EUROPEAN CENTRAL BANK

EUROSYSTEM

WORKING PAPER SERIES

NO 1734 / SEPTEMBER 2014

PENSIONS AND FERTILITY BACK TO THE ROOTS

The introduction of Bismarck's pension scheme and the European fertility decline

Robert Fenge and Beatrice Scheubel

In 2014 all ECB publications feature a motif taken from the €20 banknote.

Acknowledgements
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.

Robert Fenge
University of Rostock; e-mail: robert.fenge@uni-rostock.de

Beatrice Scheubel (corresponding author)
European Central Bank; e-mail: beatrice.scheubel@ecb.europa.eu

Address	Kaiserstrasse 29, 60311 Frankfurt am Main, Germany
Postal address	Postfach 16 03 19, 60066 Frankfurt am Main, Germany
Telephone	+49 69 1344 0
Internet	http://www.ecb.europa.eu

ISSN	1725-2806 (online)
ISBN	978-92-899-1142-9
EU Catalogue No	QB-AR-14-108-EN-N (online)

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 provides 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 at a jurisdictional level for a time when comprehensive social security was introduced, we estimate the effects predicted by the model. We find that beyond the traditional determinants of the first demographic transition, a lower internal rate of return of the pension system is associated with a higher birth rate. This result is robust to including the traditional determinants of the first demographic transition as controls as well as to other policy changes at the time.

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

JEL-Codes: C21, H31, H53, H55, J13, J18, J26, N33

Non-technical summary

An ageing population is considered one of the major challenges for developed economies. To deal with population change, its causes have to be understood. One major cause for population change – the existence of the welfare state – has received comparatively little attention in the recent academic debate. This paper tries to fill this gap by analysing the link between social security and fertility in a theoretical model and by testing the model implications with historical data.

To most economists, it is clear that social insurance provision as well as social insurance contributions trigger changes in behaviour, for example in the labour supply decision. This link between social security and individual behaviour has been postulated as the so-called social security hypothesis which states that the individual provision for the major risks of life – sickness, accidents, poverty – declines whenever the state provides insurance against these risks. Therefore, it may seem surprising that the link between social security and other changes in individual behaviour, such as fertility, has received less attention in broader discussions on the fertility decline in advanced economies.

In the public finance literature, it is well-established that the link between fertility and the public provision of pension insurance can be considered a special case of the social security hypothesis. However, testing this link is more difficult than testing for example labour market effects of social insurance, since social security and in particular pension systems have been in place for over a century in most advanced welfare states. Exogenous changes are rare. In addition, social insurance also affects the savings decision, which complicates the analysis even further.

This paper provides two major contributions to understanding the impact of social security on fertility and thus to the understanding of the causes of population ageing. First, we model the fertility decision jointly with the savings decision and allow for endogenous labour supply, developing further previous microeconomic models on the link between pensions and fertility. With this model, we derive a testable hypothesis on the link between pensions and fertility. Second, we use a novel data set on the introduction of pension insurance in Imperial Germany to test the hypothesis derived from the model. While the data set suffers from the usual constraints regarding the availability of specific variables that are inherent to historical data, it provides a unique opportunity for analysis since it covers the period of the *introduction* of social security in Imperial Germany.

The model used in this paper is a simple overlapping generations (OLG) 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. One of the crucial assumptions in the model is the reduction of labour supply whenever a household decides to have children. This assumption implies that there is an opportunity cost of having children in terms of foregone lifetime income. Since a higher contribution to the pension system reduces the net wage, it also reduces this opportunity cost, having ceteris paribus a positive effect on the birth rate. At the same time a higher contribution rate also implies a higher income. Consequently, if the internal rate of return of the pension system is high, it is more costly to have children instead of paying into the pension system, such that fertility is reduced. In equilibrium, these effects are traded off against each other. The relation between income and pension claims – the internal rate of return of the pension system – determines which effect dominates.

In the empirical analysis for the jurisdictions of Imperial Germany we test the hypothesis of reduced fertility when the internal rate of return of the pension system is higher by using historical information on the internal rate of return of the Bismarckian pension system in the late 1890s and early 1900s. We show that a higher internal rate of return of the pension system is associated with a lower birth rate.

This paper thus provides a theoretical underpinning and an empirical confirmation of the negative relationship between statutory old-age insurance or more broadly statutory social insurance and fertility. The effect amounts to a total reduction of approximately 1.7 marital births per 1000 between 1895 and 1907. Since we also test for the other determinants of the first demographic transition which have been identified in the literature, we can compare the impact of pension insurance to other factors. For example, th impact of pension insurance is comparable to the impact of an increase in urbanisation by 10-20%.

Considering that the impact of social security on people's lives has increased rather than decreased since the early nineteenth century, the impact of social security on current levels of fertility is likely to be even larger. Therefore, the impact of social security on the current ageing problem should not be underestimated. In particular in the context of strained public finances and a widespread need for structural reforms, re-evaluating the design of the welfare state seems a promising area of development.

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). 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, 1988, 1991) have received more attention in the context of the first demographic transition. 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 comparatively little attention. Early work incorporated population growth into growth models, providing a macroeconomic perspective on the fertility decline (e.g. Leibenstein 1957). As fertility behaviour was increasingly incorporated in microeconomic models of individual behaviour (e.g. Neher 1971; Nugent 1985), these models were combined with another strand of the literature concerned with the labour-leisure choice (e.g. Becker 1965) and with the analysis of how social security affects retirement patterns and saving (e.g. Feldstein 1974), either in a development context (Hohm 1975), a sophistication of endogenous growth models (e.g. Abio et al. 2004) or in overlapping generation models (e.g. Prinz 1990; Cigno 1993; Cigno and Rosati 1996; Sinn 2004; Fenge and Meier 2005; Cremer et al. 2008). We have developed our model in the spirit of Cigno (1993); thus it can be considered part of the literature focusing on microeconomic explanations for fertility behaviour.

[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.

Empirical studies on the pensions-fertility nexus are less prevalent, partly because most public pension systems have been in place for years such that exogenous variation in key pension system determinants is difficult to find. Some rely on cross-country variation (e.g. Ehrlich and Zhong 1998; Boldrin et al. 2005), which always entails the caveats of country-specific initial conditions and development paths. Other studies focus on specific countries or exogenous changes within a specific pension system (e.g. Cigno and Rosati 1992; Cigno et al. 2003; Billari and Galasso 2009). Cigno and Werding (2007) give an overview of the work on the pensions-fertility nexus in the contemporary context. Guinnane (2011) provides a summary of more recent empirical research on the historical fertility decline. The studies which analyse the connection between the generosity of the pension system and fertility find that a less generous pension system has positive effects on fertility.

In this paper, we provide more evidence on the pensions-fertility nexus in the historical 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 novel set of historical data.

This study shows that a higher internal rate of return is associated with a lower birth rate. Moreover, even after controlling for the traditional determinants of the first demographic transition, inter alia industrialisation, education and urbanisation, and a time trend, we find that on average, the pension system had contributed a little less than 1/6 of the total decline in birth rates between 1895 and 1907. Even when controlling for the introduction of other pillars of social insurance as well as other policy reforms, the effect is persistent. 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.

Pension insurance provided 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. Being unable to work because of age was only recognised if a worker reached the age of 70[3] while 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). Therefore and since 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, we interpret the distinction between disability pensions and old age pensions as mainly semantic. In other words, the disability pension was the relevant pension for a worker considered 'old' at the time in most cases.

The pension system of the 1890s was neither a pure pay as you go pension scheme nor a fully-funded pension scheme (Scheubel 2013). 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.[4]

The pension system of 1891 was a partially funded pension system that was mandatory only for parts of the population (Scheubel 2013). 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).[5] As a consequence, about 20-25% of the population were covered by pension insurance (Scheubel 2013).

The pension level depended on contributions, such that the pension system can be classified as a defined-contribution system (Scheubel 2014). 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. The average old-age pension in Imperial Germany was 21.88% of the average

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

[4] Refer to the publication of the law in *Reichsgesetzblatt* (RGbl) 1899/33.

[5] Also refer to the published law in *Reichsgesetzblatt* (RGbl) 1889/13.

annual wage in rail track supervision and maintenance and the average disability pension was 21.36% of the average annual wage in that sector (Lotz 1905).[6]

The administration of the pension system was decentralised and administered by regional authorities, the so-called Regional Insurance Agencies (*Landesversicherungsanstalten*). These Regional Insurance Agencies (RIAs) already administered the health insurance system and 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, 1992; 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 a focus of economic research (e.g. McNicholl 1980; Becker and Murphy 1988; Smith 1989; Guinnane and Ogilvie 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 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 Bismarckian PAYG pension system with pensions of a generation which are proportional to the their contributions.

[6] 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). Note that detailed regional information on wages is only available for selected professions.

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.[7] 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.[8] 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.[9] 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

[7] 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).

[8] 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.

[9] How such transfers from adults to their elderly parents can be enforced is subject of an extended literature about implicit contracts within the family, see e.g. Cigno (2006), Cigno et al.(2006), Sinn (2004).

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

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 individual pension, p_{t+1}^{BIS}, is higher than the average pension,

[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.

$(1+\overline{n_{t+1}})\tau w_{t+1}(1-f(\overline{n_{t+1}}))$, 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 $(1+\overline{n_{t+1}})\frac{w_{t+1}}{w_t}\frac{1-f(\overline{n_{t+1}})}{1-f(\overline{n_t})}$:

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

We assume that individuals take this effect into account when deciding on fertility. In a Bismarckian system pensions are proportional to individual wage income. If raising children reduces working time it should be obvious for rational individuals that it reduces pensions.

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

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

We assume this marginal price to be positive at an inner solution of the 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:

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

Now we consider the fertility decision in a PAYG pension 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:

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

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:

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

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,

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

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

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

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 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

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

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, 2004). 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.[12]

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 becomes sufficiently high, the fertility rate always decreases with an extension of the pension system, i.e. a higher contribution rate. However, in a dynamically efficient economy where the internal rate of return is lower than the capital market interest rate but the return of children in the pension system is higher than in the intra-family transfer system: $R_{t+1}>\Omega_{t+1}>B_{t+1}/\tau w_t f'(n_t)$, fertility falls only if the income effect is larger than the price effect, and vice versa. Hence, we can state:

Proposition 1 ***Price effect and income effect*** *In a dynamically efficient economy, the overall effect of a PAYG pension system on fertility is negative if the income effect overcompensates the price effect, and vice versa.*

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 refer to the analysis in Appendix B. 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 will be partially crowded out if the relative return of the pension system is higher relative to capital market savings and 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

[12]Note that without intra-family transfers ($B_t=B_{t+1}=0$) the price of a child increases and is always positive. The only effect of excluding such tranfers from the model is a stronger income effect.

