CESifo Working Papers

www.cesifo.org/wp

Pensions and Fertility: Back to the Roots

The Introduction of Bismarck's Pension Scheme and the European Fertility Decline

Robert Fenge
Beatrice Scheubel

CESIFO WORKING PAPER NO. 4383
CATEGORY 1: PUBLIC FINANCE
SEPTEMBER 2013

CESifo *Working Paper No. 4383*

Pensions and Fertility: Back to the Roots
The Introduction of Bismarck's Pension Scheme and the European Fertility Decline

Abstract

Fertility has long been declining in industrialised countries and the existence of public pension systems is considered as one of the causes. This paper is the first to provide detailed evidence based on historical data on the mechanism by which a public pension system depresses fertility. Our theoretical framework highlights that the effect of a public pension system on fertility works via the impact of contributions in such a system on disposable income as well as via the impact on future disposable income that is related to the internal rate of return of the pension system. Drawing on a unique historical data set which allows us to measure these variables a jurisdictional level for a time when comprehensive social security was introduced, we estimate the effects predicted by the model. We find that beyond a general depressing effect of social security on birth, a lower internal rate of return of the pension system is associated with a higher birth rate and a higher contribution rate is associated with a lower birth rate.

JEL-Code: C210, H310, H530, H550, J130, J180, J260, N330.

Keywords: public pension, fertility, transition theory, historical data, social security hypothesis.

Robert Fenge
Department of Economics
University of Rostock
Ulmenstraße 69
Germany – 18057 Rostock
robert.fenge@uni-rostock.de

*Beatrice Scheubel**
European Central Bank
Kaiserstraße 29
Germany – 60311 Frankfurt am Main
beatrice.scheubel@ecb.europa.eu

*corresponding author

This version: August 2013
We would like to thank Kathrin Weny for valuable research assistance. This paper should not be reported as representing the views of the European Central Bank (ECB). The views expressed are those of the author and do not necessarily reflect those of the ECB.

1 Introduction

In the mid to late nineteenth century, fertility in Europe began to drop and never rose again. As much as the exact definition of the onset of this decline is disputed, so are the causes for its persistence.[1] Regarding the definition of the onset of the fertility decline, Coale (1965) was one of the first researchers to observe that fertility would never rise again once it had declined by more than 10% from a previous plateau. Coale then heuristically defined the onset of the fertility decline as the point in time when fertility first declined by at least 10%. Regarding the causes, the Princeton European Fertility Project[2] concluded that innovations, e.g. in the area of birth control, and the diffusion of the new technologies caused the fertility decline rather than changed economic and social conditions. This is often termed the 'cultural diffusion hypothesis' or the 'Princeton View'.

Not surprisingly, the results of the Princeton European Fertility Project have been challenged, both on grounds of the quality of the data set (e.g. Galloway et al. 1994) and on grounds of the methodology (e.g. Richards 1977; Brown and Guinnane 2007; Goldstein and Klüsener 2010). Recently, the heterogeneity of the historical experience has been stressed, which also contradicts the Princeton View. For example, Hirschman (2001) notes that pre-decline fertility levels were much lower in Europe than in other regions of the world.

Instead, the effects predicted by economic theory (e.g. Becker 1960) have received more attention in the context of the first demographic transition.[3] These effects are also considered as the demand theory of fertility, according to which the marginal benefit of rearing a child in terms of intrinsic utility and the child's contribution to current and to future income have to be equal to the marginal cost, including the cost of child-rearing and the opportunity cost related to reduced income.

Among the economic explanations for the fertility decline, the reduced necessity for having children as a provision for old age has received surprisingly little attention. It is clear that people have children to provide for old age (e.g. Leibenstein 1957; Neher 1971; Nugent 1985; Cigno 1993). Social security contributions and social security benefits affect disposable income, and thereby life-cycle consumption smoothing. If they affect disposable income, and are moreover linked to labour income, it is natural to assume that the labour supply decision is affected as well. The few studies that analyse the connection between the generosity of the pension system and fertility find that a less generous pension system has positive effects on fertility (e.g. Billari and Galasso 2009).

In this paper, we provide more evidence on the pensions-fertility nexus in the histori-

[1] Cleland and Wilson (1987) give an overview of the debate in classic demographic transition theory and link this to early descriptive studies, inter alia of historical data. Arroyo and Zhang (1997) give a comprehensive overview of dynamic microeconomic models and the derivation of reduced-form models for estimation. Therefore they provide an important connection between theoretical advances and the empirical tests of the theories.

[2] Coale and Watkins 1986 provide a summary.

[3] Guinnane (2011) gives a nice summary on more recent empirical research on the historical fertility decline.

cal context. For one, the introduction of social security has only recently been considered as one of the causes of the first demographic transition (Guinnane 2011). For another, analysing the introduction of social security instead of changes in the configuration of the social security system facilitates the identification of the effect.

To show the effects of the introduction of social security on fertility, we first establish a simple theoretical framework on the pensions-fertility nexus and then provide evidence for the hypotheses derived from the model using historical data. To establish a theoretical framework, we construct a simple overlapping-generations model in the spirit of Cigno (1993) to show that the external provision of old-age income triggers a portfolio-rebalancing of individual investment. Thereby, our study also renders support to the social security hypothesis (Feldstein 1974). Depending on the internal rate of return of the pension system in relation to the rate of return (and accessibility) of capital markets, fertility can be negatively affected.

Since reliable demographic data combined with reliable data on social security is scarce for the late nineteenth and early twentieth century, we restrict our analysis to Imperial Germany, for which such data exist. Imperial Germany was the first European country that enacted an irreversible transition into a welfare state. The authorities collected information on several key variables of social insurance from the beginning. We explore the effect on aggregate fertility at the provincial level using a newly compiled data set of historical data.

This study shows that a lower internal rate of return is associated with a higher birth rate, while a higher level of contributions is associated with a lower birth rate. However, a change in the growth of the internal rate of return has much stronger effects than a change in the growth of contributions. Moreover, even after controlling for the economic determinants of fertility as well as the impact of pension insurance and a time trend, we find that about 25% of the decline between 1891 and 1914 took place during the late 1890s and the early 1900s, which is when pension insurance was introduced as the last element of comprehensive social insurance. Our results therefore also point to a general effect of social insurance on fertility that goes beyond pure consumption-related aspects.

Section 2 provides institutional details on Germany and social policy in the late nineteenth century. Section 3 then presents the theoretical model and section 4 derives the identification strategy from the theoretical framework, provides information on the data set as well as considerations on econometric issues. Section 5 presents a descriptive analysis and multivariate results as well as sensitivity analyses. Section 6 concludes.

2 Institutional Background

The introduction of comprehensive social insurance in Germany took place between 1883 and 1891. Health insurance was introduced in 1883 and accident insurance in 1884. The law on pension insurance was adopted in 1889 and came into force in 1891.

While Chancellor Otto von Bismarck was extensively involved in promoting the idea of comprehensive social insurance and pension insurance in the early 1880s, his role in developing the draft text on pension insurance was only marginal, even though he defended the draft in parliament in 1887 (von Bismarck 1894).[4] Pension insurance provided for so-called disability pensions and old-age pensions. Disability pensions were provided if a worker was unable to work because of physical conditions; old-age pensions were provided if a worker was unable to work because of age. Both disability pension and old-age pension were designed as a supplementary income that was paid when workers were unable to earn their income due to disability or when they were unable to make a living because of age. Neither the disability pension nor the old-age pension were designed as the only source of income, but as a supplement. A worker was expected to earn as much as he (physically) could.

The conditions to qualify for pension payments differed between disability and old age pensions. Workers could apply for a disability pension whenever they had paid contributions for at least 5 years. When they applied for a disability pension, the local authority could decide whether the worker was disabled on a case-by-case basis. Accordingly, the disability pension was effectively means-tested. If a worker reached the age of 70 – at a time when average life expectancy for a boy born in Prussia between 1865 and 1867 was 32.5 years (Marschalck 1984) and average life expectancy for a child born between 1881 and 1890 in Imperial Germany was 42.3 years (Marschalck 1984) – they automatically qualified for an old age pension.[5]

The pension system of 1891 was a partially mandatory, partially funded defined-contribution pension system. For workers in specific occupational categories with an annual income below 2000 Reichsmark pension insurance was mandatory; for people in other occupations it was voluntary (Verhandlungen des Reichstages 1887/88 and RGbl 1889/13). As a consequence, about 20-25% of the population were covered by pension insurance (Scheubel 2013a). Contrary to what is commonly believed, Bismarck's pension system was not what is understood to be a Bismarckian pension system today (Scheubel 2013b).[6] However, the pension system of the 1890s was neither a pure pay as you go pension scheme nor a fully-funded pension scheme. While the system was based on current contributions financing current pensions, it was also supposed to accumulate a capital stock. The set-up contained considerably more funded than pay-as-you-go elements. This set-up changed when the law was revised in 1899, coming into effect in 1900. The pension system became a fully-fledged pay as you go system (RGBl 1899/33). The pension level depended on contributions, such that the pension system can be classified as a defined-contribution system (Scheubel 2013b). Workers paid contributions according to income; there were four income categories. A fifth category was introduced with the revision of the law in 1899, which divided the previous category IV in two new categories. By paying contributions, workers could earn entitlement stamps, which would augment their basic (lump-sum) pension entitlement.[7]

[4] Bismarck had to leave the office as Chancellor in 1890, one year before the law on pension insurance came into force.

[5] After 1900 the definition of old age changed slightly and every worker who reached the age of 65 was automatically classified as disabled.

[6] The term 'Bismarckian pension system' is usually used to refer to a pay as you go type of pension system.

[7] The average old-age pension in Imperial Germany was 21.88% of the average annual wage in rail

The administration of the pension system was decentralised and administered by regional authorities, the so-called Regional Insurance Agencies (*Landesversicherungsanstalten*). These Regional Insurance Agencies already administered the health insurance system when pension insurance was introduced. They enjoyed discretion with regard to setting contribution rates within certain limits and to approving pension applications.

3 Theoretical analysis of effects of pension systems on fertility and savings

Microeconomic theories of fertility choice were developed by Becker and others (Becker 1960, 1965, 1988, 1991; Schultz 1969; Barro and Becker 1986, 1888, 1989; Easterlin 1975; Becker and Tomes 1976; Cigno and Ermisch 1989). These approaches to an (economic) theory of fertility are often referred to as the demand model of fertility, because children are modelled as a consumption good and fertility is considered as the demand for children. The marginal benefit of an additional child has to be equal to the marginal cost of rearing the child in equilibrium.

More recently, the microeconomic theories were related to economic growth (Barro and Becker 1989; Becker et al. 1990; Becker 1992). This provided the missing link between the microeconomic theories and the macroeconomic view on the fertility decline that was adopted by its early observers. The impact of institutions on fertility has also become the focus of economic research (e.g. McNicholl 1980; Becker and Murphy 1988; Smith 1989; Guinnane and Oglivie 2008). The impact of institutions has, however, not been discussed extensively in the context of the demographic transition in nineteenth century Europe. Guinnane (2011) goes into some detail with regard to considering children as a means for the provision for old age, and the existence of institutions and social security in particular as a possibility to substitute away from this.

We discuss several possible channels how the introduction or extension of a pension system may affect the fertility and savings decisions of the population. For this we use a simple two-period overlapping generations model which combines three options to provide for old age: private savings, an intra-family transfer from children to parents when they no longer work and a public pension system. We analyse two types of the public pension systems. The first type is a fully-funded system in which the pensions are financed by the accumulated capital out of the savings that the government enforces. This is a compulsory savings system. The second type is a pay-as-you-go (PAYG) pension system in which the working generations finance the pensions of the retired generations by their contributions in the same period. In particular, we investigate a PAYG pension

track supervision and maintenance and the average disability pension was 21.36% of the average annual wage in that sector (Lotz 1905). Both the pension level and the average annual wage in rail track supervision and maintenance differed across regions. The average old-age pension in that sector in Baden was 18.81% of the local average annual wage and the average disability pension was 18.49% of the local average annual wage. After 30 to 50 years of contribution, this fraction could increase to about half of a worker's wage in the lowest category and to about 40% of a worker's wage in the middle category (Reichsversicherungsamt 1910).

system with pensions of a generation which are proportional to the their contributions.

