PROGRAM ON THE GLOBAL DEMOGRAPHY OF AGING

Working Paper Series

The Long-Run Determinants of Fertility: One Century of Demographic Change 1900-1999

Dierk Herzer, Holger Strulik, Sebastian Vollmer

October 2010

PGDA Working Paper No. 63
http://www.hsph.harvard.edu/pgda/working.htm

 The Program on the Global Demography of Aging receives funding from the National Institute on Aging, Grant No. 1 P30 AG024409-06.

The Long-Run Determinants of Fertility: One Century of Demographic Change 1900-1999*

Dierk Herzer†

Holger Strulik‡

Sebastian Vollmer§

First version: October 2010. Revised version: May 2011.

Abstract. We examine the long-run relationship between fertility, mortality, and income using panel cointegration techniques and the available data for the last century. Our main result is that mortality changes and growth of income per capita account for a major part of the fertility change characterizing the demographic transition. The fertility reduction triggered by falling mortality, however, is not enough to overcompensate the positive effect of falling mortality on population growth. This means that growth of income per capita is essential to explain the observed secular decline of population growth. These results are robust against alternative estimation methods, potential outliers, sample selection, different measures of mortality, and the sample period. In addition, our causality tests suggest that fertility changes are both cause and consequence of economic development.

Keywords: fertility; mortality; economic development; panel cointegration.

JEL: J1, J13, C23.

*We would like to thank David Reher for sharing the data and David Canning, Carl-Johan Dalgaard, Michael Funke, and two anonymous referees for helpful comments.

† Helmut-Schmidt-University Hamburg, Department of Economics, Holstenhofweg 85, 22043 Hamburg, Germany. Email: herzer@hsu-hh.de.

‡ University of Hannover, Wirtschaftswissenschaftliche Fakultaet, Koenigsworther Platz 1, 30167 Hannover, Germany. Email: strulik@vwl.uni-hannover.de.

§ Harvard University, Center for Population and Development Studies, 9 Bow Street, Cambridge, MA 02138, USA. Email: svollmer@hsph.harvard.edu.

1. Introduction

Every successfully developing country runs through two one-time transformations, an industrial revolution, characterized by a secular take-off of income per capita, and a demographic transition, characterized by decreasing mortality and fertility rates. Although there are also important issues of timing – to which we turn later – the most salient observation is that both transformations happen so closely to each other chronologically that "our instincts suggest that there is some underlying connection between these events" (Clark, 2005).

The most debated question in this respect is probably whether the fertility decline is mainly caused by declining mortality – this would be the typical demographer's view – or whether declining fertility is essentially caused by technological change and the associated secular rise of income per capita – the typical economist's view. Moreover, neoclassical growth theory (Solow, 1954, Mankiw et al., 1992) argues in favor of an impact of mortality and fertility on income per capita through population growth and capital dilution while unified growth theory (Galor, 2005) argues that fertility changes are both cause and consequence of economic development.

The objective of this paper is to examine empirically the long-run effects of mortality and income on fertility and to disentangle the intricate problems of causality. For that purpose we take the available data for the demographic and economic evolution over the last century for a panel of countries and employ panel cointegration techniques. Panel cointegration estimators are robust under cointegration to a variety of estimation problems that often plague empirical work, including omitted variables and endogeneity (see, e.g., Banerjee, 1999; Baltagi and Kao, 2000; Pedroni, 2007). Moreover, panel cointegration methods can be implemented with shorter data spans than their time-series counterparts.

Because the demographic transition is an inherently dynamic phenomenon, the most interesting quest for causality is probably along the time-dimension. To be specific, we ask if and to what extent an observable fertility change should be seen as a response to a preceding change of mortality or as a response to a preceding change of income. To tackle these questions cointegration techniques and Granger causality appear to be the most appropriate tools because the whole idea of causality in the Granger sense is that the cause occurred before the effect.[1]

[1]Nevertheless it could be that Granger causality fails to identify true causality. It could be that the cointegrated variables are driven by another neglected process. This, however, would not affect the identified stationary relationship between the cointegrated variables.

The remainder of the paper is composed of four sections. In Section 2, we discuss theoretical background and empirical evidence. Section 3 sets out the basic empirical model and describes the data. Section 4 presents the econometric implementation and our main results. It documents that economic growth as well as declining mortality explain large parts of the fertility decline observed during the last century, that declining mortality per se is insufficient to explain the secular decline of population growth, and that fertility changes are both cause and consequence of successful economic development. Section 5 concludes. A detailed description of the data and of our econometric tests can be found in the Appendix.

2. Theoretical Background and Empirical Evidence

2.1. **Theory.** Most of the available theories of the demographic transition focus either on the impact of mortality or on the impact of income and economic growth. Demographers seem to emphasize the mortality channel while economists emphasize the income channel broadly understood, i.e. with rising income per capita functioning as a proxy for technological change and productivity growth. Among the most prominent explanations for the mortality channel put forward by demographers are physiological mechanisms (the link between breastfeeding and fecundity) and the concept of an ideal family size (implying the wish for replacement of deceased children). While these channels establish a negative association between fertility and mortality they are insufficient to explain the demographic transition understood as the secular decline of *net* fertility, i.e. of the number of surviving children per family and thus the secular decline of population growth.

In order to establish the mortality channel as sufficient for the demographic transition several refinements of the theory have been proposed. Most well-known is probably the idea of precautionary child-bearing of risk-averse parents (Sah, 1991, Kalemli-Ozcan, 2002, see Doepke, 2005, for a critique). More complex theories involve the interaction between extrinsic survival conditions and child health (Strulik, 2008) and the impact of adult longevity on fertility (Soares, 2005, Cervelatti and Sunde, 2007).

The basic challenge of economic demographic theory is to explain a negative association between income and fertility without abandoning the assumption of children as "normal goods". A common element is that a generally positive income effect is dominated by an accompanying negative substitution effect. Theories differ with respect to their motivation of the substitution

effect. Gary Becker has contributed two theories to that end, one based on time allocation (children are more time-intensive than other consumption goods; Becker, 1965), the other based on the quantity-quality trade-off (preferences and or constraints are such that households prefer to substitute fertility with child expenditure as income rises; Becker, 1960, Becker and Lewis, 1973).