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). Kaschke and Sniegs (2001) is the primary source for data on the functioning of the pension system while the Annual Yearbooks of Statistics are the primary source for control variables. Figure 1 shows the regional entities after harmonising the data sets.[13]

[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 inferred from historical data since individual-level data is hardly available. Typical fertility indices which are used in cross-country studies, like the total fertility rate (TFR)[14] require statistics on the age of the mother. However, the annual yearbooks of statistics provide information on crude (marital) birth rates[15] annually and on the age distribution in the population for some years (1871, 1885, 1890), which allows us to compute two different indices of fertility. First, we compute the crude marital birth rate (CMBR)[16] for all years. Second, we compute the Total Fertility Index (TFI) developed by Coale (1965, 1969) which is slightly more sophisticated as it takes into account natural fertility.[17] The TFI can

[13]The regional entities had to be made consistent, because the Annual Yearbook of Statistics covers 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.

[14]$TFR_t = \sum_{age=15}^{age=49} \frac{\left(Number\ of\ births_t^{age}\right)}{Women_t^{age} \cdot 1000}$

[15]In fact, the annual yearbooks of statistics provide information on total birth rates and on illegitimacy rates. We use both the total birth rate and the marital birth rate in our regression analyses and do not find significant differences except for the obvious influence of marriages on marital births and of health care availability on illegitimate births. The results are discussed with the other sensitivity analyses below.

[16]$CMBR = \frac{Number\ of\ births_t}{1000}$.

[17] The term *natural* fertility was coined by Henry (1961) and describes fertility in the absence of any deliberate birth control. The values for natural fertility used for the computation of the Total Fertility

take a maximum value of 1 if a society practices no birth control; otherwise it is always smaller than 1.

Despite being a comparatively crude measure, the CMBR maps fertility developments well. For the years for which we can compute the TFI, the regional distribution of the CMBR in Imperial Germany corresponds to the regional distribution of the TFI. Moreover, both the TFI and the CMBR are broadly in line with the information in Knodel (1974).

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. The fertility decision in our model is determined by the numerator of equation (6), the sign of which hinges on the difference between returns on savings R_{t+1} and the internal rate of return of the pension system Ω_{t+1}. Therefore, we focus our empirical model on identifying the effect of the internal rate of return of the pension system (taking the return on savings as given under certain conditions which are discussed below). However, we have to make sure that at the same time we control for all other factors that have shaped the decline in birth rates, which in our model are the main elements of the utility function. 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 how our econometric specification can capture the concept of additinve utility in our model and second, we have to consider the variables with which we can proxy the determinants of fertility.

First, consider the econometric model. Since we assume it to be additively separable in its arguments, we also assume that the main determinants of the fertility decision enter our econometric model additively. Taking into account that our data are available at the jurisdictional level, we can write the specification to be estimated:

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

The measure $n_{i,t}$ refers to the crude marital birth rate (CMBR) (or in our sensitivity analysis, the Total Fertility Index, TFI) in jurisdiction i in year t; we assume that the utility from the pleasure of having children is the same for all households, it enters the intercept y_{0_i}; T_t are time-specific effects; $\boldsymbol{x}_{\mathbf{i,t}}$ is a vector of variables which proxy current consumption (detailed below); $\boldsymbol{p}_{\mathbf{i,t}}$ is a vector of variables which proxy future consumption (detailed below); α_i refers to time-invariant region-specific effects and $\varepsilon_{i,t}$ is an i.i.d. error term.

Index are those in Henry (1961). $TFI = I_{t,i} = \frac{Number\ of\ births_t}{n_{g,i}F_{g,i}}$, where $n_{g,i}$ is the number of women in age group g in province i and $F_{g,i}$ is the natural fertility for age group g.

Second, consider the determinants of fertility that should be included. In our empirical specification of the model, we select those variables as proxies for current and future consumption that have been found to be the main determinants of the First Demographic Transition.[18] The theoretical representation of the factors that determine the decision to have children is consistent with those factors that have previously been found to be the main determinants of the First Demographic Transition: a general (child) mortality decline which increased returns to child quality (as more children survived the investment in their education became more valuable), innovation in contraception and the changed availability of contraception (which was spread by urbanisation and better communication), changes to the direct costs of children (e.g. due to compulsory schooling laws or laws banning child labour which reduced the household income), changes in the opportunity cost of children (since the labour market participation rate of women steadily increased), and the reduced value of children as an insurance against risk due to the introduction of comprehensive social insurance (Guinnane 2011 gives a nice summary of those factors, other studies are Galloway et al. 1994, 1998; Richards 1977; Brown and Guinnane 2007 and in particular Knodel 1974 for Germany). This should not be surprising given the fact that modern fertility theory (e.g. Becker 1960, 1965, 1988, 1991; Schultz 1969; Barro and Becker 1986, 1988, 1989; Easterlin 1975; Becker and Thomes 1976; Cigno and Ermisch 1989) has emerged from earlier, mostly empirical studies on the determinants of fertility, also in the historical context (e.g. the Princeton Fertility Project, refer to Coale 1965; Coale and Watkins 1986). Scheubel (2013) gives an account of the empirical studies of the First Demographic Transition, the determinants of the fertility decline that emerge, and how these studies link with the theoretical models of fertility.

According to our model, for capturing current consumption, we should include a measure for intrinsic utility of having children ($U(n)$), 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. To capture the intrinsic utility from having children, we include information on the number of marriages, the gender imbalances ratio (capturing migration) and (lagged) information on the share of Catholics relative to Protestans as proxies for the intrinsic utility of having children. The gender imbalances ratio measures the number of married women relative to married men. If it is significantly larger than 1, it provides and indication that relatively more men than women have (temporarily) left the province, most often to work in one of the quickly industrialising areas (e.g. Haines 1976; Nugent 1995). Therefore, it is a measure for work-related migration.

The diffusion of Catholicism, one of the standard control variables in demographic transition theory, can be expected to work largely though what in the terminology of the model is the intrinsic utility from children: not using use contraception is deeply rooted in Catholicism and motives such as providing a better education to fewer children or having

[18]This approach also helps us to reproduce previous findings on the First Demographic Transition, which shows that our proxies capture the main determinants that have been identified in the literature. Scheubel (2013) provides an extensive analysis for each of the variables used in this study.

a smaller family to have a larger disposable income were less prevalent among Catholics (e.g. Scheubel 2013). However, in contrast to other work on the topic (e.g. Galloway et al. 1994), we do not expect a variable on e.g. the share of Catholics in a provice to have a large effect for two reasons. First, despite the important effect of religion on culture and behaviour (e.g. Becker and Wössmann 2009) it is not obvious that this effect can be captured by annual (or even quinquennial as in our data) variation in the share of the population that is Catholic or Protestant or has a different religion. Since identification in a fixed effects panel estimation, which is the workhorse model in the literature (e.g. Galloway et al. 1994, 1998), would be derived from the annual (or as available here the quinquennial) variation in Catholicism or Protestantism, the coefficient in the model would reflect the underlying trends of this variation. However, it is probable that this variation is driven rather by migration. To instead capture the effect of absolute (largely time-invariant) numbers of Catholics, some authors have used OLS including e.g. the level of Catholics (Goldstein and Klüsener 2010). However, exactly this influence of the level of Catholics should already be well-captured in a fixed effects model: if the level of Catholics is invariant over the time span analysed, the influence is captured in the unobserved fixed effects. Alternatively, if the variation in the share of Catholics does not only reflect migration, it could simply be related to the birth rate being higher among Catholics since an increase in the number of Catholics in a province is also driven by the number of children born into a Catholic family. In other words, measures of the diffusion of religion are probably highly endogenous to the birth rate. Our approach is to assume that level effects are captured by the fixed effects in our model while we control for short-term variation in the share of Catholics relative to Protestants (which were the major religious groups) by including a variable that measures the difference between the number of Catholics and the number of Protestants relative to the overall number of Catholics with a 10 year lag to avoid the endogeneity issues.

As control variables for current consumption $\boldsymbol{x}_{\mathbf{i,g,t}}$ and the cost of having children, we use total contributions to the pension system and the diffusion of education.

Contributions to the pension system serve as a proxy for income since these contribution were, like progressive taxes, directly related to income. Like in our model, having children implied foregone wage income, both because of fewer hours worked and because of larger household expenses. Even though children could later on contribute to household income, compulsory schooling lowered this contribution. To proxy the diffusion of (mandatory) basic education, we include the share of recruits without at least basic education.

The old age dependency ratio helps to capture the cost of intra-family transfers (B_t) and related effects on the budget constraint. To capture the fact that children would also contribute to household income, particularly in working class families, we compute the share of workers in each province on the basis of the respective population census.

In addition to these variables which are also related to the microeconomic theories of fertility, we add variables that help us to capture general effects of industrialisation and urbanisation. For industrialisation, we use the share of the population working in the primary, the secondary, and the tertiary sector respectively. For urbanisation, we add three variables: the (lagged) number of people per building, the number of cities with 5.000-20.000 inhabitants and the number of cities with more than 20.000 inhabitants. Regarding these urbanisation variables, the number of births also directly affects the number of people in a building. Therefore, we also use this variable with a lag. As the information on the number of people in a building is only available for 1880 and 1885, this makes a 15-20 year lag when using it in an estimation for the years 1895 and 1907, which are the key years in our analysis since most variables are available for those two years. A lag of at least 15 years is important to ensure that we are not measuring population growth, but living conditions. Endogeneity should be less of an issue when considering the number of cities with more than 20.000 inhabitants and the number of cities with 5.000-20.000 inhabitants. For this variable we use the observations 'closest' to the main years of the analysis; 1880 and 1905. Note that the number of workers in a province is not endogenous to the birth rate since the population censuses report the number of people in certain professions and their dependants separately. As industrialisation was also related to increasing female labour force participation we add the share of contributions to the pension system in category I (the lowest category) relative to contributions in all other categories since category I was considered the 'shrews' category[19]

For capturing future consumption, we have to 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}. Since pension insurance was introduced as the last pillar of social insurance (Scheubel 2013) we have to take into account fertility effects related to the provision of social security and not to the particular effects of public pension insurance, and the provision of health insurance in particular. First, to capture the enrollment effect in social security we include the number of insured in the pension system in 1895 and 1907, which was collected by the regional insurance agencies as part of the data collection for the population censuses during these years. Second, to capture the impact of health insurance, we include the number of health insurance agencies per 1000, which helps to proxy health insurance coverage.

