

Sharing Demographic Risk – Who is Afraid of the Baby Bust?*

Alexander Ludwig[†] Michael Reiter[‡]

August 2008

Abstract

We model the optimal reaction of a public PAYG pension system to demographic shocks. We compare the ex-ante first best and second best solution of a Ramsey planner with full commitment to the outcome under simple third best rules that mimic the pension systems observed in the real world. The model, in particular the pension system, is calibrated to the German economy. The objective of the social planner is calibrated such that the size of the German pension system was optimal under the economic and demographic conditions of the 1960s. We find that the German system comes relatively close to the second-best solution, especially when labor market distortions are correctly modelled. Furthermore, the German system and a constant contribution rate lead to a lower variability of lifetime utility than does the second best policy. The recent baby-boom/baby-bust cycle leads to welfare losses of about 5% of lifetime consumption for some cohorts. We argue that it is crucial for these results to model correctly the labor market distortions arising from the pension system.

JEL classification: E62, H3, H55

Keywords: social security; pension design; optimal fiscal policy; demographic uncertainty

*We thank Alan Auerbach, Axel Börsch-Supan, Juan Carlos Conesa, Peter Diamond, Martin Gervais, Ashok Kaul, Dirk Krüger, Alice Schoonbroodt and seminar participants at SCE 2006 in Cyprus, Netspar 2007 Workshop in Tilburg, MIT, Institute of Advanced Studies in Vienna, Universitat Pompeu Fabra, University of Frankfurt, University of Mannheim, LMU Munich and University of Würzburg for helpful comments and discussions. Alex Ludwig gratefully acknowledges financial support from the German National Research Foundation (DFG) through SFB 504, the State of Baden-Württemberg and the German Insurers Association (GDV).

[†]Mannheim Research Institute for the Economics of Aging (MEA); Universität Mannheim; L13, 17; 68131 Mannheim; Germany; Email: ludwig@mea.uni-mannheim.de.

[‡]Institute for Advanced Studies (IHS); Department of Economics & Finance; Stumpergasse 56; 1060 Vienna; Austria; Email: michael.reiter@ihs.ac.at.

1 Introduction

All over the world, demographic transition processes are characterized by falling fertility rates and increasing life-expectancy which leads to sharp decreases of working-age population ratios and corresponding increases in old-age dependency ratios. This will put a strain on public budgets, especially on existing public pay-as-you-go (PAYG) pension systems, as is highlighted by the ongoing reform processes in many countries.

From an analytical perspective, it is useful to separate two distinct aspects of demographic change. First, the continuous increase in life expectancy will lead to a continuous increase in the old-age dependency ratio (OADR) – the ratio of the population aged 65 and older to the working age population aged 20 to 64 – which will make current PAYG systems unsustainable unless the retirement age is significantly increased. There is a large literature dealing with this problem (cf. Krüger and Ludwig (2007) and references there). A second issue are the fluctuations in the fertility rate, notably the recent baby-boom/baby-bust cycle, which lead to significant *fluctuations* in the OADR. In this paper we focus on the second aspect of demographic change, and we therefore abstract from secular changes in longevity.

The fluctuations of the OADR caused by changing fertility patterns affect different cohorts in an unequal way, through their effects on factor prices and on the contributions and benefits of pension systems. From an ex-ante perspective, these fluctuations are stochastic, and the question arises as of how to design public pension systems such that they contribute to an efficient sharing of the risks that arise through demographic fluctuations (Bohn 2001). It is well known that a decentralized OLG economy cannot allocate all the risks efficiently, since living generations cannot write contracts with generations not yet born (Gordon and Varian 1988). Fiscal policy, in particular a public pension system, may then have a role in making the intergenerational allocation more efficient. This is the subject of this paper.

We address the issue of optimal pension policy in a standard overlapping generations model with a realistic demographic structure. We compare three different types of policy: (i) the first-best efficient response to demographic shocks, (ii) the second best optimal policy, defined as the optimal Ramsey policy under commitment where the setup is similar to Erosa and Gervais (2002) and (iii) several simple rules that give a stylized description of pension systems currently in place in developed countries. Since the German government has recently enacted a pension reform that explicitly reacts to changes in demographic variables, our model is calibrated to German data, and we investigate the efficiency aspects of the German pension system.

We place special emphasis on the distortionary effects of real-world pension systems on household labor supply. These distortions are often exaggerated in the literature by the assumption that pension contributions are tied to labor income, whereas benefits are lump sum. We model the German pension system correctly by assuming that, within a cohort, pension benefits are proportional to life-time contributions. According to this setup, the pension system is distortionary only to the extent that the rate of return in the pension system is lower than the market return on financial assets, such that households would

prefer not to contribute to the public system, but are forced to do so. The distortionary effect of the pension system is aggravated by the fact that it adds to the already existing labor supply distortions of general income taxation. Since the marginal excess burden is rising in the tax level, the undesirable side effects of public pensions are the more severe the higher is the “background distortion” from the general tax system. We highlight the importance of this form of background distortion in this paper. We also show that a realistic calibration of the labor supply elasticity is very important.

Our paper is most closely related to Bohn (2001) who analyzes the risk sharing properties of alternative pension policies and characterizes what an efficient response would be. A main finding is that favourable factor price movements make the members of small cohorts, the baby bust generation, better off, even after accounting for the additional burden that these cohorts face by having to finance the pensions of the baby boom generation. Contrary to conventional wisdom, an efficient response would require to even *raise* the benefit levels of the baby boom generations. This is confirmed by Sanchez-Marcos and Martin (2006) in a more realistic OLG setting. We show that this conclusion is reversed once the distortionary nature of the public pension system is appropriately accounted for.

We evaluate simple policy rules by their distance from second-best outcomes in policy space and also by the entailed welfare loss, in terms of a standard Utilitarian welfare criterion. We find that a constant benefit rule is relatively close to the second-best policy in most of our parameterizations. Moreover, realistic simple rules outperform the Ramsey solution if a different, but plausible criterion is used, namely to minimize the variability of lifetime welfare. This suggests that, in the political process, other criteria than ex-ante efficiency are dominant.

In our setup, the existence of a pension system reflects the preferences of a social planner who wants to redistribute resources from future to current generations. Therefore, our results are consistent with Krueger and Kubler (2006), who compare economies with and without social security systems in the presence of aggregate technology shocks. They find that only the early generations benefit from the introduction of a PAYG system. For future generations, the crowding-out of private capital formation due to PAYG systems by far outweighs the favorable effects of risk-sharing.¹ For this paper, we accept the social preferences that lead to a systematic redistribution from future to current generations. We identify the respective weights of different cohorts in the social welfare function which rationalize the actual size of the PAYG systems. Given those preferences, we ask how the parameters of the pension system should be modified optimally in response to aggregate shocks.

In a stylized two-period OLG model with stochastic production, Gottardi and Kubler (2008) characterize the conditions for a social security system to provide an ex-ante Pareto improvement. In their framework, the government can provide small stochastic lump sum transfers from the young to the old, which are required to be non-negative. We differ from their contribution by focusing on the distortionary aspects of pension systems, and

¹Similar results have been reported for models featuring only idiosyncratic shocks by Conesa and Krueger (1999).

by searching for optimal pension systems.

The outline of the paper is as follows. Section 2 presents the main elements of our model. In Section 3 we discuss the properties of first best and second best allocations. We also investigate the equivalence between a PAYG pension system and a system with government debt. In Section 4 we then extend the model presented in Section 2 to a fully specified model which can be used to analyze the quantitative questions we are interested in. Section 5 presents our main results. Section 6 concludes.

2 The OLG Economy

2.1 Time and Risk

Time is discrete and extends from $t = 0, \dots, \infty$. The only source of aggregate uncertainty is the stochastic fertility rate. To keep the analysis focussed, we abstract from all other potential sources of aggregate risk.

2.2 Demographic Processes

The exogenous driving force in our model is a time and age specific stochastic fertility process. For simplicity, we take age-specific mortality rates as constant over time and we abstract from migration. Starting from an initial age distribution of population in year 0, the demographic distribution in each year t evolves according to

$$N_{t,i} = \begin{cases} N_{t-1,i-1}\varsigma_{i-1} & \text{for } i > 0 \\ \sum_{i=0}^I f_{t-1,i}N_{t-1,i} & \text{for } i = 0 \end{cases} \quad (1)$$

where ς_i denotes age-specific mortality rates and $f_{t,i}$ denotes age and time specific fertility rates.

2.3 The Production Sector

Firms are assumed to operate in a perfectly competitive market, using the constant returns to scale production function

$$Y_t = F(K_{t-1}, Z_t L_t) \quad (2)$$

where Y_t is output, K_{t-1} denotes the capital stock at the beginning of period t , and L_t is aggregate effective labor input in period t . Z_t is the technology level in period t that grows at the constant rate g , hence $Z_t = (1 + g)^t Z_0$.

Profit maximization of the representative firm implies that the aggregate wage rate, w_t , and the rate of return to capital, r_t , are given by

$$w_t = F_L(K_{t-1}, Z_t L_t) \quad (3)$$

$$r_t = F_K(K_{t-1}, Z_t L_t) - \delta \quad (4)$$

where δ is the depreciation rate, and F_K and F_L denote the partial derivative of the production function with respect to the first and second argument, respectively.

2.4 The Household

The preferences of a household born in period t over distributions of consumption, $C_{t,i}$, and leisure, $1 - L_{t,i}$, are represented by the expected life-time utility

$$\mathcal{U}_t = \mathbb{E}_t \sum_{i=i^A}^I \beta^i \pi_i U(C_{t+i,i}, L_{t+i,i}), \quad (5)$$

where β is the pure time discount factor of the individual household and where expectations are taken over the shocks to fertility rates. In addition to pure time discounting, households discount future utility with the probability of surviving to age i , which we denote by π_i . Notice that $\pi_i = \prod_{j=0}^{i-1} \varsigma_j$, where ς_j is the conditional probability to survive from j to $j+1$.

In this paper we analyze the role of the pension system for intergenerational risk-sharing, not its merit in providing annuity markets, although this may be an important aspect in reality. We therefore assume perfect annuity markets within cohorts (Yaari 1965). By this idealized setup individual households are insured against the idiosyncratic mortality risk. Accordingly, we define the adjusted net interest rate, $r_{t,i}^n$, including annuitization and capital taxation at rate τ_t^c as

$$1 + r_{t,i}^n = \frac{1 + r_t(1 - \tau_t^c)}{\varsigma_{i-1}} \quad (6)$$

Households, indexed by i , receive labor income during their working period and pension income $p_{t,i}$ when they are retired. Government taxes labor income at rates $\tau^l + \tau_t^p$, whereby τ_t^p are contributions to the pension system and τ^l is used to finance a lump sum redistribution within the cohort, $T_{t,i}$. Net labor income is accordingly given by

$$w_{t,i}(1 - \tau^l - \tau_t^p)L_{t,i}, \quad (7)$$

where $w_{t,i} = \epsilon_i w_t$ is the age-specific gross wage, and ϵ_i is age-specific productivity, which displays a hump-shaped profile. Notice that $\epsilon_i = 0$ for $i \geq i^R$, where i^R is the retirement age.

Collecting these components and denoting the financial wealth at the beginning of period t by $A_{t-1,i-1}$, maximization of the household's inter-temporal utility is subject to a dynamic budget constraint given by

$$A_{t,i} + C_{t,i} = (1 + r_{t,i}^n)A_{t-1,i-1} + w_{t,i}(1 - \tau^l - \tau_t^p)L_{t,i} + p_{t,i} + T_{t,i}. \quad (8)$$

Households start adult life with $A_{t-1,i^A-1} = 0$, because we abstract from intended or accidental bequests. The government lump-sum transfer $T_{t,i}$, and pension income $p_{t,i}$ will be described in the next section.

2.5 The Government Sector

The government has access to a set of fiscal policy instruments and a commitment technology to implement its fiscal policy. It uses these instruments to implement efficient risk sharing between generations. Pension payments, $p_{t,i}$, are financed by payroll taxes on labor income, τ_t^p . In addition, the government has access to government debt, D_t , and capital income taxes, τ_t^c . Denoting by D_{t-1} the amount of outstanding government debt issued as bonds in period $t - 1$, the government budget constraint is accordingly given by

$$D_t = (1 + r_t)D_{t-1} + \sum_{i=i^A}^I (p_{t,i} - \tau_t^p w_{t,i} L_{t,i} - \tau_t^c r_t A_{t-1,i-1}) N_{t,i} \quad \text{for all } t. \quad (9)$$

Recall that $w_{t,i} = 0$ for $i \geq i^R$ and $p_{t,i} = 0$ for $i < i^R$. We describe the details of the pension system linking payments $p_{t,i}$ to contributions $\tau_{t,i}^p$ in Sections 3.3 and 4.2.

Exogenous to the government is the “background distortion” in the labor market, captured by the labor tax rate τ^l . We assume it is constant and finances a within-cohort transfer $T_{t,i}$:

$$T_{t,i} = \tau^l w_{t,i} L_{t,i}, \quad i = i^A, \dots, i^R - 1. \quad (10)$$

The presence of this transfer system allows us to calibrate the degree of labor market distortions in the economy. Background distortions are a metaphor for the effects of redistribution within the general tax system, which reflects preferences of society that we do not model explicitly. They should also capture the effects of distortionary taxation to finance public consumption. For the purpose of studying pension policy, we assume the government cannot change τ^l .

2.6 Market Structure

Generations born in t can obviously not write an insurance contract with generations born in $t + \tau$, $\tau > 0$. This is the fundamental form of market incompleteness in OLG models that was pointed out by Gordon and Varian (1988). We assume additional market incompleteness by focusing on an economy with only one asset, the physical capital stock K_t , which was also assumed, e.g., by Krueger and Kubler (2006). Markets in this economy are spot markets for consumption, labor, capital, all of which are assumed to be perfectly competitive. We assume the economy to be closed.

