The Trade-off between Fertility and Education:
Evidence from before the Demographic Transition

Sascha O. Becker

Francesco Cinnirella

Ludger Woessmann

Online at http://www.economics.stir.ac.uk

The Trade-off between Fertility and Education:
Evidence from before the Demographic Transition[*]

Sascha O. Becker[†]
U Stirling, Ifo, CESifo, and IZA

Francesco Cinnirella[‡]
Ifo Institute and CESifo

Ludger Woessmann[§]
U Munich, Ifo, CESifo, and IZA

Abstract

The trade-off between child quantity and education is a crucial ingredient of unified growth models that explain the transition from Malthusian stagnation to modern growth. We present first evidence that such a trade-off indeed existed before the demographic transition, exploiting a unique census-based dataset of 334 Prussian counties in 1849. Estimating two separate instrumental-variable models that instrument education by landownership inequality and distance to Wittenberg and fertility by previous-generation fertility and sex-imbalance ratio, we find that causation between fertility and education runs both ways. Furthermore, education in 1849 predicts the fertility transition in 1880-1905.

This version: August 26, 2009

[†] Division of Economics, Stirling Management School, University of Stirling, Stirling FK9 4LA, United Kingdom; sascha.becker@stir.ac.uk.

[‡] Ifo Institute for Economic Research, Poschingerstr. 5, 81679 Munich, Germany; cinnirella@ifo.de.

[§] Department of Economics, University of Munich and Ifo Institute for Economic Research, Poschingerstr. 5, 81679 Munich, Germany; woessmann@ifo.de.

1. Introduction

The trade-off between the number of children and the human capital invested in each child—i.e., between fertility and education—is a crucial ingredient of unified growth theory, which models the transition from Malthusian stagnation to sustained growth in a unified framework (see Galor 2005a for a seminal review). In most of these models, the new technologies that emerged during the process of industrialization increased the demand for education, which in turn triggered the demographic transition at the end of the 19th century (Galor 2005b). But due to a dearth of historical data, empirical evidence that the so-called child quantity-quality (Q-Q) trade-off indeed existed before the demographic transition is missing so far. As Guinnane (2008, p. 30) points out, "we lack detailed empirical studies of the relationship between … education and the demand for children in the relevant period." Most existing evidence supporting the Q-Q trade-off is based on modern data, and the earliest analysis we are aware of stems from 1910 (Bleakley and Lange 2009), after the main phase of the fertility transition. However, the trade-off may well have emerged only during or after the demographic transition. Historians often question the thought that people in the pre-modern age had a strong tendency to follow such sophisticated "economic" thoughts as those underlying the Q-Q trade-off. What is more, it is well conceivable that the transition itself, which started at the end of the 19th century and brought fundamental changes to demographics, *created* the very trade-off that unified growth models take as an assumption for the transition to occur.

This paper aims to test the existence and strength of the child Q-Q trade-off before the demographic transition. Based on a unique micro-regional dataset of more than 330 Prussian counties stemming from full Population and Education Censuses conducted in 1849 by the Prussian Statistical Office, we investigate the link between education and fertility behavior for a period *before* the demographic transition. We find that a significant Q-Q trade-off indeed existed in Prussia long before the demographic transition, that the direction of causation runs both ways—from education to fertility and from fertility to education—and that differences in education in 1849 predict the cross-county variation in the fertility transition at the turn from the 19th to the 20th century.

Prussia constitutes an interesting case not only because of its unique data situation, but particularly because as early as 1763, Frederick the Great enacted the General Elementary School Regulation, which contained an explicit general invitation to attend school that is often

interpreted as quasi-compulsory primary education. Yet, this law was not strictly enforced, as evidenced by the fact that full enrolment in primary school was achieved only towards the end of the 19th century. Therefore, Prussian families were able to decide whether or not to send their children to school. In fact, across Prussian counties, primary-school enrolment rates still varied from 33% to 99% in 1849. As we will show below, a significant part of this substantial variation in education can be traced back to variation in landownership inequality and in religious denomination.

The Q-Q trade-off depicts a mutual association between education and fertility. Given a budget constraint, exogenous changes to the number of children will affect the optimal amount of education invested in the children, and exogenous changes to the amount of education will affect the optimal number of children. Thus, although most existing literature focuses on only one direction of causation, causation in the child Q-Q trade-off should actually run both ways. Our empirical analysis models *both* of these directions of causality.

Since the decisions how many children to have and how much to invest in each child are simultaneously taken, fertility and education are endogenous with respect to one another. This simultaneity renders empirical identification of causal links a major issue. Furthermore, fertility and education may be affected by common third factors such as income, urbanization, and life expectancy, so that identification may in addition be particularly vulnerable to bias from omitted variables. Estimating two separate equations (Wooldridge 2002, p. 222), we employ instrumental-variable (IV) strategies to overcome the endogeneity problems that adhere to the two variables.

To identify the effect of education on fertility, we employ exogenous variation in education that stems from cross-county differences in landownership inequality and distance to Wittenberg. As Galor, Moav, and Vollrath (2009) show, landownership inequality depresses human-capital accumulation because large landowners have little interest in public schooling due to low land-human capital complementarity. Distance to Wittenberg provides another source of exogenous variation in education across Prussia because it predicts Protestants' urge for literacy to read the Bible (Becker and Woessmann 2009). The identifying assumption is that neither landownership inequality nor distance to Wittenberg is directly related to fertility. Given one of these assumptions, we can test the second one by over-identification restrictions.

To identify the effect of fertility on education, we employ exogenous variation in fertility that stems from cross-county differences in previous-generation fertility and in the sex ratio (the relative numbers of women and men). Since fertility has a strong hereditary component (Rodgers et al. 2001), fertility in a county in 1816 is significantly associated with fertility in the same county in 1849, even after controlling for educational enrolment in 1816. Sex ratios in the adult population affect the tightness of the marriage market (Angrist 2002) and are thus associated with average fertility rates. Again, the identifying assumption is that neither previous-generation fertility nor the sex ratio is directly related to education (after controlling for previous-generation education), and we can apply over-identification tests to the two instruments.

We find evidence of significant effects of child quality on child quantity and vice versa. The results are robust to a rich set of controls for the level of development and to the incidence of mortality. Finally, we study the long-run impact of educational differences for the demographic transition using census data from the turn of the 19th to the 20th century. Our results suggest that the fertility transition in Prussia at the turn of the century was more pronounced in those counties with higher educational enrolment in 1849.

Our findings on the negative relationship between education and fertility before the fertility transition are consistent with the theoretical predictions of evolutionary models in the spirit of Galor and Moav (2002), who characterize the Q-Q trade-off as existing all along. In these models, the fraction of parents emphasizing quality of their offspring over quantity is low in early stages of development, but they ultimately gain an evolutionary advantage. Galor and Moav (2002, p. 1144) point out that careful empirical analyses of the relevance of child quality in the pre-industrial revolution era is not available so far, due to the lack of corresponding historical data. Our analysis aims to fill this gap.

The remainder of the paper is structured as follows. Section 2 places our contribution in the existing literature. Section 3 describes our new dataset and provides descriptive statistics. Section 4 develops our estimation strategy. Section 5 presents results on the child Q-Q trade-off before the demographic transition. Section 6 documents the long-run effects of investments in education in 1849 on the fertility transition in 1880-1905. Section 7 summarizes and concludes.

2. Related Literature

The industrial revolution was characterized by major developments in the economic and demographic environment. Between late 18th century and early 20th century, most European countries witnessed the transition from stagnation to growth. Furthermore, in later phases of the industrialization, while income increased, fertility decreased and gave rise to the demographic transition. Whereas much previous research saw these transitions as discontinuous changes, recently, much scholarly attention has been devoted to explain the transition from stagnation to growth within a unified growth theory.[1] The main contribution of unified growth theory is to combine the elements of a typical Malthusian economy—displaying decreasing returns to labor, high fertility rates, and a positive correlation between income and population size—with the factors that characterize developed economies with high income and low fertility. In establishing the micro-foundations of the transition between these two phases, scholars were faced with the problem of explaining how the Malthusian positive link between income and population, for some countries, turned negative towards the end of the 19th century. In the majority of unified growth models, investments in human capital provide the explanation.[2]

In their seminal contributions, Galor and Weil (1999, 2000) and Galor and Moav (2002) stress the importance of human capital in explaining the transition from stagnation to growth. Their models characterize the child Q-Q trade-off as one of the fundamental trade-offs that exist in nature. In the economics literature, the Q-Q trade-off has originally been modeled by Becker (1960), who noted that the Malthusian model of a positive link between income and fertility did not take into account the role of child quality. He argued that the elasticity of child quantity with respect to income is usually small compared to the elasticity of child quality with respect to income, implying that with rising income emphasis shifts towards child quality. In his model, the optimal choice of a child Q-Q combination depends on income, prices, and tastes. Holding

---

[1] See Galor and Weil (1999, 2000), Galor and Moav (2002), Hansen and Prescott (2002), Lucas (2002, ch. 5), Doepke (2004), Galor (2005b), and Clark (2007) for examples of leading contributions.

