

Fenge, Robert; Scheubel, Beatrice

Conference Paper

Pensions and Fertility: Back to the Roots

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2016: Demographischer Wandel -
Session: Pensions and Labor Market Outcomes, No. G05-V3

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Fenge, Robert; Scheubel, Beatrice (2016) : Pensions and Fertility: Back to the Roots, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2016: Demographischer Wandel - Session: Pensions and Labor Market Outcomes, No. G05-V3, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at:

<https://hdl.handle.net/10419/145689>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

PENSIONS AND FERTILITY: BACK TO THE ROOTS

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

Robert Fenge[†] Beatrice Scheubel[‡]

This version: 11 February 2016

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 on the mechanism by which a public pension system depresses fertility, based on historical data. Our theoretical framework highlights that the effect of a public pension system on fertility is ex ante ambiguous while its size is determined by the internal rate of return of the pension system. We identify an overall negative effect of the introduction of pension insurance on fertility using regional variation across 23 provinces of Imperial Germany in key variables of Bismarck's pension system, which was introduced in Imperial Germany in 1891. The negative effect on fertility is robust to controlling for the traditional determinants of the first demographic transition as well as to other policy changes.

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

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

*We would like to thank Kathrin Weny for valuable research assistance. We are also grateful to Tobias Jopp for helpful comments and suggestions.

[†]Department of Economics, University of Rostock, Ulmenstr. 69, 18057 Rostock, Germany; email: robert.fenge@uni-rostock.de.

[‡]Corresponding author. European Central Bank. Address: ECB, 60604 Frankfurt am Main, Germany. Email: beatrice.scheubel@ecb.europa.eu. This paper should not be reported as representing the views of the European Central Bank (ECB). The views expressed are those of the author and do not necessarily reflect those of the ECB.

1 Introduction

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 (Becker 1965; Feldstein 1974) or labour migration (e.g. Borjas 1999), also in the historical context (e.g. Khoudor-Cásteras 2008). This link between social security and individual behaviour has been postulated as the so-called social security hypothesis (Feldstein 1974), 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 both in the seminal economic literature (e.g. Becker 1960, 1988, 1991) and in broader discussions on the fertility decline in advanced economies.¹

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 (Bental 1989; Prinz 1990; Cigno 1993; Cigno and Rosati 1996; Sinn 2004; Fenge and Meier 2005; Cigno and Werding 2007; Cremer et al. 2008). Recently, the link between pensions and fertility has received increasing attention, also in the historical context of the first demographic transition (Guinnane 2011). 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, and cross-country variation is rare (Ehrlich and Zhong 1998 and Boldrin et al. 2015 are exceptions). Moreover, there are only few exogenous changes within a given pension system (Cigno and Rosati 1992; Cigno et al. 2003; Billari and Galasso 2009). 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 better understanding of the causes of population ageing. First, we further develop microeconomic overlapping generation models in the spirit of Cigno (1993), to illustrate that the effect of pension insurance on fertility is ex ante ambiguous and determined by the extent to which pension insurance affects lifetime income. We derive testable hypotheses from the model on this link between pensions and fertility. Second, we use the historical data set on the introduction of pension insurance in Imperial Germany developed by Scheubel (2013) to test the hypotheses derived from

¹ The literature explaining the decline in fertility has put relatively more emphasis on labour market institutions affecting female labour supply (e.g. Ahn and Mira 2002), the interaction between the tax system and family policy (e.g. Apps and Rees 2004; Egger and Radulescu 2012), and maternity leave legislation (e.g. Berger and Waldfogel 2004).

the model. We quantify the effect of pension insurance on fertility using cross-regional variation in 23 provinces of Imperial Germany, for which we construct a pseudo panel of two periods. As this implies that the total number of observations in our study is only 46, we show that our results hold for different proxies and that they do not depend on the specific construction of the pseudo panel.

Imperial Germany was the first European country that enacted an irreversible transition into a welfare state. Company-level pension insurance for workers in certain professions was already common during the mid-nineteenth century (Jopp 2013). By the end of the nineteenth century, the concept of pension insurance had become statutory for almost all workers at the national level (Scheubel 2013). The authorities collected information on several key variables of the system since its inception. We use this data for our analysis, which 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 and a public pension system. One of the key 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, whenever the return from contributions to the pension system is lower than the return from the other possibilities of saving for old age, in other words, when there is an implicit tax in the pension system, a higher contribution rate also implies lower disposable income, which has a negative effect on fertility. In equilibrium, these effects are traded off against each other. The internal rate of return of the pension system determines the size of the overall effect. In our empirical analysis we show that the overall effect of a higher enrolment rate 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 0.5 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 (e.g. Richards 1977; Galloway et al. 1994; Brown and Guinnane 2007), we can compare the impact of pension insurance to other factors. For example, the impact of a 1% increase in pension insurance coverage is approximately equivalent to 10 times the impact of a 1% increase in education (proxied by the share of recruits with at least basic schooling) and to half the impact of 1 additional person per building (which is our proxy for urbanisation).

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.

Section 2 provides institutional details on social policy in late nineteenth century Germany. 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.

The pension system of 1891 consisted of both funded and pay as you go elements and was turned into a pure pay as you go (PAYG) system by a law adopted in 1899 and coming into effect in 1900. 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. However, there was a general fear that the government would touch the capital stock, not least because it seemed that some regions ran into financing difficulties. In addition, the 1891 set-up was not perceived as socially fair. As a consequence the pension system became a fully-fledged pay as you go system when the law was revised in 1899, coming into effect in 1900.²

The pension system 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).³ As a consequence, about 20-25% of the population were covered by pension insurance.

Pension insurance provided so-called invalidity pensions and old-age pensions. Invalidity 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. A worker had to prove that they met one of these conditions by either reaching the age of 70 or by proving that they were not able to earn at least the average day labourer's wage.⁴ Both invalidity pension and old-age pension were designed as a supplementary income. As 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), the invalidity pension was by far the more important supplementary income. Since both types of pension were paid when workers were unable to earn their income due to disability

² *Reichsgesetzblatt* (RGBl) 1899/33.

³ Also refer to the published law in *Reichsgesetzblatt* (RGBl) 1889/13.

⁴ After 1900 the definition of old age changed slightly and every worker who reached the age of 65 was automatically classified as disabled.

(either age-related or not related to age), we interpret the distinction between invalidity pensions and old age pensions as mainly semantic. This interpretation is in line with the common understanding of ageing in the historical context, when not only biological age was considered as a qualifying condition for a pension, but also physical deficiencies (Bourdieu and Kesztenbaum 2007). In other words, the disability pension was the relevant pension for a worker to be considered as ‘old’ at the time.

Contribution rates only differed between 4 wage categories, which implies that workers paid contributions according to income, but the link between income and contribution was not direct. Between 1891 and 1900 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 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).⁵

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 with regard to approving pension applications.

3 Theoretical analysis of the pension system and fertility

Microeconomic theories of fertility choice were developed by Becker and others (Becker 1960, 1965, 1988, 1991, 1992; Schultz 1969; Barro and Becker 1986, 1988, 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. In equilibrium, the marginal benefit of an additional child has to be equal to the marginal cost of rearing the child.

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; Fenge and Meier 2009; Fenge and von Weizsäcker 2010). The impact of institutions has, however, not been discussed extensively in the context of the demographic transition in nineteenth century Europe (refer to Guinnane 2011 for a review).

Our model which provides a framework for analysing the impact of pension insurance on the first demographic transition is also linked to the literature on the social

⁵ 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.

security hypothesis which postulates that institutions have an impact on behaviour, such as labour supply reactions, but more recently has also been linked to old age provision (Caldwell 1978; Willis 1979; Bental 1989). With our model we investigate how the introduction or expansion of public pension schemes reduces fertility. The model combines three options to provide for old age in a simple two-period overlapping generations setting. The first option is to voluntarily accumulate savings in order to form a capital stock from which private pensions can be drawn during the retirement period. The second option is an intra-family transfer that children give to their parents because they derive utility from the well-being of their parents (altruistic preferences). The third option is to contribute to a PAYG pension system⁶ which provides a public pension in the old age period. The decisions about fertility, savings and the intra-family transfer are endogenous. We discuss several possible channels how the introduction or extension of a pension system may affect the decisions of a generation.

The set-up and results of our model are consistent with earlier findings in the literature. Boldrin et al. (2015) calculate the quantitative effects of the public provision of old-age pensions in calibrated models based on the old-age security motive for children. They find that in such models there is a sizeable negative effect on fertility which is consistent with our empirical results. Moreover, they show that an improved access to capital markets reduces the incentives for childbearing. Our model is consistent with these findings, as illustrated in Appendix B. Ehrlich and Kim (2007) also find that social security contributions and benefits generate incentives to reduce fertility, but in contrast to our model they analyse a PAYG pension system in which the number of children neither affects labour supply nor wage income. They show that these effects cannot be fully compensated by inter-temporal or intergenerational transfers within families. Puhakka and Viren (2012) show that also Beveridgean PAYG pension systems (i.e. with lump-sum contributions and pensions) reduce fertility. In contrast, Hirazawa et al. (2014) show that in a model with specific log-linear utility functions the effect of Bismarckian PAYG pension scheme on fertility vanishes.

3.1 The Model

We consider the impact of a pension system on fertility, savings and intra-family transfers in a two-period overlapping generations model which is 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 $\frac{N_{t+1}}{N_t} = 1 + \bar{n}_{t+1}$. We analyse the decisions of a household on the number of children n_t , savings s_t and the intra-family transfer b_t in period t . Note that the number of children of an atomistic household has no effect on population growth. In other words, the fact that more children imply more contributors to the pension system is not internalised by the household which leads to the well-known positive externality

⁶ We analyse a 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 what is known in the literature as a Bismarckian PAYG pension system in which pensions are proportional to contributions.

in PAYG pension systems, also known as the fiscal externality of children in a PAYG pension system. The number of children in a family and the growth rate of the population only coincide in equilibrium if 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 needed for rearing a child $f(n_t)$ 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, which corresponds to the historical set-up which only had four wage categories. 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:

$$c_t = w_t(1 - f(n_t))(1 - \tau) - s_t - \pi_t n_t - b_t. \quad (1)$$

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 b_{t+1} paid by the children. The budget constraint in the second period is:

$$z_{t+1} = p_{t+1} + R_{t+1}s_t + b_{t+1}n_t. \quad (2)$$

The old-age consumption of the parental generation $t - 1$,

$$z_t = p_t + R_t s_{t-1} + b_t n_{t-1}, \quad (3)$$

enters the utility of the children's generation t . Since s_{t-1} , n_{t-1} and p_t are determined in the past the only determinant of z_t in period t is the intra-family transfer b_t paid by the household of generation t . The utility of the household depends on own consumption in both periods, on the old-age consumption of the parents and the individual number of children. The function $U(c_t, z_{t+1}, z_t, n_t)$ is increasing in all four arguments, strictly concave and additively separable: $U_{cz} = U_{cn} = U_{zz} = U_{zn} = 0$.

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 for having 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

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. Sinn (2004), Cigno (2006), Cigno et al. (2006).

create a model set-up which corresponds to the set-up of pension insurance in Imperial Germany. During the first 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 also present theoretical results on the behavioural effect of the transfer channel.