2.7 Market Clearing and Aggregation

The aggregate resource constraint of the economy in period t is

$$Y_t = F(K_{t-1}, Z_t L_t) \quad (11a)$$

$$= K_t - (1 - \delta_t)K_{t-1} + C_t \quad (11b)$$

where

$$C_t = \sum_{i=0}^I C_{t,i} N_{t,i}, \quad L_t = \sum_{i=0}^I \epsilon_i L_{t,i} N_{t,i}, \quad K_t = \sum_{i=0}^I A_{t,i} N_{t,i} - D_t. \quad (12)$$

3 First and Second-Best Allocations

3.1 Efficient Allocations

Two concepts of Pareto efficiency have been defined in the literature on OLG economies (Demange 2002). The first is the concept of *conditional* or *interim* Pareto optimality, which evaluates allocations conditional on the time period of birth of an agent. The second concept is the concept of efficiency in the stronger sense of *ex-ante* efficiency which we adopt here:

Definition 1. Let $\mathbb{E}_0 \mathcal{U}_t$ be the time 0 expectation of the life-time discounted utility of an agent. A feasible allocation, x , is said to be ex-ante optimal if there is no other feasible allocation, \tilde{x} , such that $\mathbb{E}_0 \mathcal{U}_t(\tilde{x}) \geq \mathbb{E}_0 \mathcal{U}_t(x)$ for all t with the inequality being strict for at least one t .

This strong notion of efficiency requires that all the insurance possibilities between generations that exist at $t = 0$ are exploited. As Gordon and Varian (1988) have pointed out, a decentralized equilibrium can in general not achieve ex-ante efficiency, because insurance contracts with generations that are not yet born cannot be written.

Following Bohn (2001), we consider ex-ante efficient symmetric allocations.² As in Samuelson (1968), Atkinson and Sandmo (1980) and Erosa and Gervais (2002), an ex-ante efficient allocation maximizes the Utilitarian social welfare function

$$\mathbb{E}_0 \sum_{t=-I}^{\infty} \Omega_t N_{t,0} \mathcal{U}_t \quad (13)$$

subject to the aggregate resource constraints (11) and (12) and the specifications of the exogenous stochastic process for fertility.³ Here, Ω_t is the welfare weight of generation t

²Obviously, not all ex-ante efficient allocations solve (13), only those allocations that are symmetric in the sense that they assign the same consumption and leisure to each born member of a generation. From our ex-ante perspective, it is probably best to reinterpret the model of stochastic (but exogenous) population as a model with deterministic population in the following way. Say that in each period t there is a fixed set of households, some of which are randomly selected to be born and some of which are not. In the latter case, households receive utility zero and are not affected by economic allocations. They do therefore not appear in (13). Formula (13) gives a fixed weight to each born household. The concept of Pareto-optimality is more problematic in the case of endogenous fertility, cf. Golosov et al. (2007).

³Social welfare functions of the above type with discounting of future generations have more recently also been adopted in dynastic models by Bernheim (1989), Caplin and Leahy (2004) and Farhi and Werning (2006, 2008). See Stern (2006) for a critical review of the discussion regarding the ethical justifications for discounting in an intergenerational context.

where $\Omega_0 = 1$. These weights reflect the relative preferences of a social planner for present and future generations. Different weights give different efficient allocations.

The next proposition characterizes optimal allocations:

Proposition 1. *An allocation is a solution to the optimization problem (13) if and only if it satisfies (11), (12) and the following first order conditions:*

$$-U_L(C_{t,i}, L_{t,i}) = Z_t \epsilon_i F_L(K_{t-1}, Z_t L_t) U_C(C_{t,i}, L_{t,i}) \quad (14a)$$

$$U_C(C_{t,i}, L_{t,i}) = \beta \mathbb{E}_t (1 + F_K(K_t, Z_{t+1} L_{t+1}) - \delta) U_C(C_{t+1,i+1}, L_{t+1,i+1}) \quad (14b)$$

$$\frac{U_C(C_{t,i}, L_{t,i})}{U_C(C_{t,j}, L_{t,j})} = \frac{\Omega_{t-j}}{\Omega_{t-i}} \beta^{j-i}, \quad \forall i, j \in (0, \dots, I) \quad (14c)$$

Proof. See Appendix A.3. □

Conditions (14a) and (14b) are the first order conditions of the household problem in an undistorted decentralized equilibrium with perfect annuity markets. Condition (14c) illustrates the perfect risk-sharing between cohorts. The ratio of marginal utilities from consumption of different generations is constant. This means that the effects of a shock (past or present) are spread equally across generations in t . In a decentralized equilibrium, there is no mechanism to achieve this. However, a government can implement such an allocation if it has stochastic lump sum taxes at its disposal.

3.2 The Risk-Free Economy on the Balanced Growth Path

We next concentrate on the implications of our model along the balanced growth path of a risk-free economy, in which variables grow at constant rates. We call such a path a deterministic steady state. Along the balanced growth path, labor input is constant, capital and output grow at rates $g + n$, where n is the steady state population growth rate. The marginal product of capital is constant at $1 + \bar{r} = 1 + F_K(K_0(1 + g + n)^t, L_0, 1) - \delta$ and the wage rate grows at the constant rate g .

To get balanced growth, we make the following homogeneity assumptions:

Assumption 1. *The marginal utility of consumption, $U_C(C_{t,i}, L_{t,i})$, is homogeneous of degree $-\theta$ in $C_{t,i}$.*

Assumption 2. *Social weights are $\Omega_t = \omega^t$.*

With these assumptions we can derive the following relationship to hold between the exogenous growth rate of technology, g , the steady state interest rate, \bar{r} , and the welfare weights, ω :

Proposition 2. *Consider the steady state of a risk-free economy that maximizes (13) and denote the steady state interest rate by \bar{r} . Then $\omega = \left(\frac{(1+g)^\theta}{1+\bar{r}} \right)$.*

Proof. See Appendix A.3. □

The relationship between the government discount factor and the interest rate in Proposition 2 is also referred to as the *modified golden rule* property (Samuelson 1968) and has some interesting implications. First, it shows that a competitive deterministic economy without government, and an efficient allocation that maximizes (13), have the same steady state if and only if

$$\omega_n = \frac{(1+g)^\theta}{1+\bar{r}_d} \quad (15)$$

where \bar{r}_d is the interest rate of the competitive economy. This means that a government with a discount rate given in Equ. (15) is inactive in steady state; it does not interfere with the private allocation. We call such a government “neutral”:

Definition 2. A government with a discount factor ω_n satisfying (15) is called a *neutral* government. A government with a discount factor $\omega < \omega_n$ is called an *impatient* government.

An impatient government gives relatively higher weight to currently living generations and will use its policy instruments to redistribute from future to current generations. We do not consider the case of a “patient” government with $\omega > \omega_n$.

More generally, in an allocation maximizing (13), the interest rate is given by

$$\bar{r} = \frac{(1+g)^\theta}{\omega} - 1. \quad (16)$$

Changes of any parameter other than g , θ and ω have no effect on the optimal interest rate. For example, if we change household preferences in a way that generates higher savings, the government sets policy instruments such that it crowds out private savings until the condition (16) is met.

3.3 The Fiscal Policy Setup: Debt vs. Pensions

It is well known that a pension system and government debt are basically equivalent policy instruments, at least in deterministic models or models with complete markets. Here we derive conditions under which equivalence holds in a stochastic incomplete markets model with distortive taxation. This means that all taxes and transfers that a household has to pay are linearly related to the household decisions on labor supply and saving. There are no lump sum taxes or transfers. For reasons of space, we confine ourselves to the case of two-period lives, where households work in the first period and retire in the second. Similar results hold with many cohorts, if the government has age-dependent taxes available; this is investigated in detail in Ludwig and Reiter (2008).

In a stochastic setup, the pension system serves two purposes. First, it shifts resources from future to current generations, which we assume reflects the preferences of the policy maker. Second, it helps to spread the effects of shocks efficiently between different generations. To highlight both aspects, we distinguish between three different fiscal policy regimes:

1. Debt policy (**DP**). The instruments are government debt, a labor income tax at rate τ_t^d and the capital tax τ_t^c . The capital tax is state dependent, i.e., the tax rate that applies in period t is only determined in t . Government debt follows

$$(1 + r_t)d_{t-1,0}N_{t-1,0} - d_{t,0}N_{t,0} = \tau_t^d w_{t,0}L_{t,0}N_{t,0} + \tau_t^c r_t(k_{t-1,0} + d_{t-1,0})N_{t-1,0} \quad \text{for all } t. \quad (17)$$

Here, $k_{t,0}$ and $d_{t,0}$ denote asset holdings in the form of capital and debt, respectively, such that $k_{t,0} + d_{t,0} = A_{t,0}$. As only the young are savers in our simplified two generations economy, we have that $K_t = k_{t,0}N_{t,0}$ and $D_t = d_{t,0}N_{t,0}$. We assume a no-Ponzi condition on the government.

The household budget constraints are

$$k_{t,0} + d_{t,0} + C_{t,0} = w_{t,0}(1 - \tau_t^d - \tau^l)L_{t,0} + T_{t,0} \quad (18a)$$

$$C_{t+1,1} = \frac{1 + r_{t+1}(1 - \tau_{t+1}^c)}{\varsigma_{t,0}}(k_{t,0} + d_{t,0}) \quad (18b)$$

and the household first order conditions are

$$U_C(C_{t,0}, L_{t,0}) = \beta \varsigma_{t,0} \mathbb{E}_t \left[\frac{1 + r_{t+1}(1 - \tau_{t+1}^c)}{\varsigma_{t,0}} U_C(C_{t+1,1}, 0) \right] \quad (19a)$$

$$-U_L(C_{t,0}, L_{t,0}) = w_{t,0}(1 - \tau_t^d - \tau^l)U_C(C_{t,0}, L_{t,0}) \quad (19b)$$

2. Pension policy with a predetermined pension factor (**PP1**). The instruments are pension contributions τ_t^p , the capital tax τ_t^c , and a pension benefit factor \tilde{b}_t . The capital tax is specified as in **DP**. Pension income of the old in period $t + 1$ is given by

$$p_{t+1,1} = \tau_{t,0}^p w_{t,0} L_{t,0} \frac{1 + r_{t+1}(1 - \tau_{t+1}^c)}{\varsigma_{t,0}} \tilde{b}_t. \quad (20)$$

It is the product of three components: the past contributions to the pension system, $\tau_{t,0}^p w_{t,0} L_{t,0}$; the adjusted interest on contributions, $\frac{1 + r_{t+1}(1 - \tau_{t+1}^c)}{\varsigma_{t,0}}$; the pension factor, \tilde{b}_t , which is already fixed in period t (policy instruments carry a tilde if they are predetermined, which means that the rate applied in $t + 1$ is already determined in t). The government budget is balanced in every period:

$$\tilde{b}_{t-1} \frac{1 + r_t(1 - \tau_t^c)}{\varsigma_{t-1,0}} \tau_{t-1,0}^p w_{t-1,0} L_{t-1,0} N_{t,1} = \tau_{t,0}^p w_{t,0} L_{t,0} N_{t,0} + \tau_t^c r_t k_{t-1,0} N_{t-1,0} \quad \text{for all } t. \quad (21)$$

The household budget constraints are

$$k_{t,0} + C_{t,0} = w_{t,0}(1 - \tau_t^p - \tau^l)L_{t,0} + T_{t,0} \quad (22a)$$

$$C_{t+1,1} = \frac{1 + r_{t+1}(1 - \tau_{t+1}^c)}{\varsigma_{t,0}} \left[k_{t,0} + \tilde{b}_t \tau_t^p w_{t,0} L_{t,0} \right] \quad (22b)$$

and the household first order conditions are

$$U_C(C_{t,0}, L_{t,0}) = \beta_{\varsigma_{t,0}} \mathbb{E}_t \left[\frac{1 + r_{t+1}(1 - \tau_{t+1}^c)}{\varsigma_{t,0}} U_C(C_{t+1,1}, 0) \right] \quad (23a)$$

$$\begin{aligned} -U_L(C_{t,0}, L_{t,0}) = w_{t,0} & \left\{ (1 - \tau_t^p - \tau^l) U_C(C_{t,0}, L_{t,0}) \right. \\ & \left. + \beta_{\varsigma_{t,0}} \tau_t^p \tilde{b}_t \mathbb{E}_t \left[\frac{1 + r_{t+1}(1 - \tau_{t+1}^c)}{\varsigma_{t,0}} U_C(C_{t+1,1}, 0) \right] \right\} \quad (23b) \end{aligned}$$

3. Pension policy with state-dependent pension factor (**PP2**). In this case, the capital tax to be paid in $t+1$ is already fixed in t , and denoted by $\tilde{\tau}_t^c$. Now the pension factor b_{t+1} is state-dependent, it is only determined in $t+1$. Formulas are the same as in (**PP1**), except for the timing of $\tilde{\tau}_t^c$ and b_{t+1} , and are given in Appendix A.3.

Proposition 3. *The three fiscal policy regimes **DP**, **PP1** and **PP2** are equivalent in the sense that every real allocation that can be implemented in one policy regime can also be implemented in any of the other two regimes.*

*The equivalence between **DP** and **PP1** and between **DP** and **PP2** holds true except for a set of allocations that has measure zero, namely when $\tau_t^d = -d_{t,0}/(w_{t,0}L_{t,0})$ and $d_{t,0} \neq 0$. It is certainly true if $\tau_t^d \geq 0$ and $d_{t,0} \geq 0$ for all t and all realizations of shocks.*

Proof. See Appendix A.3. □

The equivalence between **DP** and **PP1** is quite intuitive; we just set

$$\tau_t^d = \tau_t^p(1 - \tilde{b}_t) \quad (24a)$$

$$d_{t,0} = \tau_t^p \tilde{b}_t w_{t,0} L_{t,0}. \quad (24b)$$

The fraction \tilde{b}_t of the pension contribution is the fraction that the household will receive with interest during retirement. It is therefore like a credit from the household to the government, which is expressed in (24b). The fraction $1 - \tilde{b}_t$ is like a tax on labor income, which explains (24a).

Efficiency requires that the effect of a shock be shared also by old households. In **DP** and **PP1**, this is achieved by a state-dependent capital tax. As the equivalence between **PP1** and **PP2** shows, a state-dependent pension factor b_{t+1} can take over the role of a shock absorber instead of the capital income tax. Therefore, the capital income tax of period $t+1$ can already be determined in period t . To have both the pension factor and the capital tax determined only in $t+1$ would be redundant.

