t the government keeps the debt level at the constant level, i.e. $D_{t+1} = \bar{D}$. Thus, the tax rate given the initial value of debt D_t follows from the government budget (4), in which $\tilde{F}_t^F = 0$. In the case of asset capture in period t, the evolution of government debt and the taxes is as follows. Because the funded pillar in period $t-1$ accumulates with the interest rate rather than the payroll growth rate, the need to honour the funded entitlements of the retired generates a deficit in the PAYG pillar in period $t+1$, which is covered by saving some of the captured assets. The remainder of the captured assets is used to lower lump-sum taxes. Concretely, in case of $\theta_{t-1} = 2$ and $\theta_t = 1$, we have $\tilde{F}_t^F = \sum_{j=1}^{J-1}(f_{j,t} + \tau^F w_t)$, while taxes follow from (4) and :

$$\theta_{t-1} = 2 \quad \wedge \quad \theta_t = 1 \quad : D_{t+1} = \bar{D} - \frac{(1+g)(1+r)\tau^F w_{t-1} - (1+g)(1+g)\tau^F w_{t-1}}{(1+r)},$$

$$\text{otherwise} \quad : D_{t+1} = \bar{D}. \qquad (10)$$

Individuals vote on the social security arrangement and not on the government debt. Hence, we assume that the initial government debt is $\bar{D}$ and that it returns to this level after a change in social security.

Individual decisions For a given state $(s_{j,t}, S_t)$, prices, social security arrangement $\theta_t = \theta(\theta_{t-1}) \in \mathcal{P}(\theta_{t-1})$ and government policies G_t and $\Upsilon(S_t, \theta_t)$, an individual solves:

$$V_{j,t}(s_{j,t}, S_t) = \max_{(c_{j,t}, a_{j+1,t+1})} \{u(c_{j,t}) + \delta V_{j+1,t+1}(s_{j+1,t+1}, S_{t+1})\}, \qquad (11)$$

subject to the budget constraint (5). Social security benefits in the budget constraint are determined by the specific arrangement in place.

Definition of equilibrium We employ the concept of Markov perfect equilibrium to analyze our economy. First, strategies are based on the current state of the economy and not on the whole histories. Second, agents while casting their vote about maintaining the existing social security take the results of subsequent votes as given, because no commitment to future votes is feasible. Voters in each period vote as long as in the past they did not decide to (partially or fully) abandon the arrangement. Following the literature on the stability of PAYG systems, we assume that what has been dismantled cannot be reinstated. This notion of equilibrium is illustrated by Figure 2.

Figure 2: Repeated voting on (components of) the social security arrangement
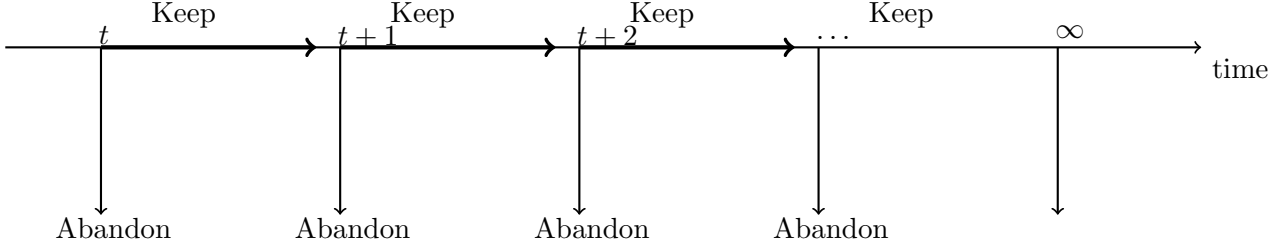
Definition 1. *A Markov-Perfect Equilibrium is a policy function $\theta(\theta_{t-1})$, lump-sum taxes $\Upsilon(S_t, \theta_t)$, contribution rates $(\tau_t^P(S_t, \theta_t), \tau_t^F(S_t, \theta_t))$, pensions $(b_{J,t}^P(S_t, \theta_t), b_{J,t}^F(S_t, \theta_t))$, evolution of debt $D_{t+1}(S_t, \theta_t)$, an individual's policy function $(c_{j,t}(s_{j,t}, S_t), a_{j+1,t+1}(s_{j,t}, S_t))$ and a value function $V_{j,t}(s_{j,t}, S_t)$, assuming that $D_{t-1} = \bar{D}$, such that:*

- *$\theta_t = \theta(\theta_{t-1}) \in \mathcal{P}(\theta_{t-1})$ is a policy selected in pure majority voting when pitched against each of the available alternatives, given government policy $(\tau_t^P(S_t, \theta_t), \tau_t^F(S_t, \theta_t))$, $(b_{J,t}^P(S_t, \theta_t), b_{J,t}^F(S_t, \theta_t))$ and $\Upsilon(S_t, \theta_t)$.*

- *the policy and value functions, $(c_{j,t}(s_{j,t}, S_t), a_{j+1,t+1}(s_{j,t}, S_t))$ and $V_{j,t}(s_{j,t}, S_t)$ respectively, solve the individual's problem (11);*

- *social security contributions and pensions are given by (7) - (9).*

- *lump-sum taxes satisfy the government budget (4) given the evolution of debt (10).*

2.2 Results

We present three main theoretical results. First, we show that (under weak conditions) the political stability of the PAYG pillar found by Cooley and Soares (1999a) carries over to the current setting. Second, we show that a funded pillar is not politically stable and will be abolished in a majority vote. Finally, we demonstrate the existence of an equilibrium with a single PAYG social security pillar, while an equilibrium two social security pillars does not exist. The alternatives available in the votes are determined by the assumptions above. In particular, once a social security pillar has been abolished, it cannot be re-installed.

The first result is analogous to that in Cooley and Soares (1999a):

Proposition 1. *Assume that the initial arrangement is one with a PAYG social security pillar only ($\theta_{t-1} = 1$). If $r < 1 + 2g$, the arrangement is politically stable, i.e. voters will always vote to keep it.*

The proof is straightforward and contained in Appendix A. It is based on the congruence of the interests between the retired and the middle-aged. To demonstrate the congruence it suffices to show how different arrangements affect the budget constraint of the different cohorts alive. Retirees would lose if the PAYG system is abolished, because they have already paid their contributions during their working life and cannot recover them, while they would lose their PAYG retirement benefit. Middle-aged individuals face a trade-off between dismantling the PAYG pillar and losing the contributions made so far, or benefiting from a higher return $r > g$ on putting new contributions into a savings account instead. Unless the interest rate is extremely high, the former option dominates. Young

individuals prefer to abandon social security, because the return on their PAYG contributions is dominated by the return on their private savings. However, they are outnumbered by the retired and the middle-aged.

Second, we show that for a funded pillar the analogue of Cooley and Soares (1999a) does not hold. In fact, there is now a congruence of interests of the young and the old to capture the funded pillar assets and, hence, the funded pillar is politically not stable:

Proposition 2. *Assume that the initial arrangement features both a PAYG and a funded pillar ($\theta_{t-1} = 2$). Then, in a vote whether or not to maintain the funded pillar, the latter is abolished.*

Again, the proof is straightforward and contained in Appendix A. The intuition for abolishing the funded pillar, while maintaining the PAYG pillar is as follows. Following equation (9), all previously assumed retirement benefits are honored, i.e. the assigned pensions are protected in terms of value in the period in which the funded pillar is abolished. The retired consume their savings in the funded pillar, while they benefit from the reduced taxation afforded by the asset capture. Because $r > g$, young individuals expect to a lower total retirement benefit. However, the tax reduction more than compensates the reduction of their retirement benefit. The asset capture shifts the burden of financing the retirement benefits of the current young to future cohorts, who lose, but are not represented in the vote. The middle-aged lose the assets they have accumulated in the first period of their life, but they benefit from the tax reduction. Their net benefit depends on the return differential $r - g$ between the funded and the PAYG pillar. However, their vote would always be dominated by the coalition between the young and the old.

The above propositions do not inform us about the outcome when all possible arrangements are pitched against each other. The next proposition states what *is* an equilibrium:

Proposition 3. *Let $r < 1+2g$. Starting from single-pillar PAYG social security or social security with both a PAYG and a funded pillar, the unique voting outcome will be an arrangement with single-pillar PAYG social security. Hence, there exists a Markov perfect equilibrium with a single-pillar PAYG social security arrangement, but no Markov perfect equilibrium with both a PAYG and a funded social security pillar.*

Appendix A contains the complete proof, of which we provide a sketch here. We demonstrate that the proposed equilibrium is an equilibrium indeed by resorting to the "one-shot deviation principle", i.e. for $\theta_{t-1} \in \{1, 2\}$ an agent finds a one-shot deviation from $\theta_t = \theta(\theta_{t-1}) = 1$ not optimal.

We start with the case of $\theta_{t-1} = 2$. For this case we show that regardless of the sequence in which pairs of arrangements are put to majority vote, the eventual outcome is always that voters choose to abolish the funded pillar, i.e. capture assets (and shift contributions), but not to abolish the PAYG pillar.

First, compare no asset capture ($\theta_t = 2$) with asset capture ($\theta_t = 1$). By the one-shot deviation principle, not capturing assets this period means that they are captured next period. For the young, asset capture implies a reduction in the retirement benefit (discounted to the present). However, this reduction is dominated by the present value of the loss resulting from the postponement of the tax reduction. The old individuals prefer asset capture in the current period, so as to benefit from the tax reduction, instead of asset capture after they have died. Hence, the voting outcome is asset capture.

The second comparison is asset capture ($\theta_t = 1$) with completely abandoning social security ($\theta_t = 0$). For the middle-aged, both policies are associated with a cut in current taxes, though the cut differs between $\theta_t = 1$ and $\theta_t = 0$ due to the fact that in case of abolishing social security there is no need to keep some of the funds to finance future deficits in PAYG pillar. In addition, when $\theta_t = 0$, they gain the abolishment of the current contribution to PAYG social security, but they lose their discounted PAYG benefit. On net, they lose, implying that they prefer to capture assets rather than abolish social security completely. In comparison with the middle-aged, the old do not benefit from abolishing the contribution payment in return for giving up their PAYG retirement benefit and, hence, they also prefer to capture assets and shift contributions rather than to abandon social security completely.

The third comparison is keeping both pillars intact ($\theta_t = 2$) versus abandoning social security completely ($\theta_t = 0$). In the latter case, the middle-aged lose the discounted value of their social security benefit, but gain from not having to pay the contribution in this period. They also gain a current reduction in taxation, but they forego the discounted reduction in taxation associated with (only) the elimination of the funded pillar in the next period as prescribed by the equilibrium path. On net they lose. The old clearly prefer to keep both pillars intact to abandoning social security completely.

The final step in the proof is to start from $\theta_{t-1} = 1$. From Proposition 1 it follows that the vote results into $\theta_t = \theta(\theta_{t-1}) = 1$.

In effect, starting from a two-pillar scheme, assets are captured by a coalition of the young and the old with funded-pillar contributions being redirected to the PAYG pillar, while the PAYG pillar is sustained by a coalition of the middle-aged and the old. A number of further remarks are warranted. First, we can show that the sequence in which the alternatives are voted upon does not affect the final outcome, which is a Condorcet winner of pure majority voting. Second, we can show that the above propositions hold in the presence of elastic labor supply and labor taxation. Third, Cooley and Soares (1999a) argue that the political stability of a PAYG pillar crucially depends on the general equilibrium effects implied by the presence of social security, in particular their influence on the capital stock and the return to private capital. However, our analysis shows that the political viability of the PAYG pillar goes through also in a stylized small open economy context with an exogenous asset return r and wage growth g. The remainder of this paper explores the political stability of a funded social security pillar in a fully-fledged quantitative small-open economy framework. Finally, other stylized setups in the literature emphasize the role of potential coalitions between heterogeneous agents (including Cooley and Soares 1999b). Specifically, political outcomes may be affected by a preference for redistribution from high to low productivity individuals. Therefore, in our fully-fledged setup below we will also consider ex-ante intra-cohort heterogeneity.

3 A quantitative model

This section constructs a fully-fledged small-open economy model that allows for realistic calibration and evaluation of the political stability of social security arrangements. The model features a large number of overlapping cohorts that are shrinking as they become older due to mortality which rises with age. With a fully-fledged general equilibrium setup we are able to realistically calibrate the main

elements of our stylized theoretical framework and show that the findings for that framework carry over to a more realistic setting. In particular, we study the political stability of funded social security when labor supply, fiscal policy and social security arrangements are realistically modeled.

In the earlier literature on the political stability of PAYG social security, intra-cohort heterogeneity and coalitions across generations were of crucial importance (in the spirit of Aaron 1966, Cooley and Soares 1999a, among others). In particular, in this literature low-productivity individuals would form a political coalition with elderly cohorts to support redistribution.[11] We allow for a realistic and substantial amount of dispersion in individual productivity and time preferences to explore the scope for similar coalition formation in policies to unfund social security. In addition, the presence of such intra-cohort heterogeneity allows us to study in a meaningful way how changes in social security arrangements affect old-age poverty, which is a major political concern.

3.1 International capital markets

Households have access to international capital markets in which they can borrow or lend at the interest rate r_t. As proposed by Schmitt-Grohe and Uribe (2003), the domestic interest rate r_t equals the world interest rate r_t^*, adjusted for a risk premium according to:

$$r_t = r_t^* + \xi \frac{D_t^*}{Y_t}$$

where D_t^* is the level of net foreign debt of the economy, Y_t is output (GDP) and ξ is a constant. Hence, an increase in the ratio of net foreign debt to GDP raises the interest rate at which domestic individuals can borrow in international markets.[12]

3.2 Households

Each individual lives up to a maximum of J periods, with age denoted by $j \in \{1, 2, ..., J\}$.[13] Individuals face a non-zero probability of dying in each period. The unconditional probability that an individual of age j is still alive at time t is denoted by $\pi_{j,t}$. This probability is homogeneous within the cohort. Rising longevity is modeled through rising survival probabilities over time. Individuals of the same age are heterogeneous with respect to their productivity and degree of time preference. Each combination of productivity and time preference corresponds to a different type $\kappa \in 1, ..., \mathcal{K}$. The set of different types is the same in each cohort and the share of each type group is the same across the cohorts. The size of subcohort κ of age j in period t is denoted by $N_{j,\kappa,t}$. The discount factor of an individual of type κ is δ_κ.

---

[11]The literature has also identified "unlikely" coalitions in other contexts. For example, Epple and Romano (1996) demonstrate that an "ends-against-the-middle" result may arise in the context of the simultaneous provision of private and public schooling. High-income individuals make no use of public schooling and, hence, prefer a combination low taxes and low spending on public schooling. Low-income individuals are less willing than the rich to see consumption replaced by public school spending and, hence, also prefer low taxes and public school spending. By contrast, middle-income individuals use public schools and prefer them to be of high quality and, hence, are willing to be taxed for that.

[12]Not allowing for the risk premium would result into implausibly large swings in the net foreign asset position. Since we calibrate the model to the case of Poland, an emerging economy, if we were to make the assumption that the economy is closed, the (endogenous) interest rate would over time fall due to the convergence of the productivity level to that of advanced economies. However, Poland, being part of the EU internal market, is better characterized as an open economy. Our main findings go through under the closed-economy assumption. The detailed results are available upon request.

[13]In our simulations $j = 1$ corresponds to 21 years and $j = J$ to 100 years of age in the "real world", as mortality tables usually go up to 100 years. Hence, we set $J = 80$.

Individuals choose consumption $c_{j,\kappa,t}$, labor $l_{j,\kappa,t}$, for which they receive a wage w_t per efficiency unit, and assets $a_{j,\kappa,t}$, which earn a rate of return r_t, to maximize the following Bellman equation:

$$V_{j,\kappa,t}(s_{j,\kappa,t}) = \max_{c_{j,\kappa,t},l_{j,\kappa,t}} \{u_\kappa(c_{j,\kappa,t},l_{j,\kappa,t}) + \delta_\kappa \frac{\pi_{j+1,t+1}}{\pi_{j,t}}V_{j+1,\kappa,t+1}(s_{j+1,\kappa,t+1})\}, \tag{12}$$

where $s_{j,\kappa,t}$ denotes the individual state, determined by assets $a_{j,\kappa,t}$ as well as accumulated social security contributions $f^P_{j,\kappa,t}$ and $f^F_{j,\kappa,t}$, and where we specify $u_\kappa(c_{j,\kappa,t},l_{j,\kappa,t}) = \ln c_{j,\kappa,t} + \phi \ln(1 - l_{j,\kappa,t})$ with $\phi \geq 0$.

Individuals face the following budget constraint:

$$(1 + \tau_c)c_{j,\kappa,t} + a_{j+1,\kappa,t+1} + \Upsilon_t = (1 - \tau_l)\Psi_{j,\kappa,t} + (1 + r_t(1 - \tau_k))a_{j,\kappa,t} + \mathcal{B}_{j,\kappa,t}, \tag{13}$$

where τ_c is the consumption tax rate, τ_l the labor income tax rate, τ_κ the capital income tax rate, Υ_t lump-sum taxes and $\mathcal{B}_{j,\kappa,t}$ unintentionally bequeathed assets. Bequests of voluntarily accumulated assets are evenly distributed among the remaining members of subcohort κ of age j. Further, $\Psi_{j,\kappa,t}$ is current-period gross income from labor or social security, which is given by:

$$\Psi_{j,\kappa,t} = \begin{cases} (1 - \tau)w_t\omega_\kappa l_{j,\kappa,t}, & \text{for } j < \bar{J} \\ b_{j,\kappa,t} & \text{for } j \geq \bar{J} \end{cases}, \tag{14}$$

where, as before, τ denotes the social security contribution rate, which will be constant irrespective of the specific social security arrangement in place, $\bar{J}$ the retirement age, ω_κ productivity of an individual in subcohort κ and $b_{j,\kappa,t}$ the social security benefit, which we will discuss below.

3.3 Firms

We assume a perfectly-competitive production sector that uses capital K_t and labor to produce output Y_t with the Cobb-Douglas technology: $Y_t = K_t^\alpha(z_t L_t)^{1-\alpha}$, where L_t is aggregate effective labor input and z_t captures exogenous labor-augmenting technological progress. Hence, profit is maximized when the net return on capital $r_t - d$, where d denotes the depreciation rate of capital, equals the marginal product of capital, i.e., $r_t = \alpha K_t^{\alpha-1}(z_t L_t)^{1-\alpha} - d$, and the wage rate equals the marginal product of labor, i.e. $w_t = (1 - \alpha)K_t^\alpha z_t^{1-\alpha} L_t^{-\alpha}$.

3.4 Social security arrangements

We model the evolution of social security arrangements after Poland and other CEE countries, which initially featured single-pillar PAYG-DB arrangements, but reformed these into two-pillar DC schemes with a PAYG and a funded pillar over the course of the 1990s and 2000s (see Holzmann and Stiglitz 2001, Orszag and Stiglitz 2001). In practice, after the reform, PAYG-DB retirement benefits were typically kept unchanged for cohorts close to retirement at the moment of the reform. For example, in Poland cohorts older than 40 years at the moment of the reform continued to receive benefits according to the pre-reform rules.[14] Hence, in our model we assume that cohorts older than 40 years

---

[14]For details on the social security reform in Poland, see Chlon et al. (1999).

at the moment of the reform vote continue to receive retirement benefits according to equation (15), while all younger cohorts receive benefits according to equations (19) below.

The initial social security arrangement is a single-pillar PAYG-DB arrangement with benefits given by:

$$b_{j,\kappa,t}^{PAYG-DB} = \begin{cases} 0, & \text{for } j < \bar{J}, \\ \rho\omega_\kappa w_{t-1}\bar{l}_{j-1,\kappa,t-1}, & \text{for } j = \bar{J}, \\ (1+g_t)b_{j-1,\kappa,t-1}^{PAYG-DB} & \text{for } \bar{J} < j \leq J, \end{cases} \tag{15}$$

where ω_κ is the productivity of subcohort κ. Hence, the benefits in the first year of retirement of subcohort κ are expressed in terms of the replacement rate ρ of the last pre-retirement earnings of the subcohort. Further, $\bar{l}_{j-1,\kappa,t-1}$ is the subcohort average labor supply in period $t-1$. Because in reality the link between an individual's labor supply and her future DB social security benefit was weak, we approximate this situation by assuming that such a link is entirely absent and that the benefit simply depends on the average labor supply of the whole subcohort. In this system, the contributions of the current workers pay for the retirement benefits of the current retirees, while a potential deficit needs to be plugged with a government subsidy $Subs_t$. Hence, the budget constraint of the PAYG-DB arrangement is:

$$\sum_{\kappa=1}^{\mathcal{K}}\sum_{j=\bar{J}}^{J} N_{j,\kappa,t}b_{j,\kappa,t}^{PAYG-DB} = \tau w_t L_t + Subs_t. \tag{16}$$

Shortly after the start of our economy, the social security system is unexpectedly subjected to a vote that turns it into a two-pillar DC system with a PAYG and a funded pillar. The entitlements to the new PAYG-DC pillar grow at the rate of growth of the total wage bill $g_t = w_t L_t/(w_{t-1}L_{t-1}) - 1$. Contributions to the funded pillar are invested against a return r_t. Hence, during their working life, the individuals' accounts in the PAYG-DC and funded pillars evolve as:[15]

$$f_{j+1,\kappa,t+1}^{P} = \frac{\pi_{j,t}}{\pi_{j+1,t+1}}(1+g_t)\big(f_{j,\kappa,t}^{P} + \tau_t^{P}\omega_\kappa w_t l_{j,\kappa,t}\big) \tag{17}$$

$$f_{j+1,\kappa,t+1}^{F} = \frac{\pi_{j,t}}{\pi_{j+1,t+1}}(1+r_t)\big(f_{j,\kappa,t}^{F} + \tau_t^{F}\omega_\kappa w_t l_{j,\kappa,t}\big). \tag{18}$$

where τ_t^{P} and τ_t^{F} are the contribution rates to the PAYG-DC and funded pillars, respectively. At retirement, each account is converted into an annuity. However, the difference in indexation rates between the two pillars remains. The social security benefits paid under the PAYG-DC and funded pillars are:

$$b_{j,\kappa,t}^{P} = \begin{cases} 0, & \text{for } j < \bar{J} \\ f_{\bar{J},\kappa,t}^{P}/LE_{\bar{J},t}, & \text{for } j = \bar{J} \\ (1+g)b_{j-1,\kappa,t-1}^{P} & \text{for } \bar{J} < j \leq J \end{cases} \quad \text{and} \quad b_{j,\kappa,t}^{F} = \begin{cases} 0, & \text{for } j < \bar{J} \\ f_{\bar{J},\kappa,t}^{F}/LE_{\bar{J},t}, & \text{for } j = \bar{J} \\ (1+r)b_{j-1,\kappa,t-1}^{F} & \text{for } \bar{J} < j \leq J \end{cases} \tag{19}$$

with $LE_{\bar{J},t} = \sum_{s=0}^{J-\bar{J}} \frac{\pi_{\bar{J}+s,t+s}}{\pi_{\bar{J},t}}$ denoting life expectancy at retirement and where the payouts are such

---

[15]The entitlements and assets of those agents who die prior to claiming benefits are added to the accounts of the survivors, i.e. they are distributed equally within each birth cohort and, hence, remain within the respective social security pillars. This is captured by the first term on the right-hand sides of the expressions below.

that their discounted sum equals the entitlements at the moment of retiring (as it is straightforward to verify).