The household determines the number of children and savings by maximising utility subject to the budget constraints (1), (2) and (3). Substituting these constraints for the consumption variables in the utility function results in a maximisation problem of a function depending on n_t , s_t and b_t :

$$\begin{aligned} \max_{n_t, s_t, b_t} V(n_t, s_t, b_t) &= U(w_t(1 - f(n_t))(1 - \tau) - s_t - \pi n_t - b_t, \\ &\quad p_{t+1} + R_{t+1}s_t + b_{t+1}n_t, p_t + R_t s_{t-1} + b_t n_{t-1}, n_t). \end{aligned} \quad (4)$$

Hence, we can write the first-order conditions of the maximisation problem as:

$$V_n = -U_{c_t}((1 - \tau)w_t f'(n_t) + \pi_t) + U_{z_{t+1}} \left(\frac{\partial p_{t+1}}{\partial n_t} + b_{t+1} \right) + U_{n_t} = 0, \quad (5)$$

$$V_s = -U_{c_t} + U_{z_{t+1}} R_{t+1} = 0, \quad (6)$$

and

$$V_b = -U_{c_t} + U_{z_t} n_{t-1} = 0. \quad (7)$$

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

In the following we analyse the impact of increasing the contribution rate τ in period t on fertility n_t in a PAYG pension system. In order to focus the paper on this key effect we present the other effects on savings s_t and the intra-family transfer b_t in Appendix B. Furthermore, Appendix B presents also the model results for an economy without a functioning capital market.

3.2 The effect of a Bismarckian pay-as-you-go pension system on fertility

In a 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:

$$p_{t+1}^{BIS} = (1 + \overline{n_{t+1}}) \tau w_{t+1} (1 - \overline{f(n_{t+1})}) \frac{\tau w_t (1 - f(n_t))}{\tau w_t (1 - \overline{f(n_t)})}, \quad (8)$$

where $(1 - \overline{f(n_t)})$ 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, $\tau w_t (1 - \overline{f(n_t)})$, the individual pension, p_{t+1}^{BIS} , is higher than the average pension,

$(1 + \overline{n_{t+1}})\tau w_{t+1}(1 - \overline{f(n_{t+1})})$, by the same proportion. Since the wage rate and the contribution rate are identical for all households we may write the proportionality factor as $\frac{1-f(n_t)}{1-\overline{f(n_t)}}$ and call it the Bismarck factor. 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: $1 - \overline{f(n_t)} = 1 - f(n_t)$ in the case of homogeneous households.

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

$$\frac{\partial p_{t+1}^{BIS}}{\partial n_t} = -(1 + \overline{n_{t+1}})\tau w_t f'(n_t) \frac{w_{t+1}}{w_t} \frac{1 - \overline{f(n_{t+1})}}{1 - f(n_t)} < 0. \quad (9)$$

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 raising children also reduces pensions.

Second period consumption is given by:

$$z_{t+1} = (1 + \overline{n_{t+1}})\tau w_{t+1}(1 - \overline{f(n_{t+1})}) \frac{1 - f(n_t)}{1 - \overline{f(n_t)}} + R_{t+1}s_t + b_{t+1}n_t \quad (10)$$

and the intertemporal budget by:

$$\begin{aligned} R_{t+1}c_t + z_{t+1} &= R_{t+1}[(1 - \tau)w_t(1 - f(n_t)) - \pi_t n_t - b_t] \\ &+ (1 + \overline{n_{t+1}})\frac{w_{t+1}(1 - \overline{f(n_{t+1})})}{w_t(1 - f(n_t))}\tau w_t(1 - f(n_t)) + b_{t+1}n_t. \end{aligned} \quad (11)$$

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

$$\begin{aligned} \Pi_{t+1}^{BIS} &= R_{t+1}((1 - \tau)w_t f'(n_t) + \pi_t) \\ &+ (1 + \overline{n_{t+1}})\frac{w_{t+1}(1 - \overline{f(n_{t+1})})}{w_t(1 - f(n_t))}\tau w_t f'(n_t) - b_{t+1}. \end{aligned} \quad (12)$$

If this marginal price is positive there is an inner solution of the fertility decision. We assume it to be positive in the following.

Moreover, we denote the internal rate of return of contributions to the PAYG pensions system in equilibrium by:

$$\Omega_{t+1} \equiv p_{t+1}/\tau w_t(1 - f(n_t)). \quad (13)$$

If contribution rates are constant as we assume this is equal to the payroll growth factor:

$$\Omega_{t+1} = (1 + \overline{n_{t+1}})\frac{w_{t+1}}{w_t}\frac{1 - \overline{f(n_{t+1})}}{1 - f(n_t)}. \quad (14)$$

Now we consider the fertility decision in a PAYG pension system of the Bismarckian type. The fertility effect is given by:

$$\frac{\partial n}{\partial \tau} = - \frac{\begin{vmatrix} V_{n\tau} & V_{ns} & V_{nb} \\ V_{s\tau} & V_{ss} & V_{sb} \\ V_{b\tau} & V_{bs} & V_{bb} \end{vmatrix}}{\begin{vmatrix} V_{nn} & V_{ns} & V_{nb} \\ V_{sn} & V_{ss} & V_{sb} \\ V_{bn} & V_{bs} & V_{bb} \end{vmatrix}}. \quad (15)$$

Due to the second-order conditions for a maximum the denominator is negative as shown in Appendix B. In order to calculate the sign of the numerator of (15) we need the second derivatives of utility with respect to the contribution rate:

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

$$V_{s\tau} = w_t(1 - f(n_t))[U_{cc} + U_{zz}\Omega_{t+1}R_{t+1}] < 0, \quad (17)$$

and

$$V_{b\tau} = U_{cc}w_t(1 - f(n_t)) < 0. \quad (18)$$

The numerator of equation (15) can be calculated as:

$$\begin{aligned} \begin{vmatrix} V_{n\tau} & V_{ns} & V_{nb} \\ V_{s\tau} & V_{ss} & V_{sb} \\ V_{b\tau} & V_{bs} & V_{bb} \end{vmatrix} &= (R_{t+1} - \Omega_{t+1}) [w_t f'(n_t) U_{z_{t+1}} \\ &\quad (n_{t-1}^2 U_{c_t c_t} U_{z_t z_t} + R_{t+1}^2 U_{c_t c_t} U_{z_{t+1} z_{t+1}} + n_{t-1}^2 R_{t+1}^2 U_{z_t z_t} U_{z_{t+1} z_{t+1}}) \\ &\quad + w_t(1 - f(n_t)) U_{c_t c_t} U_{z_t z_t} U_{z_{t+1} z_{t+1}} n_{t-1}^2 \\ &\quad (R_{t+1}((1 - \tau)w_t f'(n_t) + \pi_t) - (b_{t+1} - \Omega_{t+1}\tau w_t f'(n_t)))]. \end{aligned} \quad (19)$$

The sign of the numerator is ambiguous and we have to consider the separate effects in turn. Using (14), the marginal price of children from equation (12) 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 The first summand on the RHS of equation (19) is the effect of the contribution rate via the marginal price of a child. It is positive for the following reason. Raising 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 raises 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 . Therefore, a higher contribution rate lowers the opportunity cost of having a child in the first period, but raises the opportunity cost of having a child in the second period in terms of pension entitlements. If $R_{t+1} > \Omega_{t+1}$

the total opportunity cost falls. Partial derivation of (12) with respect to τ shows that the price of a child decreases with a higher contribution rate:

$$\frac{\partial \Pi_{t+1}^{BIS}}{\partial \tau} = -(R_{t+1} - \Omega_{t+1}) w_t f'(n_t) < 0. \quad (20)$$

Since children become relatively cheaper as a provision for old-age, the number of children increases.

The income effect The second summand on the RHS of equation (19) is the effect of the contribution rate via a change in lifetime income. This income effect reduces fertility. By using the definition of the payroll growth factor (14) the lifetime budget constraint (11) can be written as:

$$R_{t+1}c_t + z_{t+1} = w_t(1 - f(n_t)) [R_{t+1} - \tau(R_{t+1} - \Omega_{t+1})] - (R_{t+1}\pi_t - b_{t+1})n_t. \quad (21)$$

The derivation of the RHS of (21) with respect to τ shows that a higher contribution rate reduces lifetime income by

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

Lifetime income is reduced because PAYG pension system imposes an implicit tax on wage income if $R_{t+1} > \Omega_{t+1} \forall t$ (e.g. Barro and Becker 1988; Fenge and Werding 2004; Sinn 2000, 2004), since in this case compulsory contributions to the pension system mean a loss in lifetime income as investing the same amount of contributions in the capital market instead would yield a higher rate of return. The implicit wage tax rate can be written as $\tau(R_{t+1} - \Omega_{t+1}) > 0$. A higher contribution rate raises this implicit tax and reduces lifetime income. This 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.¹⁰

The total effect on fertility is negative if the income effect is larger than the price effect, and vice versa. The scale of both effects depends on the factor $(R_{t+1} - \Omega_{t+1})$ with the internal rate of return assumed to be lower than the capital market interest rate. Hence, the size of the total fertility effect is larger the smaller the internal rate of return of the pension system $\Omega_{t+1} \equiv p_{t+1}/\tau w_t(1 - f(n_t))$. We can state:

Proposition 1 *Fertility effect* *If the internal rate of return is lower than the capital market interest rate, 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. Furthermore, the fertility effect is stronger the smaller the internal rate of return of the pension system.*

From the theoretical result we can derive the following proceeding for the empirical investigation. The fertility effect as combination of the price effect and the income effect

¹⁰ Note that without the 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 transfers from the model is a stronger income effect.

is ambiguous. If the income effect is larger than the price effect, fertility declines with an introduction or extension of the PAYG pension system. Since the theoretical model does not provide a definite answer to how fertility responds to the PAYG system we analyse the fertility effect empirically in order to get a definite understanding of which partial effect prevails.

We can summarise the findings in our main hypotheses:

HYPOTHESIS 1: TOTAL FERTILITY EFFECT IN A BISMARCKIAN PAY-AS-YOU-GO PENSION SYSTEM

Under the condition $R_{t+1} > \Omega_{t+1}$, if the PAYG pension system is introduced and the income effect is higher than the price effect then fertility declines.

HYPOTHESIS 2: PRICE AND INCOME EFFECT

Assume $R_{t+1} > \Omega_{t+1}$. Then a rising contribution rate of the PAYG pension system has two effects. The opportunity cost of a child decreases which has a positive effect on fertility (price effect). Lifetime income decreases due to a higher implicit tax which has a negative effect on fertility (income effect).

HYPOTHESIS 3: MAGNITUDE OF THE FERTILITY EFFECT

The smaller the internal rate of return Ω_{t+1} , the stronger all three effects, especially the higher is the fertility effect.

4 Data, identification strategy, and econometric considerations

4.1 The data set

To test our hypotheses we use regional historical data from the time of the introduction of the first comprehensive pension system and the first demographic transition. The introduction of the first comprehensive pension system in Germany towards the end of the nineteenth century is well-suited for an analysis of the impact of pension insurance on fertility because there is well-documented regional variation in key variables of the pension system which we can use for identification. In addition, fertility developments have also been well-documented for most German provinces.

The data on fertility, population and a set of control variables is taken from the Imperial Annual Yearbook of Statistics (*Statistisches Jahrbuch für das Deutsche Reich*).¹¹

The data on the pension system, first collected by Kaschke and Sniegs (2001), is taken from the Annual Reports of the RIAs. As they were largely autonomous in administering

¹¹ Population numbers were reported annually until 1895, but afterwards only during census years, i.e. in 1895, 1899, 1900, 1905, 1909, 1910. We use the extrapolated population numbers from Scheubel (2013) for the missing years.

the pension system, it should not be surprising that each RIA also collected detailed statistics on how it managed the pension system.