The equivalence between **DP** and the two pension regimes can break down for certain combinations of τ_t^d and $d_{t,0}$. From (24) we get $\tau_t^p = \tau_t^d + d_{t,0}/(w_{t,0}L_{t,0})$. Thus it is possible to get $\tau_t^p = 0$ although $d_{t,0} \neq 0$. Then (24b) is not satisfied. Although this happens only under knife-edge conditions, namely if $\tau_t^d = -d_{t,0}/(w_{t,0}L_{t,0})$ and $d_{t,0} \neq 0$, it nevertheless points to the fact that a pension system is not a natural policy regime for all kinds of intergenerational redistributions. It is natural if there is a systematic redistribution from

future to current generations: if $\tau_t^d \geq 0$ and $d_{t,0} \geq 0$ always, then the above problem cannot occur.

The results of this section motivate the policy setup that we are going to investigate quantitatively from Section 4 onward. First, we abstract from government debt. Although debt and a pension system are not completely equivalent in a model with many cohorts and age-independent taxes (Ludwig and Reiter 2008), they are still almost equivalent. Then if we were allowing debt and a pension system simultaneously in the model, the result would be an optimal policy that consists of taking very large offsetting positions in the two instruments. Such a policy is both unrealistic and hard to interpret. It is therefore better to shut down one of the instruments.

Second, we abstract from capital taxes altogether. In fact, with our calibrations, the optimal tax on capital income would be positive in the steady state. We find that it is close to 10 percent. Nevertheless, in the numerical examples below, we don't allow capital taxes for various reasons. We have shown above that capital taxes and the pension factor b_t can both play the role of a shock absorber, in an equivalent way. The purpose of the present paper is to show how the pension system, not the capital tax, optimally reacts to demographic shocks. This is what seems relevant for practical policy purposes. The tax system in many countries is designed to react to demographic developments, while the capital tax appears to be governed by different considerations. Having both instruments active would make it hard to interpret the results. Moreover, if we allow for capital taxes, it is not clear where the tax revenues should go. If they enter the general budget from which pensions are paid, this may create effects in the model that are unrealistic, since pensions in reality are mostly paid by payroll taxes.

Third, we focus exclusively on a Bismarck pension system, where benefits are linearly linked to contributions, and ignore systems of the Beveridge type, where pension payments are independent of past contributions. In the latter case, the full pension contribution rate τ_t^p acts as a distortive tax, such that the first order condition for labor supply equals $U_L(C_{t,0}, L_{t,0}) = w_{t,0} U_C(C_{t,0}, L_{t,0}) (1 - \tau_t^p - \tau^l)$. It is obvious that such a pension system is not equivalent to **DP**, **PP1** or **PP2**, because it introduces an additional labor market distortion. A Beveridge pension system may have a role to play if intra-generational redistribution is important. But in a model such as ours where people of the same cohort are all alike, a Beveridge pension system is pointless.

With no government debt, no capital taxes, but background distortions of the form (10), the golden rule of Proposition 2 does not hold exactly.⁴ We will see in Section 5.1 that it holds approximately.

⁴In the presence of capital taxes, the *modified golden rule*, cf. proposition 2, continues to hold because the inter-temporal household Euler equations do not impose a restriction on the government (Erosa and Gervais 2002). However, in the absence of capital taxes, the modified golden rule generally breaks down because then the first-order condition of the government with respect to capital involves an additional term reflecting the inter-temporal household Euler equations since the interest rate is a function of capital. Furthermore, in our model with background distortions, cf. equation (10), the modified golden rule breaks down if $\tau^l \neq 0$. The reason is that wages entering in (10) are a function of capital which is also reflected in the first-order condition of the government with respect to capital.

4 Pension Policy in a Calibrated OLG Economy

We now present a calibrated version of the model of Section 2. In particular, we describe the institutional details of a PAYG pension system, which is the instrument by which the government shares demographic risk between generations.

4.1 Age Limits

A model period stands in for 5 actual years. Accordingly, we set $I = 20$ reflecting an actual age limit of 100 and $i^R = 14$ for an actual retirement age of 65. Adulthood in our model starts at $i^A = 5$ for an actual age of 20. Hence, there are four periods of inactive childhood.

4.2 The Pension System

Pension income, $p_{t,i}$, is the product of two components, a time-specific pension factor common to all households, which we denote by b_t and an individual pension stock, $P_{t,i}$, which is accumulated during the working period:

$$p_{t,i} = b_t P_{t,i} \quad \text{for } i \geq i^R. \quad (25a)$$

The government chooses in each period the pension benefit factor b_t . The rules on how the pension stock $P_{t,i}$ accumulates over the life-cycle are fixed and given by

$$P_{t,i} = \begin{cases} (1 + r_{t,i}^n) \left(P_{t-1,i-1} + \frac{w_{t,i}}{w_t} L_{t,i} \right) & \text{for } 0 < i < i^R \\ P_{t-1,i-1} & \text{for } i \geq i^R \\ 0 & \text{for } i = 0. \end{cases} \quad (25b)$$

Labor supply $L_{t,i}$ enters linearly into the pension formula, and it is weighted by the relative wage $w_{t,i}/w_t$.⁵ In other words, one “earnings point” is credited if the individual wage earnings equal the average wage. After retirement, the pension stock is held constant, such that pension income only grows with b_t . The pension stock accumulation formula in (25b) is a mix between the German earnings point system and the Swedish system of notional defined contributions.

The Ramsey planner takes individuals’ optimizing behavior as given and chooses its fiscal instruments so as to maximize the objective function (13). In particular, it chooses in each period an optimal pension factor b_t . In addition to this second best policy, we also

⁵In contrast to the specification in subsection 3.3, contributions to the pension stock do no longer depend on the pension contribution rate τ_t^p . This is important for the stability of the system, for the following reason. If the contribution rate enters into the pension stock formula, a shock in period t forces an increase of the contribution rate τ_t^p , such that pension claims in future periods automatically increase. This in turn requires an increase of future contribution rates which tends to destabilize the system if the pension factor is not chosen optimally but follows one of the simple rules specified below.

consider simple rules that are supposed to be a representation, if stylized, of actual pension systems. Concretely, we assume for b_t that

$$b_t = \text{const} \cdot a_t \cdot \left(\frac{\sum_{i=i^R}^I N_{t,i}}{\sum_{i=i^A}^{i^R-1} N_{t,i}} \right)^{-\alpha_b}. \quad (26)$$

The term in parenthesis is the old-age dependency ratio. With $\alpha_b > 0$, pension payments are reduced if there are more retirees compared to the working population. As shown in Appendix A.1, the case $\alpha_b = 0.25$ is a close approximation to the current legislation in Germany, after the ‘‘Rürup-reform’’ implemented in spring 2004.

The factor a_t indexes the increase in the aggregate pension factor b_t to the growth of wages. We consider two alternative indexation schemes. We take as a benchmark that pensions are indexed to the growth of net wages, which is the practice in Germany since 1992. Net wage adjustment (NWA) implies automatic stabilization of the pension system’s budget because increases in contribution rates simultaneously lead to decreases of benefit levels. In the second scheme (GWA), indexation is to the growth of gross wages, as it was done in Germany before 1992. We therefore have

$$a_t = \begin{cases} w_t(1 - \tau^l - \tau_t^p) & \text{NWA: net wage adjustment} \\ w_t & \text{GWA: gross wage adjustment.} \end{cases} \quad (27)$$

An important special case in (26) is $\alpha_b = 0$. The pension factor b_t then varies only with net or gross wages. We refer to this rule as the *constant benefit rule* where contribution rates τ_t^p endogenously adjust to balance the budget. Notice that this definition of a constant benefit rule is not equivalent to what is called a defined benefit plan in the literature. The other extreme is the *constant contribution system* where $\tau_t^p = \bar{\tau}^p$. Here it is the pension benefit factor b_t that adjusts to balance the budget.

Of course we assume that households perfectly understand the pension rules. For the labor supply decision, the household therefore takes into account that the pension stock gets increased by working more and thus contributing to the pension system. The first order condition can then be written as

$$-U_L(C_{t,i}, L_{t,i}) = w_{t,i} \left[(1 - \tau_t^p - \tau^l) U_C(C_{t,i}, L_{t,i}) + \frac{1}{w_t} V_{t,i} \right] \quad (28)$$

where the value of the pension stock, $V_{t,i}$, follows the recursive equation

$$V_{t,i} = \begin{cases} 0 & \text{for } i = I + 1 \\ b_t U_C(C_{t,i}, 0) + \beta \varsigma_i \mathbb{E}_t V_{t+1, i+1} & \text{for } i^R \leq i \leq I \\ \beta \varsigma_i (1 + r_{t,i}^n) \mathbb{E}_t V_{t+1, i+1} & \text{for } i < i^R. \end{cases} \quad (29)$$

4.3 The Fertility Process

The exogenous driving force in our model is the total fertility rate, f_t , which shifts the entire distribution of age-specific fertility rates:

$$f_{t,i} = \frac{\bar{f}_i}{\sum_{i=i^A}^{i^F} \bar{f}_i} \cdot f_t, \quad i = i^A, \dots, i^F$$

We assume that the shape of the fertility distribution over the life cycle is determined by the constants \bar{f}_i and is not affected by shocks. In our model, fertility is zero for $i < i^A = 4$, where i^A denotes the age when model households become economically active and corresponds to 20 years. Fertility equals zero also for $i > i^F = 10$, which corresponds to an age of 50 years. We specify the log of the total fertility rate as an AR(1) process:

$$\ln(f_t) = c + \rho^f \ln(f_{t-1}) + \epsilon_t^f, \quad (30)$$

where c is a constant, $0 < \rho^f < 1$ is the autocorrelation coefficient and ϵ_t^f is a normal i.i.d. shock, $\epsilon_t^f \sim \mathcal{N}(0, \sigma_f^2)$.

We calibrate (30) using German data on age-specific fertility rates as taken from HMD (2008). Before 1950, fluctuations in fertility are dominated by the effects of two world wars and the great depression. After the second world war, we observe a regime of high fertility (until the late 1960s) and a regime of low fertility afterwards, the well-known baby-boom-bust cycle. The demographics before 1950 are probably not relevant for the current situation, and history since 1950 is rather short for obtaining a reliable estimate of the fertility process. We choose ρ^f so that it takes 20 periods (100 years) for the process to return halfway to its unconditional mean. This gives a ρ^f of $0.9659^{0.2} = 0.993$ annually, close to the point estimate of 0.991 for the sample period 1950 – 2002 (the 95 percent confidence interval is [0.94, 1.03]). We set the standard deviation of the fertility shock to $\sigma_f = 0.0525$ per period (the point estimate for the period 1950-2002 is 0.037 annually). This implies that the difference between a total fertility rate of 2.1, which is the replacement fertility, and 1.4, which is roughly the mean for the years 1975-2002, equals two standard deviations of the total fertility rate.

In our quantitative analysis we then consider two alternative scenarios for the steady state demographics. In both cases the steady state total fertility rate is assumed to be at a replacement level of roughly 2.1. The first scenario uses the mortality rates measured by HMD (2008) for the years 1956–1970. This yields an old-age dependency ratio (OADR), defined as the ratio of the old-age population (65+) to the working-age population (20–64), of about 27%. The second case uses estimated mortality rates of recent years (1998–2002) and gives an old-age dependency ratio of about 40%.⁶ We refer to these two scenarios as “low OADR” and “high OADR”, respectively.

⁶Since we focus on stationary populations, the OADR in our model is higher than the OADR reported in actual German data. Both, the low and high OADR steady state of our model reflect steady states that were not yet reached in the respective periods.

4.4 Parameters of the economic Model

Household utility is of the CRRA type:

$$U(C, l) \begin{cases} [C^{1-\theta}(1-l)^{\eta(1-\theta)} - 1]/(1-\theta) & \text{for } \theta \neq 1 \\ \ln C + \eta \ln(1-l) & \text{for } \theta = 1. \end{cases} \quad (31)$$

Since we stress the importance of labor supply distortions, a crucial parameter in the model is the leisure share, η , which also measures the Frisch elasticity of labor supply. There is no consensus yet on what this elasticity is in reality. The early empirical literature on labor supply elasticities reports low values, around 0.2 (Pencavel 1987). This is due to the fact that those studies have typically focused on the hours decision of prime age men (intensive margin), which is now known to be the most inelastic part of labor supply. The more recent literature not only examines hours worked, but also the participation decision, of both men and women (extensive margin). Haefke and Reiter (2006) find that a value of 0.6 is compatible with both micro and macro evidence. Proponents of real business cycle analysis typically use even higher values of η , around 1.5, cf., e.g., Prescott (2004). We therefore choose $\eta = 0.6$ as a benchmark, but also report results for the cases $\eta = 0.2$ and $\eta = 1.5$. In our benchmark calibration, we set the coefficient of relative risk aversion to $\theta = 1$ (log utility). In the sensitivity analysis we also consider the case $\theta = 2$.

We use the CES production function

$$F(K_t, Z_t L_t) = \begin{cases} [\alpha K_t^{-\xi} + (1-\alpha)(Z_t L_t)^{-\xi}]^{-1/\xi} & \text{for } \xi \neq 0 \\ K_t^\alpha (Z_t L_t)^{1-\alpha} & \text{for } \xi = 0, \end{cases} \quad (32)$$

where the elasticity of substitution between capital and labor is given by $1/(1+\xi)$. Our benchmark specification is the Cobb-Douglas production function, $\xi = 0$. We also consider a substitution elasticity equal to 4, in which case factor prices react less in response to a demographic shock. Throughout, we set the production elasticity of capital such that the capital share in gross output equals $1/3$.

Our target value of the gross investment to GDP ratio is 0.235, which is the average of 1960-1990. Since the capital output ratio slightly trends upwards during this time period, we choose the value at the end of the period (1990) which equals 3. With these values and under the assumption of a constant population and an annual productivity growth of $g = 0.015$, we obtain a depreciation rate of $\delta = 0.345$ over 5 years.

In the steady state that represents the German economy after 1990, we calibrate the background distortion such that the labor tax burden equals 0.57. This is the value of the “tax wedge” for the year 2000 in the OECD institutional database, updated by Baker et al. (2003). We interpret this number as including all the social security contributions. In our model, the actual labor supply distortion is lower, because only part of the pension contributions are effectively a tax. To match the labor tax burden of 0.57 we set background distortions to $\tau^l = 0.3497$. For 1960, the database reports a wedge of 0.42. Our “low OADR” calibration then targets this labor tax rate, which requires a background distortion of $\tau^l = 0.2892$.