In parallel with the stylized theoretical model, the total social security contribution rate is kept at the original contribution rate τ under the initial PAYG-DB scheme, i.e. $\tau_t^P + \tau_t^F = \tau$. Hence, the part τ_t^F of the contribution originally used to finance the PAYG-DB benefits now goes into the funded pillar. During the transition period, the PAYG pillar pays out DB retirement benefits (15) to those not affected by the reform and keeps accounts of the DC benefits of those affected by the reform (so-called "notionally defined contribution"). Hence, there is a transitory deterioration in the balance of the PAYG pillar. The deterioration is offset with a subsidy, as implied by equation (16). Once the transition is complete, $Subs_t$ becomes permanently zero.

3.5   The government

The government collects revenues to finance exogenous government purchases G_t, the deficit in the non-funded part of the social security and the servicing of the public debt D_t. Government revenues excluding the lump-sum transfers are:

$$T_t = \tau_{l,t}\Big[(1-\tau)w_t L_t + \sum_{\kappa=1}^{\mathcal{K}} \sum_{j=\bar{J}}^{J} b_{j,\kappa,t} N_{j,\kappa,t}\Big] + \sum_{\kappa=1}^{\mathcal{K}} \sum_{j=1}^{J} (\tau_{c,t} c_{j,\kappa,t} + \tau_{k,t} r_t a_{j,\kappa,t}) N_{j,\kappa,t}. \qquad (20)$$

The first term on the right-hand side is the sum of the taxation of labor income, after social security contributions have been paid, plus the taxation of retirement benefits, while the second term is the sum of the taxation of consumption and capital income. Hence, compared with voluntary savings, and in line with common practice in many countries including Poland (OECD 2018), contributions to the funded social security pillar receive favorable tax treatment, while the benefits are taxed when they are paid out. The government budget constraint is:

$$G_t + Subs_t + r_t D_t = T_t + (D_{t+1} - D_t) + \sum_{\kappa=1}^{\mathcal{K}} \sum_{j=1}^{J} N_{j,\kappa,t} \Upsilon_t, \qquad (21)$$

The final term is the lump-sum net transfer aggregated over all groups. Whenever a budget deficit emerges, it is absorbed by an adjustment in public debt. However, if debt exceeds its steady-state value, this also triggers an adjustment of the consumption tax rate. Define the ratio of public debt over output by $\mathcal{D} = D/Y$. The consumption tax rate is assumed to evolve as:

$$\tau_{c,t} = (1-\varrho)\tau_c^{final} + \varrho\tau_{c,t-1} + \varrho_D(\tilde{\mathcal{D}}_t - \mathcal{D}^{final}), \qquad (22)$$

where ϱ captures the smoothing of the consumption tax rate and ϱ_D the strength of the reaction to a deviation of the government debt ratio of GDP from its steady-state value. Here, the superscript $final$ is used to denote the steady-state value. After a social security reform the steady state changes and τ_c^{final} and $\mathcal{D}^{final}$ adjust accordingly. Our fiscal rule has $\tilde{\mathcal{D}}_t$ as a moving average of the debt ratio over the past two, the current and the next two years. This allows consumption taxes to react to debt in the near future. We adjust consumption tax rate, rather than lump-sum taxes, as in reality consumption taxes make up a larger part of government revenues and historically have proven responsive to the

budgetary situation. Via the parameters in (22) we can regulate this responsiveness.

3.6 Closing the model

The model is closed with market clearing conditions for:

$$\text{the labor market:} \quad L_t = \sum_{\kappa=1}^{\mathcal{K}} \sum_{j=1}^{\bar{J}-1} N_{j,\kappa,t} \omega_\kappa l_{j,\kappa,t}, \tag{23}$$

$$\text{the capital market:} \quad K_t + D_t = \sum_{\kappa=1}^{\mathcal{K}} \sum_{j=1}^{J} N_{j,\kappa,t} (a_{j,\kappa,t} + f_{j,\kappa,t}^{F}) + D_t^*, \tag{24}$$

$$\text{the goods market:} \quad \sum_{\kappa=1}^{\mathcal{K}} \sum_{j=1}^{J} N_{j,\kappa,t} c_{j,\kappa,t} + G_t + K_{t+1} + NX_t = Y_t + (1-d)K_t, \tag{25}$$

where NX_t denotes net exports and D_t^* denotes net foreign debt. Net foreign debt evolves over time according to:

$$D_{t+1}^* - D_t^* = r_t D_t^* - NX_t. \tag{26}$$

4 Calibration

We calibrate the model to match the features of the Polish economy as an example of a country that replaced its single-pillar PAYG-DB system with a two-pillar DC system, and maintained this with essentially unchanged features for over a decade, after which it effectively abolished the funded pillar. To avoid biases due to business cycle fluctuations, the calibration of the economic variables in the initial steady state is based on the averages for Poland over the period 1995 - 2010.[16]

For the initial steady state, we calibrate the demography to the year 1999. This is consistent with the economic data used for the calibration, while it is also the moment of the shift from the single-pillar PAYG-DB to the two-pillar DC arrangement. Specifically, we use the mortality rates for different ages of 1999 and birth-cohort sizes consistent with the demographic data for the computation of the initial steady state. This demographic structure is selected to adequately calibrate the parameters of the initial social security PAYG-DB system: the contribution rate, the replacement rate, and the deficit in this system. All three values are observable in the macroeconomic data.

We use the detailed demographic projection released by the Aging Working Group (AWG) of the European Commission to reproduce the arrival of new cohorts to the economy, as well as the annual survival probabilities of each subsequent birth cohort. The projection is available until 2080. We further assume that mortality and birth rates are constant after 2080. The specific values of the parameters imply a population decline at a rate of 1% per year in the very long run.

We also use the projection for the exogenous technological progress from the AWG as of 2010, whereas for the years between 1999 and 2010 we use the moving-average (to smooth out cyclical fluctuations) growth rate of total factor productivity (TFP) obtained from the data for this period. The AWG scenario for productivity assumes a gradual convergence to the annual average EU TFP growth rate of 1.54% between 2010 and 2040, after which it remains constant at this rate.

---

[16]Reliable data are not available before 1995, while after 2010 the capture of assets from the funded social security pillar and diversion of contributions from this pillar started.

The retirement age $\bar{J}$ is calibrated to the data on the effective retirement age collected by the OECD. It equals 61 in 1999 and remains constant throughout the projection (and across the scenarios we consider).[17] The replacement rate in the initial PAYG-DB system is set so as to replicate the share of retirement benefits in GDP prior to the reform of 1999. Knowing the replacement rates, we set the overall contribution rate τ to replicate the social security deficit of 0.8% of GDP observed in the years before the switch to two-pillar DC. The split between τ^{PAYG} and τ^F after the switch to the two-pillar DC scheme follows the proportions set by the Polish legislation.

Table 1: Calibrated parameters for the initial steady state

	Model parameter	Value	Target	Data (source)	Model outcome homogeneity	heterogeneity
α	capital share	0.300				
ϕ	preference for leisure	0.894	average hours	56% (LFS)	56%	56%
δ	discount rate	0.996	interest rate	5.7%	5.7%	5.7%
τ_l	labor tax rate	0.073	revenue/GDP	5.2% (OECD)	5.2%	5.2%
τ_c	consumption tax rate	0.200	revenue/GDP	11.3% (OECD)	11.3%	11.3%
τ_k	capital income tax rate	0.190	*de jure*			
ρ	replacement rate	0.216	benefits/GDP	5.0% (SIF)	5.0%	5.0%
τ	social security contr. rate	0.060	SIF deficit/GDP	0.8% (SIF)	0.8%	0.8%
d	depreciation rate	0.0325	investment/GDP	21% (NA)	21%	21%
ξ	sensitivity of r to B/Y	0.033	domestic inv./GDP	16% (NA)	16%	16%
$\mathcal{D}$	government debt/GDP	45%	$\mathcal{D}$	45% (NA)	45%	45%

Notes: The world interest rate r^* is assumed to be 3%. Tax shares in GDP are obtained from the OECD Tax Revenue database. "LFS" stands for the Polish Labor Force Survey and has been averaged over the 1995-2010 period, "SIF" stands for the annual reports of the Social Insurance Fund (in Polish: Fundusz Ubezpieczen Spolecznych), and "NA" stands for the National Accounts. Further, "homogeneity" refers to absence of intra-cohort heterogeneity and "heterogeneity" refers to its presence. Finally, the final two columns differ slightly from each other when we increase the precision of the reported values.

We use the data on the employment rate from the Labor Force Survey to calibrate the aggregate preference for leisure ϕ. Subsequently, we seek the discount factor δ and the capital depreciation rate d that would be consistent with the national savings rate of approximately 16% of GDP and the investment rate of approximately 21% of GDP, as observed on average between 1990 and 1999. We calibrate the global real interest rate to 3% on an annual basis. Following Schmitt-Grohe and Uribe (2003) we set the coefficient ξ governing the debt-elastic interest-rate premium at 0.034. While this number may be on the high side, if anything, raising the return on domestic investment stacks the setting against our main result that a funded pillar is not politically stable.

Based on the standard figure used in the literature, we assume that the elasticity of output with respect to capital is $\alpha = 30\%$. The share of government expenditure of GDP is set at 20%, to replicate the actual proportion. The capital income tax rate τ_k is set at the *de jure* rate of 19%. The labor income tax rate τ_l is calibrated to replicate the ratio between labor income tax revenue and labor income in the national accounts, thus at its effective rate, because there is no single nominal rate. We calibrate the value of the consumption tax rate τ_c to match the value-added tax revenues as a share of GDP. We also assume that the initial and eventual steady-state government debt-to-GDP ratios equal

---

[17]The eligible retirement age was 55 for women and 60 for men. In 2009 the government terminated early retirement for the majority of the work force, although some occupations remained entitled to early retirement. In 2012 the government imposed a gradual increase of the eligible age to 67 for both genders, but in 2015, before this legislation had effectively come into force, the increase was already reversed, resulting into a current eligible retirement age of 60 for women and 65 for men, with an entitlement to claim partial benefits from the ages of 58, respectively 63.

45%, which corresponds to the actual debt ratio in the late 1990s in Poland. By calibrating the final debt-to-GDP ratio equal to the initial one, we eliminate the effects associated with a change in the long-run debt ratio. Finally, we set the lump-sum tax to close the government budget constraint given by equation (21). The parameter values are summarized in Table 1. To assure a smooth evolution of the consumption tax rate, along the transition path we set $\rho = 0.9$ and $\rho_D = 0.050$.

4.1 Calibrating the intra-cohort heterogeneity

We calibrate the intra-cohort heterogeneity in productivity and time preferences using individual level data.[18]

Individual productivity ω_κ We use linked employer-employee data to reduce biases, such as those associated with self-reporting, inherent in labor force surveys. Therefore, we use the Structure of Earnings Survey, designed by Eurostat and collected biennially by the national statistical offices.[19] This survey covers roughly 700,000 employees, for which all individual characteristics such as age, gender, earnings, sector and occupation, as well as actual hours worked, are reported by the employer. Since in our model individual labor productivity is determined once for the entire lifetime, we use the early (first five) career years to obtain the distribution of these productivities. We estimate a standard Mincerian wage regression with education levels, occupation, industry and region controls, as well as the form of contract (fixed term or permanent) and the form of employment (part-time, full-time, weekends, etc.). We use total hourly wages, including overtime and bonuses. We take the logarithm of the fitted value of the hourly wage and obtain the final distribution of individual productivities depicted in Figure B.2 in Appendix B, which we discretize to the set of values $\omega_\kappa \in \{0.70, 0.76, 0.84, 0.93, 0.98, 1.03, 1.08, 1.14, 1.20, 1.26\}$ relative to a mean of 1.[20]

The discount factor δ_κ The economy-wide discount factor δ is calibrated to replicate the interest rate, given the depreciation rate. As far as we are aware, there is no empirical literature to guide us in calibrating the heterogeneity in the discount factor across agents. We assume symmetric departures from the economy-wide discount factor and calibrate the spread around the latter such that the wealth inequality is replicated by our model. Davies et al. (2011) and a household wealth survey by the National Bank of Poland [21] yield a range for the Gini coefficient of wealth of between 65.7 in 2000 and 57.9 in 2014. We target the mid-point of these two values, assuming that 40% of the individuals have the average discount factor, while 30% of the individuals have a lower- or higher-than-average discount factor. Concretely, $\delta_\kappa \in \langle 0.988\delta, 1.0\delta, 1.012\delta \rangle$, where δ is the economy-wide discount factor reported in Table 1. This calibration yields a Gini coefficient of wealth of 62 in the initial steady state.

---

[18]The two dimensions of heterogeneity are assumed to be independent. Our setup allows for any potential correlation between the different types of heterogeneity. However, empirical research identifying such correlations is scarce. For a discussion on the estimates for household rather than individual data, see, for example, Hénin and Weitzenblum (2005), McGrattan and Prescott (2013), Kindermann and Krueger (2014).

[19]We use data from the 1998 wave to calibrate the distribution of individual productivities.

[20]Whether we express the distribution relative to the mean (log) wage or its median value is quantitatively irrelevant for the distribution of productivities. The results for individual productivity relative to the median are available upon request.

[21]This is a standardized tool developed within the Household Finance and Consumption Network of the European System of the Central Banks.

4.2 Solving the model

Once for a given set of prices and policy parameters the individual's problem is solved for each κ-type of agent at each age j, we apply the Gauss-Seidel algorithm using the updated aggregate variables to obtain the general equilibrium solution. The procedure is repeated until the difference between the paths of aggregate capital from subsequent iterations has become negligible.[22] We set the length of the transition path in order to assure that the latest birth cohort in the simulation lives its entire life in the new steady state.

5 Results

This section studies the political stability of funded social security in our fully-fledged calibrated model. Our analysis will be guided by the event line in Figure 3. In 1999 the initial arrangement of PAYG-DB arrangement in Poland was replaced by a two-pillar DC scheme, after which a transition period started. During this transition, the funded pillar would be built up gradually, while existing PAYG-DB entitlements of the older cohorts would be fully respected. However, in 2012 the unfunding of social security was put into force with the capture of funded pension assets. This also corresponds roughly to the moment when most other CEE countries rolled back their funded social security pillars. Hence, in line with the actual course of events, our analysis studies an unfunding vote in 2012.

In principle, such vote can be counterfactually considered in any year. We report two interesting cases: once the demographic transition is over and once the macroeconomic adjustments are conluded. For the former, five generations from 1999, the demographic transition is concluded $(1999 + 2 \times J)$, which makes 2162 an interesting period to study. By this time, all the original PAYG-DB entitlements have expired and all individuals have lived their entire life-span under the two-pillar social security. Moreover, cohorts that are born later live longer and are less numerous, which should raise the benefit from funding relative to PAYG and thus raise the bar for unfunding. For the latter, by 2192, the macroeconomy and all the budgetary variables under the two-pillar social security have converged to their steady state with two-pillar social security, which makes this year a second interesting case to study for an unfunding vote.[23]

---

[22]Specifically, this is the case as soon as the L_1-norm of the difference between the subsequent iterations of the vector describing the time path of the capital stock falls below 10^{-12}.

[23]Since these votes are counterfactual, they are not a statement that the two-pillar social security lasts unchanged for nearly two hundred years. Rather we study what would have happened in terms of political economy, had the two-pillar social security been still in place by that time.

Figure 3: Event line to guide the structure of our analysis
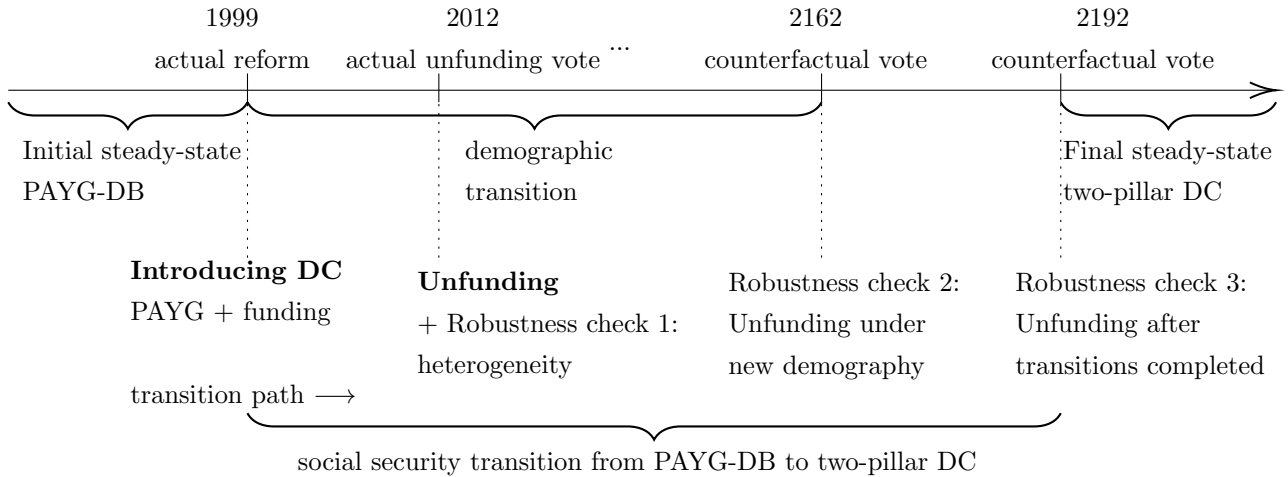
The roadmap to the remainder of the analysis is as follows. Subsection 5.1 describes the political process and the evaluation of the welfare effects of changes in social security. Subsection 5.2 demonstrates that the results obtained with the simple setup in Section 2 carry over to the fully-fledged calibrated model. Using the same equilibrium concept, it demonstrates that a one-shot deviation in which asset capture is postponed from some period t until period $t + k$ is rejected in a vote. We demonstrate this for different voting periods t and different lengths of postponement k. Naturally, the one-shot deviation analysis is able to account for only part of the dynamics in the social security arrangements. Hence, Subsection 5.2 constitutes a bridge between the simple setup in Section 2 and the analysis in Subsection 5.3. In this subsection we demonstrate that, given the absence of initial funding assets, the shift from PAYG-DB to a two-pillar DC system is politically feasible, but we also show that, once a stock of funding assets has been accumulated, the funded pillar is politically not stable when pitched against a shift towards more PAYG-DC social security, irrespective of the moment of the vote. This subsection also delves into the macro-economic consequences of such a shift. Finally, Subsection 5.4 explores the robustness of our findings for variations on the baseline assumptions and extensions in which we allow for intra-cohort heterogeneity and altruism. The presence of intra-cohort heterogeneity allows us to additionally explore how the unfunding social security impinges on poverty.

5.1 The political process and evaluation of welfare effects

Policy outcomes are based on a pure majority vote. In line with Phelan (2006), we assume that individuals do not vote strategically in that (i) they only evaluate alternative outcomes in terms of their own utility and (ii) they do not expect a new vote in the future. Hence, individuals vote for a proposed once-and-for-all policy change if this yields higher individual utility than the *status quo*.[24] Individuals know their type when voting. The welfare effect of a policy change for a type κ individual of age $j \geq 1$ is obtained in terms of a compensating variation. Denote for someone of age j by $V^S_{j,\kappa,t}$

---

[24]Voting is costless in our setup. The literature pays increasing attention to the cost of voting and how it affects electoral turnout (e.g. Borgers 2004, Ghosal and Lockwood 2009, Chakravarty et al. 2018). Voting costs are potentially relevant when they affect participation unevenly across age groups or in other dimensions, such as relative productivity. However, we will see that a substantial fraction of the subcohorts prefer the same policy outcome once we introduce intra-cohort heterogeneity. Hence, introducing the costs of voting is unlikely to alter the political outcome and the evolution of the economy.

the remaining lifetime utility under the status-quo scenario and by $V_{j,\kappa,t}^R$ utility under a reform policy to be introduced in period t. Appendix C derives, under the assumption of logarithmic utility, the compensating variation of someone of age j as the fractional increase in consumption in each of the remaining periods of her life brought about by the reform in period t when compared to the status-quo policy:

$$W_{j,\kappa,t} = 1 - \exp\left\{ \frac{V_{j,\kappa,t}^S - V_{j,\kappa,t}^R}{\sum_{s=0}^{J-1} \delta_\kappa^s \pi_{1+s,t+s}} \right\}, \tag{27}$$

Hence, if $W_{j,\kappa,t} > 0$, the individual benefits from the reform. Appendix C also derives the compensating variation $W_{1,\kappa,t'}$ of someone of subcohort κ entering the economy in period t' after the reform, i.e. $t' \geq t$. Finally, we calculate the (aggregate) welfare effect $W_{t'}$ in some future period $t' \geq t$ as the aggregate over all sub-cohorts κ, where the sub-cohorts are weighed by their shares in the total number of newborns in t:

$$W_{1,t'} = 1 - \exp\left\{ \sum_{\kappa=1}^{\mathcal{K}} \left(\frac{N_{\kappa,t'}}{N_{t'}} \cdot \frac{V_{1,\kappa,t'}^S - V_{1,\kappa,t'}^R}{\sum_{s=0}^{J-1} \delta_\kappa^s \pi_{1+s,t'+s}} \right) \right\}. \tag{28}$$

This is also equal to the expected welfare effect of an individual who enters the economy in period t' but does not yet know her type. We use the same formula to obtain the compensating variation in the final steady state.

5.2 Political stability of funded social security

In this subsection we show that capturing assets is an equilibrium strategy. The equilibrium concept is the same as in Definition 1 in Section 2: voting individuals compare their utility on the equilibrium path in which the funded pillar assets are captured immediately in year t or in period $t+k$, that is with a delay of k years. If individuals prefer to capture assets from the funded pillar immediately, then this confirms that asset capture is an equilibrium, analogous to what we found for the stylized theoretical model. Here, as in Section 2, asset capture by the government from the funded pillar results into the complete termination of this pillar.

