UNIVERSITÄT
BERN

Faculty of Business, Economics
and Social Sciences

Department of Economics

Evaluating pay-as-you-go social security systems

Andreas Bachmann
Kaspar Wüthrich

13-10

November 2013

DISCUSSION PAPERS

Schanzeneckstrasse 1
Postfach 8573
CH-3001 Bern, Switzerland
http://www.vwi.unibe.ch

Evaluating pay-as-you-go social security systems *

Andreas Bachmann † — University of Bern
Kaspar Wüthrich‡ — University of Bern

November 19, 2013

Abstract

This paper proposes a new method for welfare analysis of unfunded social security systems. Based on an overlapping generations model with endogenous labor supply, we derive a formula for the evaluation of existing pay-as-you-go social security systems that depends on impulse response functions and projected growth rates only. We propose an implementation strategy based on reduced form estimates of a VAR model that is valid under weak assumptions about the deep structure of the model. Our method is related to the sufficient statistic approach (Chetty, 2009). For the current system in the United States, we find that a transitory increase in the payroll tax rate along with higher pension benefits leads to a welfare increase mainly due to welfare gains of today's retirees. A scenario analysis demonstrates the robustness of this result.

JEL Classification: E62, H55

Keywords: unfunded social security system; sufficient statistic; overlapping generations; reduced form VAR

*Acknowledgments: We are grateful to Fabrice Collard, Gita Gopinath, Stefan Leist, Klaus Neusser, Dirk Niepelt, Philipp Wegmüller, and seminar participants at the University of Bern for helpful suggestions and comments. All errors are of course our own.

†University of Bern, Department of Economics, Schanzeneckstrasse 1, CH-3001 Bern, Switzerland, phone: +41 31 631 4777, email: andreas.bachmann@vwi.unibe.ch.

‡University of Bern, Department of Economics, Schanzeneckstrasse 1, CH-3001 Bern, Switzerland, phone: +41 31 631 3189, email: kaspar.wuethrich@vwi.unibe.ch.

1 Introduction

Unfunded pay-as-you-go (PAYGO) social security systems play an important role in many developed countries' social insurance programs. Since demographic changes and the associated growing fraction of retirees in the population cause increasing financial stress for these systems, the question of how to design social security systems optimally becomes more and more relevant.

Social security systems are typically studied in the context of structural overlapping generations (OLG) models (examples include: Auerbach and Kotlikoff, 1987; Imrohoroglu et al., 1995; Kotlikoff et al., 1999, 2007; Nickel et al., 2008; Fehr et al., 2012; McGrattan and Prescott, 2013, among many others).[1] Welfare analysis in these models proceeds in two steps: first, the deep structure of the model (preferences and production function) is parameterized and estimated. Second, the effect of different policies and alternative social security systems on social welfare is computed using simulation methods. This approach features two main drawbacks: first, even flexible functional form assumptions might be arbitrary and hard to justify and second, it is typically difficult to estimate and identify all deep parameters in an empirically compelling manner.

This paper contributes to the literature by proposing a new method for welfare analysis of social security systems. Based on a Ramsey problem in a standard OLG model (cf. Diamond, 1965) augmented with endogenous labor supply, we derive a formula for the welfare consequences of changes in payroll taxes used to finance transfers in PAYGO systems. This formula does not depend on the deep structure of the model but rather on few high level quantities, which allows us to identify and estimate marginal welfare changes avoiding important parametric assumptions. In particular, we do not require specifying the functional form of the aggregate production function nor a full parameterization of household preferences.

Because our formula is a function of high level quantities rather than the deep structure of the model, it can be interpreted as a sufficient statistic (in the sense of Chetty, 2009). The sufficient statistic approach to welfare analysis has recently become important in the public economics literature (see e.g. Chetty, 2009, for a review). It provides a middle course between structural models and reduced-form methods. From the structural approach, it borrows the ability to make predictions about welfare, but avoids the problem of having to estimate or

[1] An important focus of these studies has been on shifts from a primarily unfunded system towards mixed systems that combine PAYGO with investment based personal retirement accounts.

calibrate the deep parameters of the model. From the reduced-form approach, it borrows the advantage of transparent and credible identification. To the best of our knowledge, the sufficient statistic approach has not yet been used to analyze social security systems in the context of macroeconomic dynamic general equilibrium models.

The central insight of our theoretical analysis is that a change in the tax rate affects welfare through three distinct channels: first, the direct effect of getting more transfers and paying more taxes, second, the effect through changes in factor prices, and third, the change in transfers due to the labor adjustment of the subsequent generations.

To implement the sufficient statistic formula empirically, we propose two different approaches. First, we consider approximate consumption equivalent impacts on each generation. The advantage of this approach is that we do not require any assumptions on preferences. Second, to be able to evaluate the total impact of a change in the payroll tax, we obtain a money metric of the welfare effect of a marginal change in the tax rate by an appropriate standardization. Because the overall effect inherently requires a comparison of weighted marginal utilities of different generations, we need to impose arguably weak assumptions on preferences and welfare weights. We show that for both approaches, welfare changes can be stated as functions of impulse response functions and predictions of future growth rates only. This allows for an empirical implementation based on the reduced form estimates of a vector autoregressive (VAR) model. Our empirical strategy therefore differs from the sufficient statistic literature (Card et al., 2007; Chetty, 2008, 2009) that merely relies on cross sectional estimates which is not possible in this paper given the dynamic general equilibrium nature of the theoretical model.

We illustrate our approach by assessing the PAYGO system of the United States. We find that in terms of approximate consumption equivalents, a marginal increase in the payroll tax raises welfare of today's retirees substantially and of today's workers slightly, while it reduces the welfare for future generations. A decomposition by theoretical channels reveals that factor price effects (i.e. induced changes in wage and interest rates) are the most important determinants of welfare changes. In terms of the overall effect, our findings indicate that for a broad range of values for the coefficient of relative risk aversion and the individual discount factor a marginal increase in the payroll tax is welfare increasing. Consistent with the approximate findings for the generationwise welfare changes, we find that the overall effect is mainly driven by the factor price effect. A scenario analysis confirms the robustness of our empirical findings.

Our paper differs from the aforementioned literature on structural OLG models in two main respects. First, it provides a new and complementary method for evaluating PAYGO systems. Under weak assumptions, we derive a sufficient statistic formula for welfare evaluation that can be implemented empirically. Second, our paper has a different focus. While a large strand of the literature analyzes the welfare consequences of shifts in the structure of the pension system, we focus on the evaluation of existing PAYGO systems.

The analysis in this paper is also related to studies focusing on globally optimal PAYGO system (e.g. Feldstein, 1985; Imrohoroglu et al., 1995). However, the analysis in this paper has a somewhat different focus on local welfare improvements due to small changes in the payroll tax.

The remainder of the paper is structured as follows. Section 2 presents the model and derives a formula for the welfare analysis of a change in the payroll tax. In section 3, we use this formula to assess the welfare consequences of a change in the payroll tax for the United States. Section 4 concludes.

2 Theory

We consider a deterministic OLG model with endogenous labor supply. The framework is closely related to the setups considered by Breyer and Straub (1993), Nourry (2001), Fanti and Spataro (2006), Gonzales-Eiras and Niepelt (2008) and Lopez-Garcia (2008). First, we discuss the problems of the household and the representative firm. Second, we characterize the competitive equilibrium. Third, we analyze the Ramsey problem of the benevolent government assuming fixed labor supply for the purpose of illustrating our procedure and comparing it to the literature. Then, we extend the Ramsey problem to the more relevant case of elastic labor supply and derive a formula for the welfare consequences of a change in the payroll tax. Finally, we propose two approaches to empirically implement this formula.

2.1 Demographics, preferences and technology

We consider a deterministic, perfectly competitive economy inhabited by an infinite sequence of overlapping generations.[2] Each generation lives for two periods. In the first period,

[2]There is an initial old generation at the beginning.

households supply labor elastically, $0 \leq n_t \leq \bar{n}$.[3] In the second period, they retire. Population grows at an exogenously given rate. Let L_t denote the size of the labor force (i.e. the size of the young generation) in period t and define $\chi_{t,z} \equiv \frac{L_z}{L_t} - 1$ as the working age population growth rate between two periods t and z.