With the rise of unified growth theory (see Galor, 2005, for a survey) the economic analysis of fertility has been reframed in a dynamic context. The focus shifted away from the association between fertility and income (across countries) towards the association between fertility *change* and income *growth* (within countries over time). Moreover, the time-cost idea and the child quality-quantity trade off have been refined in several new ways. For example, it has been proposed that the prospect of higher future returns on education induces a child quantity-quantity substitution (Becker et al., 1990) that rising income – as a proxy for technological progress – is associated with a reduction of the comparative advantage of men in production and thus rises the opportunity cost of fertility for women (Galor and Weil, 1996), that technological progress is skill-biased and raises the importance of human capital (education, child quality) vis a vis raw labor in production (Galor and Weil, 2000, Galor and Moav, 2002), and that technological progress changes the structural composition of the economy toward manufacturing and thus raises the relative price of nutrition, i.e. the relative price of child quantity (Koegel and Prskawetz, 2001, Strulik and Weisdorf, 2008).

A common element of these income-based theories is that – without further assumptions – mortality plays no role in explaining the fertility transition. Indeed if child mortality is added in a standard fashion in these frameworks it cancels out in the computation of optimal *net* fertility (see Doepke, 2005). Without further augmentation these models thus predict that a change of mortality leads to a one-to-one response of fertility and has no consequences on population growth.

A micro-foundation of net fertility and mortality can be established by abandoning the assumption of homothetic utility. Based on this idea we next present a simple model providing a theoretical motivation of our main empirical findings. The model predicts that fertility is negatively associated with income and positively associated with mortality whereas net fertility is *negatively* associated with mortality, implying the prediction that declining mortality is not

sufficient to explain the phenomenon of declining population growth during the demographic transition.

Suppose life is divided into three periods: childhood, young adulthood and old age. Let c_1 and c_2 denote consumption of manufactured goods at young and old age, n the fertility rate, and π_1 the child survival rate such that $\pi_1 n$ denotes the number of surviving children. All decisions are made by young adults. A young individual maximizes life-time utility received from goods consumption now, from expected consumption in old age and from the number of surviving children such that $u = \log(c_1 + \alpha) + \beta\pi_2 \log(c_2) + \gamma \log(\pi n - \bar{n})$. Here β is the time discount rate, π_2 the survival probability from young adulthood to old age, and γ is the weight of children in utility. Following Greenwood et al. (2005), the parameter α captures (subsistence) goods produced at home. It is further assumed that there exists a number $\bar{n}$ of children that is regarded as a basic need, i.e. below which marginal utility from an additional child is infinite (see Eckstein et al., 1999). For simplicity (and without loss of generality) we assume that old people do not operate the subsistence technology any longer and normalize the interest rate to zero.

Suppose that young individuals divide their time between supplying labor for the production of manufactured goods and child bearing and rearing, that each individual is endowed with one unit of time, and that each born child needs b units of time.[2] Old individuals neither work nor multiply. Let y denote potential market income such that the budget constraint is given by $(1 - bn)y = c_1 + c_2$. Solving the first order conditions with respect c_1, c_2, and n to fertility provides the solution

$$n = \frac{\alpha\gamma\pi_1 + (\gamma\pi_1 + b\bar{n} + \beta b\bar{n}\pi_2)y}{(1 + \gamma + \beta\pi_2)b\pi_1 y}.$$

The derivatives with respect to y, π_1, and π_2 are

$$\frac{\partial n}{\partial y} = -\frac{\alpha\gamma}{(1 + \gamma + \beta\pi_2)by^2} < 0,$$
$$\frac{\partial n}{\partial \pi_1} = -\frac{\bar{n}(1 + \beta\pi_2)}{(1 + \gamma + \beta\pi_2)\pi_1^2} < 0, \qquad \frac{\partial n}{\partial \pi_2} = -\frac{\beta\gamma\left[\alpha\pi_1 + (\pi_1 - b\bar{n})y\right]}{(1 + \gamma + \beta\pi_2)^2 b\pi_1 y} < 0.$$

The model thus suggests a negative association between fertility and income and a negative association between fertility and survival i.e. a positive association between fertility and child

[2] A more elaborate model would arrive at similar conclusions by assuming that home production takes also time but less than manufacturing, that leisure rises utility, and that time spend on child bearing depends on fertility while time spend on child rearing depends on the number of surviving children.

mortality. These correlations are driven by a hierarchy of needs. The income elasticity of fertility is $-\infty$ when income earned on the labor market goes to zero, indicating that under subsistence conditions rising market wages would pre-dominantly rise labor supply and demand for market goods. On the other hand, as income goes to infinity the income elasticity of fertility goes to zero and the substitution effect in household time-allocation levels off. Similarly the reaction of fertility on child survival prospects is largest when survival probabilities are low, indicating that the need to replace deceased children is largest when the prospects of survival are low.

The impact of adult mortality on fertility may look ambiguous but it is not. The optimal solution provides positive consumption in old age only if $\alpha\pi_1 + (\pi_1 - b\bar{n})y > 0$. Otherwise young adults are too poor and choose the corner solution for savings in favor for fertility, which is the more important need. Thus, whenever there is old age consumption, better adult survival depresses fertility.[3] The result is very intuitive and provides another channel for declining fertility: better survival prospects make saving for old age more attractive and young adults substitute savings for fertility.

Declining child mortality, however, is not sufficient to explain the demographic transition. To see this, multiply n by π_1 to get net fertility and take the derivative with respect to π_1

$$\frac{\partial(\pi_1 n)}{\partial \pi_1} = \frac{\gamma(\alpha + y)}{(1 + \gamma + \beta\pi_2)by} > 0.$$

With contrast to precautionary child-bearing (which predicts a negative association) and with contrast to the standard quality-quantity trade-off (which predicts no association), the model predicts a *positive* association between child survival and net fertility.