[2] Cervellati and Sunde (2005) provide a unified growth model where human capital is crucial not through its association with fertility, but through its association with life expectancy. For an example of a unified growth model in which human capital does not play a role, see Strulik and Weisdorf (2008).

income constant, a natural trade-off between quantity and quality arises to the extent that both dimensions require time and money resources.[3]

Econometric studies of the Q-Q trade-off nearly uniformly use modern data (see Schultz 2008 for a detailed critical review). Some studies express the Q-Q trade-off as a model where child quality is a function of child quantity (e.g., Hanushek's 1992 study of the dependence of test scores on family size), while others express the function the other way round (e.g., Doepke's 2004 macro simulation of the fertility decline). For econometric identification, it is important to note that both quantity and quality of children are endogenous variables in the Q-Q model. Rosenzweig and Wolpin (1980) point out that the distinguishing feature of Q-Q models is that the shadow price of quality per child depends on the number of children, while the shadow price of quantity depends on the level of quality chosen. Since these shadow prices cannot be observed directly, the Q-Q trade-off can only be uncovered when exogenous variations in fertility are shown to affect education per child or when exogenous variations in (returns to) education are shown to affect fertility. In the sense that natural or quasi-natural variation in fertility or education is used, one can speak of a causal effect of family size on the quality of children or, vice versa, of a causal effect of child quality on child quantity.

The first paper exploiting exogenous variations in fertility to identify the effect of child quantity on child quality is Rosenzweig and Wolpin (1980), who analyze household data from India. They use multiple births to measure exogenous increases in family size and show the resulting decrease in child quality. Angrist, Lavy, and Schlosser (2005) use dummies for multiple births at second birth and a dummy for same-sex sibling pairs in families with two or more children as instrumental variables. The latter instrument makes use of parental preferences for a mixed sibling-sex composition. Their results, using Israeli data, suggest that exogenous increases in family size at second and higher births appear to have little effect on education levels of first and second born children. In a similar vein, Black, Devereux, and Salvanes (2005) use a large sample of administrative records in Norway to look at the effects of multiple births on schooling and earnings. Controlling for birth order, the occurrence of a multiple birth has no effect on these outcomes. Qian (2009) uses regional and time variation in China's one-child policy, as well as multiple births, to estimate the effects of family size on school enrolment in China. Perhaps

---

[3] See Becker and Lewis (1973), Willis (1973), and Moav (2005) for extensions of the micro-foundations of the Q-Q model.

surprisingly, her estimates suggest that relaxation of the one-child policy increased the enrolment rates of first-born children, a finding she explains by differential parental investment in first-born and later-born children. Along the same lines, Rosenzweig and Zhang (2009), re-examining China's one-child policy, cast doubt on studies using multiple births, in general.

The opposite direction of causation, from education to fertility, is analyzed as well. While a number of papers is concerned with the effect of mother's education on her fertility (e.g., Breierova and Duflo 2004), few studies use micro data to analyze the effect of exogenous changes in (returns to) education of children on the number of children born. Bleakley and Lange (2009) exploit the introduction of a hookworm disease eradication policy in the American South which is argued to establish a shock to the price of quality because hookworm reduced the return to human capital investment, had a low case-fatality rate, and was hardly prevalent among adults. Hookworm eradication thus exogenously increased the returns to education and reduced the price of child quality. They show that the resulting increases in school attendance and literacy were accompanied by a significant fertility decline.

Bleakley and Lange (2009) is also the study going furthest back in time, to 1910, a period when the demographic transition was well under way. The Princeton European Fertility Project, which used 19th-century data to study the fertility transition (see Coale and Watkins 1986), largely ignored economic factors and did not address the child Q-Q trade-off. We are not aware of micro-econometric studies of the Q-Q trade-off before the 20th century.

3. Data

We address the child Q-Q trade-off in historical perspective using Prussian county data. In order to explore this relation, we analyze regional variation across more than 330 counties (*Kreise*). Prussia was one of the largest European countries and, with a population of about 24.6 million, accounted for 60% of the population of the German Empire founded in 1871. Over the 19th century, the Prussian Statistical Office collected a remarkable amount of information on vital statistics (e.g., age distribution by gender, marital status, and religion). Demographers have found county-level data for 19th-century Prussia to be a unique source of highest-quality data for analyses at a disaggregated level (see Galloway, Hammel, and Lee 1994).

Our main data sources are the Population, Education, and Occupation Censuses conducted in 1849. The Population Census contains information on age structure that allows us to infer

fertility and mortality. The Education Census contains detailed information on enrolment rates in public primary schools as well as on the number of schools. The Occupation Census contains the employment structure and reveals the stage of industrialization, i.e., employment in industry vs. agriculture.[4] These censuses allow us to construct a coherent database with a detailed amount of demographic and socio-economic information in 1849. To our knowledge, these data have not yet been used for micro-econometric analyses.

We also use data from the first Prussian Census in 1816 (see Becker and Woessmann 2008). The 1816 Census contains demographic and educational information that we use as instrumental variable and controls to identify exogenous variation in our variables of interest in 1849.

To measure fertility behavior, we use the child-woman ratio, defined as the number of children aged 0-5 per woman of child-bearing age (15-45).[5] Rather than using births, the number of children aged 0-5 captures surviving children, which we consider the relevant measure in the joint fertility-education decision (see Galor 2005b). Below, we ascertain that results are robust to controls for mortality differences, as measured by life expectancy at different ages and by infant mortality. We measure education as enrolment rates in public primary school in 1849, defined as the number of children enrolled in school divided by the number of children aged 6-14.[6]

The covariates used in the regression analysis are: (i) a proxy for the level of industrialization specified as the share of people employed in manufacturing, (ii) the share of people making their living of agriculture, (iii) the share of urban residents (defined as the share of the population living in towns with city rights), (iv) population density, (v) the share of married women, (vi) the share of Protestants, (vi) a geographical dummy variable for the provinces with a majority of Slavic-language people (*Poland*),[7] (vii) life expectancy at ages 0 and 5 (calculated from age-specific population and mortality data using the standard approach as depicted, e.g., in Jayachandran and Lleras-Muney 2009),[8] (viii) the number of public primary schools per 100 children in school age (6-14), and (ix) a measure for temporary male migration (constructed as

---

[4] Unfortunately, the occupation data do not distinguish between male and female labor-force participation.

[5] In 1816, the numerator is the number of children aged 0-7.

[6] Consistent with the definition of mandatory schooling at the time, we consider as primary schools both elementary schools (*Elementarschulen*) and middle schools (*Mittelschulen*).

[7] These are the regions of East Prussia, West Prussia, Poznan, and Silesia.

[8] Specifically, our life expectancy measure is calibrated on male mortality tables as they are more detailed in the Prussian data than female mortality tables.

the difference between the total number of married males and the total number of married females, divided by the total number of married females).

Table 1 provides descriptive statistics of the variables. The first thing to note is the high level of school enrolment. In 1849, already 80% of the children aged 6-14 were enrolled in primary school on average. For comparison, in 1816, 58% of the children aged 6-14 were enrolled in primary school. The enrolment data reveals that the quasi-compulsory primary-school attendance regulation enacted in Prussia in the 18th century was not strictly enforced. One reason for this was a lack of resources which were directly linked to the taxes levied on landlords (Cubberley 1920; Galor and Moav 2006), which may also explain the cross-county variation in education. The minimum and maximum enrolment rates are 3% and 95%, respectively, in 1816 and 33% and 99%, respectively, in 1849. A direct comparison with other European countries is not possible due to lack of statistics for the same years. Yet, Galor and Moav (2006, p. 89) report that in England "the proportion of children aged 5-14 in primary schools increased from 11% in 1855 to 25% in 1870". In France, the proportion of children aged 5-14 enrolled in primary school increased from 51% to 86% in the period 1850-1901 (Flora, Kraus, and Pfenning 1983).

Figure 1 shows the geographical distribution of school enrolment rate and child-woman ratio in 1849. The relatively industrialized western area (Rhineland) shows high levels of enrolment rates. The agricultural East and West Prussia shows comparatively high levels of fertility.

In the final part of our analysis, we study the long-run impact of the accumulation of human capital in the mid-19th century on the demographic transition at the turn to the 20th century. Data from a number of Prussian Population Censuses in the late 19th and early 20th century allow us to compute changes in crude birth rates between 1880 and 1905 and in marital fertility rates between 1890 and 1905, which we use to analyze the role played by investments in education for the demographic transition. The crude birth rate is measured as the number of births per 1000 people per year. The marital fertility rate is measured as the number of births per 1000 married women aged 15-49. Appendix A describes all our data sources in greater detail.