3.1 The Model

We consider the impact of a pension system on fertility and savings in a two-period overlapping generations model (similar to Fenge and Meier, 2005). In period t the size of the working population is N_t. By convention, we denote the working generation in period t as generation t. The growth of population is given by the factor $\frac{N_{t+1}}{N_t} = 1 + \overline{n_{t+1}}$. We analyse the decisions of a household on the number of children n_t and savings s_t in period t. Note that the number of children of an atomistic household has no effect on population growth. The number of children in a family and the growth rate of the population only coincide in equilibrium, since all households are identical.

In the first period the labour supply of the household depends on the number of children. Children reduce the time available for labour.[8] Normalising total time to unity, working time is given by $1 - f(n_t)$ with $f'(n_t) > 0$ and $f''(n_t) \geq 0$. Hence, the time demand of a child increases with the number of children.[9] The wage rate is w_t. The household pays contributions from wage income at the rate τ into the pension system. We assume the contribution rate to be constant. The direct cost of raising a child is π_t. Furthermore, we consider an intra-family old-age provision from the children to the parents. Each grown-up child pays a transfer B_t in her working period to the parents in retirement. Young children participate in consumption c_t in the first period, which is determined by the following budget constraint:

$$c_t = w_t(1 - f(n_t))(1 - \tau) - s_t - \pi_t n_t - B_t. \tag{1}$$

In the second period the household retires and consumes z_{t+1}. Old-age consumption can be financed via the pension p_{t+1}, the returns on savings with interest factor $1+r_{t+1} = R_{t+1}$ and the intra-family transfer. The budget constraint in the second period is:

$$z_{t+1} = p_{t+1} + R_{t+1}s_t + B_{t+1}n_t. \tag{2}$$

The utility of the household depends on consumption in both periods and the individual number of children. The function $U(c_t, z_{t+1}, n_t)$ is increasing in all three arguments, strictly concave and additively separable: $U_{cz} = U_{cn} = U_{zn} = 0$. Since fertility enters the utility function, having children is induced by a consumption motive. The consumption motive is a way of modelling the intrinsic motivation to have children. Furthermore, children provide a transfer to their parents in old-age, which constitutes an investment motive for children. This investment motive is important to create a model set-up which corresponds to the set-up of pension insurance in Imperial Germany. During the first ten

[8]Note that this assumption can be relaxed. It does, however, correspond to the fact that at the time when the pension system was introduced, unmarried women were supposed to be working, while married women were still supposed to stay at home and care for the children (Kohl 1894).

[9]Note that this assumption can easily be relaxed by e. g. assuming a u-shaped time cost of children. This would imply that with a certain number of children the cost of rearing each single one diminishes, because the older children can care for the younger children.

years, the pension system set-up could be considered fully funded, such that we expect behavioural effects via the reduced importance of the transfer channel only between 1891 and 1900. We present our theoretical considerations on the behavioural effect of the transfer channel in section 6.2.

The household determines the number of children and savings by maximising utility subject to the budget constraints (1) and (2). Substituting these constraints for the consumption variables in the utility function results in a maximisation problem of a function depending on n_t and s_t:

$$\max_{n,s_t} V(n_t, s_t) = U(w_t(1-f(n_t))(1-\tau) - s_t - \pi n_t - B_t, p_{t+1} + R_{t+1}s_t + B_{t+1}n_t, n_t). \quad (3)$$

This is the key equation for the empirical identification of an effect.

The pension is affected by fertility via the pension budget constraint as becomes clear in the next section. Hence, we can write the first-order conditions of the maximisation problem as:

$$V_n = -U_c((1-\tau)w_t f'(n_t) + \pi_t) + U_z \left(\frac{\partial p_{t+1}}{\partial n_t} + B_{t+1} \right) + U_n = 0 \quad (4)$$

and

$$V_s = -U_c + U_z R_{t+1} = 0. \quad (5)$$

The second-order conditions for a maximum are satisfied (see Appendix A).

In the following we analyse the impact of a higher contribution rate on fertility and savings for a pay-as-you-go and a fully-funded pension system. The fertility effect is given by:

$$\frac{\partial n}{\partial \tau} = -\frac{V_{n\tau}V_{ss} - V_{ns}V_{s\tau}}{V_{nn}V_{ss} - V_{ns}V_{sn}} \quad (6)$$

3.2 Fertility effect in a pay-as-you-go pension

In a pay-as-you-go (PAYG) system pensions of generation t are financed by the contributions of generation $t+1$. If the PAYG pension is of the Bismarckian type the individual pension is identical to the average pension weighted by an individual factor which relates the individual pension contribution payment of a household of generation t to the generation's average[10]:

$$p_{t+1}^{BIS} = (1 + \overline{n_{t+1}})\tau w_{t+1}(1 - f(\overline{n_{t+1}}))\frac{\tau w_t\,(1 - f(n_t))}{\tau w_t\,(1 - f(\overline{n_t}))}, \quad (7)$$

where $(1-f(\overline{n_t}))$ denotes the average labour supply of generation t and the growth factor of the population, $1+\overline{n_{t+1}} = \frac{N_{t+1}}{N_t}$, is equal to the average number of children of generation t. If the individual contribution, $\tau w_t\,(1-f(n_t))$, is above average, $\tau w_t\,(1-f(\overline{n_t}))$, the

[10]The pension system that was introduced by Bismarck was very similar to the institutional setting in Germany today. As a main feature, current pension claims were paid from current contributions. See also section 2.

individual pension, p^{BIS}_{t+1}, is higher than the average pension, $(1+\overline{n_{t+1}})\tau w_{t+1}(1-f(\overline{n_{t+1}}))$, by the same proportion. Since the wage rate and the contribution rate are identical for all households we may write the proportionality factor as $\frac{1-f(n_t)}{1-f(\overline{n_t})}$ and call it the Bismarck factor.

In the Bismarckian case a higher number of children reduces the pension claims proportional to the payroll growth factor $(1+\overline{n_{t+1}})\frac{w_{t+1}}{w_t}\frac{1-f(\overline{n_{t+1}})}{1-f(\overline{n_t})}$:

$$\frac{\partial p^{BIS}_{t+1}}{\partial n_t} = -(1+\overline{n_{t+1}})\tau w_t f'(n_t)\frac{w_{t+1}}{w_t}\frac{1-f(\overline{n_{t+1}})}{1-f(\overline{n_t})} < 0 \tag{8}$$

Second period consumption is given by

$$z_{t+1} = (1+\overline{n_{t+1}})\tau w_{t+1}(1-f(\overline{n_{t+1}}))\frac{1-f(n_t)}{1-f(\overline{n_t})} + R_{t+1}s_t + B_{t+1}n_t. \tag{9}$$

and the intertemporal budget by:

$$\begin{aligned} R_{t+1}c_t + z_{t+1} &= R_{t+1}\left[(1-\tau)\, w_t(1-f(n_t)) - \pi_t n_t - B_t\right] \\ &\quad +(1+\overline{n_{t+1}})\frac{w_{t+1}(1-f(\overline{n_{t+1}}))}{w_t(1-f(\overline{n_t}))}\tau w_t(1-f(n_t)) + B_{t+1}n_t. \end{aligned} \tag{10}$$

The marginal price of children in present value terms of period $t+1$ is

$$\begin{aligned} \Pi^{BIS}_{t+1} &= R_{t+1}((1-\tau)w_t f'(n_t) + \pi_t) \\ &\quad +(1+\overline{n_{t+1}})\frac{w_{t+1}(1-f(\overline{n_{t+1}}))}{w_t(1-f(\overline{n_t}))}\tau w_t f'(n_t) - B_{t+1} \end{aligned} \tag{11}$$

We assume this marginal price to be positive at an inner solution of fertility decision.

In equilibrium, the average population growth factor is identical to individual fertility: $\overline{n_t} = n_t$ and, hence, average labour supply is identical to individual labour supply: $1-f(\overline{n_t}) = 1-f(n_t)$ in the case of homogeneous households. In what follows we denote the internal rate of return of contributions to the PAYG pensions system in equilibrium by

$$\Omega_{t+1} \equiv p_{t+1}/\tau w_t\left(1-f\left(n_t\right)\right). \tag{12}$$

In the case of constant contribution rates this is equal to the payroll growth factor:

$$\Omega_{t+1} = (1+\overline{n_{t+1}})\,\frac{w_{t+1}}{w_t}\frac{1-f(\overline{n_{t+1}})}{1-f(\overline{n_t})}. \tag{13}$$

In the following analysis of the fertility effect the relation between this internal rate of return and the capital market interest rate will be crucial. A steady state equilibrium which is dynamically efficient satisfies $R_{t+1} > \Omega_{t+1}\ \forall\ t$. In transitional periods of the economy the relation may be reversed. However, $R_{t+1} > \Omega_{t+1}$ can be justified for the period from 1878 to 1914 since the economy moved out of mildly deflationary environment i n approximately 1895 (Wehler 2008).

Now we consider the fertility decision in a PAYG pensions system of the Bismarckian type. In order to calculate the sign of the numerator of (6) we need the second derivatives of utility with respect to the contribution rate:

$$\begin{aligned} V_{n\tau} &= w_t f'(n_t) U_z (R_{t+1} - \Omega_{t+1}) + w_t (1 - f(n_t)) \\ & \left[U_{cc}((1-\tau) w_t f'(n_t) + \pi_t) + U_{zz} \left(B_{t+1} - \Omega_{t+1} \tau w_t f'(n_t) \right) \Omega_{t+1} \right] \end{aligned} \tag{14}$$

and

$$V_{s\tau} = w_t (1 - f(n_t)) [U_{cc} + U_{zz} \Omega_{t+1} R_{t+1}] < 0 \tag{15}$$

The numerator of equation (6) can be calculated as:

$$\begin{aligned} V_{n\tau} V_{ss} - V_{ns} V_{s\tau} &= (R_{t+1} - \Omega_{t+1}) \left[w_t f'(n_t) U_z (U_{cc} + U_{zz} R^2_{t+1}) \right. \\ & + w_t (1 - f(n_t)) U_{cc} U_{zz} \\ & \left. \left(R_{t+1}((1-\tau) w_t f'(n_t) + \pi_t) - \left(B_{t+1} - \Omega_{t+1} \tau w_t f'(n_t) \right) \right) \right] \end{aligned} \tag{16}$$

The sign of the numerator is ambiguous and we have to consider the separate effects in turn. Using (13), the marginal price of children from equation (11) can be written as $R_{t+1}((1-\tau) w_t f'(n_t) + \pi_t) - (B_{t+1} - \Omega_{t+1} \tau w_t f'(n_t))$ which is positive.

The price effect: Increasing the contribution rate reduces the opportunity cost of having children in terms of foregone lifetime income. A higher contribution rate reduces the net wage income in the first period so that the opportunity cost of a child is reduced by $w_t f'(n_t)$. Moreover, a higher contribution rate increases the pension entitlement in the second period. This implies that the reduction of the Bismarck pension due to another child increases. This increase of the opportunity cost of a child in the second period is expressed by $\frac{\Omega_{t+1}}{R_{t+1}} w_t f'(n_t)$ in present values of period t. Thus, a higher contribution rate lowers the opportunity cost of having a child in the first period, but increases the opportunity cost of having a child in the second period in terms of pension entitlements. In a dynamically efficient economy, the total opportunity cost falls. Partial derivation of (11) with respect to τ shows that the price of a child decreases with a higher contribution rate,

$$\frac{\partial \Pi^{BIS}_{t+1}}{\partial \tau} = - \left(R_{t+1} - \Omega_{t+1} \right) w_t f'(n_t) < 0, \tag{17}$$

Since children become relatively cheaper than savings as a provision for old-age, more children are substituted against less savings which increases consumption and utility in the first period. The number of children increases at the expense of savings[11].