The model offers more precise predictions with respect to the subsequent empirical analysis if we assume that the probabilities of surviving to young adulthood and of surviving to old age are correlated. Specifically, let π denote a measure of aggregate survival probability and assume that $\pi_1 = \rho_1\pi$ and $\pi_2 = \rho_2\pi$, $\rho_i \geq 0$. The implied total population of the economy is $1 + \rho_1\pi + \rho_1\rho_2\pi^2$ times the number of births. At any point of time the total number of deaths is $(1 - \rho_1\pi) + (1 - \rho_2)\rho_1\pi + \rho_1\rho_2\pi^2$ times the number of births. The implied crude death rate is given by $d(\pi) = 1/(1 + \rho_1\pi + \rho_1\rho_2\pi^2)$. It is negatively correlated with the aggregate measure of survival π, $\partial d/\partial\pi < 0$.

[3]The solution for second period consumption is $c_2 = \beta\pi_2 \left[\alpha\pi_1 + (\pi_1 - b\bar{n})y\right] / \left[\pi_1(1 + \gamma + \beta\pi_2)\right]$.

Substituting π into fertility n and net fertility $\pi_1 n$ and taking the derivative with respect to π provides:

$$\frac{\partial n}{\partial \pi} = -\frac{\alpha\beta\gamma\rho_1\rho_2\pi^2 + \left\{\beta\gamma\rho_1\rho_2\pi^2 + b\bar{n}\left[\gamma + (1+\beta\rho_2\pi)^2\right]\right\}}{\beta\rho_1\pi^2\left[1+\gamma+\beta\rho_2\pi\right]^2 y} < 0$$

$$\frac{\partial(\pi_1 n)}{\partial \pi} = \frac{\alpha\gamma(1+\gamma)\rho_1 + \left[\rho_1(1+\gamma) + \beta b\bar{n}\rho_2\right]\gamma y}{b(1+\gamma+\beta\rho_2\pi)^2 y} > 0.$$

The model thus predicts that an increasing crude death rate increases fertility and lowers net fertility. This means that the model predicts that the mortality channel cannot explain the secular decline of population growth observed along the demographic transition. Income per capita growth is needed to explain declining population growth.

2.2. **Evidence.** There exists still surprisingly little macro-econometric evidence on the determinants of fertility in modern (i.e. post-Malthusian) times.[4] Overall, the available literature provides a mixed and inconclusive picture. Across countries Brander and Dowrick (1993) document a negative association between fertility and economic growth, Schultz (1997) finds that income per adult is negatively associated with mortality and positively with fertility, and Ahituv (2001) finds a negative association between fertility and income per capita. Lorentzen et al. (2008) find a positive association between fertility and mortality and (indirectly) a negative association between fertility and economic growth.[5]

More closely related to our approach is the work of Wang et al. (1994), Eckstein et al. (1999) and Angeles (2010). Wang et al. use a structural VAR model and US data from the second half of the twentieth century and document the endogeneity of fertility in a cointegrated system together with output and employment. The impact of mortality is not investigated. Eckstein et al. use long-run Swedish data from 1751-1990 to fit a five-period overlapping generation model, which takes child mortality and income as (exogenous) determinants of fertility. They identify a negative impact of income on fertility and child mortality as the most important factor explaining

[4] There exists a relatively large literature on fertility in pre-modern times, i.e. times for which Malthusian theory predicts a positive association between fertility and income and a negative association between population density and income. See, among others, Eckstein et al. (1984), Galloway (1988), Lee and Anderson, 2002, Nicolini (2007), and Ashraf and Galor (2011). Microeconometric evidence is compiled in Schultz (1997).

[5] Without explicitly considering fertility, Acemoglu and Johnson (2007) observe a negative impact of life-expectancy on income per capita. Cervellati and Sunde (2009) demonstrate that this result depends heavily on the selected sample. In particular for countries that have already initiated the fertility transition, they document a positive effect of improving life-expectancy on economic growth.

the fertility decline. Interestingly, they also find that child mortality is not sufficient to explain the secular fall of net fertility. For that rising income is essential.

Like us Angeles (2010) tries to resolve endogeneity problems and to identify causality. For that purpose he uses the Arellano and Bond (1991) difference-GMM estimator and finds that a fall in mortality induces a significant reduction in fertility while the impact of GDP per capita on fertility is statistically insignificant. The GMM approach, however, needs strong assumptions for consistent estimates, assumptions which have not been verified in Angeles' study. The difference-GMM estimator instruments the right-hand-side variables with lagged values of the original regressors. Lagged levels, however, are weak instruments for a regression in differences when the variables are persistent. The fact that current fertility does not directly affect past mortality does not resolve the endogeneity concerns because autocorrelated mortality could imply that current mortality is correlated with past fertility. More importantly, it is well-known (and in detail documented by Reher, 2004) that the lag structure between mortality decline and fertility decline differs wildly across the world. For a sample of more than one hundred countries we thus expect problems arising from the assumption of a common lag structure.

We are therefore confident that our cointegration approach advances the state-of-the art empirical research on the long-run determinants of fertility.

3. Empirical Model and Data

Since it may take a long time before changes in mortality and the standard of living are reflected in changes in fertility, we adopt an empirical specification that captures the long-run relationship between these variables. In this section, we present the empirical specification, discuss some econometric issues, and describe the data.

3.1. **Empirical Specification and Econometric Issues.**

We assume that the correct specification of the long-run relationship between fertility, mortality, and economic development is given by

$$fert_{it} = a_i + \beta_1 \cdot mort_{it} + \beta_2 \cdot \log(gdp_{it}) + e_{it} \tag{1}$$

where $i = 1, 2, \ldots, N$ and $t = 1, 2, \ldots, T$ are country and time indices, $fert_{it}$ is fertility, measured by the crude birth rate (births per thousand population), and $mort_{it}$ stands for mortality, measured by the crude death rate (deaths per thousand population). We use the crude death

rate and not infant or child mortality because the crude death rate captures more effectively the full effect of mortality on fertility including effects from adult longevity (later on we check robustness of our results by substituting infant mortality into the regression). Another advantage of focusing on the crude death is that we can readily infer from our estimates the impact of mortality decline on population growth.