The households' discount factor β is chosen such that, in steady state, the capital-output ratio matches the target of 3. The discount factor of the government, ω , is chosen to match the size of the pension system. More precisely, the optimal pension size, measured as old-age cash benefits to GDP, matches the target 0.087, which we take from the OECD's Social Security Database as the average during the 90th. The environment for which we do this calibration should reflect the German economy before 1970, with a labor tax rate of 42 percent. We then obtain a private discount rate of 0.9893 annually, and a government discount rate of 0.9774 annually. This difference implies, for example, that the marginal utility of a currently 20 year old household counts only 54.4 percent of the marginal utility of a currently 70 year old. This difference is big, but not implausible.

Re-calibration in experiments

We report results for a variety of parameter combinations. When we change a basic parameter of the model, which does not reflect a change in policy, or a change in the environment over time, we recalibrate β , ω and α so as to match the capital output ratio, labor share and the size of the pension system under the low OADR scenario. This is the case when we use a CES rather than a Cobb-Douglas production function, or risk aversion $\theta = 2$ rather than log utility.

An exception is the labor supply elasticity. When changing η , we keep β , ω and α unchanged, because we are interested in how the labor supply elasticity affects the optimal size of the pension system for fixed government preferences and technology. Similarly, for the comparison of the "low OADR" to the "high OADR" population scenario, and for comparing different pension systems, we do not recalibrate β , ω or α .

4.5 Solution Method

We model the government problem as a Ramsey problem under full commitment. To solve it, we adopt the recursive contracts approach of Marcet and Marimon (1998). In this approach, the state vector is augmented by the Lagrange multipliers of the dynamic households' first order conditions, which carry the information about the planner's past commitments. In period 0, where the Ramsey planner starts the optimization without being bound by any past commitments, the Lagrange multipliers start at the value 0.

Since there are many cohorts, our dynamic system is very large (it has about 200 variables, out of which about 100 are state variables). We can therefore only solve for a linear approximation around the deterministic steady state. Note that the steady state is unique because we deal with OLG households (Erosa and Gervais 2002), contrary to many optimal fiscal policy problems with infinitely-lived households. To compute the linearized model, we use a symbolic differentiation tool to differentiate the government's Lagrangian with respect to all the variables, and output the result in Matlab syntax. The linearized system is then solved using the Sims (2001) package. There is a unique stable solution of the linearized model.

For impulse responses and simulations, we start the system at the steady state, including the steady state values of the Lagrange multipliers. This type of solution to the policy problem is now often called “policy from a timeless perspective” in the literature on optimal monetary policy (cf. for example Woodford (2003)).

5 Numerical Results

5.1 Steady State Analysis

In the steady state analysis, we focus on how the existence of a pension system affects economic variables such as the capital stock, the interest rate etc. We also investigate how changes in parameters or the demographic environment affect the size of the optimal pension system, keeping government preferences (ω) fixed.

The second column of Table 1 refers to our calibration for the “low OADR” scenario with a Frisch labor supply elasticity of 0.6. If the government is neutral and does not operate a pension system, the labor tax rate is at 28.92 percent. With an impatient government, total pension payments are 8.72 percent of GDP, and the labor tax rate, including social security contributions, has to rise to 42 percent. The tax wedge however, which measures the distortionary aspect of labor taxes (cf. Appendix A.2), is only 34.42 percent, which is roughly in the middle of 28 and 42 percent. One can therefore say that about half of the pension contributions are seen as tax, and half as an investment for retirement income. The distortion results from the fact that an investment in the capital market yields a return of 3.9 percent, whereas the internal rate of return of the pension system is only 1.5 percent (it equals the productivity growth rate, since population growth is zero in steady state).

The inter-generational redistribution through the pension system leads to a substantial crowding out of capital so that the steady state capital stock is about 14 percent lower than in the economy without pensions. Consequently, the annual real interest rate is 3.9 percent rather than 3.11 percent. Output is lower by 6 percent. The negative effects of this crowding out on utility in the steady state is clearly higher than any gain from risk sharing that we report in Section 5.6 below. Given the results of Krueger and Kubler (2006), this is not surprising. However, we do not regard this finding as an argument against the pension system. In our setup, the lower utility of future generations is just the price to pay for the welfare gains of earlier generations.

The first and third columns of the table show results for different values of the labor supply elasticity η . Higher η leads to a larger pension system in steady state, for the following reason. The government tries to implement the modified golden rule, at least approximately (cf. Section 3.3). If labor input varies elastically, a larger reduction of capital is necessary to achieve a certain reduction in the capital-labor ratio and therefore the interest rate. This in turn requires more intergenerational redistribution.

The bottom part of the table provides information for different levels of the background distortion. The main finding is that the optimal size of the pension system increases with the background distortion. This is similar to the effect of higher labor supply elasticity: if

the labor supply distortion is already high, labor reacts more strongly to a further increase of the distortion, such that again a stronger reduction in capital is necessary to implement the golden rule.

Columns 4-6 of the table show results for the “high OADR” scenario. In this case, the optimal size of the pension system is substantially higher: pension payments now amount to almost 15 percent of GDP, and the labor tax rate is increased to roughly 51 percent. To understand this result, notice that we hold the weight of each household in each cohort constant across experiments. With more retirees, this implies that the total weight of the old cohorts in the welfare function increases. Higher size of the pension system is therefore just what the ex-ante optimality criterion of our Utilitarian welfare function requires. One would expect the same result in a model of political equilibrium, where their larger size increases the political power of the old cohorts, cf. Gonzalez-Eiras and Niepelt (2005).

5.2 Optimal Risk-Sharing in the Undistorted Economy

In this section, we analyze the optimal policy response to a negative shock in the fertility rate when the government is neutral (cf. Definition 2) and when there is no background distortion. In the deterministic steady state, tax rates are zero and the government discount rate is calibrated such that there is no pension system. The government only becomes active in response to shocks, to spread the effects of this shock efficiently between generations.

We consider the case where the OADR is low. The exogenous time paths of the fertility rate and the corresponding OADR are displayed in the top panels of figure 1. The size of the shock is one standard deviation of the fertility rate, $\sigma_f = 0.0525$, cf. Section 4.3. Twenty years after the shock, the smaller cohorts enter the labor force so that the OADR starts increasing. 65 years after the shock, when the first small cohort retires, the OADR reaches its peak. Afterwards it reverts back to the long-run steady state level.

In Figure 2, we compare the response of three different policy regimes to this demographic shock. As we analyze deviations from a non-distorted steady state, all the policies fluctuate around the same deterministic steady state. The three policy regimes are the first best allocation (“first best”), the second best pension system (“optimal b”), and the allocation without government intervention (“no gov.”). The four variables displayed in the four different panels of the figure are the labor tax rate, the lifetime welfare gain of each cohort, the labor tax wedge, and the real interest rate. The tax wedge is obviously zero in the no-government and in the first best case. The welfare gain of a cohort is measured as the change in utility caused by the shock, compared to steady state utility. The utility change is expressed as a consumption equivalent variation, that is, the necessary percentage change of consumption to make the household indifferent. For households already alive at the time of the shock, we express the welfare change as the necessary change in consumption for all periods after the shock has occurred.

We see in Figure 2 that the second best allocation is close to the first-best allocation. This is so because there is no government activity in the steady state. The labor tax fluctuates around zero, so that the distortionary aspect of the second-best solution is less important.

The most interesting result is that the first best response is very different from the outcome without government intervention. The welfare results depicted in Panel (b) of Figure 2 are in line with the results of Bohn (2001): generations born after the negative fertility shock who belong to small cohorts benefit in the economy without government, while generations already alive loose. This is the consequence of the expected factor price movements: young cohorts enjoy high wages while working and old cohorts face lower interest rates which diminishes the value of their retirement savings. The first best policy reverses this pattern of gains and losses: by increasing the pension contribution rates after the shock, the social planner increases the lifetime welfare of the generations already alive, and decreases the utility of the future small cohorts. This may come as a surprise: should not the optimal insurance policy consist in spreading the burden of the demographic shock across different generations? We would expect optimal policy to level out the welfare gains across generations, but not to reverse the pattern of gains and losses.

Figure 3 helps to explain this apparent puzzle. It shows the consumption profile of different cohorts, measured in percentage deviations from the steady state. Each line refers to a different cohort, starting and ending in different time periods. The top panel displays the situation in the absence of government. About 20 years after the shock, consumption profiles start sloping downward, because the real interest rate is then below its steady state value. Since a fertility shock is felt in the labor market only after 20 years, the generation born 50 years before the shock is hardly affected by the change. For the generation born 25 years after the shock the changes in factor prices are strongly beneficial, because gains from higher consumption today exceed losses from lower consumption in the future. Afterwards the beneficial effect declines.

What happens in the first best solution? At a given point in time, the proportional consumption deviation is identical for all cohorts as shown in the middle panel of Figure 3. This is what perfect risk sharing implies for a log utility function, cf. Equ. (14c). However, optimal consumption fluctuates over time, as a consequence of the change in the real interest rate. In particular, the fall in the interest rate about 20 years after the shock causes a steep decline in optimal consumption. Cohorts participate in this decline more or less depending on when they live. This illustrates that the concept of ex-ante efficiency requires a redistribution between cohorts depending on the path of the interest rate. This can also be seen directly from the first-order conditions of the social planner in (14). Combining (14b) and (14c) we get

$$U_C(C_{t,i}, L_{t,i}) = \mathbb{E}_t \left[(1 + F_K(K_t, Z_{t+1}L_{t+1}) - \delta) \frac{\Omega_{t-i+1}}{\Omega_{t-i}} U_C(C_{t+1,i}, L_{t+1,i}) \right].$$

5.3 Responses of the Impatient Government to Fertility Shocks

We now analyze the optimal policy of an impatient government, which runs a pension system even in the deterministic steady state. That means, we calibrate the government's discount factor such as to match the actual size of the pension system. We here contrast two scenarios, one with low and one with high overall distortions.

With an impatient government, the first best allocation has a different steady state than the allocation under a second best policy. Since we compare different policy regimes according to how they adjust to the demographic processes outside the steady state, we want the steady state itself to be identical for all these regimes. Therefore, we now only look at second and third best policies (simple rules). We compare four different policy schemes: the second best pension policy (“optimal b ”), and three simple rules: a pension system with constant pension factor (“constant b ”), with a constant contribution rate (“constant τ ”), and the pension system where the benefit rate varies inversely with the OADR as in Germany since 2004 (“ b adj. OADR”).

5.3.1 Environment With Low Distortions

We start the analysis by looking at an environment with very low labor supply distortions. We set the labor supply elasticity to $\eta = 0.2$, assume there is no background distortion ($\tau^l = 0$), and analyze the “low OADR” scenario, reflecting the German economy before 1975. Figure 4 displays the corresponding impulse responses to a negative fertility shock.

In Panel (a) of Figure 4 we observe an optimal policy response that is again consistent with the analysis of Bohn (2001): about 20 years after the shock to fertility, when the smaller cohorts enter the labor market, the optimal response is to *increase* the pension payments, as measured by the pension factor b . With constant b , the contribution rate – depicted in Panel (c) of the figure – has to increase, because of the increase in the OADR, but it does not increase as much as in the second best response. Keeping the contribution rate constant leads to a decrease in the pension factor, exactly the opposite of what is called for according to the optimal policy. The adjustment of the pension factor as implied by the recent German reform is somewhere in between and closer to the constant b scenario.

The two graphs in Panels (b) and (d) of Figure 4 shed some new light on this. In the distorted economy, the persistent decrease in fertility leaves all cohorts losing. In relative terms, however, keeping the contribution rate constant favors future generations, again because of favourable factor price movements. For the reasons that we discussed in Section 5.2 (expected low interest rates in the decades ahead), the optimal policy favors currently living generations.

5.3.2 Environment With High Distortions

Let us now turn to what we take as a realistic description of the current and future economic environment in Germany and some other continental European countries. The labor supply elasticity equals $\eta = 0.6$, the old-age dependency is high and the general tax system is already creating a high level of labor supply distortion, $\tau^l = 0.345$. The exogenous time paths of the fertility rate and the OADR are displayed in the middle panels of figure 1. Figure 5 shows the impulse responses to a negative fertility shock in this setup. The optimal pattern of fiscal policy now differs from Figure 4 in that the pension factor is *reduced* at the time when the OADR peaks, about 65 years after the shock. Observe that the optimal policy now moves closer to the German “ b adj. OADR” rule when compared

to the low-distortion environment analyzed above. Key for these findings is the existing labor supply distortion. The optimal policy implies an increase in labor tax rates once the small cohorts enter the labor market. The efficiency cost of the increase in the tax wedge is higher if labor supply is more elastic, or if the background distortion is higher. This is because the excess burden is a convex function of tax rates. Once we get close to the maximum of the Laffer curve, increasing contribution rates further yields very little extra revenue. Then it becomes more important to smooth tax rates.

An interesting feature of Figure 5, which is also present in Figure 4, is that the simple adjustment rules (“constant b” and “b adj. OADR”) seem to be more “egalitarian” than the Ramsey planner solution, in the sense that they redistribute the burden more equally across cohorts. They cause smaller variations in lifetime utility than the Ramsey solution. We will document and discuss this issue more systematically in Section 5.6.

5.3.3 The Role of Relative Price Reactions

So far we have concentrated on a Cobb-Douglas production function with an implied substitution elasticity of one, where changes in relative factor supplies trigger relatively strong reactions of wages and interest rates. We next analyze the case of a CES production function, calibrated with a substitution elasticity of 4 so that fertility shocks generate much smaller fluctuations of relative factor prices. We interpret this as a kind of substitute for modeling an open economy, where factor prices are determined to a large extent on global, not domestic markets. In this setup, the fluctuations in interest rates are smaller, and therefore there is less reason to favor current generations after a negative shock to fertility. The results are shown in Figure 6. Now the optimal policy almost perfectly tracks the German “b adj. OADR” rule until roughly 65 years after the shock, which is when the OADR peaks. This underlines again the importance of interest rate movements in shaping optimal policy. When those movements are smaller, a policy that increases the contribution rate of small cohorts becomes less attractive.

5.4 Distance from Optimal Policy

Following up the analysis of the impulse response functions, Table 2 documents, for a wide range of parameter combinations, how far the third-best policies are away from the Ramsey solution. We measure distance in policy space by the root mean squared deviations of impulse responses of the labor tax rate to different shocks, discounted by the discount rate of the social planner:

$$\mathcal{D}(\mathcal{P}) = \sqrt{\frac{1}{1-\omega} \sum_{t=0}^{\infty} \omega^t (IR_t^{\mathcal{P}} - IR_t^{SecondBest})^2}. \quad (33)$$

Here, $\mathcal{D}(\mathcal{P})$ is our distance measure for the simple policy \mathcal{P} , and $IR_t^{\mathcal{P}}$ is the impulse response of \mathcal{P} at time t to the fertility shock.