As the states and provinces recorded in the Annual Yearbook of Statistics did not fully overlap with the set-up of the RIAs, we use the matched data set developed by Scheubel (2013). While the Annual Yearbook of Statistics provides information at the state and province level, of which there were in total 44, some RIAs covered more than one state or province and one state could also be covered by more than one RIA (for example large states such as the Kingdom of Bavaria). In the matched data set, regions, provinces and RIAs are matched based on their geographical location which implies that observations for those states or provinces which are covered by one RIA are averaged and observations for those RIAs which cover a part of the same state or province are also averaged. As a consequence, the combined data set consists of 24 cross-sectional observations of which we however drop one outlier as detailed below. Figure 1 shows the regional entities in the harmonised data set.

[Figure 1 about here.]

4.2 Identification strategy

Mechanism

To identify an effect of social insurance on fertility, we follow our theoretical model in looking at the coverage of pension insurance. Our model suggests that the overall effect $\frac{\partial n}{\partial \tau}$ is composed of a price effect related to the reduction in disposable income caused by contributing to the pension system and by an income effect related to lower lifetime income caused by the implicit tax in the pension system. The sign of the total effect is determined by the larger of the two effects and its size is determined by the internal rate of return of the pension system.

As the sustained decline in birth rates pictured in the left panel of figure 2 only appeared across all provinces long after the onset of industrialisation, we hypothesise that the overall effect of pension insurance on fertility is negative. In the framework of our model this would mean that the income effect dominates the price effect. The fertility decline seems to have started fully only in the 1890s when pension insurance was introduced, and it increased its pace after 1899 when more pay as you go elements were introduced. The challenge for our identification strategy is to choose appropriate proxies for the overall effect $\frac{\partial n}{\partial \tau}$ at the regional level which help us to identify the overall effect of pension insurance by using regional variation.

[Figure 2 about here.]

Choice of proxies

Our choice of proxies is determined by the functioning of Bismarck's pension system and by data availability. RIAs collected a battery of variables on the functioning of the pension system from which we calculate five proxies which we believe help us best to

gauge the effect of pension insurance and the behavioural mechanisms underlying this effect.

As we are first and foremost interested in the total effect of pension insurance, $\frac{\partial n}{\partial \tau}$, we use the coverage of compulsory insurance, the *share of the population insured*, as a proxy. As unfortunately, the contribution rate τ did not vary much across RIAs¹² and there was no change foreseen in contribution rates at least for the first ten years, we use one of the main particularities of Bismarck's pension system to proxy the effect of τ : insurance was compulsory, but only for a certain group of people. Only those people were required to contribute to pension insurance who based on their job description were considered as workers. As the share of people classified as workers differed across regions, also the share of people who had to participate in pension insurance differed across regions. Consequently, only a certain share of the population experienced the decline in lifetime income which was caused by having to pay contributions to pension insurance, and which in the terms of our model corresponds to raising contribution rates from 0 to τ for the part of the population covered by pension insurance.

While compulsory insurance should proxy well the overall effect of pension insurance on lifetime income, we also use this variable weighted with the likelihood of receiving a pension. The decentralised set up of the pension system introduced some uncertainty with regard to receiving a pension, which should have intensified the effect of compulsory insurance on lifetime income. As the RIAs enjoyed some discretion in the maintenance of the system's administration, particularly with regard to the eligibility criteria for receiving a pension (Kaschke and Sniegs 2001),¹³ prospective pensioners had to apply for receiving a pension. They had to prove that they had paid contributions at least for the minimum period as well as prove that they were unable to earn a subsistence level income. Particularly the latter criterion involved considerable judgement by the RIA official dealing with the application. The right panel of figure 2 shows that there has been considerable variation in this *approval rate* across provinces. As the approval rate can hence be considered as the probability of receiving a pension, we weigh the share insured with the approval rate to create a refined proxy for the overall effect of pension insurance.

To test whether the mechanism suggested by our model framework is present in the data, we also construct proxies for the income effect and the price effect. To proxy the income effect we construct a measure of the implicit tax in the pension system. The implicit tax is normally given by the difference between contributions and discounted future pension. We calculate the *implicit tax* as the difference between the average pension contribution per insured (i.e. $\tau(1 - f(n))w$) and the average pension (i.e. p_t). To proxy the price effect we choose a variable which captures particularly the reduction in first-period disposable income. As the *average contribution per insured* is inversely related to disposable income, we consider it a good proxy.

¹² In fact, contribution rates only varied between the four contribution categories, but not between provinces and not over time.

¹³ Refer to the 1889 law on pension insurance (*Reichsgesetzblatt 1889/13*) and the 1899 revision (*Reichsgesetzblatt 1899/33*).

Finally, to further examine whether the mechanism suggested by our model framework is visible in the data, we construct a proxy for the *internal rate of return* that helps us to investigate whether the magnitude of the overall effect is affected by the internal rate of return. We proxy the internal rate of return by dividing the average future pension level by the average contribution, but we also use the ratio between current pension level and current average contribution as a proxy.

Identifying assumptions

As RIAs collected the data independently, we can base our identification of an effect on cross-regional variation on three main identifying assumptions. First, the effect of pension insurance needs to be observable to the population in order to instigate an immediate change in the fertility rate. Second, pension insurance (and the variables which we choose as proxies) needs to be independent of the fertility decline which took place during the second half of the nineteenth century to distinguish the effects of pension insurance. Third, the variables we would like to measure and the respective proxies we construct need to reflect regional differences in the pension system rather than other regional differences. We look at these assumptions in turn.

First, there are several reasons that lead us to assume that people in the street were able to observe that the pension system was working. If young couples perceived that the pension system was working reliably and that pensions were sufficient to make up for savings or transfers from children, they could be induced to have fewer children. Since the pension system operated locally, it is likely that people in the street could form an opinion on the coverage of pension insurance because either they were insured themselves or because they observed their fellow citizens' participation in the pension system. Participation was observable, because weekly contributions were paid at the post office and so-called *Klebeurkunden*, a form of stamps, had to be collected in a book similar to a collector's album. Consequently, even those not insured could observe participation rates as well as the level of average contributions. Moreover, the average pension level should have been clear from the beginning. The system entailed some transitional arrangements, which meant that pensions were paid as of 1892 to some groups of the population which were too old or too unable to earn enough to accumulate the number of stamps needed for applying for a pension. In fact, the number of pensioners was particularly large in 1892. It is fair to assume that even those unrelated to a pensioner would be able to hear about the pension level from the pensioner's co-workers. As applications for and payment of pensions was administered locally, it is also probable that people were informed about the approval rate of pension applications and based on this formed an opinion about the probability of receiving a pension.

Second, it is essential that the variation in the share insured, the average contribution, the average pension and the approval rate is not caused by the fertility decline itself. In particular, should changes in the birth rate drive differences in one of these key pension system variables, it would not be possible to causally relate the cross-regional differences in those variables to differences in fertility. One particular concern in this regard may be the fact that the share insured is related to having an occupation which required compulsory insurance. As insurance was intended for the 'working class', it would be

difficult to relate any changes in fertility to pension insurance if the ‘working class’ had a fertility rate which was significantly different from other parts of the population. For example, it is likely that rural workers displayed a higher birth rate as children contributed to household income. However, the ‘working class’ as defined in the law on pension insurance included workers from all sectors of the economy, including sectors for which we would expect a lower fertility rate. For example, particularly the workers in mining were likely to have a lower birth rate since miners’ associations provided pension insurance long before the introduction of comprehensive health insurance at the union level (Jopp 2013). To further ensure that our proxy ‘share insured’ does not pick up any fertility development which might be particular to the ‘working class’, we add a variable which measures the share of workers and which has been developed by Scheubel (2013) based on the job description in the Annual Yearbook of Statistics (based on the 1895 and 1907 occupational census). If the fertility rate was different among workers, this variable should pick up this difference. Similarly, it is unlikely that the discretion which RIA officials exercised in approving pension applications was related to the birth rate of a particular group of people rather than by individual motives. One exception to this may have been the discrimination against Slav minorities which had been widespread in provinces with a large Slav population (Kasche and Sniegs 2001), should fertility have been different among Slavs. This cannot be ruled out as Knodel (1974) found higher fertility rates in regions with a larger number of Poles and Galloway et al. (1994) find a significant positive effect of a large Slav population on fertility. We discuss the implications for our analysis when presenting our descriptive results below.

Third, when estimating the effect of pension insurance on fertility we need to make sure that the effect is not confounded by other developments, such as industrialisation. It is obvious that pre-pension system differences between the states or provinces, such as the number of the elderly, the degree of migration or the level of industrialisation would affect both the birth rate and also the pension system indicators. For example, a high number of elderly *ceteris paribus* should result both in a lower birth rate and in a higher number of approved pensions and thus lower pensions. A higher level of industrialisation should result in more working women and thus lower fertility while it would also imply that more people would be insured as there were more workers in the industrialising areas. However, these factors only constitute a problem for identification if we cannot control for them. Hence, we have added an extensive set of proxies which lead us to assume that once controlling for the confounding factors, the variation in key pension system variables across provinces is indeed exogenous.

Choice of the dependent variable

We choose the crude marital birth rate (CMBR) as the main measure of fertility.¹⁴ Since the calculation of fertility indices which are more widespread in the analysis of fertility in a non-historic context requires information on the age of the female population, which in our data set is only available for years 1885 and 1890, the CMBR is easy to compute and available as a long time series. Moreover, the CMBR is well-suited to analyse main cross-

¹⁴ The CMBR can be computed as $CMBR = (1 - illegitimacy\ rate_t) * \frac{Number\ of\ births_t}{1000}$ for all years.

jurisdictional developments in fertility since it maps broadly the same developments as other fertility indices.¹⁵ We also use other fertility measures to check the robustness of our model in section 5.3.3.

Choice of control variables

Particularly the third identifying assumption rests on an appropriate selection of control variables. We choose the variables describing both current and future consumption to reflect earlier empirical studies on the determinants of the first demographic transition.¹⁶ The factors that have previously been found to be the main determinants of the First Demographic Transition (Guinnane 2011 gives a comprehensive overview, other studies are Galloway et al. 1994, 1998; Richards 1977; Brown and Guinnane 2007; and in particular Knodel 1974 for Germany) are consistent with a consumption-based model of fertility like the one we use as a motivation for our study. 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).

The determinants of the first demographic transition include a general (child) mortality decline which increased returns to child quality (since more children survived, the investment in their education became more valuable), which has been found to be associated with a smaller family size. Innovation in contraception and the changed availability of contraception (which was spread by urbanisation and better communication) improved the success of attempts to control fertility. As compulsory schooling laws or laws banning child labour were introduced, the direct costs of children who previously contributed to the household income rose. One factor frequently mentioned in the literature, the higher opportunity cost of children due to increased labour market participation of women, is one of the main mechanisms in our model. Finally, the introduction of comprehensive social insurance reduced the value of children as an insurance against risk.

[Table 1 about here.]

Table 1 details how we have proxied these developments with the variables available in our data. We provide summary statistics for all variables in table 2.

[Table 2 about here.]