The level of economic development is represented by GDP per capita, gdp_{it}, measured in logs, as is common practice in the related empirical literature. Moreover, the measurement in logs has important implications with respect to the underlying test of demo-economic theory. To see this, differentiate (1) and obtain the change of fertility $\mathrm{d}fert_{it}$ as a function of the change of mortality $\mathrm{d}mort_{it}$ and of the growth rate of GDP per capita, $\mathrm{d}gdp_{it}/gdp_{it}$. With respect to the demographic transition equation (1) thus stipulates that fertility *change* is associated with income *growth* as suggested by unified growth theory.

The β coefficients in (1) capture the long-run effects of mortality and income on fertility. Because our principal interest is on long-run effects, it is not essential to be concerned about the variable lags through which mortality and per capita income affect fertility. Finally, we include country-specific fixed effects, a_i, to control for country-specific factors that are relatively stable over time, such as geography and culture.

Equation (1) assumes a long-run trivariate relationship between permanent movements in the crude birth rate, the crude death rate, and the log level of GDP per capita. Necessary conditions for this assumption to hold – and thus for our model to be a correct description of the data – are that the individual time series for fertility, mortality, and per capita income are nonstationary or, more specifically, integrated of the same order and that $fert_{it}$, $mort_{it}$, and $\log(gdp_{it})$ form a cointegrated system.

A specific advantage of the cointegration framework is that a regression consisting of cointegrated variables has a stationary error term, implying that no relevant integrated variables are omitted. Any omitted non-stationary variable that is part of the cointegrating relationship would enter the error term e_{it}, thereby producing non-stationary residuals and failure to detect cointegration. If, on the other hand, there is cointegration between a set of variables, then the same stationary relationship exists also in an extended variable space (see, e.g., Johansen, 2000); if the variables are nonstationary and not cointegrated, the error term is nonstationary as well, and any inferences are spurious.

These features are in particular important with respect to education as an omitted variable. Theory, in particular unified growth theory, as well as other other empirical studies (e.g. the work of Schultz, 1997, and Angeles, 2010) suggest that education is an important explanatory variable for fertility besides mortality and income. Here we have not considered education because of lacking data for the complete last century. It is thus important to emphasize that the cointegration tests described below verify that omitted education does not bias our results on the long-run relationship between mortality, fertility, and income. In other words, education (human capital) could potentially be in our set of cointegrated variables *instead* of income but it cannot be in it *on top of* income.[6]

Another assumption inherent in Equation (1) is that fertility is endogenous in the sense that, in the long run, changes in mortality and per capita income cause changes in fertility. The fact that the existence of cointegration implies long-run Granger-causality in at least one direction, however, does not exclude the possibility of long-run causality running from fertility to GDP per capita and mortality.

According to neoclassical growth theory, for example, high population growth due to increased fertility lowers income per capita because capital is spread more thinly over the population. On the other hand, population growth plays quite a different role in many R&D-based models of endogenous growth (Romer, 1990, Jones, 1995). Strictly interpreted, i.e. in the sense that "more people means more Isaac Newtons and therefore more ideas" (Jones, 2003), these theories predict that higher population growth leads to higher economic growth and thus to higher income per capita. Finally, an increase in fertility may also lead to an increase in mortality because a larger number of children per household could entail fewer resources available to invest in health of each child (Strulik, 2008).

The overall empirical implication is that it is not only crucial to examine the time-series properties of the variables and to test whether the variables are cointegrated, but it is also important to deal with these endogeneity problems and to investigate the direction of causality.

3.2. **Data and Descriptive Statistics.** The analysis of the long-run relationship between fertility, mortality, and income requires the use of data over a long time window. Therefore,

[6]There are, of course, potentially several others factors conceivable that influence fertility (such as, for example, government policy and health status). Since the cointegration property is invariant to extensions of the information set, adding further variables may result in further cointegrating relationships; it would, however, not destroy the original cointegrating relationship. As discussed in detail by Lütkepohl (2007), this property justifies to consider "subsystems" like the cointegrating relationship between $fert_{it}$, $mort_{it}$, and $\log(gdp_{it})$.

we select a sample of countries for which continuous data are available over a 100-year period from 1900 to 1999. Data on birth and death rates are from the database compiled by David Reher (2004) and data on (real) per capita GDP are from Maddison (2003), available at http://dx.doi.org/10.1787/456125276116. Since Reher's data are averaged over five years, we use five-year averages of all variables, implying that we have 20 time series observations per country. We include all countries with complete time series, resulting in a balanced panel with 400 observations and 20 countries. As illustrated in Figure A1 in the Appendix, these countries are geographically dispersed around the world, located in North America (Canada), Central America (Mexico), South America (Argentina, Chile, Colombia, Uruguay, and Venezuela), Europe (Belgium, Denmark, Finland, France, Italy, Netherlands, Norway, Portugal, Spain, Sweden, and Switzerland), South Asia (Sri Lanka), and East Asia (Japan).

In the figures in Appendix A1 we show the data for each country separately over the period 1900-1999. As can be seen, fertility and mortality exhibit a decreasing trend in all cases except for Denmark where mortality declined from 1900 to 1954 and then rose between 1955 and 1999 (see Figure A.2, row 2, column 2). Real GDP per capita, in contrast, exhibits a strong upward trend in all countries. Overall, the time-series evolution is consistent with the possibility that $fert_{it}$, $mort_{it}$, and $\log(gdp_{it})$ are nonstationary and cointegrated, an observation which we confirm by several panel unit root test and panel cointegration tests (Appendix A1 and A2).