In our model, the sign of the fertility effect of the pension system in equation (4), i.e. $\frac{\partial n}{\partial \tau}$, depends on the determinants in equation (16): the reduction in the opportunity cost of having children in terms of foregone lifetime income (which 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). The price effect is related to income which we proxy by the total amount of contributions. The income effect is determined by the internal rate of return of the pension system. For estimating the effect of the internal rate of return of the pension system Ω, we include the net assets accumulated by a Regional Insurance Agency (RIA) divided by the total number of existing and expected pensioners, i.e. the number of existing pension claims plus the number of approved pension claims. Since RIAs also differed in their ability to accumulate assets (Kaschke and Sniegs 2001), we also include the total amount of net assets as a separate variable.

[19]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.

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, such interactions would require a simultaneous equations estimation approach, which we cannot pursue since we do not have a reliable time-varying measure of savings for all provinces of Imperial Germany. However, note that according to equations (4) the optimum number of children is determined by $U_z\left(\frac{\partial p_{t+1}}{\partial n_t} + B_{t+1}\right) + U_n - U_c\left[(1-\tau)w_t f'(n_t) + \pi_t\right]$. At the same time, we know from equation (5) that $U_c = U_z R_{t+1}$ 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, assuming that the diffusion of private saving opportunities across households did not change over the time span we analyse, this requires a measure of the capital market rate of return. Since we do not have a measure of the capital market rate of return, we include the productivity in agriculture, as the productivity in agriculture is typically positively linked with industrialisation and the productivity of capital (O'Brien and Prados de la Escosura 1992) as well as economic growth in general (Murphy et al. 1989; Gollin et al. 2002).

All regions differ with respect to population size, therefore all variables not expressed in percentage terms or other shares are normalised to population size. We provide the summary statistics for all variables in table 1.

[Table 1 about here.]

4.3 Econometric considerations

Since we use historical data, we have to deal with variables which have not been collected by the Imperial Statistical Office for all years.[20] The pension system variables as well as the birth rate are available for all years from 1891 to 1914. Variables on the demographic structure or the share of the population working in the primary, secondary or tertiary sector, and the number insured in the pension system have only been collected in the population censuses and are thus only available for two points in time.[21] Other variables, such as age structure were only available during earlier years. For example, demographic information was mostly collected in connection with the occupational censuses, which were only conducted every five years. As we adjust most variables to population size to make the numbers comparable, we extrapolate population figures for the years for which

[20] The data collected for Imperial Germany by the Imperial Statistical Office is not as detailed as Prussian data, which has been used for similar analyses before (Becker and Wössmann 2009; Becker et al. 2010, 2011; Hornung 2014). One of the reasons for the different level of detail is that information had to be harmonised for all parts of Imperial Germany, not all of which collected data as detailed as the data collected by the Prussian Statistical Office. Scheubel (2013) provides an overview of the history of data collection and Sniegs (1998) gives a detailed account of the role of historical statistics in Germany.

[21] Between the introduction of pension insurance in 1891 and 1914, only two population censuses were conducted.

population figures are not available. [22] Variables that capture urbanisation, like the size of cities, are available only with large time spans in between. In addition, some of the proxies we use are likely to be endogenous to fertility, at least the contemporaneous observations, in particular religion and the variables that can be related to family size, such as the number of people in a building. Table 2 details the data availability.

[Table 2 about here.]

To control for unobserved province-specific effects while taking into account the limited availability for some variables, we follow an approach that is common in the literature using historical data, known as pooled cross-section time series methods (e.g. Galloway et al. 1994, 1998; Becker et al. 2010, 2011). We construct a data set consisting of two periods, t and $t - x$ where $x \in 1, 2, 3, ...$ We choose years t and $t - x$ such that most variables are available in both years. For those variables which are not available in t or $t - x$ we use the observations for other years instead. Since this can introduce biases if the time span between the observations is too large, we discuss the selection of variables and the years of availability in detail below. Moreover, we provide several robustness checks regarding the selection of years.

Another complication arises from the fact that the data are collected at the province level. This has two major implications. First, we have to account for unobserved province-specific effects. Second, standard errors may be spatially correlated. The typical approach in the literature is a fixed effects panel specification (e.g. Galloway et al. 1994). To account for the invariant region-specific effects, we use a fixed effects model with standard errors adjusted for serial correlation. Second, 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 also use a model robust to spatial correlation.

5 Results

5.1 Descriptive Analysis

A sustained fertility decline started in Imperial Germany only during the 1890s, which is also when the pension system was introduced. While fertility declined in some provinces of Imperial Germany already in the 1880s, fertility rose again in some provinces in the late 1880s and early 1890s before it declined more sustainedly towards the end of the 1890s (figure 2). This is remarkable, first because the industrialisation process had started earlier and second because these points in time coincide with the introduction of the pension system. The pension system came into effect in 1891 as a partially funded system and the changes that turned it into a pay as you go system came into effect in 1900 (Scheubel 2013).

[Figure 2 about here.]

[22] Refer to Scheubel (2013) for the derivation of the extrapolated numbers.

The internal rate of return, proxied by the net assets accumulated by a Regional Insurance Agency (RIA) relative to current plus expected number of pensioners, for four representative provinces shows two main trends (figure 3). First, in some provinces, the net assets per pensioner increased more or less steadily as the RIAs accumulated assets, like in Westpreußen or in Pfalz. Second, in other provinces,like Baden or Hessen-Nassau, net assets per pensioner dropped towards the end of the 1890s and increased again only after approximately 1903. The drop of net assets per pensioner towards the end of the 1890s in these provinces probably reflects the system switch to a fully funded system with a regional equalisation scheme between RIAs. As inherent in equalisation schemes, some RIAs were more affected than others, depending on the demographic structure in the respective region. For example, the demographic structure in all of the Prussian provinces in the East (including Westpreußen) was disadvantageous, so provinces there benefitted comparatively more from an equalisation scheme between RIAs. The fact that after 1904 developments in net assets per pensioner were very similar across provinces lends additional support to this interpretation.

[Figure 3 about here.]

When plotted against net assets per pensioner (figure 3), it becomes apparent that fertility fell as net assets per pensioner increased. While fertility had declined constantly in some provinces already since the 1880s, a sustained fertility decline, i.e. a real change in trend in all provinces, can only be inferred from approximately 1902 (figure 2). This coincides with the time when there was a sustained increase in net assets per pensioner (figure 3). Moreover, the temporary increase in birth rates during the late 1890s in some provinces took place during a time when there was a temporary drop in net assets per pensioner. This lends support to our initial hypothesis that the pension system had an impact on fertility behaviour.

5.2 Multivariate Analysis

The negative relationship between net assets per pensioner and birth rates persists when controlling for other determinants of the first demographic transition. Table 3 shows four specifications to illustrate this.

Column (1) gives the correlation between the crude marital birth rate (CMBR) and marriages and total contributions per insured, which capture the basic elements of the consumption model of fertility, i.e. intrinsic utility from children (here proxied by marriage) and income (here proxied by contribution revenues).[23]

Column (2) adds demographic variables to this: the level of education proxied by the share of recruits without at least basic schooling, a proxy for migration, the old age dependency ratio, the share of workers in each province, the share of Catholics and a proxy for health insurance coverage.

[23] The traditional approach in the literature is to use a measure of tax revenues. Since there is no reliable data for tax collection at the federal level, we use total contributions to the pension system instead since contributions were proportional to income and thus to tax revenues. (Scheubel 2013)

Column (3) adds those variables which have been found to be the main determinants of the first demographic transition and which are particularly related to industrialisation. Urbanisation is captured by the number of people per building and the number of large and medium-sized cities. Moreover, we add information on the share of the population working in the primary, the secondary and the tertiary sector. To capture increasing productivity, both of capital and labour, we calculate a crop yield index as increases in agricultural productivity, being a pre-condition for industrialisation (Murphy et al. 1989), have been found to closely correlate with overall productivity (e.g. Dowrick and Gemmell 1991) and with growth (O'Brien and Prados de la Escosura 1992; Gollin et al. 2002). Since female labour force participation has increased in response to industrialisation, we also include a measure of female labour force participation: we include the share of contributions that were collected in contribution category I (relative to contributions in other categorties), since category I was the women category (Haerendel 2001; Scheubel 2013). To account for the existence of social insurance in general, we include the share enrolled in pension insurance, since pension insurance also provided some health-related services (Kaschke and Sniegs 2001).

Column (4) adds three main indicators of pension insurance coverage: the share of the population covered by pension insurance, the internal rate of return (measured as net assets of a regional insurance agency per pensioner) and the net assets of each regional insurance agency.[24]

[Table 3 about here.]

Table 3 shows that there is a significant negative effect of higher net assets per pensioner on fertility, even when including the other determinants of the first demographic transition. This effect is equivalent to a reduction of marital fertility by approximately 1.7 births per 1000 if net assets per pensioner increase by 1000 Mark between 1895 and 1907. Since net assets per pensioner varied between 1000 and 5000 Mark in 1907, a reduction of marital fertility by approximately 2 births per 1000 would have required a doubling of net assets per pensioner in some provinces and at least an increase of 20% in others. In 1895, the average net assets per pensioner for Imperial Germany was approximately 1100 Mark per pensioner. This figure increased to approximately 1400 Mark in 1907. Therefore, since on average for Imperial Germany births per thousand decreased by 6 births, our estimates suggest that the pension system had contributed a reduction of 0.8 births per thousand to the overall decline in birth rates in Imperial Germany. To put this into perspective, take the degree of urbanisation, proxied by number of persons per building, which increased from approximately 9.6 persons on average for Imperial Germany in 1880 to 9.9 persons in 1885. Therefore, the associated decrease in birth rates should have been approximately 0.33 births per 1000. In other words, column (4) suggests that the impact of changes in net assets per pensioner was approximately twice the effect of urbanisation. The effect of industrialisation measured by the share of the

[24] We have also tested specifications with more variables on the functioning of the pension system, such as average pensions; however, these do not add information and remain insignificant in all tested models.

population working in the secondary sector is approximately 1/6 of the effect of changes in net assets per pensioner.