¹⁵ Scheubel (2013) illustrates this by comparing the CMBR to other fertility indices which take into account natural fertility and the age structure of women. As information on age structure is only available for years 1871, 1885, and 1890 while pension insurance was introduced in 1891, we cannot use other fertility measures for the analysis in this paper.

¹⁶ 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.

4.3 Econometric considerations

Estimation approach

While we would prefer estimating a panel model, not all variables have been collected by the Imperial Statistical Office for all years. The data collected for Imperial Germany by the Imperial Statistical Office is, for example, 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.¹⁷ The CMBR are available for almost all years. However, variables on the demographic structure or the share of the population working in the primary, secondary or tertiary sector have only been collected during a population or occupational census.¹⁸ In addition, not every variable has been recorded during every population census. For example, information on age structure was not collected after 1890. Occupational information was only collected in the 1871, 1882, 1895 and 1907 occupational censuses. Unfortunately, this also impacts some of the pension system variables. While the level of pensions and the approval rate have been collected almost in every year and for every RIA since 1891, the share insured – which is based on occupational information – has only been collected during the occupational census of 1895 and during the occupational census of 1907. Hence, our main proxy is only available for 1895 and 1907.

The main complication for our empirical specification thus arises from the fact that there is no year during which all variables are available. As our identification strategy builds on variation between jurisdictions, we consider it essential that we are able to control for province-specific effects. Given that our main proxy is available for two points in time, 1895 and 1907, we aim to use panel techniques on a panel of $t = 2$. However, not all control variables are available for 1895 and 1907. While other authors have imputed or extrapolated values if they were missing for some variables for some years (e.g. Becker et al. 2010, 2011a,b), the possibility to do so is limited if a variable is only available for $t = 2$ and $n = 23$. Hence, we resort to a solution used in previous studies (e.g. Galloway et al. 1994, 1998).

In particular, we construct a panel of two periods, r and s . As our main proxy is available for 1895 and 1907, $r = 1895$ and $s = 1907$. For variables which are not available in r and s but for two other years, we use the first year for cross section r and the second year for cross section s . For example, information on the age structure of the population is only available for 1871, 1885 and 1890. Hence, for all proxies based on the age structure, such as the old age dependency ratio, $r = 1885$ and $s = 1890$. Table 1 gives an account of the years of availability for each variable we use in the model and also lists the years which we use for the construction of the two panels. While the data set resulting from this approach is not a clear pseudo panel consisting of two pooled cross

¹⁷ For further details on the quality of Imperial German statistics, refer to Sniegs (1998) and Scheubel (2013).

¹⁸ Before 1895 population censuses were conducted almost every year. After 1895 population censuses were conducted in 1899, 1900, 1905, 1909 and 1910.

sections because most variables are included for the same observations for two different points in time, we think of it as a pseudo panel since we do include some control variables for different points in time.

Using our pseudo panel can introduce biases if r and s are very different. We adjust those variables not expressed in percentage terms to the population size in the year from which they are taken to make the numbers comparable. Moreover, we provide several robustness checks regarding the selection of years.

After constructing the pseudo panel, we run a panel estimator on the two cross-sections to allow us to account for province-specific effects. To control for those unobserved province-specific effects we use a fixed effects estimator, similar to e.g. Galloway et al. 1994, who also constructed their pseudo panel in a similar way. To account for the invariant region-specific effects, we use standard errors adjusted for serial correlation. In addition, as errors can be correlated across adjacent provinces (spatial correlation), we also use standard errors which are robust to spatial correlation.

Model specifications

In line with our identifying assumptions and corresponding to our theoretical considerations, we estimate a model in which the share insured and the main determinants of the fertility decision enter our econometric model additively. Our empirical specification reads:

$$n_{i,t} = y_0 + T_t + \beta_\tau \tau_{i,t} + \mathbf{x}_{i,t} \beta_x + \mathbf{z}_{i,t} \beta_z + \alpha_i + \varepsilon_{i,t}. \quad (22)$$

Note that this specification corresponds in spirit to our model (e.g. equation 4), assuming that the fertility decision is determined by the four main elements of the utility function: the pleasure of having children and the pleasure of supporting the elderly, the impact on current consumption, and the impact on future consumption. The measure $n_{i,t}$ refers to the crude marital birth rate (CMBR) (or in our sensitivity analysis, to the Marital Fertility Index, MFI, and the Total Fertility Index, TFI) in jurisdiction i in year t . T_t is a time-specific effect, i.e. in most specifications a dummy for year 1907. We capture the overall effect of pension insurance by the share insured which we label $\tau_{i,t}$. Hence, the coefficient $\beta_{\tau_{i,t}}$ should give an estimate of the total effect, i.e. of $\frac{\partial n}{\partial \tau}$ for the case when the share insured is raised from 0 to τ . $\mathbf{x}_{i,t}$ is a vector of demographic variables which affect the pleasure of having children, the pleasure of supporting the elderly, as well as current and future consumption (including the consumption of the elderly who are part of the household); $\mathbf{z}_{i,t}$ is a vector of variables related to industrialisation which affect current and future consumption; α_i refers to time-invariant region-specific effects and $\varepsilon_{i,t}$ is an i.i.d. error term.

To test the robustness of our results, we reproduce specification (22) with a different proxy for the overall effect of pension insurance, the share insured weighted with the approval rate, as shown in equation (23). The approval rate is denoted by a and the weighted share insured is denoted by τ^w :

$$n_{i,t} = y_0 + T_t + \beta_\tau \tau_{i,t} + \beta_{\tau^w} \tau_{i,t}^w + \beta_a a_{i,t} + \mathbf{x}_{i,t} \beta_x + \mathbf{z}_{i,t} \beta_z + \alpha_i + \varepsilon_{i,t}. \quad (23)$$

To test the relevance of income and price effect, we estimate a model with the implicit tax, denoted by $t_{i,t}$, and the average contribution, denoted by $c_{i,t}$, as additional explanatory variables:

$$n_{i,t} = y_0 + T_t + \beta_\tau \tau_{i,t} + \beta_t t_{i,t} + \beta_c c_{i,t} + \mathbf{x}_{i,t} \boldsymbol{\beta}_x + \mathbf{z}_{i,t} \boldsymbol{\beta}_z + \alpha_i + \varepsilon_{i,t}. \quad (24)$$

Finally, to evaluate the impact of the internal rate of return, we add the proxy for it, denoted by Ω , as well as an interaction term with the share insured, denoted by Ω^τ to the model:

$$n_{i,t} = y_0 + T_t + \beta_\tau \tau_{i,t} + \beta_\Omega \Omega_{i,t} + \beta_{\Omega^\tau} \Omega_{i,t}^\tau + \mathbf{x}_{i,t} \boldsymbol{\beta}_x + \mathbf{z}_{i,t} \boldsymbol{\beta}_z + \alpha_i + \varepsilon_{i,t}. \quad (25)$$

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 (refer to figure 2). The decline became particularly pronounced around 1900 when the pension system was turned from a partially funded system into a full pay as you go system. In our sample, the crude marital birth rate fell from more than 33 births per thousand in 1895 to less than 30 births per thousand in 1907.

This decline in birth rates is correlated with the change in the share insured. While the average share of the population which was insured in pension insurance only rose marginally from 21.3% to 21.6% between 1895 and 1907, this small difference hides substantial increases in some provinces. The left panel of figure 3 shows that the change in the share insured between 1895 and 1907 ranged between -4 and 3 percentage points, but in most provinces the share insured rose by between 1 and 2 percentage points. Figure 3 also highlights the negative cross-regional correlation between the change in the share insured and the change in the crude marital birth rate and hence supports our hypothesis that the pension system has contributed to the decline in birth rates.

[Figure 3 about here.]

The correlation between the change in the share insured and the change in crude marital birth rates would be even stronger if we disregarded the most rural provinces. The Kingdom of Bavaria (*Bayern*) is the most obvious outlier in this respect, being largely rural and one of the largest provinces. Also most provinces in the Eastern part of Prussia were largely rural *Posen*, *Westpreußen* and *Ostpreußen*.

In addition, the left panel of figure 3 shows that the change in birth rates was not significantly different in those East Prussian provinces with large Slav minorities. The provinces with large Slav minorities are highlighted. While the change in birth rates in the provinces with large Slav minorities has been at the lower end of the range, they are not significantly below the range of the other provinces.

To render further support to our hypothesis that the introduction of pension insurance contributed to the fertility decline in Imperial Germany, we illustrate the negative relationship between the change in the weighted share insured and the change in marital birth rates in the right panel of figure 3. The right panel of figure 3 illustrates that the negative correlation between the change in the weighed share insured and the change in birth rates persists.

However, the right panel of figure 3 also shows that in terms of weighted share insured, the East Prussian provinces with large Slav minorities differed significantly from other provinces. While the change in the birth rate was not significantly lower than in other provinces, the change in the weighted share insured was at the lower end of the range in those provinces with large Slav minorities. This supports the observation by Kaschke and Sniegs (1998) that RIA officials in those provinces with large Slav minorities discriminated against Slav when deciding on the approval of a pension application. We acknowledge this in our multivariate analysis by excluding Ostpreussen as the most obvious outlier and by checking the robustness of our results with regard to excluding the provinces with large Slav minorities altogether.

5.2 Multivariate Analysis

Our multivariate analysis indicates that the negative relationship between the share insured and the birth rate persists when controlling for other determinants of the first demographic transition. Table 3 shows four specifications to illustrate that the overall effect of pension insurance on fertility was negative. Two additional specifications confirm that our model framework is applicable to the data.

[Table 3 about here.]

Column (1) is equivalent to the left panel of figure 3 and confirms the significant correlation between the change in the CMBR and the change in the share insured. The coefficient indicates that a change in the enrolment in pension insurance by 1 percentage point is associated with an average reduction of fertility by approximately 0.54 marital births per thousand. This is quite substantial considering that on average, the standard deviation of the share insured in our sample is 3% and the standard deviation of marital births per thousand in our sample is 4.3 and in view of the fact that in our sample marital births per thousand only fell from 33.38 marital births per thousand in 1895 to 29.79 in 1907.

Columns (2) and (3) show that our results continue to hold if we add control variables, the coefficients on which are in line with standard demographic transition theory. Column (2) adds basic demographic information. As the consumption value of children typically rises with marriage, we expect that 1 marriage per thousand leads to approximately 1 more birth per thousand. By contrast, the need to care for the elderly, which we proxy by the old age dependency ratio, reduces disposable income and may thus lead a couple to have fewer children. Finally, Catholicism has been associated with higher birth rates. We look at the number of Protestant inhabitants in a province relative to

the number of Catholic inhabitants. Correspondingly, we expect this proxy to have a negative effect on fertility.

Adding demographic information in column (2) confirms that the negative effect of pension insurance persists. However, contrary to our expectation, the coefficient on marriages is not significant in specification (2). This may be related to unobserved correlation with variables which are not included in the specification in column (2) since the coefficient turns significant and of the expected magnitude in our sensitivity analyses in table 4. Similarly, neither the old age dependency ratio nor the share of Protestants are significant, also suggesting potential omitted variable bias from other determinants of the first demographic transition which are not included in column (2).