Table 1 lists the countries along with the average values for $fert_{it}$, $mort_{it}$, and $\log(gdp_{it})$ over the period of observation. As expected, there are large cross-country differences in the values of these parameters. Mexico is the country with the highest fertility rate, followed by Venezuela, Colombia, and Chile, while Belgium ranks at the bottom of the fertility scale. Mexico is also the country with the highest mortality rate, followed by Chile, Sri Lanka, and Colombia. Average income is highest in Switzerland, and lowest in Sri Lanka, Colombia, and Mexico. Altogether, it appears that countries with higher mortality rates and lower per capita income tend to have higher fertility rates, suggesting a positive relationship between fertility and mortality and a negative relationship between fertility and income.

The last column in Table 1 reports the year of the onset of the fertility transition as identified by Reher (2004). In all but two countries (Sweden and Uruguay) the fertility transition began in 20th century, indicating that we focus indeed on the most interesting century of demographic change. 12 countries experienced the onset in the first half of the last century while 6 countries

Table 1: Countries and Country Summary Statistics

	crude birth rate	crude death rate	log GDP per capita	onset of transition
Argentina	27.12	11.67	3.71	1910
Belgium	16.61	12.94	3.84	1905
Canada	22.64	9.50	3.88	1915
Chile	34.02	17.70	3.59	1960
Colombia	38.57	16.14	3.34	1965
Denmark	18.89	11.02	3.88	1910
Finland	20.57	12.85	3.69	1915
France	16.70	13.93	3.81	1900
Italy	20.46	13.66	3.71	1925
Japan	24.14	13.19	3.59	1950
Mexico	40.88	19.92	3.45	1970
Netherlands	20.89	9.96	3.86	1910
Norway	18.55	10.92	3.77	1905
Portugal	24.11	14.78	3.46	1925
Spain	22.32	14.28	3.57	1910
Sri Lanka	33.33	16.64	3.16	1960
Sweden	16.64	11.62	3.83	1865
Switzerland	17.47	11.40	3.96	1910
Uruguay	22.89	10.44	3.63	1890
Venezuela	39.01	15.54	3.63	1965

Numbers for birth rates and death rates (in per thousand) and for GDP are country averages 1900-1999. Onset of the transition is the year of onset of the fertility transition as identified by Reher (2004).

experienced it in the second half. The huge variation of the onset of the transition across countries could be one explanation for the problem of earlier studies (by focussing on individual countries or across countries on a single year or on a shorter time period) in identifying a general pattern for the long-run determinants of fertility.

4. Empirical Analysis

The pre-tests for unit-roots and cointegration, which are reported in the Appendix, suggest that the variables are nonstationary and cointegrated, as assumed in Equation (1). In this section, we provide estimates of the cointegrating relationship between fertility, mortality, and income, test the robustness of the estimates, and investigate the direction of causality between the three variables.

4.1. **Long-run Relationship.** In order to estimate the long-run elasticities of fertility with respect to mortality and per capita income, we use the dynamic ordinary least squares (DOLS) estimator. This estimator is asymptotically equivalent to Johansen's (1988) system estimator. It generates unbiased and asymptotically efficient estimates of the long run relationship, even

with endogenous regressors (see, for example, Stock and Watson, 1993), thus allowing us to control for the potential endogeneity of mortality and per capita income. In addition, it is well-known that in small T samples (like ours) the DOLS estimator performs better than other available estimators, like, for example, the FIML estimator of Johansen (1988) or the fully modified ordinary least squares (FMOLS) estimator of Phillips and Hansen (1990). This is true for time series models as well as for panel data models (see, e.g., Stock and Watson, 1993; Kao and Chiang, 2000; Wagner and Hlouskova, 2010). Following Kao and Chiang (2000), the within-dimension-based DOLS model for our research question is given by (2).

$$fert_{it} = a_i + \beta_1 \cdot mort_{it} + \beta_2 \cdot \log(gdp_{it}) + \sum_{j=-k}^{k} \Phi_{1ij} \Delta mort_{it-j} + \sum_{j=-k}^{k} \Phi_{2ij} \Delta \log(gdp_{it-j}) + \epsilon_{it} \quad (2)$$

where Φ_{1ij} and Φ_{2ij} are coefficients of lead and lag differences which account for possible serial correlation and endogeneity of the regressors, thus yielding unbiased estimates of β_1 and β_2. The results of this estimation procedure are presented in the first row of Table 2 where, for brevity, we report only the estimated β coefficients. The coefficient on $mort_{it}$ is highly significant and positive, while the GDP per capita variable has a highly significant negative coefficient.

More precisely, the elasticity of fertility with respect to mortality is estimated to be 0.378, implying that, in the long-run, a one-standard-deviation increase in the mortality variable is associated with an increase in the fertility variable equal to 25 percent of a standard deviation in that variable. The coefficient on $\log(gdp_{it})$, in contrast, is -5.246, indicating that a one-standard-deviation increase in this variable reduces the fertility rate by 42 percent of a standard deviation in the fertility variable.[7]

In other words, these results imply that an increase of GDP per capita by $1000 and a decrease of the mortality rate by 0.5 percentage points both decreases the fertility rate by about 0.19 percentage points. In conclusion, both mortality changes as well as income changes have a large impact on fertility reductions and account for a major part of the fertility change characterizing the demographic transition.

Our estimates imply furthermore that a reduction of the mortality rate by 0.5 percentage points is associated with an *increase* of the population growth rate by 0.5-0.19=0.31 percentage points holding GDP constant. From that we conclude that declining mortality is insufficient

[7] The standardized coefficients are calculated by multiplying the unstandardized coefficients (the β coefficients) by the ratio of the standard deviations of the independent and dependent variables. The standard deviation of $fert_{it}$ is 9.992, the standard deviation of $mort_{it}$ is 6.627, and the standard deviation of $\log(gdp_{it})$ is 0.797.

to explain the declining population growth observed along the path of demographic transition. Although mortality is identified as an important driver of decreasing fertility, GDP growth is essential in order to explain the secular decline of population growth.