The change in net assets per pensioner varied substantially between provinces between 1895 and 1907, indicating that the contribution of pension insurance to the decline in birth rates must have varied as well. Net assets per pensioner increased by almost 90% in Braunschweig, but decreased by almost 21% in Wuerttemberg. During the same period, birth rates declined between 8 and 0.7 births per 1000. However, the considerable variation across provinces also suggests that the other determinants of the demographic transition had a varying impact, too.

In fact, our results are in line with standard demographic transition theory. Column (1) renders support to the consumption theory of fertility. The coefficients suggests that one marriage per 1000 leads to approximately 1 birth per 1000. A higher contribution capacity in a province is also associated with a higher birth rate: a higher contribution capacity by 1 Mark per insured is associated with 0.4 more births per 1000. Note that contributions per insured ranged from 6.50 Mark in Ostpreußen to 42.6 Mark in Oldenburg in 1895. This effect is reduced if we add demographic variables in column (2), albeit remaining significant even when adding pension variables in column (4). Column (2) confirms standard demographic variables as determinants of fertility. For example, the old age dependency ratio appears as a main determinant of the birth rate, rendering support to the hypothesis of the importance of inter-generational transfers for old age provision. While both columns (2) and (3) suggest that a 1% increase in the old-age dependency ratio reduces birth rates by 1 per 1000 (which may be related to the need to devote more resources to intra-family transfers), the effect is much smaller and insignificant in column (4). This suggests that the variable may have also captured ageing dynamics which are reflected in the dynamics of net assets per pensioner as well. In our view this underpins the information content of the variable measuring the internal rate of return of the pension system (net assets per pensioner).

Turning to the variables related to industrialisation, the share of workers in a province is associated with a higher birth rate; one percentage point increase in the share of workers would lead to 0.08 more births per 1000. The share of people working in the secondary sector also has a negative effect, which is however only significant in column (4).[25] Urbanisation, captured by persons per building has a significant negative effect on the number of births, both in columns (3) and (4).

It is reasonable to assume that not only pension insurance changed people's behaviour, but that in fact the major game changer was the whole package of social insurance introduced at the time. If pension insurance has an impact on private insurance for old age, then reducing the risk of poverty due to inability to work more in general should also have an effect. Therefore, it would make sense to assume that other insurance like health care coverage should also have an effect on fertility. If we do not control

[25] This negative coefficient may also be related to the fact that miners' associations provided pension insurance before the introduction of comprehensive health insurance (Jopp 2013). In other words, people working in this sector had been exposed to pension insurance before; thus it should not be surprising that the share of the population working in this sector is negatively associated with the birth rate.

sufficiently for this effect, then any effect we find for pension insurance may overestimate the true pensions-fertility nexus. In our baseline model in columns (1)–(4) we have added the number of insured as recorded by the RIAs in 1895 and 1907 to account for this. However, it could not be sufficient to capture health care effects. Hence, we add another measures of health care coverage in column (5): the number of health insurance agencies per 1000 as a measure of health care coverage. The internal rate of return is significant and of the same magnitude as in the baseline specification. At the same time, the number of health insurance agencies per 1000 is not significant.

5.3 Sensitivity

5.3.1 Estimation approach

While it may seem straightforward to use a fixed effects estimator with standard errors adjusted for serial correlation for the case presented in this paper, it may be helpful to illustrate the robustness of the results to the use of other approaches.[26] Table 4 presents an OLS model in column (1), our baseline model in column (2) and a first differences estimator in column (3).

[Table 4 about here.]

A standard OLS model would suffer from several endogeneity issues, such as clustered standard errors and serial, potentially also spatial correlation. Presenting the OLS model (with standard errors robust to at least serial correlation and some clustering) in this context helps to illustrate the importance of controlling for the unobserved fixed effects. In particular, note that the OLS estimates differ in two important respects from our baseline model. First, the coefficients from our baseline model tend to be either overestimated or underestimated by the OLS approach. Second, even though standard errors are adjusted for some clustering as well as for serial correlation, the OLS model sometimes indicates significant estimates while the fixed effects model does not. At the same time, the OLS model is able to indicate the relative size of the different effects fairly well.

In theory, first differencing should yield exactly the same inference as a fixed effects model when the fixed effects model is applied to only two time periods. This is illustrated when comparing columns (2) and (3). The coefficients are the same while standard errors are larger in the model in first differences. This should not be surprising given that the first differences model is less efficient. Losing a degree of freedom in a model with only a small number of cross-sectional observations potentially has a big impact on the precision of the estimates. Therefore, we present estimates with a reduced number of covariates in column (4) to illustrate that while coefficients potentially suffer from omitted variable bias in this model, they are not substantially different from our baseline model and as conjectured, the precision of estimates is increased somewhat.

[26] We refrain from discussing the option of using a random effects model here; it is obvious that we have to control for non-random unobserved province-specific effect. This notion is also confirmed by a simple Hausman test.

Assuming that the province-specific unobserved effects are well-captured in a fixed effects model, the model may not sufficiently control for spatial correlation. For example, if the decline in birth rates is correlated for adjacent provinces, this will lead to a correlation between the province-specific effects $\alpha_{i,g}$ with the error term $\varepsilon_{i,g,t}$. One option to deal with this potential endogeneity issue is introducing a spatial lag and adjusting the standard errors accordingly (e.g. Anselin 1988). Another option is to correct standard errors using non-parametric techniques (e.g. Driscoll and Kraay 1998, Conley and Molinari 2007). However, given the small sample size and the limited effective time dimension ($T = 2$), standard asymptotics do not hold.[27] Alternatively, considering the correlation between provinces as a form of cluster correlation, we can use techniques for cluster-robust inference (e.g. Wooldridge 2003; Cameron et al. 2008; Cameron and Miller 2010; Cameron and Gelbach 2011), some of which are better behaved in small samples (e.g. Donald and Lang 2007).

Using additional cluster-robust inference methods confirms our main results. The approaches we pursue here are parametric approaches since our sample is too small for non-parametric methods. To implement these methods, we need to have an idea about the type of clustering that takes place. There are three different types of clusters we investigate: clustering according to membership in the Kingdom of Prussia, clustering according to having a large Slav minority and clustering according to the type of region (rural, industrial or mixed). The latter two types of clusters have already been identified by Kaschke and Sniegs (2001).[28] For all three types of clustering, we pursue an exercise in which we collapse data to the cluster level (Bertrand et al. 2004). Essentially, the main effects are the same as those in our baseline model.

5.3.2 Other policy changes

Since the late 1890s and the early 1900s were a time of industrial change, but also of cultural and political changes, it is important to rule out that we measure a time trend that captures other effects than pension insurance. For example, stricter child labour laws reduce the scope for current consumption as children go to school instead of contributing to household income. This should lead to a lower number of children ceteris paribus. In fact, there were some reforms of child labour laws in the period we study in this paper (Boentert 2007).[29] Moreover, there were a few changes to the pension

[27] We ran the basic model only with the variables which are available for more than just a few periods (marriages, agricultural productivity, education, share of contributions in category I) to implement a Driscoll and Kraay (1998) adjustment of the standard errors. While the model obviously suffers from omitted variable bias, the magnitude of the variables of interest, in particular of net assets per pensioner, is broadly the same.

[28] As mentioned in Kaschke and Sniegs (2001) RIAs differed structurally in terms of being more agricultural or more industrial based. Kaschke and Sniegs (2001) define RIAs as agricultural, industrial or mixed. We use their definition for defining the clusters according to type of region.

[29] There were three major changes to legislation during the period we study: changes to the *Gewerbeordnungsnovelle* (amendments to the Industrial Code) in 1878 and 1891 and a law banning child labour in 1903 (Boentert 2007). Importantly, the amendments to the Industrial Code did not affect child labour in all areas of production. The 1878 amendment prohibited children below the age of 14 to work in fac-

system itself. In 1900 the system switched to a pure pay as you go system. A new wage category was added to the pension system. In 1903 and 1904, the associated introduction of a financial equalisation scheme between RIAs prompted the Federal Insurance Agency to conduct a review of RIAs' code of conduct (Kaschke and Sniegs 2001). To address concerns that these changes – which happened between 1895 and 1907 – drive the effects we measure, we select other years between 1890 and 1912 to show that the effect persists even when we select two periods before 1900.[30] Even though we use different years for the estimates in table 5, some control variables are only available for a few years and thus it should not be surprising to see the coefficients of the control variables vary.

[Table 5 about here.]

Except for the coefficients on the share insured, the share employed in the trade sector and the number of persons per building, results are qualitatively consistent irrespective of the choice of years. The share employed in trade (and in other sectors) and the share insured are only available for 1895 and 1907; the number of people per building only in 1871, 1880, and 1885.

Regarding the share insured, for the specifications in columns (1) – (5) we only use the numbers for 1895 and an estimated number for 1882.[31] We do this to be particularly cautious since the 1903 or 1904 birth rate could also be correlated with the share insured in 1907 or the share working in a particular sector. The latter variables are potentially also correlated with eligibility criteria for pension insurance since eligibility for pension insurance and the functioning of the system changed. As we cannot measure the change in eligibility criteria, we would introduce endogeneity into the model. This implies that the coefficient on the share of insured and the share in trade should be interpreted with caution, particularly when our dependent variable is from the 1900s and the control variable is from the 1880s.

Regarding the number of people per building, we use this variable with a lag since contemporaneous values are likely to be endogenous to the birth rate. However, this also means that the appropriate lag of at least 15 years is only given for years 1900 or later. Thus, it should not be surprising to see the expected negative effect mainly for years after 1900.

The other control variables used in the model paint a consistent picture of birth rates being positively influenced by new marriages and higher incomes, and negatively influenced by those factors related to industrialisation (such as the move to industrialising areas, employment in the secondary sector or urbanisation).