The income effect: By using the definition of the payroll growth factor (13) the lifetime budget constraint (10) can be written as:

$$R_{t+1} c_t + z_{t+1} = w_t (1 - f(n_t)) \left[R_{t+1} - \tau \left(R_{t+1} - \Omega_{t+1} \right) \right] - \left(R_{t+1} \pi_t - B_{t+1} \right) n_t \tag{18}$$

The derivation of the RHS of (18) with respect to τ shows that a higher contribution rate reduces lifetime income by

$$\left(R_{t+1} - \Omega_{t+1} \right) w_t (1 - f(n_t)).$$

[11] The formal treatment of the savings decision can be found in the Appendix A2.

The reason is that the PAYG pension system incurs a implicit tax on wage income. In a dynamically efficient equilibrium, i.e. $R_{t+1} > \Omega_{t+1}\ \forall\ t$, compulsory contributions mean a loss in lifetime income, because compulsory contributions to the pension system mean a loss in lifetime income since investing the same amount of contributions in the capital market instead would yield a higher rate of return. The lower rate of return in the pension system implies that the Bismarck pension system involves an implicit wage tax, $\tau\left(R_{t+1} - \Omega_{t+1}\right) > 0$ (e.g. Barro and Becker 1988; Sinn 2000, 2004b). A higher contribution rate increases this implicit tax and reduces lifetime income. With normal goods, consumption in both periods is reduced. The reduction of lifetime income is partially compensated by decreasing the number of children. Each child less lowers the reduction by its price $\Pi_{t+1}^{BIS} = R_{t+1}((1-\tau)w_t f'(n_t) + \pi_t) - (B_{t+1} - \Omega_{t+1}\tau w_t f'(n_t)) > 0$. Hence, due to the income effect fertility decreases with rising contribution rates.

Proposition 1 ***Price effect and income effect*** *If the price effect overcompensates the income effect, the overall effect of a PAYG pension system on fertility is positive. If the income effect overcompensates the price effect, the overall effect of a PAYG pension system on fertility is negative.*

The price effect and the income effect depend on the opportunity cost of having children and thus on the internal rate of return of the pension system $\Omega_{t+1} \equiv p_{t+1}/\tau w_t\left(1 - f\left(n_t\right)\right)$. *If the internal rate of return of the pension system increases (decreases), the price effect is lower (higher) and the income effect overcompensates (is lower than) the price effect: fertility falls (rises).*

Furthermore, we can show that savings are a partial substitute to children under the following conditions on the net return of children. On the one hand assume the intra-family transfer of children in the second period is higher than the cost of children due to the reduced Bismarckian pension. Then having more children would increase the consumption in the second period. If, on the other hand, the discounted intra-family transfer is lower than the cost of children in the first period, a higher number of children decreases consumption in the first period. Smoothing the consumption profile leads to a reduction of savings. Combining both effects implies that savings are substituted for a higher number of children. For details see the analysis in Appendix A. Hence, if higher contribution rates increase fertility, the effect of the Bismarck pension system on savings is negative.

Proposition 2 ***Crowding out of savings in a PAYG system*** *Savings may be partially crowded out depending on the relative return of the pension system in relation to capital market savings and in relation to children.*

Thus we can summarise the findings in our main hypothesis:

> Hypothesis 1: Fertility effect in a pay-as-you-go Bismarckian pension system
>
> In a dynamically efficient economy the introduction or expansion of a pay-as-you-go public pension scheme of the Bismarck type sets incentives to reduce

(increase) the number of children if the income effect is higher (lower) than the price effect on fertility. The relation between these effects is determined by the internal rate of return of the pension system.

4 Data, identification strategy and econometric considerations

4.1 Data

Showing the impact of social insurance and, in particular, pension insurance on fertility for the late nineteenth century requires reliable data. Our empirical analysis is based on a regional data set for Imperial Germany which we combined from two primary data sources, the Imperial Annual Yearbook of Statistics and the Annual Reports of the Regional Insurance Agencies, which were collected by Kaschke and Sniegs (2001). We also use data from Mombert (1909) matched with these data sources for some analyses. The regional entities had to be made consistent, because data in the Annual Yearbook of Statistics has been collected at the state level, while one Regional Insurance Agency could cover a region larger than a state, or if a state was large (like Bayern) there could be more than one Regional Insurance Agency in that state. Figure 1 shows the regional entities after harmonising the data sets. We use the German names for these regions, because for some regional names there is no English equivalent. However, when we refer to a broad region, like for example the Kingdom of Prussia, we use the English names. Therefore, as a rule, when we use English names we refer to a region, while when we use German names, we refer to a unit of observation.

[Figure 1 about here.]

Measuring fertility in the historical context is as complex as finding a suitable data set. Individual-specific measures which are common in event-history analysis like the individual birth history of a woman or a household cannot be derived from historical data. Individual-level data is hardly available. Even detailed measures of fertility are difficult to derive for the aggregate population, because measures like the total fertility rate (TFR)[12] require cohort-specific fertility rates, and the size of each cohort. Information on the age structure of the population is scarce for late nineteenth century Germany. If it is available, it is only available for census years. Information on crude birth rates is, however, available annually. Therefore, we can construct the crude birth rate (CBR) and also the crude marital birth rate (CMBR) as number of births per mill and number of marital births per mill respectively, because German authorities also collected data on illegitimacy rates.[13]

[12] The TFR is defined as $TFR_t = \sum_{age=15}^{age=49} \frac{(BIRTHS_t^{age})}{WOMEN_t^{age} \cdot 1000}$. That is to say, the TFR in year t is equal to the sum of all cohort-specific birth rates in year t.

[13] The regional distribution of the CBR in Imperial Germany corresponds to the regional distribution of the total fertility index in those years for which we can compute the total fertility index. The information is broadly in line with the information in Knodel (1974).

More importantly, however, some covariates are not as frequently available as e.g. the birth rate.[14] We account for this fact in our econometric approach, detailed in section 4.3.

4.2 Identification strategy

Our theoretical model gives us an indication how to best identify an effect of compulsory saving in a public pension system on the number of children. According to our model, a household takes a simultaneous decision on the number of children and the amount of capital market savings, depending on the amount that has to be contributed to the public pension system. As depicted in equation (3), the utility function's three elements consist of the utility from consumption in the current period, utility after retiring and utility from having children. In translating this into an econometric model, we have to consider first, the form of the utility function, second, the determinants of the utility function which should also enter the econometric model and third, the fact that we have data available at the jurisdictional level, while we have written down our model at the household level.

First, consider the form of the utility function. Since we assume it to be additively separable in its arguments, we can also write our econometric model such that the main determinants of the fertility decision enter additively. Thus, our multivariate model links the number of children to its main determinants additively:

$$y_{HH,t} = y_{0_{H}H} + \boldsymbol{U}_{\mathbf{cHH,t}}\boldsymbol{\beta}_c + \boldsymbol{U}_{\boldsymbol{z}\mathbf{HH,t}}\boldsymbol{\beta}_z + \varepsilon_{HH,t}.$$

The number of children in a household y_{HH} depends on the utility from current consumption $\boldsymbol{U}_{\mathbf{cHH}}$ and on the utility from future consumption $\boldsymbol{U}_{\boldsymbol{z}\mathbf{HH}}$. Note that if we assume that the utility from the pleasure of having children is the same for all households, it enters the intercept $y_{0_{H}H}$ as depicted here. However, we can also assume this utility to be household-specific and would then have to include household-specific dummies in the regression. $\varepsilon_{HH,t}$ is an i.i.d. error term.

Second, consider the determinants of fertility that should be included. According to our model, for capturing current consumption, we should include a measure of disposable labour income, represented by $(1-\tau)(w_t(1-f(n_t)))$, a measure of the cost of children, represented by $\pi_t n_t$, and a measure of the intra-family transfer B_t which has to be paid to the parent generation. For capturing future consumption, we should include measures of the pension level p_{t+1}, of the amount of savings $R_{t+1}s_t$ and the intra-generational transfer B_{t+1}. We modify the representation of our econometric model as $y_{HH,t} = y_{HH} + \boldsymbol{x}_{\mathbf{HH,t}}\boldsymbol{\beta}_x + \boldsymbol{p}_{\mathbf{HH,t}}\boldsymbol{\beta}_p + \varepsilon_{HH,t}$ where $\boldsymbol{x}_{\mathbf{HH,t}}$ represents the vector of variables for current consumption and $\boldsymbol{p}_{\mathbf{HH,t}}$ represents the vector of variables for future consumption.

Note that it is not clear ex ante where to best include a measure of savings. Our model illustrates that not accounting for the fact that children and savings can be substitutes to a certain extent can lead to biased estimates. The pension system's crowding out of fertility may only appear once a crowding out of savings has taken place. In practice,

[14] Data availability is detailed in Annex B.

such interactions would require a simultaneous equations estimation approach, which we cannot pursue since there is no reliable time-varying measure of savings. However, note that according to equations (4) the optimum number of children is determined by $U_z\left(\frac{\partial p_{t+1}}{\partial n_t}+B_{t+1}\right)+U_n-U_c\left[(1-\tau)w_tf'(n_t)+\pi_t\right]$. At the same time, we know from equation (5) that $U_c=U_zR_{t+1}$ in the optimum. In other words, we only have to make sure that we include a measure of how current consumption can be traded off against future consumption in the optimum. In the historical context, that requires a measure of the capital market rate of return, but also the diffusion of private saving opportunities across households. We therefore include the productivity in agriculture, which should also be linked to the return on capital, in our regressions as well as the diffusion of savings books in 1900.

Third, consider the availability of the data only at the jurisdictional level. Since our model is constructed for identical agents, it is even better suited to describe the average reaction of the population in a certain jurisdiction rather than a household's fertility decision since we assume that the labour participation rate is reduced by having children. While this is not necessarily applicable to a household of two, it is certainly true in aggregate. Therefore, we rewrite our model to represent the jurisdictional structure of the data:

$$y_{i,t,g}=y_{0_i}+T_t+\boldsymbol{x}_{\mathbf{i},\mathbf{g},\mathbf{t}}\boldsymbol{\beta}_x+\boldsymbol{p}_{\mathbf{i},\mathbf{g},\mathbf{t}}\boldsymbol{\beta}_p+\alpha_{i,g}+\varepsilon_{i,g,t}. \tag{19}$$

The measure $y_{i,t,g}$ refers to the crude marital birth rate (CMBR) in jurisdiction i in year t in jurisdictional cluster g; $\alpha_{i,g}$ refers to cluster-specific effects in cluster g and $\varepsilon_{i,g,t}$ is an i.i.d. error term.

As control variables for current consumption $\boldsymbol{x}_{\mathbf{i},\mathbf{g},\mathbf{t}}$, we include information on the number of stillbirths and the number of marriages as proxies for the cost of having children; population density, the share of the population working in the primary, secondary and tertiary sector as well as the share of illiterate recruits, the number of horses per 1000 inhabitants and the change in the number of large cities as proxies for average income in a jurisdiction. Moreover, we include the share of contributions in category I and the relative share of contributions in category I relative to category IV or V to proxy the share of working women in a jurisdiction.[15] To proxy the size and likelihood of the payment of an intra-generational transfer, we use data on the number of persons per household. Since the number of persons in a household is of course endogenous to current family size, we use this variable with a 10-years lag. Note that these variables also correspond to the main determinants previously identified as the key drivers of the first demographic transition in the literature. We provide the summary statistics for these variables in Table 1 in Appendix B.

In our model, the sign of the fertility effect in equation (4), i.e. of the pension system on fertility $\frac{\partial n}{\partial \tau}$ depends on the determinants in equation (16): the reduction in the opportunity cost of having children in terms of foregone lifetime income (which

[15] Category I was considered the women's category since only very low-paying jobs would be included in this category. This was one of the reasons why there was not separate category for women. In the same vein, it is reasonable to assume that there were no women contributing in the higher wage categories.

we termed the price effect) versus the total reduction in lifetime income depending on internal rate of return of the pension system (which is the income effect). To capture these effects, we make use of the rich data set on pension insurance variables. To capture the price effect, we include the share of contributions in each wage category, the amount of contributions in each wage category, and the change in the latter. To capture the income effect, we include administrative costs, revenues from other sources than contributions, expenditures for pensions, assets, expenditures for other purposes. Therefore, the effect of the pension system on fertility, $\frac{\partial n}{\partial \tau}$, is not captured by a single coefficient, but rather represented by the vector of coefficients $\boldsymbol{\beta}_p$.