Table 2: Estimates of the Long-run Effects on Fertility

	$mort_{it}$	$\log(gdp_{it})$
Within-dimension DOLS estimator Kao and Chiang (2000)	0.378^{**} (7.40)	-5.246^{**} (-10.18)
DOLS mean group estimator Pedroni (2001)	0.747^{**} (9.04)	-5.489^{**} (-12.83)
CCE mean group estimator Pesaran (2006)	0.880^{**} (10.45)	-4.456^{**} (-11.23)
2-step estimator Breitung (2005)	0.855^{**} (11.33)	-8.455^{**} (-19.87)
Diagnosis tests		
JB	18.73 [0.998]	
RESET	45.07 [0.268]	
HET	36.60 [0.624]	
LM(1)	39.09 [0.511]	
LM(3)	46.42 [0.225]	
STABILITY	42.69 [0.356]	

The dependent variable is $fert_{it}$, $**$ indicate significance at the 1% level. t-statistics in parentheses. The DOLS regression was estimated with one lead and one lag. All statistics presented in the diagnostics are Fisher (1932) statistics, which are based on the country-specific diagnostic tests of the respective DOLS model; the Fisher statistic is distributed as χ^2 with $2 \times N$ degrees of freedom; the numbers in brackets are the corresponding p-values.

Finally, in the bottom part of Table 2, we present the results of some diagnostic tests. JB is a Jarque-Bera test for normality, RESET is a Ramsey RESET test for general nonlinearity and functional form misspecification, HET stands for a Breusch-Pagan-Godfrey test for heteroscedasticity, LM(k), $k = 1, 3$, are Lagrange Multiplier tests for autocorrelation based on one and three lags, and STABILITY is an Lc type panel test for parameter instability in the style of Hansen (1992). All statistics presented in the diagnostics are Fisher (1932) statistics, defined as $\lambda = -2\sum_i \log(p_i)$, where p_i is the p-value of the country-specific diagnostic test of the respective DOLS model; the Fisher statistic is distributed as χ^2 with $2 \times N$ degrees of freedom.

As can be seen, all test statistics reject the respective null hypothesis, suggesting that neither obvious nonlinearity nor misspecification is present, that the residuals show no signs of non-normality, autocorrelation or autoregressive heteroscedasticity, and that the estimated parameters are stable. Since parameter constancy may imply a cointegrating relationship, whereas

parameter instability and structural change can lead to the finding of no cointegration, the finding of stability is consistent with the finding that $fert_{it}$, $mort_{it}$, and $\log(gdp_{it})$ are cointegrated without a structural break in the cointegrating vector.

4.2. **Robustness Checks.** To assess the robustness of our conclusions, we perform several sensitivity checks. First, we investigate whether the estimates are robust to alternative estimation methods. Specifically, a potential problem with the above estimation procedure could be that it assumes homogeneous β coefficients, which may be empirically incorrect. Countries differ widely in terms of economic structure, institutions, government policy, and other characteristics, implying that the effects of mortality and income on fertility could also differ across countries. To allow the slope coefficients to vary across countries, we use the between-dimension, group-mean panel DOLS estimator suggested by Pedroni (2001). This estimator involves estimating separate DOLS regressions for each country and averaging the long-run coefficients, $\hat{\beta} = N^{-1} \sum_{i=1}^{N} \hat{\beta}_i$. The t-statistic for the average coefficient is calculated as the sum of the individual t-statistics divided by the root of the number of cross-sectional units, $t_{\hat{\beta}} = \sum_{i=1}^{N} t_{\hat{\beta}_i} / \sqrt{N}$. We present the DOLS group-mean point estimates of the effects of mortality and income on fertility in the second row of Table 2.

Because the DOLS estimates could be biased in the presence of cross-sectional dependence, we also report (in the third row) the result of the common correlated effects (CCE) mean group estimator suggested by Pesaran (2006).[8] Compared to the use of common time dummies (to control for cross-sectional dependence through common time effects), as is common practice in panel studies, the CCE mean group estimator has the advantage that it allows for cross-sectional dependencies arising from multiple unobserved common factors, and that it permits the individual responses to the common factors to differ across countries.[9] It augments the cointegrating regression with the cross-sectional averages of the dependent variable and the observed regressors as proxies for the unobserved factors (see Equation (A.8) in the Appendix).

[8]Cross sectional dependence can arise due to several factors, such as omitted observed common factors, unobserved common factors, or spatial spillover effects. For example, the data may be in part driven by common global business cycles or health shocks. Shocks affecting fertility and mortality (and income) in several countries at the same time include major influenza epidemics, the spread of HIV/AIDS, the introduction of new vaccines, and the diffusion of antibiotics and contraceptives.

[9]The use of time dummies (or cross-sectionally demeaned data) implicitly assumes that the form of the dependency is such that it is driven by a single common source, and that individual countries respond in a similar fashion (Pedroni, 2007).

For completeness, we also present estimates obtained using the two-step estimator suggested by Breitung (2005). While the estimators discussed above are single equation techniques, the vector error-correction model (VECM) estimator of Breitung is a system approach. It involves estimating the Johansen (1988) VECM (given by Equation A.9 in the Appendix) separately for each country to obtain the country-specific error-correction coefficients a_i. In the second step, the estimated a_is are used to estimate the cointegration matrix β by running a pooled regression of $\hat{z}_{it}$ on $y_{t-1}^{(2)}$, where $\hat{z}_{it} = (\hat{\alpha}_i' \widehat{\sum}_i^{-1} \hat{\alpha}_i)^{-1} \hat{\alpha}_i' \widehat{\sum}_i^{-1} \Delta y_{it} - y_{t-1}^{(1)}$ and $y_{it} = [y_{t-1}^{(1)}, y_{t-1}^{(2)}]$. $y_{t-1}^{(1)}$ and $y_{t-1}^{(2)}$ are $r \times 1$ and $(p-r) \times 1$ sub-vectors of y_{it}.