Households have preferences over consumption in both periods and leisure $l_t = \bar{n} - n_t$. Consumption in the first period, c_t^y, equals post tax labor income, $n_t w_t(1 - \tau_t)$ where τ_t denotes the payroll tax, minus savings s_{t+1}. In the second period, households consume c_{t+1}^o, which is equal to the gross returns on savings, $R_{t+1}s_{t+1}$, plus lump sum social security benefits, T_{t+1}, and profits of the firm, $\frac{\Pi_{t+1}}{L_t}$, where Π_{t+1} is the overall profit of firms which is distributed across the capital owners. Preferences are summarized by the utility function $u(c_t^y, \bar{n} - n_t, c_{t+1}^o)$ with $u_{c^y}(\cdot) > 0$, $u_l(\cdot) > 0$, $u_{c^o}(\cdot) > 0$, $u_{c^y c^y}(\cdot) < 0$ and $u_{c^o c^o}(\cdot) < 0$. The intertemporal decision of the household solves

$$\max_{n_t, c_t^y, c_{t+1}^o, s_{t+1}} \quad u(c_t^y, \bar{n} - n_t, c_{t+1}^o)$$

$$\text{s.t.} \quad c_t^y + s_{t+1} = n_t w_t (1 - \tau_t)$$

$$c_{t+1}^o = R_{t+1} s_{t+1} + T_{t+1} + \frac{\Pi_{t+1}}{L_t}$$

For later reference, define λ_t^y and λ_{t+1}^o to be the Lagrange multipliers associated with the budget constraint of the household when young and old, respectively. The first order conditions of the household maximization problem read:

$$\lambda_t^y = u_{c^y}(c_t^y, \bar{n} - n_t, c_{t+1}^o) \tag{1}$$

$$\lambda_t^y w_t (1 - \tau_t) = u_l(c_t^y, \bar{n} - n_t, c_{t+1}^o) \tag{2}$$

$$\lambda_{t+1}^o = u_{c^o}(c_t^y, \bar{n} - n_t, c_{t+1}^o) \tag{3}$$

$$\lambda_t^y = R_{t+1} \lambda_{t+1}^o \tag{4}$$

The saving's decision of a household in cohort t is given by the usual consumption Euler equation:

$$u_{c^y}(c_t^y, \bar{n} - n_t, c_{t+1}^o) = R_{t+1} u_{c^o}(c_t^y, \bar{n} - n_t, c_{t+1}^o). \tag{5}$$

[3] $\bar{n}$ denotes the number of available hours in a time period that can be split between leisure and work.

The labor supply is described by:

$$u_{c^y}(c_t^y, \bar{n} - n_t, c_{t+1}^o) w_t (1 - \tau_t) = u_l(c_t^y, \bar{n} - n_t, c_{t+1}^o). \tag{6}$$

(5) and (6) combined with the household's budget constraints map factor prices w_t and R_{t+1}, policy variables τ_t and T_{t+1}, and profits $\frac{\Pi_{t+1}}{L_t}$ into savings and labor:

$$s_{t+1} = S(w_t, \tau_t, T_{t+1}, R_{t+1}, n_t, \frac{\Pi_{t+1}}{L_t}) \tag{7}$$

$$n_t = N(w_t, \tau_t, T_{t+1}, R_{t+1}, s_{t+1}, \frac{\Pi_{t+1}}{L_t}) \tag{8}$$

The firm sector is characterized by a set of competitive firms that can be represented by an aggregate production function, $F(K_t, H_t E_t)$, that maps inputs of capital K_t and hours worked $H_t = L_t n_t$ into output. E_t denotes exogenous labor efficiency. The problem of the firm is static. In each period, the representative firm solves

$$\max_{K_t, H_t} \quad F(K_t, H_t E_t) - w_t H_t - r_t K_t.$$

The first order conditions of the firm problem imply:

$$w_t = F_{HE}(K_t, H_t E_t) E_t \tag{9}$$

$$r_t = F_K(K_t, H_t E_t) \tag{10}$$

We impose the following standard assumption on the aggregate production function.

Assumption 1. $F(K_t, H_t E_t)$ *exhibits constant returns to scale.*

Assumption 1 and the Euler theorem imply zero profits ($\Pi_t = 0$) in equilibrium.

2.2 Competitive equilibrium without pension system

Before turning to the planner's problem, we characterize a competitive equilibrium in this economy absent a PAYGO system, i.e. $\tau_t = 0$ and $T_t = 0$ for all $t = 0, .., \infty$. A competitive equilibrium in this OLG economy is a sequence of aggregate capital stocks, household consumption, labor supply, and factor prices $\{K_t, c_t^y, c_t^o, n_t, R_t, w_t\}_{t=0}^{\infty}$ such that firms' input and output choices are profit maximizing given factor prices, households' consumption and labor supply decisions are utility maximizing given factor prices and firms' profits, and the allocation is feasible. Factor market clearing implies:

$$K_t = s_t L_{t-1} \tag{11}$$

$$H_t = n_t L_t \tag{12}$$

Since capital is the only asset that households may save in, the gross return on savings corresponds to

$$R_t = 1 - \delta_t + r_t, \tag{13}$$

where δ_t denotes the depreciation rate of the capital stock. Combining the households' budget constraints with (9), (10) and (13) implies the goods market equilibrium:

$$F(K_t, H_t E_t) = L_t c_t^y + L_{t-1} c_t^o + K_{t+1} - (1 - \delta_t) K_t. \tag{14}$$

2.3 A Ramsey problem with inelastic labor supply

We consider a benevolent government that seeks to maximize social welfare, W, subject to technological and competitive equilibrium constraints. Based on the problem of a Ramsey planner, our approach is to analyze the welfare change induced by a change of the payroll tax, $dW/d\tau$. In contrast to the standard procedure, we do not solve for the optimal sequence of payroll taxes. Instead, we are interested in the welfare effect of a one-time change in the current tax rate. This focus is motivated by the two period nature of our model where a sensible choice of the period length is multiple decades.[4] A straightforward alternative would be to consider a permanent change in the payroll tax. Our goal is to derive an empirically implementable expression for $dW/d\tau$ that is a function of empirically estimable high level elasticities.

To illustrate our procedure and to compare it to the literature, we first consider an economy with fixed labor supply.

Assumption 2. *Individual labor supply n_t is inelastic and normalized to* 1.

For the ease of notation, we define $u(c_t^y, 1, c_{t+1}^o) \equiv u(c_t^y, c_{t+1}^o)$. Notice that Assumption 2

[4]Following the literature, we will choose a period length of 30 years for the empirical implementation.

implies that $L_t = H_t$. We extend our analysis to the empirically more relevant case of elastic labor supply in the next section.

Under Assumption 2, following Gonzales-Eiras and Niepelt (2007, 2008), the program of the government with commitment – the Ramsey program – for a given sequence of welfare weights $\{\xi_t\}$ at $t = 0$ is given by

$$\max_{0 \leq \{\tau_t\}_{t=0}^{\infty} \leq 1} W = \sum_{t=0}^{\infty} \xi_t u(c_t^y, c_{t+1}^o) + \xi_{-1} u\left(c_{-1}^y, c_0^o\right)$$

$$\text{subject to} \begin{cases} s_0, c_{-1}^y \quad \text{given} \\ s_{t+1} = S(w_t, \tau_t, T_{t+1}, R_{t+1}, 1, 0) \quad t \geq 0 \\ \text{household budget constraints} \end{cases} \tag{15}$$

We assume throughout that the sequence of welfare weights $\{\xi_t\}$ is declining sufficiently fast for the problem to be well defined. The pay-as-you-go character of the social security system implies that retiree pensions paid are equal to the taxes collected:

$$L_{t-1} T_t = w_t \tau_t L_t$$

Using the envelope conditions of the household maximization problem, the effect of a marginal increase in the current payroll tax rate τ_0 on the Ramsey planner's objective function is given by

$$\frac{dW}{d\tau_0} = -w_0 u_{c^y}(c_0^y, c_1^o)\xi_0 + w_0(1 + \chi_{-1,0}) u_{c^o}(c_{-1}^y, c_0^o)\xi_{-1} + \Psi_0 \tag{16}$$

where Ψ_0 summarizes the general equilibrium effects of a change in τ_0. The total effect can be decomposed into three components. The first term in (16) measures the direct welfare loss of the young generation caused by an increase in tax payments. The second term reflects the direct welfare gain of the old generation in period $t = 0$ due to an increase in social security transfers. Using the FOC of the household optimization problem, the net social benefit of transferring one unit of resources in period $t = 0$ from young to old can be expressed as

$$w_0(1 + \chi_{-1,0}) u_{c^o}(c_{-1}^y, c_0^o)\xi_{-1} - w_0 u_{c^y}(c_0^y, c_1^o)\xi_0 \tag{17}$$