The contribution of the internal rate of return of the pension system to the decline in birth rates is consistent, albeit decreasing. The coefficient on the internal rate of return

tories. After 1891, this prohibition was extended to workshops and production at home, such as spinning and weaving. The general law from 1903 extended this also to agricultural production. Probably, the changes in 1891 had the comparatively largest impact on household income. However, birth rates only started their sustained decline during the 1900s in all provinces.

[30] We exclude years after 1912 as the years preceding the First World War were affected by increased military spending and the mobilisation of troops.

[31] The procedure for estimating this number is described in Scheubel (2013).

of the pension system is larger than in our baseline specification in columns (1) and (2), which comprise the 1890s. In contrast, it is smaller than in our baseline specification in columns (3) to (5) and insignificant in column (6). This may be an indication that the effect was stronger in the 1890s. Another explanation could be the introduction of the financial equalisation scheme between RIAs that was introduced during the early 1900s, which implies that there was less variation in the asset structure between provinces after the financial equalisation scheme was put into place. Consequently, identification through variation between provinces would also be more difficult. In addition, we find a significant effect even when limiting the time horizon to years that exclude changes in child labour legislation, such as 1895 and 1899 and 1903 and 1907. Given that the effect is visible in the data even when varying the time horizon and even when looking at a period when no other policy changes took place lead us to conclude that the impact of pension insurance on fertility is robust.

As gradual policy changes or time trends, such as the effects of increasing industrialisation, are more difficult to measure, we check for the impact of excluding large industrialised cities or provinces instead. The impact of excluding cities or regions depends on the importance of a province or city, but generally just excluding one observation does not make a significant difference. In contrast, excluding a full region (say Eastern Prussia for example) changes results. However, this should not be surprising given the small sample size.

5.3.3 Measuring fertility

While we have already discussed that the marital birth rate is an accepted measure of fertility, especially in the historical context, we show that other measures of fertility give comparable results for the periods for which we can compute these alternative measures. One caveat to this is that we cannot include all control variables from our baseline specification since these were not available for the years 1880 and 1885, which are the years for which the age structure of the population is available. However, the results even for a reduced model are broadly in line with our baseline model and help to illustrate that the use of the crude marital birth rate instead of more sophisticated fertility indicators yields reliable results.

Table 6 shows a model of fertility for the years 1890 and 1885. Thus we can compute the Total Fertility Index and the Marital Fertility Index. Column (1) shows the crude birth rate and column (2) the corresponding Total Fertility Index. Column (3) shows the marital birth rate used throughout this paper and column (4) shows the corresponding Marital Fertility Index. The model includes all those determinants of fertility which are available for 1885 and 1890. It is obvious that models (1) and (2) as well as models (3) and (4) are comparable in terms of the variables which they indicate as important determinants of fertility. Therefore, we conclude that using the crude marital birth rate in our model gives results that do not need to be qualified by the fact that we cannot control for the age structure of mothers.

[Table 6 about here.]

To show that the similarity is not driven by the specification of the model being too inflexible, we show that the model can indentify different determinants in case the dependent variable measures something different: column (5) shows the same model, but uses the share of non-marital births as dependent variable. As expected, the model identifies other variables as the main determinants of the non-marital birth rate. For example, the model predicts correctly that the share of non-marital births should fall as the number of marriages rises.

6 Conclusions

Our paper provides a theoretical underpinning and an empirical confirmation of the negative relationship between statutory old-age 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. 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 negative effect of a higher 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 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 1.7 marital births per 1000 between 1891–1899 for an increase between 20% and 100% of the internal rate of return.

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 more than 10% 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

Abio, G., Mathieu, G., and Patxot, C. (2004). On the optimality of payg pension systems in an endogenous fertility setting. *Journal of Pension Economics and Finance*, 3:35–62.

Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers, Dordrecht.

Arroyo, C. R. and Zhang, J. (1997). Dynamic microeconomic models of fertility choice: A survey. *Journal of Population Economics*, 10:23–65.

Barro, R. J. and Becker, G. S. (1986). Altruism and the economic theory of fertility. *Population and Development Review*, 12:69–76.

Barro, R. J. and Becker, G. S. (1988). A reformulation of the economic theory of fertility. *Quarterly Journal of Economics*, 103:1–25.

Barro, R. J. and Becker, G. S. (1989). Fertility choice in a model of economic growth. *Econometrica*, 57:481–501.

Becker, G. S. (1960). An economic analysis of fertility. In of Economic Research, N. B., editor, *Demographic and Economic Change in Developed Countries*, pages 225–256. Princeton University Press, Princeton.

Becker, G. S. (1965). A theory of the allocation of time. *Economic Journal*, 75:493–517.

Becker, G. S. (1988). Family economics and macro behavior. *American Economic Review*, 78:1–13.

Becker, G. S. (1991). *A treatise on the family*. Harvard University Press, Harvard.

Becker, G. S. (1992). Fertility and the economy. *Journal of Population Economics*, 5:185–201.

Becker, G. S. and Tomes, N. (1976). Child endowments and the quantity and quality of children. *Journal of Political Economy*, 84:S143–S162.

Becker, S. O., Cinnirella, F., and Wössmann, L. (2010). The trade-off between fertility and education: evidence from before the demographic transition. *Journal of Economic Growth*, 15:177–204.

Becker, S. O., Cinnirella, F., and Wössmann, L. (2011a). The effect of investment in children's education on fertility in 1816 prussia. *Cliometrica*, 6:1–16.

Becker, S. O., Hornung, E., and Wössmann, L. (2011b). Education and catch-up in the industrial revolution. *American Economic Journal: Macroeconomics*, 3:92–126.

Becker, S. O. and Wössmann, L. (2009). Was weber wrong? a human capital theory of protestant economic history. *Quarterly Journal of Economics*, 124:531–596.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics*, 119:249–275.

Billari, F. C. and Galasso, V. (2009). What explains fertility? evidence from italian pension reforms. CESifo Working Paper.

Boentert, A. (2007). *Kinderarbeit im Kaiserreich 1871-1914).* Schöningh, Paderborn.

Boldrin, M., Nardi, M. D., and Jones, L. E. (2005). Fertility and social security. NBER Working Paper.

Brown, J. C. and Guinnane, T. W. (2007). Regions and time in the european fertility transition: problems in the princeton project's statistical methodology. *Economic History Review*, 60:574–595.

Cameron, A. C. and Gelbach, J. B. (2011). Robust inference with multiway clustering. *Journal of Business and Economic Statistics*, 29:238–249.

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics*, 90:414–427.

Cameron, A. C. and Miller, D. L. (2012). Robust inference with clustered data. In Ullah, A. and Giles, D. E. A., editors, *Handbook of Empirical Economics and Finance*, pages 1–28. Chapman & Hall/CRC, Boca Raton.

Cigno, A. (1993). Intergenerational transfers without altruism. family, market and state. *European Journal of Political Economy*, 9:505–518.

Cigno, A., Casolaro, L., and Rosati, F. C. (2003). The impact of social security on saving and fertility in germany. *FinanzArchiv*, 59:189–211.

Cigno, A. and Ermisch, J. (1989). A microeconomic analysis of the timing of births. *European Economic Review*, 33:737–760.

Cigno, A. and Rosati, F. C. (1992). The effects of financial markets and social security on saving and fertility behaviour in italy. *Journal of Population Economics*, 5:319–341.

Cigno, A. and Rosati, F. C. (1996). Jointly determined saving and fertility behaviour: Theory, and estimates for germany, italy, uk and usa. *European Economic Review*, 33:1561–1589.

Cigno, A. and Werding, M. (2007). *Children and Pensions.* MIT Press, Cambridge, Massachusetts.

Cleland, J. and Wilson, C. (1987). Demand theories of the fertility transition: An iconoclastic view. *Population Studies*, 41:5–30.

Coale, A. and Watkins, S. (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.

Coale, A. J. (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 Trussell, J. (1974). Variations in model fertility schedules: The age structure of childbearing in human populations. *Population Index*, 40:185–258.

Conley, T. G. and Molinari, F. (2007). Spatial correlation robust inference with errors in location or distance. *Journal of Econometrics*, 140:76–96.

Cremer, H., Gahvari, F., and Pestieau, P. (2008). Pensions with heterogenous individuals and endogenous fertility. *Journal of Population Economics*, 21:961–981.

Donald, S. G. and Lang, K. (2007). Inference with difference-in-differences and other panel data. *Review of Economics and Statistics*, 89:221–233.

Dowrick, S. and Gemmell, N. (1991). Industrialisation, catching up and economic growth: A comparative study across the world's capitalist economies. *Economic Journal*, 101:263–275.

Driscoll, J. C. and Kraay, A. C. (1998). Consistent covariance matrix estimation with spatially depedent panel data. *Review of Economics and Statistics*, 80:549–560.

Easterlin, R. A. (1975). An economic framework for fertility analysis. *Studies in Family Planning*, 6:54–63.

Ehrlich, I. B. and Zhong, J.-G. (1998). Social security and the real economy. *American Economic Review*, 88:151–157.

Feldstein, M. (1974). Social security, induced retirement, and aggregate capital accumulation. *Journal of Political Economy*, 82:905–926.

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

Galloway, P. R., Hammel, E. A., and Lee, R. D. (1994). Fertility decline in prussia, 1875-1910: A pooled cross-section time series analysis. *Population Studies*, 48:135–158.

Galloway, P. R., Hammel, E. A., and Lee, R. D. (1998). Urban versus rural: Fertility decline in the cities and rural districts of prussia, 1875 to 1910. *European Journal of Population*, 14:209–264.

Goldstein, J. R. and Klüsener, S. (2010). Culture resurfaces – a geographic analysis of fertility decline in prussia. MPIDR working paper 2010-012.

Gollin, D., Parente, S., and Rogerson, R. (2002). The role of agriculture in development. *American Economic Review*, 92:160–164.

Guinnane, T. (2011). The historical fertility transition: A guide for economists. *Journal of Economic Literature*, 49:589–614.

Guinnane, T. and Ogilvie, S. (2008). Institutions and demographic responses to shocks: Württemberg, 1634-1870. University of Yale Economics Department Working Paper No. 44.

Haerendel, U. (2001). *Die Anfänge der gesetzlichen Rentenversicherung in Deutschland. Die Invaliditäts- und Alterssicherung von 1889 im Spannungsfeld von Reichsverwaltung, Bundesrat und Parlament.* Forschungsinstitut für öffentliche Verwaltung, Speyer.