4. Empirical Model

We aim to assess whether there is indeed a trade-off between fertility and education before the demographic transition, a key ingredient of unified growth theory. In doing so, we distinguish between the two directions of causation, where fertility is a function of education and

education is a function of fertility. Following Wooldridge (2002, p. 222), we estimate the two equations separately because of the greater robustness of single-equation methods:

$$fertility_i = \alpha \cdot education_i + \mathbf{X}_{i1}\delta_1 + e_{i1} \qquad (1)$$

$$education_i = \beta \cdot fertility_i + \mathbf{X}_{i2}\delta_2 + e_{i2} \qquad (2)$$

where the variable *education* is the enrolment rate in public primary schools and *fertility* is the child-woman ratio for county i, both measured in 1849. The coefficients α and β are our parameters of interest. The $\mathbf{X}$'s are vectors of control variables as described above.

As pointed out above, fertility behavior and decisions about children's education are taken simultaneously, so that simple ordinary least squares (OLS) estimates of equations (1) and (2) are subject to endogeneity bias and cannot depict causal effects. To the extent that there is a reinforcing feedback loop between fertility and education, OLS estimates of α and β will be biased downwards, towards (in absolute terms) larger negative estimates. The fertility-education association may be additionally affected by omitted variable bias. To the extent that third factors that affect both fertility and education, such as income, urbanization, and life expectancy, are not perfectly measured so that the $\mathbf{X}$ vectors are not fully specified, the error terms of OLS models will be correlated with the dependent variables. A pure income effect would make such omitted factors positively related to both fertility and education, so that the omitted variable bias is likely to bias OLS estimates of α and β upwards, towards less negative estimates. To address the issues of simultaneity and omitted variables, we employ instrumental-variable (IV) methods, separately in the two equations.

In the fertility equation (1), we use landownership inequality and distance to Wittenberg as instruments for primary education to identify the effect of education on fertility. The idea of using landownership inequality builds on Galor, Moav, and Vollrath (2009), who present a theoretical model where inequality in the distribution of landownership negatively affects the implementation of human-capital-promoting institutions. This is due to the fact that large landowners would not benefit from the accumulation of human capital given the low complementarity between land and human capital. Galor, Moav, and Vollrath (2009) also provide empirical evidence for the United States in the 20th century which corroborates this prediction. For their empirical evidence, they use the share of land held by large landowners.

Similarly, we have information on the number and size of estates[9] which allows us to construct an index of landownership inequality. The index is constructed as the ratio of the largest land holdings (greater than 600 *Morgen*) over the total number of land holdings. Results are perfectly robust to using the two largest categories (greater than 300 *Morgen*) in the numerator.[10]

We also use distance to Wittenberg as a second instrument for primary education. Becker and Woessmann (2009) observe that at the times of Martin Luther, Protestantism in Prussia had a tendency to spread in circles around Wittenberg, where Luther preached that every Christian should be able to read the Bible. They show that as a consequence, distance to Wittenberg gives rise to a decreasing prevalence of education in Prussia, and that Protestantism is unlikely to have had substantial economic effects besides its indirect effect through education. We follow their approach and use distance to Wittenberg, in addition to landownership inequality, to identify exogenous variation in Prussian enrolment rates.

The exclusion restrictions of the two instruments are that landownership inequality and distance to Wittenberg are not directly related to fertility. Becker and Woessmann (2009) show that distance to Wittenberg is orthogonal to several measures of pre-Lutheran economic and educational development, which suggests that it is unrelated to possible important pre-existing correlates of fertility. Independent of its effect through education, landownership inequality—if anything—should likely depress fertility, to the extent that the average family has less ownership to support bigger families. Thus, the IV estimates based on these two instruments may well be viewed as lower bounds of the true effect of education on fertility. As long as the exclusion restrictions are accepted for one of the two instruments, over-identification tests allow us to test the validity of the second instrument.

In the education equation (2), we use fertility in the previous generation (namely in 1816) and the sex-imbalance ratio as instruments for fertility in 1849. The use of fertility in the previous generation as an instrument is based on the finding that fertility has a significant hereditary component (Rodgers et al. 2001). In order for the exclusion restriction to hold, there

---

[9] The categories are: less than 5 Morgen, 5-30 Morgen, 30-300 Morgen, 300-600 Morgen, and greater than 600 Morgen. 1 Morgen is ca. 0.25 hectars.

[10] We also have information on the extension of *agricultural* land. Yet, by using this as denominator of our index for landownership inequality, we lose information on 15 counties. However, results are robust to the use of this alternative denominator. Landownership inequality data were missing for two counties in 1849. We imputed the data based on landownership inequality data available in 1858. Given the high correlation of 0.96 of the 1858 data with the 1849 data, we regressed the 1849 data on the 1858 data and predicted the two missing 1849 values.

should be no direct effect of fertility in 1816 on school enrolment in 1849. Since fertility behavior of the previous generation might have influenced educational choices of the previous generation, we control for school enrolment in 1816 in equation (2).

The sex-imbalance ratio in 1849 serves as an additional instrument for fertility behavior in 1849. The sex ratio constitutes a measure of marriage market tightness affecting marriage rates and fertility (see Angrist 2002). It is defined as the number of males aged 15-45 over the number of females aged 15-45. The identifying assumption for this instrument is that the sex ratio is exogenously determined by differential birth and death rates and that it affects fertility behavior only through its influence on the probability of finding a mate, but is otherwise unrelated to education. A potential threat to this assumption comes from gender-specific migration rates, a concern which we address below.

5. Education versus Fertility before the Demographic Transition

5.1 The Descriptive Association between Education and Fertility

In Table 2, we present standard OLS models in order to show the trade-off between education and fertility descriptively. The dependent variable is the child-woman ratio in 1849. The bivariate regression (column 1) shows that, indeed, education and fertility are significantly negatively correlated. In the subsequent columns, we progressively introduce control variables. The occupational structure and urbanization are strongly associated with fertility. As already shown in the literature (Galloway, Lee, and Hammel 1998), in urban and more densely populated environments fertility rates tend to be lower, whereas in counties with a larger share of employment in manufacturing, fertility rates are higher.[11] Controlling for the share of married women increases the explanatory power of the model and the estimated coefficient on education. Fertility is lower in Protestant areas as well as in counties with a majority of Slavic population. As expected, life expectancy is negatively associated with fertility.

The education coefficient remains statistically significantly negative throughout the different specifications. Furthermore, the association between education and fertility is economically non-trivial. Considering the richest specification in column 6, an increase in the enrolment rate by 10

---

[11] When adding an interaction term between urbanization and education, it is statistically significantly negative, suggesting that the Q-Q trade-off is somewhat stronger in urban counties.

percentage points is associated with a decrease in the child-woman ratio by about 1.8 children per 100 women in child-bearing age. Yet, given the simultaneity of the decisions about the quantity and quality of children and possible omitted variable biases, these coefficients do not have a causal interpretation.

5.2 Instrumental-Variable Estimates of the Effect of Education on Fertility

In Table 3, we present estimates of the causal effect of education on fertility. School enrolment rates in 1849 are instrumented by landownership inequality and distance to Wittenberg. Columns 1-3 show the first-stage estimates, where the instruments first enter separately (columns 1 and 2) and then jointly (column 3). The first-stage estimates suggest that landownership inequality and distance to Wittenberg are strongly associated with education, as mirrored by the large values of the partial F-statistics of the instruments in the first stage. When both instruments are used jointly, the Sargan-Hansen test of over-identification restrictions does not reject the hypothesis that, as long as one instrument is valid, the other one is as well.

Columns 4-6 present our main results, namely the corresponding second-stage estimates. School enrolment rates have a statistically significant and large negative impact on fertility. The IV estimates are notably larger than the OLS estimates of Table 2, suggesting that the dominating bias of the latter stems from omitted variables that drive them towards zero.[12] As suggested above, such bias in the OLS estimates could stem from imperfection of the income measures which gives rise to a positive association between fertility and education. In the IV models, an increase in the enrolment rate by 10 percentage points causes a decline in the child-woman ratio by about 5 children per 100 women in child-bearing age. Thus, education had a large and significant effect on fertility in Prussia already before the demographic transition.[13]

Table 4 presents several robustness checks that show that the IV coefficients for education are robust to different model specifications. To isolate the effect of parental preferences for sending their children to school from the availability of schools, column 1 adds the number of

---

[12] In addition, the bias towards zero may stem from measurement error in the enrolment data.

[13] While our result is consistent with the child Q-Q model that models a *contemporary* trade-off between education and fertility, it is in principle also consistent with models that stress the link between *parental* education and their fertility. However, when we add our measure of school enrolment in 1816 as a proxy for education of the previous generation to the models of Table 3 (not shown), this does not enter significantly and does not change the qualitative results, supporting a child Q-Q model interpretation.

schools per 100 children in school age to account for differences in school supply. This does not change our coefficient of interest.