The vote takes place, unexpectedly, in $t = 2012$, which coincides with the actual asset capture occurring in Poland and roughly the moment when most other CEE countries rolled back their funded social security pillars. We also illustrate the one-shot deviation strategy once the demographic transition is over (i.e., a vote in $t = 2162$) and once the fiscal adjustments related to the introduction of the two-pillar DC-scheme have come to an end (i.e., a vote in $t = 2192$).
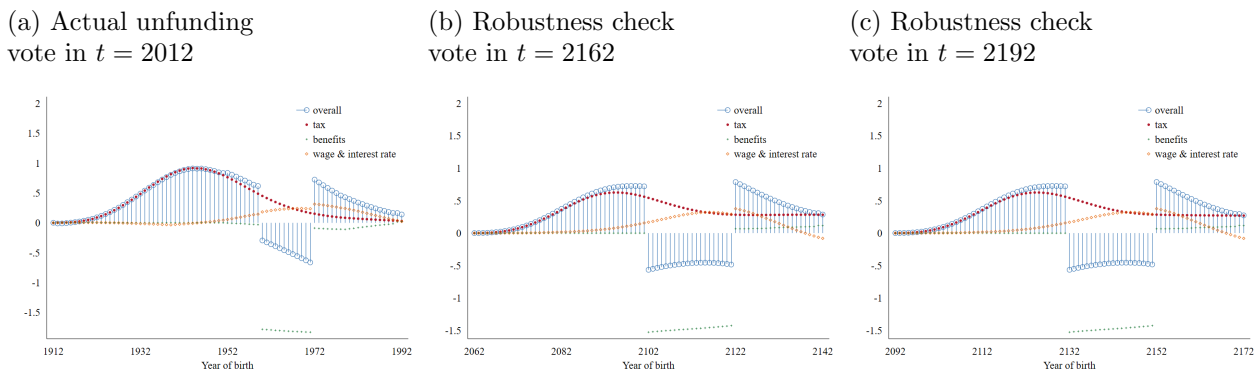
The introduction of the funded pillar leads to costs associated with the need to simultaneously finance the retirement benefits of the current retirees and the accumulation of the own social security assets by the workers. The existing literature has argued that relative to PAYG-DB social security, political stability of funded social security is eventually achieved by the majority of living cohorts having become better off with the two-pillar arrangement once these transition costs have been fully taken care of by the previous cohorts. We thus study the political stability of social security funding when the transition associated with the introduction of the two-pillar DC scheme has been completed. In this subsection we assume that intra-cohort heterogeneity is absent.

Figure 4 depicts for each birth cohort the total welfare effect of capturing assets from the funded

pillar in year t instead of in year $t + k$; welfare is expressed in terms of compensating variation as the percent change in consumption over the remaining lifetime. The figure also decomposes the total welfare effect into the effects attributable to tax changes and benefit changes, which also underlie the result of Proposition 3 based on our simple theoretical framework, and general equilibrium effects. To calculate the component attributable to tax changes, we take the consumption tax from a scenario when assets are captured in $t + k$ and all other variables from a scenario when assets are captured t. Analogously, we obtain the component associated with changes in benefits, which include the cumulative effect of changes in indexation. To obtain the "interest rate & wage" component, i.e., the general equilibrium effect, we calculate the welfare effect of taking the interest rate and the wage rate under asset capture in $t + k$ and all other macroeconomic variables under asset capture in t.

The retirement benefit channel is economically negligible for old individuals, i.e. those born before 1950, who are already retired on voting date t, while the channel is negative for the middle-aged and young individuals. Further, all cohorts currently alive benefit from the reduction in taxes made possible by the asset capture. Quantitatively, the trade-off between taxes and benefits is of key importance, the general equilibrium channel gains significance for younger cohorts.[25] Overall, older and younger individuals always prefer the unfunding policy, whereas middle-aged individuals prefer the unfunding policy depending on the specific features of the economy.

Figure 4: Welfare effects and their components under a one-shot deviation, $k = 20$

(a) Actual unfunding vote in $t = 2012$ (b) Robustness check vote in $t = 2162$ (c) Robustness check vote in $t = 2192$

Note: The figure portrays under intra-cohort homogeneity the welfare effects expressed in terms of compensating variation as the percent change in consumption over the remaining lifetime for the one-shot deviation strategy, in which assets are captured in the indicated year t. We decompose the overall welfare effect into the components attributable to tax changes, benefit changes and general equilibrium (i.e., "wage & interest rate") effects. Their construction is described in the main text. Due to the curvature of the utility function the total welfare effect is not a simple sum of the partial effects.

Table 2 summarizes the political support among those alive for unfunding social security in period t rather than in period $t+k$. It does so for the baseline of $t = 2012$ and $k = 20$ as well as for robustness checks associated with votes on a later date and different delays of unfunding, k. Table 2 shows that for all cases considered a majority of the voters on date t prefer to unfund social security immediately,

---

[25]In general, the voters prefer living in an economy with a large capital stock, assuming the economy remains dynamically efficient. A higher capital stock raises the marginal productivity of labor. However, the benefits from further capital accumulation differ by age at the moment of the vote. Older individuals – those retired or close to retirement – care more about the return on capital than about the wage rate, because their consumption depends more on income from holding capital. Unfunding lowers the capital-to-labor ratio as a result of a fall in the capital stock and raises the labor supply among the younger cohorts who need to save more for their old age. The return on capital increases. Even if these effects are transitory, they raise the support for unfunding among the older cohorts. Meanwhile, younger individuals, who still expect to work for many years at the time of vote, suffer from the reduced marginal productivity of their labor caused by the unfunding.

Table 2: Political support for immediate and complete unfunding of social security

Political support for immediate unfunding % of those alive	Actual unfunding in $t = 2012$	Robustness check vote in $t = 2162$	Robustness check vote in $t = 2192$
$k = 4$	79%	94%	93%
$k = 10$	84%	86%	86%
$k = 20$	84%	73%	73%
$k = 30$	87%	96%	97%

Note: The table reports the fraction of individuals alive in year t who prefer to unfund social security immediately rather than in period $t+k$. The political support in t is calculated according to $\sum_{s=1}^{J} \mathbf{1}\left\{W_{\kappa,t-(s-1)} > 0\right\} \frac{N_{s,\kappa,t}}{N_t}$, where s counts over the cohorts alive at the time of vote and $\mathbf{1}\left\{W_{\kappa,t-(s-1)} > 0\right\} = 1$ if the condition in brackets holds, and zero, otherwise.

rather than delay this to period $t+k$. The reason is that immediate unfunding allows for an immediate reduction in taxes, while a delay to $t+k$ would delay the tax reduction, which dominates the benefit of a somewhat higher social security benefit from delay.

5.3 Unfunding in Poland: macroeconomic effects, welfare and political support

We now explore the macroeconomic and welfare effects of, and political support for, social security unfunding in Poland during early 2010s. Poland provides a typical example of the pattern of changes enacted in the social security systems of the CEE countries. In this subsection we continue to assume intra-cohort homogeneity. The next subsection turns to the case of intra-cohort heterogeneity.

As we discuss below, one can demonstrate that there is a majority support for introducing funding. The majority hinges on the stock of social security assets being zero at the moment of this decision. However, the accumulation of these assets changes the trade-off faced by the voters. Therefore, starting from the *status-quo* of the two-pillar DC scheme introduced earlier, we pitch continuation with this status-quo against three alternative arrangements to unfund social security.

First, we study the case of diverting contributions from the funded pillar to the PAYG-DC pillar, while keeping the existing stock of funded-pillar assets untouched. We refer to this alternative as "Policy 1". It mimics the type of changes that have been temporarily or permanently implemented by all CEE countries in the aftermath of the global financial crisis. Because in most of these countries part of the contributions continued to be transferred to the funded pillar, we assume that the adoption of Policy 1 changes the shares of the total contribution to the two pillars of social security: prior to the vote, $\tau_t^P = 2\tau_t^F$, whereas after a successful vote the new proportions (indicated by a tilde) become $\tilde{\tau}_t^P = 9\tilde{\tau}_t^F$, while the total contribution rate is kept at the level before the vote, i.e. $\tau_t^P + \tau_t^F = \tilde{\tau}_t^P + \tilde{\tau}_t^F = \tau$.[26] Policy 1 assumes that this shift will be permanent, resulting into a permanently smaller funded pillar and larger PAYG pillar.

"Policy 2" keeps the contributions to each of the two pillars unchanged, but transfers the funded-pillar assets to the PAYG pillar. Since the government balances the PAYG social security pillar, this asset capture effectively amounts to a large transfer from the funded social security pillar to the government budget constraint. Such changes have been implemented by Hungary, Bulgaria, Poland and Slovakia.[27] This policy is transitory in the sense that, while the accumulated assets are captured,

---

[26]This shift in the proportions for the two pillars vote replicates the Polish case. The analysis can easily be extended to any shift of policy relevance.

[27]The share of the assets captured varies between these countries. In Hungary the nationalization of the accumulated
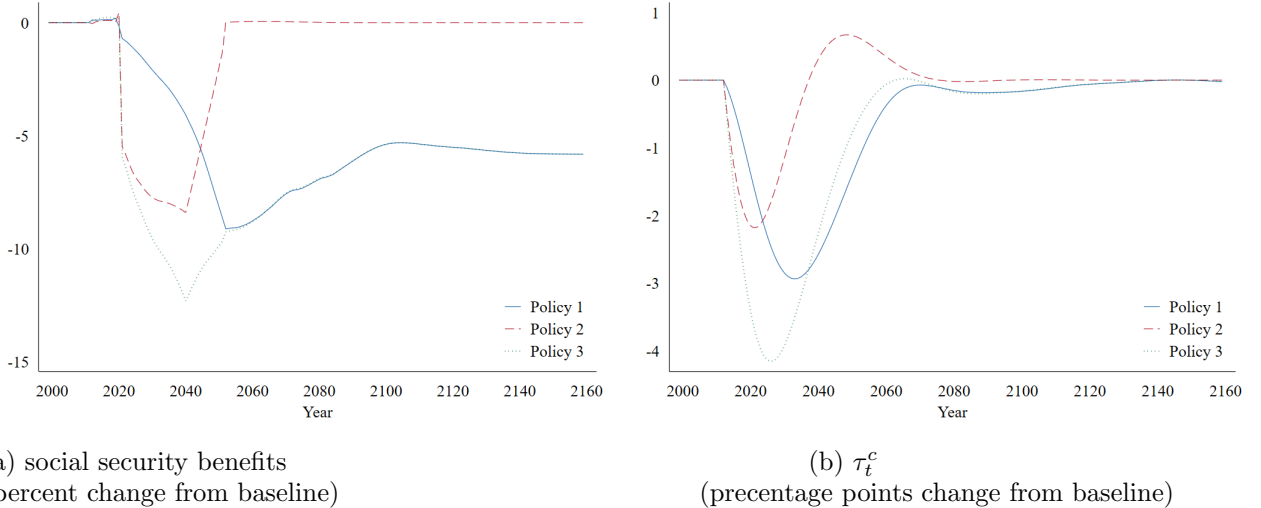
new asset accumulation in the funded pillar resumes as of the next period.

Because in some countries Policies 1 and 2 were implemented simultaneously, "Policy 3" combines Policies 1 and 2, hence it combines the shift in contributions and the asset capture in the same vote, thereby combining permanent with transitory effects.[28] For the same reasons as laid out in Subsection 5.2, we study the effects of a vote not only in the year $t = 2012$ of the actual start of the unfunding, but also in $t = 2162$ and $t = 2192$; recall Figure 3.

Before we study the political support for switching to Policies 1 – 3, we discuss their implications for the public budget and the economy (for more details see Appendix E). Policy 1 reduces the current deficit of the PAYG scheme. With general government consumption fixed, a reduction in $Subs_t$ allows for a reduction in taxes relative to the current two-pillar DC scheme. Meanwhile, since the return of the PAYG pillar is lower than of the funded pillar, Policy 1 results in permanently lower future retirement benefits compared to the current two-pillar DC scheme. Policy 2 reduces immediately the debt of the government. This allows current and future taxes to be reduced, because also the subsequent costs of servicing the public debt fall. Again, this comes at the cost of lower retirement benefits.

Figure 5 portrays the adjustments in retirement benefits and taxes. Relative to the current two-pillar DC scheme, the adjustment in taxes is rather sizable, but the decline in benefits is reasonably moderate, because it is effectively determined by the difference in the return on the funded social security assets and the growth of the wage bill. Moreover, as we discuss below, lower retirement benefits stimulate voluntary private savings and the labor supply, thereby ameliorating the fall in benefits.

Figure 5: The effects of unfunding on benefits and taxes relative to the current two-pillar DC scheme

(a) social security benefits
(percent change from baseline)

(b) τ_t^c
(precentage points change from baseline)

Note: Results from a policy shift in 2012 (results for alternative shift years are available upon request). Policy 1 is a permanent shift of contributions from the funded to the PAYG pillar. Policy 2 is a one-off capture of the funde-pillar assets. Policy 3 combines Policies 1 and 2. The levels of retirement benefits and taxes as well as debt adjustments are reported in Figure E.7 in Appendix E.

Lower benefits due to the reduction of the funded pillar result in higher voluntary savings. However,

---

assets was immediate and complete, whereas in the other countries it was partial and gradual.

[28]The experiment in Subsection 5.2 effectively corresponds to Policy 3, with the additional constraints that $\tilde{\tau}_t^P = \tau$ and $\tilde{\tau}_t^F = 0$.

since voluntary savings are taxed and cannot be converted into an annuity at retirement age, which yields a higher return due to those who die, this increase fails to make up for the effect of unfunding on the capital stock – see Figure 6 for a policy shift in 2012.

Figure 6: Adjustments in capital K_t (left) and labor supply L_t (right)
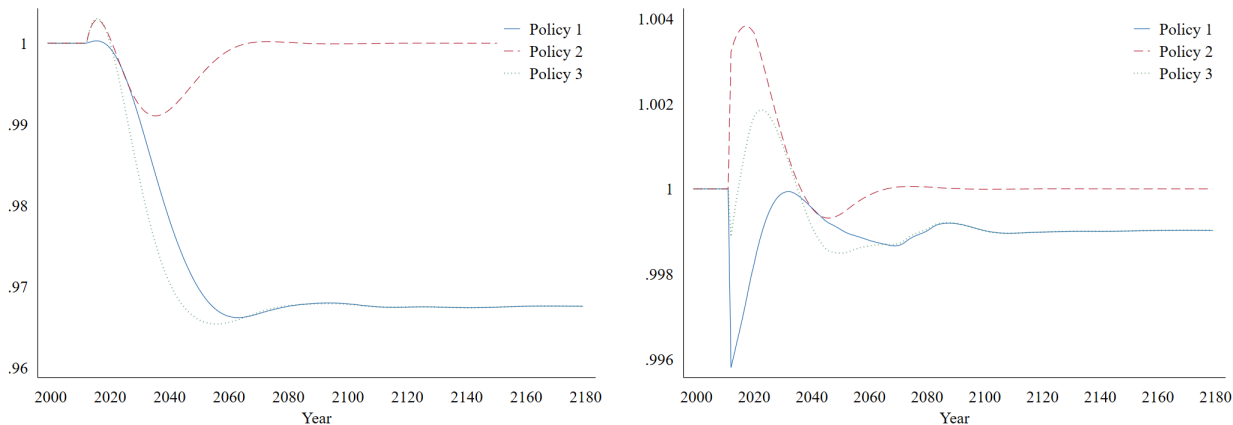
Notes: Adjustments are expressed as ratios relative to the current two-pillar DC scheme, assuming the unfunding policy switch takes place in 2012. Policy 1 is a permanent shift of contributions from the funded to the PAYG pillar. Policy 2 is a one-off capture of the funde-pillar assets. Policy 3 combines Policies 1 and 2. Results for alternative switching years are available upon request.

The impact on the labor supply is twofold. First, lower consumption taxes distort less the intra-temporal choice between consumption and leisure, which tends to increase the labor supply. Second, due to the shift away from funding, social security contributions yield a lower return in terms of future benefits, hence the incentive to work is reduced. Depending on which of the effects is stronger, the labor supply rises or falls. For Policy 1 the second effect dominates. The labor supply drops rapidly following its introduction. For Policy 2 the first effect is stronger and the labor supply increases. Finally, for Policy 3 the effect of the lower consumption tax rate dominates initially, as the decline in taxes is largest in the initial years after Policy 3 is introduced, but over time this effect is overtaken by the effect of the distortion associated with the second effect; see Figure 6 for a policy shift in 2012.

Table 3 summarizes the political support for each of the Policies 1 – 3 in a vote against the current two-pillar DC scheme, as well as the aggregate welfare effect of the switch to these policies. The static comparison exhibits the lifetime welfare effect of introducing Policies 1 – 3 in the final steady state reached under the current two-pillar DC scheme. This long-run welfare effect is a lifetime consumption loss of -0.45% for Policies 1 and 3, while Policy 2 has no long-term welfare effect, because it involves a one-off asset capture from the funded pillar in year t. Even though Policies 1 and 3 reduce the lifetime utility of the newborns in the long-run when compared to maintaining the current two-pillar DC scheme, at the moment of the vote these policies improve the situation of a majority of the then-alive at the expense of future cohorts. This is also the case for Policy 2. Hence, ultimately, the funded pillar never becomes politically stable.

For alternative voting years 2012 and 2192 Figure 7 decomposes by cohort the overall welfare effect into the effects from changes in taxes and retirement benefits, and a general equilibrium effect. The vote in 2012 takes place during the early phase of the transition under the current two-pillar DC scheme,

Table 3: Political support for the three alternatives to the current two-pillar DC scheme

Year of vote t	2012	2162	2192	final steady state
Political support in % of those alive	Against current two-pillar DC scheme			
for Policy 1 - shifting contributions (permanent)	98	100	100	
for Policy 2 - capturing assets (one-off)	55	65	65	
for Policy 3 - combination of Policies 1 and 2	100	100	100	
Long-term welfare effect of Policy 1				-0.45%
Long-term welfare effect of Policy 2				0.00%
Long-term welfare effect of Policy 3				-0.45%

Notes: (i) Under Policy 1 the share of contributions to the funded pillar is reduced, while the overall contribution rate stays the same. Policy 2 is a one-off capture in period t of the assets accumulated in the funded pillar. Policy 3 combines Policies 1 and 2, i.e. it amounts to a permanent reduction of the contributions to the funded pillar and a one-off capture of the assets in the funded pillar.
(ii) The long-term welfare effect is the lifetime welfare effect of introducing Policies 1 – 3 in the final steady state reached under the current two-pillar DC scheme.

while in 2192 the steady state would have been reached that corresponds to the eventual stationary demographic structure (with a shrinking population) and completion of the budgetary transition from the introduction of the current two-pillar DC scheme. The decomposition is analogous to that in the previous subsection.

Figure 7: Welfare effect over remaining lifetime and its components by birth year for vote in 2012 or 2192

(a) 2012, Policy 1: permanent shift contributions
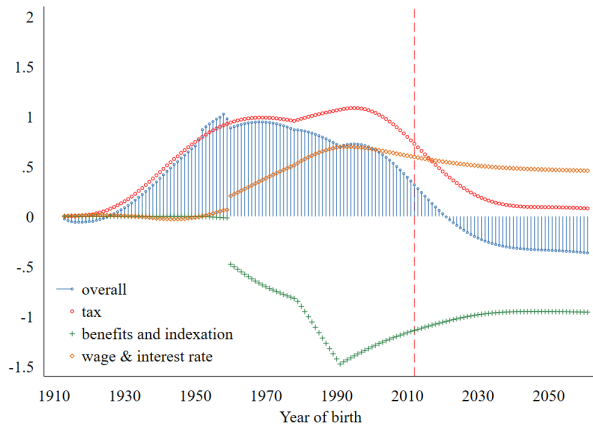
(b) 2192, Policy 1: permanent shift contributions
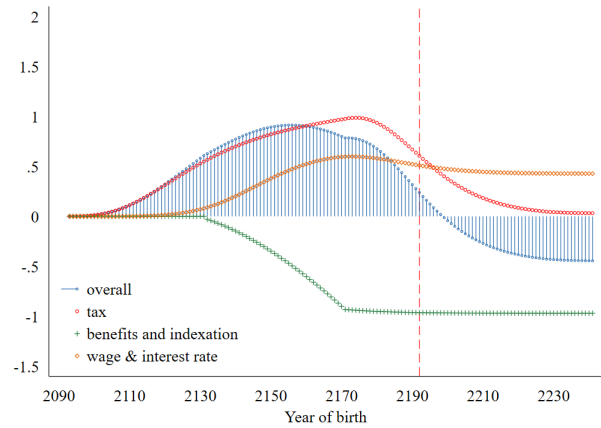
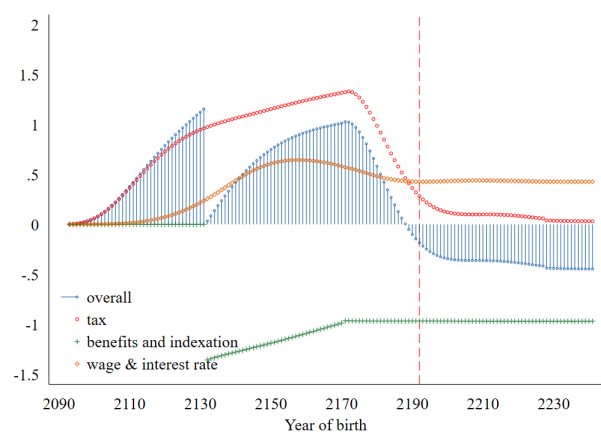(c) 2012, Policy 2: one-off asset capture
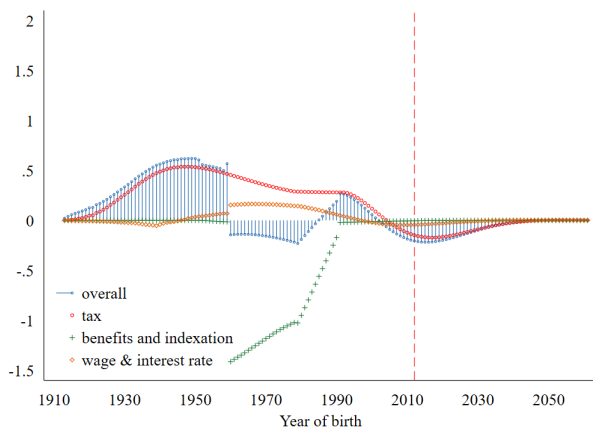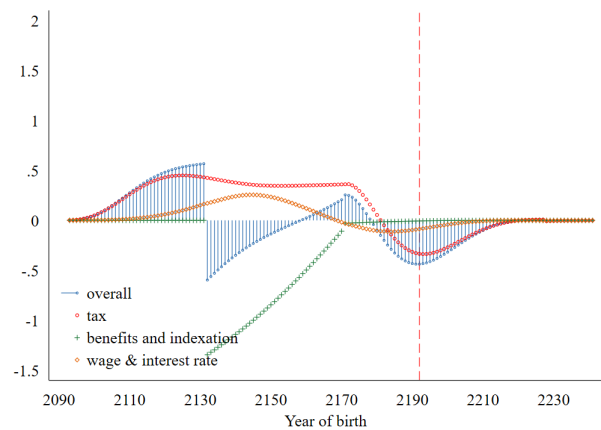
(d) 2192, Policy 2: one-off asset capture