Haines, M. R. (1976). Population and economic change in nineteenth-century eastern europe: Prussian upper silesia, 1840-1913. *Journal of Economic History*, 36:334–358.

Hirschman, C. (2001). Globalization and theories of fertility decline. *Population and Development Review*, 27:116–125.

Hohm, C. F. (1975). Social security and fertility: An international perspective. *Demography*, 12:629–644.

Hornung, E. (2014). Immigration and the diffusion of technology: the huguenot diaspora in prussia. *American Economic Review*, 104:84–122.

Jopp, T. (2013). *Insurance, Find Size, and Concentration.* Akademie Verlag, Berlin.

Kaschke, L. and Sniegs, M. (2001). *Kommentierte Statistiken zur Sozialversicherung in Deutschland von ihren Anfängen bis in die Gegenwart. Band 1: Die Invaliditäts- und Alterssicherung im Kaiserreich (1891-1913).* Scripta Mercature Verlag, St. Katharinen.

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.

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 Eversley, D. and Köllmann, W., editors, *Population Change and Social Planning*, pages 62–87. Erward Arnold, London.

McNicoll, G. (1980). Institutional determinants of fertility change. *Population and Development Review*, 6:441–462.

Meier, V. and Wrede, M. (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.

Murphy, K. M., Shleifer, A., and Vishny, R. (1989). Income distribution, market size, and industrialization. *Quarterly Journal of Economics*, 104:537–564.

Neher, P. A. (1971). Peasants, procreation, and pensions. *American Economic Review*, 61:380–389.

Nugent, J. B. (1985). The old-age security motive for fertility. *Population and Development Review*, 11:75–97.

O'Brien, P. K. and de la Escosura, L. P. (1992). Agricultural productivity and european industrialization 1890-1980. *Economic History Review*, 45:514–536.

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. (2013). *Bismarck's Institutions - A Historical Perspective on the Social Security Hypothesis*. Mohr Siebeck, T'ubingen.

Scheubel, B. D. (2014). Was bismarck's pension system bismarckian? a note on the design of the first comprehensive pension system. Working Paper.

Schultz, T. P. (1969). An economic model of family planning and fertility. *Journal of Political Economy*, 77:153–180.

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

von Auer, L. and Büttner, B. (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 Kohl, H., editor, *Die politischen Reden des Fürsten Bismarck, Zwölfter Band*, pages 597–616. Verlag der J. G. Cottaschen Buchhandlung, Stuttgart.

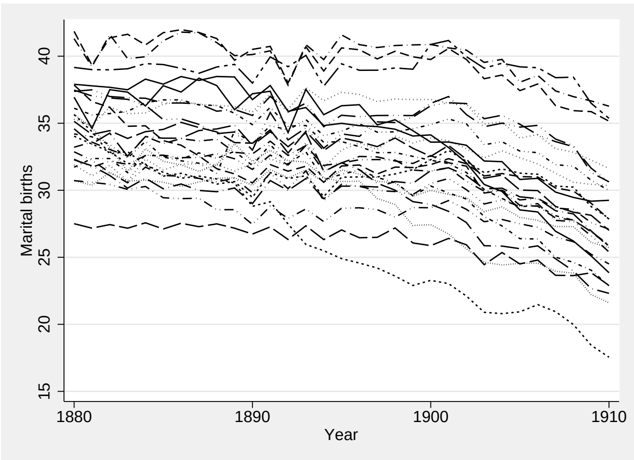
Wooldridge, J. M. (2003). Cluster-sample methods in applied econometrics. *American Economic Review, Papers and Proceedings of the One Hundred Fifteenth Annual Meeting of the American Economic Association*, 93:133–138.

Appendix A

Figures

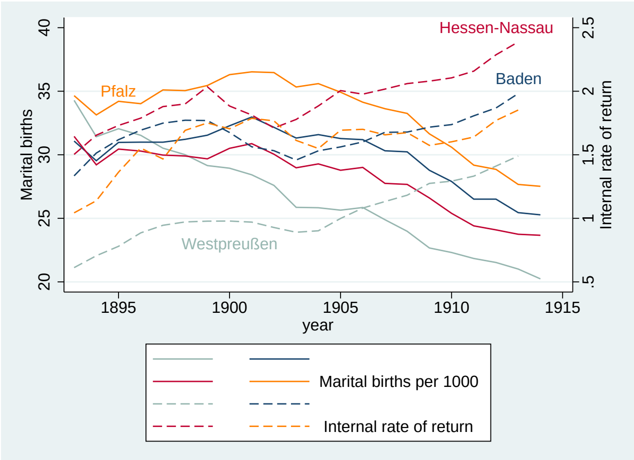
Figure 1: Regions in Imperial Germany

Figure 2: Marital birth rates in Imperial Germany



Notes: Marital birth rates for all regional entities in the data set, expressed in per 1000. For the sake of illustrating trends, region names are suppressed.

Figure 3: Marital birth rates and the internal rate of return



Notes: Selection of provinces only for illustrative purposes; trends are similar in all provinces. Internal rate of return expressed in assets per pensioner (unit: 1000 Mark). Birth rates expressed in per 1000.

Tables

Table 1: Summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Marital births (p 1000)	827	32.22228	4.548677	20.64291	43.36475
Marriages (p 1000)	851	7.795968	0.8512005	0.8228995	22.10484
Revenues from pension contrib. (Mark)	529	13.36139	11.79899	0.5695652	89.53026
Recruits without basic educ. (%)	851	0.618891	1.605311	0	10.96346
Gender imbalances ratio	851	100.2023	1.710657	88.23055	103.1515
Old age dep. ratio	851	13.4691	2.130164	7.391825	18.29854
Share workers (%)	91	56.77949	24.80694	22.56086	85.39279
Diff. Cath. & Prot. pop.	851	-1541.431	2618.02	-11571.2	72.50089
Share contrib. cat. I	506	20.79672	23.65952	0.4977029	131.7939
Persons per building	851	7.976165	1.599643	5.8	11.2
Cities > 20.000	851	7.310223	7.777653	0	47
Cities $5.000 - 20.000$	851	31.75911	29.90683	1	141
Share farming (%)	91	35.15728	15.98899	1.865982	65.37634
Share mining (%)	91	18.61956	8.769134	6.180529	45.48942
Share trade (%)	91	5.479508	3.267425	1.993503	23.89369
Agric. productivity	781	3.760986	3.237734	1.399423	57.24372
Share insured (%)	851	22.01405	3.486248	14.25401	37.92206
Internal rate of return	484	1.5499	1.019078	0.2111418	6.217336
Net RIA assets	529	35301.28	37861.3	350.007	231055
No. health insurance agencies	483	0.4229023	0.2646751	0.0831041	1.21369

Table 2: Availability of variables

Variable name	Fertility determinant	Proxy (Unit)	Description	Availability (years)	Years used in model
maritalbirthspt	-	–	Marital births	1871, 1878-1914	
ehespt	Marriage	–	Marriages per 1000	1871; 1897-1914	
relbeitr_mark	Income	Contributions to pension system (instead of tax revenues)	Revenues from contributions per insured (Pfennige)	1891-1893, 1895-1913	1895, 1907
edu1neu	Education	Share of recruits without basic education	Recruits without basic education over total number of recruits * 100 (Unit: %)	1871; 1879-1911	1895; 1907
sexratio_cont	Migration	Gender imbalances ratio	Married women over married men * 100 (Unit: married women for 100 married men)	1871; 1880; 1885; 1890; 1900; 1905; 1910	1885; 1890

oldagedepratio_cont	Demographic structure of the population, need for intra-family transfers (budget constraint)		Share of population older than 60/share of population aged 15-20[32]	1871; 1885; 1890	1885; 1890
relworkers	Income, child labour	Working class	Share of workers in a provice	1871; 1882; 1895; 1907	1895; 1907
relbeitr1_3	Female labour force participation rate	Relative importance of revenues in the 'women's category'	Share of contributions in category I over contributions in all other categories	1891-1893; 1895-1913	1895; 1907
diff_cath_prot_lag	Religion	Share of Catholics	Difference between Catholic and Protestant population over Catholic population * 100 (Unit: %)	1871; 1880; 1885; 1890	1885; 1890
persgeb_cont	Urbanisation	Persons per building	Persons per building	1871; 1880; 1885	1880; 1885
grorte_cont	Urbanisation	Cities $\geq$ 20.000 inhabitants	Cities $\geq$ 20.000 inhabitants	1871; 1880; 1900; 1905; 1910	1880; 1905
klorte_cont	Urbanisation	Cities 5.000-20.000 inhabitants	Cities 5.000-20.000 inhabitants	1871; 1880; 1900; 1905; 1910	1880; 1905
relfarming	Industrialisation	Share of people working in farming	Number / 1000 (excludes depedants)	1871; 1882; 1895; 1907	1895; 1907
relmining	Industrialisation / Pension insurance	Share of people working in mining	Number / 1000 (excludes depedants)	1871; 1882; 1895; 1907	1895; 1907
reltrade	Changed availability of contraception, diffusion of knowledge	Share of people working in trade	Number / 1000 (excludes depedants)	1871; 1882; 1895; 1907	1895; 1907
prodindex	Growth, productivity increases, returns to capital	Index combined of the yield (tons) per hectar for different crops (wheat, barley, rye, oats, potatoes, hey)	Tons / hectar	1880-1897; 1899-1914	1895;1907
verspp	Pension insurance	Share of population enrolled in pension insurance	Share enrolled (%)	1895; 1907	1895; 1907
int_rate_return_2	Internal rate of return	Internal rate of return	Net RIA assets (1000 Mark)/(approved pension applications + existing pension claims); unit: Assets per pensioner (1000 Mark)	1892-1913	1895; 1907
reinvermoegen_1000m	Efficiency of RIA	Ability to accumulate assets	Net RIA assets (1000 Mark)	1891-1913	1895; 1907

[32] Value for 1890 is estimated since only information on the age bracket age 10-20 was collected, also refer to Scheubel (2013).

anz_kk_pt2	Social security coverage	Share insured in health insurance	Number in health insurance / 1000	1888-1893; 1895-1899; 1913	1895; 1913

Data sources: Annual Yearbooks of Statistics, Annual Reports of RIAs.