Besides the direct redistribution effects, a change in τ_0 has general equilibrium effects that are captured in Ψ_0. The policy change in period $t = 0$ causes a change in savings and, thus, in the capital stock, which in turn affects future wages and interest rates and, consequently, social welfare (for the ease of notation, we replace the marginal utilities of the households by the respective multipliers):

$$\begin{aligned}\Psi_0 &= \sum_{t=1}^{\infty} \frac{dw_t}{d\tau_0}(1-\tau_t)\lambda_t^y \xi_t + \sum_{t=0}^{\infty} \left(\frac{dR_{t+1}}{d\tau_0} s_{t+1} + \frac{dT_{t+1}}{d\tau_0} \right) \lambda_{t+1}^o \xi_t \\ &= \sum_{t=1}^{\infty} \frac{dw_t}{d\tau_0}(1-\tau_t)\lambda_t^y \xi_t + \sum_{t=0}^{\infty} \left(\frac{dR_{t+1}}{d\tau_0} s_{t+1} + (1+\chi_{t,t+1})\tau_{t+1} \frac{dw_{t+1}}{d\tau_0} \right) \lambda_{t+1}^o \xi_t\end{aligned}$$

Due to constant returns to scale, there is a direct relation between $\frac{dR_t}{d\tau_0}$ and $\frac{dw_t}{d\tau_0}$. Totally differentiating $F(K_t, H_t E_t) = w_t H_t + r_t K_t$ with respect to τ_0 implies[5]

$$F_K(K_t, H_t E_t) \frac{dK_t}{d\tau_0} + F_{HE}(K_t, H_t E_t) E_t \frac{dH_t}{d\tau_0} = \frac{dw_t}{d\tau_0} H_t + w_t \frac{dH_t}{d\tau_0} + \frac{dr_t}{d\tau_0} K_t + r_t \frac{dK_t}{d\tau_0}$$

Using (9), (10), (11), (12) and (13), this simplifies to

$$\frac{dR_t}{d\tau_0} = \frac{dr_t}{d\tau_0} = -\frac{dw_t}{d\tau_0} \frac{H_t}{K_t} = -\frac{dw_t}{d\tau_0} \frac{n_t(1+\chi_t)}{s_t} \tag{18}$$

Using the optimality condition (4) of the household problem, we can replace λ_t^y by $\lambda_{t+1}^o R_{t+1}$. The overall welfare change given a change in τ_0 can therefore be expressed as follows:

$$\begin{aligned}\frac{dW}{d\tau_0} =& - w_0 \lambda_1^o R_1 \xi_0 + w_0 (1+\chi_{-1,0}) \lambda_0^o \xi_{-1} \\ &+ \sum_{t=0}^{\infty} \frac{dw_{t+1}}{d\tau_0} (1-\tau_{t+1}) \left(\lambda_{t+2}^o R_{t+2} \xi_{t+1} - (1+\chi_{t,t+1}) \lambda_{t+1}^o \xi_t \right)\end{aligned} \tag{19}$$

2.4 Elastic labor supply

The results in the previous section were derived assuming inelastic labor supply. Clearly, this assumption is not satisfactorily because labor supply is likely to respond directly or indirectly to demographic developments (Gonzales-Eiras and Niepelt, 2008). Moreover, distortions induced by the payroll tax can have important welfare consequences that must be considered in the evaluation of changes in tax rates.

With elastic labor supply, the program of the Ramsey planner at $t = 0$ for a given sequence

[5] Labor efficiency E_t is exogenous and therefore $\frac{dE_t}{d\tau_0} = 0$.

of welfare weights $\{\xi_t\}$ reads

$$\max_{0\leq\{\tau_t\}_{t=0}^{\infty}\leq 1} W = \sum_{t=0}^{\infty} \xi_t u(c_t^y, \bar{n} - n_t, c_{t+1}^o) + \xi_{-1} u\left(c_{-1}^y, \bar{n} - n_{-1}, c_0^o\right)$$

$$\text{subject to} \begin{cases} s_0, c_{-1}^y, n_{-1} \quad \text{given} \\ s_{t+1} = S(w_t, \tau_t, T_{t+1}, R_{t+1}, n_t, 0) \quad t \geq 0 \\ n_t = N(w_t, \tau_t, T_{t+1}, R_{t+1}, s_{t+1}, 0) \quad t \geq 0 \\ \text{household budget constraints} \end{cases} \tag{20}$$

where transfers are given by $T_t = n_t w_t \tau_t (1 + \chi_{t-1,t})$. The effect of a marginal increase in the current payroll tax rate τ_0 on welfare is given by

$$\frac{dW}{d\tau_0} = w_0 n_0 ((1 + \chi_{-1,0})\lambda_0^o \xi_{-1} - \lambda_0^y \xi_0) + \Psi_0, \tag{21}$$

where Ψ_0 summarizes the effects of adjustments in n_t, w_t and R_t due to the change in τ_0. In contrast to the case of inelastic labor supply, w_0 is not predetermined by K_0 any more. Instead, it also depends on H_0, which may be affected by a change in τ_0. Ψ_0 is given by:

$$\begin{aligned} \Psi_0 = &\sum_{t=0}^{\infty} \xi_t \left(\lambda_t^y \frac{dw_t}{d\tau_0} n_t (1 - \tau_t) + \lambda_{t+1}^o \left(\frac{dR_{t+1}}{d\tau_0} s_{t+1} + \frac{dT_{t+1}}{d\tau_0} \right) \right) \\ &+ \xi_{-1} \lambda_0^o \left(\frac{dR_0}{d\tau_0} s_0 + \tau_0 (1 + \chi_0) \left(n_0 \frac{dw_0}{d\tau_0} + w_0 \frac{dn_0}{d\tau_0} \right) \right) \end{aligned}$$

Using

$$\begin{aligned} \frac{dT_{t+1}}{d\tau_0} &= \frac{dn_{t+1}}{d\tau_0} w_{t+1} \tau_{t+1} (1 + \chi_{t,t+1}) + \frac{dw_{t+1}}{d\tau_0} n_{t+1} \tau_{t+1} (1 + \chi_{t,t+1}), \\ \frac{dR_{t+1}}{d\tau_0} &= -\frac{dw_{t+1}}{d\tau_0} \frac{n_{t+1}(1 + \chi_{t,t+1})}{s_{t+1}}, \\ \lambda_t^y &= R_{t+1} \lambda_{t+1}^o, \end{aligned}$$

we get the following overall welfare effect:

$$
\begin{aligned}
\frac{dW}{d\tau_0} =& w_0 n_0(-\lambda_1^o R_1 \xi_0 + (1+\chi_{-1,0})\lambda_0^o \xi_{-1}) \\
&+ \sum_{t=0}^{\infty} \frac{dw_t}{d\tau_0} n_t (1-\tau_t)(\lambda_{t+1}^o R_{t+1} \xi_t - (1+\chi_{t-1,t})\lambda_t^o \xi_{t-1}) \\
&+ \sum_{t=0}^{\infty} \frac{dn_t}{d\tau_0} w_t \tau_t (1+\chi_{t-1,t})\lambda_t^o \xi_{t-1}
\end{aligned} \tag{22}
$$