As can be seen from Table 2, all four estimators provide qualitatively similar results, suggesting that both the positive effect of mortality and the negative effect of GDP per capita on fertility are neither due to potentially restrictive homogeneity assumptions, nor due to possible cross-sectional dependence, nor due to the single-equation specification. As expected, the between-dimension DOLS estimator produces larger estimates (in absolute value) than its within-dimension counterpart, a result that is in line with the findings of Pedroni (2001). More specifically, the magnitude of the mortality effect is about half as large for the within-dimension DOLS estimator compared to the other three estimators. For GDP per capita, the coefficients are fairly similar across the two DOLS and the CEE models, while the GDP per capita coefficient obtained by the two-step system estimator is substantially larger in magnitude. Thus, our main conclusions still hold, albeit less strongly for the inferred impact on population growth. According to the alternative estimators a 0.5 percentage point reduction of the death rate leads to a reduction of the birth rate by about 0.4 percentage points so that the mortality reduction per se, i.e. holding income constant, is still associated with a mild increase of population growth.

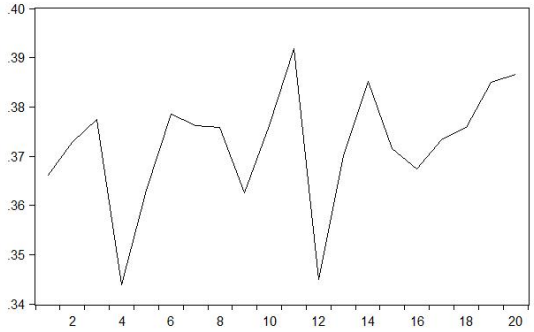
Given the limited number of time-series observations in our sample, the mean group results (which are based on individual time-series regressions) should be interpreted with caution. In addition, the CCE mean group estimator is intended for the case in which the regressors are exogenous, so that we lose the ability to account for the likely endogeneity of mortality and per capita GDP. Also, it is worth mentioning that there is evidence to suggest that the efficiency gains from pooling are likely to offset the potential biases due to individual heterogeneity (see, e.g., Baltagi and Griffin, 1997). In addition, Wagner and Hlouskova (2010) found that the pooled DOLS estimator outperforms all other estimators-both single equation and system estimators.

We are thus convinced that the pooled within-dimension panel DOLS estimator is the most appropriate one and continue our robustness analysis for this estimator.

In order to verify that the positive and negative coefficients on $mort_{it}$ and $\log(gdp_{it})$ are not due to potential outliers we re-estimate the DOLS regression excluding one country at a time from the sample. The sequentially estimated coefficients and their t-statistics are presented in Figure 1. They indicate that the coefficients on $mort_{it}$ are always significantly positive (and relatively stable between 0.344 and 0.392) and that the coefficients on $\log(gdp_{it})$ are always significantly negative (and relatively stable between -5.581 and -4.882). We conclude that our results are robust to potential outliers.

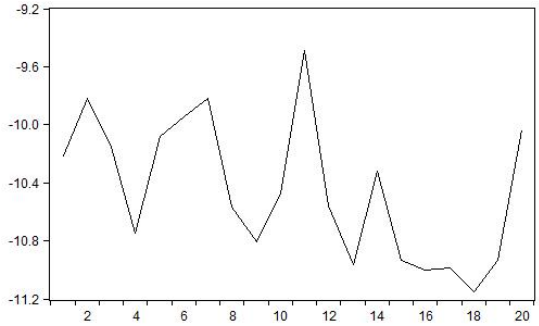
FIGURE 1: DOLS ESTIMATION WITH SINGLE COUNTRY EXCLUDED FROM THE SAMPLE

coefficients on $mort_{it}$



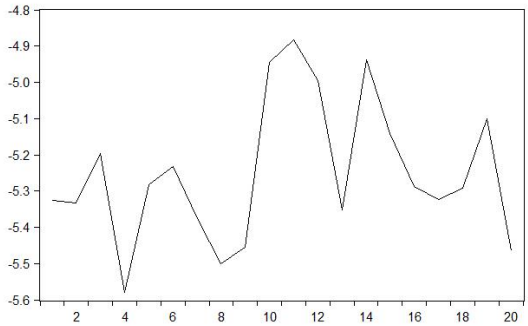
No. of omitted country

t–statistics of coefficients on $mort_{it}$



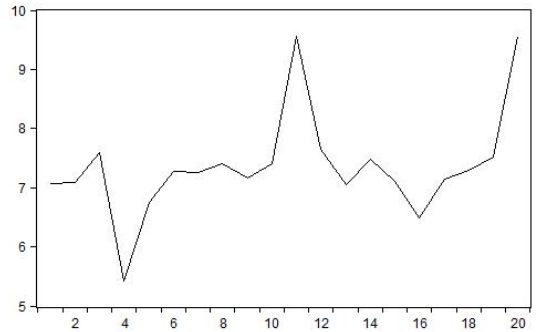
No. of omitted country

coefficients on $\log(gdp_{it})$



No. of omitted country

t–statistics of the coefficients on $\log(gdp_{it})$



No. of omitted country

Next, we examine whether the positive relationship between fertility and mortality, as well as the negative relationship between fertility and income are due to sample-selection bias. Sample-selection bias occurs when the selected sample is not random and thus not representative. Admittedly, a potential problem with our sample could be that it includes only 20 countries. We therefore re-estimate the DOLS regression for a second sample with 1190 observations on 119 countries over the period from 1950 to 1999 (again using five-year averages). The sample of countries (listed in Appendix A4) is now much more heterogenous and includes also the latecomers of the demographic transition from Asia and Africa.

The results based on this sample are reported in Table 3. The second row shows the estimated coefficients on the crude death rate and log GDP per capita (the original variables), while the third row presents DOLS estimates using the infant mortality rate, labeled $infantmort_{it}$, in place of the crude death rate to examine also the robustness of the results to alternative measures of mortality. The data sources are the same as described above. As can be seen from the table, the long-run effects of mortality and GDP per capita are still positive and negative, respectively, regardless of which sample and mortality measure is used, indicating that the results are robust to different samples and measures of mortality.