Households' decisions to invest in children's education may have been affected by differences in mortality and thus in the expected length of productive life, which affects the potential returns to education and, indirectly, the desirable number of surviving children (e.g., Cervellati and Sunde 2005, 2009; Soares 2005; Jayachandran and Lleras-Muney 2009). Thus, columns 2 and 3 control for life expectancy at ages 0 and 5, respectively. Life expectancy at age 0 mirrors cross-county differences in mortality rates across all age groups, including infant mortality. Life expectancy at age 5 is expected to mirror mortality differences beyond infant mortality. In both cases, the coefficient on education remains statistically significant and of similar magnitude. The same holds when life expectancy is measured at age 25 and when survival rates at different ages are used as controls (not shown). An additional concern is that our measure of fertility, the child-woman ratio, which has the number of children aged 0-5 as its numerator, may be affected by infant mortality patterns. However, controlling directly for the infant mortality rate also does not affect the reported results (not shown).

As another robustness check, column 4 estimates the effect of school enrolment on the marital fertility rate in 1849, which considers married women rather than all women in child-bearing age in the denominator.[14] Using the marital fertility rate, the coefficient is still negative and highly significant: An increase in school enrolment by 10 percentage points is associated with a decrease in the marital fertility rate by about 7 children per 100 married women.

As for the control variables in the model, high levels of industrial employment are associated with high levels of fertility, which may capture a positive fertility effect of the income level or of the higher demand for child labor due to industrialization. Fertility rates are lower in urban areas. The share of married women is strongly associated with fertility, in line with the fact that children were predominantly born in wedlock (Hajnal 1965).

5.3 Instrumental-Variable Estimates of the Effect of Fertility on Education

Table 5 investigates the opposite direction of the Q-Q trade-off, from fertility to education. We estimate equation (2) instrumenting the child-woman ratio in 1849 by the child-woman ratio

---

[14] Unfortunately, the number of married women is not specified by age structure.

in 1816 and the sex-imbalance ratio in 1849.[15] The child-woman ratio in 1816 is a measure of the fertility behavior of the previous generation. To exclude the possibility that previous-generation fertility may have affected current education by its association with previous-generation education, the specifications that use previous-generation fertility as an instrument control for school enrolment in 1816. In the first stage, the child-woman ratio of the previous generation is significantly positively associated with fertility in 1849 (column 1), though the partial *F*-statistic stays somewhat below the conventional threshold of 10. This is partly due to the fact that controlling for school enrolment in 1816 picks up part of the variation in previous-generation fertility (while at the same time increasing the explained variance in the second stage). The sex-imbalance ratio in column 2 is significantly associated with our measure of fertility at a comforting partial *F*-statistic. When using the two instruments simultaneously, the Sargan-Hansen test does not reject the over-identification restrictions.

In the second stage, the IV coefficients have the expected negative sign and are statistically significant. In terms of magnitude, the IV estimate using both instruments implies that an increase of the child-woman ratio by 1 child per 10 women is associated with a decrease of the enrolment rate by about 4 percentage points.

Table 6 presents a set of robustness checks which confirm our results. Controlling for the number of schools per 100 children in school age (column 1) in order to focus on the effect of parental preferences given the availability of schools does not alter the results. As one would expect, the more schools are available—i.e., the higher the school density and the easier it is to send children to school—the higher enrolment. Again, controlling for life expectancy (column 2) does not alter the effect of fertility on education.

The sex-ratio instrument may be affected by gender-specific migration patterns, which might introduce endogeneity. For example, male migrant workers with a specific education pattern might migrate to work in factories in other counties to be able to afford a family at home. To exclude bias from such migration, we add a measure of temporary male migration proxied by the share of married men currently not in the county as a control variable. Again, results are qualitatively unaffected (column 3). The same is true when the number of migrants per capita (available for 1862) is used as an alternative measure of migration (not shown).

---

[15] As county borders changed after 1816, standard errors are clusterd at the level of the 280 units of observation in the 1816 data.

The analysis conducted to this point suggests that causality between education and fertility runs in both directions. These results corroborate the thesis put forward by unified growth theory that there is a child Q-Q trade-off even before the demographic transition.

6. Human Capital and the Demographic Transition

Unified growth theory and recent empirical studies (Galor 2005a; Galor and Moav 2006; Goldin and Katz 2008) suggest that during the second phase of the industrial revolution—which corresponds roughly to the second half of the 19th century in Prussia—returns to education increased due to complementarities between technological change and human capital, and households responded by investing more in children's education and by limiting fertility. Following these theoretical predictions, in this section we investigate the long-run effect of investments in human capital on the fertility transition by estimating the predictive power of school enrolment in 1849 for the decline in the crude birth rate (CBR) and in the marital fertility rate (MFR) during the Prussian demographic transition at the turn to the 20th century.[16]

An extensive literature focuses on explaining the determinants of the demographic transition. For instance, the Princeton European Fertility Project (EFP) of the 1960s/70s aimed at studying the fertility patterns of most Western European countries (Coale and Watkins 1986). In particular, Knodel (1974) documented the demographic transition in Germany through an analysis of family reconstitutions. The EFP concluded that the spread of new moral and cultural norms together with birth control technology were responsible for the fertility decline in Europe. This view has been strongly criticized. Numerous studies have successively argued for a significant role played by economic factors in triggering the fertility transition (e.g., Galloway, Hammel, and Lee 1994; Brown and Guinnane 2007; Bleakley and Lange 2009).

Similarly, we suggest that the economic factor of the child Q-Q trade-off and thus the accumulation of human capital may play a notable role in explaining the fertility transition in Prussia. Figure 2 shows the geographic distribution of school enrolment rates in 1849 alongside the change in MFR in 1890-1905. The highest enrolment rates are concentrated in the central counties of Prussia (left panel); the same regions later tend to experience the steepest decline in

---

[16] Another interesting question to look at would be whether *changes* in education at the end of the 19th century are related to the fertility decline. Unfortunately, there is no Prussian school enrolment data similar to 1849 at the end of the 19th century which would allow such an analysis.

MFR (right panel). In order to test the hypothesis that human-capital accumulation was important for the fertility transition, we estimate the predictive power of school enrolment in 1849 for the decline in the CBR and MFR during the Prussian demographic transition:

$$rate\ of\ change(fertility)_{i,1880/90-1905} = \gamma \cdot education_{i,1849} + \mathbf{X}_{i3}\delta_3 + e_{i3} \qquad (3)$$

where the dependent variable is, in a first specification, the percentage change in the CBR between 1880 and 1905, and, in a second specification, the percentage change in the MFR between 1890 and 1905.[17]

Given the time lag between the dependent variable measured in 1880-1905 and the explanatory variable measured in 1849, there is no direct simultaneity in this specification. To address remaining issues raised by possible omitted variables that may create simultaneity bias and by the intergenerational correlation of fertility, we add the 1849 level of fertility to the control vector $\mathbf{X}_3$ in specification (3), in addition to the previous controls. The time span of about 50 years between 1849 and the turn of the century amounts to a difference of 1-2 generations. Estimated effects are thus likely to measure how education of the generation of parents and grandparents influenced fertility behavior of children and grandchildren.

Results are reported for the CBR in Table 7 and for the MFR in Table 8. The estimates show that those counties in which the investments in education were higher already in 1849 experienced a steeper fertility decline in terms of both CBR and MFR. Given the already high levels of primary school enrolment rates in 1849, this suggests that the demographic transition was faster in those counties where the demand for post-primary education was subsequently stronger. The estimates are particularly robust and significant for the MFR, but only marginally significant in the more extensive specifications of the CBR model.

Regarding the interpretation of the coefficients, if in 1849 the school enrolment rate would have been 10 percentage points higher, the MFR would have declined by an *additional* 1.1% (column 4 of Table 8). These results are robust to the choice of different truncation years for the dependent variables.[18] In terms of explanatory power, in the richer models investments in human capital explain 57% of the variation in changes in the MFR.

---

[17] The fact that we use slightly different time intervals for the two dependent variables is exclusively determined by data availability. The dependent variable is not truncated at zero, so that both increases and decreases in fertility rates in the specified time period are considered.

[18] Data on CBR are available until 1914.

7. Conclusion

Understanding the dynamics which allowed developed economies to escape the Malthusian world and enter a path of rapid economic growth is important in its own right. It becomes even more relevant as many developing countries are nowadays experiencing a demographic transition similar to that faced by western economies at the turn from the 19th to the 20th century. Much effort has been spent on modeling such dynamics. Most of the theoretical literature assigns a major role to demography and its interaction with human capital. Yet, historical econometric evidence which tests and quantifies these theoretical predictions is rather scarce.