Of course, one may argue that it is not possible to control for all determinants of the birth rate at such a high level of aggregation, especially when the quality of the proxies is rather poor. To account for this concern, we do not estimate a simple multivariate model, but we augment our model with another term that captures the residual effect of the introduction of the pension system. To do so, we make use of a feature of the pension system that makes the set-up similar to a natural experiment.

As part of the particular features of the pension system, only workers with an annual income equal to or below 2000 Mark were compulsorily insured. The share of workers that were covered by pension insurance therefore varied across provinces.As the data collected by the Regional Insurance Agencies includes a variable on coverage by province, this variable essentially measures the overall impact of pension insurance in each province. Unfortunately, this variable is only available for the two years of the population census: 1895 and 1907. Therefore, we can only include it as an additional explanatory variable in a multivariate regression on a panel with N=25 and T=2. While doing this to establish the importance of the variable as a determinant of fertility, we cannot add many control variables due to the small number of degrees of freedom. To establish more robust results, we instead compare the change in fertility after the introduction of pension insurance in provinces with a large share of the population insured and in provinces with a small share of the population insured. This approach is very similar to the well-known difference-in-differences estimation approach used for natural experiments.

4.3 Econometric considerations

The fact that we would like to control for the determinants of fertility while also including pension system variables in our model poses practical challenges since in historical data, only few variables are available for all years. For example, demographic information was mostly collected in connection with the occupational censuses, which were only conducted every five years.[16] Therefore, we resort to a change and level model, in which we can include all years of available data for those variables for which we have more than one year of observations while at the same time being able to include those variables for which we only have one or two years of observations.

[16] As we adjust most variables to population size to make the numbers comparable, we extrapolate population figures for the years for which population figures are not available. Annex B details how we derived the extrapolated numbers.

Since we have a panel of provinces over time, it is of the essence to account for province-specific effects. The usual easy way to circumvent biased estimates due to unobserved province-specific effects is a fixed-effects OLS estimator. In order to apply a such a fixed effects OLS estimator, we transform the information on the share of insured into a binary variable. We use the moments of the distribution of these shares and we define provinces with a share of insured that is one standard deviation above the mean as 'treated'. These provinces are identified by the variable $D_{g,t}$, which identifies a province to be in the 'treated' group g in time period t. $g = 1$ if a province is defined as 'treated', and $g = 0$ otherwise. An dummy variable $DD_{g,t}$ indicates the interaction between the group variable and a term identifying years of 'treatment'.

However, we cannot use this estimator on a change and level model as the fixed effects would automatically capture the time-invariant variables. The opportunity of differencing out province-specific effects by using first differences (FD) remains, even though we lose a degree of freedom by using this technique:

$$\begin{aligned} \Delta y_{i,t,g} &= y_{0_i} + T_t \\ &\quad + \boldsymbol{\Delta x^c}_{\mathbf{i,g,t}} \boldsymbol{\beta}_{x^c} + \boldsymbol{x^l}_{\mathbf{i,g,t}} \boldsymbol{\beta}_{x^l} + \boldsymbol{\Delta p^c}_{\mathbf{i,g,t}} \boldsymbol{\beta}_{p^c} + \boldsymbol{p^l}_{\mathbf{i,g,t}} \boldsymbol{\beta}_{p^l} + \alpha_{i,g} + \varepsilon_{i,g,t}, \end{aligned} \tag{20}$$

where the superscript c denotes the variables that enter the regression in first differences (i.e. the stillbirth change year on year, marriages, and the index of agricultural productivity, the spatial lag, the share of contributions in each contribution category and the measures for the internal rate of return of the pension system) and the superscript l denotes the variables that enter the regression in levels (i.e. the number of savings banks books in 1900, the share of Catholics, the share in farming, trade, and mining, the number of persons per household and the increase in localities larger than 20.000 inhabitants). T_t denotes a set of time dummies. Note that in a model in first differences, the interaction terms previously referred to as $DD_{g,t}$ are in fact captured by the time dummies since we estimate a model already in differences.

It is surprising that in the models on the European fertility decline it is not very common to use the lagged version of the explanatory variables. Although we lose a whole period of observations, this is closer to the microeconomic foundations of fertility models. In general, given the time lag of at least nine months between the decision to have a child and the observation of a birth, the variables from the previous year are likely to play a more important role. However, we cannot rule out that factors in year t played a role for births late in t, we include both current explanatory variables and the first lag in the set of covariates where applicable.

Errors can be correlated across adjacent provinces, also known as spatial correlation. This may significantly affect both estimated coefficients and the corresponding standard errors. We follow Goldstein and Klüsener (2011) to account for that by including a spatial lag. We calculate the average birth rate over all adjacent provinces and include this as an explanatory variable, and in a different approach the birth rate of the neighbouring province with the highest birth rate.

Since migration can be an important factor in driving such spatial correlations in the birth rate, we include the sex imbalances ratio as a good proxy for migration (e.g. Knodel

1974). In addition, we calculate the share of foreign-born persons in each province.[17] Then we derive the province of origin for the majority of foreign-born persons. We use the birth rate in this province as an additional spatial lag that helps to take into account migration.

5 Results

5.1 Descriptive Analysis

To establish a negative relationship between pension insurance coverage and fertility, we relate the share of insured persons in each province to the crude marital birth rate in that province. In 1895, 49.26% of the working population were insured and 51.35% of the working population were insured in 1907. Based on the methodology described in Verhandlungen des Reichstages (1888), we calculate a measure of insurance coverage based on the 1882 occupational census which represents the number of people potentially covered by pension insurance before it was introduced.

Figure 2 shows the correlation between the CMBR and the share of insured persons for the 1882 projection and for the 1895 and 1907 data. The left panel shows the correlation with the 1892 CMBR and the right panel shows the correlation with the 1907 CMBR. We do not find and do not expect a relationship between the measures and the 1892 CMBR shown in the left panel, because the pension system only came into effect in 1891 and the introduction of comprehensive social insurance was only started gradually in 1882.

The right panel shows that the projected share insured based on 1882 census numbers is weakly inversely related to the birth rate in 1907. As the 1882 approximation reflects the sectoral composition of a province's workforce rather than the insured population, finding only a weak correlation is reassuring regarding the concern that the share insured might measure sectoral composition rather than insurance coverage. We examine this argument further below. The negative relationship between the 1895 share of insured persons and the 1907 CMBR is more pronounced. The negative relationship is also strong for the 1907 share of insured persons and the 1907 CMBR.

[Figure 2 about here.]

Some simple OLS regressions substantiate the finding that a more strongly declining marital birth rate is significantly correlated with a higher share of the population enrolled in social insurance.

[Table 1 about here.]

Table 1 shows three specifications. Column (1) gives a raw correlation of the CMBR and the share of insured persons. Note that we can use all observations from both years, 1895 and 1907 for this exercise. The correlation is highly significant. When including

[17] We only use information on *internal* migration, i.e. the resident population born in another province of Imperial Germany.

both a set of economic determinants of fertility and a selection of pension system variables in column (2) the coefficient on the share of insured persons becomes much smaller, but is still highly significant. Even when adding information on sectors of the economy, religion, and saving in column (3), the coefficient remains significant. Thus, the inverse relationship holds even when we account for other confounding factors.

The time series plots in figure 3 help to examine whether there was a noticeable difference in the CMBR between provinces with a high share of the population enrolled in statutory social insurance. As 1891 and 1900 were the years of most important changes. These are shown by the vertical red lines in all panels. Diverging trends can be traced for provinces in the same region which displayed the same CMBR before 1891. For example, the CMBR in the lower left panel of figure 3 was very similar for Schlesien and Ostpreußen prior to 1893, but after that year the regions show a significant divergence in the CMBR. In the upper right panel, the same phenomenon can be observed for Hannover and Oldenburg. Here, the divergence starts as early as 1893. This is even more remarkable considering that the mid-1890s were the start of an economic boom period. A boom period should increase income and thus relax the budget constraint of the household, which should actually be an incentive to have more children. However, while in Oldenburg, only 157 of 1000 were covered by social insurance, 194 of 1000 were covered in Hannover. We can observe the same phenomenon with other provinces that shared a boundary and displayed the same CMBR before the mid-1890s, but a diverging CMBR after the mid-1890s. In the lower left panel, the difference is most apparent for Pommern and Brandenburg. Pommern had 212 of 1000 inhabitants insured, while Brandenburg had 227 of 1000 inhabitants insured. The fact of diverging trends between provinces that were comparable in the dimension of the CMBR, but not in the dimension of insurance coverage renders support to the hypothesis that the differences were driven by social insurance.

[Figure 3 about here.]

5.2 Multivariate Analysis

To establish our results, we compare two different periods. First we look at the period between 1891 and 1899. This period covers the first years of statutory pension insurance during which the system was a partially funded system. Second, we look at the period after 1900. This period covers the years of major amendments to pension insurance before World War I, including the move towards a pure pay as you go system. The results for the period 1891-1899 are shown in table 2 and the results for the period 1900-1914 are shown in table 3.

[Table 2 about here.]

[Table 3 about here.]

In establishing the effect of pension insurance on fertility, we follow the identification strategy derived from the theoretical model and described in section 4.2. The first column in both tables presents the basic model, in which we have included the determinants of current consumption in first differences. Moreover, this first specification also includes year dummies besides a time trend, which in a first differences model are equivalent to the interaction term in a differences-in-differences model. In column (2), we augment the model with a level term on the share of savings books in a jurisdiction in 1900. This proxy is supposed to control the crowding out effect on savings. Since we only have this information for 1900, we include the same proxy for both time periods. In column (3), we include the other level terms which are additional controls for current consumption.

To test HYPOTHESIS 1, we present two methods of capturing the internal rate of return of the pension system in columns (4) and (5). In column (4) we add the share of pension contributions in each wage category, which were proportional to the pensions. If the income effect is larger than the price effect, our model suggests that fertility should decrease with higher contribution rates. Therefore, a higher share of contributions in category III or IV/V should lead to a lower birth rate. If the price effect is higher than the income effect, the share of contributions in category III or IV/V should lead to a higher birth rate. Since these effects are likely to be confounded by the share of working women in each jurisdiction, we control an additional measure of the share of contributions in category I relative to all other categories, since category I was considered the women's category (Haerendel 2001).

In column (5) we present an alternative option to test hypothesis HYPOTHESIS 1 by including a proxy for the internal rate of return of the pension system. We calculate the number of approved pension applications (by pension type) over revenues from contributions, which gives the number of approved pension applications per Mark collected in revenues. The more pension applications get approved in relation to revenues, the lower the internal rate of return of the pension system and the higher should be the birth rate. To make sure that the variable reflects expectations about the future return of the pension system (i.e. the degree to which current revenues have to be used for current expenditures), we add a variable on current pension payments to control for that.

Regarding the baseline model in columns (1) of both tables, the number of stillbirths and the number of marriages are significant determinants of the birth rate. In particular the effect of the number of stillbirths, though smaller between 1900-1914 than between 1891-1899, is highly stable across all specifications. This is similar for the time dummies, which in total amount to a total reduction of approximately 1.8 births per mill between 1891–1899 and a total reduction of approximately 4 births per mill between 1900 and 1914. Since the average birth rate per mill ranged from almost 40 births per mill to around 25 births per mill across provinces in the late 1880s, this is a total reduction of around 25%. Note that these effects measure the differences between the differences in the birth rate in t and $t-1$ *between* the provinces since we estimate a model in first differences. This implies that the effect measured by these dummies is not a simple time effect. The difference in the differences between provinces is largest in 1892 – which is one year after pension insurance came into effect – and in 1903 – when stricter child labour laws were introduced. While the effects between 1891–1900 suggest that there is an effect of social insurance beyond the mere internal rate of return of the pension system (which renders support to the social security hypothesis in its general form), the effects between 1900–1914 suggest that it is also important to control for policy changes that affected other determinants of current consumption. For example, stricter child labour laws reduce the scope for current consumption and should therefore lead to a lower number of children ceteris paribus.[18]

Moreover, it is important to include such time dummies that capture residual effects, since we cannot completely capture all effects which we would like to measure with our proxies. In particular, while the number of savings books has the expected negative sign and is significant in some specifications for the period between 1891–1899, it is positive and significant when including pension system information for the period between 1900–1914. We draw two conclusions from this. First, the economic effect of the number of savings books in a jurisdiction is small, which implies that the power of this variable as a proxy is limited for measuring a real crowding out of private savings. Second, the fact that the sign and significance of this variable changes when we add information on the pension system indicates that there is a relationship between savings and pensions, even in this noisy measure of savings.