Hence, we add in column (3) proxies for industrialisation which have been found to be key determinants of the first demographic transition. These include a measure of the share of workers developed by Scheubel (2013), as discussed above, the gender imbalances ratio, which is a proxy for migration (refer to table 1), the share of recruits with at least basic schooling to measure the diffusion of education, the share of revenues in contribution category I relative to the other categories to proxy the share of working women,¹⁹ the lagged number of persons per building to proxy urbanisation, the share of the population working in trade to proxy the diffusion of knowledge and an index measuring the average harvest per hectare based on data for five different types of crops (refer to table 1).²⁰ Adding information on industrialisation in column (3) confirms the negative effect of pension insurance on fertility in addition to confirming the main determinants of the first demographic transition. While the share of workers in a province is not significant,²¹ the gender imbalances ratio has a significantly negative effect on the birth rate; an increase in the gender imbalances ratio by 10% is associated with a decrease of 1.7 births per 1000.

The proxy for urbanisation indicates a negative and significant impact on the birth rate which is fairly consistent also across other specifications. It suggests that an additional person in a building is associated with a reduction of approximately 1 marital birth per thousand.

The proxy for the diffusion of knowledge is highly significant in specification (3) and suggests that an additional 1% of the population working in trade is associated with 0.5 fewer births per thousand. While this variable is not significant in all specifications in table 3, its negative effect is confirmed in the sensitivity analysis in table 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 in-

¹⁹ Women's wages were lower such that almost only women contributed in the lowest contribution category (Haerendel 2001).

²⁰ Increases in agricultural productivity were a pre-condition for industrialisation (Murphy et al. 1989) and 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).

²¹ As discussed above, this may be related to the fact that our measure of workers includes those working in mining. Miners' associations provided pension insurance before the introduction of comprehensive health insurance (Jopp 2013). Hence, any positive correlation between the share of workers and birth rates may be confounded by the negative correlation between the share working in mining and the birth rate.

insurance introduced at the time. Therefore, it would make sense to assume that other insurance like health care coverage should also have an effect on fertility. Hence, we also add a measure of health care coverage in column (3): the share of the population covered by the previously introduced health insurance. In fact, the coefficient suggests that health insurance coverage has a positive effect on births. This may be related to health insurance reducing the mortality of both mothers and children.

That being said, the insight we gain from column (3) is an important one: our model confirms previous findings from the demographic transition literature, but it also shows that pension insurance had a significant additional impact. For example, column (3) implies that the effect of an increase of the share insured by 1% is approximately equivalent to an increase in the gender imbalances ratio by 3%.

In column (4) we confirm the negative effect of pension insurance for a different proxy which shows that the significant negative coefficient on the share insured is related to pension insurance instead of picking up e.g. some particular characteristics of the group of insured. In particular, we add the weighted share insured, which is the share insured weighted with the approval rate. As we also have to add the approval rate, the total marginal effect of the share insured can be derived as $\beta_\tau + \beta_{\tau w} * a$. If all pension applications would be approved, i.e. if $a = 1$, the total effect of the share insured would amount to a reduction in the birth rate of 1.05 per thousand. At the average approval rate, the total marginal effect of the share insured is equivalent in magnitude to the effect of the unweighted share insured in column (3).

In addition, we illustrate in column (5) that the underlying behavioural mechanisms are in line with the framework of our model. In column (5) we add to the basic specification from column (3) two variables which we consider as best available proxies of an income and a price effect. In line with our expectations, adding the two proxies for the income effect and the price effect reduces the coefficient on the share insured and both proxies are significant and of the expected sign. However, the magnitude of the price effect proxy coefficient is larger than the income effect proxy coefficient and the coefficient on the share insured is only halved and remains significant. Hence, we consider specification (5) as a confirmation of our model framework, but as also highlighting that for one the model framework may have its limitations in explaining the dynamics of Bismarck's pension system which after all has been a mixed system for the first 10 years of its existence and for another the proxies may be imperfect and picking up unobserved differences in the pension system which we cannot measure. For example, it cannot be ruled out fully that a higher contribution per insured could be picking up a higher wage level in a province.

In column (6) we test an additional element of our model which tells us that the overall effect of pension insurance should be stronger the lower the internal rate of return of the pension system (hypothesis 3). We add the proxy for the internal rate of return as well as an interaction with the share insured to our baseline model. Both the share insured and the interaction term are significant in this specification. Again, the coefficient on the share insured can be calculated as $\beta_\tau + \beta_{\Omega\tau}\Omega$. At the average level of the internal rate of return in our sample 103.721 the total effect of the share insured

on the birth rate would be equivalent to a reduction by 0.4 births per thousand at a 1% increase in the share insured, which corresponds to the coefficient in our baseline specification.

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, we illustrate a comparison with a simple OLS model and with a model in first differences in the supplementary appendix C.²²

Assuming that the province-specific unobserved effects are well-captured in a fixed effects model, the model may however 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 α_i with the error term $\varepsilon_{i,t}$. 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 ($T = 2$), these methods cannot be used for our small sample. At the same time, when running the basic model only with the variables which are available for more than just a few periods (such as marriages, agricultural productivity, education, share of contributions in category I), implementing a Driscoll and Kraay (1998) adjustment of the standard errors gives coefficients of broadly the same magnitude as our small-sample 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 effects other than pension insurance. There are three major changes which are of particular interest. First, the pension system was reformed in 1899; the law came into effect in 1900. This change turned the previously mixed system into a full pay as you go system. A new contribution category was introduced. In addition, a new financial equalisation scheme between RIAs was introduced. Second, in 1903 there was a major amendment to child labour laws (Boentert 2007) which rendered children more costly in the sense that stricter child labour laws reduced the scope for current consumption as children went to school instead of contributing to household income.²³ Third, in 1903 and particularly in 1904,

²² 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.

²³ 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

the earlier 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) which may have led to more restrictive approval practices and thus have lowered the probability of receiving a pension.

To test whether there was a major difference in coefficients if we do not use 1907 as a reference year we run our baseline specification (column (3) in table 3) comparing the year 1895 to the key years 1899, 1900, 1903 and 1904. As a placebo check we also compare the year 1904 to the year 1907. Also for this sensitivity analysis we construct a panel of $t = 2$ from a pooled cross section. As some control variables are only available for 2 years, we have to use the same observations for all pseudo panels for these control variables, as indicated in the notes to table 4. Particularly, this applies to the share insured, the share of Protestants, the share of workers, and the share working in trade.

[Table 4 about here.]

Column (1), which compares years 1895 and 1899 indicates a significant negative effect of the share insured on the birth rate. However, the coefficient is not as large as in column (3) in table 3. This gives some support that at least half of the effect we have seen in column (3) of table 3 is driven by the introduction of pension insurance.

The coefficient on the share insured is marginally not significant in column (2) which compares years 1895 and 1900 while being marginally significant and comparable in magnitude to the baseline specification when we compare years 1895 and 1903. This effect persists when comparing years 1895 and 1904 in column (4). While the significant coefficient in column (3) could be interpreted as child labour laws having an effect on household disposable income and thus reinforcing the impact of compulsory pension insurance, the evidence from column (4) suggests that the review of the code of conduct of RIAs in 1903/1904 may have raised awareness about the pension system among the population and hence may have intensified any behavioural reaction. The placebo comparison in column (5) which compares years 1904 and 1907 confirms that the main effect we show in our baseline model is driven by the years before 1904.

The magnitude of the control variables in table 4 is broadly in line with our baseline specification. Similar to the baseline specification, marriages have a positive effect on the birth rate while a higher relative share of Protestants is associated with a lower birth rate. Migration, education, female labour force participation and a high share of people working in trade also reduce the birth rate. The coefficient on the share of workers is positive in some of the specifications in table 4 which is in line with our initial hypothesis. While the proxy for urbanisation, the number of people per building, has the expected negative effect in most specifications, we suspect that the positive effect in column (1)

in all areas of production. The 1878 amendment prohibited children below the age of 14 to work in factories. 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.

may be related to the lag being too small for the years used in that specification.²⁴ While the coefficient on the crop yield index is consistent with the baseline model in the specification in which it is significant, we relate the inconclusive behaviour of this variable in table 4 to the fact that we had to extrapolate some values, particularly around 1900.

5.3.3 Measuring fertility

While we have already discussed that the CMBR is a meaningful measure of fertility, especially in the historical context, we show that other measures of fertility give comparable results for the years 1885 and 1890 for which we can compute these alternative measures. Typical fertility indices, which are used in cross-country studies, are the total fertility rate (TFR), or the Total Fertility Index (TFI) and the Marital Fertility Index (MFI) developed by Coale (1965, 1969), which are slightly more sophisticated as it takes into account natural fertility.²⁵

One caveat to looking at other measures of fertility is that we cannot include the share insured and the proxy for the share of working women from our baseline specification. As pension insurance had not yet been introduced, these were not available for the years 1880 and 1885. However, a regression only using the fertility determinants available for 1880 and 1885 is broadly in line with our baseline model and helps to illustrate that the use of the CMBR instead of more sophisticated fertility indicators yields reliable results.

Table 5 shows such a specification which compares years 1885 and 1890. Column (1) shows the CMBR and column (2) the crude birth rate (CBR). Column (3) shows the corresponding MFI and column (4) the corresponding TFI.

It is obvious that models (1) and (2) as well as models (3) and (4) are comparable in terms of the variables which they confirm as important determinants of fertility. Therefore, we conclude that using the CMBR 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 5 about here.]

To show that the similarity is not driven by the specification of the model being too inflexible, we additionally provide evidence that the model identifies different determinants in case the dependent variable measures something different: column (5) shows the model predictions when using the share of non-marital births as dependent variable. The results in column (5) differ from the other columns where we expect them to, e.g. marriages reduce the illegitimacy rate, but the illegitimacy rate is not affected if a large share of wives and husbands is separated due to migration.

²⁴ 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.

²⁵ 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 TFI are those in Henry (1961). The definitions used for the computation are provided in table 1.

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 provide further evidence on a well-known theoretical concept in public economics, the social security hypothesis. In addition, we use a 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.

The theoretical model, adding to the literature on overlapping generation models, highlights the effects of compulsory contributions to a pension system on fertility when labour supply is endogenous. Our crucial assumption is that labour supply is reduced when a household has children, which translates into an opportunity cost of having children in terms of foregone lifetime income. This gives rise to two counterbalancing effects in equilibrium: (i) a higher contribution rate to the pension system reduces this opportunity cost, leading to a *positive* effect on fertility (which we name the price effect) and (ii) a higher contribution rate to the pension system lowers lifetime income to the extent that there is an implicit tax in the pension system, leading to a *negative* effect on fertility (which we name the income effect). While the sign of the overall effect is determined by the larger of the two, the size of the overall effect is determined by the internal rate of return of the pension system.

Our empirical results confirm that a higher enrolment rate in pension insurance leads to a lower fertility rate. We use a historical data set which covers the introduction of the Bismarckian pension system at the end of the nineteenth century in Imperial Germany. This data set allows us to exploit cross-jurisdictional variation in the regional enrolment rate for identification.

The results are robust even when controlling for other determinants of the first demographic transition, confirming the residual effect of pension insurance on the fertility decline. When controlling for those determinants, an increase of the share insured by 1% translates into a total reduction of approximately 0.5 marital births per thousand. This corresponds to a contribution of 15% of the total decline in birth rates between 1895 and 1907.