At this point, it is worthwhile discussing the different channels of the overall welfare effect in more detail. There are three basic components, each of which corresponds to a line in equation (22). First, there is the direct effect of getting more transfers and paying higher taxes. The initial old generation benefits while the initial young households are worse off. Second, there are indirect welfare effects owing to the impact of the policy change on factor prices. The effect of factor prices consists of three components: changes in the wage for the young household, changes in the interest rate and changes in the wage that affect the old households through transfers. Third, for each generation there is an indirect welfare effect due to the labor adjustment of the subsequent generation. Grouping the overall welfare change by generations, equation (22) can be rewritten as follows:

$$
\begin{aligned}
\frac{dW}{d\tau_0} =& \xi_{-1}\lambda_0^o \Bigg(w_0 n_0 (1+\chi_{-1,0}) - \frac{dw_0}{d\tau_0} n_0 (1-\tau_0)(1+\chi_0) + \frac{dn_0}{d\tau_0} w_0 \tau_0 (1+\chi_{-1,0}) \Bigg) \\
&+ \xi_0 \lambda_1^o \Bigg(-R_1 w_0 n_0 + R_1 \frac{dw_0}{d\tau_0} n_0 (1-\tau_0) - \frac{dw_1}{d\tau_0} n_1 (1-\tau_1)(1+\chi_{0,1}) \\
&\qquad + \frac{dn_1}{d\tau_0} w_1 \tau_1 (1+\chi_{0,1}) \Bigg) \\
&+ \sum_{t=1}^{\infty} \xi_t \lambda_{t+1}^o \Bigg(R_{t+1} \frac{dw_t}{d\tau_0} n_t (1-\tau_t) - \frac{dw_{t+1}}{d\tau_0} n_{t+1} (1-\tau_{t+1})(1+\chi_{t,t+1}) \\
&\qquad + \frac{dn_{t+1}}{d\tau_0} w_{t+1} \tau_{t+1} (1+\chi_{t,t+1}) \Bigg)
\end{aligned} \tag{23}
$$

Clearly, only the initial old and young generation are directly affected by higher transfers and higher tax rates, respectively, while factor price and labor supply changes have an indirect impact on all generations.

At this point, the standard approach to proceed in the literature is to parameterize the deep structure of the model, i.e. impose functional form assumptions on preferences and the aggregate production function and express $dW/d\tau_0$ as a function of a set of primitives that need to be estimated from the data. In this paper, we follow an alternative strategy. We make use of the fact that equation (23) is a function of (i) $dw_t/d\tau_0, \quad t = 0, ..., \infty$, (ii) marginal utilities and (iii) predicted economic quantities such as labor supply and wages. Put it differently, in order to identify and estimate $dW/d\tau_0$ it is sufficient know quantities (i) to (iii). The key point is that knowledge of the deep parameters that generate this quantities is not required and, thus, a full specification of the deep structure of the model can be avoided.

We argue in section 3 that (i) can be identified and estimated using a reduced form VAR and that empirical predictions can be used to compute (iii). The remaining challenge is to identify the marginal utilities empirically. We propose two alternative approaches. In section 2.5, we show that it is possible to compute first order approximations for the consumption equivalent impact of a marginal change in the payroll tax for each generation without further assumptions on preferences. Identification and estimation of the overall effect $dW/d\tau_0$ requires more assumptions because comparisons of marginal utilities between different generations are involved. In particular, we partly need to parameterize household preferences as discussed in section 2.6.

2.5 Consumption equivalent impact on each generation

Suppose there is a hypothetical increase in c^o by ϕ percent. Utility of the generation born in t would then be given by $u(c_t^y, \bar{n} - n_t, (1 + \phi_t)c_{t+1}^o)$. A first order Taylor approximation around $\phi_t = 0$ yields

$$u(c_t^y, \bar{n} - n_t, (1 + \phi_t)c_{t+1}^o) \approx u(c_t^y, \bar{n} - n_t, c_{t+1}^o) + u_{c^o}(c_t^y, \bar{n} - n_t, c_{t+1}^o)c_{t+1}^o\phi_t$$

The change in utility can therefore be approximated by

$$u(c_t^y, \bar{n} - n_t, (1 + \phi_t)c_{t+1}^o) - u(c_t^y, \bar{n} - n_t, c_{t+1}^o) \approx \lambda_{t+1}^o c_{t+1}^o \phi_t.$$

The change in a generation's utility due to the change in τ_0 is linear in this generations marginal utility (cf. equation (23)). In particular, the change in utility takes the form $\lambda_{t+1}^o\Omega_t$, which is then weighted by ξ_t and summed up in order to get the overall welfare change. For each generation, we can approximatively calculate the (hypothetical) percentage change in consumption when retired which would make this generation equally good off as the policy change:

$$\lambda^o_{t+1} c^o_{t+1} \phi_t \approx \lambda^o_{t+1} \Omega_t \tag{24}$$

$$\Rightarrow \phi_t \approx \frac{\Omega_t}{c^o_{t+1}} \tag{25}$$

This yields the approximative welfare effect of the policy change in terms of consumption for each generation.

Proposition 1. *Consider an OLG economy as described in section 2.1 with a PAYGO social security system. Suppose that Assumption 1 is satisfied. Then, the impact of a one-time marginal change in τ_0 on the welfare of the generation born in t is equivalent (up to a first-order approximation) to an increase in this generations consumption when retired by ϕ_t, where $\phi_t = \frac{\Omega_t}{c^o_{t+1}}$ and*

$$\Omega_{-1} = w_0 n_0 (1 + \chi_{-1,0}) - \frac{dw_0}{d\tau_0} n_0 (1 - \tau_0)(1 + \chi_{-1,0}) + \frac{dn_0}{d\tau_0} w_0 \tau_0 (1 + \chi_{-1,0}), \tag{26}$$

$$\Omega_0 = -R_1 w_0 n_0 + R_1 \frac{dw_0}{d\tau_0} n_0 (1 - \tau_0) - \frac{dw_1}{d\tau_0} n_1 (1 - \tau_1)(1 + \chi_{0,1}) + \frac{dn_1}{d\tau_0} w_1 \tau_1 (1 + \chi_{0,1}), \tag{27}$$

$$\Omega_t = R_{t+1} \frac{dw_t}{d\tau_0} n_t (1 - \tau_t) - \frac{dw_{t+1}}{d\tau_0} n_{t+1} (1 - \tau_{t+1})(1 + \chi_{t,t+1}) + \frac{dn_{t+1}}{d\tau_0} w_{t+1} \tau_{t+1} (1 + \chi_{t,t+1}), \tag{28}$$

for $t > 0$.

Observe that equation (26), (27), and (28) contains impulse response functions with respect to a transitory shock in the payroll tax rate, $\{\frac{dw_t}{d\tau_0}, \frac{dn_t}{d\tau_0}\}_{j=0,\ldots,\infty}$, predictions of economic quantities, and projected working age population growth. This allows for an implementation of equation (30) based on a reduced form VAR model. We provide a more detailed discussion of the implementation in section 3.

2.6 Overall welfare effect

Proposition 1 allows for an approximate welfare analysis by generation in terms of consumption equivalences. However, for a throughout policy evaluation this information is not sufficient because the Ramsey planner cares about a weighted sum of all future generations' utilities. Thus, knowledge of the overall effect of a change in the payroll tax is imperative. Because utility is not quasi-linear, we need to convert $\frac{dW}{d\tau_0}$ into a money metric (Chetty, 2009). We obtain an intuitive metric by normalizing the welfare change given an increase in the payroll tax rate by the welfare gain from a hypothetical additional unit of income, a_0, of the initial old household ($\frac{dW}{da_0} = \xi_{-1}\lambda_0^o$):

$$\begin{aligned}\frac{\frac{dW}{d\tau_0}}{\frac{dW}{da_0}} =& w_0 n_0 \left(-\frac{\lambda_1^o}{\lambda_0^o} R_1 \frac{\xi_0}{\xi_{-1}} + (1+\chi_{-1,0}) \right) \\ &+ \sum_{t=0}^{\infty} \frac{dw_t}{d\tau_0} n_t (1-\tau_t) \left(\frac{\lambda_{t+1}^o}{\lambda_0^o} R_{t+1} \frac{\xi_t}{\xi_{-1}} - (1+\chi_{t-1,t}) \frac{\lambda_t^o}{\lambda_0^o} \frac{\xi_{t-1}}{\xi_{-1}} \right) \\ &+ \sum_{t=0}^{\infty} \frac{dn_t}{d\tau_0} w_t \tau_t (1+\chi_{t-1,t}) \frac{\lambda_t^o}{\lambda_0^o} \frac{\xi_{t-1}}{\xi_{-1}}\end{aligned} \tag{29}$$

Equation (29) shows that an assessment of the overall effect $dW/d\tau_0$ requires aggregating the generation-specific effects. This aggregation inherently includes a comparison of weighted marginal utilities between different generations and, therefore, requires some additional structure on marginal utilities, λ_{t+1}^o, and welfare weights, ξ_t. We impose the following assumption on the household utility function.