From Table 2 we observe the following patterns. The higher the labor supply distortion is (higher η , higher τ^l , high OADR),

- the better it is to use net-wage rather than gross-wage adjustment
- the better is the relative performance of the German benefit adjustment rate, although the constant-benefit rule still looks better in all cases
- the better is the performance of the constant-contribution rule, although it is furthest away from the optimal policy in all cases.

Furthermore, under the CES production function that reduces factor price fluctuations, the German rule comes closer to the optimal policy. Finally, using a higher coefficient of relative risk aversion of $\theta = 2$ moves the second best policies closer to the optimal policy in all cases.

5.5 The Welfare Loss From Fluctuations

So far the comparison between the Ramsey solution and the simple-rule pension systems was done in policy space. Next we want to see whether the optimal policy brings significant welfare gains over the simple rules. The solution technique that we use – log-linear approximation about the deterministic steady state – allows us to obtain a correct second-order approximation to the welfare function around the steady state.⁷ Notice that the concept of the steady state of a Ramsey solution also implies that the Lagrange multipliers (co-states) take their steady state values. This means that the Ramsey planner respects the commitments undertaken in earlier periods.

Table 3 displays the welfare measure for a variety of model variants. More precisely, it measures the value function of the stochastic planner’s problem at the steady state, in deviation from the value function of the deterministic model. All numbers are expressed as percentage equivalent variation of permanent consumption. Positive numbers mean welfare gains, negative ones welfare losses, always compared to the deterministic steady state.

It may appear puzzling that fluctuations in fertility cause “welfare gains”, that is, an increase in the objective function (13) compared to the deterministic steady state.⁸ However, this comparison is not very meaningful, because the steady state and the case with demographic fluctuations involve different sets of people. What is relevant in the table is the difference in the value function across policies, for a given process of fertility. The most obvious conclusion from the table is that the welfare differences between second and third best policies are very small. A typical difference is in the range of 0.01 percent of consumption. Being familiar with Lucas (1987)-type calculations, this is what one would

⁷This is shown by Debortoli and Nunes (2006), generalizing an idea of Benigno and Woodford (2006).

⁸The reason is that higher fertility increases the total number of households, and at the same time increases the utility of the typical household. Number of households and utility are therefore positively correlated, and since the objective function is essentially the number of people multiplied with their utility, this induces a natural convexity of the objective function w.r.t. fertility.

expect given that consumption fluctuations are typically in the range of a few percent. It also reflects the fact that all the considered policy rules are quite reasonable.

Otherwise, the comparison of policies in value space comes basically to the same conclusions as the comparison in policy space. When labor market distortions increase, net wage adjustment performs relatively better than gross wage adjustment, and the constant contribution rate becomes somewhat more attractive. Still, among all the policies considered, the constant contribution policy performs worst in terms of ex-ante efficiency.

5.6 The Variability of Lifetime Utility

So far we have evaluated outcomes by the criterion of ex-ante efficiency, which implies the maximization of a weighted sum of ex-ante expected utilities as in (13). An alternative concept of intergenerational risk-sharing might be to share the burden of demographic shocks equally among cohorts. This would require to minimize the fluctuations of lifetime utility around the steady state. As a simple measure of intergenerational equity we therefore use the unconditional variance of lifetime utility.⁹ We have already seen in Section 5.2 that ex-ante efficiency leads to a very different ranking of alternative policies than does intergenerational equity. Efficiency requires that the intergenerational allocation of resources responds to variations in the real interest rate. For example, if expected future interest rates are low, it is relatively cheap to give more consumption to current generations, and then those generations should receive higher lifetime welfare. Equity requires to give different generations the same or similar utility, irrespective of the interest rate.

The impulse responses presented in Section 5.3 suggest that the simple rules implemented in reality lead to smaller variations in lifetime welfare than does the Ramsey policy, and in this sense are more equitable. Table 4 confirms this. It presents the unconditional variance of lifetime utility under a variety of policy rules. The policy with a constant contribution rate is the one that consistently performs best under this criterion. Furthermore, the German rule of adjusting benefits consistently reduces the variability of lifetime welfare compared to the constant benefit rule. If combined with net-wage adjustment, it always yields lower variability than the Ramsey policy.

We cannot decide here what the right objective function for policymakers is. But we think it is important to become aware of the implications of those different conceptions of optimal insurance. And we find it remarkable that real-world policy rules seem to perform well under the criterion of intergenerational equity. This criterion provides support for the German pension reform of 2004, and for the idea of keeping contribution rates constant, which also plays an important role in policy debates.

5.7 A Baby-Boom-Bust Cycle

We finally investigate the implications of our model for the recent baby boom/baby bust cycle as observed in many industrialized countries. To this end we feed into the model

⁹Auerbach and Lee (2008) use a more sophisticated concept of intergenerational equity where more weight is given to the difference in utility of adjacent generations.

the actual historical observations of German fertility rates until 1990. From then onwards, the fertility rate evolves according to the process specified in Section 4.3, that is, we again assume that the fertility rate is reverting back to the replacement level in the long-run. The time path of the fertility rate in this scenario and the corresponding old-age dependency ratio – for the high OADR steady state – are displayed in the bottom panel of figure 1.

Figure 7 shows the simulation results. The German system, which adjusts the pension payments to the OADR, is remarkably close to the optimal policy throughout a long period (until about 2040 when the OADR peaks). Welfare consequences of the simple rules “contant b” and “b adj. OADR” are close to the optimal policy. Under the German rule the utility loss compared to the steady state amounts to about 5 percent of lifetime resources for cohorts born around 2005. This is not a catastrophic, but a sizeable loss.

Obviously, these calculations are based on long-run projections that are subject to a high degree of uncertainty. We feel that the assumptions underlying this scenario are rather pessimistic in a number of ways. First, we assume that the fertility rate stays very low for a very long time. Second, we do not allow for immigration which could partially make up for the lack of children. And third, we rule out the possibility that a thorough supply side reform, with a reduction of the unemployment rate and the corresponding welfare benefits, could be used to lower the labor supply distortion of the general tax system. The lower the background distortion in the economy, the lower the welfare losses that arise from the baby bust.

6 Conclusions

How should the pension system react to a drop in the fertility rate? The intuitive response is that pension benefits should be reduced when the old-age dependency ratio is high, such that the contribution rate of the small working cohorts need not increase by too much. This idea was formalized in the German pension reform of 2004 and written into the law. Bohn (2001) comes to the opposite conclusion: he finds that ex-ante efficiency requires to raise benefit rates for the big cohorts, financed by a strong increase of contribution rates for the small cohorts. In Bohn’s analysis, the pension system is non-distortive, and the demographic change causes strong factor price reactions. What we find is decisive in generating this result is that the expected decrease in interest rate due to the shrinking work force implies an optimal path of aggregate consumption that is falling over time, which favors current over future cohorts.

We identify three lines of argument that work against Bohn’s conclusion. First, increasing the contribution rates for small cohorts becomes costly if labor supply distortions are taken into account. This includes the correct modeling of the distortionary character of the pension system, the right labor supply elasticity, and accounting for the background distortion that comes from the general tax system. Second, the reason to favor the currently old generations becomes weaker if the interest rate reacts less strongly to the demographic change than is implied by a Cobb-Douglas production function. This could be because the elasticity of substitution between labor and capital is higher than one, or, more impor-

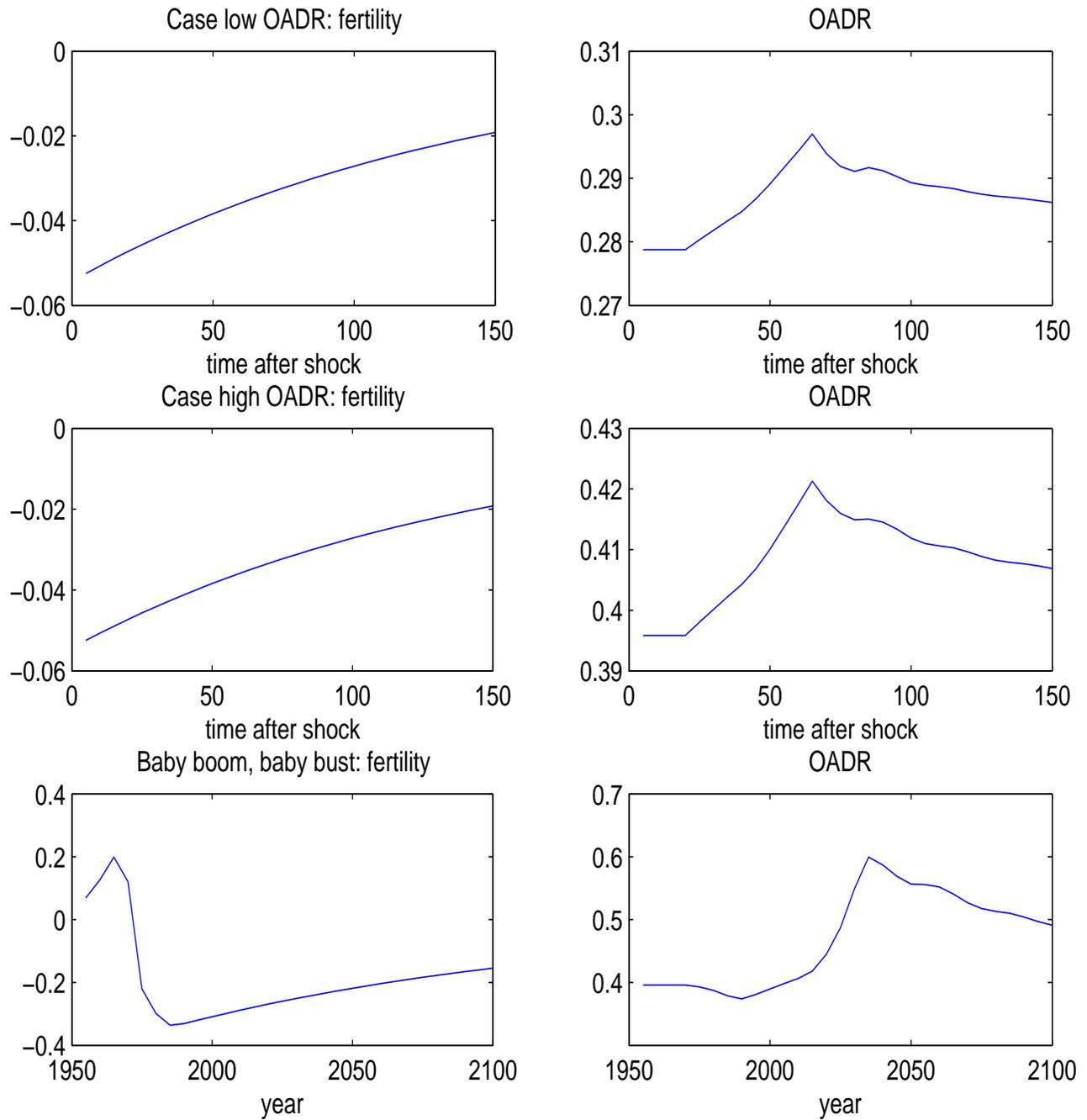
tantly, because the interest rate is dominated by global developments rather than domestic demographics. Third, politics might reject altogether the idea that resources should be redistributed between generations in response to changing interest rates. This would mean a deviation from the objective of ex-ante efficiency. A politically more attractive alternative might be to say that the pension system should spread the costs of demographic change equally between generations, so as to minimize the variability of lifetime utility. According to this criterion of intergenerational equity, the German pension reform is a clear improvement, and a policy of keeping the pension contribution rate constant would be even better. These results call for a thorough reflection on what is the right objective of pension policy, both from a positive and a normative viewpoint.

References

- Atkinson, A. B. and A. Sandmo (1980). Welfare implications of the taxation of savings. *The Economic Journal* 90, 529–549.
- Auerbach, A. J. and R. Lee (2008). Welfare and generational equity in sustainable unfunded pension systems. Manuscript.
- Baker, D., A. Glyn, D. Howell, and J. Schmitt (2003). Labor market institutions and unemployment: a critical assessment of the cross-country evidence. Center for European Studies Working Paper 98.
- Benigno, P. and M. Woodford (2006). Optimal taxation in an RBC model: A linear-quadratic approach. *Journal of Economic Dynamics and Control* 30(9-10), 1445–1489.
- Bernheim, B. D. (1989). Intergenerational altruism, dynastic equilibria and social welfare. *Review of Economic Studies* 56 (1), 119–128.
- Bohn, H. (2001). Social security and demographic uncertainty: The risk sharing properties of alternative policies. In J. Campbell and M. F. Feldstein (Eds.), *Risk Aspects of Investment Based Social Security Reform*, pp. 203–241. University of Chicago Press.
- Caplin, A. and J. Leahy (2004). The social discount rate. *Journal of Political Economy* 112 (6), 1257–1268.
- Conesa, J. C. and D. Krueger (1999). Social security reform with heterogeneous agents. *Review of Economic Dynamics* 2, 757–795.
- Debortoli, D. and R. Nunes (2006). On linear-quadratic approximations. Manuscript, Universitat Pompeu Fabra.
- Demange, G. (2002). On optimality in intergenerational risk sharing. *Economic Theory* 20, 1–27.
- Erosa, A. and M. Gervais (2002). Optimal taxation in life-cycle economies. *Journal of Economic Theory* 105(2), 338–69.
- Farhi, E. and I. Werning (2006). Inequality and social discounting. Working Paper, MIT.
- Farhi, E. and I. Werning (2008). Progressive estate taxation. Working Paper, MIT.
- Golosov, M., L. E. Jones, and M. Tertilt (2007). Efficiency with endogenous population growth. *Econometrica* 75(4), 1039–1072.
- Gonzalez-Eiras, M. and D. Niepelt (2005). Sustaining social security. CESIFO Working Paper No. 1494.
- Gordon, R. H. and H. R. Varian (1988). Intergenerational risk sharing. *Journal of Public Economics* 37(2), 185–202.
- Gottardi, P. and F. Kubler (2008). Social security and risk sharing. mimeo.
- Haefke, C. and M. Reiter (2006). Endogenous labor market participation and the business cycle. IZA DP 2029.