Because our analysis only covers the time span 1895–1907, 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. Given that the direct effect of pensions on fertility amounted to almost 15% of the overall decline between 1895 and 1907, 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 NBER, 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. *Econometrica*, 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.
- Bental, B. (1989). The old age security hypothesis and optimal population growth. *Journal of Population Economics*, 1:285–301.
- 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. (2015). Fertility and social security. *Journal of Demographic Economics*, 81:261–299.
- Borjas, G. J. (1999). Immigration and welfare magnets. *Journal of Labor Economics*, 17:607–637.
- Bourdieu, J. and Kezstenbaum, L. (2007). Surviving old age in an ageing world: Old people in france 1820-1940. *Population*, 62:183–211.
- 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.
- Caldwell, J. C. (1978). A theory of fertility: From high plateau to desestabilization. *Population and Development Review*, 4:553–577.
- 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 dependent 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.
- Egger, P. and Radulescu, D. (2012). Family policy and the number of children: Evidence from a natural experiment. *European Journal of Political Economy*, 28:524–539.
- Ehrlich, I. and Kim, J. (2007). Social security and demographic trends: Theory and evidence from the international experience. *Review of Economic Dynamics*, 10:55–77.

- 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.
- Fenge, R. and Meier, V. (2009). Are family allowances and fertility-related pensions perfect substitutes? *International Tax and Public Finance*, 16:137–163.
- Fenge, R. and von Weizsäcker, J. (2010). Mixing bismarck and child pension systems: an optimum taxation approach. *Journal of Population Economics*, 23:805–823.
- 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. (2014). Spatial analysis of the causes of fertility decline in prussia. *Population and Development Review*, 40:497–525.
- 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.
- Hirazawa, M., Kitaura, K., and Yakita, A. (2014). Fertility, intra-generational redistribution, and social security sustainability. *Canadian Journal of Economics*, 47:98–114.
- 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, Fund Size, and Concentration: Prussian Miners’ Knappschaften in the Nineteenth- and Early Twentieth-Centuries and Their Quest for Optimal Scale*. Akademie Verlag, Berlin.
- Jopp, T. (2014). What had a promise from early social insurance once been worth. mimeo.
- 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.
- Khoudor-Cásteras, D. (2008). Welfare state and labor mobility: The impact of bismarck’s social legislation on german emigration before world war I. *Journal of Economic History*, 68:211–243.
- 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.
- Puhakka, M. and Viren, M. (2012). Social security, saving and fertility. *Finnish Economic Papers*, 25:28–42.
- Reichsversicherungsamt (1910). *Das Reichsversicherungsamt und die Deutsche Arbeitsversicherung*. 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, Tübingen.
- 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. Cotta’schen Buchhandlung, Stuttgart.
- Willis, R. J. (1979). The old-age security hypothesis and population growth. NBER Working Paper.

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.

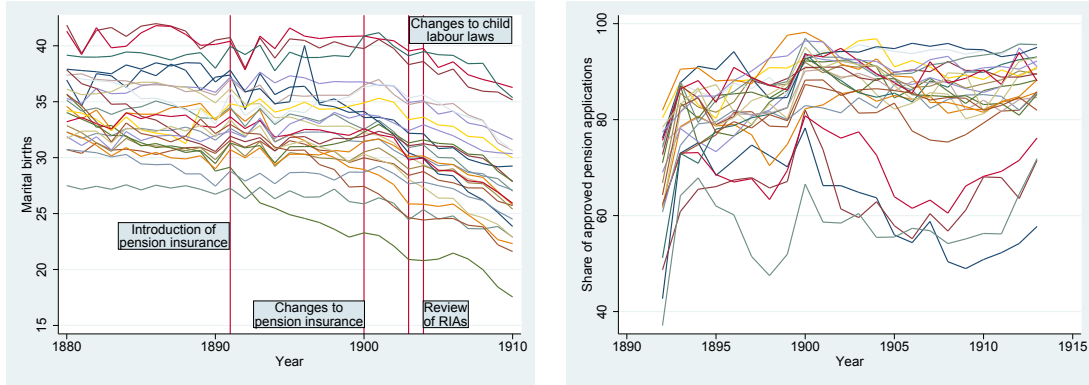
Appendix A

Figures

Figure 1: REGIONS IN IMPERIAL GERMANY

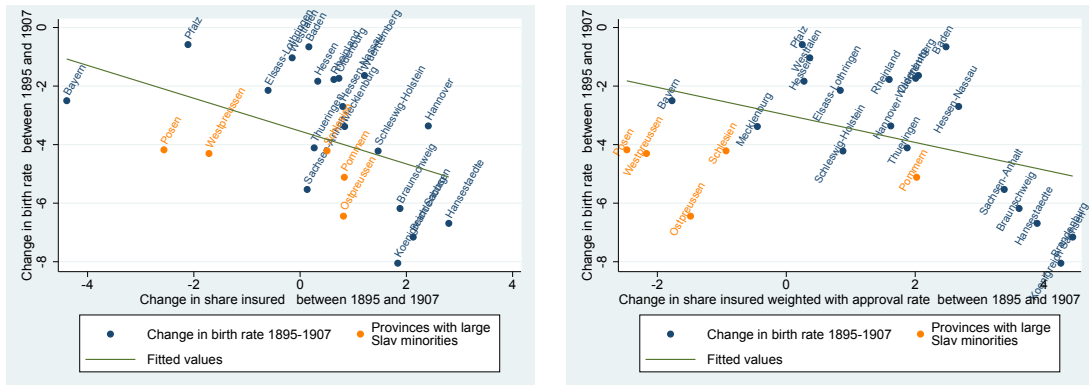


Figure 2: MARITAL BIRTH RATES AND PENSION APPLICATION APPROVAL RATES



Notes: Marital birth rates for all regional entities in the data set, expressed in per thousand. Rate of approved pension applications for all regional entities in the data set, expressed in %. For the sake of illustrating trends, region names are suppressed.

Figure 3: CHANGES IN MARITAL BIRTH RATES AND THE SHARE INSURED



Birth rates expressed in per thousand. Change in share insured expressed in percentage points.

Tables

Table 1: AVAILABILITY OF VARIABLES

Variable	Proxy	Description (Unit)	Years avail- able	Years used
<i>Fertility measures</i>				
Fertility	$Marital fertility n_t^m$	Marital births (1/1000 births)	1871, 1878- 1914	1895; 1907
Fertility n_t	$Number of children n_t$	Crude births (1/1000 births)	1871, 1878- 1914	1895; 1907
Fertility	$Total fertility index (TFI) = \frac{I_{t,i}}{n_{g,i} F_{g,i}}$ 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	The TFI can take a maximum value of 1 if a society practices no birth control; oth- erwise it is always smaller than 1	1871; 1885; 1890	1885; 1890
Fertility	$Marital fertility index MFI = \frac{I_{t,i}^m}{n_{g,i}^m F_{g,i}}$ where $n_{g,i}^m$ is the number of married women in age group g in province i and $F_{g,i}$ is the natural fertility for age group g	The MFI can take a maximum value of 1 if all married couples practice no birth control; oth- erwise it is always smaller than 1	1871; 1885; 1890	1885; 1890
Number of illegiti- mate children	Illegitimacy rate i_t	(%)	1871, 1878- 1914	1895; 1907
<i>Pension system variables</i>				
Pension insurance	$\frac{\partial n}{\partial \tau}$	Share enrolled (%)	1895; 1907	1895; 1907

Pension insurance	Approval rate of pension applications	Share of pension applications approved, average over old age and invalidity pension (%), except for weighing the rate of insured)	1895; 1907	1895; 1907
Income effect	Implicit tax in pension system	Average contribution / average pension (Mark). Average contribution calculated as total contributions divided by number of insured. Average pension has been recorded by RIAs.	1891-1893, 1895-1913	1895, 1907
Price effect	Contributions to pension system	Contributions per insured (Mark)	1891-1893, 1895-1913	1895, 1907
Internal rate of return	Future pension / current contribution	The average future pension is proxied by the current average pension.	1895; 1907	1895; 1907
<i>Demographic variables</i>				
Intrinsic utility of having children	Marriages	Marriages per capita (1/1000 marriages)	1871; 1897-1914	
	Relative share of protestants	number Protestants / number Catholics (persons) ²⁶	1871; 1880; 1885; 1890	1885; 1890

²⁶ The diffusion of Catholicism is a standard control variable in demographic transition theory (e.g. Galloway et al. 1994) and has been found to be an important determinant of culture and behaviour (e.g. Becker and Wössmann 2009). Short run changes in the number of Catholics or Protestants may reflect three main factors. First, these changes may reflect births and hence be endogenous to the birth rate. Second, they may reflect migration. Third, they may reflect cultural differences between Catholics and Protestants. As we are only interested in the third effect, we have to make sure that our proxy neither reflects births nor migration. To avoid the endogeneity between the number of Protestants and births, we include the relative share of Protestants with at least a 10-year lag. To avoid measuring migration with this variable, we add a proxy for migration separately.

Intra-family transfers / need to care for the elderly	Old age dependency ratio (%)	Population older than 60/population aged 15-20 * 100 ²⁷	1871; 1885; 1890	1885; 1890
---	------------------------------	--	------------------	------------

Industrialisation variables

Contribution of children to household income / child labour	Share workers	Workers / working population (%)	1871; 1882; 1895; 1907	1895; 1907
Migration	Gender imbalances ratio	Married women over married men * 100. If this measure is significantly larger than 100, it provides and indication that relatively more men than women have (temporarily) left a province, most often to work in one of the quickly industrialising areas (e.g. Haines 1976; Nugent 1995)	1871; 1880; 1885; 1890; 1900; 1905; 1910	1885; 1890
Direct cost of children associated with education	Recruits with basic education	Recruits with basic education / 1000 recruits (persons)	1871; 1879-1911	1895; 1907
Female labour force participation rate	Contributions in category I (in which mainly women contributed)	Contributions in category I / sum of contributions in cat. II-IV(V) * 100 (%)	1891-1893; 1895-1913	1895; 1907
Urbanisation	Persons per building	Persons per building. To avoid endogeneity with the current birth rate, we include this variable with at least a 10 year lag.	1871; 1880; 1885	1880; 1885

²⁷Value for 1890 is estimated since only information on the age bracket age 10-20 was collected.

Changed availability of contraception, diffusion of knowledge	Communication networks and cross-regional trade	People working in trade / working population (%)	1871; 1882; 1895; 1907	1895; 1907
Growth, return to capital	Index measuring the yield (tons) per hectare for different crops.	Tons / hectare, average for wheat, barley, rye, oats, potatoes, hey ²⁸	1880-1897; 1899-1914	1895;1907
<i>Social security variables</i>				
Social security coverage	Share insured in health insurance	Number in health insurance / population * 100 (%)	1888-1893; 1895-1899; 1913	1895; 1899 ²⁹

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

Table 2: SUMMARY STATISTICS

Variable	Obs	Mean	Std. Dev.	Min	Max
Marital births (1/1000)	46	31.56	4.28	23.66	41.59
Share insured (%)	46	21.44	2.97	15.79	28.18
Share of approved pension applications (%)	46	81.68	9.07	56.86	95.29
Implicit tax (contribution per insured - av. pension (Mark))	46	-134.80	19.10	-172.83	-85.24
Contribution per insured (Mark)	46	12.04	10.01	0.67	54.07
Internal rate of return (Proxy)	46	103.72	138.61	18.97	748.96
Marriages (1/1000)	46	7.91	0.44	6.85	8.90
Old-age dep. ratio (%)	46	13.60	2.13	8.52	18.30
Gender imb. ratio (%)	46	20.96	33.51	0.26	155.52
Educated recruits (1/1000)	46	60.99	21.24	24.01	85.39
Protestants (per Catholic)	46	100.07	2.07	88.23	103.15
Share revenues in cat. I (%)	46	997.22	10.29	949.31	1000.00
Share workers (%)	46	18.94	19.70	1.10	77.09
Share working in trade (%)	46	8.10	1.57	6.10	11.20
Persons per building	46	8.55	5.00	2.55	25.19
Crop yield index	46	3.79	0.66	2.76	5.06
Insured: health (%)	46	15.57	6.86	4.87	33.58

²⁸ 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).