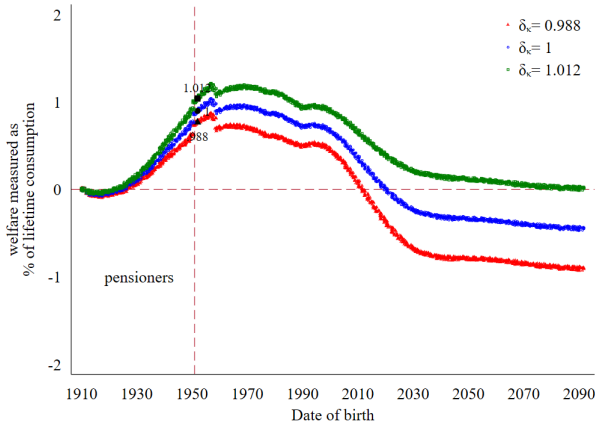(e) 2012, Policy 3: combination Policies 1 and 2
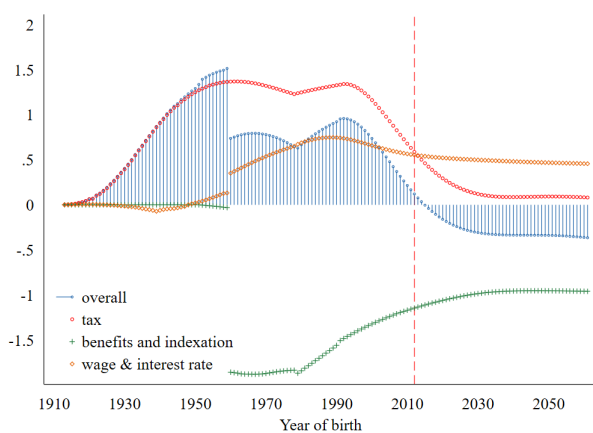
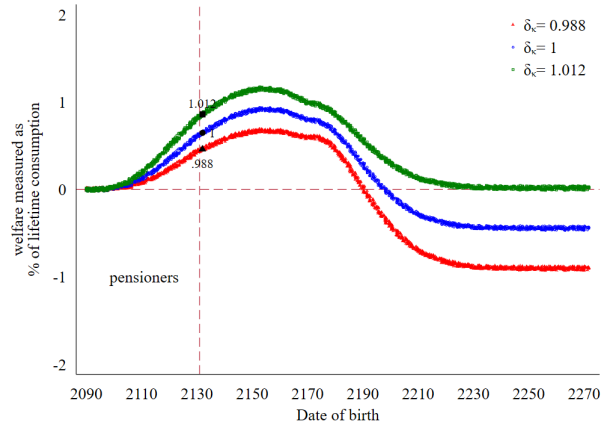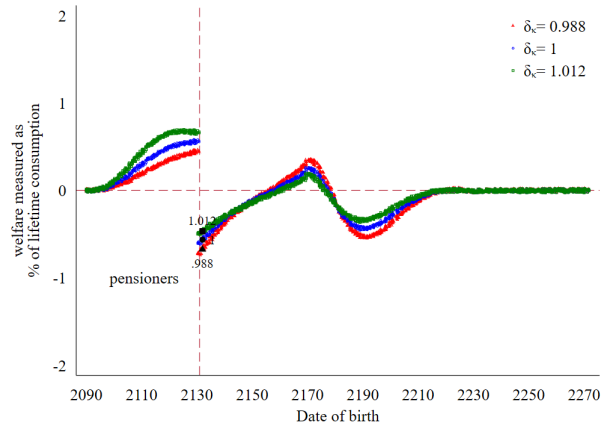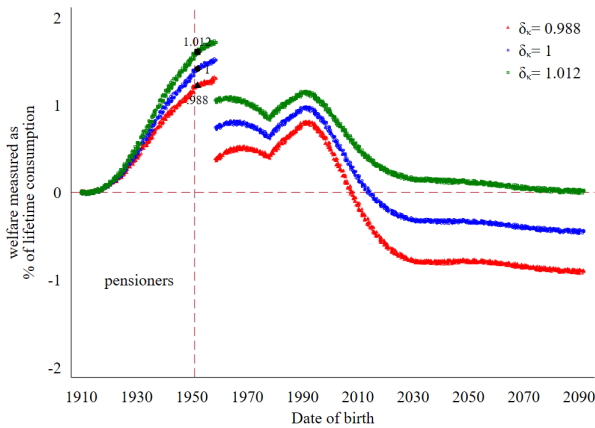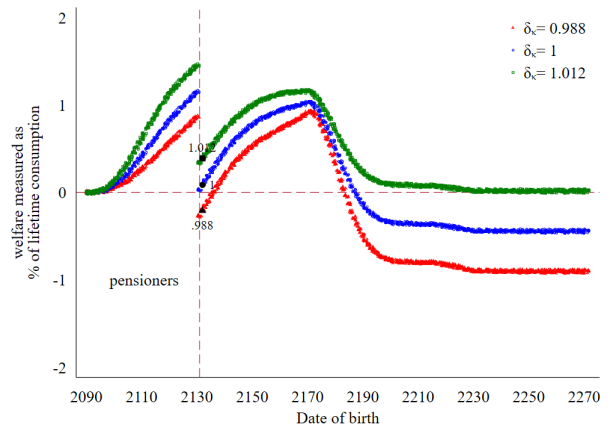(f) 2192, Policy 3: combination Policies 1 and 2

Note: The decomposition of the aggregate welfare effect is described in Subsection 5.2. Policy 1 is a permanent shift of contributions from the funded to the PAYG pillar. Policy 2 is a one-off capture of the funde-pillar assets. Policy 3 combines Policies 1 and 2.

Figure 8: Role of intra-cohort heterogeneity for the dispersion of the welfare effects, vote in 2012 or 2192

(a) 2012, Policy 1: permanent shift contributions

(b) 2192, Policy 1: permanent shift contributions

(c) 2192, Policy 2: one-off asset capture

(d) 2192, Policy 2: one-off asset capture

(e) 2012, Policy 3: combination Policies 1 and 2

(f) 2192, Policy 3: combination Policies 1 and 2

Notes: Policy 1 is a permanent shift of contributions from the funded to the PAYG pillar. Policy 2 is a one-off capture of the funded-pillar assets. Policy 3 combines Policies 1 and 2. We report the welfare effects across subcohorts of Policies 1 – 3 in terms of a compensating variation as a percentage of remaining lifetime consumption relative to the baseline of continuing with the current two-pillar DC social security. Intra-cohort heterogeneity takes the form of differences in the degree of patience (the δ multipliers) and productivity (the ω multipliers). The dispersion within each dotted plot reflects the heterogeneity of welfare effects across productivity types.

For the cohorts born after the vote, Policy 1 leads to a net welfare loss, resulting from lower retirement benefits dominating the beneficial tax and general-equilibrium effects. Policy 1 improves welfare of almost all living cohorts, and almost solely via the change in taxes. The benefit from reduced taxes dominates the cost in terms of a lower retirement benefit. If the switch to Policy 1 is in addition used to reduce the public debt, the lower debt-servicing costs may also benefit cohorts born in the near future. Adoption of Policy 2 following a vote in 2012 leads to a more varied pattern in which the youngest cohorts and the future born are worse off, while it benefits retirees and individuals who were not affected by the initial 1999 reform and have little or no assets in the funded pillar. Reduced taxation is essentially the only source of the net gains. Net losers are the cohorts with the largest funded pillar assets and who are close to retirement now, i.e. who were younger than 40 in 1999. The pattern for Policy 3 is rather similar to that for Policy 1. We observe a similar pattern for a 2192 vote.

As a final note, observe that the introduction in 1999 of DC social security with a PAYG and a funded pillar can be demonstrated to improve welfare in a Kaldor-Hicks sense and receive majority support at the moment of its introduction. The details of the analysis are found in Appendix D, in which we show that, compared to a scenario of maintaining single-pillar PAYG-DB social security, the initial reform into a two-pillar DC scheme has more than 50% support in our setup.[29] We assume that the total contribution rate to social security remains unchanged, i.e. the reform merely shifts the contributions away from the PAYG pillar to the funded pillar,[30] and honors the retirement benefits obligations to the older generations.

5.4 Robustness and extensions

This subsection addresses the robustness of our findings to some of the assumptions of the model. Of some of the variants we consider we only summarize the outcome, while the detailed results are available upon request. First, we demonstrate that the demographic path and, in particular, the fact that the population is non-stationary, also in the final steady state, is inconsequential for our finding that social security funding is politically unsustainable. Demographic projections show a declining fertility for Poland, and this decline continues beyond the end of the projection horizon in 2080. We study an alternative demographic structure with the size of the newborn cohort reproducing in each period throughout the simulation period, hence the population no longer shrinks.[31] In such setup, introducing a funded pillar still gets sufficient political support, but so does unfunding afterwards. Second, we explore whether the results are sensitive to the share of the asset capture that is used to reduce taxes versus reducing public debt, and find that even if the largest part of assets capture is used to reduce the public debt, political support for the three policies remains, because the lower debt

---

[29]The existing literature argues that, after the initial transition costs are covered, the cohorts benefiting from a shift to funded social security would eventually form a majority, thus suggesting that the two-pillar DC scheme would be politically stable in the long term. Our model replicates this prediction; we find that the share of the population benefiting from the original shift in 1999 to two-pillar DC social security declines for some time after which it picks up again. As of 2060, i.e. once a sufficient share of newborn cohorts populates the economy, again the majority of the cohorts benefit from this shift (i.e., compared to maintaining the original one-pillar DB scheme). Importantly, unfunding in our model does not happen because the support for the original shift to two-pillar DC loses a majority. Unfunding happens because the trade-offs are fundamentally altered once funding has been introduced, specifically because contributions to the funded pillar can be diverted and/or because accumulated assets can be captured.

[30]Hence, the predictions of studies such as Browning (1975) and Butler (2000) are not relevant in our case (see also Congleton et al. 2013, for an additional treatment).

[31]With each newborn cohort of the same size, the population grows slightly as longevity progresses.

translates into lower taxes at some moment in an individual's life. Third, we analyze the role of the rate of technological progress. Throughout the analysis we rely on projections which assume Poland to catch up relatively fast in terms of TFP growth and continue on a lower growth path later on. The alternative assumption that TFP growth does not slow down leaves our results unaffected.

The results present so far were based on the assumption that intra-cohort heterogeneity is absent and, hence, differences in voting preferences are determined by the specific cohort in which an individual is born. Now, we explore the political sustainability of the current two-pillar DC scheme in the presence of intra-cohort heterogeneity. At the moment of the vote, each voter knows in which subcohort she is located. The presence of such intra-hohort heterogeneity is motivated by the fact that earlier literature has demonstrated the role of coalitions across cohorts and productivity types when voting about redistribution through the design of the social security arrangement – see Aaron (1966), Cooley and Soares (1999a) and Tabellini (1991). Hence, by allowing for intra-cohort heterogeneity we can explore whether such coalition formation affects the political stability of the current two-pillar DC scheme also in our setup. As described above we introduce a rich amount of intra-cohort heterogeneity, allowing for a wide dispersion in productivity and in patience (i.e., in time preference rates).
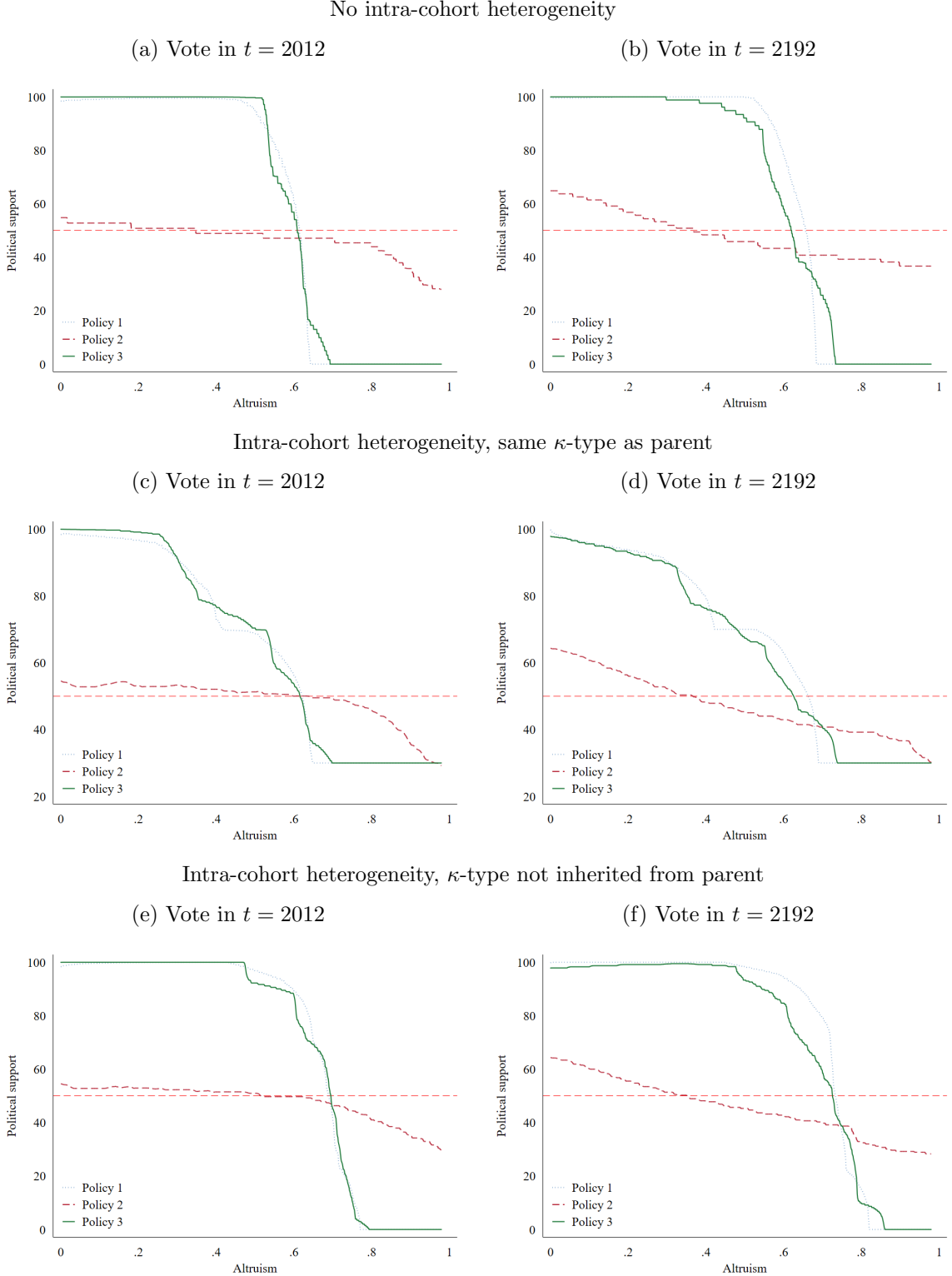
Figure 8 depicts the welfare effects of a switch from the current two-pillar DC scheme to Policies 1 - 3. The left-hand panels of the figure do so for a vote in 2012, whereas the right-hand panel does so for a vote in 2192. Of the dimensions of the intra-cohort heterogeneity, the dispersion in patience is quantitatively by far the most relevant one. This is not surprising, because the most important trade-off is that between the gains from lower current taxes and future efficiency losses and lower social security benefits. The latter obviously receive a higher weight when individual patience increases. Indeed, individuals with the highest degree of patience experience the lowest benefit, or largest loss, of the adoption of one of the Policies 1 – 3. The welfare effects of differences in individual productivity are barely discernible. To establish the outcome of the majority vote, we aggregate over all cohorts and subcohorts, each with its own weight in the total voting population. We observe that the outcome of the majority vote is the same in all cases. That is, if one of the Policies 1 – 3 is pitched against the current two-pillar DC scheme, the latter loses, irrespective of whether the vote takes place in 2012 or 2192. We also observe that the support for Policies 1 – 3 is in all cases essentially the same as in the absence of intra-cohort heterogeneity. See Table F.2 in Appendix F.

Intra-cohort heterogeneity also allows us to study the effects of reducing the funded social security pillar on poverty. We consider two measures of poverty. The first is an absolute poverty measure: we set the threshold for absolute poverty at 60% of median consumption under the initial steady state, i.e. the situation under PAYG-DB before the introduction of the two-pillar DC scheme. The second measure is a relative poverty measure, defined as the fraction of individuals living on less than 60% of median consumption in a given policy scenario. The more extensive analysis is relegated to Appendix G. Here, we confine ourselves to reporting the key findings. First, differences in absolute poverty across the different policy regimes (the original PAYG-DB system, the current two-pillar DC scheme and Policies 1 –3) are small. Second, the shift from PAYG-DB to two-pillar DC has a strong positive effect on the share of old-age households that cannot afford consumption above 60% of the median. Policies 1–3 do very little to alleviate the increase in relative poverty.

Except for the parametrization, we can also study the relevance of our behavioral assumptions. In

particular, in our baseline setting individuals are not altruistic, i.e. when voting they are not concerned about the welfare of the future cohorts (e.g. Fisman et al. 2017). Here we explore the case in which voters are altruistic to a certain extent. Concretely, individuals cast their vote on the basis of an altruistic utility function that is the sum of their own utility over consumption and leisure, as used so far, plus (with some weight) the utility over consumption and leisure of their off-spring. Figure 9 depicts how the political support for Policies 1 - 3 depends on the degree of altruism. The figure depicts the case in which intra-cohort heterogeneity is absent and the case in which there is such heterogeneity. The latter case distinguishes between the one in which the subcohort of the off-spring is unrelated to that of the parent and the one in which the off-spring inherits the parent's subcohort. One observes that the degree of altruism does affect the voting outcome: if, roughly speaking, individuals attach a weight relative to their own utility from consumption and leisure of slightly over a half or more to their offspring's utility, unfunding is no longer politically viable. Then, the combination of little or no benefit from reduced taxation and the cost of reduced social security benefits experienced by its offspring from unfunding dominates the individual's direct benefit in terms of own consumption and leisure. This finding is common across the various cases in Figure 9.

Figure 9: Political support for unfunding as a function of degree of altruism

No intra-cohort heterogeneity

(a) Vote in $t = 2012$

(b) Vote in $t = 2192$

Intra-cohort heterogeneity, same κ-type as parent

(c) Vote in $t = 2012$

(d) Vote in $t = 2192$

Intra-cohort heterogeneity, κ-type not inherited from parent

(e) Vote in $t = 2012$

(f) Vote in $t = 2192$

Notes: Denote by $U_{\kappa,t}$ the lifetime utility derived by an individual born in period t, as specified in equation (12). Let parameter $\mathcal{A}$ measure the degree of inter-generational altruism, with $\mathcal{A} = 0$ denoting no altruism. Then in the case of full persistence of heterogeneity from parent to offspring we have $U_{\kappa,t}^{a,1} = U_{\kappa,t} + \sum_{i=1}^{\infty} \mathcal{A}^i U_{\kappa,t+30i}$, which is the utility of an altruistic individual who discounts the utility of future generations with an inter-generational discount factor equal to $\mathcal{A}$ and whose offspring is born when the individual is 30 years old. If we alternatively assume that there is no transmission of heterogeneity across generations, the altruistic utility function becomes $U_{\kappa,t}^{a,2} = U_{\kappa,t} + \sum_{i=1}^{\infty} \mathcal{A}^i E(U_{\kappa,t+30i})$. For a vote in period t, we consider $U_{\kappa,t}^{a,1}$ or $U_{\kappa,t}^{a,2}$ for $t \in (t - 100, t - 20)$, because those are the individuals alive at the time of the vote. The horizontal line corresponds to 50% of the voting population. Finally, Policy 1 is a permanent shift of contributions from the funded to the PAYG pillar. Policy 2 is a one-off capture of the funded-pillar assets. Policy 3 combines Policies 1 and 2.

6 Conclusions

The political stability of social security in the literature rests on the idea that if working cohorts refuse to finance the retirement benefits of their elderly they are penalized with the loss of their own future benefits. However, while this mechanism is plausible for PAYG social security, this is not the case for funded systems. The conventional wisdom is that, although those alive at the moment of the introduction of social security funding may suffer a welfare loss, the future cohorts will eventually experience an increase in lifetime welfare, while overall the reforms can also improve welfare in a Kaldor-Hicks sense. The working cohorts need to finance the retirement benefits of the elderly, while at the same time they have to accumulate assets for their own retirement. The gains from introducing a funded pillar are related to the direct effects of more capital accumulation and the indirect general equilibrium effects through changes in wages and capital returns. Consequently, introducing funding is typically portrayed as producing immediate costs and delayed gains. The literature suggests that as time passes, the fraction of living agents benefiting from the reform increases and, hence, there will come a moment when (partially) funded social security has become politically stable. In this paper, we show that this intuition is not correct.

We show that, even if funded social security has sufficient political support at the moment of its introduction, under rather general circumstances it is politically unsustainable in a future majority vote. The policy relevance of our study is immediate. Since introducing social security funding implies an unequal distribution of the costs and benefits across cohorts, with future cohorts benefiting on net more, it is crucial for such a reform to be politically sustainable. Otherwise, the reform may be reversed even before society starts enjoying its benefits, implying a massive inefficiency: costs are incurred without experiencing any gains. As is clear from recent reversals of social security funding in countries such as Poland, such risks are not merely theoretical.

For our analysis we develop a fully-fledged and carefully calibrated OLG model of an economy undergoing a shift from PAYG-DB to DC with partial funding. Because in the model the obligations associated with the social security benefits of the cohorts in or close to retirement are honored, the reform generates costs to other cohorts currently alive. Allowing public debt to partially smooth these costs generates majority support for the transition. As time passes, the reform becomes beneficial to all cohorts at the moment of their birth: it allows for faster accumulation of capital and, thanks to the funded pillar, a rise in total social security benefits relative to the original one-pillar PAYG-DB system. On a selected date we allow individuals to vote once-and-for-all on diverting funded contributions towards the PAYG-DC pillar and/or capturing the assets accumulated in the funded pillar, while keeping the total social security contribution rate constant.