Our evidence uses a unique dataset of more than 330 county-level observations in Prussia in the mid-19th century. We find that, indeed, a trade-off between quantity and quality of children was already in place in 1849, long before the onset of the demographic transition. Our results support the core assumption of unified growth theory that the child quantity-quality trade-off existed *before* the demographic transition. Identification in our models builds on instrumental-variable strategies, exploiting exogenous variation in fertility and education to establish the two-way causality between the components of the child Q-Q trade-off. We use two instruments for fertility in 1849, the sex ratio of the adult population in 1849 and fertility levels of the previous generation to capture the intergenerational component of fertility. Our two instruments for education in 1849 are landownership inequality and distance to Wittenberg. The results support a mutual causation between education and fertility, in line with the standard trade-off concept. The findings are robust when accounting for differences in mortality patterns.

Furthermore, we show that education levels in 1849 are a strong predictor of the strength of the fertility transition between 1880 and 1905. Counties with higher school enrolment rates in 1849 show a steeper fertility decline, both in terms of crude birth rates and marital fertility rates, at the turn from the 19th to the 20th century. Thus, in line with unified growth theory, human capital played a significant role in the fertility transition and as such in the transition from Malthusian stagnation to modern growth.

References

Angrist, Joshua D. (2002). How Do Sex Ratios Affect Marriage and Labor Markets? Evidence from America's Second Generation. *Quarterly Journal of Economics* 117 (3): 997-1038.

Angrist, Joshua D., Victor Lavy, Analia Schlosser (2005). New Evidence on the Causal Link Between the Quantity and Quality of Children. NBER Working Paper 11835. Cambridge, MA: National Bureau of Economic Research.

Becker, Gary, S. (1960). An Economic Analysis of Fertility. In: Gary S. Becker (ed.), *Demographic and Economic Change in Developed Countries*, pp. 209-240. Princeton, NJ: Princeton University Press.

Becker, Gary S., H. Gregg Lewis (1973). On the Interaction between the Quantity and Quality of Children. *Journal of Political Economy* 81 (2, part 2): S279-S288.

Becker, Sascha O., Ludger Woessmann (2008). Luther and the Girls: Religious Denomination and the Female Education Gap in 19th Century Prussia. *Scandinavian Journal of Economics* 110 (4): 777-805.

Becker, Sascha O., Ludger Woessmann (2009). Was Weber Wrong? A Human Capital Theory of Protestant Economic History. *Quarterly Journal of Economics* 124 (2): 531-596.

Black, Sandra E., Paul J. Devereux, Kjell G. Salvanes (2005). The More the Merrier? The Effect of Family Size and Birth Order on Children's Education. *Quarterly Journal of Economics* 120 (2): 669-700.

Bleakley, Hoyt, Fabian Lange (2009). Chronic Disease Burden and the Interaction of Education, Fertility, and Growth. *Review of Economics and Statistics* 91 (1): 52-65.

Breierova, Lucia, Esther Duflo (2004). The Impact of Education on Fertility and Child Mortality: Do Fathers Really Matter Less Than Mothers? NBER Working Paper 10513. Cambridge, MA: National Bureau of Economic Research.

Brown, John C., Timothy W. Guinnane (2007). Regions and Time in the European Fertility Transition: Problems in the Princeton Project's Statistical Methodology. *Economic History Review* 60 (3): 574-595.

Cervellati, Matteo, Uwe Sunde (2005). Human Capital Formation, Life Expectancy, and the Process of Development. *American Economic Review* 95 (5): 1653-1672.

Cervellati, Matteo, Uwe Sunde (2009). Life Expectancy and Economic Growth: The Role of the Demographic Transition. IZA Discussion Paper 4160. Bonn: Institute for the Study of Labor.

Clark, Gregory (2007). *A Farewell to Alms: A Brief Economic History of the World*. Princeton, NJ: Princeton University Press.

Coale, Ansley J., Susan Cotts Watkins (1986). *The Decline of Fertility in Europe*. Princeton, NJ: Princeton University Press.

Cubberley, Ellwood P. (1920). *The History of Education*. Cambridge, UK: Cambridge University Press.

Doepke, Matthias (2004). Accounting for Fertility Decline during the Transition to Growth. *Journal of Economic Growth* 9 (3): 347-383.

Flora, Peter, Franz Kraus, Winfried Pfenning (1983). *State, Economy and Society in Western Europe 1815-1975*. Chicago, IL: St. James Press.

Galloway, Patrick R., Eugene A. Hammel, Ronald D. Lee (1994). Fertility Decline in Prussia, 1875-1910: A Pooled Cross-Section Time Series Analysis. *Population Studies* 48 (1): 135-158.

Galloway, Patrick R., Ronald D. Lee, Eugene A. Hammel (1998). Urban versus Rural Fertility Decline in the Cities and Rural Districts of Prussia, 1875 to 1910. *European Journal of Population* 14 (3): 209-264.

Galor, Oded (2005a). From Stagnation to Growth: Unified Growth Theory. In: Philippe Aghion, Steven N. Durlauf (eds.), *Handbook of Economic Growth*, vol. 1A, pp. 171-293. Amsterdam: Elsevier.

Galor, Oded (2005b). The Demographic Transition and the Emergence of Sustained Economic Growth. *Journal of the European Economic Association* 3 (2-3): 494-504.

Galor, Oded, Omer Moav (2002). Natural Selection and the Origin of Economic Growth. *Quarterly Journal of Economics* 117 (4): 1133-1191.

Galor, Oded, Omer Moav (2006). Das Human-Kapital: A Theory of the Demise of the Class Structure. *Review of Economic Studies* 73 (1): 85-117.

Galor, Oded, Omer Moav, Dietrich Vollrath (2009). Inequality in Land Ownership, the Emergence of Human Capital Promoting Institutions, and the Great Divergence. *Review of Economic Studies* 76 (1): 143-179.

Galor, Oded, David N. Weil (1999). From Malthusian Stagnation to Modern Growth. *American Economic Review* 89 (2): 150-154.

Galor, Oded, David N. Weil (2000). Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond. *American Economic Review* 90 (4): 806-828.

Goldin, Claudia, Lawrence F. Katz (2008). *The Race between Education and Technology*. Cambridge, MA: Harvard University Press.

Guinnane, Timothy W. (2008). The Historical Fertility Transition and Theories of Long-Run Growth: A Guide for Economists. Mimeo, Yale University (August).

Hajnal, John (1965). European Marriage Patterns in Perspective. In: David V. Glass, David E.C. Eversley (eds.), *Population in History: Essays in Historical Demography*, pp. 101-143. London: Edward Arnold.

Hansen, Gary D., Edward C. Prescott (2002). Malthus to Solow. *American Economic Review* 92 (4): 1205-1217.

Hanushek, Eric A. (1992). The Trade-Off between Child Quantity and Quality. *Journal of Political Economy* 100 (1): 84-117.

Jayachandran, Seema, Adriana Lleras-Muney (2009). Life Expectancy and Human Capital Investments: Evidence from Maternal Mortality Declines. *Quarterly Journal of Economics* 124 (1): 349-397.

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

Lucas, Robert E. (2002). *Lectures on Economic Growth*. Cambridge, MA: Harvard University Press.

Moav, Omer (2005). Cheap Children and the Persistence of Poverty. *Economic Journal* 115 (500): 88-110.

Qian, Nancy (2009). Quantity-Quality and the One Child Policy: The Only-Child Disadvantage in School Enrollment in Rural China. NBER Working Paper 14973. Cambridge, MA: National Bureau of Economic Research.

Rodgers, Joseph L., Kimberly Hughes, Hans-Peter Kohler, Kaare Christensen, Debby Doughty, David C. Rowe, Warren B. Miller (2001). Genetic Influence Helps Explain Variation in Human Fertility: Evidence From Recent Behavioral and Molecular Genetic Studies. *Current Directions in Psychological Science* 10 (5): 184-188.

Rosenzweig, Mark R., Kenneth I. Wolpin (1980). Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment. *Econometrica* 48 (1): 227-240.

Rosenzweig, Mark R., Junsen Zhang (2009). Do Population Control Policies Induce More Human Capital Investment? Twins, Birthweight, and China's 'One Child' Policy. *Review of Economic Studies*: forthcoming.

Schultz, T. Paul (2008). Population Policies, Fertility, Women's Human Capital, and Child Quality. In: T. Paul Schultz, John Strauss (eds.), *Handbook of Development Economics*, vol. 4, pp. 3249-3303. Amsterdam: Elsevier.

Soares, Rodrigo R. (2005). Mortality Reductions, Educational Attainment, and Fertility Choice. *American Economic Review* 95 (3): 580-601.

Strulik, Holger, Jacob Weisdorf (2008). Population, Food, and Knowledge: A Simple Unified Growth Theory. *Journal of Economic Growth* 13 (3): 195-216.

Willis, Robert J. (1973). A New Approach to the Economic Theory of Fertility Behavior. *Journal of Political Economy* 81 (2, part 2): S14-S64.

Wooldridge, Jeffrey M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

Appendix A: County-Level Data for Prussia in the 19th Century

1849 Census

The source of the 1849 census data is the Prussian Statistical Office, which published the data in the period 1851-1855 under the title "Tabellen und amtliche Nachrichten über den Preussischen Staat für das Jahr 1849". The census of 1849 contains seven volumes. We use Vol. 1 for population data, Vol. 2 for education and mortality data, and Vol. 6a for factory data. All data are available for 334 counties.