²⁹ For Oldenburg: 1895; 1913.

Table 3: DETERMINANTS OF THE FIRST DEMOGRAPHIC TRANSITION

	Dep. var: marital births (1/1000)					
	Hypothesis 1				Hypothesis 2	Hypothesis 3
	(1)	(2)	(3)	(4)	(5)	(6)
Pension system variables						
Share insured (%)	-.542 (.239)**	-.605 (.298)**	-.453 (.264)*	2.123 (1.666)	-.278 (.162)*	-1.646 (.711)**
Share insured weighted with approved pension applications (%)				-3.168 (1.704)*		
Share of approved pension applications (%)				.512 (.411)		
Implicit tax (contribution per in- sured - av. pension (Mark))					-.272 (.063)***	
Contribution per insured (Mark)					.475 (.099)***	
Int. rate of return proxy (av. pension / contribution per in- sured)						-.090 (.088)
Share insured * int. rate of return						.012 (.006)**
Demographic variables						
Marriages (1/1000)		1.623 (1.042)	-.616 (1.450)	.818 (1.027)	-.769 (.579)	-.503 (.923)
Old-age dep. ratio (%)		-.535 (.498)	-.621 (.659)	-.608 (.495)	.474 (.423)	-.625 (.503)
Protestants (per Catholic)		.023 (.032)	-.009 (.028)	.045 (.033)	-.091 (.022)***	-.003 (.034)
Industrialisation variables						
Share workers (%)			-.017 (.078)	.037 (.070)	-.025 (.057)	-.041 (.032)
Gender imb. ratio (%)			-.174 (.089)*	-.015 (.074)	-.187 (.062)***	-.183 (.058)***
Educated recruits (1/1000)			-.554 (.226)**	-.267 (.331)	-.113 (.107)	-.614 (.151)***
Share revenues in cat. I (%)			.018 (.053)	.047 (.040)	-.002 (.031)	-.026 (.041)
Persons per building			-1.021 (.452)**	.472 (.603)	-1.251 (.272)***	-.985 (1.340)
Share working in trade (%)			-.523 (.300)*	-.093 (.289)	-.641 (.227)***	-.359 (.323)
Crop yield index			1.547 (.722)**	.842 (.442)*	.877 (.357)**	1.230 (.530)**
Social insurance variables						
Insured: health (%)			.159 (.778)	.472 (.655)	.934 (.474)**	.253 (.572)
Obs.	46	46	46	46	46	46

Estimation with fixed effects OLS, standard errors adjusted to serial correlation. Significance level: *** : $p < 0.01$; ** : $p < 0.05$;* : $p < 0.1$. For variable definitions refer to table 1.

Table 4: SENSITIVITY: POLICY EFFECTS

	Dep. var: marital births (1/1000)				
	(1895/1899)	(1895/1900)	(1895/1903)	(1895/1904)	(1904/1907)
Pension system variables					
Share insured (%)	-.270 (.159)*	-.284 (.180)	-.623 (.243)**	-.526 (.214)**	.083 (.066)
Demographic variables					
Marriages (1/1000)	.652 (.706)	.395 (.248)	.532 (.269)**	1.494 (1.004)	1.034 (.348)***
Old-age dep. ratio (%)	.078 (.214)	-.258 (.382)	.031 (.469)	.468 (.389)	-.249 (.162)
Protestants (per Catholic)	-.0002 (.0001)**	-.0001 (.0001)	.00007 (.0002)	-.00004 (.0003)	-.003 (.001)*
Industrialisation variables					
Gender imb. ratio (%)	-.029 (.029)	-.096 (.033)***	-.778 (.304)**	-.041 (.058)	-.081 (.017)***
Educated recruits (1/1000)	-.410 (.187)**	-.539 (.340)	-.048 (.200)	-.100 (.157)	.055 (.224)
Share revenues in cat. I (%)	-.060 (.031)*	-.175 (.125)	-.061 (.057)	-.097 (.034)***	-.028 (.050)
Share workers (%)	-.019 (.018)	.070 (.028)**	.079 (.040)*	.011 (.036)	-.021 (.017)
Share working in trade (%)	-.899 (.264)***	-1.087 (.637)*	-1.901 (1.048)*	-.826 (.609)	-.069 (.079)
Persons per building	.318 (.125)**	-.454 (.179)**	-1.049 (.348)***	-.092 (.245)	-.469 (.130)***
Crop yield index	.490 (.424)	-.087 (.893)	-.586 (1.405)	2.459 (.809)***	.164 (.157)
Social insurance variables					
Insured: health (%)	.236 (.286)	.542 (.552)	.791 (.963)	.078 (.703)	.906 (.467)*
Obs.	46	46	46	46	46

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: *** : $p < 0.01$; ** : $p < 0.05$; * : $p < 0.1$.

Control variables in columns (1)-(4) not from years 1895 and 1899/1900/1903/1904: share insured (1985 and 1907), gender imbalances ratio (1885 and 1890), old age dependency ratio (1885 and 1890), share workers and in trade (1882 and 1895), relative share of Protestants (1885 and 1890), persons per building (1880 and 1885). Control variables in column (5) not from years 1904 and 1907: share insured (1985 and 1907), gender imbalances ratio (1885 and 1890), old age dependency ratio (1885 and 1890), share workers and in trade (1882 and 1895), relative share of Protestants (1885 and 1890), persons per building (1880 and 1885).

Table 5: SENSITIVITY: MEASURING FERTILITY

Dep. var	Marital births	Births	Marital fert. index	Total fert. index	Illeg. rate
Demographic variables					
Marriages (1/1000)	2.275 (1.017)**	2.699 (1.009)***	.047 (.030)	.029 (.012)**	-.555 (.376)
Old-age dep. ratio (%)	.478 (.357)	.644 (.328)**	.017 (.011)	.011 (.007)*	-.391 (.158)**
Protestants (per Catholic)	.001 (.002)	.0006 (.002)	-.0001 (.00006)**	.00004 (.00002)	.002 (.0007)***
Industrialisation variables					
Gender imb. ratio (%)	.008 (.048)	-.066 (.059)	-.006 (.002)***	-.0008 (.0008)	.018 (.023)
Educated recruits (1/1000)	-.002 (.014)	.026 (.010)***	-.001 (.0003)***	-.00005 (.0002)	-.016 (.006)***
Share workers (%)	-.043 (.023)*	.018 (.013)	-.002 (.0006)***	.00004 (.0002)	-.020 (.010)*
Share working in trade (%)	-.949 (.730)	-1.227 (.836)	-.012 (.026)	-.012 (.011)	.630 (.311)**
Persons per building	.384 (.180)**	.004 (.134)	.011 (.004)***	-.0009 (.002)	-.110 (.099)
Crop yield index	-.006 (.010)	.003 (.010)	.0007 (.0003)**	.0005 (.0002)**	-.006 (.005)
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. 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.

Appendix B: Details on the theoretical model

B.1 Second order conditions

In the model of the Bismarckian pay-as-you-go pension system the second derivatives of equations (5), (6) and (7) are given by:

$$\begin{aligned} V_{nn} = & -U_c(1-\tau)w_tf''(n_t) - U_z\Omega_{t+1}\tau w_{t+1}f''(n_t) \\ & + U_{cc}[(1-\tau)w_tf'(n_t) + \pi_t]^2 + U_{zz}[b_{t+1} - \Omega_{t+1}\tau w_{t+1}f'(n_t)]^2 + U_{nn} \end{aligned} \quad (\text{B.26})$$

$$V_{ns} = V_{sn} = U_{cc}((1-\tau)w_tf'(n_t) + \pi_t) + U_{zz}[b_{t+1} - \Omega_{t+1}\tau w_{t+1}f'(n_t)]R_{t+1}, \quad (\text{B.1})$$

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

$$V_{bn} = V_{nb} = U_{cc}((1-\tau)w_tf'(n_t) + \pi_t) < 0, \quad (\text{B.1})$$

$$V_{bs} = V_{sb} = U_{cc} < 0, \quad (\text{B.1})$$

$$V_{bb} = U_{cc} + n_{t-1}^2 U_{zz} < 0. \quad (\text{B.1})$$

The second-order conditions for a maximum of problem (4) are satisfied since V_{nn} is negative and the following conditions hold true:

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

$$\begin{aligned} \begin{vmatrix} V_{nn} & V_{ns} & V_{nb} \\ V_{sn} & V_{ss} & V_{sb} \\ V_{bn} & V_{bs} & V_{bb} \end{vmatrix} = & [U_{nn} - U_c(1-\tau)w_tf''(n_t) - U_z\Omega_{t+1}\tau w_{t+1}f''(n_t)] \\ & \cdot (R_{t+1}^2 U_{cc} U_{zz} + n_{t-1}^2 U_{zz} (U_{cc} + U_{zz}R_{t+1}^2)) \\ & + U_{cc}U_{zz}U_{zz}n_{t-1}^2 \\ & \cdot [R_{t+1}((1-\tau)w_tf'(n_t) + \pi_t) - (B_{t+1} - \Omega_{t+1}\tau w_{t+1}f'(n_t))]^2 \\ < 0. \end{aligned} \quad (\text{B.1})$$

This demonstrates that the objective function $V(n_t, s_t, b_t)$ is strictly concave in the decision variables.

B.2 The savings effect of the Bismarckian PAYG pension system

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

$$\frac{\partial s}{\partial \tau} = - \frac{\begin{vmatrix} V_{nn} & V_{n\tau} & V_{nb} \\ V_{sn} & V_{s\tau} & V_{sb} \\ V_{bn} & V_{b\tau} & V_{bb} \end{vmatrix}}{\begin{vmatrix} V_{nn} & V_{ns} & V_{nb} \\ V_{sn} & V_{ss} & V_{sb} \\ V_{bn} & V_{bs} & V_{bb} \end{vmatrix}} \quad (\text{B.1})$$

With the negative denominator savings decrease with a higher contribution rate to the PAYG system if the numerator is negative.