However, as we shift the date of the vote further into the future, the current two-pillar DC scheme never becomes politically sustainable, even once the transition from the original reform has been completed: the currently alive always benefit from unfunding and shifting its cost to the future cohorts. This finding also holds if we introduce substantial intra-cohort heterogeneity in terms of productivity and patience. Only in the presence of sufficiently strong altruism, is the result overturned and is social security funding politically sustainable.

One might have the impression that the share of contributions going to the funded pillar has simply

been set at a too high level, so rational agents adjust it downwards to their preferred, lower level, in the spirit of Browning (1975) or Cooley and Soares (1999b). However, we demonstrate that the support for unfunding is not driven by better aligning mandatory savings for retirement with individual savings preferences, but rather by shifting welfare from future to current cohorts. The transitory tax reduction is the main driving force in this regard.

In a broader context, our findings should be interpreted as supporting the view that the protection of property rights is key in maintaining multi-pillar state-run social security. Indeed, long-run credibility issues need to be addressed explicitly and ex ante, because the risk of an unfunding policy is always there. In some advanced economies, funded pension pillars are part of a tripartite agreement among employers, employees and the government.[32] Consequently, the property rights associated with pension assets are set at a par with the property rights of other financial instruments. By contrast, the 1990s wave of privatizations, with governments in a role of collecting contributions from workers and transferring these to social security funds, made the funded pillars merely an element of a social contract rather than a proper financial instrument.

---

[32]Here, stablity of funded pensions may depend critically on whether participation in the funded pillar is mandatory for employees; see Beetsma et al. (2012).

References

Aaron, H.: 1966, The social insurance paradox, *The Canadian Journal of Economics and Political Science/Revue canadienne d'Economique et de Science politique* **32**(3), 371–374.

Beetsma, R. M., Romp, W. E. and Vos, S. J.: 2012, Voluntary participation and intergenerational risk sharing in a funded pension system, *European Economic Review* **56**(6), 1310–1324.

Belan, P. and Pestieau, P.: 1999, Privatizing social security: A critical assessment, *The Geneva Papers on Risk and Insurance Issues and Practice* **24**(1), 114–130.

Boadway, R. W. and Wildasin, D. E.: 1989a, A median voter model of social security, *International Economic Review* pp. 307–328.

Boadway, R. and Wildasin, D.: 1989b, Voting models of social security determinatio, *Contributions to Economic Analysis*, Vol. 179, Elsevier, pp. 29–50.

Boldrin, M. and Rustichini, A.: 2000, Political Equilibria with Social Security, *Review of Economic Dynamics* **3**(1), 41–78.

Boll, S., Raffelhüschen, B. and Walliser, J.: 1994, Social security and intergenerational redistribution: A generational accounting perspective, *Public Choice* **81**(1-2), 79–100.

Bonoli, G. and Shinkawa, T.: 2006, *Ageing and pension reform around the world: evidence from eleven countries*, Edward Elgar Publishing.

Borgers, T.: 2004, Costly voting, *American Economic Review* **94**(1), 57–66.

Breyer, F.: 1989, On the intergenerational Pareto efficiency of pay-as-you-go financed pension systems, *Journal of Institutional and Theoretical Economics (JITE)/Zeitschrift für die gesamte Staatswissenschaft* **145**(4), 643–658.

Browning, E. K.: 1975, Why the social insurance budget is too large in a democracy, *Economic Inquiry* **13**(3), 373–388.

Butler, M.: 2000, The political feasibility of pension reform options: the case of Switzerland, *Journal of Public Economics* **75**(3), 389–416.

Casamatta, G., Cremer, H. and Pestieau, P.: 2001, Demographic shock and social security: A political economy perspective, *International Tax and Public Finance* **8**(4), 417–431.

Chakravarty, S., Kaplan, T. R. and Myles, G.: 2018, When costly voting is beneficial, *Journal of Public Economics* **167**, 33 – 42.

Chlon, A., Gora, M. and Rutkowski, M.: 1999, Shaping pension reform in Poland: security through diversity, *Social Protection Discussion Papers 20852*, The World Bank.

Conde-Ruiz, J. I. and Galasso, V.: 2005, Positive arithmetic of the welfare state, *Journal of Public Economics* **89**(5-6), 933–955.

Conesa, J. C. and Garriga, C.: 2008, Optimal fiscal policy in the design of social security reforms, *International Economic Review* **49**(1), 291–318.

Conesa, J. C. and Krueger, D.: 1999, Social Security Reform with Heterogeneous Agents, *Review of Economic Dynamics* **2**(4), 757–795.

Congleton, R., Batinti, A., Bose, F., Kim, Y. and Pietrantonio, R.: 2013, *Public choice and the modern welfare state*, Edward Elgar Publishing, pp. 362–381.

Cooley, T. F. and Soares, J.: 1996, Will social security survive the baby boom?, *Carnegie-Rochester Conference Series on Public Policy* **45**(1), 89–121.

Cooley, T. F. and Soares, J.: 1999a, A positive theory of social security based on reputation, *Journal of Political Economy* **107**(1), 135–160.

Cooley, T. F. and Soares, J.: 1999b, Privatizing social security, *Review of Economic Dynamics* **2**, 731–755.

Davies, J. B., Sandström, S., Shorrocks, A. and Wolff, E. N.: 2011, The Level and Distribution of Global Household Wealth, *Economic Journal* **121**(551), 223–254.

de Walque, G.: 2005, Voting on pensions: A survey, *Journal of Economic Surveys* **19**(2), 181–211.

Epple, D. and Romano, R. E.: 1996, Ends against the middle: Determining public service provision when there are private alternatives, *Journal of Public Economics* **62**(3), 297–325.

Fehr, H.: 2009, Computable stochastic equilibrium models and their use in pension-and ageing research, *De Economist* **157**(4), 359–416.

Fisman, R., Jakiela, P. and Kariv, S.: 2017, Distributional preferences and political behavior, *Journal of Public Economics* **155**, 1–10.

Galasso, V. and Profeta, P.: 2002, The political economy of social security: a survey, *European Journal of Political Economy* **18**(1), 1–29.

Ghosal, S. and Lockwood, B.: 2009, Costly voting when both information and preferences differ: is turnout too high or too low?, *Social Choice and Welfare* **33**(1), 25–50.

Gruber, J. and Wise, D. A.: 2009, *Social security programs and retirement around the world: Fiscal implications of reform*, University of Chicago Press.

Hagemejer, J., Makarski, K. and Tyrowicz, J.: 2015, Unprivatizing the pension system: the case of Poland, *Applied Economics* **47**(8), 833–852.

Hénin, P.-Y. and Weitzenblum, T.: 2005, Welfare effects of alternative pension reforms: Assessing the transition costs for French socio-occupational groups, *Journal of Pension Economics and Finance* **4**(03), 249–271.

Holzmann, R. and Stiglitz, J. E.: 2001, *New ideas about old age security*, Washington, DC: World Bank.

Huang, H., Imrohoglu, S. and Sargent, T. J.: 1997, Two computations to fund social security, *Macroeconomic Dynamics* **1**(01), 7–44.

Jarrett, P.: 2011, Pension reforms in Poland and elsewhere: The view from Paris, *CASE Network Studies and Analyses 425*, CASE-Center for Social and Economic Research.

Kelley, D. G.: 2014, The political economy of unfunded public pension liabilities, *Public Choice* **158**(1-2), 21–38.

Kindermann, F. and Krueger, D.: 2014, High Marginal Tax Rates on the Top 1%? Lessons from a Life Cycle Model with Idiosyncratic Income Risk, *NBER Working Papers 20601*, National Bureau of Economic Research, Inc.

Krieger, T. and Ruhose, J.: 2013, Honey, I shrunk the kids' benefits – revisiting intergenerational conflict in OECD countries, *Public Choice* **157**(1-2), 115–143.

Krusell, P. and Rios-Rull, J.-V.: 1999, On the size of us government: political economy in the neoclassical growth model, *American Economic Review* **89**(5), 1156–1181.

Makarski, K., Tyrowicz, J. and Hagemejer, J.: 2017, Efficiency of the pension reform: the welfare effects of various fiscal closures, *Macroeconomic Dynamics* **21**, 1205–1234.

McGrattan, E. R. and Prescott, E. C.: 2013, On financing retirement with an aging population, *NBER Working Paper 18760*, National Bureau of Economic Research.

Mulligan, C. B. and Sala-i Martin, X.: 2015, Political and economic forces sustaining social security, *The B.E. Journal of Economic Analysis & Policy* **4**(1), 1–56.

OECD: 2018, *OECD Pensions Outlook 2018*, OECD.

Orszag, P. R. and Stiglitz, J. E.: 2001, Rethinking pension reform: Ten myths about social security systems, *in* R. Holzman and J. E. Stiglitz (eds), *New ideas about old age security*, Washington, DC: World Bank, pp. 17–56.

Parlevliet, J.: 2017, What drives public acceptance of reforms? Longitudinal evidence from a Dutch pension reform, *Public Choice* **173**(1-2), 1–23.

Phelan, C.: 2006, Public trust and government betrayal, *Journal of Economic Theory* **130**(1), 27–43.

Samuelson, P. A.: 1958, An exact consumption-loan model of interest with or without the social contrivance of money, *Journal of Political Economy* **66**(6), 467–482.

Schmitt-Grohe, S. and Uribe, M.: 2003, Closing small open economy models, *Journal of International Economics* **61**(1), 163–185.

Schwarz, A. M.: 2011, New realities of pension policy in central europe, *Pension systems in emerging Europe: Reform in the age of austrerity.* The conference organized by the European Bank for Reconstruction and Development on April 1st, 2011, in Paris.

Schwarz, A. M.: 2014, *The Inverting Pyramid: Pension Systems Facing Demographic Challenges in Europe and Central Asia*, World Bank Publications, pp. 34–40.

Sjoblom, K.: 1985, Voting for social security, *Public Choice* **45**(3), 225–240.

Song, Z., Storesletten, K., Wang, Y. and Zilibotti, F.: 2015, Sharing high growth across generations: pensions and demographic transition in china, *American Economic Journal: Macroeconomics* **7**(2), 1–39.

Tabellini, G.: 1991, The politics of intergenerational redistribution, *Journal of Political Economy* **99**(2), 335–357.

Tabellini, G.: 2000, A positive theory of social security, *Scandinavian Journal of Economics* **102**(3), 523–545.

Online Appendices (not intended for publication)

A Proofs

It is convenient to first present a lemma specifying how much taxes are lowered in the case of capturing assets and shifting contributions.

Lemma 1. *If $\theta_{t-1} = 2$ and $\theta_t = 1$, i.e., asset capture and shifting of contributions occurs (denoted by superscript AC), lump-sum taxes Υ_t in period t are lowered by*

$$-\Delta\Upsilon_t^{AC} = \tau^F w_t \left(\frac{3 + r + 2g}{3(1+g)} - \frac{r - g}{3(1+r)} \right) > 0. \tag{A.1}$$

Proof. According to (7) and (8), for $\theta_{t-1} = 2$ and $\theta_t = 1$, $f_{1,t}^F = 0$ and $f_{2,t}^F = (1+r)\tau^F w_{t-1}$. From equation (4), the government balance adjusts by:

$$-\sum_j \Delta\Upsilon_t^{AC} = \sum_{j=1}^{J-1}(f_{j,t} + \tau^F w_t) - \frac{(1+g)(1+r)\tau^F w_{t-1} - (1+g)(1+g)\tau^F w_{t-1}}{(1+r)}.$$

Using that the total population alive has size 3 and that $w_t = w_{t-1}(1+g)$,

$$-3\Delta\Upsilon_t^{AC} = \left((1+r)\tau^F w_{t-1} + (1+1)\tau^F w_t \right) - \frac{(r-g)(1+g)\tau^F w_{t-1}}{1+r}.$$

Simplifying

$$-3\Delta\Upsilon_t^{AC} = \tau^F w_t \left(\frac{3 + r + 2g}{1+g} - \frac{r - g}{1+r} \right) > 0.$$

$\square$

Lemma 2. *If $\theta_{t-1} = 2$ and $\theta_t = 0$, i.e., social security is abandoned completely (denoted by superscript A), lump-sum taxes Υ_t in period t are lowered by*

$$-\Delta\Upsilon_t^A = \frac{3 + r + 2g}{3(1+g)} \tau^F w_t > 0. \tag{A.2}$$

Proof. If $\theta_{t-1} = 2$ and $\theta_t = 0$, then $f_{1,t}^F = 0$ and $f_{2,t}^F = (1+r)\tau^F w_{t-1}$. From equation (4), the government balance adjusts by:

$$-\sum_j \Delta\Upsilon_t^A = \sum_j^{J-1}(f_{j,t} + \tau^F w_t).$$

Using that the total population size alive is 3 and that $w_t = w_{t-1}(1+g)$, and simplifying:

$$-3\Delta\Upsilon_t^A = \frac{3 + r + 2g}{1+g} \tau^F w_t > 0$$

$\square$

Proposition 1:

Proof. Abandoning social security implies zero social security contributions and zero benefit payments. Hence, the PAYG pillar disappears.

- The intertemporal budget constraint of a young individual born in period t is:

$$c_{1,t} + \frac{c_{2,t+1}}{(1+r)} + \frac{c_{3,t+2}}{(1+r)^2} = (1-\tau)w_t - \Upsilon_t + \frac{(1-\tau)w_{t+1} - \Upsilon_{t+1}}{(1+r)} + \frac{b_{t+2}^P - \Upsilon_{t+2}}{(1+r)^2}.$$

  The resource effect for a young person in period t of abolishing the PAYG pillar of social security is given by

$$\tau w_t + \frac{\tau(1+g)w_t}{(1+r)} - \frac{b_{t+2}^P}{(1+r)^2} = \frac{(2+r+g)(1+r) - (2+g+g)(1+g)}{(1+r)^2}\tau w_t.$$

  This expression is positive if $r > g$. Hence, young agents would benefit from abandoning PAYG social security, because diverting their contributions into private savings yields a higher rate of return. Thus, young agents vote to abandon it.

- The budget constraint of a middle-aged individual can be expressed as

$$c_{2,t} + \frac{c_{3,t+1}}{(1+r)} = (1+r)a_{1,t-1} + (1-\tau)w - \Upsilon_t + \frac{b_{t+1}^P - \Upsilon_{t+1}}{(1+r)}.$$

  The resource effect of abolishing social security would be given by:

$$\tau w_t - \frac{b_{t+1}^P}{(1+r)} = \frac{r - 1 - 2g}{1+r}\tau w_t,$$

  which is negative if $r < 1 + 2g$. Therefore, for plausible values of r and g, middle-aged agents lose from abandoning PAYG social security, so they vote to keep it.

- The budget constraint of an old individual can be expressed as $c_{3,t} = (1+r)a_{2,t-1} + b_t^P - \Upsilon_t$. Abandoning social security implies a loss of resources of $b_t^P > 0$. Thus old agents would lose from abandoning PAYG social security. Thus old agents vote to keep it.

- Old and middle-aged agents outnumber the young, so the PAYG system is stable.

$\square$

Proposition 2:

Proof. We show that keeping the system intact cannot be an equilibrium outcome.

- The budget constraint of a young individual is:

$$c_{1,t} + \frac{c_{2,t+1}}{(1+r)} + \frac{c_{3,t+2}}{(1+r)^2} = (1-\tau)w_t - \Upsilon_t + \frac{(1-\tau)w_{t+1} - \Upsilon_{t+1}}{(1+r)} + \frac{b_{t+2} - \Upsilon_{t+2}}{(1+r)^2},$$

and asset capture lowers their retirement benefit by

$$\Delta b_{t+2} = (r-g)(3 + 2g + r)\tau^F w_t.$$

Compare the loss in terms of social security benefits with the taxes reduced $(\Delta \Upsilon_t^{AC})$ as per equation (A.1). The young individual supports asset capture if:

$$\Delta \Upsilon_t^{AC} > -\frac{\Delta b_{t+2}}{(1+r)^2} \iff \frac{3+2g+r}{3(1+g)}\tau^F w_t - \frac{(r-g)}{3(1+r)}\tau^F w_t > \frac{(r-g)}{(1+r)^2}(3+2g+r)\tau^F w_t$$

$$\iff (3+r+2g)(1+r)^2 > 3(r-g)(1+g)(3+2g+r) + (r-g)(1+r)(1+g),$$

which holds $\forall \quad r, g > 0.$[33]

- The budget constraint of a middle-aged individual is expressed as

$$c_{2,t} + \frac{c_{3,t+1}}{(1+r)} = (1+r)a_{1,t-1} + (1-\tau_t)w_t - \Upsilon_t + \frac{b_{t+1} - \Upsilon_{t+1}}{(1+r)},$$

and assets capture lowers their retirement benefit by

$$\Delta b_{t+1} = \frac{(r-g)(2+r+g)}{1+g}\tau^F w_t.$$

Analogously to the young individual, we use equation (A.1) to compare the gain from lowered taxes due to the asset capture with the decline in retirement benefits:

$$\Delta \Upsilon_t^{AC} \overset{?}{>} -\frac{\Delta b_{t+1}}{1+r} \iff (1+r)(3+r+2g) \overset{?}{>} (r-g)(7+3r+4g).$$

The above inequality depends on the magnitude of $r - g$.[34]

- Finally, the budget constraint of an old individual is $c_{3,t} = (1+r)a_{2,t-1} + b_t - \Upsilon_t$. As asset capture does not affect their retirement benefit ($\Delta b_t = 0$), using Lemma 1 it is clear that the old gain from asset capture.

Summarizing, old and young individuals vote to capture assets, so even if middle-aged individuals vote against asset capture, they are outvoted. Hence, a system with a funded pillar is not politically stable.

$\square$

---

[33]The computations yield a third-order polynomial in both r and g, which takes positive values for $r > 0$ and $g > 0$.

[34]If $r - g$ is sufficiently small, the middle-aged individual votes to capture assets, otherwise they are in favor of maintaining the funded arrangement as it is. Note that a small difference between r and g is equivalent to a small gain of having a funded social security over having a PAYG system.

Proposition 3:

Proof. We will show that voting results in the following rule for social security: $\theta_t = \theta(\theta_{t-1}) = 1$ for $\theta_t \in \{1, 2\}$. Recall that we assumed that, once a social security pillar has been abolished, it cannot be re-introduced, i.e. $\theta_t = \theta(\theta_{t-1}) \in \{0, 1, 2\}$ for $\theta_{t-1} = 2$, $\theta_t = \theta(\theta_{t-1}) \in \{0, 1\}$ for $\theta_{t-1} = 1$ and $\theta_t = \theta(\theta_{t-1}) \in \{0\}$ for $\theta_{t-1} = 0$. We demonstrate that the proposed equilibrium is an equilibrium indeed by resorting to the "one-shot deviation principle", i.e. for $\theta_{t-1} \in \{1, 2\}$ an agent finds a one-shot deviation from $\theta_t = \theta(\theta_{t-1}) = 1$ not optimal.

Case of $\theta_{t-1} = 2$.

Using the one-shot deviation principle we demonstrate the voting outcomes summarized in Table A.1. The table reports the majority voting outcomes for pairs of policies pitched against each other. The tables shows that regardless of the sequence in which pairs of arrangements are put to majority vote, the eventual outcome is always that voters choose to capture assets and shift contributions.

Table A.1: Majority voting outcomes for pairs of policies pitched against each other, $\theta_{t-1} = 2$

policies\age and outcome	$j = 1$	$j = 2$	$j = 3$	Voting outcome
no capture $\theta_t = 2$ vs capture $\theta_t = 1$	$\theta_t = 1$	$\theta_t = 2$	$\theta_t = 1$	$\theta_t = 1$
abandon $\theta_t = 0$ vs capture $\theta_t = 1$	?	$\theta_t = 1$	$\theta_t = 1$	$\theta_t = 1$
abandon $\theta_t = 0$ vs no capture $\theta_t = 2$	?	$\theta_t = 2$	$\theta_t = 2$	$\theta_t = 2$

- No asset capture, i.e. $\theta_t = 2$, versus asset capture $\theta_t = 1$.

  Given the proposed equilibrium policy $\theta_t = \theta(\theta_{t-1}) = 1$ for $\theta_{t-1} \in \{1, 2\}$, the one-shot deviation from this is to not capture assets in period t, but capture them instead in period $t + 1$. We demonstrate that the deviation reduces welfare for a majority of the cohorts currently alive.

  - The budget constraint of a young individual is expressed as:

  $$c_{1,t} + \frac{c_{2,t+1}}{(1+r)} + \frac{c_{3,t+2}}{(1+r)^2} = (1-\tau)w_t - \Upsilon_t + \frac{(1-\tau)w_{t+1} - \Upsilon_{t+1}}{(1+r)} + \frac{b_{t+2} - \Upsilon_{t+2}}{(1+r)^2}.$$

  Capturing assets in the next period instead of the current period raises the future retirement benefits of the young by

  $$\Delta b_{t+2} = (1+g)(r-g)\tau^F w.$$

  Since taxes are lowered in period $t + 1$ rather than in period t, the present value of the reduction in taxes shrinks by $-(-\Delta\Upsilon_t^{AC}) + \frac{-\Delta\Upsilon_{t+1}^{AC}}{1+r}$. Combining with equation (A.1) and simplifying:

  $$-(-\Delta\Upsilon_t^{AC}) + \frac{-\Delta\Upsilon_{t+1}^{AC}}{1+r} + \frac{\Delta b_{t+2}}{(1+r)^2} < 0 \iff 1 + g < 1 + r,$$

  which holds for all $r > g > 0$. Higher retirement benefits due to postponing asset capture do not compensate for the postponement of the tax cut. Thus, the young individuals prefer asset capture.

– Old individuals prefer asset capture in period t, so as to benefit from the tax reduction, instead of asset capture in the next period $t + 1$ after they have died.

– The budget constraint of a middle-aged individual is

$$c_{2,t} + \frac{c_{3,t+1}}{(1+r)} = (1+r)a_{1,t-1} + (1-\tau_t)w_t - \Upsilon_t + \frac{b_{t+1} - \Upsilon_{t+1}}{(1+r)}.$$

Postponing asset capture raises their retirement benefit by $\Delta b_{t+1} = (r-g)(2+r+g)\tau^F w$. We need to compare this with the loss of postponing the reduction in taxes:

$$-(-\Delta\Upsilon_t^{AC}) + \frac{-\Delta\Upsilon_{t+1}^{AC}}{1+r} + \frac{\Delta b_{t+1}}{(1+r)} > 0 \iff 2 + 2g + 2rg > 2g^2.$$

Hence, for $1 + 2g > r > g$ middle-aged individuals vote not to capture assets.