The 1849 education data contain information on the number of schools and students for public elementary schools (*Öffentliche Elementarschulen*) and public middle schools for boys and girls (*Öffentliche Mittelschulen für Söhne und Töchter*). We combine enrolment in elementary and middle schools to obtain primary school enrolment, matching the fact that children at recommended school age (6-14 years) could either attend elementary schools or middle schools.

We construct measures of life expectancy at different ages by calculating age-specific mortality rates from population and mortality data, which is reported in age groups of varying size, usually encompassing five years.

The 1849 factory data contain information on the number of factories and workers for the Prussian counties. 119 types of factories are distinguished by the products fabricated. Our variable for industrialization in 1849 refers to the share of population working in textile, metal, and other factories. The textile sector includes factories for spinning, weaving, dyeing, and apparel. The metal sector includes processing of metals and production of metal products and machinery, as well as manufacture of stone and glass products. The other industrial sectors include such factories as those producing food, wood, paper, wax, and rubber.

1816 Census

1816 is the earliest year for which the Prussian Statistical Office, founded in 1805, collected detailed data at the county and municipality level. The 1816 data refer to 330 counties and 172 large and medium-size towns with detailed information on the age structure by gender. Due to a change of county borders after 1816, the 1816 data had to be re-grouped to match the county delineation in 1849, so that the ultimate number of units of observation of the 1816 data is 280.

The county data provide information on the number of public elementary schools (*Öffentliche Elementarschulen*), the only school type equally available in rural areas and towns at the time, and the number of students therein. In addition, the town data report on the number of schools and students in the following school types available only in towns: private elementary schools (*Privat-Elementarschulen*), public middle schools for boys and girls (*Öffentliche Bürger- oder Mittelschulen für Söhne und Töchter*), and private middle schools (*Private Bürger- oder Mittelschulen für Söhne und Töchter*). To capture all children at recommended school age, county and town enrolment data are aggregated to compute total primary enrolment. The source of the 1816 Population Census data is Alexander A. Mützell (1825), *Neues Topographisch-Statistich-Geographisches Wörterbuch des Preussischen Staats*, Vol. 6, Halle: Karl August Kümmel.[19]

Demographic Data 1880-1905

In order to generate the variables crude birth rate (CBR) and marital fertility rate (MFR), we use data on vital statistics and population censuses, respectively. Vital statistics are available on an annual basis for the period 1862-1914. For the MFR, the information on the number of married women aged 15-45 and the number of newborns are available only for the years 1890, 1895, 1900, and 1905. To construct the CBR variable, we use data from *Preussische Statistik*, Vol. 42 and Vol. 249. For the MFR variable, we use data contained in *Preussische Statistik*, Vol. 121a, Vol. 148a, Vol. 177, and Vol. 206a.

After 1849, some bigger counties were split in two (or more) separate counties; we aggregated the post-1849 data up to the 334 counties existing in 1849. To match the 1849 data, the analysis for the late 19th/early 20th century is also restricted to Prussia in its 1849 borders, even though Prussia had annexed several territories since.

---

[19] The 1816 school enrolment data were missing for the eleven counties of the district of Cologne. We imputed the data based on school enrolment data available in 1829 for all 59 counties of the Rhine Province. Given a correlation of 0.59 of the 1829 data with the 1816 data for the 48 counties with both datasets available, we regressed the 1816 data on the 1829 data and predicted the 1816 values for the eleven Cologne counties based on their 1829 values. Data on the age structure in 1816 were also missing for the eleven counties of the district of Cologne, which report population totals only. We used the share of the female age group 0-7 and 15-45 of the contiguous provinces (Arnsberg, Duesseldorf, Aachen, Koblenz) to impute the missing child-woman ratio of 1816 for the eleven Cologne counties.

Table 1. Summary statistics

	Mean	Std. dev.	Min	Max
Census 1849				
Child-woman ratio	0.64	0.08	0.35	0.84
School enrolment rate	0.80	0.12	0.33	0.99
Share in industry	0.03	0.03	0.01	0.32
Share in agriculture	0.43	0.17	0.00	0.85
Share urban	0.24	0.19	0.00	1.00
Population density (1000 people per km^2)	0.20	1.12	0.02	14.98
Share married women	0.70	0.06	0.43	0.85
Share Protestants	0.60	0.39	0.00	1.00
Poland	0.42	0.49	0.00	1.00
Life expectancy at age 0	35.48	7.53	12.19	49.48
Life expectancy at age 5	45.84	6.92	20.05	56.62
Schools per 100 children (6-14)	0.82	0.27	0.27	1.72
Temporary male migration	0.00	0.03	-0.04	0.52
Marital fertility rate	0.70	0.06	0.43	0.85
Landownership inequality	0.01	0.01	0.00	0.08
Distance to Wittenberg (in 100 km)	3.33	1.64	0.00	7.31
Sex-imbalance ratio	0.99	0.08	0.82	1.39
Census 1816				
Child-woman ratio	0.89	0.16	0.50	2.29
School enrolment rate	0.58	0.20	0.03	0.95
Demographic data 1880-1905				
Crude birth rate 1880-1905 (% change)	-0.10	0.10	-0.46	0.20
Marital fertility rate 1890-1905 (% change)	-0.08	0.10	-0.40	0.21

Source: Data for 334 counties from the Prussian Censuses 1816 and 1849 and demographic data for different years; see main text and Appendix A for details.

Notes: Child-woman ratio is the number of children aged 0-5 (0-7 in 1816) over the number of women aged 15-45. School enrolment rate is the share of children aged 6-14 enrolled in public primary schools.

Table 2. The association between education and fertility

Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)
			OLS			
			Child-woman ratio			
School enrolment	-0.080**	-0.075**	-0.159***	-0.148***	-0.199***	-0.174***
	(0.040)	(0.037)	(0.036)	(0.033)	(0.036)	(0.037)
Share in industry		0.431***	0.420***	0.394***	0.333***	0.341***
		(0.112)	(0.101)	(0.106)	(0.104)	(0.101)
Share in agriculture		0.091***	0.097***	0.106***	0.114***	0.116***
		(0.033)	(0.033)	(0.032)	(0.031)	(0.030)
Share urban		-0.069**	-0.025	0.017	0.003	-0.024
		(0.034)	(0.029)	(0.030)	(0.029)	(0.031)
Population density		-0.010**	-0.002	-0.004	-0.004	-0.004
		(0.004)	(0.004)	(0.003)	(0.004)	(0.003)
Share married women			0.561***	0.676***	0.677***	0.658***
			(0.085)	(0.093)	(0.088)	(0.084)
Share Protestants				-0.045***	-0.043***	-0.038***
				(0.009)	(0.009)	(0.010)
Poland					-0.031***	-0.045***
					(0.009)	(0.010)
Life expectancy at age 0						-0.002***
						(0.001)
Constant	0.702***	0.665***	0.325***	0.249***	0.303***	0.367***
	(0.033)	(0.040)	(0.060)	(0.063)	(0.066)	(0.069)
Observations	334	334	334	334	334	334
R^2	0.015	0.170	0.297	0.342	0.372	0.388

Source: County-level data from the Prussian Census 1849; see main text and Appendix A for details.

Notes: Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Child-woman ratio is the number of children aged 0-5 (0-7 in 1816) over the number of women aged 15-45. School enrolment rate is the share of children aged 6-14 enrolled in public primary schools.

Table 3. The effect of education on fertility

Dependent variable:	IV first stage School enrolment			IV second stage Child-woman ratio		
	(1)	(2)	(3)	(4)	(5)	(6)
School enrolment				-0.557***	-0.570***	-0.564***
				(0.172)	(0.155)	(0.105)
Landownership inequality	-1.974***		-2.082***			
	(0.378)		(0.358)			
Distance to Wittenberg		-0.025***	-0.027***			
		(0.004)	(0.004)			
Share in industry	0.233	0.439**	0.308	0.478***	0.483***	0.481***
	(0.209)	(0.206)	(0.198)	(0.113)	(0.115)	(0.105)
Share in agriculture	-0.054	-0.060	0.002	0.067*	0.065*	0.066*
	(0.043)	(0.042)	(0.041)	(0.039)	(0.039)	(0.037)
Share urban	-0.077*	-0.189***	-0.155***	-0.077**	-0.079**	-0.078**
	(0.040)	(0.042)	(0.040)	(0.034)	(0.033)	(0.030)
Population density	0.013**	0.016***	0.015***	0.003	0.004	0.003
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)
Share married women	0.660***	0.251**	0.281**	0.779***	0.787***	0.783***
	(0.102)	(0.119)	(0.113)	(0.145)	(0.152)	(0.128)
Poland	-0.071***	-0.060***	-0.043***	-0.062***	-0.063***	-0.063***
	(0.012)	(0.012)	(0.012)	(0.018)	(0.015)	(0.013)
Constant	0.427***	0.790***	0.763***	0.540***	0.546***	0.543***
	(0.080)	(0.097)	(0.093)	(0.106)	(0.089)	(0.078)
Observations	334	334	334	334	334	334
R^2	0.360	0.370	0.430	0.139	0.124	0.131
Partial F-statistic 1st stage				18.007	34.233	27.985
Sargan-Hansen p-value						0.958

Source: County-level data from the Prussian Census 1849; see main text and Appendix A for details.