Assumption 3. *Household preferences are additively separable over time and flow utility of the old households is of CRRA type, i.e.*

$$u(c_t^y, \bar{n} - n_t, c_{t+1}^o) = u(c_t^y, \bar{n} - n_t) + \beta \frac{\left(c_{t+1}^o\right)^{1-\gamma} - 1}{1-\gamma}$$

where β is the individual discount factor and γ is the coefficient of relative risk aversion.

Notice that Assumption 3 implies that the ratio of marginal utilities is a function of gross consumption growth which can be predicted empirically.

$$\frac{u_c(c_j^o)}{u_c(c_0^o)} = \left(\frac{c_j^o}{c_0^o} \right)^{-\gamma}$$

As a final step towards the implementation of equation (29), we need to impose some structure on the welfare weights (ξ_j). Following the literature (e.g. Gonzales-Eiras and Niepelt, 2008) we make the following assumption.

Assumption 4. *The social planner's welfare weights for different generations reflect both* *the discount factor of individual households and the size of the generations:*

$$\xi_j = \beta(1 + \chi_j)\xi_{j-1},$$

with $\xi_{-1} = 1$

We are now in the position to summarize the empirically implementable formula for the welfare consequences of a change in the payroll tax in the following proposition.

Proposition 2. *Consider an OLG economy as described in ection 2.1 with a PAYGO social security system. Suppose that Assumptions 1, 3 and 4 hold. Then, the overall welfare gain from an increase in the payroll tax* τ *relative to a $1.00 increase the income of the initial old household is given by*

$$\begin{aligned}\frac{\frac{dW}{d\tau_0}}{\frac{dW}{da_0}} =& w_0 n_0 \left((1 + \chi_{-1,0}) \left(1 - \left(\frac{c_1^o}{c_0^o} \right)^{-\gamma} R_1 \beta \right) \right) \\ &+ \sum_{t=0}^{\infty} \frac{dw_t}{d\tau_0} n_t (1 - \tau_t) \left((1 + \chi_{-1,t}) \beta^t \left(\left(\frac{c_{t+1}^o}{c_0^o} \right)^{-\gamma} R_{t+1} \beta - \left(\frac{c_t^o}{c_0^o} \right)^{-\gamma} \right) \right) \\ &+ \sum_{t=0}^{\infty} \frac{dn_t}{d\tau_0} w_t \tau_t \left(\frac{c_t^o}{c_0^o} \right)^{-\gamma} \beta^t (1 + \chi_{-1,t})\end{aligned} \tag{30}$$

Apart from γ and β, the quantities in equation (30) are either observable or can be estimated. As for the approximate consumption equivalent formulae in Proposition 1, equation (30) contains impulse response functions, predictions of economic quantities, and projected population growth. In addition, consumption growth must be projected to infer the ratios of marginal utilities. Given this similarities, implementation can be based on the same reduced form VAR estimates. We refer to section 3 for more details.

3 Empirical analysis

We illustrate our method for computing the welfare consequences of a change in the payroll tax by analyzing the social security system in the United States.[6] First, the empirical implementation of Propositions 1 and 2 is discussed. Second, we describe the data and the aggregation to the frequency of the OLG model. Finally, we present the results and some robustness checks.

[6]See e.g. Feldstein (2005) for a description of the social security system in the United States.

3.1 Methodology

Proposition 1 provides formulae for the percentage change in consumption when retired which would make a generation equally good off as the change in the payroll tax rate. Proposition 2 provides a formula for the overall welfare effect of a change in the payroll tax rate. The quantities needed to estimate the equations in these propositions empirically can be divided into four groups.

Payroll tax. In the United States, payroll taxes for social security can be split into the designated purposes of Old-Age and Survivors Insurance (OASI), Disability Insurance (DI) and Hospital Insurance (HI). Given the setup of our model, only payroll taxes used to fund the PAYGO system are considered. Thus, we take τ to be the payroll tax for OASI. We evaluate equations (26) to (28) and (30) at the current level of $\tau = 0.106$.

Impulse responses of labor supply and factor prices. The welfare effect depends on the dynamic responses of labor and wages with respect to a change in the payroll tax rate. A natural way to estimate these dynamic responses with respect to an exogenous change in the payroll tax rate is to estimate impulse response functions based on a reduced form VAR model:

$$Y_t = c + \Phi_1 Y_{t-1} + \ldots + \Phi_p Y_{t-p} + \theta \tau_t + \varepsilon_t. \tag{31}$$

The impulse response to a change in τ_t is given by

$$\frac{dY_{t+j}}{d\tau_t} = \begin{cases} \theta & \text{for } j = 0 \\ \Phi_1 \frac{dY_{t-1+j}}{d\tau_t} + \ldots + \Phi_p \frac{dY_{t-p+j}}{d\tau_t} & \text{for } j = 1, 2, \ldots \end{cases} \tag{32}$$

where $\frac{dY_k}{d\tau_t} = 0$ for $k < t$. Y_t is a vector of endogenous variables, Φ_j, c and θ are matrices and vectors of coefficients. ε_t is a vector of error terms which is multivariate white noise. We estimate a four-dimensional VAR model of order 3 using quarterly data on growth rates of hours worked per capita, real wages, real consumption per capita and real GDP per capita. All of these growth rates are stationary. The VAR order of 3 is chosen based on residual autocorrelation tests and the Akaike information criterion. We include the payroll tax rate as an exogenous variable. While the payroll tax rate clearly depends on factors such as life expectancy, retirement age and the shape of the age pyramid, we consider it to be exogenous in the present setting. These factors influencing τ_t are unlikely to have a direct impact on Y_t conditional on the lagged values of Y. Moreover, because we have quarterly data and τ_t is determined by legislation, which takes some time to adjust, τ_t is likely to be independent of ε_t conditional on lagged values of Y. Therefore, our identifying assumption is that $\mathbb{E}[\varepsilon_t \tau_t | Y_{t-1}, \ldots, Y_{t-p}] = 0$. To provide some evidence for our identifying assumption, we perform a Granger causality by including τ_t as an endogenous variable in the VAR model. We find that τ_t is exogenous in the sense that it cannot be predicted using lagged values of Y.

Forecasts. The empirical implementation of equations (26) to (28) and (30) requires forecasts for present and future generations on real wages, real interest rates, labor, working age population growth, and consumption of retirees. Because long-term forecasting is a very delicate issue, we decide to conduct a scenario analysis to account for the uncertainty of future developments (see section 3.3.4). The baseline scenario is constructed as follows. Real wages, hours worked and real consumption of retirees are projected using the (geometric) mean growth rate of the observed data. The projected real interest rate is set to the mean of the observed data. For the working age population, we base the projections on the (geometric) mean growth rate according to the national population projections released by the U.S. Census Bureau.[7]

Parameters. We avoid identification and estimation of the parameters β and γ. Instead, we estimate equation (30) for various plausible parameter values.

3.2 Data

Table 1 provides a description of the data that are used for the empirical estimation described in the previous section. However, these data cannot directly be plugged into equations (26) to (28) and (30) in order to evaluate the PAYGO system. Particular attention has to be paid to the aggregation of the data in order to fit the framework of the OLG model. In the OLG model, two time periods correspond to an entire lifespan. Following the literature (e.g. Gonzales-Eiras and Niepelt, 2008; Song, 2011), we take one period in the model to be 30 years. Period 0 in equations (26) to (28) and (30) corresponds to the years 2013 to 2042, period 1 to the years 2043 to 2072, etc. Table 2 explains how the quantities needed for estimating the sufficient statistic are aggregated.

[7] http://www.census.gov/population/projections/, last accessed on November 15, 2013.