Surprisingly, there is no effect of the additional controls for current consumption we include in column (3). While the sample for the years 1891–1899 is comparatively small, there is hardly a significant effect for the 1900–1914 sample, except for the increase in the number of localities with more than 20.000 inhabitants between 1880 and 1905.

Columns (4) and (5) in tables 2 and 3 present the specifications which include proxies for the internal rate of return of the pension system. Specification (4) adds the contributions in each category as explanatory variables. In terms of the theoretical model, these variables are supposed to capture future consumption. Since the pension system was only introduced in 1891, information on its functioning was only collected from 1892 onwards. If we estimate first differences, this implies that we can only estimate a model for 1893–1899. As a consequence, the number of observations in the 1891–1900 sample is relatively small. Keeping this caveat in mind, it is interesting to find a significant negative effect of the share of contributions in category II in specification (4). Such an effect can be interpreted of the price effect overcompensating the income effect, which is plausible for lower wage categories. In contrast, the variables on the internal rate of return of the pension system included in specification (5) are not significant in this small sample.

For the larger sample for the period 1900–1914, we find a significant negative effect of the share of contributions in wage category IV in specification (4), which would again point to the price effect overcompensating the income effect. Moreover, including this information reduces the magnitude of the information on the sectoral shares of workers; the positive effect of the share of workers in farming and trading turns significant. This underpins our strategy of controlling for other factors that would affect both current and future consumption since otherwise they could mask the effect of pension insurance.

[18]Note that while in 1903 there was the reform of the child labour laws as a major policy change, there was no such change during the 1890s. Therefore, we relate the difference in the differences in the 1890s to social insurance.

Specification (5) in table 3 also shows that including variables that control for other determinants of current consumption is important. A higher share of current pension payments has a positive effect on fertility. However, we find that the indicator for future pension expenditures has an even larger effect on fertility. Both effects are significant and positive. This positive effect confirms HYPOTHESIS 1, since a higher share of future payments per collected Mark of contributions indicates a lower future rate of return of the pension system. If the change in approved old age pensions per Mark of contributions increases by one, the birth rate increases by 1.4 marital births per mill ceteris paribus. However, note that the average value of approved old age pensions per Mark is 0.29 and after 1900 it was even lower at 0.31, which implies that the internal rate of return of the pension system increased between 1891–1899 and 1900–1914. Therefore, it is more reasonable to consider an increase (decrease) in the internal rate of return of the pension system of 0.1, which would reduce (raise) the birth rate by 0.14 births per mill ceteris paribus. Importantly, neither in specification (4) nor in specification (5) adding pension system variables changes magnitude or significance of the other determinants of fertility.

5.3 Sensitivity

To address concerns that might be raised regarding the effects we measure with the time dummies and with the pension system variables, we use another estimator for some additional robustness checks. First, we illustrate that the effect measured by the time dummies in our baseline model is equivalent to an interaction term in a fixed effects model in which we define a treatment group according to the share insured in a province. Second, we show that the negative effect of contributions on fertility and a positive effect of then pension system variables related to the internal rate of return of the pension system persist even when using a different estimator.

In this section, we use a fixed effects estimator, which is in principle equivalent to the first differences estimator. However, it is more efficient since we lose less degrees of freedom. As a consequence, the number of observations is higher than in the change and level model in the previous section. We present specifications (4) and (5) from tables 2 and 3 as fixed effects models in table 3.

While the magnitude of the effects is not exactly the same, which should not be surprising due to the different sample size, the main effects presented in tables 2 and 3 remain robust. The number of stillbirths per mill is highly significant in all specifications and even higher than in the change and level model. Moreover, the positive effect of marriages is also higher for the period 1891–1899 in the fixed effects model. Like in the change and level model, it is only significant in the 1891–1899 sample.

The year effects presented in table 3 are the interaction terms of a dummy that identifies provinces with an at least one standard deviation above the average share of insured with a year dummy. It is interesting that we do not find the significant negative effect in 1892, but instead observe a significant *positive* effect for 1896. This effect could be related to the dowry effect described in Scheubel (2013a). Since women could be reimbursed their pension contributions if they married, but only after 5 years of contributing, 1896 was the first year during which such a dowry effect on fertility could appear.

For the interaction terms for the years 1900-1914 we find negative effects similar to the change and level model. Only those interaction terms for the years 1903 and later are significant, which supports our argument that child labour legislation may have been a main factor driving this effect. These effects however turn insignificant in column (4) which is likely related to adding information on the internal rate of return of the pension system and on current pension payments. The effect of current pension payments is *negative* in column (4) which suggests that the effect may be related to the negative effect captured by the interaction terms in specification (3). Importantly, like in the change and level model, our measure for the internal rate of return of the pension system indicates that a lower internal rate of return of the pension system is associated with a higher birth rate. Like in the change and level model, this effect is only significant for the period 1900-1914.

6 Conclusions

Our paper provides a theoretical underpinning and an empirical confirmation of the negative relationship between statutory old-age insurance and more broadly statutory social insurance and fertility. We thereby give evidence on a well-known theoretical concept in public economics, the social security hypothesis. At the same time, we employ a new historical data set to show that a negative relationship between pensions and fertility can already be observed for late nineteenth century Germany, where the first comprehensive welfare state in the world was introduced at that time. More broadly, our analysis is a confirmation of the fact that people react to institutional incentives.

In this paper, we provide a framework in which the existence of a public pension system can crowd out private savings for old age as well as fertility. Since the overall effect depends on the internal rate of return of the pension system, we use a new and unique historical data set which provides evidence on this internal rate of return for the Bismarckian pension system implemented at the end of the nineteenth century in Imperial Germany. Using this information in a multivariate model, we confirm a positive effect of a lower internal rate of return of the pension system on the birth rate.

In addition, our empirical analysis confirms an overall negative effect of the pension system or more generally the introduction of comprehensive social insurance on fertility, even when controlling for other determinants of fertility as derived from our theoretical model, which also correspond to the usual determinants for the first demographic transition mentioned in the literature. This additional effect amounts to a total reduction of approximately 1.8 marital births per mill between 1891–1899 and a total reduction of approximately 4 marital births per mill between 1900 and 1914. Taken together, this is a reduction of about 25% of the average 1885 level of marital births.

Because our analysis only covers the time span 1891–1914, we cannot account for the longer term impact of pension insurance on people's behaviour. After all, behavioural change mostly takes place gradually. It should, however, not be surprising that nowadays most individuals do not consider old-age provision as a motive for having children. The state had assumed this task long ago. Moreover, in a pay-as-you-go pension system, children constitute a fiscal externality (e.g. Prinz 1990; Kolmar 1997; van Groezen et al. 2003; Sinn 2004; von Auer and Büttner 2004; Fenge and Meier 2009; Meier and Wrede 2010), i.e. the incentive to have children is further reduced because other children would pay an individual's pension once there is credible enforcement by the state. Our model allows for this fiscal externality. Individuals do not take into account the effect of their fertility decision on the internal rate of return of the pension system. We leave a clear identification of this fiscal externality to future research. Given that the direct effect of pensions on fertility amounted to up to 25% of the overall decline, the contribution of statutory pension insurance to the overall decline of fertility up to the current date must be even larger.

References

ABADIE, A. (2005): "Semiparametric Difference-in-Differences Estimators," *Review of Economic Studies*, 72, 1–19.

ARROYO, C. R., AND J. ZHANG (1997): "Dynamic microeconomic models of fertility choice: A survey," *Journal of Population Economics*, 10, 23–65.

ASHENFELTER, O., AND D. CARD (1985): "Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs," *Review of Economics and Statistics*, 67, 648–660.

BARRO, R. J., AND G. S. BECKER (1986): "Altruism and the Economic Theory of Fertility," *Population and Development Review*, 12, 69–76.

——— (1988): "A Reformulation of the Economic Theory of Fertility," *Quarterly Journal of Economics*, 103, 1–25.

——— (1989): "Fertility Choice in a Model of Economic Growth," *Econometrica*, 57, 481–501.

BECKER, G. S. (1960): "An economic analysis of fertility," in *Demographic and Economic Change in Developed Countries*, ed. by N. B. of Economic Research, pp. 225–256. Princeton University Press, Princeton.

——— (1965): "A theory of the allocation of time," *Economic Journal*, 75, 493–517.

——— (1988): "Family Economics and Macro Behavior," *American Economic Review*, 78, 1–13.

——— (1991): *A treatise on the family*. Harvard University Press, Harvard.

——— (1992): "Fertility and the economy," *Journal of Population Economics*, 5, 185–201.

BERTRAND, M., E. DUFLO, AND S. MULLAINATHAN (2004): "How Much Should We Trust Differences-in-Differences Estimates?," *Quarterly Journal of Economics*, 119, 249–275.

BILLARI, F. C., AND V. GALASSO (2009): "What explains fertility? Evidence from Italian pension reforms," CESifo Working Paper.

BROWN, J. C., AND T. W. GUINNANE (2007): "Regions and time in the European fertility transition: problems in the Princeton Project's statistical methodology," *Economic History Review*, 60, 574–595.

CIGNO, A., AND J. ERMISCH (1989): "A Microeconomic Analysis of the Timing of Births," *European Economic Review*, 33, 737–760.

Cleland, J., and C. Wilson (1987): "Demand Theories of the Fertility Transition: An Iconoclastic View," *Population Studies*, 41, 5–30.

Coale, A., and S. Watkins (1986): *The Decline of Fertility in Europe*. Princeton University Press, Princeton.

Coale, A. J. (1965): *Factors associated with the development of low fertility: a historic summary*. United Nations, New York.

——— (1969): *The decline of fertility in Europe from the French Revolution to World War II*. University of Michigan Press, Ann Arbor.

Coale, A. J., and J. Trussell (1974): "Variations in Model Fertility Schedules: The Age Structure of Childbearing in Human Populations," *Population Index*, 40, 185–258.

Donald, S. G., and K. Lang (2007): "Inference with Difference-in-Differences and Other Panel Data," *Review of Economics and Statistics*, 89, 221–233.

Easterlin, R. A. (1975): "An Economic Framework for Fertility Analysis," *Studies in Family Planning*, 6, 54–63.

Feldstein, M. (1974): "Social Security, Induced Retirement, and Aggregate Capital Accumulation," *Journal of Political Economy*, 82, 905–926.

Fenge, R., and V. Meier (2005): "Pensions and fertility incentives," *Canadian Journal of Economics*, 38, 28–48.

Galloway, P. R., E. A. Hammel, and R. D. Lee (1994): "Fertility Decline in Prussia, 1875-1910: A Pooled Cross-Section Time Series Analysis," *Population Studies*, 48, 135–158.

Goldstein, J. R., and S. Klüsener (2010): "Culture Resurfaces – A Geographic Analysis of Fertility Decline in Prussia," MPIDR working paper 2010-012.

Guinnane, T. (2011): "The Historical Fertility Transition: A Guide for Economists," *Journal of Economic Literature*, 49, 589–614.

Guinnane, T., and S. Ogilvie (2008): "Institutions and Demographic Responses to Shocks: Württemberg, 1634-1870," University of Yale Economics Department Working Paper No. 44.

Hirschman, C. (2001): "Globalization and Theories of Fertility Decline," *Population and Development Review*, 27, 116–125.

Knodel, J. E. (1974): *The Decline of Fertility in Germany, 1871-1939*. Princeton University Press, Princeton, New Jersey.

Kolmar, M. (1997): "Intergenerational redistribution in a small open economy with endogenous fertility," *Journal of Population Economics*, 10, 335–356.

Leibenstein, H. M. (1957): *Economic Backwardness and Economic Growth*. Wiley, New York.