Table 3: DETERMINANTS OF THE FIRST DEMOGRAPHIC TRANSITION

	(1)	(2)	(3)	(4)	(5)
Marriages (per thousand)	1.374 (.344)***	1.136 (.672)*	.435 (.496)	.215 (1.550)	.360 (1.348)
Contributions (*Mark* per insured)	.404 (.104)***	.269 (.051)***	.121 (.060)**	.176 (.094)*	.149 (.102)
Share of recruits without basic schooling		.676 (1.025)	2.565 (2.468)	1.405 (2.571)	.895 (2.445)
Married women (per 100 married men)		.009 (.059)	-.021 (.147)	-.066 (.121)	-.032 (.161)
Old age dependency ratio		-.994 (.280)***	-.748 (.900)	-.456 (.774)	-.321 (.643)
Share workers (%)		.080 (.050)	.015 (.240)	.119 (.174)	.178 (.211)
(Catholics-Protestants)/Catholics * 100		-.0002 (.0003)	-.0002 (.0003)	-.00003 (.0004)	.00004 (.0004)
Contributions in cat. I (rel. to cat. II-V)			.007 (.064)	.040 (.037)	.059 (.047)
Persons (per building)			-.736 (.468)	-1.041 (.366)***	-1.109 (.408)***
Cities $\geq$ 20.000			.009 (.073)	-.045 (.075)	-.083 (.093)
Cities 5.000–20.000			.049 (.105)	.126 (.075)*	.165 (.103)
Agriculture (per thousand)			-.106 (.263)	.277 (.324)	.367 (.312)
Mining and quarrying (per thousand)			.016 (.050)	-.244 (.121)**	-.287 (.107)***
Trade (per thousand)			.490 (1.632)	.672 (1.521)	1.167 (1.796)
Crop yield index (tons per hectar)			.132 (.252)	.070 (.488)	.251 (.542)
Insured (per thousand)			-.285 (.582)	-.089 (.380)	.030 (.497)
Net RIA assets per pensioner (1000 *Mark*)				-1.733 (.734)**	-1.910 (.792)**
Net RIA assets (1000 *Mark*)				-.00004 (.00003)	-.00004 (.00003)
Health insurance agencies (per thousand)					-7.634 (13.553)
Obs.	47	47	47	47	47

Estimation with fixed effects OLS robust to serial correlation. Significance level: *** : $p < 0.01$; ** : $p < 0.05$; * : $p < 0.1$.

Note: Income level proxied by contributions to pension system; diffusion of primary education proxied by share of recruits without basic education; migration proxied by the gender imbalances ratio; the need for intra-family transfers proxied by the old-age dependency ratio; reliance on child labour proxied by size of working class; share of working women proxied by share of revenues in category I (i.e. the women category); diffusion of religion proxied by difference between Catholic and Protestant population; urbanisation proxied by the number of persons per building, the number of large cities and the number of smaller cities; industrialisation proxied by the share working in the primary or the secondary sector; changed availability of contraception and diffusion of knowledge proxied by share working in the tertiary sector; growth and returns to capital proxied by crop yield index; social insurance proxied by number enrolled in pension insurance and by the number of health insurance agencies; internal rate of return proxied by net Regional Insurance Agency (RIA) assets per pensioner.

Table 4: Sensitivity: estimation approach

	(1)	(2)	(3)	(4)
Marriages (per thousand)	1.435	.215	.215	.605
	(1.116)	(1.550)	(2.788)	(1.668)
Contributions (*Mark* per insured)	-.004	-.004	-.004	.227
	(.033)	(.094)*	(.170)	(.097)**
Share of recruits without basic schooling	.789	1.405	1.405	–
	(.287)***	(.287)***	(4.624)	–
Married women (per 100 married men)	-.126	-.066	-.066	-.051
	(.128)	(.121)	(.218)	(.194)
Old age dependency ratio	-1.108	-.456	-.456	-.597
	(.412)***	(.774)	(1.391)	(1.230)
Share workers (%)	.075	.119	.119	.186
	(.092)	(.174)	(.313)	(.201)
(Catholics-Protestants)/Catholics * 100	.00007	-.00003	-.00003	.0001
	(.0001)	(.0004)	(.0007)	(.0003)
Contributions in cat. I (rel. to cat. II-V)	.054	.040	.040	.049
	(.030)*	(.037)	(.067)	(.048)
Persons (per building)	-.180	-1.041	-1.041	-1.055
	(.043)***	(.366)***	(.659)	(.515)**
Cities $\geq$ 20.000	.004	-.045	-.045	-.063
	(.065)	(.075)	(.134)	(.085)
Cities 5.000–20.000	.080	.126	.126	.149
	(.056)	(.075)*	(.135)	(.088)*
Agriculture (per thousand)	.130	.277	.277	.433
	(.150)	(.324)	(.583)	(.280)
Mining and quarrying (per thousand)	-.108	-.244	-.244	-.304
	(.099)	(.121)**	(.218)	(.172)*
Trade (per thousand)	.149	.672	.672	.304
	(.473)	(1.521)	(2.735)	(1.872)
Crop yield index (tons per hectar)	-.240	.070	.070	–
	(1.291)	(.488)	(.877)	–
Insured (per thousand)	-.441	-.089	-.089	.054
	(.174)**	(.380)	(.684)	(.509)
Net RIA assets per pensioner (1000 *Mark*)	-.350	-1.733	-1.733	-1.846
	(.644)	(.734)**	(1.320)	(1.127)
Net RIA assets (1000 *Mark*)	-.00005	-.00004	-.00004	-.00004
	(.00003)	(.00003)	(.00005)	(.00004)
Year: 1907	-4.351	-3.177	–	–
	(2.287)*	(7.273)	–	–
Obs.	47	47	23	

Estimation with OLS in column (1), standard errors adjusted for clustering and serial correlation. Baseline fixed effects model in column (2), standard errors adjusted for serial correlation. Model in first differences in columns (3) and (4), standard errors adjusted for serial correlation. Significance level: *** : $p < 0.01$; ** : $p < 0.05$; * : $p < 0.1$. Note: Income level proxied by contributions to pension system; diffusion of primary education proxied by share of recruits without basic education; migration proxied by the gender imbalances ratio; the need for intra-family transfers proxied by the old-age dependency ratio; reliance on child labour proxied by size of working class; share of working women proxied by share of revenues in category I (i.e. the women category); diffusion of religion proxied by difference between Catholic and Protestant population; urbanisation proxied by the number of persons per building, the number of large cities and the number of smaller cities; industrialisation proxied by the share working in the primary or the secondary sector; changed availability of contraception and diffusion of knowledge proxied by share working in the tertiary sector; growth and returns to capital proxied by crop yield index; social insurance proxied by number enrolled in pension insurance and by the number of health insurance agencies; internal rate of return proxied by net Regional Insurance Agency (RIA) assets per pensioner.

Table 5: Sensitivity: policy effects

	(1)	(2)	(3)	(4)	(5)	(6)
Marriages (per thousand)	.235	3.390	-.022	-.006	1.790	2.564
	(.836)	(1.223)***	(.154)	(.150)	(.645)***	(.950)***
Contributions (*Mark* per insured)	.374	.180	-.071	.146	.035	.015
	(.188)**	(.082)**	(.160)	(.045)***	(.035)	(.022)
Share of recruits without basic schooling	-.123	1.301	-.771	4.148	2.898	-.883
	(.239)	(1.120)	(4.195)	(6.979)	(2.489)	(2.484)
Married women (per 100 married men)	.153	.046	.135	-.063	-.194	-.035
	(.236)	(.135)	(.424)	(.102)	(.093)**	(.112)
Old age dependency ratio	.431	-.138	-.656	-.193	-.391	-.248
	(.194)**	(.123)	(.314)**	(.158)	(.155)**	(.518)
(Cath.-Prot.)/Cath. * 100	-.00009	.0002	-.001	.0002	.0001	.0001
	(.00004)**	(.0001)**	(.001)	(.0002)	(.00009)	(.0006)
Contrib. in cat. I (rel. to cat. II-V)	.084	-.165	.013	-.002	-.090	-.002
	(.062)	(.052)***	(.043)	(.024)	(.057)	(.043)
Persons (per building)	1.840	.340	-1.335	-.965	-.371	-.533
	(.903)**	(.349)	(.314)***	(.143)***	(.125)***	(.766)
Cities $\geq$ 20.000	-.067	-.125	.020	-.049	-.095	-.064
	(.053)	(.041)***	(.076)	(.012)***	(.040)**	(.067)
Cities 5.000–20.000	.011	.037	.154	.024	-.076	.024
	(.015)	(.019)*	(.137)	(.024)	(.047)	(.046)
Agriculture (per thousand)	-.189	.065	-.057	.019	.056	-.112
	(.157)	(.072)	(.193)	(.054)	(.040)	(.178)
Mining and quarrying (per thousand)	-.900	-.241	.322	-.076	.074	.050
	(.398)**	(.448)	(.377)	(.216)	(.106)	(.190)
Trade (per thousand)	-.217	-1.287	.338	.600	.435	-.035
	(.633)	(.390)***	(.453)	(.230)***	(.128)***	(1.110)
Crop yield index (tons per hectar)	-.717	.193	-.136	-.166	.402	-.736
	(.683)	(.327)	(.696)	(.463)	(.317)	(1.191)
Insured (per thousand)	.009	-.014	.023	.010	.005	.007
	(.008)	(.007)*	(.021)	(.005)**	(.004)	(.017)

	(1)	(2)	(3)	(4)	(5)	(6)
Net RIA assets per pensioner (1000 *Mark*)	-2.345	-3.988	-.829	-.805	-.526	-.250
	(1.239)*	(2.312)*	(.759)	(.452)*	(.190)***	(1.906)
Net RIA assets (1000 *Mark*)	-.0001	4.96e-06	-.00002	.00004	4.80e-06	1.00e-05
	(.00004)***	(.00002)	(.00003)	(1.00e-05)***	(1.00e-05)	(.00003)
Obs.	48	47	43	46	46	49

Estimation with FE OLS, standard errors adjusted for clustering and serial correlation. Years in model: 1892 and 1895 in column (1), 1895 and 1899 in column (2), 1900 and 1907 in column (3), 1903 and 1907 in column (4), 1904 and 1907 in column (5) and 1907 and 1912 in column (6). Significance level: $^{***}: p < 0.01$; $^{**}: p < 0.05$; $^{*}: p < 0.1$.