Variable	Description	Sample period
n_t	Number of hours worked per capita per quarter. The series is constructed using an index of hours worked in the business sector (source: Bureau of Labor Statistics), scaled up to match total hours worked in the base year and divided by the number of people between 15 and 65 years old (source: Bureau of Labor Statistics).	1948q1 to 2013q1
w_t	Real hourly wage (in 2010 dollars). In correspondance with n_t, we use an index of real hourly compensation in the business sector (source: Bureau of Labor Statistics), scaled to match average hourly earnings of all employees in the total private sector in 2010.	1947q1 to 2013q1
R_t	Real interest rate. The series is constructed using the *10-Year Treasury Constant Maturity Rate* (source: Board of Governors of the Federal Reserve System) as nominal long-term interest rate. Our inflation measure is based on the *Consumer Price Index for All Urban Consumers* (source: Bureau of Labor Statistics).	1962q1 to 2012q2
χ_t	Working age population growth rate (source: Bureau of Labor Statistics, and Census Bureau).	1952q2 to 2013q1
c_t^o	Real consumption of retirees (in 2010 dollars). The series is constructed by deflating yearly data on total average expenditures of people over age 65 (source: Bureau of Labor Statistics, Consumer Expenditure Survey).	1988 to 2011
real consumption per capita growth	Growth rate of real *Private Final Consumption Expenditure in United States* (source: OECD National Accounts Statistics) divided by total population (source: U.S. Department of Commerce, Census Bureau).	1952q1 to 2013q1
real GDP per capita growth	Growth rate of real GDP (source: U.S. Bureau of Economic Analysis).	1947q1 to 2013q1

Table 1: Description of the Data

Variable	Aggregation
n_t	Quarterly data are added up over 30 years in order to get the number of hours worked per capita over 30 years.
w_t	The average real hourly wage over 30 years is computed.
R_t	The real long-term gross interest rate per annum is projected to 30 years (i.e. $R_{30y} = R_{p.a.}^{30}$).
$\chi_{t,z}$	For each 30 year window, the average size of the working age population is computed. This series is then used to compute the growth rate of the working age population.
c_t^o	Annual data are added up over 30 years in order to get total real consumption per capita over 30 years.

Table 2: Aggregation of higher frequency data to OLG frequency of 30 years.

3.3 Results

3.3.1 VAR

Figure 1 plots cumulative impulse response functions with respect to a increase in the payroll tax rate by one percentage point for 120 quarters, which corresponds to an increase in the payroll tax rate for one generation in the OLG model. Both hours worked and real wages are negatively affected in the long-run. When the payroll tax rate returns to its initial level after 120 quarters, the impulse responses of the growth rates go back to zero whereas the levels of hours work and real wages remain below their initial values.

The impulse responses match economic intuition. An increase in taxes reduces the income of households when young. As a consequence, these households reduce consumption in both periods of their life which implies that, for given transfers and factor prices, savings are reduced. Thus, we intuitively expect that this policy change leads to a decline in capital. This should depress wages and increase interest rates, which correspond to their marginal product, respectively. Given higher payroll taxes and this downward pressure on wages, it is not surprising that hours worked decline.

In order to implement equations (26) to (28) and (30), we need the impulse responses $\frac{dw_j}{d\tau_0}$ and $\frac{dn_j}{d\tau_0}$. These responses of wages and hours are computed by subtracting the projected path for these variables absent any shock from the projected path given the increase in the payroll tax rate (which is constructed using the cumulative impulse response functions in figure 1). The resulting level responses are shown in figure 2. Hours worked decline by a small amount. The maximum response amounts to minus 700 hours for a generation, which means a reduction of labor by 2 hours per month. While the response of labor becomes smaller in absolute values over time, the response of wages explodes. This is due to the fact that when the shock has died out and wage growth returns to its long-run rate, the level of wages is lower than it would be absent the change in taxes. As time goes by, this difference in the wage level increases exponentially with the long-run growth rate of real wages.

3.3.2 Consumption equivalent impact

The impulse response functions estimated above and predicted economic quantities allow for calculating the impact of the payroll tax change on each generation. Proposition 1 shows how the change in utility of each generation can be approximately expressed in terms of a

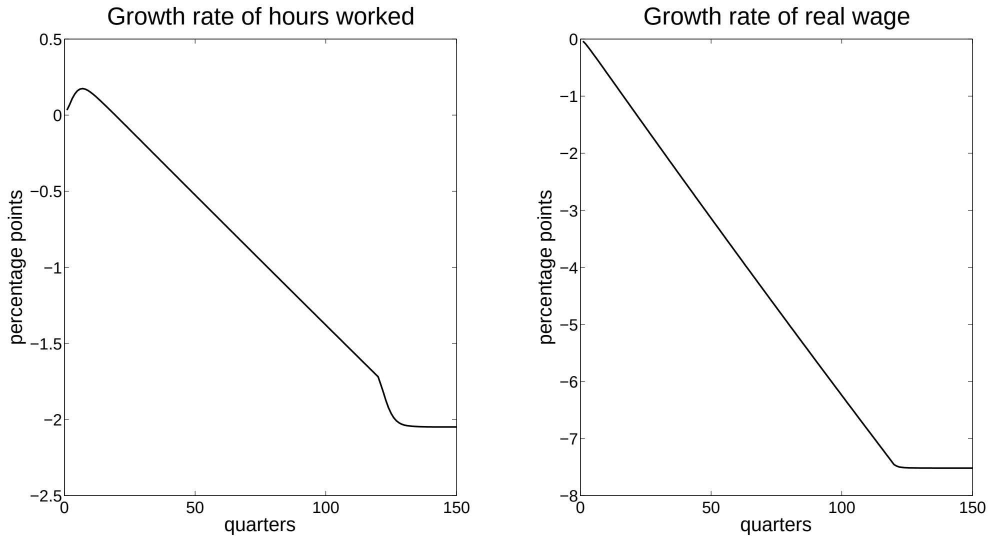

Figure 1: Cumulative impulse response of the growth rates of hours worked and real wages to an increase in the payroll tax rate for 120 quarters (i.e. one generation)

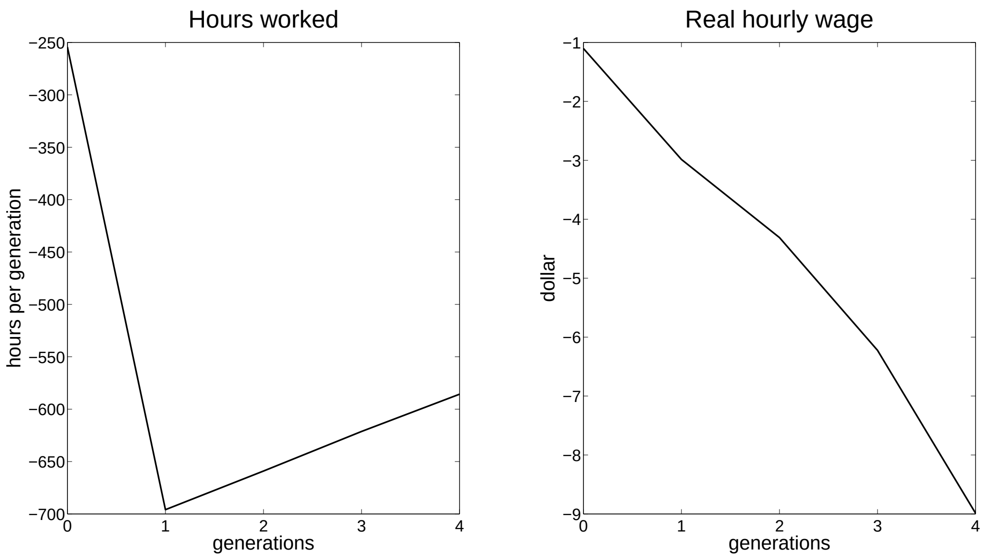

Figure 2: Impulse response of hours worked and real wages to an increase in the payroll tax rate for one generation

percentage change in consumption during retirement. The results are summarized in table 3.

Several results deserve closer attention. First, the payroll tax change has a negative impact on labor and therefore decreases transfers, but this *labor effect* seems to be of minor relevance. Second, the *factor price effects* are large compared to both the direct and the labor effects. The factor price effect can be further decomposed into the effect due to changes of wages when young, wages when old and interest rates. This decomposition reveals that changes in the interest rate and in the wage level when young have a large impact whereas changes in the wage level when old, which affect the level of transfers paid to retirees and therefore also utility, are less important.

	generation 0	generation 1	generation 2	generation 3	generation 4
direct effect	0.95	-1.28	0.00	0.00	0.00
factor price effect	3.38	1.75	-2.44	-2.64	-2.87
labor effect	-0.07	-0.20	-0.22	-0.24	-0.26
total effect	4.26	0.27	-2.66	-2.89	-3.13

Table 3: Percentage retiree's consumption change with equivalent welfare effect as the policy change.