In the case of the Bismarckian pension system the numerator is given by

$$\begin{aligned} & \begin{vmatrix} V_{nn} & V_{n\tau} & V_{nb} \\ V_{sn} & V_{s\tau} & V_{sb} \\ V_{bn} & V_{b\tau} & V_{bb} \end{vmatrix} \\ &= w_t(1 - f(n_t)) (U_{n_t n_t} - U_{c_t}(1 - \tau)w_t f''(n_t) - U_{z_{t+1}}\Omega_{t+1}\tau w_{t+1} f''(n_t)) \\ & \quad [n_{t-1}^2 U_{c_t c_t} U_{z_t z_t} + \Omega_{t+1} R_{t+1} (U_{c_t c_t} U_{z_{t+1} z_{t+1}} + n_{t-1}^2 U_{z_t z_t} U_{z_{t+1} z_{t+1}})] \\ & \quad - U_{z_{t+1}} w_t f'(n_t) (R_{t+1} - \Omega_{t+1}) [U_{c_t c_t} U_{z_t z_t} n_{t-1}^2 ((1 - \tau)w_t f'(n_t) + \pi_t) \\ & \quad + (U_{c_t c_t} U_{z_{t+1} z_{t+1}} + n_{t-1}^2 U_{z_t z_t} U_{z_{t+1} z_{t+1}}) R_{t+1} (b_{t+1} - \Omega_{t+1} \tau w_t f'(n_t))] \\ & \quad + U_{c_t c_t} U_{z_{t+1} z_{t+1}} U_{z_t z_t} w_t (1 - f(n_t)) n_{t-1}^2 \\ & \quad [R_{t+1} ((1 - \tau)w_t f'(n_t) + \pi_t) - (b_{t+1} - \Omega_{t+1} \tau w_t f'(n_t))] \\ & \quad [\Omega_{t+1} (w_t f'(n_t) + \pi_t) - b_{t+1}]. \end{aligned} \quad (\text{B.1})$$

Since the price of a child is positive, $R_{t+1}((1 - \tau)w_t f'(n_t) + \pi_t) - (b_{t+1} - \Omega_{t+1} \tau w_t f'(n_t)) > 0$, the numerator is negative if the following condition for the intra-family transfer b_{t+1} holds: $\tau w_t f'(n_t) < \frac{b_{t+1}}{\Omega_{t+1}} < w_t f'(n_t) + \pi_t$. If this condition holds, savings decrease with a higher contribution rate in the Bismarckian system. The condition is equivalent to: $-\frac{\partial p_{t+1}^{BIS}}{\partial n_t} < b_{t+1} < \Omega_{t+1}((1 - \tau)w_t f'(n_t) + \pi_t) - \frac{\partial p_{t+1}^{BIS}}{\partial n_t}$. Note that the Bismarckian pension decreases in the number of children because the pension is proportional to contributions and income which decreases with more children. Thus, we have according to (9): $\frac{\partial p_{t+1}^{BIS}}{\partial n_t} < 0$. We can rewrite the condition for a negative savings effect as $\Omega_{t+1}((1 - \tau)w_t f'(n_t) + \pi_t) > b_{t+1} + \frac{\partial p_{t+1}^{BIS}}{\partial n_t} > 0$ which can be interpreted as follows.

Assume that a higher contribution rate reduces the number of children, i.e. the income effect is larger than the price effect. The second part of the inequality condition means that the loss of intra-family transfer due to fewer children in the second period is higher than the gain of a larger Bismarckian pension. Thus having fewer children reduces income and decreases consumption in the second period. Then the first part of the condition implies that the discounted reduction of (opportunity) costs for children

in the first period is higher than the loss of income in the second period due to fewer children. Hence, a lower number of children increases income and consumption in the first period by more than it reduces consumption in the second period. This implies that the parents react with lower savings in order to re-establish their preferred consumption profile and compensate the negative effect of the contribution rate on the number of children. If saved costs of fewer children in the first period are higher than the income loss in the second period a lower number of children induces lower savings.

Proposition 2 *Savings effect* *The introduction or expansion of the PAYG system reduces savings if the lower number of children raises income in the first period to a larger extent than it lowers income in the second period.*

B.3 The effect of a Bismarckian PAYG pension system on the intra-family transfer

The effect of a higher contribution rate on the intra-family transfer is given by:

$$\frac{\partial b_t}{\partial \tau} = - \frac{\begin{vmatrix} V_{nn} & V_{ns} & V_{n\tau} \\ V_{sn} & V_{ss} & V_{s\tau} \\ V_{bn} & V_{bs} & V_{b\tau} \end{vmatrix}}{\begin{vmatrix} V_{nn} & V_{ns} & V_{nb} \\ V_{sn} & V_{ss} & V_{sb} \\ V_{bn} & V_{bs} & V_{bb} \end{vmatrix}} \quad (\text{B.1})$$

The numerator can be calculated as:

$$\begin{vmatrix} V_{nn} & V_{ns} & V_{n\tau} \\ V_{sn} & V_{ss} & V_{s\tau} \\ V_{bn} & V_{bs} & V_{b\tau} \end{vmatrix} = (U_{n_t n_t} - U_{c_t}(1 - \tau)w_t f''(n_t) - U_{z_{t+1}}\Omega_{t+1}\tau w_{t+1} f''(n_t)) \\ w_t(1 - f(n_t))U_{c_t c_t}U_{z_{t+1} z_{t+1}}R_{t+1}(R_{t+1} - \Omega_{t+1}) \\ - U_{z_{t+1}}w_t f'(n_t)(R_{t+1} - \Omega_{t+1})R_{t+1}U_{c_t c_t}U_{z_{t+1} z_{t+1}} \\ [R_{t+1}((1 - \tau)w_t f'(n_t) + \pi_t) - (b_{t+1} - \Omega_{t+1}\tau w_t f'(n_t))] \quad (\text{B.1})$$

With $R_{t+1} > \Omega_{t+1}$ and a positive price of a child, $R_{t+1}((1 - \tau)w_t f'(n_t) + \pi_t) - (b_{t+1} - \Omega_{t+1}\tau w_t f'(n_t)) > 0$, the parents reduce the intra-family transfer if the PAYG system is extended: $\frac{\partial b_t}{\partial \tau} < 0$. The intuition for this result is that a higher contribution rate together with $R_{t+1} > \Omega_{t+1}$ reduces lifetime income since the contribution contains an implicit tax on wage income. The parents reduce their transfer to the grandparents in order to compensate for this loss in lifetime income. This reduces the old-age consumption of the grandparents.

Proposition 3 *Effect on intra-family transfer* *The introduction or expansion of the PAYG system induces the parents to reduce the intra-family transfer to the grandparents.*

B.4 Lack of capital markets

B.4.1 The fertility effect

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 (8). Again the first-order condition (5) holds. The implicit function theorem yields:

$$\frac{\partial n}{\partial \tau} = - \frac{V_{n\tau} V_{bb} - V_{nb} V_{b\tau}}{V_{nn} V_{bb} - V_{nb} V_{bn}},$$

and $V_{nn} < 0$ by (B.26) and

$$\begin{aligned} & V_{nn} V_{bb} - V_{nb} V_{bn} \\ = & [-U_c(1 - \tau)w_t f''(n_t) - U_{z_{t+1}} \Omega_{t+1} \tau w_{t+1} f''(n_t) \\ & + U_{z_{t+1} z_{t+1}} (b_{t+1} - \Omega_{t+1} \tau w_t f'(n_t))^2 + U_{nn}] \end{aligned} \quad (\text{B.1})$$

$$\begin{aligned} & (U_{cc} + n_{t-1}^2 U_{z_t z_t}) + n_{t-1}^2 U_{z_t z_t} U_{cc} ((1 - \tau)w_t f'(n_t) + \pi_t)^2 \\ > & 0 \end{aligned} \quad (\text{B.1})$$

satisfy the second-order condition. Hence, the fertility response with respect to an introduction or extension of the pension system is determined by the sign of $V_{n\tau} V_{bb} - V_{nb} V_{b\tau}$:

$$\begin{aligned} V_{n\tau} V_{bb} - V_{nb} V_{b\tau} &= (R_{t+1} - \Omega_{t+1})w_t f'(n_t)U_{z_{t+1}} (U_{cc} + n_{t-1}^2 U_{z_t z_t}) \\ &+ w_t(1 - f(n_t)) \\ &[U_{cc} U_{z_t z_t} n_{t-1}^2 ((1 - \tau)w_t f'(n_t) + \pi_t) \\ &+ U_{z_{t+1} z_{t+1}} (b_{t+1} - \Omega_{t+1} \tau w_t f'(n_t)) \Omega_{t+1} (U_{cc} + n_{t-1}^2 U_{z_t z_t})] \end{aligned} \quad (\text{B.1})$$

If $R_{t+1} > \Omega_{t+1}$ a higher contribution rate τ decreases the marginal price of a child which incites more children:

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

The second summand on the RHS is again the income effect. 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.

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.*

B.4.2 The effect on intra-family transfer

In this case the effect on intra-family transfer is given by:

$$\frac{\partial b_t}{\partial \tau} = -\frac{V_{nn}V_{b\tau} - V_{n\tau}V_{bn}}{V_{nn}V_{bb} - V_{nb}V_{bn}}.$$

With a positive denominator (B.1) the effect of intra-family transfer by a larger PAYG system depends on the sign of the numerator:

$$\begin{aligned} & V_{nn}V_{b\tau} - V_{n\tau}V_{bn} \\ = & \left[-U_c(1-\tau)w_t f''(n_t) - U_{z_{t+1}}\Omega_{t+1}\tau w_{t+1} f''(n_t) \right. \\ & \left. + U_{z_{t+1}z_{t+1}}(b_{t+1} - \Omega_{t+1}\tau w_t f'(n_t))^2 + U_{nn} \right] w_t(1-f(n_t))U_{cc} \\ & - \left[(R_{t+1} - \Omega_{t+1})w_t f'(n_t)U_{z_{t+1}} + U_{z_{t+1}z_{t+1}}(b_{t+1} - \Omega_{t+1}\tau w_t f'(n_t))\Omega_{t+1} \right] \\ & U_{cc}((1-\tau)w_t f'(n_t) + \pi_t). \end{aligned} \tag{B.1}$$

The sign is ambiguous, in particular if $b_{t+1} > \Omega_{t+1}\tau w_t f'(n_t)$.

Appendix C: Supplementary (online) table on a comparison of estimators

Table C.1 presents an OLS model in column (1), our baseline model in column (2) and a first differences estimator in column (3).

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 clustering at the province level) 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. However, the coefficients in the first differences model are not substantially different from our baseline model and as conjectured.

Table C.1: SENSITIVITY: ESTIMATION APPROACH

	OLS (1)	FE OLS (2)	FD OLS (3)
Share insured (%)	-.457 (.270)*	-.453 (.264)*	-.453 (.331)
Marriages (1/1000)	1.692 (1.465)	-.616 (1.450)	-.616 (1.814)
Old-age dep. ratio (%)	-1.015 (.494)**	-.621 (.659)	-.621 (.825)
Educated recruits (1/1000)	-.096 (.020)***	-.554 (.226)**	-.554 (.283)**
Share workers (%)	.064 (.036)*	-.017 (.078)	-.017 (.097)
Gender imb. ratio (%)	-.128 (.142)	-.174 (.089)*	-.174 (.112)
Protestants (per Catholic)	-.014 (.015)	-.009 (.028)	-.009 (.035)
Persons per building	.287 (.434)	-1.021 (.452)**	-1.021 (.565)*
Share revenues in cat. I (%)	.055 (.029)*	.018 (.053)	.018 (.066)
Share working in trade (%)	-.391 (.135)***	-.523 (.300)*	-.523 (.376)
Crop yield index	-.440 (.589)	1.547 (.722)**	1.547 (.903)*
Insured: health (%)	-.136 (.122)	.159 (.778)	.159 (.974)
Obs.	46	46	23

Significance level: *** : $p < 0.01$; ** : $p < 0.05$; * : $p < 0.1$. For variable definitions refer to table 1. Model in first differences estimated by first calculating the differences and then applying an OLS estimator to the differences; hence the number of observations is lower.