– With young and old individuals voting in favor of asset capture in period t instead of period $t + 1$, the outcome of the vote is asset capture in period t.

• Complete abandonment of social security, i.e. $\theta_t = 0$, versus asset capture only, i.e. $\theta_t = 1$.

Again, we demonstrate that the deviation reduces welfare for a majority of the cohorts currently alive.

– The budget constraint of a young individual is

$$c_{1,t} + \frac{c_{2,t+1}}{(1+r)} + \frac{c_{3,t+2}}{(1+r)^2} = (1-\tau_t)w_t - \Upsilon_t + \frac{(1-\tau_{t+1})w_{t+1} - \Upsilon_{t+1}}{(1+r)} + \frac{b_{t+2} - \Upsilon_{t+2}}{(1+r)^2}.$$

Equations (A.1) and (A.2) imply that complete dismantling of social security affects the present discounted value of the young's resources as:

$$-\Delta\Upsilon_t^A - (-\Delta\Upsilon_t^{AC}) + \tau w_t + \frac{\tau w_{t+1}}{(1+r)} - \frac{b_{t+2}}{(1+r)^2}$$
$$= \frac{(r-g)}{3(1+r)}\tau^F w + \left(\frac{(2+r+g)(1+r)}{(1+r)^2} - \frac{2(1+g)^2}{(1+r)^2}\right)\tau w.$$

Since $r > g$, the young individual prefers to capture assets rather than abandon social security entirely.

– The budget constraint of a middle-aged agent can be expressed as

$$c_{2,t} + \frac{c_{3,t+1}}{(1+r)} = (1+r)a_{1,t-1} + (1-\tau)w_t - \Upsilon_t + \frac{b_{t+1} - \Upsilon_{t+1}}{(1+r)}.$$

The difference is, using equations (A.1) and (A.2):

$$-\Delta\Upsilon_t^A - (-\Delta\Upsilon_t^{AC}) + \tau w_t - \frac{b_{t+1}}{(1+r)} = \frac{(r-1-2g)}{(1+r)}\tau^P w + \frac{r-3-4g}{3(1+r)}\tau^F w.$$

Hence, for $r < 1 + 2g$ middle-aged individuals prefer to capture assets rather than to abandon social security completely.

- The budget constraint of an old individual is $c_{3,t} = (1 + r)a_{2,t-1} + b_t - \Upsilon_t$. In comparison with the middle-aged, the old do not benefit from abolishing the contribution payment in return for giving up their PAYG retirement benefit and, hence, they also prefer to capture assets and shift contributions rather than to abandon social security completely.

- We conclude that the outcome of the vote between capturing assets and diverting contributions versus abandoning social security completely is the former.

• Keeping both pillars intact, i.e. $\theta_t = 2$, versus complete abandonment of social security $\theta_t = 0$.

We now compare two one-shot deviations. The first one is to not capture assets in period t and the second one is to completely abandon social security in period t. Recall that once social security is completely abandoned, it cannot be restored.

- The budget constraint of a young individual is given by:

$$c_{1,t} + \frac{c_{2,t+1}}{(1+r)} + \frac{c_{3,t+2}}{(1+r)^2} = (1-\tau)w_t - \Upsilon_t + \frac{(1-\tau)w_{t+1} - \Upsilon_{t+1}}{(1+r)} + \frac{b_{t+2} - \Upsilon_{t+2}}{(1+r)^2}.$$

Completely abandoning social security in period t means that they lose future retirement benefits and future lower taxation, but gain by not paying contributions in period t and $t+1$ as well as from lower taxes in period t. Using equations (A.1) and (A.2), their budget constraint is affected as follows:

$$\tau w_t + \frac{\tau w_{t+1}}{(1+r)} - \frac{b_{t+2}}{(1+r)^2} + (-\Delta\Upsilon_t^A) - \frac{(-\Delta\Upsilon_{t+1}^{AC})}{(1+r)}$$
$$= \tau^F w_t \frac{r-g}{3(1+r)^2}\left(\frac{g(2+g) - r(2(2+g)+r) - 1}{(1+g)(1+r)}\right) + \tau^P w_t\left(\frac{r(3+r) - g(3+2g-r)}{(1+r)^2}\right),$$

which is positive for $r > g$, hence young agents prefer abandoning social security completely to capturing assets.

- The budget constraint of a middle-aged individual is

$$c_{2,t} + \frac{c_{3,t+1}}{(1+r)} = (1+r)a_{1,t-1} + (1-\tau_t)w_t - \Upsilon_t + \frac{b_{t+1} - \Upsilon_{t+1}}{(1+r)}.$$

Completely abandoning the social security in period t means that they lose future retirement benefits and future lower taxation, but gain by not paying contributions in period t as well as experiencing lower taxes in period t. Using equations (A.1) and (A.2), the budget constraint is affected as follows:

$$\tau w_t - \frac{b_{t+1}}{(1+r)} + (-\Delta\Upsilon_t^A) - \frac{(-\Delta\Upsilon_{t+1}^{AC})}{(1+r)}$$
$$= \tau^F w_t\left(1 - \frac{2+r+g}{1+g} + \frac{r-g}{3(1+r)}\left(\frac{2+g^2 - r(4+r+2g)}{(1+g)(1+r)}\right)\right) + \tau^P w_t\left(\frac{r-1-2g}{1+r}\right).$$

Since $0 < g < r < 1 + 2g$, this expression implies that middle-aged individuals prefer to capture assets rather than abandon social security altogether.

- The old clearly prefer keeping both pillars intact to abandoning social security completely.

– We conclude that the voting outcome is to keep both pillars versus abandoning social security completely.

The case of $\theta_{t-1} = 1$.

It follows from Proposition 1 that voting results into $\theta_t = \theta(\theta_{t-1}) = 1$ for $\theta_{t-1} = 1$.

$\square$

B Calibration

This appendix provides further detailed information on the inputs used in the calibration.

Figure B.1: Number of 20-year-old arriving in each period (upper left), one year survival probability over time for a selected cohort (upper right), age structure of population in the initial steady state (lower left) and labor-augmenting technological progress (lower right).
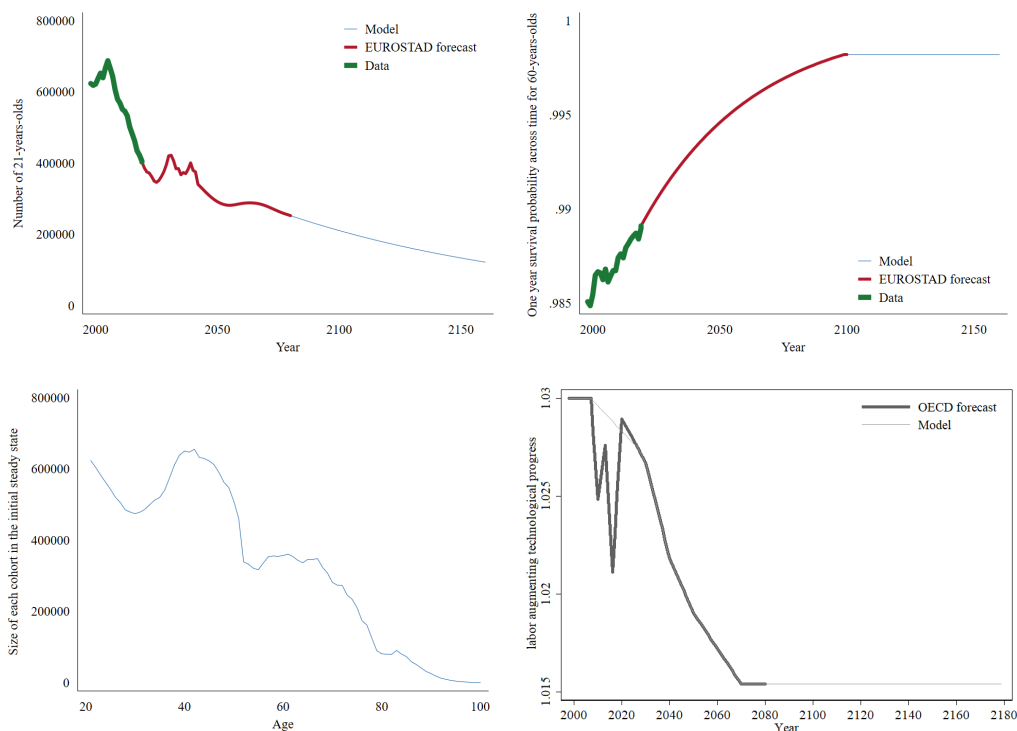
Source: Demographic forecasts until 2080 are based on EUROSTAT, the rate of technological progress is based on the forecasts of the Ageing Working Group of the Economic Policy Committee and the European Commission, and the total factor productivity data represented by the thick solid line are from the OECD. The latter are smoothed by a moving-average for the cyclical component.

Figure B.2: Distribution of ω multipliers for individual productivity
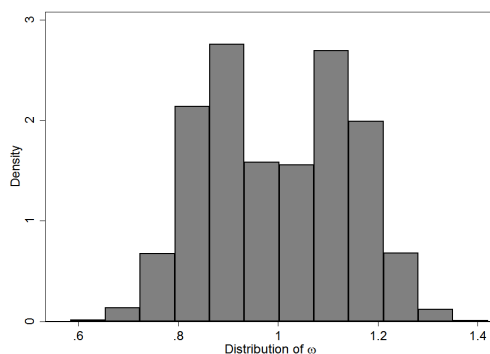
Note: Estimates of the distribution are obtained using data from the Structure of Earnings Survey, wave of 1998.

Figure B.3: Interest rate and growth rate of total wage bill when continuing with PAYG-DB (left panel) and in transition to two-pillar DC
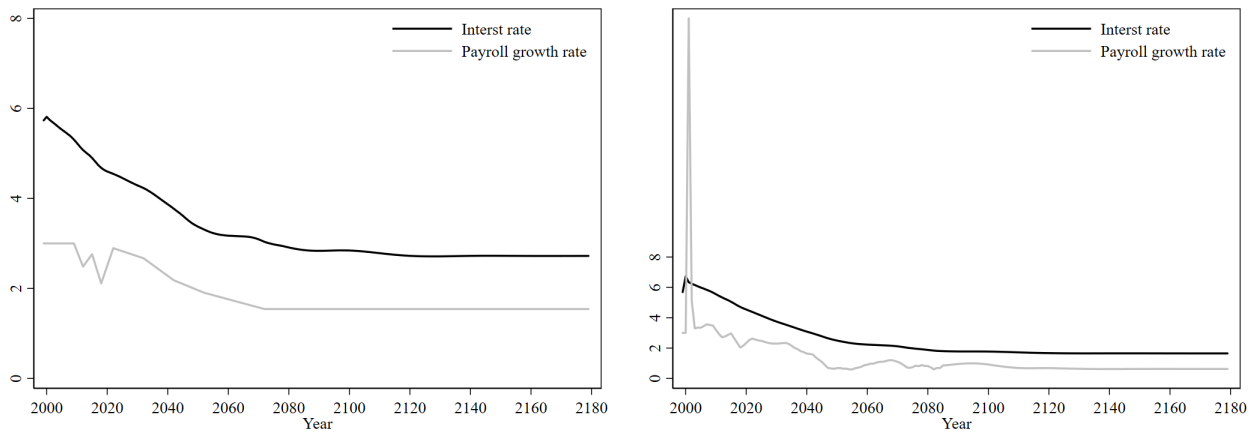
Note: The figure depicts the model-based paths for the interest rate and the growth rate of the total wage bill for the case with intra-cohort homogeneity; the left-hand panel for continuing with PAYG-DB and the right-hand panel for a transition to two-pillar DC social security. The analogous figures with intra-cohort heterogeneity are virtually identical and are available upon request.

C Derivation of the welfare effect of reform

Cohorts alive at the time of the reform The remaining lifetime utility of an individual of type κ who is j years old in period t can be written as:

$$V_{j,\kappa,t} = \sum_{s=0}^{J-j} \delta_\kappa^s \pi_{j+s,t+s} u\left(c_{j+s,\kappa,t+s}, l_{j+s,\kappa,t+s}\right). \tag{C.3}$$

Define $V_{j,\kappa,t}^S$ and $V_{j,\kappa,t}^R$ as the remaining lifetime utilities of such an individual under respectively the status-quo and the reform scenarios:

$$
\begin{aligned}
V_{j,\kappa,t}^S &= \sum_{s=0}^{J-j} \delta_\kappa^s \pi_{j+s,t+s} u\left(c_{j+s,\kappa,t+s}^S, l_{j+s,\kappa,t+s}^S\right), \\
V_{j,\kappa,t}^R &= \sum_{s=0}^{J-j} \delta_\kappa^s \pi_{j+s,t+s} u\left(c_{j+s,\kappa,t+s}^R, l_{j+s,\kappa,t+s}^R\right),
\end{aligned}
$$

where $c_{j+s,\kappa,t+s}^S$ and $l_{j+s,\kappa,t+s}^S$ are consumption and labor in the status-quo scenario and similarly $c_{j+s,\kappa,t+s}^R$ and $l_{j+s,\kappa,t+s}^R$ refer to consumption and labor in the reform scenario. The welfare effect $W_{j,\kappa,t}$ that brings the remaining lifetime utility of a κ-type individual of age j under the reform scenario to the remaining lifetime utility under the status-quo scenario is given by:

$$\sum_{s=0}^{J-j} \delta_\kappa^s \pi_{j+s,t+s} \left(u(1-W_{j,\kappa,t})c_{j+s,\kappa,t+s}^R, l_{j+s,\kappa,t+s}^R\right) = \sum_{s=0}^{J-j} \delta_\kappa^s \pi_{j+s,t+s} u\left(c_{j+s,\kappa,t+s}^S, l_{j+s,\kappa,t+s}^S\right). \tag{C.4}$$

For a logarithmic utility function we can rewrite the left-hand side (LHS) of equation (C.4) as follows:

$$
\begin{aligned}
LHS &= \sum_{s=0}^{J-j} \delta_\kappa^s \pi_{j+s,t+s} \left[\ln\left((1-W_{j,\kappa,t})c_{j+s,\kappa,t+s}^R\right) + \phi \ln\left(l_{j+s,\kappa,t+s}^R\right)\right] \\
&= \sum_{s=0}^{J-j} \delta_\kappa^s \pi_{j+s,t+s} \left[\ln(1-W_{j,\kappa,t}) + \ln\left(c_{j+s,\kappa,t+s}^R\right) + \phi \ln\left(l_{j+s,\kappa,t+s}^R\right)\right] \\
&= V_{j,\kappa,t}^R + \ln(1-W_{j,\kappa,t}) \sum_{s=0}^{J-j} \delta_\kappa^s \pi_{j+s,t+s}
\end{aligned}
$$

Substituting into equation (C.4) we obtain:

$$\ln(1-W_{j,\kappa,t}) \sum_{s=0}^{J-j} \delta_\kappa^s \pi_{j+s,t+s} = V_{j,\kappa,t}^S - V_{j,\kappa,t}^R.$$

Therefore, the welfare effect can be expressed as:

$$W_{j,\kappa,t} = 1 - \exp\left\{\frac{V_{j,\kappa,t}^S - V_{j,\kappa,t}^R}{\sum_{s=0}^{J-j} \delta_\kappa^s \pi_{j+s,t+s}}\right\}. \tag{C.5}$$

Cohorts born after the reform For an individual of κ type who arrives into the economy in period $t' \geq t$ the welfare gain from the reform can analogously and straightforwardly be derived as:

$$W_{1,\kappa,t'} = 1 - \exp \left\{ \frac{V^S_{1,\kappa,t'} - V^R_{1,\kappa,t'}}{\sum_{s=0}^{J-1} \delta_\kappa^s \pi_{1+s,t'+s}} \right\}. \tag{C.6}$$

50

D Introduction of two-pillar DC social security in 1999

The emergence of a two-pillar DC social security arrangement in 1999 can be rationalized as the outcome of a vote in which individuals weigh a number of factors. First, progressing longevity would imply higher PAYG-DB benefits relative to the alternative of a DC arrangement, *ceteris paribus*. Second, with a PAYG-DB system increasing longevity would require a permanent subsidy from the government, because a longer life at retirement would raise the amount of retirement benefits to be paid out. The resulting deficit would need to be financed with higher consumption taxes and debt according to the fiscal rule given by equation (22). Replacement by a DC scheme would eventually render social security balanced, producing lower taxes in the long run when compared with continuation of PAYG-DB. Third, with the return on assets exceeding the growth rate of the wage bill, the benefits paid out by the funded social security pillar would exceed those paid out of the component of the PAYG-DB pillar that is replaced by the funded pillar. Fourth, during the transition period to the two-pillar DC scheme, a deficit would emerge in social security, necessitating a transitory increase in taxes, which would be smoothed by a transitory increase in the public debt. Finally, there would be general equilibrium effects, stemming from an adjustment in private voluntary savings and in the labor supply. The tighter link between contributions and benefits in the DC system would strengthen labor supply, thereby increasing social security benefits.

We discuss first the case in which intra-cohort heterogeneity is absent. Replacing a PAYG-DB social security system with a DC system with partial funding has different effects across cohorts. Overall there is a majority among those alive in 1999 in favor of this reform – see Figure D.4. Based on welfare evaluations, which account for direct and general equilibrium effects, we can demonstrate which cohorts are in support of turning the PAYG-DB system into a DC system with partial funding. The retired are sheltered from the reform in any case and they benefit from the fact that introducing a DC system strengthens the incentives for labor supply (see Figure D.5). The increase in the wage bill raises the indexation rate of the social security benefits. The same is the case for the elderly workers who thus also benefit from the reform. The calculation for younger cohorts is more complicated. These cohorts enjoy the benefit from the higher return on the funded pillar relative to the part of the PAYG-DB it replaces. However, they are negatively affected by the fact that consumption taxes have to be raised for a while, since directing some of the contributions to the funded pillar necessitates fiscal adjustment in order to pay for the elders' retirement benefits. Moreover, the shift of part of the PAYG-DB to PAYG-DC has a negative effect on the benefits obtained from the PAYG part of social security, because the switch results into a balanced social security, while at the moment of the reform the original PAYG-DB system already features a deficit that will be growing further with rising life expectancy. This deficit is reduced by a corresponding reduction of retirement benefits from the new PAYG-DC pillar, which is only partially compensated by the increase in benefits coming from the return on the funded pillar dominating the growth rate of the wage bill. The reform eventually improves welfare for newborn cohorts at the moment of their birth. They benefit from increased overall efficiency and lower taxation.

Partial funding eventually raises the welfare gains from having the two-pillar DC system instead of the pure PAYG-DB system. Figure D.5 depicts for each year the percentage of individuals who

benefit from the reform in 1999 when compared with maintaining the single-pillar DB scheme. The losing cohorts pass away over time, whereas the newborn benefit from increased overall efficiency and lower taxation. Eventually, all cohorts alive at a given moment would benefit from the reform. The aggregate long-term welfare effect of the reform amounts to 1%, so that if all losers were compensated by the winners, the latter would on net still gain (i.e., a Kaldor-Hicks welfare improvement).

Figure D.4: Distribution of the welfare effects by cohort from introducing two-pillar DC system in 1999
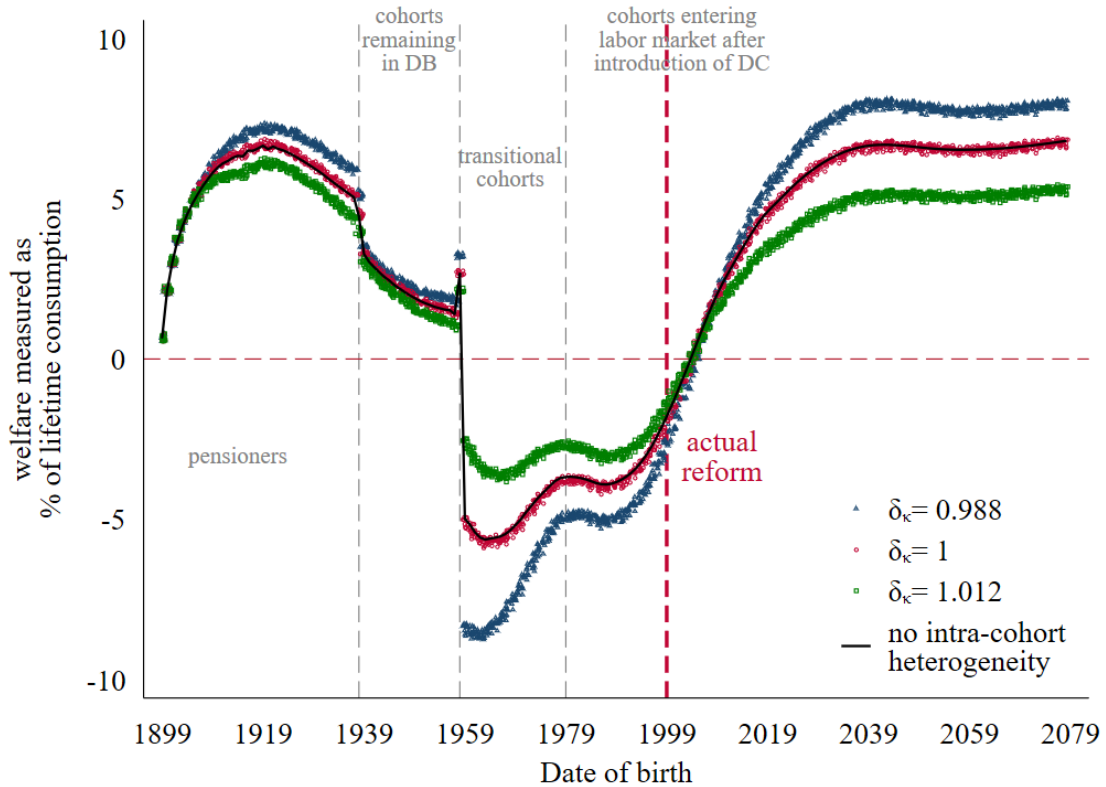
Note: Lines depict welfare effect in terms of percent increase in the remainder of lifetime consumption relative to continuation of the PAYG-DB system for each cohort, which is indicated by the birth year, at the moment of the reform. Of the future-born the figure thus depicts the effect over the full lifetime. The solid line obtains in the absence of intra-cohort heterogeneity, while the other lines correspond to the presence of intra-cohort heterogeneity. Here, the dispersion within each non-solid line reflects the heterogeneity of welfare effects across productivity types.