Liang, K.-L., and S. L. Zeger (1986): "Longitudinal data analysis using generalized linear models," *Biometrika*, 73, 13–22.

Lotz, W. (1905): "Der Fiskus als Arbeitgeber im deutschen Staatsbahnwesen," *Archiv für Sozialwissenschaft und Sozialpolitik*, 21, 612–658.

Marschalck, P. (1982): "The Federal Republic of Germany with an explanatory hypothesis," in *Population Change and Social Planning*, ed. by D. Eversley, and W. Köllmann, pp. 62–87. Erward Arnold, London.

McNicoll, G. (1980): "Institutional Determinants of Fertility Change," *Population and Development Review*, 6, 441–462.

Meier, V., and M. Wrede (2010): "Pensions, fertility, and education," *Journal of Pension Economics and Finance*, 9, 75–92.

Moulton, B. R. (1990): "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units," *Review of Economics and Statistics*, 72, 334–338.

Neher, P. A. (1971): "Peasants, Procreation, and Pensions," *American Economic Review*, 61, 380–389.

Prinz, A. (1990): "Endogenous fertility, altruistic behavior across generations, and social security systems," *Journal of Population Economics*, 3, 179–192.

Reichsversicherungsamt (1910): *Das Reichsversicherungsamt und die Deutsche Arbeiterversicherung*. Verlag von Behrend & Co., Berlin.

Richards, T. (1977): "Fertility Decline in Germany: An Econometric Appraisal," *Population Studies*, 31, 537–553.

Scheubel, B. D. (2013a): *Bismarck's Institutions - A Historical Perspective on the Social Security Hypothesis*. Mohr Siebeck, Tübingen.

——— (2013b): "Was Bismarck's Pension System Bismarckian? A Note on the Design of the First Comprehensive Pension System," Working Paper.

Sinn, H.-W. (2004): "The pay-as-you-go system as a fertility insurance and enforcement device," *Journal of Public Economics*, 88, 1335–1357.

Smith, H. L. (1989): "Integrating Theory and Research on the Institutional Determinants of Fertility," *Demography*, 26, 171–184.

van Groezen, B., T. Leers, and L. Meijdam (2003): "Social security and endogenous fertility: pensions and child allowances as siamese twins," *Journal of Public Economics*, 87, 233 – 251.

von Auer, L., and B. Büttner (2004): "Endogenous Fertility, Externalities, and Efficiency in Old-Age Pension Systems," *Journal of Institutional and Theoretical Economics JITE*, 160, 294–310.

von Bismarck, O. (1894): "Die Alters- und Invaliditätsversicherung," in *Die politischen Reden des Fürsten Bismarck, Zwölfter Band*, ed. by H. Kohl, pp. 597–616. Verlag der J. G. Cottaschen Buchhandlung, Stuttgart.

Watkins, S. C. (1981): "Regional Patterns of Nuotiality in Europe, 1870-1960," *Population Studies*, 35, 199–215.

Wehler, H.-U. (2008): *Deutsche Gesellschaftsgeschichte 1849-1914*. C.H. Beck, München.

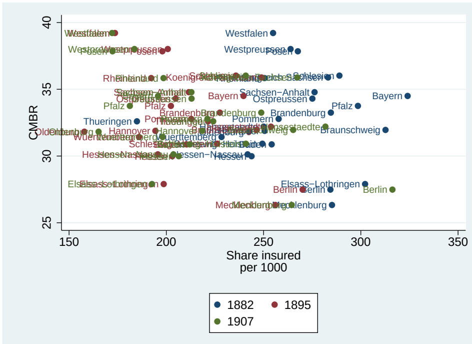
Wooldridge, J. M. (2003): "Cluster-Sample Methods in Applied Econometrics," *American Economic Review, Papers and Proceedings of the One Hundred Fifteenth Annual Meeting of the American Economic Association*, 93, 133–138.

Appendix

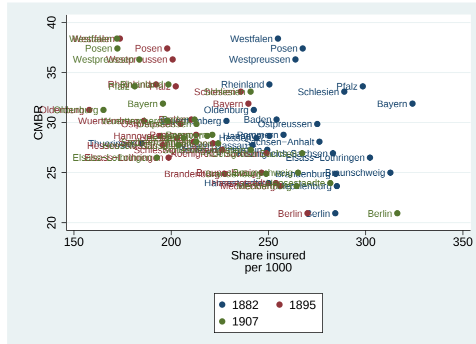
Figures

Figure 1: Regions in Imperial Germany

Figure 2: CMBR and share of insured persons

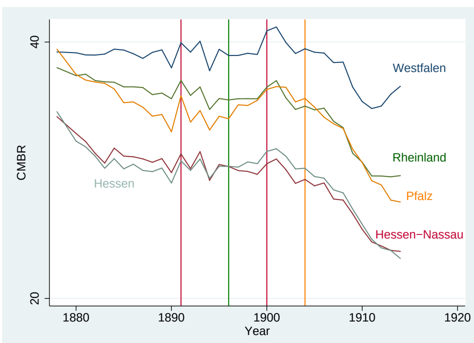


CMBR 1892

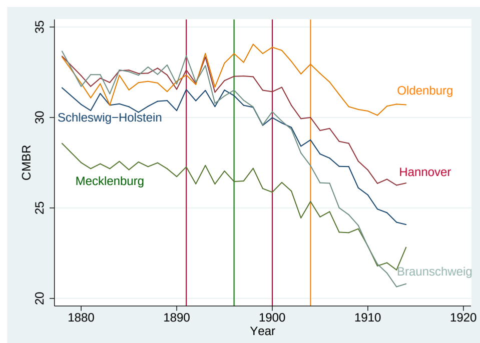


CMBR 1907

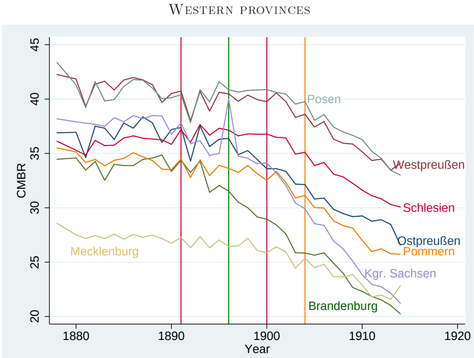
Figure 3: Crude marital birth rate by region



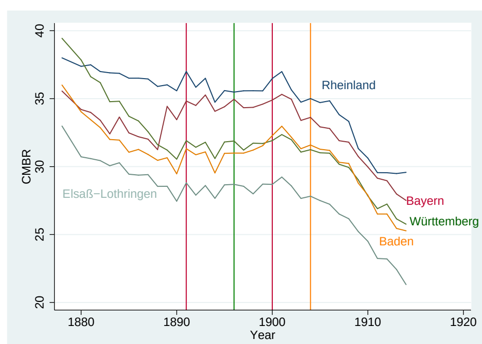
Western provinces



Northern provinces



Eastern provinces



Southern provinces

Tables

Table 1: Share of insured persons persons and fertility

	(1)	(2)	(3)
CMBR			
Insured p. 1000	-.086 (.017)***	-.006 (.002)***	-.005 (.002)**
Stillbirths change yoy		.040 (.013)***	.047 (.015)***
Stillbirths change yoy (L1)		-.013 (.009)	-.004 (.007)
Concl. marriages pT		.152 (.559)	.265 (.450)
Concl. marriages pT (L1)		.309 (.239)	.034 (.261)
Index of agric. productivity		-.007 (.143)	-.043 (.123)
Index of agric. productivity (L1)		-.055 (.193)	-.077 (.186)
Assets per cap. (L1)		-.066 (.282)	-.452 (.355)
Net disability pension entitlements (L1)		.007 (.112)	.024 (.099)
Year: 1907 (D)		-1.581 (.169)***	-1.236 (.259)***
Share in farming			.020 (.010)**
Share in trade			-.021 (.022)
Share in mining			.018 (.014)
Share Catholic			-.0006 (.003)
Savings bank books p. 100 (1900)			-.001 (.0005)**
Obs.	50	44	44

Estimation with OLS. Explanatory variables are also included as first lag in all columns, pension variables only as first lag. L1 indicates coefficient on first lag. Contribution variables only in columns (2) and (4). Level variables included in specification (2). Significance level: *** : $p < 0.01$; ** : $p < 0.05$; * : $p < 0.1$.

Table 2: Results 1891-1899

	(1)	(2)	(3)	(4)	(5)
CMBR					
Δ stillbirths	4.588	4.598	4.566	5.698	5.077
	(1.079)***	(.784)***	(.818)***	(2.289)**	(2.494)*
Δ marriages	0.480	0.479	0.418	0.178	.302
	(.261)*	(.260)*	(.251)	(.342)	(.431)
Δ Productivity	0.002	0.002	0.001	0.005	.025
	(.005)	(.004)	(.004)	(.009)	(.246)
Δ spatial lag	0.299	0.295	0.290	0.338	.439
	(.011)***	(.008)***	(.114)**	(.174)**	(.191)*
Year : 1892	-1.071	-1.078	-1.090	–	–
	(.180)***	(.184)***	(.200)***	–	–
Year : 1893	.404	.405	.397	.732	–
	(.139)***	(.140)***	(.146)**	(.302)**	–
Year : 1894	-.987	-.993	-1.004	–	–
	(.262)***	(.265)***	(.282)***	–	–
Year : 1895	.537	.539	.544	–	–
	(.149)***	(.149)***	(.156)***	–	–
Year : 1896	.449	.449	.454	.414	.031
	(.447)	(.448)	(.460)	(.270)	(.180)
Savings books		-.0003	-.0004	-.006	.003
		(.0001)	(.001)	(.003)	(.001)**
Share farming			.012	-0.031	.027
			(.008)	(.048)	(.036)
Share trading			.004	-0.123	.039
			(.027)	(.131)	(.108)
Share mining			.025	-0.099	.038
			(.017)	(.110)	(.076)
Localities			-0.013	0.012	.002
			(.023)	(.042)	(.012)
Percentage Catholic			0.001	0.004	.004
			(.001)	(.003)	(.003)
Persons per household			.113	-0.222	-.095
			(.107)	(.320)	(.105)
Δ Share contrib. cat. I				.104	
				(.251)	
Δ Share contrib. cat. II				-.132	
				(.064)*	
Δ Share contrib. cat. III				.193	
				(.122)	
Δ Contrib. cat. I/IV(V)				-6.097	
				(10.787)	
Δ Approved old-age pensions/Mark					-3.760
					(2.555)
Δ Approved disability pensions/Mark					-2.698
					(2.012)
Δ Current pension payments per mill					14.051
					(10.901)
Time trend	YES	YES	YES	YES	YES
Proxies for current consumption	NO	NO	YES	YES	YES
Proxies for future consumption	NO	NO	NO	YES	YES
Obs.	161	161	161	92	69

Estimation with OLS, correction for unobserved heterogeneity with first differencing. CMBR, stillbirths and marriages calculated per mill. Spatial lag calculated as the birth rate in the province where most immigrants are from. Productivity in agriculture is measured as the harvest per square kilometre. Figures on the share of the population in farming, mining and trade from occupational census 1895. Variable on localities with a population above 20.000 measures the change between census years 1871 and 1880. The percentage of Catholics in the population is from 1890. The share of contributions in category IV is dropped because of collinearity. Data on savings books from year 1900. Significance level: $^{***}: p < 0.01$; $^{**}: p < 0.05$; $^{*}: p < 0.1$.