Note: Income level proxied by contributions to pension system; diffusion of primary education proxied by share of recruits without basic education; migration proxied by the gender imbalances ratio; the need for intra-family transfers proxied by the old-age dependency ratio; reliance on child labour proxied by size of working class; share of working women proxied by share of revenues in category I (i.e. the women category); diffusion of religion proxied by difference between Catholic and Protestant population; urbanisation proxied by the number of persons per building, the number of large cities and the number of smaller cities; industrialisation proxied by the share working in the primary or the secondary sector; changed availability of contraception and diffusion of knowledge proxied by share working in the tertiary sector; growth and returns to capital proxied by crop yield index; social insurance proxied by number enrolled in pension insurance and by the number of health insurance agencies; internal rate of return proxied by net Regional Insurance Agency (RIA) assets per pensioner.

Control variables in column (1) not from years 1892 and 1895: gender imbalances ratio (1885 and 1890), old age dependency ratio (1885 and 1890), agriculture (1882 and 1895), mining and quarrying (1882 and 1895), trade (1882 and 1895), large and small cities (1880 and 1900), relative share of Catholics (1871 and 1885), persons per building (1871 and 1880), share workers (1882 and 1895).

Control variables in column (2) not from years 1895 and 1899: gender imbalances ratio (1885 and 1890), old age dependency ratio (1885 and 1890), agriculture (1882 and 1895), mining and quarrying (1882 and 1895), trade (1882 and 1895), large and small cities (1880 and 1900), relative share of Catholics (1871 and 1885), persons per building (1871 and 1880), share workers (1882 and 1895).

Control variables in column (3) not from years 1900 and 1907: gender imbalances ratio (1885 and 1890), old age dependency ratio (1885 and 1890), agriculture (1882 and 1895), mining and quarrying (1882 and 1895), trade (1882 and 1895), large and small cities (1880 and 1900), relative share of Catholics (1871 and 1885), persons per building (1880 and 1885), share workers (1895 and 1907).

Control variables in column (4) not from years 1903 and 1907: gender imbalances ratio (1885 and 1890), old age dependency ratio (1885 and 1890), agriculture (1882 and 1895), mining and quarrying (1882 and 1895), trade (1882 and 1895), large and small cities (1880 and 1900), relative share of Catholics (1871 and 1885), persons per building (1880 and 1885), share workers (1895 and 1907).

Control variables in column (5) not from years 1904 and 1907: gender imbalances ratio (1885 and 1890), old age dependency ratio (1885 and 1890), agriculture (1882 and 1895), mining and quarrying (1882 and 1895), trade (1882 and 1895), large and small cities (1880 and 1900), relative share of Catholics (1871 and 1885), persons per building (1880 and 1885), share workers (1895 and 1907).

Control variables in column (6) not from years 1907 and 1912: gender imbalances ratio (1885 and 1890), old age dependency ratio (1885 and 1890), agriculture (1895 and 1907), mining and quarrying (1895 and 1907), trade (1895 and 1907), large and small cities (1880 and 1900), relative share of Catholics (1890 and 1900), persons per building (1880 and 1885), share workers (1895 and 1907).

Table 6: Sensitivity: measuring fertility

	(1)	(2)	(3)	(4)	(5)
Marriages (per thousand)	.027	2.176	.081	2.236	-.639
	(.014)*	(1.112)*	(.027)***	(1.029)**	(.327)*
Married women (per 100 married men)	.001	.223	-.005	.233	-.091
	(.003)	(.259)	(.007)	(.291)	(.063)
Old age dependency ratio	.013	.600	.019	.033	-.349
	(.006)*	(.348)*	(.011)*	(.311)	(.219)
Share workers (%)	.0004	.018	-.0006	-.055	-.0005
	(.0003)	(.025)	(.001)	(.028)**	(.012)
Persons (per building)	.005	.046	.035	.728	-.068
	(.006)	(.503)	(.013)***	(.490)	(.163)
Cities $\geq$ 20.000	-.001	-.248	.003	.210	-.0006
	(.002)	(.179)	(.006)	(.199)	(.080)
Cities 5.000–20.000	-.003	-.082	-.009	-.090	-.045
	(.002)	(.166)	(.004)*	(.165)	(.044)
Agriculture (per thousand)	-.001	-.153	-.0009	-.096	.027
	(.002)	(.145)	(.004)	(.163)	(.041)
Mining and quarrying (per thousand)	-.006	-.727	-.022	-.661	.262
	(.004)	(.304)**	(.008)***	(.333)**	(.108)**
Trade (per thousand)	-.010	-.811	-.030	-.669	.684
	(.010)	(.831)	(.018)	(.721)	(.214)***
Crop yield index (tons per hectar)	.0003	-.013	-.0002	-.025	-.005
	(.0002)	(.012)	(.0004)	(.013)*	(.007)
Obs.	46	46	46	46	46

Estimation with FE OLS for years 1885 and 1890 (i.e. years for which age structure is available), standard errors adjusted for clustering and serial correlation. Variable on persons per building omitted since a sufficient lag is not available. Dependent variable: total fertility index in column (1), crude birth rate in column (2), martial fertility index in column (3), crude marital birth rate in column (4), share of illegitimate births in column (5). Significance level: *** : $p < 0.01$; ** : $p < 0.05$; * : $p < 0.1$. Note: Not all explanatory variables from the baseline model included because of limited data availability for earlier years. In particular, variables on the pension system are only available after its inception in 1891. Migration proxied by the gender imbalances ratio; the need for intra-family transfers proxied by the old-age dependency ratio; reliance on child labour proxied by size of working class; urbanisation proxied by the number of persons per building, the number of large cities and the number of smaller cities; industrialisation proxied by the share working in the primary or the secondary sector; changed availability of contraception and diffusion of knowledge proxied by share working in the tertiary sector; growth and returns to capital proxied by crop yield index.

Appendix B: Details on the theoretical model

B.1 Second order conditions

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

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

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

in the Bismarckian case,

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

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

in the Beveridgean case and

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

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

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

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

holds.

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:

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

in the Bismarckian case,

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

in the Beveridgean case and

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

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.

B.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

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

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

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

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

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

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 $t + 1$:

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

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 impact of an extended pension system on savings is given by:

$$\frac{\partial s_t}{\partial \tau} = -\frac{V_{nn}V_{s\tau} - V_{n\tau}V_{sn}}{V_{nn}V_{ss} - V_{ns}V_{sn}}. \tag{B.15}$$

The effect of a higher contribution rate on savings depends on the sign of the numerator on the RHS of (B.15). By using the second derivatives (B.5) and (B.6) and the second derivatives with respect to the contribution rate:

$$\begin{aligned} V_{n\tau} &= w_t(1-f(n_t)) \\ & \left[U_{cc}((1-\tau)w_t f'(n_t)+\pi_t)+U_{zz}\left[B_{t+1}-R_{t+1}\tau w_t f'(n_t)\right]R_{t+1}\right] \end{aligned} \quad \text{(B.16)}$$

$$V_{s\tau} = w_t(1-f(n_t))\left[U_{cc}+U_{zz}R_{t+1}^2\right] < 0 \quad \text{(B.17)}$$

this numerator is given by:

$$\begin{aligned} V_{nn}V_{s\tau}-V_{n\tau}V_{sn} &= w_t(1-f(n_t)) \\ & \left[\left(U_{nn}-U_c w_t f''(n_t)\right)\left(U_{cc}+U_{zz}R_{t+1}^2\right)\right. \\ & \left.+U_{cc}U_{zz}\left(R_{t+1}(w_t f'(n_t)+\pi_t)-B_{t+1}\right)^2\right] \\ &> 0. \end{aligned} \quad \text{(B.18)}$$

By employing (B.10) 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 (B.14) nor the lifetime income from (B.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.

B.3 Savings decision in a pay as you go pension system

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

$$\frac{\partial s_t}{\partial \tau} = -\frac{V_{nn}V_{s\tau} - V_{n\tau}V_{sn}}{V_{nn}V_{ss} - V_{ns}V_{sn}}. \tag{B.19}$$

The denominator is positive for all three pension types.

In the case of the Bismarckian pension system we have

$$\begin{aligned} V_{nn}V_{s\tau} - V_{n\tau}V_{sn} \;=\; & w_t(1-f(n_t))(U_{cc} + U_{zz}\Omega_{t+1}R_{t+1}) \\ & \big(U_{nn} - U_c(1-\tau)w_t f''(n_t) - U_z\Omega_{t+1}\tau w_{t+1}f''(n_t)\big) \\ & -U_z w_t f'(n_t)(R_{t+1} - \Omega_{t+1})[U_{cc}((1-\tau)w_t f'(n_t) + \pi_t) \\ & +U_{zz}R_{t+1}\left(B_{t+1} - \Omega_{t+1}\tau w_t f'(n_t)\right)] \\ & +U_{cc}U_{zz}w_t(1-f(n_t)) \\ & \big[R_{t+1}((1-\tau)w_t f'(n_t) + \pi_t) - \big(B_{t+1} - \Omega_{t+1}\tau w_t f'(n_t)\big) \\ & \big(\Omega_{t+1}(w_t f'(n_t) + \pi_t) - B_{t+1}\big)\big]. \end{aligned} \tag{B.20}$$

This numerator is positive if the following condition for the intra-family transfer B_{t+1} holds: $-\frac{\partial p^{BIS}_{t+1}}{\partial n_t} < B_{t+1} < \Omega_{t+1}((1-\tau)w_t f'(n_t) + \pi_t) - \frac{\partial p^{BIS}_{t+1}}{\partial n_t}$. 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.

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

$$\begin{aligned} V_{n\tau} &= w_t f'(n_t) U_z (R_{t+1} - \Omega_{t+1}) + w_t (1 - f(n_t)) \\ & \quad \left[U_{cc}((1-\tau) w_t f'(n_t) + \pi_t) + U_{zz} \left(B_{t+1} - \Omega_{t+1} \tau w_t f'(n_t) \right) \Omega_{t+1} \right]. \end{aligned} \quad \text{(B.21)}$$

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: 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.