We turn now to the case of intra-cohort heterogeneity. The driving force behind differences in the evaluation of the reform across the subcohorts are differences in the discount factor δ_κ. In the long run all subcohorts gain due to lower taxation. Among the transition cohorts, i.e., those younger than 40 years in 1999, nearly all subcohorts experience a decline in welfare due to lower overall social security benefits and higher taxation. Exempted from the decline in welfare are the most patient transition subcohorts among the transition cohorts.

Differences in the individual productivity endowment have virtually no effect on the reform evaluation – see the plots of welfare in Figure D.4. This also implies that there is no room for coalition building across cohorts based on individual productivity endowments, such as between the elderly and high-productivity individuals in the transition cohorts; tax and general equilibrium effects dominate

Figure D.5: Political support in each year for introduction of two-pillar social security in 1999 (left) and gains specifically attributable to funded pillar (right)
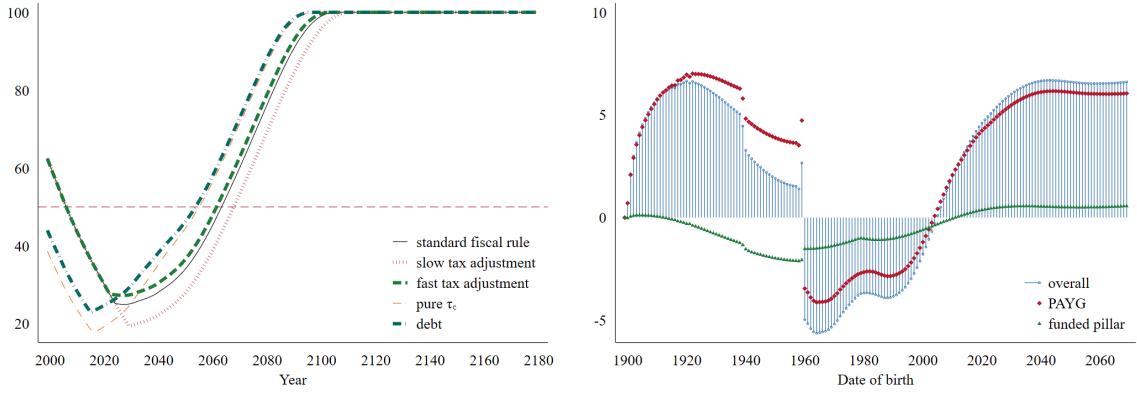
Notes: Results from simulations with intra-cohort homogeneity. Those for intra-cohort heterogeneity are virtually the same. The left-hand panel depicts as a function of the tax rule the percent of subcohorts living in each period who gain from introducing the two-pillar DC social security, as described above. Hence, for each year on the horizontal axis it gives the percent of political support for the reform that took place in 1999 versus continuing with single-pillar PAYG-DB. The solid black line, denoted as 'standard fiscal rule' refers to the fiscal rule as described by equation (22) and parametrized in Table 1. The first alternative specification considers instantaneous consumption tax adjustment (denoted as 'pure τ_c'). Combining equation (20) and (21), with $D_{t+1} = D_t$ we have:

$$\tau_{c,t} = \frac{G_t + Subs_t + r_t D_t - \tau_{l,t}\big[(1-\tau)w_t L_t + \sum_{\kappa=1}^{\mathcal{K}} \sum_{j=\bar{J}}^{J} b_{j,\kappa,t} N_{j,\kappa,t}\big] - \sum_{\kappa=1}^{\mathcal{K}} \sum_{j=1}^{J} N_{j,\kappa,t}(\tau_{k,t} r_t a_{j,\kappa,t} + \Upsilon_t)}{\sum_{\kappa=1}^{\mathcal{K}} \sum_{j=1}^{J} N_{j,\kappa,t} c_{j,\kappa,t}}.$$

The second alternative considers full accommodation through public debt keeping the consumption tax rate initially unchanged. Subsequently, after the transition cohorts have passed away debt is gradually reduced using the fiscal rule described in equation (22) (denoted as 'debt'). We trigger adjustment in consumption taxes as soon as $\mathcal{D}_t$ exceeds 90%. Third, we study slower tax adjustment than described in Table 1 by setting parameter ϱ_D to 0.03 in the fiscal rule (22) (denoted as 'slow tax adjustment'), and faster tax adjustment by setting parameter ϱ_D to 0.07 in (22) (denoted as 'fast tax adjustment'). The right-hand panel depicts the decomposition of the overall effect of a switch from PAYG-DB to a two-pillar DC scheme into an effect attributable to a switch from PAYG-DB to single-pillar PAYG-DC and an effect attributable to the introduction of a funded pillar.

the role of productivity differences.[35]

This latter finding points to the relevance of the fiscal adjustments accompanying social security reforms: if lump-sum taxes were used to finance the transition costs of the reform, the utility of the elderly cohorts would be harmed, leading them to oppose the reform. However, the fiscal rule portrayed in equation (22) partly raises consumption taxes, but mostly smooths the costs of the reform to the future cohorts via the public debt. Overall, the cohort distribution of these welfare effects is such that in 1999 approximately 54.7% of the population living at the moment of the reform would support it; see the left panel of Figure D.5.

To disentangle the effect of a switch from PAYG-DB to a two-pillar DC scheme into an effect attributable to a switch from DB to DC and an effect attributable to the introduction of a funded pillar, we run an additional simulation with $\tau^P = \tau$ and $\tau^F = 0$. Subtraction of the resulting welfare effect from the overall welfare effect of the switch to the two-pillar DC scheme yields the welfare

---

[35]Earlier literature on introducing (redistributive) social security has studied coalition formation between elderly and low-productivity young agents, who have a common interest in a highly redistributive arrangement. For example, see Tabellini (1991).

effect effect of the introduction of the funded pillar. The right-hand panel of Figure D.5 depicts the decomposition for intra-cohort homogeneity. For intra-cohort heterogeneity, the results are virtually identical. The welfare effect of introducing a funded pillar starting from a single-pillar PAYG-DC scheme is negative for all cohorts born until approximately 2010, after which it becomes positive for all subsequently-born cohorts, indicating that eventually the gains from a faster-growing and larger economy exceed the costs of financing the transitory deficit in public social security.

Figure D.6 depicts the development of the interest rate and the consumption tax rate for the three cases under intra-cohort homogeneity: maintaining the initial single-pillar PAYG-DB, the single-pillar PAYG-DC and the two-pillar DC. The corresponding figure under intra-cohort heterogeneity is essentially identical. Single-pillar PAYG-DB features the lowest capital stock, hence is responsible for the highest interest rate. Vice-versa, the highest capital stock under the two-pillar DC regime is responsible for the lowest interest rate. A single-pillar DC forces individuals to save more, in order to make up for the loss of retirement benefits compared to single-pillar DB. However, the additional savings do not make up for the absence of a funded pillar, hence the interest rate in this case is in-between that of the other two cases. The growing deficit under the single pillar PAYG-DB regime produces a strong increase in the consumption tax rate relative to that in the other two cases.

Figure D.6: Interest rates (left) and consumption tax rates (right) under the DB and DC social security
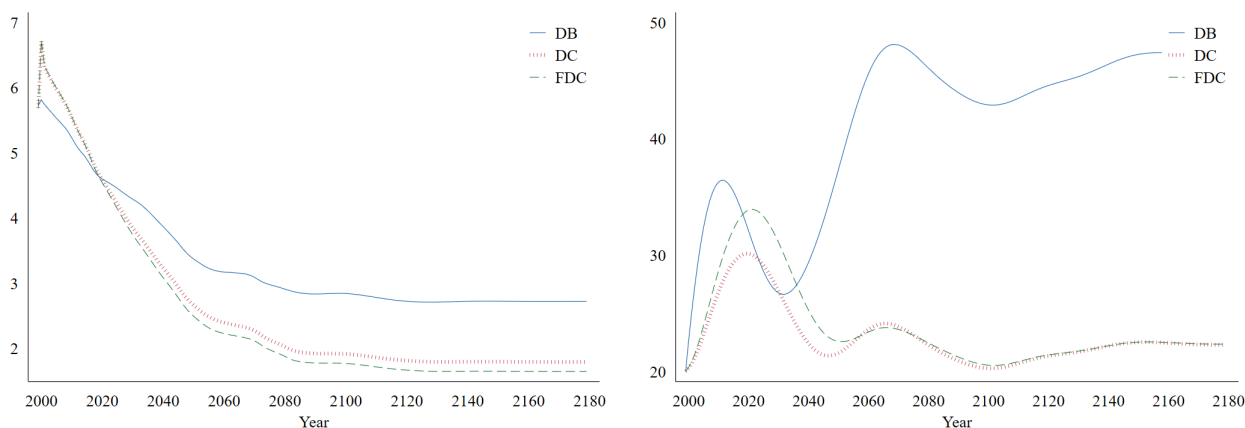
Notes: "FDC" denotes the two-pillar DC system with partial funding, while "DC" denotes a DC system financed entirely on a PAYG basis. "DB" denotes a DB system fully based on PAYG financing. The initial steady state is always the PAYG-DB system. Each line denotes a separate simulation on an economy with the same initial steady state. In the scenarios of a change to DC (with or without funding) the reform is introduced in 1999. The transition is gradual, as described in Section 3.4 in the main text.

E   Macroeconomic effects of reducing the funded pillar of social security

We discuss the effects of capturing social security assets in 2012. This year is suitable for illustration purposes, because around this date most of the countries that earlier introduced partially-funded DC systems started to unfund their social security systems. Results for alternative years for the vote to unfund social security are available upon request.

The impact of the three analyzed policies on social security benefits differs. Policy 1 redirects contributions from the funded pillar to the PAYG-DC pillar. This shift lowers the benefits for two reasons. First, because the interest rate exceeds the growth rate of the wage bill, during working age entitlements in the PAYG pillar accumulate at a lower rate than value of the assets in the funded pillar. Second, benefits-in-payment from the PAYG pillar are also indexed at a rate lower than benefits paid from the funded pillar. Under Policy 2, a one-off asset capture, social security benefits fall for cohorts working during the asset capture. However, because the asset capture is one-off, Policy 2 has no long-run effects. Under Policy 3, the effects of Policy 1 and Policy 2 add up and, hence, social security benefits are lower than under either Policy 1 or Policy 2. The trajectories of the total social security benefit payouts and the deviations from the status-quo of the current two-pillar DC scheme are depicted in Figure E.7. The interest rate and wage rate under the current two-pillar DC scheme are depicted in Figure B.3, while the effects of a policy shift away from the current two-pillar DC scheme are found in Figure E.8.

In addition to the effect on social security benefits, there is an effect through taxes. The shift of contributions under Policy 1 results in a temporarily lower consumption tax rate. Similar to Policy 1, under Policy 2, the deficit is reduced in the PAYG pillar, which benefits current cohorts as it allows to reduce the consumption tax rate. Under Policy 3, as the effects of Policy 1 and Policy 2 add up, the decline in the consumption tax rate is larger than under either of the alternative policies. The trajectories of the consumption tax rate and its deviation from the status-quo of maintaining the two-pillar DC are depicted in Figure E.7.

Lower benefits due to the reduction of the funded pillar result in higher voluntary savings. However, since voluntary savings are taxed and cannot be converted into an annuity at retirement age,[36] this increase fails to make up for the effect of the reduction of the funded pillar on the capital stock – see Figure 6. Nevertheless, with the reaction of private voluntary savings, there is an increase in the overall amount of capital income taxation, because social security assets were not subject to such taxation.

The impact on the labor supply is twofold. First, lower consumption taxes distort less the intra-temporal choice between consumption and leisure, which tends to increase the labor supply. Second, due to the shift away from funding, social security contributions yield a lower return in terms of future benefits, hence the incentive to work is reduced. Depending on which of the effects is stronger, the labor supply rises or falls. For Policy 1 the second effect dominates. The labor supply drops rapidly at the time of introducing the policy change. For Policy 2 the first effect is stronger, and the labor

---

[36]Recall that the return on the annuity exceeds the interest rate, because the funded pillar assets of individuals who die remain in this pillar.

supply increases. In the case of Policy 3, the effect of lower consumption taxes dominates initially, as the decline in taxes is largest in the initial years after the reversal of the original reform, but over time this effect is overtaken by the effect of the distortion associated with the second effect; see Figure 6.

The interest rate and the wage rate under the status-quo of the two-pillar DC scheme were already depicted in Figure B.3, while the effects of the policy shift are found in Figure E.8. The lower capital stock under Policies 1 and 3 compared to Policy 2 results into a higher interest rate under the former.

Figure E.8: Adjustments in the interest rate r_t (left) and the wage rate w_t (right)
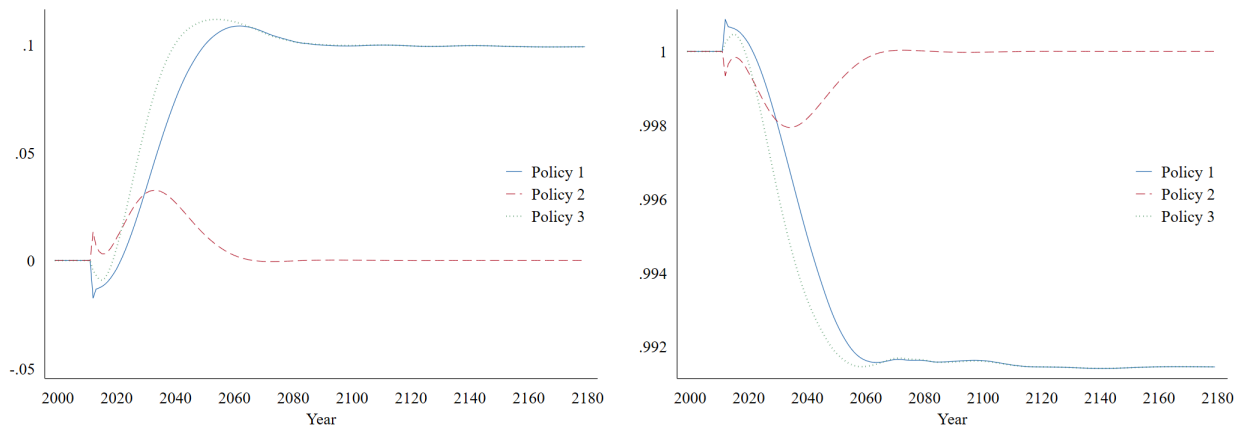
Notes: the interest rate is expressed as the difference relative to the baseline of two-pillar DC (in percentage points). The wage rate is expressed as a ratio relative to the baseline of two-pillar DC. Reported are the results from a vote in 2012. Policy 1 is a permanent shift of contributions from the funded to the PAYG pillar. Policy 2 is a one-off capture of the funde-pillar assets. Policy 3 combines Policies 1 and 2. The results for alternative voting years are available upon request.

Figure E.7: The effect of unfunding on social security benefits and fiscal variables
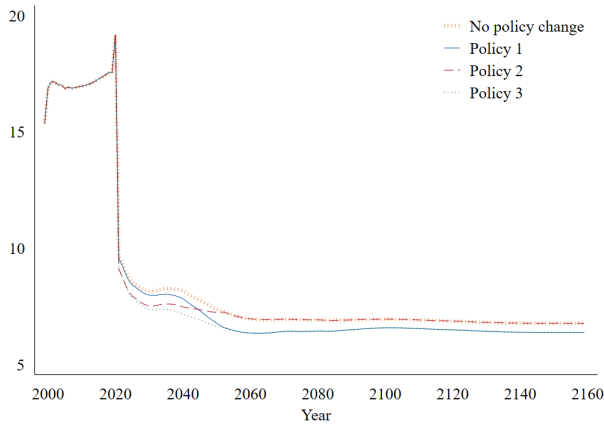
(a) social security benefits
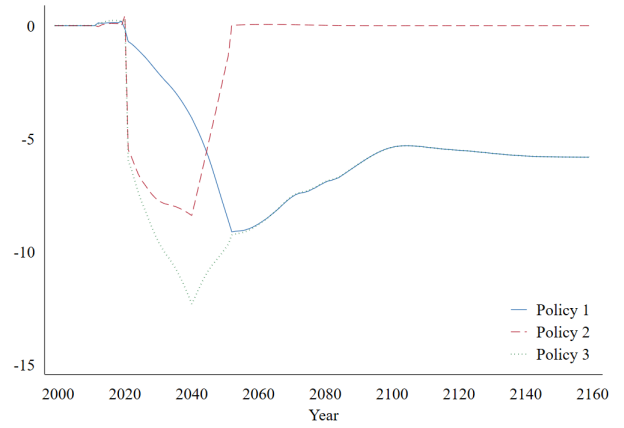
(b) social security benefits
(percent relative to baseline)
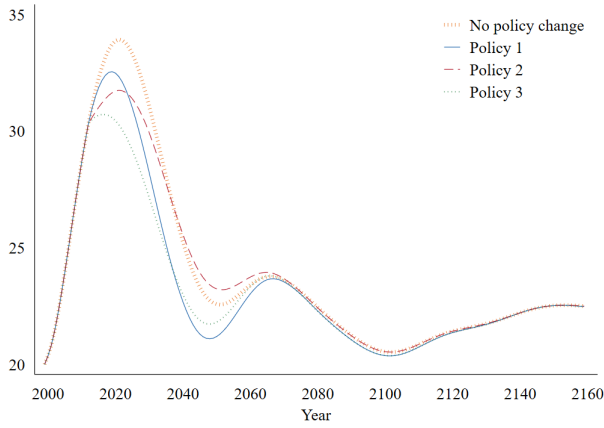
(c) τ_c rate (percent)
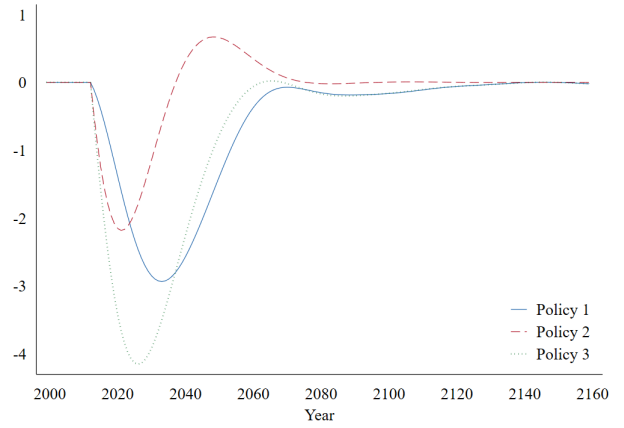
(d) τ_c rate
(percentage point relative to baseline)
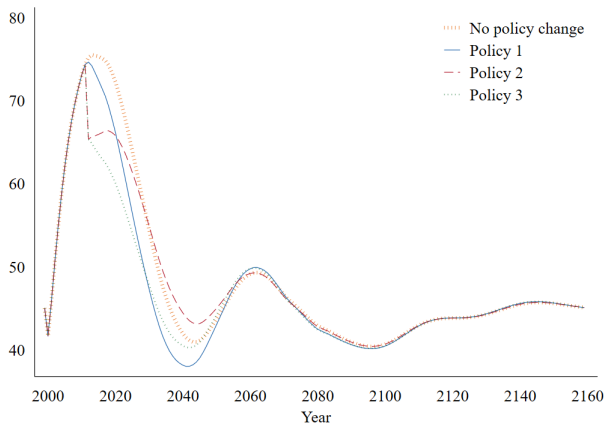
(e) debt share in GDP (percent)
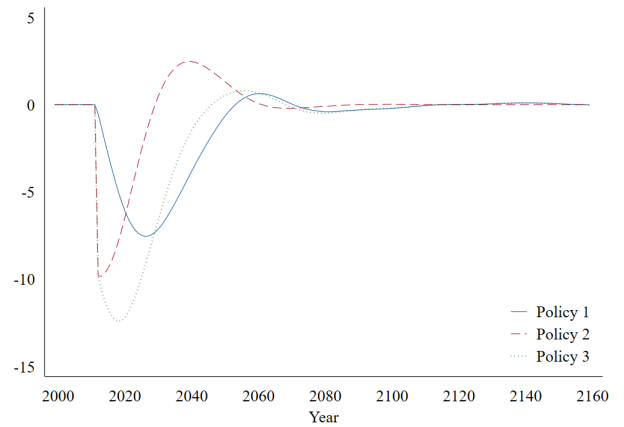
(f) debt share
(percentage point relative to baseline)

Notes: results from a vote in 2012, individuals are uniform within birth cohort. The results from an alternative voting year are available upon request. "No policy changes" and "baseline" refers to the current two-pillar DC scheme. Policy 1 is a permanent shift of the contributions from the funded pillar to the PAYG pillar. Policy 2 is a one-off capturing of the assets accumulated in the funded pillar. Policy 3 combines Policy 1 and Policy 2. The top left panel depicts the total social security benefits paid out in a given year in model units, while the top right panel depicts the deviation from the baseline of no policy change in percent.

F   Intra-cohort heterogeneity: macroeconomic and welfare effects

This appendix reports the relevant tables and figures for the case in which we allow for intra-cohort heterogeneity.

Table F.2: Political support for the three alternatives to the current two-pillar DC

Year of vote	2012	2162	2192	final steady state
Political support in % of those alive		Against two-pillar DC		
Model with intra-cohort heterogeneity				
for Policy 1 - shifting contributions (permanent)	98	100	100	
for Policy 2 - capturing assets (one-off)	54	64	64	
for Policy 3 - combination of Policies 1 and 2	100	98	98	
Long-term welfare effect of Policy 1				-0.34%
Long-term welfare effect of Policy 2				0.00%
Long-term welfare effect of Policy 3				-0.34%

Note: see Notes to Table 3.