Notes: Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 4. The effect of education on fertility – robustness checks

	IV second stage			
Dependent variable:	Child-woman ratio			Marital fertility rate
	(1)	(2)	(3)	(4)
School enrolment	-0.566***	-0.568***	-0.552***	-0.721***
	(0.097)	(0.097)	(0.103)	(0.157)
Share in industry	0.479***	0.486***	0.494***	0.733***
	(0.107)	(0.107)	(0.107)	(0.157)
Share in agriculture	0.067	0.066	0.062	0.104*
	(0.041)	(0.041)	(0.040)	(0.058)
Share urban	-0.078***	-0.086***	-0.094***	-0.133***
	(0.030)	(0.032)	(0.030)	(0.046)
Population density	0.003	0.004	0.004	0.002
	(0.004)	(0.004)	(0.004)	0.007)
Share married women	0.786***	0.789***	0.791***	-0.277
	(0.120)	(0.119)	(0.118)	(0.234)
Poland	-0.063***	-0.068***	-0.071***	-0.095***
	(0.012)	(0.012)	(0.011)	(0.016)
Schools per 100 children	-0.0022	-0.0003	0.0028	0.002
	(0.020)	(0.019)	(0.019)	(0.026)
Life expectancy at age 0		-0.001		-0.001
		(0.001)		(0.001)
Life expectancy at age 5			-0.001	
			(0.001)	
Constant	0.544***	0.565***	0.595***	1.730***
	(0.076)	(0.078)	(0.077)	(0.121)
Observations	334	334	334	334
R^2	0.130	0.129	0.155	0.105
Partial F-statistic 1st stage	39.135	46.505	41.271	46.505
Sargan-Hansen p-value	0.969	0.814	0.605	0.981

Source: County-level data from the Prussian Census 1849; see main text and Appendix A for details.

Notes: Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. School enrolment is instrumented with landownership inequality and distance to Wittenberg. Marital fertility rate is the number of children aged 0-5 over the number of married women.

Table 5. The effect of fertility on education

Dependent variable:	IV first stage Child-woman ratio			IV second stage School enrolment		
	(1)	(2)	(3)	(4)	(5)	(6)
Child-woman ratio				-0.365^{*}	-1.087^{***}	-0.430^{**}
				(0.213)	(0.364)	(0.189)
Child-woman ratio 1816	0.193^{**}		0.184^{**}			
	(0.077)		(0.078)			
Sex-imbalance ratio		0.268^{***}	0.135^{**}			
		(0.059)	(0.053)			
School enrolment 1816	-0.104^{***}		-0.090^{***}	0.180^{***}		0.170^{***}
	(0.026)		(0.028)	(0.050)		(0.049)
Share in industry	0.457^{***}	0.340^{***}	0.424^{***}	0.478^{**}	0.718^{***}	0.502^{**}
	(0.089)	(0.101)	(0.088)	(0.200)	(0.232)	(0.195)
Share in agriculture	0.058^{*}	0.061^{*}	0.052^{*}	-0.034	0.012	-0.029
	(0.029)	(0.032)	(0.030)	(0.054)	(0.063)	(0.053)
Share urban	-0.068^{**}	-0.108^{***}	-0.089^{***}	-0.173^{***}	-0.254^{***}	-0.179^{***}
	(0.034)	(0.029)	(0.032)	(0.043)	(0.057)	(0.044)
Population density	-0.008^{***}	-0.015^{***}	-0.010^{***}	0.004	-0.003	0.003
	(0.002)	(0.005)	(0.002)	(0.008)	(0.009)	(0.008)
Share Protestants	0.019	-0.022^{**}	0.021	0.011	0.037	0.012
	(0.013)	(0.011)	(0.013)	(0.033)	(0.036)	(0.033)
Poland	-0.037^{***}	-0.021^{**}	-0.031^{***}	-0.073^{***}	-0.108^{***}	-0.076^{***}
	(0.008)	(0.010)	(0.009)	(0.020)	(0.027)	(0.020)
Landownership inequality	0.582^{*}	1.481^{***}	0.653^{*}	-0.884	-0.725	-0.835
	(0.339)	(0.359)	(0.340)	(0.842)	(1.101)	(0.847)
Constant	0.500^{***}	0.370^{***}	0.373^{***}	1.008^{***}	1.563^{***}	1.054^{***}
	(0.079)	(0.062)	(0.109)	(0.149)	(0.243)	(0.134)
Observations	334	334	334	334	334	334
R^2	0.454	0.306	0.467	0.377	0.042	0.359
Partial F-statistic 1st stage				6.369	20.428	4.583
Sargan-Hansen p-value						0.334

Source: County-level data from the Prussian Censuses 1816 and 1849; see main text and Appendix A for details.

Notes: Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Standard errors are clustered at the level of the 280 units of observation in the 1816 data due to a change of borders after 1816.

Table 6. The effect of fertility on education – robustness checks

Dependent variable:	IV second stage School enrolment		
	(1)	(2)	(3)
Child-woman ratio	-0.454**	-0.412**	-0.415**
	(0.186)	(0.187)	(0.188)
School enrolment 1816	0.144***	0.144***	0.143***
	(0.044)	(0.043)	(0.043)
Share in industry	0.587***	0.564**	0.571**
	(0.213)	(0.218)	(0.220)
Share in agriculture	-0.069	-0.076	-0.073
	(0.063)	(0.061)	(0.062)
Share urban	-0.154***	-0.128***	-0.128***
	(0.049)	(0.047)	(0.047)
Population density	0.001	0.002	0.002
	(0.007)	(0.007)	(0.007)
Share Protestants	-0.004	-0.009	-0.009
	(0.036)	(0.035)	(0.035)
Poland	-0.068***	-0.054**	-0.055**
	(0.018)	(0.023)	(0.023)
Landownership inequality	-1.110	-1.030	-1.054
	(0.726)	(0.694)	(0.696)
Schools per 100 children	0.089**	0.089**	0.090**
	(0.038)	(0.038)	(0.038)
Life expectancy at age 0		0.001	0.001
		(0.001)	(0.001)
Temporary male migration			-0.099*
			(0.060)
Constant	1.030***	0.944***	0.946***
	(0.133)	(0.153)	(0.154)
Observations	334	334	334
R^2	0.377	0.395	0.395
Partial F-statistic 1st stage	4.751	4.408	4.480
Sargan-Hansen p-value	0.336	0.393	0.430

Source: County-level data from the Prussian Censuses 1816 and 1849; see main text and Appendix A for details.

Notes: Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Standard errors are clustered at the level of the 280 units of observation in the 1816 data due to a change of borders after 1816. The child-woman ratio is instrumented with the child-woman ratio 1816 and the sex-imbalance ratio.

Table 7. The long-run effect of education on the fertility transition: crude birth rates

Dependent variable:	OLS			
	Crude birth rate 1880-1905 (% change)			
	(1)	(2)	(3)	(4)
School enrolment	-0.132***	-0.079*	-0.048	-0.069
	(0.043)	(0.047)	(0.047)	(0.049)
Child-woman ratio (1849)	-0.227***	-0.299***	-0.315***	-0.290***
	(0.078)	(0.070)	(0.069)	(0.069)
Share in industry	0.110	0.010	-0.068	-0.081
	(0.165)	(0.169)	(0.176)	(0.172)
Share in agriculture	-0.008	0.021	0.061	0.057
	(0.040)	(0.037)	(0.041)	(0.040)
Share urban	-0.330***	-0.254***	-0.270***	-0.242***
	(0.036)	(0.036)	(0.037)	(0.040)
Population density	0.017***	0.007*	0.008*	0.008*
	(0.004)	(0.004)	(0.004)	(0.005)
Share Protestants		-0.118***	-0.102***	-0.106***
		(0.013)	(0.014)	(0.014)
Poland		-0.016	-0.021*	-0.008
		(0.011)	(0.011)	(0.013)
Schools per 100 children			-0.067***	-0.066***
			(0.023)	(0.023)
Life expectancy at age 0				0.002**
				(0.001)
Constant	0.223***	0.278***	0.300***	0.232***
	(0.065)	(0.069)	(0.068)	(0.072)
Observations	309	309	309	309
R^2	0.243	0.426	0.443	0.451

Source: County-level data from the Prussian Census 1849 and demographic data for different years; see main text and Appendix A for details.