Table 3: Results 1900-1914

	(1)	(2)	(3)	(4)	(5)
CMBR					
Δ stillbirths	3.985	3.980	3.876	3.321	3.370
	(.878)***	(.877)***	(.840)***	(.820)***	(.798)***
Δ marriages	-0.022	-0.022	-0.021	-.030	-.024
	(.014)	(.015)	(.015)	(.014)**	(.012)*
Δ Productivity	-0.096	-0.097	-0.100	-.053	-.080
	(.040)***	(.050)*	(.050)*	(.048)	(.049)
Δ births neighbour		0.021	0.025	.237	.018
		(.017)***	(.017)	(.101)**	(.016)
Year : 1901	-.094	-.089	-.078	–	-.161
	(.147)	(.147)	(.145)	–	(.141)
Year : 1902	-.664	-.659	-.650	-.535	-.764
	(.157)***	(.158)***	(.161)***	(.228)**	(.150)***
Year : 1903	-1.084	-1.078	-1.069	-.837	-1.249
	(.094)***	(.093)***	(.096)***	(.223)***	(.105)***
Year : 1904	.031	.034	.040	.018	-.109
	(.119)	(.119)	(.119)	(.144)	(.101)
Year : 1905	-.581	-.576	-.567	-.464	-.748
	(.109)***	(.108)***	(.110)***	(.183)**	(.118)***
Year : 1906	.061	.064	.069	.030	-.084
	(.108)	(.108)	(.110)	(.134)	(.079)
Year : 1907	-.783	-.779	-.772	-.645	-.961
	(.110)***	(.109)***	(.112)***	(.178)***	(.072)***
Year : 1908	-.163	-.160	-.154	-.133	-.373
	(.010)	(.122)	(.101)	(.125)	(.082)
Year : 1909	-.620	-.617	-.612	-.417	-.873
	(.157)***	(.157)***	(.161)***	(.146)***	(.106)***
Year : 1910	-.465	-.463	-.460	-.346	-.769
	(.151)***	(.151)***	(.155)***	(.154)**	(.105)**
Year : 1911	-.614	-.612	-.609	-.452	-.915
	(.138)***	(.138)***	(.141)***	(.153)***	(.121)***
Year : 1912	.341	.342	.347	.252	–
	(.125)**	(.125)**	(.128)**	(.142)*	–
Year : 1913	-.174	-.172	-.169	-.107	-.587
	(.146) (.146)	(.148)	(.143)	(.129)***	
Savings books		-0.0003	0.0001	.0002	.0000
		(.0002)*	(.0002)	(.0001)	(.0002)
Share farming			0.009	.016	.003
			(.007)	(.008)*	(.008)
Share trading			0.34	.049	.031
			(.020)	(.022)**	(.020)
Share mining			-0.002	.004	-.014
			(.015)	(.015)	(.010)
Localities > 20.000			0.009	.010	.006
			(.003)**	(.004)**	(.004)
Percentage Catholic (1890)			0.001	-.0001	.0006
			(.001)	(.001)	(.001)
Horses per mill			0.001	.002	.0002
			(.002)	(.002)	(.002)
Persons per household				-0.222	-.095
				(.320)	(.105)

	(1)	(2)	(3)	(4)	(5)
CMBR					
Δ Share contrib. cat. I				-.034 (.033)	
Δ Share contrib. cat. II				-.009 (.023)	
Δ Share contrib. cat. III				-.022 (.022)	
Δ Share contrib. cat. IV				-.021 (.011)*	
Δ Share contrib. cat. V				.003 (.021)	
Δ Contrib. cat. I/IV(V)				.256 (.859)	
Δ Approved old-age pensions/Mark					1.443 (.506)***
Δ Approved disability pensions/Mark					.022 (.066)
Δ Current pension payments per mill					.999 (.534)*
Time trend	YES	YES	YES	YES	YES
Proxies for current consumption	NO	NO	YES	YES	YES
Proxies for future consumption	NO	NO	NO	YES	YES
Obs.	344	344	344	322	321

Estimation with OLS, correction for unobserved heterogeneity with first differencing. CMBR, stillbirths and marriages calculated per mill. Spatial lag calculated as the birth rate in the province where most immigrants are from. Productivity in agriculture is measured as the harvest per square kilometre. Figures on the share of the population in farming, mining and trade from occupational census 1905. Variable on localities with a population above 20.000 measures the change between census years 1880 and 1905. The percentage of Catholics in the population is from 1900. The number of horses measures the difference between years 1897 and 1892 when this figure was reported and is calculated per mill. Data on savings books from year 1900. Significance level: $^{***}: p < 0.01$; $^{**}: p < 0.05$; $^{*}: p < 0.1$.

Table 3: Sensitivity: Fixed effects model

	(1)	(2)	(3)	(4)
CMBR				
Stillbirths per mill	7.839 (1.412)***	7.503 (1.644)***	7.224 (.994)***	5.211 (.891)***
Marriages per mill	1.166 (.297)***	1.164 (.315)***	.003 (.041)	.014 (.035)
Productivity	.009 (.014)	.067 (.242)	-.055 (.119)	-.145 (.100)
Spatial lag	.052 (.023)**	.047 (.027)*	.356 (.068)***	.233 (.062)***
Year: 1891	.803 (.515)	.450 (.549)		
Year: 1892	.073 (.480)	.399 (.533)		
Year: 1893	-.055 (.479)	.236 (.520)		
Year: 1894				
Year: 1895	-.268 (.481)			
Year: 1896	1.622 (.485)***	1.684 (.516)***		
Year: 1897	.152 (.485)	.227 (.516)		
Year: 1898	–	–		

	(1)	(2)	(3)	(4)
Year: 1901			-.189 (.457)	-.085 (.372)
Year: 1902			-.569 (.463)	.026 (.375)
Year: 1903			-.854 (.471)*	-.012 (.387)
Year: 1904			-.955 (.464)**	-.075 (.383)
Year: 1905			-1.130 (.462)**	-.094 (.386)
Year: 1906			-1.168 (.461)**	-.065 (.384)
Year: 1907			-1.469 (.465)***	-.274 (.385)
Year: 1908			-1.228 (.464)***	-.106 (.385)
Year: 1909			-1.345 (.473)***	-.143 (.386)
Year: 1910			-1.305 (.473)***	-.086 (.387)
Year: 1911			-1.486 (.470)***	-.188 (.389)
Year: 1912			-1.285 (.467)***	.062 (.391)
Year: 1913			-1.301 (.465)***	.211 (.395)
Year: 1914			-1.329 (.463)***	–
Share contrib. cat. I	.028 (.073)		-.070 (.126)	
Share contrib. cat. II	-.027 (.041)		-.054 (.126)	
Share contrib. cat. III	.046 (.046)		-.035 (.126)	
Share contrib. cat. IV			-.074 (.124)	
Share contrib. cat. V			-.073 (.126)	
Approved old-age pensions/Mark		-.011 (.332)		2.525 (.429)***
Approved disability pensions/Mark		.035 (.152)		.0007 (.130)
Current pension payments per mill		1.607 (1.443)		-2.199 (.502)***
Time trend	YES	YES	YES	YES
Years	1891–1899	1891–1899	1900–1914	1900–1914
Obs.	160	137	345	322

Estimation with OLS, correction for unobserved heterogeneity with within transformation. CMBR, stillbirths and marriages calculated per mill. Spatial lag calculated as the birth rate in the province where most immigrants are from. Productivity in agriculture is measured as the harvest per square kilometre. Figures on the share of the population in farming, mining and trade from occupational census 1905. Variable on localities with a population above 20.000 measures the change between census years 1880 and 1905. The percentage of Catholics in the population is from 1900. The number of horses measures the difference between years 1897 and 1892 when this figure was reported and is calculated per mill. Data on savings books from year 1900. Significance level: $^{***}: p < 0.01$; $^{**}: p < 0.05$; $^{*}: p < 0.1$.

Appendix A: Details on the theoretical model

6.1 Second Order Conditions

The second derivatives of equations (4) and (5) are given by:

$$\begin{aligned} V_{nn} &= -U_c(1-\tau)w_t f''(n_t) - U_z\Omega_{t+1}\tau w_{t+1} f''(n_t) \\ &\quad + U_{cc}\left[(1-\tau)w_t f'(n_t) + \pi_t\right]^2 + U_{zz}\left[B_{t+1} - \Omega_{t+1}\tau w_t f'(n_t)\right]^2 + U_{nn} < 0 \end{aligned} \quad (1)$$

$$V_{ns} = U_{cc}((1-\tau)w_t f'(n_t) + \pi_t) + U_{zz}\left[B_{t+1} - \Omega_{t+1}\tau w_t f'(n_t)\right] R_{t+1} = V_{sn} \quad (2)$$

in the Bismarckian case,

$$\begin{aligned} V_{nn} &= -U_c(1-\tau)w_t f''(n_t) + U_{cc}\left[(1-\tau)w_t f'(n_t) + \pi_t\right]^2 \\ &\quad + U_{zz}B_{t+1}^2 + U_{nn} < 0 \end{aligned} \quad (3)$$

$$V_{ns} = U_{cc}((1-\tau)w_t f'(n_t) + \pi_t) + U_{zz}R_{t+1}B_{t+1} = V_{sn} \quad (4)$$

in the Beveridgean case and

$$\begin{aligned} V_{nn} &= -U_c w_t f''(n_t) + U_{cc}\left[(1-\tau)w_t f'(n_t) + \pi_t\right]^2 \\ &\quad + U_{zz}\left[B_{t+1} - R_{t+1}\tau w_t f'(n_t)\right]^2 + U_{nn} < 0 \end{aligned} \quad (5)$$

$$\begin{aligned} V_{ns} &= U_{cc}((1-\tau)w_t f'(n_t) + \pi_t) + U_{zz}R_{t+1}\left[B_{t+1} - R_{t+1}\tau w_t f'(n_t)\right] \\ &= V_{sn} \end{aligned} \quad (6)$$

in the fully-funded pensions system. In all pension systems holds

$$V_{ss} = U_{cc} + U_{zz}R_{t+1}^2 < 0. \quad (7)$$

The second-order conditions for a maximum of problem (3) are satisfied under all three pension systems since V_{nn} is negative and the following conditions hold true:

$$\begin{aligned} V_{nn}V_{ss} - V_{ns}V_{sn} &= (U_{cc} + U_{zz}R_{t+1}^2)\left[U_{nn} - U_c(1-\tau)w_t f''(n_t) - U_z\Omega_{t+1}\tau w_{t+1} f''(n_t)\right] \\ &\quad + U_{cc}U_{zz}\left[R_{t+1}((1-\tau)w_t f'(n_t) + \pi_t) - \left(B_{t+1} - \Omega_{t+1}\tau w_{t+1} f'(n_t)\right)\right]^2 \\ &> 0 \end{aligned} \quad (8)$$

in the Bismarckian case,

$$
\begin{aligned}
V_{nn}V_{ss} - V_{ns}V_{sn} &= (U_{cc} + U_{zz}R_{t+1}^2)\left[U_{nn} - U_c(1-\tau)w_t f''(n_t)\right] \\
&\quad + U_{cc}U_{zz}\left[R_{t+1}((1-\tau)w_t f'(n_t) + \pi_t) - B_{t+1}\right]^2 \\
&> 0 \qquad (9)
\end{aligned}
$$

in the Beveridgean case and

$$
\begin{aligned}
V_{nn}V_{ss} - V_{ns}V_{sn} &= (U_{cc} + U_{zz}R_{t+1}^2)\left[U_{nn} - U_c w_t f''(n_t)\right] \\
&\quad + U_{cc}U_{zz}\left[R_{t+1}(w_t f'(n_t) + \pi_t) - B_{t+1}\right]^2 \\
&> 0 \qquad (10)
\end{aligned}
$$

in the fully-funded case. This demonstrates that in each case the objective function $V(n_t, s_t)$ is strictly concave in the decision variables.

6.2 Crowding out of savings in a fully funded system

In a fully-funded pension system, contributions during the working period are invested in the capital market, yield the interest factor R and are paid out as pensions in the retirement period. Hence, the pension of a household of generation t is given by

$$p_{t+1}^{FF} = R_{t+1}\tau w_t(1 - f(n_t)). \qquad (11)$$

Note that in a fully-funded pension system another child reduces the pension proportional to the interest factor:

$$\frac{\partial p_{t+1}^{FF}}{\partial n_t} = -\tau w_t f'(n_t) R_{t+1} < 0. \qquad (12)$$

The intertemporal budget constraint is given by substituting (11) in (2) and combining this individual budget constraint in the second period with (1):

$$R_{t+1}c_t + z_{t+1} = R_{t+1}\left[w_t(1 - f(n_t)) - \pi_t n_t - B_t\right] + B_{t+1}n_t. \qquad (13)$$

Lifetime consumption in second period units on the LHS is financed by lifetime income on the RHS. Evaluating the effect of an additional child on lifetime income by differentiating lifetime income with respect to n_t yields the marginal price of children in present value terms of period $t+1$:

$$\Pi_{t+1}^{FF} = R_{t+1}(w_t f'(n_t) + \pi_t) - B_{t+1} \qquad (14)$$

An additional child causes opportunity costs by reducing wage income by $w_t f'(n_t)$ and direct costs of π_t. However, a child pays an intra family transfer of B_{t+1} which reduces the marginal price. For the sake of a well-defined decision problem with a finite number of children we assume this price to be positive.