The results in table 3 indicate a distributional conflict across generations. We find that the initial old generation benefits, not only directly through higher transfers but also indirectly due to favorable factor price changes. In particular, they benefit from higher interest rates on their savings. This indirect effect even exceeds the direct impact. For the initial young generation, there are counteracting effects. While there is obviously a negative direct effect of having to pay higher taxes, there is a positive indirect effect owing to changes in factor prices, which leaves the overall effect close to zero. The decomposition of the factor price effect reveals that the benefit from higher interest rates outweighs the loss due to lower wages. Finally, the future generations which are not directly affected by the policy change are worse off. The negative effect is mostly due to the lower wage level during their working age.

The findings in table 3 have important implications for the welfare evaluation of PAYGO systems. It is not sufficient to consider only the direct redistribution effects of a policy change in the PAYGO system because there are indirect effects of the tax change on wages and interest rates which seem to have substantial welfare impacts. Moreover, not only the directly affected generations must be taken into account but also future generations which are indirectly affected.

3.3.3 Overall welfare effect

The previous section allowed for an analysis of the approximate welfare impact for each generation. Since some generations are better and others are worse off, it is important to have a measure which aggregates the utility changes of all generations in order to evaluate the policy change. Proposition 2 provides such a measure. In this section, we use equation (30) to evaluate the PAYGO social security system of the United States.

The unobservable parameters γ and β are set to a variety of values covering the range of parameter values commonly used in the literature.[8] Table 4 shows the result for a selection of parameter values. The sign of the welfare change due to an increase in the payroll tax crucially depends on these parameters. The larger β, the more relative weight is given to future generations. As table 3 shows, the current generations benefit from the policy change whereas the future generations are negatively affected. Thus, the overall welfare effect is positive for low values of β and negative for large values of β. Although γ is conceptually very different from β, it also matters for the relative welfare weights across generations. Since we project consumption to grow, future generations have lower marginal utility. How much marginal utility shrinks over time depends on γ. A decrease in γ shifts relative weight to future generations, similar as an increase in β.

	$\beta = 0.9^{30}$	$\beta = 0.925^{30}$	$\beta = 0.95^{30}$	$\beta = 0.975^{30}$	$\beta = 0.995^{30}$
$\gamma = 0.0$	5.5	5.5	5.1	-0.1	-1'012.3
$\gamma = 0.5$	5.5	5.5	5.2	2.1	-258.5
$\gamma = 1.0$	5.5	5.5	5.3	3.3	-73.4
$\gamma = 1.5$	5.5	5.5	5.4	4.0	-23.1
$\gamma = 2.0$	5.5	5.5	5.4	4.4	-7.3

Table 4: Overall welfare effect (in 10'000) according to (30) as a function of the parameters γ and β.

To understand the driving factors of the numbers in table 4, it is instructive to decompose the values by channels and generations. In section 2, we have shown that the welfare consequences of a change in the payroll tax is the sum of three different components: the direct change in taxes and transfers, the welfare impact via changes in the wage and interest rate, and the welfare impact via changes in labor. As a benchmark case, table 5 shows the results

[8] In particular, we analyze the welfare change for values $\gamma \in [0, 2]$ and $\beta \in [0.9^{30}, 0.995^{30}]$. The later interval stems from assuming an individual yearly discount factor of at least 0.9 and strictly less than 1.

of this decomposition for $\beta = 0.95^{30}$ and $\gamma = 1$. Of course, the sign of the numbers in table 5 are identical with the sign of the numbers in table 3. However, in contrast to table 3, future generations are now downweighted because $\beta < 1$ and because of the projection $\frac{c_t^o}{c_0^o} > 1$ for $t > 0$, which implies lower marginal utility for future generations.

The sum of the labor effects across generations is negative for each considered combination of parameters γ and β. The sum of the factor price effects is positive except for $\beta = 0.995^{30}$. The same holds for the sum of the direct welfare effects. Thus, the sign of the overall welfare effect depends crucially on the direct and on the factor price effect, which may be positive or negative depending on the parameter values.

	generation 0	generation 1	generation 2	generation 3	generation 4	total
direct effect	12.2	-4.4	0	0	0	7.9
factor price effect	43.7	6.0	-2.0	-0.5	-0.1	47.0
labor effect	-0.9	-0.7	-0.2	-0.0	-0.0	-1.9
total	55.1	0.9	-2.2	-0.6	-0.2	53.1

Table 5: Decomposition of the overall welfare effect (in 1'000) by generations and channels, based on parameter values $\gamma = 1$ and $\beta = 0.95^{30}$.

It is important to recall that our results are only locally valid. Equation (30) captures the marginal effect of a change in the payroll tax as a function macroeconomic variables and, in particular, of τ. Thus, the results presented in this section provide a welfare evaluation of the current PAYGO system in the United States (with $\tau = 0.106$). This is especially important when comparing our result to the findings in the literature based on structural models, which typically do not analyze marginal changes. For example, Auerbach and Kotlikoff (1987) analyze the welfare effects of the introduction of an unfunded social security system with 60 percent benefit-to-earnings replacement rate under different tax regimes. Similar to our results, they find gains for the older generations and losses for the younger and future generations. Using an applied general equilibrium model, Imrohoroglu et al. (1995) find the optimal replacement rate of an unfunded social security system to be 30% (as opposed the empirically more realistic rate of 60%). Moreover, their results indicate that even with an empirically realistic replacement rate of 60% a social security system can be welfare enhancing. Kotlikoff et al. (2007) consider different alternative policies to mitigate the problems of the demographic transition in the United States. One such policy consists of a 50% benefit reduction which helps to limit the (endogenous) growth in the payroll tax. Their simulations show welfare losses for the older and the present generations and welfare gains for the future generations. Because in our model a benefit reduction is directly linked to a tax cut through the government budget constraint, these results are qualitatively comparable to our welfare projections. At this point it is important to note that instead of focussing on the payroll tax rate as the policy instrument, our analysis could alternatively be based on the benefit rates.

3.3.4 Robustness checks

Detailed knowledge of the future development of real wages, hours worked, real interest rates, consumption of retirees, and working age population growth is crucial for the evaluation of PAYGO systems based on equations (26) to (28) and (30). As a baseline scenario, we use means or mean growth rates of the available data sample for forecasting.[9] In the light of structural breaks, this might be an inappropriate forecast to use given that we need projections for several future generations. Therefore, this section conducts a sensitivity analysis on how our findings depend on the projected paths for the relevant variables. Using scenarios to account for the uncertainty of future developments is quite common in the literature on social security systems (e.g. Pecchenino and Utendorf, 1999; Kotlikoff et al., 2007; Imrohoroglu and Kitao, 2009; McGrattan and Prescott, 2013).

In addition to the baseline scenario, we consider a higher and a lower future development of each variable of interest. Figrue 3 depicts the alternative paths. We cover a wide range of possible developments. For time series with a positive (negative) trend in levels,[10] the high (low) scenario is constructed assuming a 50% higher growth rate compared to the baseline. The low (high) scenario consists in eliminating the trend and assuming the series to be constant. For working age population growth, which the U.S. Census Bureau predicts to be lower in the future than it was in the past, the high scenario consists in fixing growth at the positive rate projected for the next generation. The low scenario is created by reducing the (yearly) growth rate from the baseline case by additional 0.1 percentage points per generation.[11] Finally, since the real interest rate has no trend, the high (low) scenario is constructed by adding (subtracting) 0.1 percentage points to the (yearly) baseline real interest rate per

[9] An exception is the working age population growth rate, for which the forecast is based on the national population projections from the U.S. Census Bureau.

[10] Real wages and real consumption per capita of retirees show a positive trend, hours worked per capita show a negative trend.

[11] Note that figrue 3 depicts working age population growth over generations of 30 years. For example, the long-run baseline forecast of 12.8% would translate into an average yearly growth rate of 0.4%.

generation.[12]

For each scenario, we compute the consumption equivalent impact (equations (26) to (28)) and the overall welfare effect (equation (30)) for various parameter values β and γ (i.e. we recompute the tables 3 and 4).