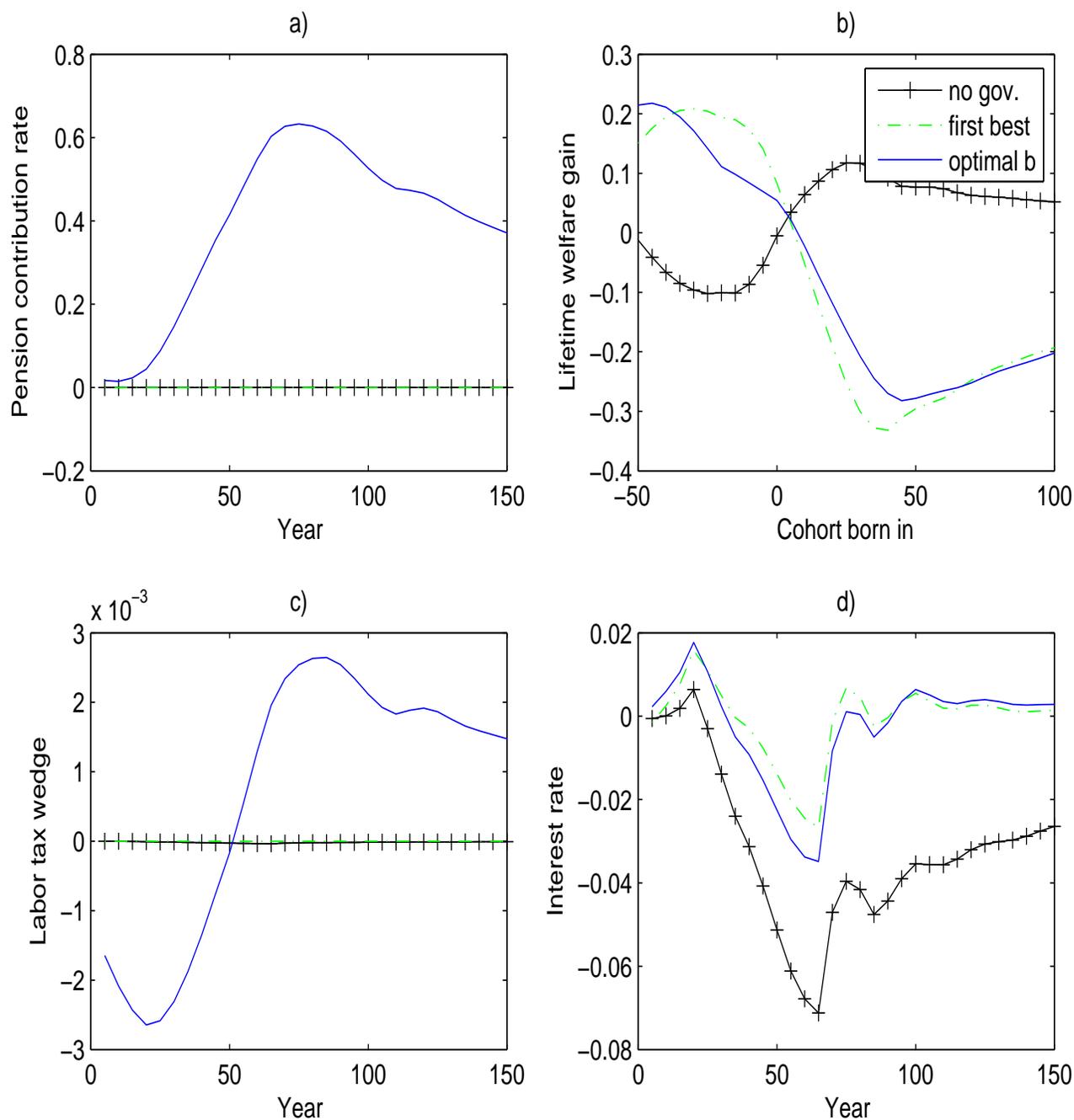
- HMD (2008). The human mortality database. <http://www.mortality.org/>.
- Krueger, D. and F. Kubler (2006). Pareto improving social security reform when financial markets are incomplete? *American Economic Review* 96(3), 737–755.
- Krüger, D. and A. Ludwig (2007). On the consequences of demographic change for rates of return to capital and the distribution of wealth and welfare. *Journal of Monetary Economics* 54(1), 49–87.
- Lucas, R. E. (1987). *Models of Business Cycles*. Oxford: Basil Blackwell.
- Ludwig, A. and M. Reiter (2008). On the equivalence of pensions and debt. Manuscript.
- Marcet, A. and R. Marimon (1998). Recursive contracts. Manuscript.
- Pencavel, J. (1987). Labor supply of men: A survey. In O. A. . R. Layard (Ed.), *Handbook of Labor Economics, vol. 1*. Elsevier.
- Prescott, E. (2004). Why do Americans work so much more than Europeans? *Federal Reserve Bank of Minneapolis Quarterly Review* 28, 2–13.
- Samuelson, P. A. (1968). The two-part golden rule deduced as the asymptotic turnpike of catenary motions. *Western Economic Journal* 6, 85–89.
- Sanchez-Marcos, V. and A. S. Martin (2006). Can social security be welfare improving when there is demographic uncertainty? *Journal of Economic Dynamics and Control* 30, 1615–1646. Working Paper.
- Sims, C. A. (2001). Solving linear rational expectations models. *Computational Economics* 20(1-2), 1–20.
- Stern, N. (2006). The stern review on the economics of climate change.
- Woodford, M. (2003). *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton: Princeton University Press.
- Yaari, M. (1965). Uncertain lifetime, life insurance, and the theory of the consumer. *The Review of Economic Studies* 32 (2), 137–150.

Figure 1: Impulse response of demography: TFR and OADR



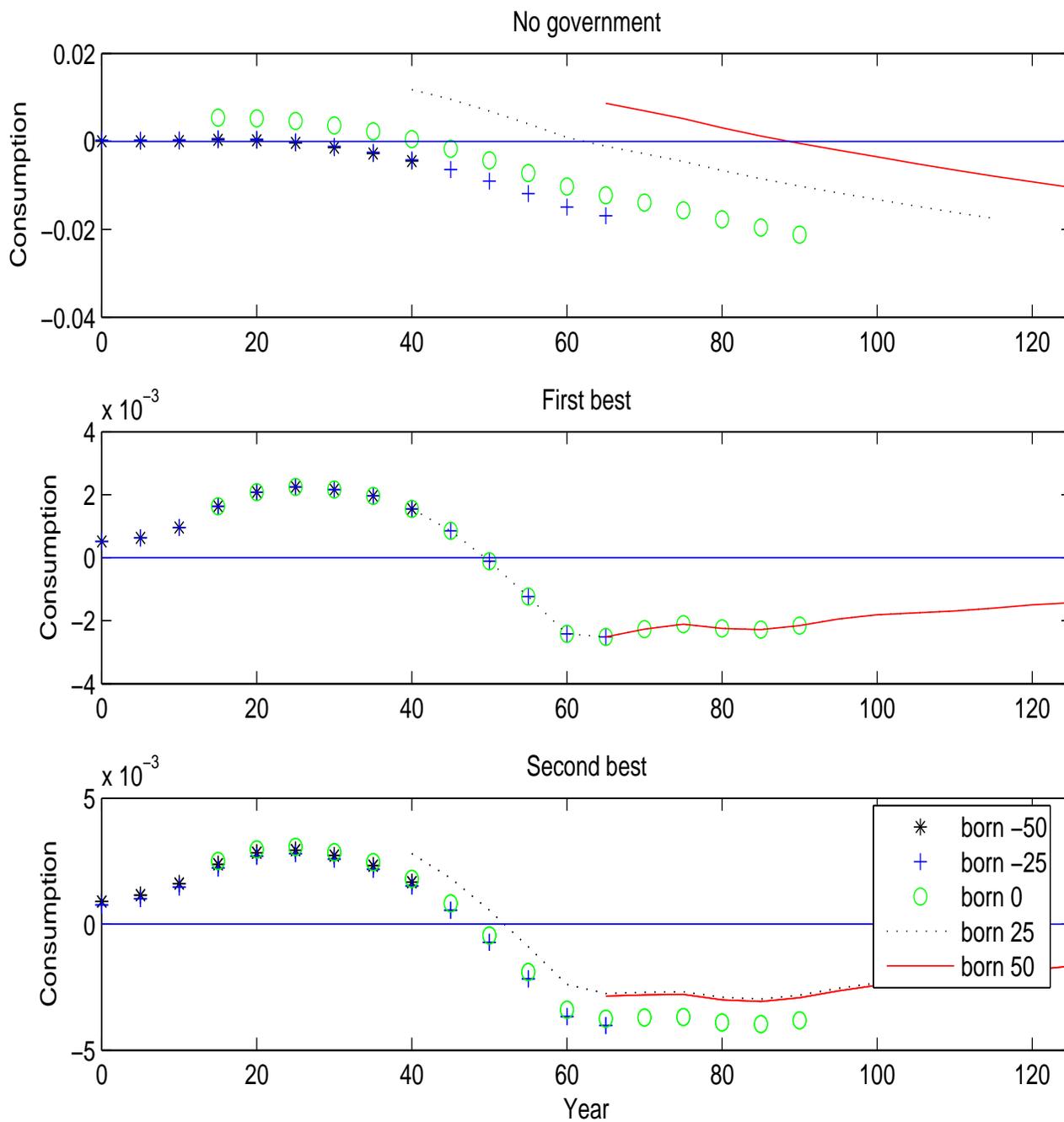
Notes: This graph shows the impulse response functions of the total fertility rate (left column) and the corresponding OADR (right column) in three population scenarios: low OADR (top row), high OADR (middle row) and baby-boom-bust cycle (bottom row).

Figure 2: Impulse response, patient gov.: $\tau^l = 0$, low OADR, $\eta = 0.6$



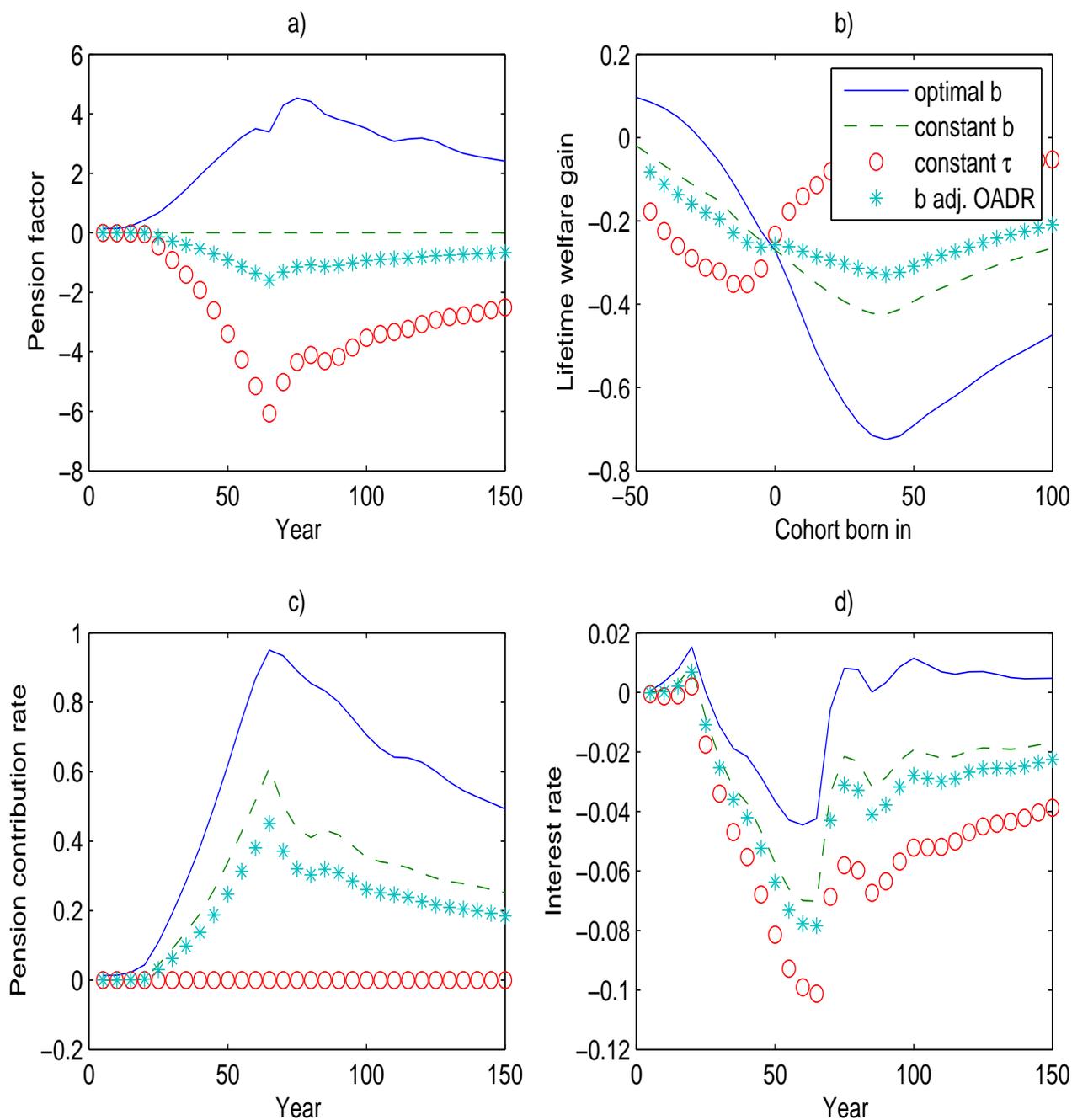
Notes: This graph shows impulse response functions for a shock to the fertility rate for the patient government where $\tau^l = 0$, the OADR is low and $\eta = 0.6$. Panel a: pension contribution rate, deviations in percentage points. Panel b: welfare gain, consumption equivalent variation. Panel c: labor tax wedge as described in Appendix A.2. Panel d: interest rate, deviations in percentage points.