Figure F.9: The effect of a successful vote against the current two-pillar DC on retirement benefits and taxes in the presence of intra-cohort heterogeneity

(a) retirement benefits
(percent change from baseline)

(b) τ_t^c
(percentage points change from baseline)
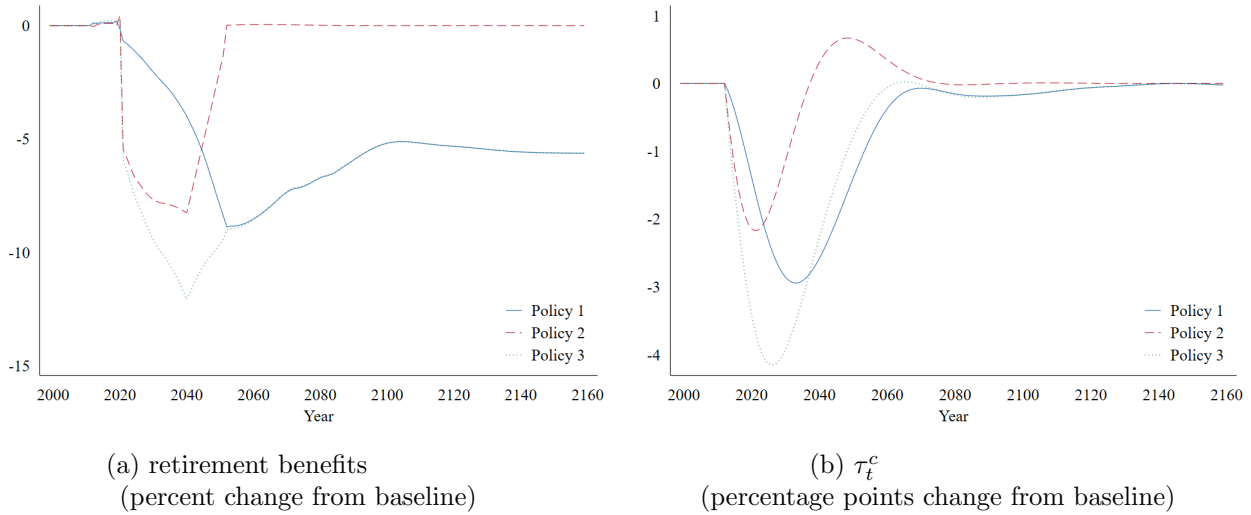
Note: See Notes to Figure E.7.

Figure F.10: Welfare effects and its decomposition of a policy change for a vote in 2012 by birth year and with intra-cohort heterogeneity

(a) Policy 1: shifting contributions (permanent)
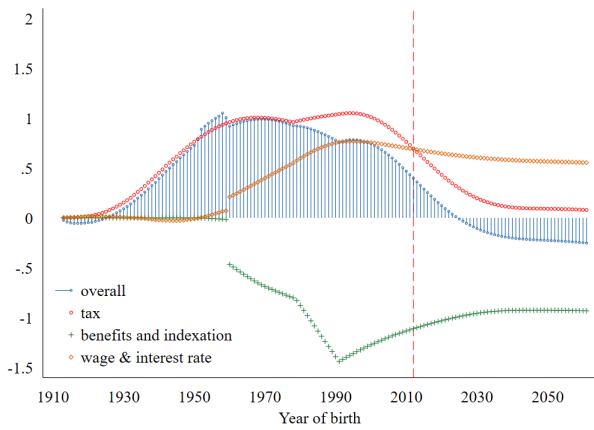
(b) Policy 2: capturing assets (one-off)
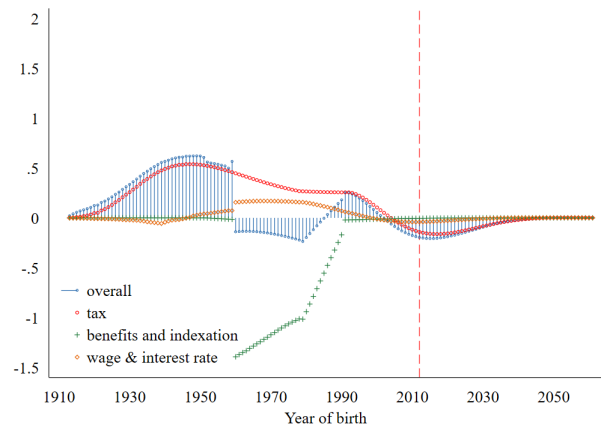
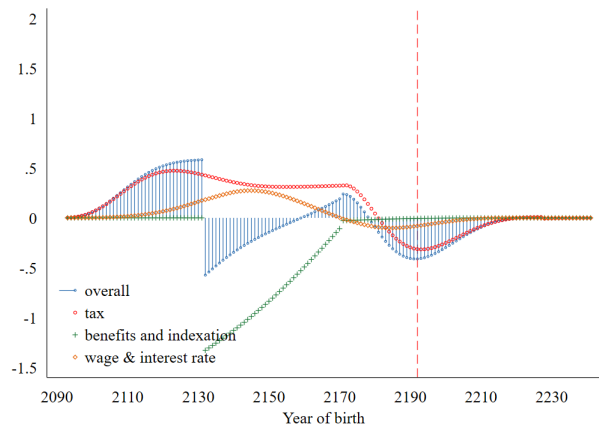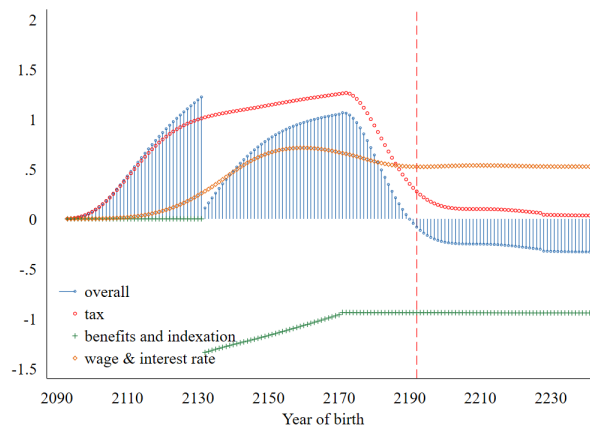(c) Policy 3: combination of Policy 1 and Policy 2
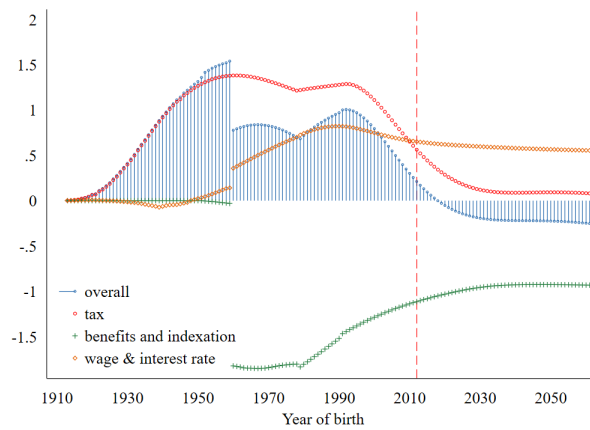

Notes: see Figure 7.

Figure F.11: Welfare effects and its decomposition of a policy change for a vote in 2192 by birth year and with intra-cohort heterogeneity

(a) Policy 1: shifting contributions (permanent)
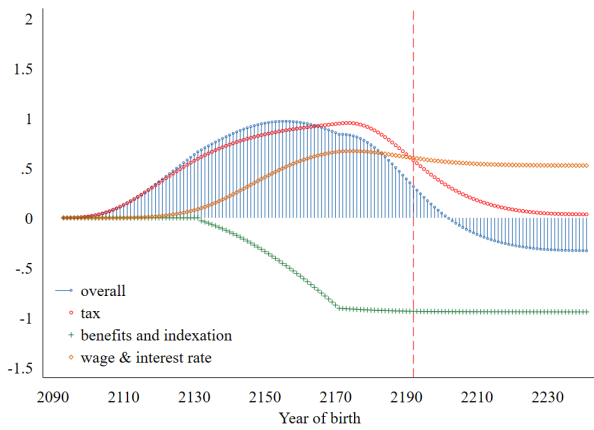
(b) Policy 2: capturing assets (one-off)

(c) Policy 3: combination of Policy 1 and Policy 2

Notes: see Figure 7.

Figure F.12: The effect of reducing the funded pillar on retirement benefits and fiscal variables in the presence of intra-cohort heterogeneity
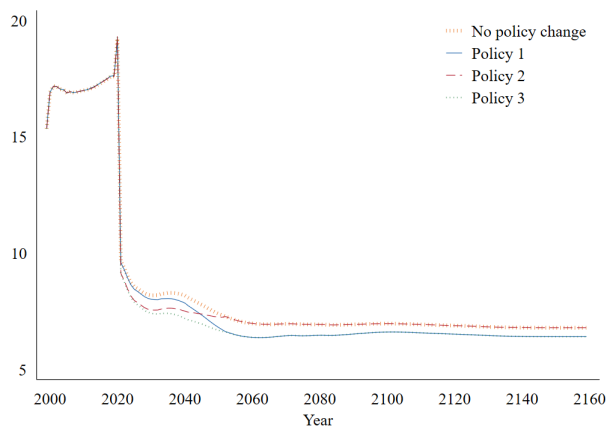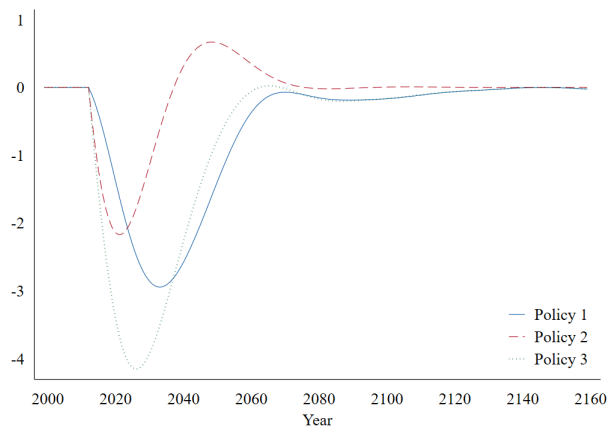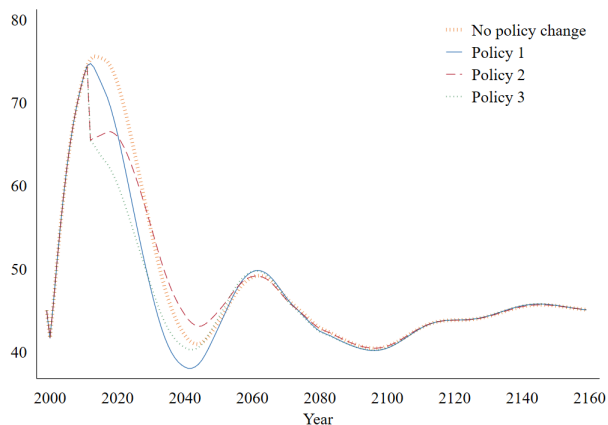

(a) social security benefits
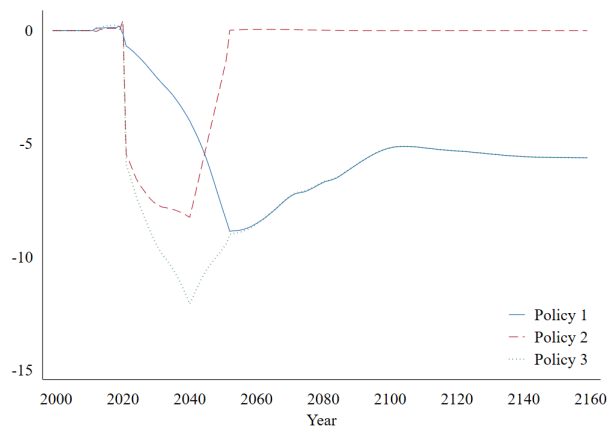
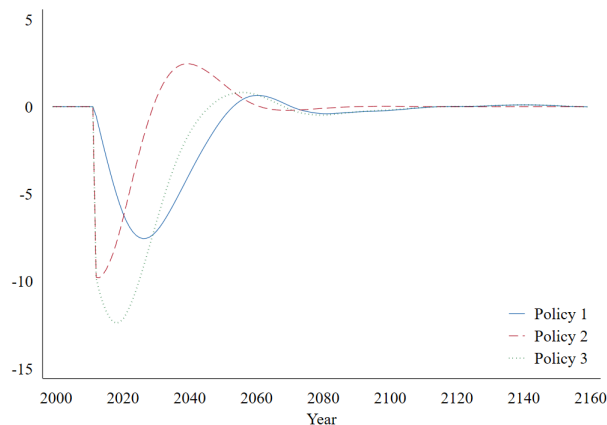(b) social security benefits
(percent change from baseline)
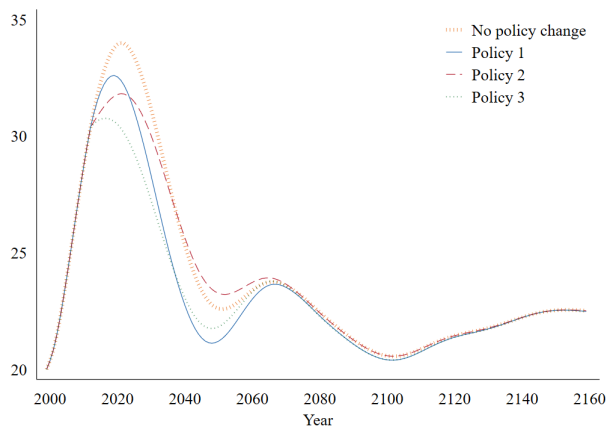
(c) τ_c rate (percent)

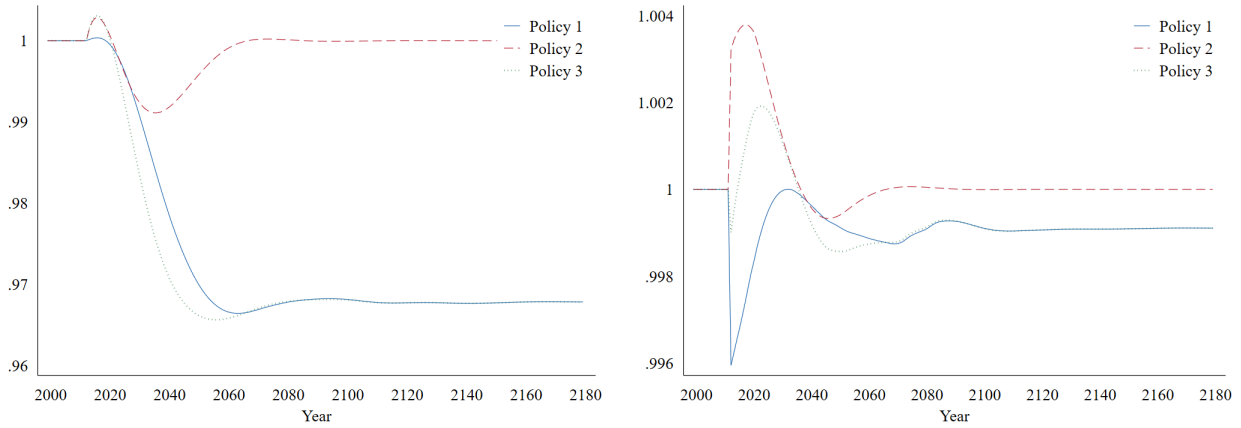(d) τ_c rate (percentage point change from baseline)

(e) debt share in GDP

(f) debt share in GDP
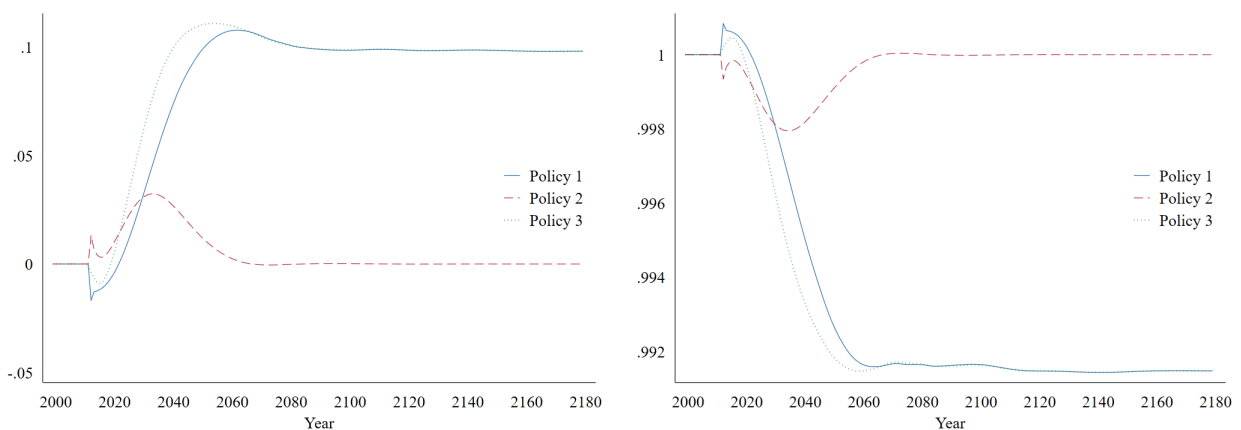(percentage point change from baseline)

Notes: results from a vote in 2012 with intra-cohort heterogeneity. Further, see Notes to Figure E.7.

Figure F.13: Adjustments in capital K_t (left) and labor supply L_t (right) in the presence of intra-cohort heterogeneity

Notes: measures expressed in terms of ratio relative to baseline of two-pillar DC. Reported are the results from a vote in 2012. The results for alternative voting years are available upon request.

Figure F.14: Adjustments in the interest rate r_t (left) and the wage rate w_t (right) in the presence of intra-cohort heterogeneity


Notes: the interest rate is expressed as the difference relative to the baseline (in percentage points). The wage rate is expressed as a ratio relative to the baseline. Reported are the results from a vote in 2012. The results for alternative voting years are available upon request.

G   The effects on poverty

We also show the aggregate effects on both absolute and relative poverty. We define relative poverty as consumption below 60% of current median consumption in any given year. We define absolute poverty as consumption below 60% of consumption in the initial PAYG-DB steady state (adjusting for technological progress and population change).[37] The evolution of the poverty measures is displayed in Figure G.15.

Figure G.15: Evolution of poverty

(a) relative poverty: all

(b) relative poverty: elderly

(c) absolute poverty: all

(d) absolute poverty: elderly
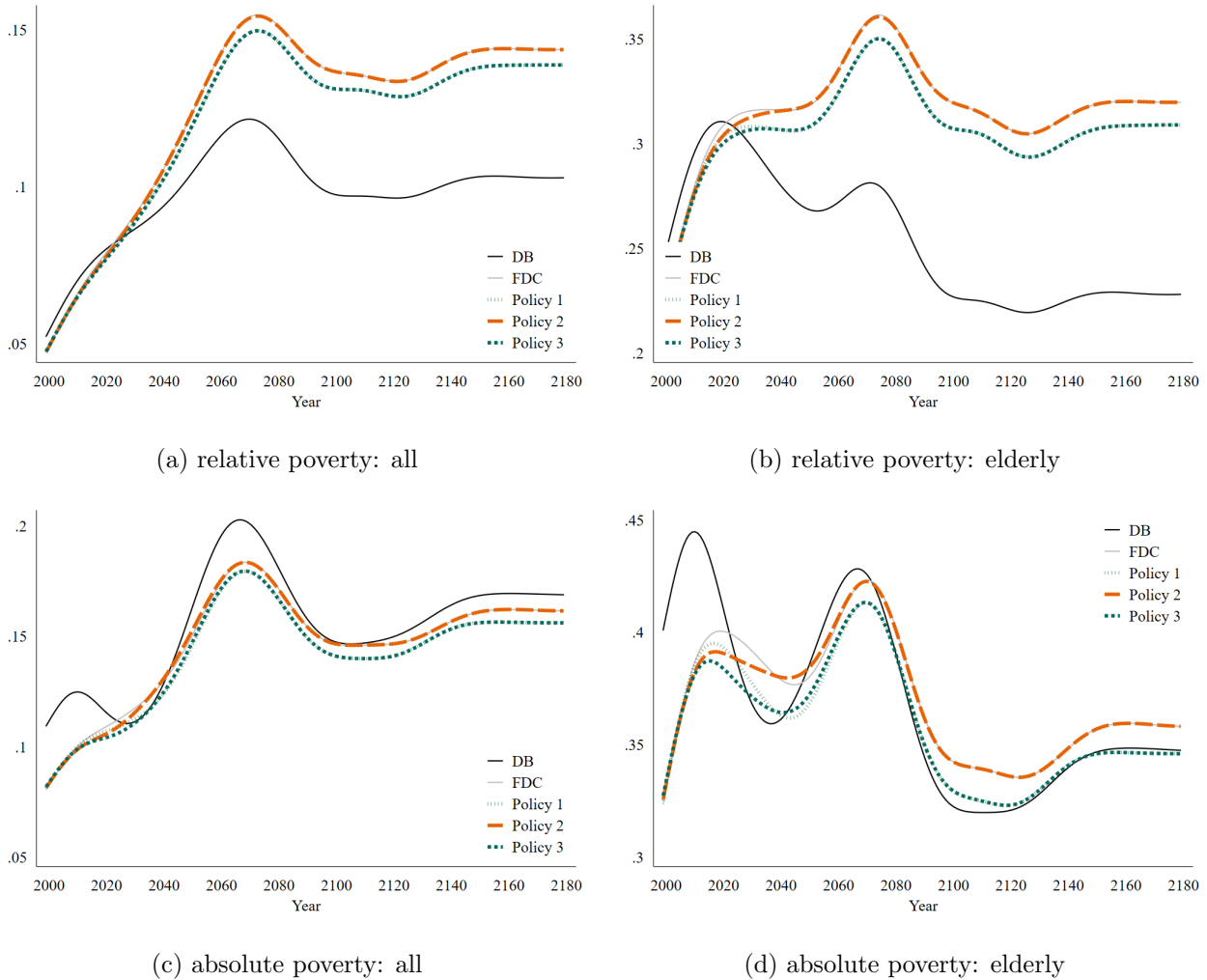
Note: relative poverty is defined as the fraction of the population with consumption below 60% of median consumption in a given year, while absolute poverty is defined as the fraction of the population with consumption below 60% of median consumption in the initial steady state. Old-age poverty measures the fraction of poor elderly as a share of the total group of elderly. Policy 1 is a permanent shift of the contributions from the funded pillar to the PAYG-DC pillar. Policy 2 denotes a one-off capturing of the assets accumulated in the funded pillar. Policy 3 combines Policies 1 and 2.

Poverty, both absolute and relative, would increase the least if the PAYG-DB system was continued. This is because the majority of the poor households are in old-age: the higher retirement benefits under the PAYG-DB system imply that a larger fraction of them can afford sufficient consumption.

---

[37]That is, we scale back quantities by the amount of technological progress and the change in the size of population relative to the initial steady state.

Moreover, the higher consumption tax under this system implies lower average consumption, hence fewer households falling below the relative poverty threshold.

The shift from PAYG-DB to single-pillar PAYG-DC has a strong positive effect on the share of old-age households that cannot afford consumption above 60% of the median in a given year. Hence, relative poverty among this group and, therefore, the population at large, rises strongly. There are two driving factors here. The main driving factor is the fall in social security benefits.

The direct consequence of the asset capture under Policy 2 is a decline of overall consumption, relative to the current two-pillar DC scheme. Overall consumption falls, because consumption needs to given up to stem the fall in social security benefits. This decline lowers the threshold above which a household is no longer defined as poor, hence relative poverty declines. The development of relative poverty under Policies 1 and 3 is very similar to that under Policy 2.