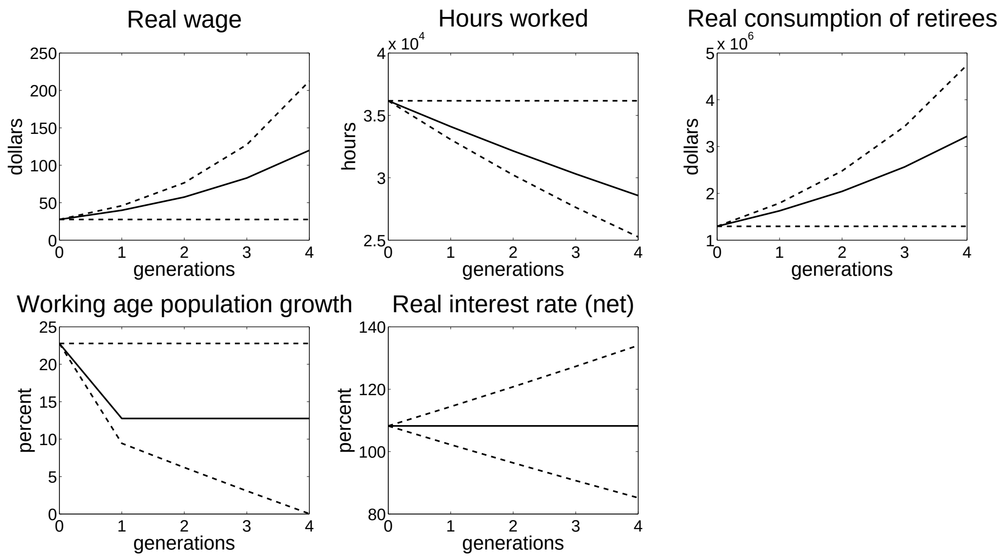

Figure 3: Different scenarios for the future development of variables affecting the welfare change. The solid line represents the baseline, the dashed lines indicate the high and low scenario.

	real wages	hours worked	consumption	population growth	real interest rate
high	5.7	5.5	5.4	5.5	5.3
baseline	5.3	5.3	5.3	5.3	5.3
low	4.6	5.2	5.1	5.2	5.3

Table 6: Values (in 10'000s) for the overall welfare effect according to (30) for different scenarios.

Overall, our results are very robust regarding changes in the future development of macroeconomic variables. For all scenarios, the sign of the consumption equivalent impact is identical to table 3 for almost each channel and generation. Moreover, although there are some quantitative differences, the overall welfare effect is stable over many of the scenarios. Table 6 summarizes the overall welfare effect for the different scenarios for the case of $\beta = 0.95^{30}$ and $\gamma = 1$. For this combination of parameter values, all scenarios yield virtually the same overall welfare effect.

[12]Note that figrue 3 depicts the net real interest rate over generations of 30 years. For example, the long-run baseline forecast of net 108.2% interest would translate into an average yearly net interest rate of 2.5%.

For other combinations of parameter values, especially for large values of β, the overall welfare effect is less stable. In particular, the evolution of consumption of retirees may be crucial. Consumption of retirees matters for marginal utility and hence for the relative welfare weight of different generations. The more consumption grows, the lower is the weight of future generations. Since future generations are negatively affected by the change in the payroll tax rate whereas the initial old generation is positively affected, higher consumption growth is associated with a more positive overall welfare effect.

4 Conclusion

Old-age provision constitutes an essential element of many developed countries' social insurance programs. As demographic changes cause increasing financial stress for PAYGO systems, reforms of existing systems become more and more relevant.

This paper proposes a new approach for the welfare analysis of PAYGO social security systems. Based on a standard OLG model with endogenous labor supply, we derive a formula for the local welfare consequences of a change in the payroll tax that can be viewed as a sufficient statistic (cf. Chetty, 2009). We propose two approaches to implement this formula based on predictions for different key quantities of the model and the reduced form estimates of a VAR. Using data for the United States, we find substantial positive welfare effects for today's retirees, small welfare gains for today's workers and welfare losses for future generations. To this end, our findings indicate that changing the payroll tax induces a distributional conflict across generations. The overall welfare effect of increasing the payroll tax is estimated to be positive provided that the coefficient of relative risk aversion and the household discount factor lie within a plausible range. A detailed decomposition by channels and generations sheds light on the driving forces behind this result. In particular, we find that the effect through factor prices outweighs the direct effect (through higher taxes and higher benefits) and the labor effect. Robustness checks confirm that our findings do not crucially depend on the precise predictions of key quantities.

Compared to the traditional approach to welfare analysis based on calibrated and estimated structural models, our method does not require the knowledge of the deep structure of the model. In particular, we do not require a full parameterization of household preferences and a specification of the aggregate production function. Instead, the welfare consequences of a change in the payroll tax can be deduced from reduced form estimates under weaker and arguably more credible assumptions.

Our approach shares two important limitations of the sufficient statistic approach (cf. Chetty, 2008, 2009): First, our results are only locally valid. In particular, the analysis of real world payroll tax changes would require additional assumptions due to the discrete (and not infinitesimal) nature of these policy changes. Second, a new sufficient statistic needs to be derived for every research question. Importantly for our application, it is not possible to use the same sufficient statistic for the analysis of mixed social security systems or to compare different pension systems (e.g. funded and unfunded systems). In the light of these limitations, we consider our method to be complementary to the structural approach, because it allows for a weakening of required assumptions on the one hand, but it only applies to the specific question of welfare consequences of payroll tax changes in PAYGO systems on the other hand.

Regarding the sufficient statistics literature, our analysis extends the range of applications to macroeconomic dynamic general equilibrium models and highlights the challenges associated with deriving and implementing sufficient statistics in these models. The basic idea of deriving sufficient statistics in dynamic general equilibrium models and estimating them using time series analysis can be applied in different settings and is an interesting direction for further research.

References

Auerbach, A., Kotlikoff, L., 1987. Dynamic fiscal policy. Cambridge University Press.

Breyer, F., Straub, M., 1993. Welfare effects of unfunded pension systems when labor supply is endogenous. Journal of Public Economics 50, 77–91.

Card, D., Chetty, R., Weber, A., 2007. Cash-on-hand and competing models of intertemporal behavior: New evidence from the labor market. Quarterly Journal of Economics 122(4), 1511–1560.

Chetty, R., 2008. Moral hazard vs. liquidity and optimal unemployment insurance. Journal of Political Economy 116, 173–234.

Chetty, R., 2009. Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods. Annual Review of Economics 1, 451–488.

Diamond, P., 1965. National debt in a neoclassical growth model. American Economic Review 41, 1126–1150.

Fanti, L., Spataro, L., 2006. Endogenous labor supply in Diamond's (1965) OLG model: A reconsideration of the debt role. Journal of Macroeconomics 28, 428–438.

Fehr, H., Kallweit, M., Kindermann, F., 2012. Pension reform with variable retirement age - a simulation analysis for Germany. Journal of Pension Economics and Finance 11(3), 389–417.

Feldstein, M., 1985. The optimal level of social security benefits. The Quarterly Journal of Economics 100(2), 303–320.

Feldstein, M., 2005. Structural reform of social security. The Journal of Economic Perspectives 19 (2), 33–55.

Gonzales-Eiras, M., Niepelt, D., 2007. The future of social security. Study Center Gerzensee Working Paper.

Gonzales-Eiras, M., Niepelt, D., 2008. The future of social security. Journal of Monetary Economics 55, 197–218.

Imrohoroglu, A., Imrohoroglu, S., Joines, D., 1995. A life cycle analysis of social security. Economic Theory 6, 83–114.

Imrohoroglu, S., Kitao, S., 2009. Labor supply elasticity and social security reform. Journal of Public Economics 93 (7-8), 867–878.

Kotlikoff, L., Smetters, K., Walliser, J., 1999. Privatizing social security in the United States - comparing the options. Review of Economic Dynamics 2, 532–574.

Kotlikoff, L., Smetters, K., Walliser, J., 2007. Mitigating America's demographic dilemma by pre-funding social security. Journal of Monetary Economics 54, 247–266.

Lopez-Garcia, M.-A., 2008. On the role of public debt in an OLG model with endogenous labor supply. Journal of Macroeconomics 30, 1323–1328.

McGrattan, E., Prescott, E., 2013. On financing retirement with an aging population. NBER working paper series, No. 18760.

Nickel, C., Rother, P., Theophilopoulou, A., 2008. Population ageing and public pension reform in a small open economy. European Central Bank Working Paper Series, No. 863.

Nourry, C., 2001. Stability of equilibria in the overlapping generations model with endogenous labor supply. Journal of Economic Dynamics & Control 25, 1647–1663.

Pecchenino, R. A., Utendorf, K. R., 1999. Social security, social welfare and the aging population. Journal of Population Economics 12 (4), pp. 607–623.

Song, Z., 2011. The dynamics of inequality and social security in general equilibrium. Review of Economic Dynamics 14, 613–635.