Figure 3: Consumption



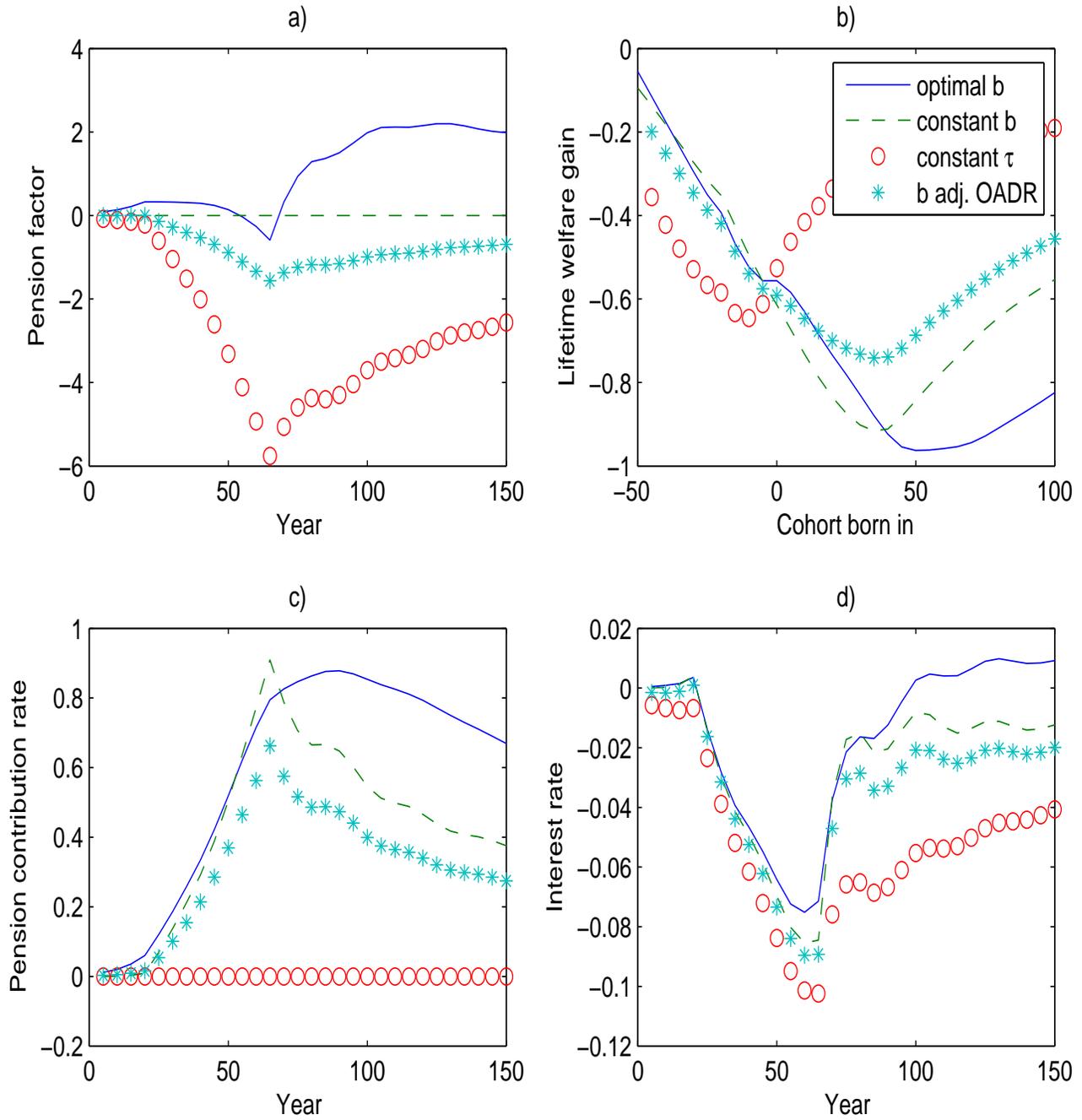
Notes: This graph shows the consumption responses for different cohorts in the case without government (top panel), first best (middle panel) and the second best scenario of Section 5.3.1 (bottom panel).

Figure 4: Impulse response, impatient gov.: $\tau^l = 0$, low OADR, $\eta = 0.2$



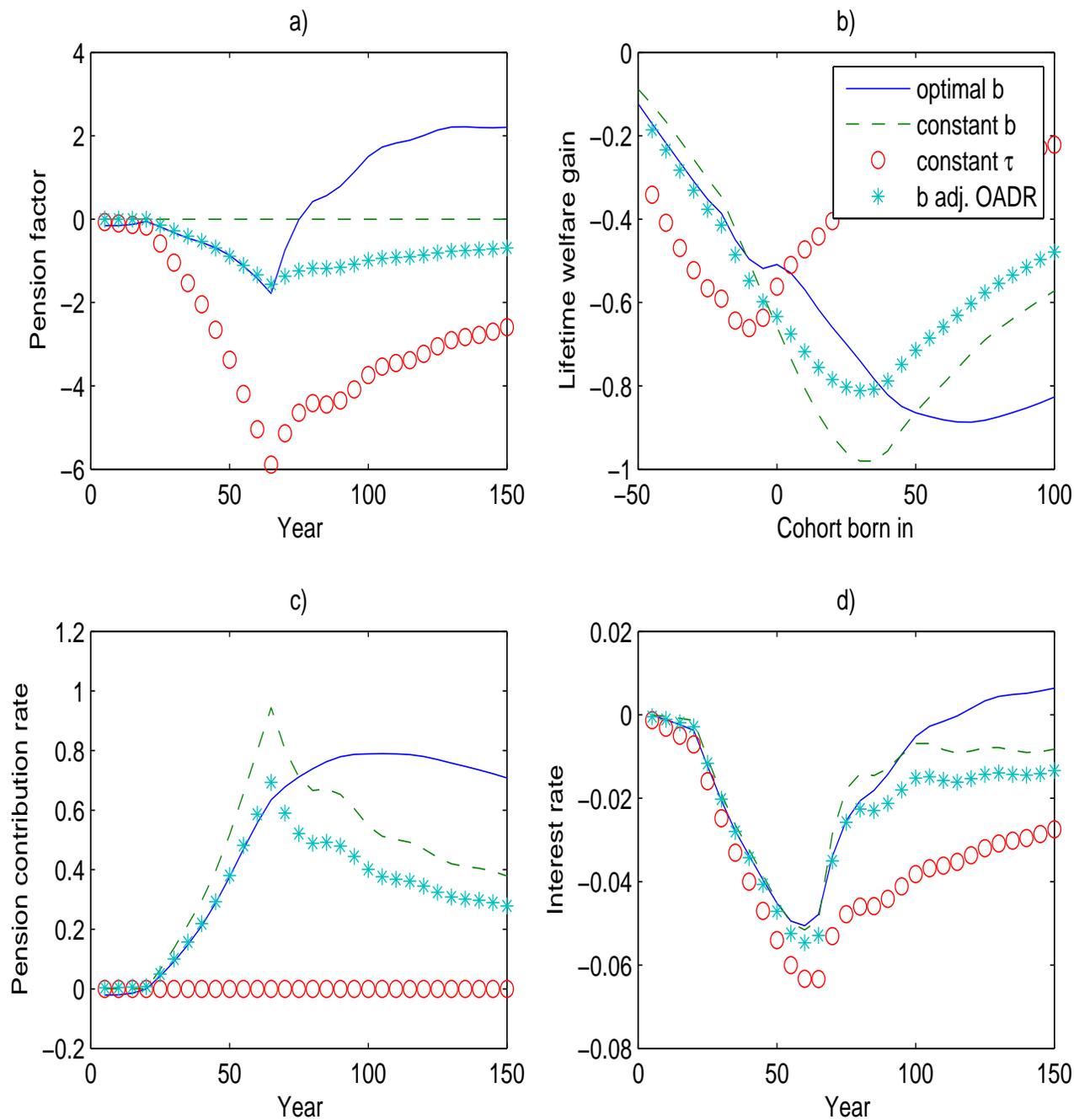
Notes: This graph shows impulse response functions for a shock to the fertility rate for the impatient government where $\tau^l = 0$, the OADR is low and $\eta = 0.2$. Panel a: pension factor, deviations in percent. Panel b: welfare gain, consumption equivalent variation. Panel c: pension contribution rate, deviations in percentage points. Panel d: interest rate, deviations in percentage points.

Figure 5: Impulse response, impatient gov.: $\tau^l = 0.345$, high OADR, $\eta = 0.6$



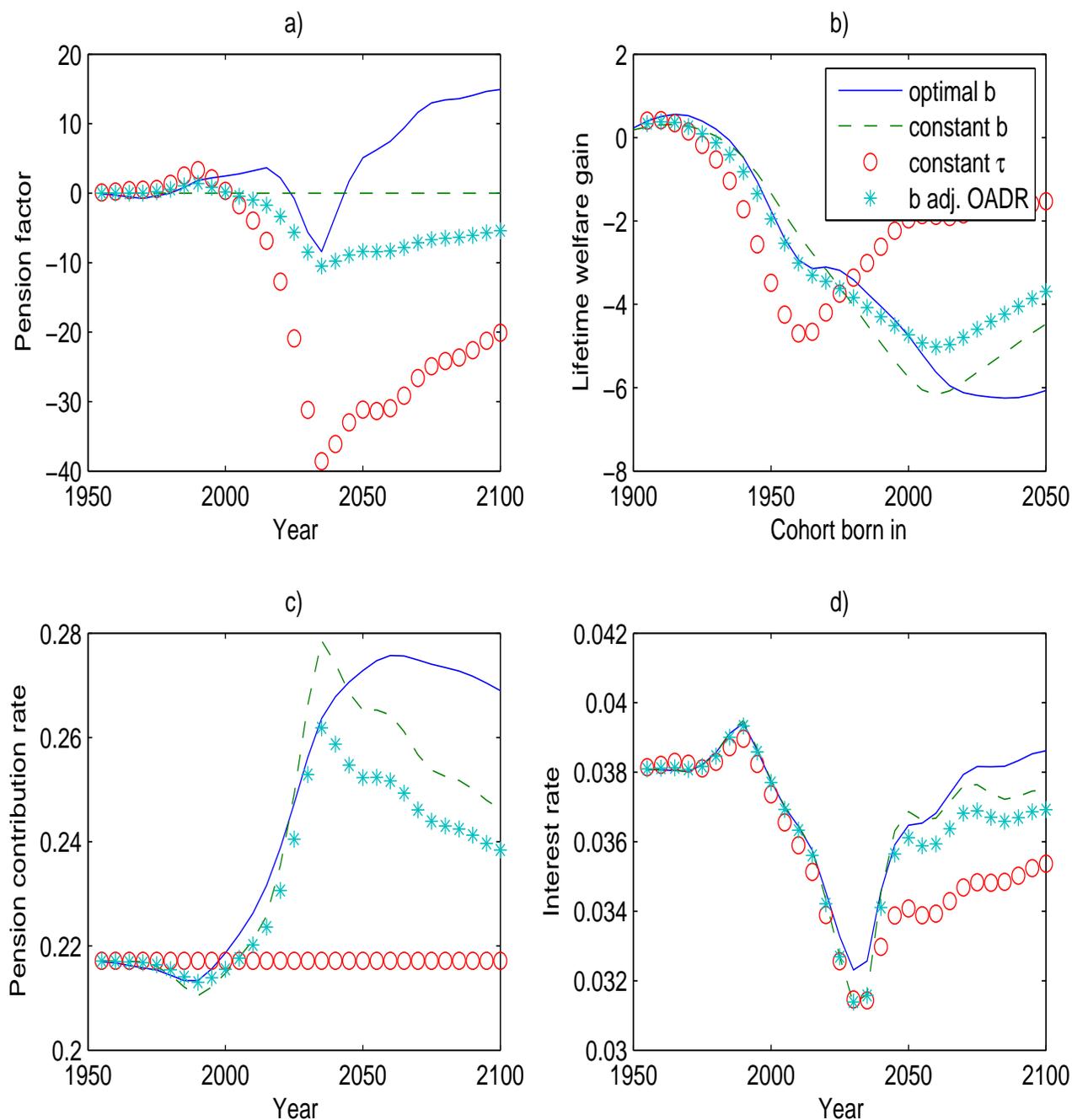
Notes: This graph shows impulse response functions for a shock to the fertility rate for the impatient government where $\tau^l = 0.345$, the OADR is high and $\eta = 0.6$. Panel a: pension factor, deviations in percent. Panel b: welfare gain, consumption equivalent variation. Panel c: pension contribution rate, deviations in percentage points. Panel d: interest rate, deviations in percentage points.

Figure 6: Impulse response, impatient gov. for CES production function: $\tau^l = 0.345$, high OADR, $\eta = 0.6$



Notes: This graph shows impulse response functions for a shock to the fertility rate for the impatient government using a CES production function with a substitution elasticity equal to 4 where $\tau^l = 0.345$, the OADR is high and $\eta = 0.6$. Panel a: pension factor, deviations in percent. Panel b: welfare gain, consumption equivalent variation. Panel c: pension contribution rate, deviations in percentage points. Panel d: interest rate, deviations in percentage points.

Figure 7: Baby boom/bust, impatient gov.: $\tau^l = 0.345$, high OADR, $\eta = 0.6$



Notes: This graph shows alternative policy reactions to the baby boom bust in Germany where $\tau^l = 0.345$, the OADR is high and $\eta = 0.6$. Panel a: pension factor. Panel b: welfare gain, consumption equivalent variation. Panel c: pension contribution rate. Panel d: interest rate.