Notes: Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Crude birth rate is defined as the number of births (in 1000s) over the total population.

Table 8. The long-run effect of education on the fertility transition: marital fertility rates

Dependent variable:	OLS			
	Marital fertility rate 1890-1905 (% change)			
	(1)	(2)	(3)	(4)
School enrolment	-0.165***	-0.100***	-0.107***	-0.113***
	(0.036)	(0.032)	(0.034)	(0.037)
Child-woman ratio (1849)	-0.022	-0.109**	-0.105**	-0.097*
	(0.070)	(0.052)	(0.053)	(0.054)
Share in industry	-0.138	-0.259	-0.243	-0.247
	(0.180)	(0.194)	(0.191)	(0.191)
Share in agriculture	0.045	0.081**	0.072*	0.071*
	(0.039)	(0.034)	(0.038)	(0.038)
Share urban	-0.246***	-0.153***	-0.149***	-0.141***
	(0.035)	(0.032)	(0.032)	(0.035)
Population density	0.010***	-0.002	-0.002	-0.002
	(0.003)	(0.003)	(0.003)	(0.003)
Share Protestants		-0.143***	-0.147***	-0.148***
		(0.011)	(0.012)	(0.012)
Poland		-0.019**	-0.018**	-0.014
		(0.009)	(0.009)	(0.010)
Schools per 100 children			0.014	0.015
			(0.019)	(0.019)
Life expectancy at age 0				0.001
				(0.001)
Constant	0.103*	0.169***	0.164***	0.143***
	(0.061)	(0.049)	(0.048)	(0.053)
Observations	309	309	309	309
R^2	0.265	0.566	0.567	0.568

Source: County-level data from the Prussian Census 1849 and demographic data for different years; see main text and Appendix A for details.

Notes: Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Marital fertility rate is defined as the number of births (in 1000s) over the number of married women aged 15-49.
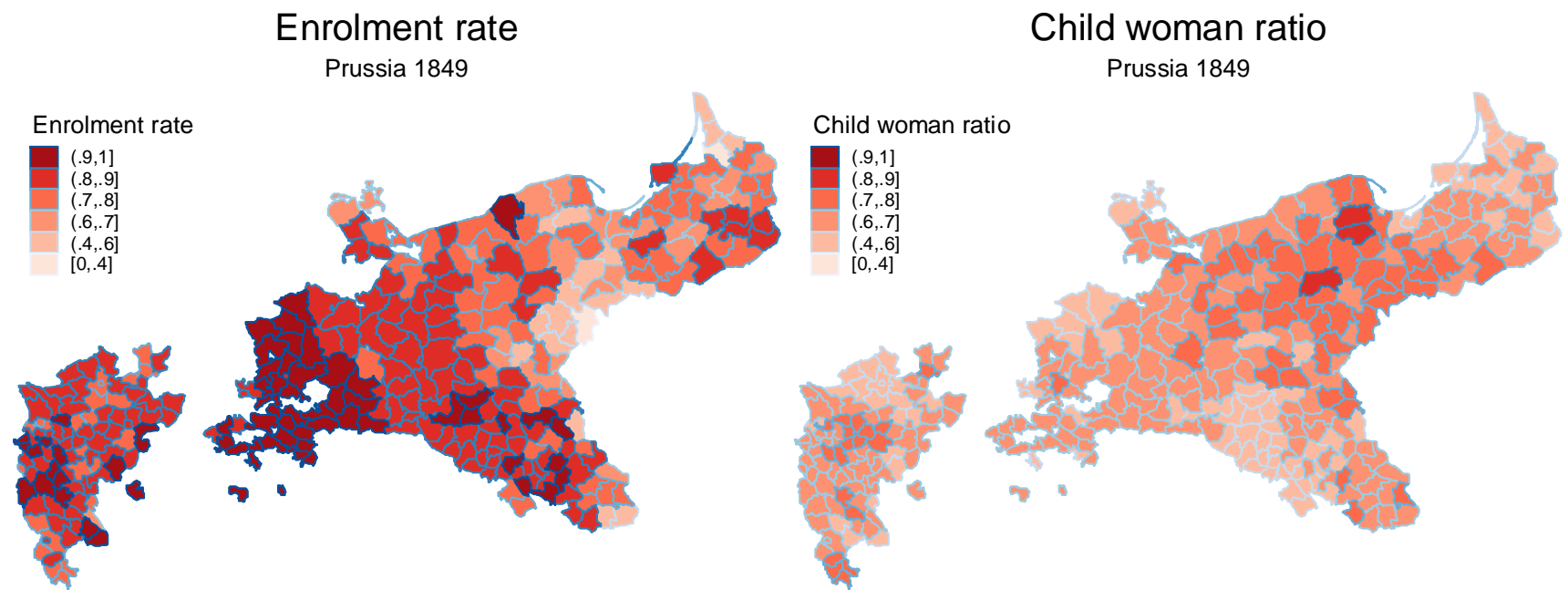
Enrolment rate
Prussia 1849

Enrolment rate
- (.9,1]
- (.8,.9]
- (.7,.8]
- (.6,.7]
- (.4,.6]
- [0,.4]

Child woman ratio
Prussia 1849

Child woman ratio
- (.9,1]
- (.8,.9]
- (.7,.8]
- (.6,.7]
- (.4,.6]
- [0,.4]

Figure 1. The geographic distribution of primary school enrolment and fertility

Source: County-level data from the Prussian Census 1849; see main text and Appendix A for details.

Enrolment rate

Prussia 1849

Enrolment rate
- (.9,1]
- (.8,.9]
- (.7,.8]
- (.6,.7]
- (.4,.6]
- [0,.4]

Change in marital fertility rate

Prussia 1890-1905

% change
- (.3,.5]
- (0,.3]
- (-.1,0]
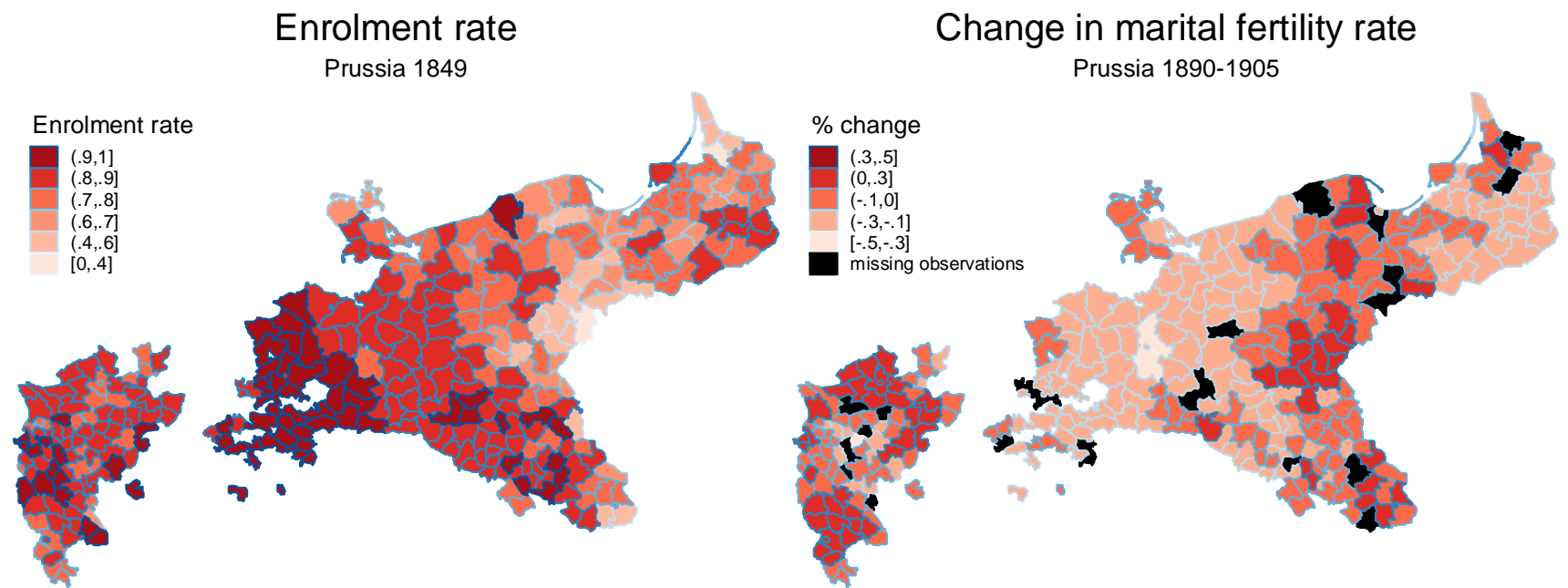- (-.3,-.1]
- [-.5,-.3]
- missing observations

Figure 2. The geographic distribution of education in 1849 and the fertility transition in 1890-1905

Source: County-level data from the Prussian Census 1849 and vital statistics for different years; see main text and Appendix A for details.