We start by analysing the savings decision under a fully-funded pension system. The effect of a higher contribution rate on savings depends on the sign of the numerator on the RHS of (**??**). By using the second derivatives (5) and (6) from Appendix A and the second derivatives with respect to the contribution rate:

$$
\begin{aligned}
V_{n\tau} &= w_t(1-f(n_t)) \\
&\quad \left[U_{cc}((1-\tau)w_t f'(n_t)+\pi_t)+U_{zz}\left[B_{t+1}-R_{t+1}\tau w_t f'(n_t)\right]R_{t+1}\right] \qquad (15)
\end{aligned}
$$

$$
V_{s\tau} = w_t(1-f(n_t))\left[U_{cc}+U_{zz}R_{t+1}^2\right] < 0 \qquad (16)
$$

this numerator is given by:

$$
\begin{aligned}
V_{nn}V_{s\tau}-V_{n\tau}V_{sn} &= w_t(1-f(n_t)) \\
&\quad \left[\left(U_{nn}-U_c w_t f''(n_t)\right)\left(U_{cc}+U_{zz}R_{t+1}^2\right)\right. \\
&\quad \left.+U_{cc}U_{zz}\left(R_{t+1}(w_t f'(n_t)+\pi_t)-B_{t+1}\right)^2\right] \\
&> 0 \qquad (17)
\end{aligned}
$$

By employing (10) in Appendix A we find that

$$
\frac{\partial s_t}{\partial \tau} = -w_t(1-f(n_t)).
$$

This means that private savings are reduced exactly by the amount at which forced savings increase in the fully-funded system. In the presence of perfect capital markets this is the well-known result of complete savings crowding-out.

The fertility decision within this pension system is determined by the numerator of the RHS of equation (6). Using the second derivatives from above the numerator reduces to zero: $V_{n\tau}V_{ss}-V_{ns}V_{s\tau}=0$. A fully-funded pension system has no effect on fertility. The reason is that neither the marginal price of children of (14) nor the lifetime income from (13) is affected by the contribution rate. Increasing forced savings for old-age is completely compensated by changes in private savings so that the optimal amount of effective savings remains unchanged with a perfect capital market. The intertemporal budget set is the same as without a fully-funded pension and the optimal allocation of the number of children and consumption is unaltered.

Note that this result rests on the assumption of an interior solution with perfect capital markets. As soon as we assume credit constraints, fertility may be negatively affected by funded pension schemes. In the case where contributions to the pension system reduce the budget by an amount larger than the optimal level of savings in the absence of pension insurance, the credit constraint may be binding and the expenditures for children have to be reduced. Here we have a pure income effect on fertility which reduces fertility as a normal good. The same holds true in the case of lacking capital markets so that private savings cannot compensate the fully-funded pension.

Put differently, in a fully-funded system, we only observe a negative income effect on fertility if credit constraints are binding. Otherwise, there is a full substitution of private savings by forced public savings.

As the link between contributions and pensions is perfect in this fully-funded case, the pension system acts as a quasi private investment. This is why we do not observe opportunity cost effects. If the internal rate of return of the pension system differs from the capital market rate of return and children reduce labour supply, we observe opportunity cost effects.

6.3 Savings decision in a PAYG pension system

The impact of extending the pension system on savings is given by:

$$\frac{\partial s_t}{\partial \tau} = -\frac{V_{nn}V_{s\tau} - V_{n\tau}V_{sn}}{V_{nn}V_{ss} - V_{ns}V_{sn}}. \tag{18}$$

The denominator is positive for all three pension types (see Appendix A1).

In the case of the Bismarckian pension system we have

$$\begin{aligned} V_{nn}V_{s\tau} - V_{n\tau}V_{sn} \quad = \quad & w_t(1-f(n_t))(U_{cc} + U_{zz}\Omega_{t+1}R_{t+1}) \\ & \big(U_{nn} - U_c(1-\tau)w_t f''(n_t) - U_z\Omega_{t+1}\tau w_{t+1}f''(n_t)\big) \\ & -U_z w_t f'(n_t)(R_{t+1} - \Omega_{t+1})[U_{cc}((1-\tau)w_t f'(n_t) + \pi_t) \\ & +U_{zz}R_{t+1}\left(B_{t+1} - \Omega_{t+1}\tau w_t f'(n_t)\right)] \\ & +U_{cc}U_{zz}w_t(1-f(n_t)) \\ & \big[R_{t+1}((1-\tau)w_t f'(n_t) + \pi_t) - \big(B_{t+1} - \Omega_{t+1}\tau w_t f'(n_t)\big) \\ & \big(\Omega_{t+1}(w_t f'(n_t) + \pi_t) - B_{t+1}\big)\big] \end{aligned} \tag{19}$$

This numerator is positive if the following condition for the intra-family transfer B_{t+1} holds: $-\frac{\partial p_{t+1}^{BIS}}{\partial n_t} < B_{t+1} < \Omega_{t+1}((1-\tau)w_t f'(n_t) + \pi_t) - \frac{\partial p_{t+1}^{BIS}}{\partial n_t}$. This condition can be simplified to $\tau w_t f'(n_t) < \frac{B_{t+1}}{\Omega_{t+1}} < w_t f'(n_t) + \pi_t$. If this condition holds, savings decrease with a higher contribution rate in the Bismarckian system.

The first part of the inequality condition means that the intra-family transfer of children in the second period is higher than the cost of children due to the reduced Bismarckian pension. Having more children would increase the consumption in the second period. The second part of the condition implies that the discounted intra-family transfer is lower than the cost of children in the first period. A higher number of children decreases consumption in the first period. In other words, a higher number of children reduces labour supply. Both effects together imply that savings will be reduced. Since $V_{ns} < 0$ is met with this inequality condition, the fall in wage income is partially offset by lower savings.

6.4 Lack of capital markets

If we assume that individuals have no possibility to provide for old age by savings the budget constraints in both periods are given by:

$$\begin{aligned} c_t &= w_t(1-f(n_t))(1-\tau) - \pi_t n_t - B_t \\ z_{t+1} &= p_{t+1} + B_{t+1} n_t \end{aligned}$$

where the pension in a Bismarckian system is determined by (7). Again the first-order condition (4) holds. The implicit function theorem yields

$$\frac{\partial n}{\partial \tau} = -\frac{V_{n\tau}}{V_{nn}}$$

and $V_{nn} < 0$ is given by (1). Hence, the fertility response with respect to an introduction or extension of the pension system is determined by the sign of $V_{n\tau}$:

$$\begin{aligned} V_{n\tau} &= w_t f'(n_t) U_z (R_{t+1} - \Omega_{t+1}) + w_t(1-f(n_t)) \\ & \quad \left[U_{cc}((1-\tau) w_t f'(n_t) + \pi_t) + U_{zz}\left(B_{t+1} - \Omega_{t+1}\tau w_t f'(n_t)\right)\Omega_{t+1}\right] \end{aligned} \tag{20}$$

Again in a dynamically efficient economy a higher contribution rate τ decreases the marginal price of a child which incites more children:

$$w_t f'(n_t) U_z (R_{t+1} - \Omega_{t+1}) > 0$$

A higher contribution rate decreases income in the first period by $w_t(1-f(n_t))$ and raises pension income in the second period by $\Omega_{t+1} w_t (1-f(n_t))$. Reducing the number of children compensates the income loss in period 1 by the expenditure $(1-\tau)w_t f'(n_t) + \pi_t$ per child and decreases the income in period 2 if $B_{t+1} > \Omega_{t+1}\tau w_t f'(n_t)$, in other words, if the intra family transfer is larger than the Bismarck pension loss due to another child. Smoothing consumption across periods increases utility of the household so that due to the income effect fertility decreases with a higher contribution rate:

$$U_{cc}((1-\tau) w_t f'(n_t) + \pi_t) + U_{zz}\left(B_{t+1} - \Omega_{t+1}\tau w_t f'(n_t)\right)\Omega_{t+1} < 0$$

Hence, the size of the intra family transfer determines the income effect and whether it is larger than the first (price) effect in which case fertility decreases with a higher contribution rate.

Corollary 5: Constrained Investment effect in a pay-as-you-go Bismarckian pension system

In economies with lacking capital markets to provide for old-age the introduction or expansion of a Bismarckian pay-as-you-go pension scheme reduces the number of children if the intra-family transfers are sufficiently large.

Appendix B: Data

6.1 The Data Set

The data set is combined from two sources. The first source is the Annual Yearbook of Statistics for Imperial Germany (*Statistisches Jahrbuch für das Deutsche Reich*), which was published by the Imperial Statistical Office (Kaiserliches Statistisches Amt 1880–1914). The first Annual Yearbook of Statistics was published in 1871, but only after 1880 it was officially called the Annual Yearbook of Statistics for Imperial Germany (before: Statistics of Imperial Germany).

The Annual Yearbook of Statistics is an invaluable source when it comes to long time series information on key indicators for the states of Imperial Germany. While the regional statistical offices collected and published information at lower jurisdictional levels, the information in the Annual Yearbook of Statistics is either aggregated at the federal level or at the state level.

Details regarding the matching of the data as well as the jurisdictions included in the data set can be found in Scheubel (2013a).

6.2 Summary statistics

In this section we provide summary statistics for the variables used in this analyses. Not every variable is available for every year in the data. Scheubel (2013a) provides an overview of the availability of each variable by year.

Table 1 shows the summary statistics averaged over all years for all variables used in our analyses.

Table 1: Summary statistics

Variable	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
Births (p 1000)	925	35.0	4.9	18.8	47.1
Share of illegitimate births	900	8.9	3.5	1.3	58.2
Stillbirths (p 1000)	899	1.2	0.3	0.7	2.0
Marriages (p 1000)	925	7.9	1.0	0.8	22.1
Share in farming	99	33.8	16.8	0.2	65.4
Share in mining	99	19.0	8.8	6.2	45.5
Share in trade	99	5.7	3.4	2.0	23.9
Catholic population (%)	175	28.7	26.4	0.2	81.0
Persons per household	100	4.6	0.4	2.4	5.3
Horses (p 1000)	218	80.7	50.4	16.6	245.2
Localities $>$ 20.000	125	7.9	8.7	0.0	47.0
Savings books (p 1000)	25	50.0	70.0	0.0	345.9
Revenues: other (Mark pc)	575	0.6	0.6	0.0	3.9
Revenues: category I (Mark pc)	550	278.5	225.4	-18.8	923.4
Revenues: category II (Mark pc)	550	721.8	260.0	117.6	1845.7
Revenues: category III (Mark pc)	550	642.4	358.0	76.2	2398.9
Revenues: category IV (Mark pc)	550	452.4	425.8	26.4	3585.7
Revenues: category V (Mark pc)	350	461.4	520.5	23.6	3681.8

Expenditures: administration (Mark pc)	575	0.1148	0.0657	0.0167	0.5191
Expenditures: pensions (Mark pc)	575	0.9998	0.5765	0.0816	3.2159
Expenditures: other (Mark pc)	600	0.035	0.034	-0.244	0.347
Existing pension entitlements (old age) (pc)	566	0.0027	0.0015	0.0006	0.0077
Existing pension entitlements (disability) (pc)	540	0.0082	0.0055	0.0001	0.0236
Approved disability applications (pc)	550	0.0016	0.0007	0.0000	0.0040
Approved old age pension applications (pc)	575	0.0004	0.0006	0.0001	0.0061
Average disability pension (Mark)	550	149.8	23.1	112.8	214.4
Average old age pension (Mark)	574	151.3	19.3	109.7	199.2
Number insured (p 1000)	75	225.1	40.1	142.5	423.2