Table 1: Steady state results

	Low OADR, $\eta =$			High OADR, $\eta =$		
	0.2	0.6	1.5	0.2	0.6	1.5
$\bar{\tau}^l = 0.2892$						
K , NoPens	24.83	17.73	10.75	26.94	19.19	11.58
ΔK , WithPens	-13.77	-13.70	-14.23	-23.73	-22.39	-21.36
ΔY , WithPens	-5.23	-6.00	-7.77	-9.23	-9.77	-11.41
ΔL , WithPens	-0.65	-1.90	-4.36	-0.97	-2.71	-5.98
annual r , NoPens	2.99	3.11	3.26	2.27	2.49	2.79
annual r , WithPens	3.86	3.90	3.94	3.84	3.85	3.88
labor tax rate, NoPens	28.92	28.92	28.92	28.92	28.92	28.92
labor tax rate, WithPens	40.70	42.00	44.57	50.60	51.38	53.13
tax wedge, WithPens	33.76	34.42	35.71	38.81	39.28	40.32
PensionPaym/ C^{old} , WithPens	53.29	59.32	71.11	78.17	81.11	87.66
PensionPaym/GDP, WithPens	7.85	8.72	10.43	14.45	14.97	16.14
$\bar{\tau}^l = 0$						
ΔK , WithPens	-13.24	-12.26	-11.09	-23.51	-21.59	-19.05
ΔL , WithPens	-0.16	-0.55	-1.42	-0.08	-0.47	-1.70
labor tax rate, WithPens	11.25	11.70	12.69	21.53	21.98	22.86
PensionPaym/ C^{old} , WithPens	50.86	52.91	57.37	77.65	79.29	82.56
$\bar{\tau}^l = 0.3455$						
ΔK , WithPens	-13.96	-14.20	-15.36	-23.79	-22.67	-22.26
ΔL , WithPens	-0.85	-2.42	-5.52	-1.32	-3.51	-7.47
labor tax rate, WithPens	46.46	47.99	51.03	56.16	57.00	59.09
PensionPaym/ C^{old} , WithPens	53.97	61.06	75.04	77.95	81.11	88.92

Notes: ΔK etc.: percent change, with pension system relative to a scenario without pension system. WithPens: optimal pension system with impatient government. The cohort weights are fixed across all experiments. NoPens: background distortions, no pension system.

Table 2: Distance of simple rules from optimal policy (deviations of impulse responses)

Parameters			const. τ^p	constant b		b , adj. OADR		
η	τ^l	OADR		NetW	GrossW	NetW	GrossW	
0.2	0.345	H	0.533	0.163	0.093	0.258	0.133	
0.6	0.345	H	0.470	0.125	0.188	0.201	0.114	
1.5	0.345	H	0.425	0.147	0.351	0.172	0.197	
0.6	0.289	H	0.501	0.129	0.153	0.216	0.116	
0.6	0.0	H	0.593	0.153	0.071	0.267	0.177	
0.6	0.345	L	0.409	0.153	0.105	0.217	0.172	
0.6	0.289	L	0.426	0.166	0.117	0.232	0.191	
0.6	0.345	H	$\theta = 2$	0.455	0.109	0.138	0.194	0.093
0.6	0.345	H	CES	0.405	0.140	0.286	0.163	0.165

Notes: CES: model with a CES production function where the elasticity of substitution is equal to 4; $\theta = 2$: model with a coefficient of relative risk aversion of 2; OADR: high or low.

Table 3: Welfare gains of fluctuations

Parameters			Opt. b	const. τ^p	constant b		b , adj. OADR		
η	τ^l	OADR			NetW	GrossW	NetW	GrossW	
0.2	0	H	0.667	0.631	0.664	0.666	0.659	0.665	
0.6	0	H	0.880	0.849	0.877	0.874	0.874	0.878	
1.5	0	H	1.193	1.167	1.189	1.169	1.189	1.185	
0.6	0	H	CES	0.944	0.926	0.941	0.932	0.941	0.939
0.6	0	H	$\gamma = 2$	1.076	0.990	1.071	1.069	1.061	1.073

Notes: CES: model with a CES production function where the elasticity of substitution is equal to 4; $\theta = 2$: model with a coefficient of relative risk aversion of 2; OADR: high or low.

Table 4: Variability of life time utility

Parameters			Opt. b	const. τ^p	constant b		b , adj. OADR		
η	τ^l	OADR			NetW	GrossW	NetW	GrossW	
0.2	0.345	H	0.082	0.019	0.052	0.091	0.037	0.058	
0.6	0.345	H	0.077	0.021	0.061	0.110	0.044	0.070	
1.5	0.345	H	0.085	0.026	0.082	0.164	0.059	0.103	
0.6	0.289	H	0.079	0.020	0.060	0.102	0.043	0.066	
0.6	0.0	H	0.092	0.019	0.062	0.089	0.043	0.057	
0.6	0.345	L	0.038	0.007	0.020	0.028	0.014	0.018	
0.6	0.289	L	0.037	0.006	0.018	0.025	0.013	0.016	
0.6	0.345	H	$\theta = 2$	0.067	0.018	0.046	0.092	0.031	0.054
0.6	0.345	H	CES	0.069	0.024	0.068	0.121	0.050	0.080

Notes: CES: model with a CES production function where the elasticity of substitution is equal to 4; $\theta = 2$: model with a coefficient of relative risk aversion of 2; OADR: high or low.

A Appendix

A.1 The German Pension System

The adjustment formula of the German pension system is given by¹⁰

$$b_t = b_{t-1} \frac{a_{t-1}}{a_{t-2}} \left(\alpha_b \left(1 - \frac{OADR_{t-1}}{OADR_{t-2}} \right) + 1 \right) \quad (34)$$

where the indexation by factor a_{t-1}/a_{t-2} is according to the net wage adjustment, cf. equation (27). Let γ_t denote the growth rate of the OADR, hence $\frac{OADR_{t-1}}{OADR_{t-2}} = 1 + \gamma_t$. We can therefore approximate the term in brackets in equation 34 as

$$\begin{aligned} \alpha_b \left(1 - \frac{OADR_{t-1}}{OADR_{t-2}} \right) + 1 &= 1 - \alpha_b \gamma_t \\ &\approx \left(\frac{OADR_{t-2}}{OADR_{t-1}} \right)^{\alpha_b}. \end{aligned}$$

With this approximation equation (34) can be rewritten as

$$b_t \approx b_{t-1} \frac{a_{t-1}}{a_{t-2}} \left(\frac{OADR_{t-2}}{OADR_{t-1}} \right)^{\alpha_b}.$$

Recursively substituting out terms in the above gives

$$b_t \approx \bar{b} \frac{1}{\bar{a}} \overline{OADR} \cdot a_{t-1} \cdot (OADR_{t-1})^{-\alpha_b}.$$

Changing timing then results in the approximation of the German pension adjustment rules as given in equation (26).

A.2 Tax Wedges

Participation in the PAYG pension system is compulsory, households would normally prefer to save in financial markets where they obtain a higher return. Pension contributions therefore distort the labor supply decision, although the distortion is smaller than the contribution rate because future pension payments are related to the labor supply decision, cf. equation (41b) below. We measure the overall distortion by the difference between the marginal utility of consumption and the marginal utility of leisure, which we call the tax wedge τ_t :

$$\tau_t = 1 + \frac{U_L/U_C}{w} \quad (35)$$

To interpret (35), note that the wedge is equal to the labor tax rate τ^l when this tax is the only distortion, because then the the household first-order condition is given by $U_L = -w(1 - \tau^l)U_C$.

¹⁰The actual German rule features an indexation to the economic dependency ratio – the ratio of pensioners to workers – rather than the OADR as we do here for reasons of simplification.

A.3 Proofs

Proof of proposition 1. Form the Lagrangian

$$\mathcal{L} = \mathbb{E}_0 \sum_{t=-I}^{\infty} \Omega_t \left[N_{t,0} \sum_{i=0}^I \beta^i \pi_{t,i} U(C_{t+i,i}, L_{t+i,i}) + \lambda_t \left(-K_{t+1} + (1 - \delta_t)K_t + F \left(K_t, Z_t \sum_{i=i^A}^{i^{R-1}} \epsilon_i L_{t,i} N_{t,i} \right) - \sum_{i=0}^I C_{t,i} N_{t,i} \right) \right]. \quad (36)$$

We get the first order conditions

$$\frac{\partial \mathcal{L}}{\partial C_{t,j}} = \Omega_{t-j} N_{t-j,0} \beta^j \pi_{t-j,j} U_C(C_{t,j}, L_{t,j}) - \Omega_t \lambda_t N_{t,j} = 0 \quad (37a)$$

$$\frac{\partial \mathcal{L}}{\partial L_{t,j}} = \Omega_{t-j} N_{t-j,0} \beta^j \pi_{t-j,j} U_L(C_{t,j}, L_{t,j}) + \Omega_t \lambda_t F_L(t) Z_t \epsilon_j N_{t,j} = 0 \quad (37b)$$

$$\frac{\partial \mathcal{L}}{\partial K_{t+1}} = -\Omega_t \lambda_t + \mathbb{E}_t [\Omega_{t+1} \lambda_{t+1} (1 - \delta + F_K(t+1))] = 0 \quad (37c)$$

We show that (37) implies (14). (14a) is immediate from combining (37a) and (37b). Since $N_{t-j,0} \pi_{t-j,j} = N_{t,j}$, we get from (37a) that

$$\Omega_t \lambda_t = \Omega_{t-j} \beta^j U_C(C_{t,j}, L_{t,j}), \quad \forall j = 0, \dots, I \quad (38)$$

Applying (38) twice in (37c), first choosing $j = i$, then choosing $j = i + 1$, we obtain (14b). Using (38) with varying $j \in (0, \dots, I)$, we obtain (14c).

This shows that the first order conditions (14) are necessary. The sufficiency follows from the concavity of the programming problem. \square

Proof of proposition 2. In the deterministic steady state we get from (14c) and (14b) that

$$\begin{aligned} \frac{U_C(C_{t-i,0}, L_{t-i,0})}{U_C(C_{t,0}, L_{t,0})} &= [\beta(1 + \bar{r})]^i \frac{U_C(C_{t,i}, L_{t,i})}{U_C(C_{t,0}, L_{t,0})} \\ &= (1 + \bar{r})^i \frac{\Omega_t}{\Omega_{t-i}} \end{aligned}$$

and $C_{t,0} = (1 + g)^i C_{t-i,0}$. By assumption (1) it then follows that

$$\frac{U_C(C_{t-i,0}, L_{t-i,0})}{U_C(C_{t,0}, L_{t,0})} = ((1 + g)^\theta)^i$$

and therefore

$$\frac{\Omega_t}{\Omega_{t-i}} = \left(\frac{(1 + g)^\theta}{1 + \bar{r}} \right)^i$$

and, setting $i = t$, we finally have that

$$\Omega_t = \omega^t = \left(\frac{(1+g)^\theta}{1+\bar{r}} \right)^t.$$

□

Proof of Proposition 3. In **PP2**, the government budget constraint is

$$b_t \frac{1+r_t(1-\tilde{\tau}_{t-1}^c)}{\varsigma_{t-1,0}} \tau_{t-1,0}^p w_{t-1,0} L_{t-1,0} N_{t,1} = \tau_{t,0}^p w_{t,0} L_{t,0} N_{t,0} + \tilde{\tau}_{t-1}^c r_t k_{t-1,0} N_{t-1,0} \quad \text{for all } t. \quad (39)$$

The household budget constraints are

$$k_{t,0} + C_{t,0} = w_{t,0}(1 - \tau_t^p - \tau^l) L_{t,0} + T_{t,0} \quad (40a)$$

$$C_{t+1,1} = \frac{1+r_{t+1}(1-\tilde{\tau}_t^c)}{\varsigma_{t,0}} [k_{t,0} + b_{t+1} \tau_t^p w_{t,0} L_{t,0}] \quad (40b)$$

and the household first order conditions are

$$U_C(C_{t,0}, L_{t,0}) = \beta_{\varsigma_{t,0}} \mathbb{E}_t \left[\frac{1+r_{t+1}(1-\tilde{\tau}_t^c)}{\varsigma_{t,0}} U_C(C_{t+1,1}, 0) \right] \quad (41a)$$

$$\begin{aligned} -U_L(C_{t,0}, L_{t,0}) &= w_{t,0} \left\{ (1 - \tau_t^p - \tau^l) U_C(C_{t,0}, L_{t,0}) \right. \\ &\quad \left. + \beta_{\varsigma_{t,0}} \tau_t^p \mathbb{E}_t \left[b_{t+1} \frac{1+r_{t+1}(1-\tilde{\tau}_t^c)}{\varsigma_{t,0}} U_C(C_{t+1,1}, 0) \right] \right\} \end{aligned} \quad (41b)$$

Equivalence between DP and PP1: first set the state-dependent capital taxes equal in both regimes. Then with the choices in (24), simple algebra shows that both the household budget constraints and first order conditions are satisfied in both regimes, for the same capital holdings. It can be easily checked that, under (24), the government budget constraints (17) and (21) are equivalent as well. Notice that this already follows from aggregate feasibility and the fact that HH budget constraints are satisfied in each period.

Equivalence between PP1 and PP2 requires that the contribution rate to the pension system, τ_t^p , is the same in both regimes and that the following conditions hold:

$$\mathbb{E}_t \left[\frac{1+r_{t+1}(1-\tau_{t+1}^c)}{\varsigma_{t,0}} U_C(C_{t+1,1}, 0) \right] = \mathbb{E}_t \left[\frac{1+r_{t+1}(1-\tilde{\tau}_t^c)}{\varsigma_{t,0}} U_C(C_{t+1,1}, 0) \right] \quad (42)$$

$$\frac{1+r_{t+1}(1-\tau_{t+1}^c)}{\varsigma_{t,0}} [k_{t,0} + b_t \tau_t^p w_{t,0} L_{t,0}] = \frac{1+r_{t+1}(1-\tilde{\tau}_t^c)}{\varsigma_{t,0}} [k_{t,0} + b_{t+1} \tau_t^p w_{t,0} L_{t,0}] \quad (43)$$

$$\mathbb{E}_t \left[b_t \frac{1+r_{t+1}(1-\tau_{t+1}^c)}{\varsigma_{t,0}} U_C(C_{t+1,1}, 0) \right] = \mathbb{E}_t \left[b_{t+1} \frac{1+r_{t+1}(1-\tilde{\tau}_t^c)}{\varsigma_{t,0}} U_C(C_{t+1,1}, 0) \right] \quad (44)$$

First choose $\tilde{\tau}_t^c$ such that (42) is satisfied. Given that, choose b_{t+1} for each state of the world in $t + 1$ such that (43) is satisfied. Then (42) and (43) together imply (44). To see this, multiply (43) on both sides by $U_C(C_{t+1,1}, 0)$ and form conditional expectations. From (42) we see that the term involving $k_{t,0}$ cancels on both sides of (43). The remainder implies (44).

To show that the two government budget constraints (21) and (39) are equivalent (again, this already follows from feasibility and the individual budget constraints), just subtract (39) from (21), add $\frac{1+r_t}{s_{t-1,0}}k_{t-1,0}$ on both sides of the resulting equation and use $N_{t-1,0} = N_{t,1}/s_{t,0}$. Then (43) implies that this difference is zero.

The argument for the case when the equivalence breaks down was already given in the text after Proposition 3. \square