Moreover, the fact that the estimated coefficients for the period 1950 to 1999 are strikingly similar to those for the period 1900 to 1999 (0.42 and -5.8 in Table 3 compared to 0.38 and -5.2 in Table 2) suggests that our results are not sensitive to the sample period. This finding is consistent with the stability test result presented in Table 2. Given, however, that the number of time series observation (10 per country) is possibly too small to generate reliable cointegration estimates, the results in Table 3 should be interpreted with caution. We therefore prefer the results in Table 2.

Table 3: DOLS Estimates: 119 Countries 1950-1999

$infantmort_{it}$	$mort_{it}$	$\log(gdp_{it})$
	0.420^{**} (13.74)	-5.829^{**} (-11.46)
0.141^{**} (21.20)		-3.029^{**} (-9.06)

The dependent variable is $fert_{it}$, $**$ indicate significance at the 1% level. t-statistics in parentheses. The DOLS regression was estimated with one lead and one lag.

Finally, we investigate whether our results are driven by developed or relatively rich countries. To this end, we split both the 20-country and 119-country samples into two sub-samples:

developed (OECD) and developing (non-OECD) countries. The resulting coefficients are listed in Table 4. Regardless of which sub-sample is chosen, the coefficient on $mort_{it}$ is significantly positive, while the coefficient on $\log(gdp_{it})$ is significantly negative. Remarkably, there appear to be no significant differences in the effects of mortality and economic development on fertility between rich and poor countries.

TABLE 4: DOLS ESTIMATES FOR SUBSAMPLES

	$mort_{it}$	$\log(gdp_{it})$	No. of countries in subsample
20-country sample			
Developed countries	0.623** (6.25)	−4.757** (-8.22)	12
Developing countries	0.470** (5.45)	−4.021** (-3.50)	8
119-country sample			
Developed countries	0.502** (5.96)	−5.567** (-4.03)	16
Developing countries	0.487** (10.45)	−4.987** (-9.83)	103

The dependent variable is $fert_{it}$, ** indicate significance at the 1% level. t-statistics in parentheses. The DOLS regression was estimated with one lead and one lag. A country is classified as a developing (non-OECD) country if it was between 1961 (when the OECD was founded) and 1999 less than 75% of the time a member of the OECD.

4.3. **Causality.** Standard growth models predict that higher fertility lowers per capita GDP because physical capital is spread more thinly over the population. An increase in fertility may also lead to an increase in mortality because a larger number of children entails less resources available per child for nutrition and health. Consequently, causality may run from $mort_{it}$ and $\log(gdp_{it})$ to $fert_{it}$, from $fert_{it}$ to $\log(gdp_{it})$ and from $fert_{it}$ to $mort_{it}$.

To test the direction of causality, we use a two-step procedure. In the first step, we employ the (within) DOLS estimate of the long-run relationship to construct the disequilibrium term

$$ec_{it} = fert_{it} - [\hat{a}_i + 0.378 \cdot mort_{it} - 5.246 \log(gdp_{it})]\,. \tag{3}$$

In the second step, we estimate the vector error correction model (VECM)

$$\begin{pmatrix} \Delta fert_{it} \\ \Delta mort_{it} \\ \Delta \log(gdp_{it}) \end{pmatrix} = \begin{pmatrix} c_{1i} \\ c_{2i} \\ c_{3i} \end{pmatrix} + \sum_{j=1}^{k} \Gamma_j \begin{pmatrix} \Delta fert_{it-j} \\ \Delta mort_{it-j} \\ \Delta \log(gdp_{it-j}) \end{pmatrix} + \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} ec_{it-1} + \begin{pmatrix} \epsilon_{1it} \\ \epsilon_{2it} \\ \epsilon_{3it} \end{pmatrix} \tag{4}$$

where the error-correction term ec_{it-1} represents the deviation from the equilibrium and the adjustment coefficients a_1, a_2, and a_3 capture how $fert_{it}$, $mort_{it}$, and $\log(gdp_{it})$ respond to deviations from the equilibrium relationship.

From the Granger representation theorem we know that at least one of the adjustment coefficients must be nonzero if a long-run relationship between the variables exists. A significant error-correction term also suggests long-run Granger causality, and thus long-run endogeneity (see, e.g., Hall and Milne, 1994), whereas a non-significant adjustment coefficient implies weak exogeneity and no long-run Granger causality running from the independent to the dependent variable(s).

In the following, we test for weak exogeneity of fertility, mortality, and the level of economic development-and thus for long-run Granger noncausality between $fert_{it}$, $mort_{it}$, and $\log(gdp_{it})$. We begin with eliminating the insignificant short-run dynamics in the model successively according to the lowest t-values (until the remaining variables are significant at the five-percent level). Then we test the significance of the adjustment coefficients. This approach has been used by Hendry (1995, Chapter 16), Urbain (1995), Juselius (2001), Lütkepohl and Wolters (1998, 2003), and Herzer (2008), among others, to reduce the number of estimated parameters (according to Hendry's general-to-specific methodology) and, thus, to increase the precision of the weak exogeneity tests on the a-coefficients.[10] Since all variables in the model, including ec_{it-1}, are stationary (because the level variables are integrated of order 1 and cointegrated), a conventional likelihood ratio test can be used to test the null hypothesis of weak exogeneity, $H_0 : a_{1,2,3} = 0$.

Table 5 presents the results. The error correction terms are significantly different from zero in each equation, implying that the null hypothesis of weak exogeneity can be rejected for $fert_{it}$, $mort_{it}$, and $\log(gdp_{it})$ at least at the 5 % level. Thus, the weak exogeneity tests suggest that all variables are endogenous in the long run, from which it can be concluded that the statistical long-run causality indeed runs from $mort_{it}$ and $\log(gdp_{it})$ to $fert_{it}$, from $fert_{it}$ to $\log(gdp_{it})$, and from $fert_{it}$ to $mort_{it}$.

To test the robustness of this conclusion, we calculate generalized impulse responses from the full VAR-VECM system (with two lags).[11] Unlike traditional impulse response analysis (see, e.g.

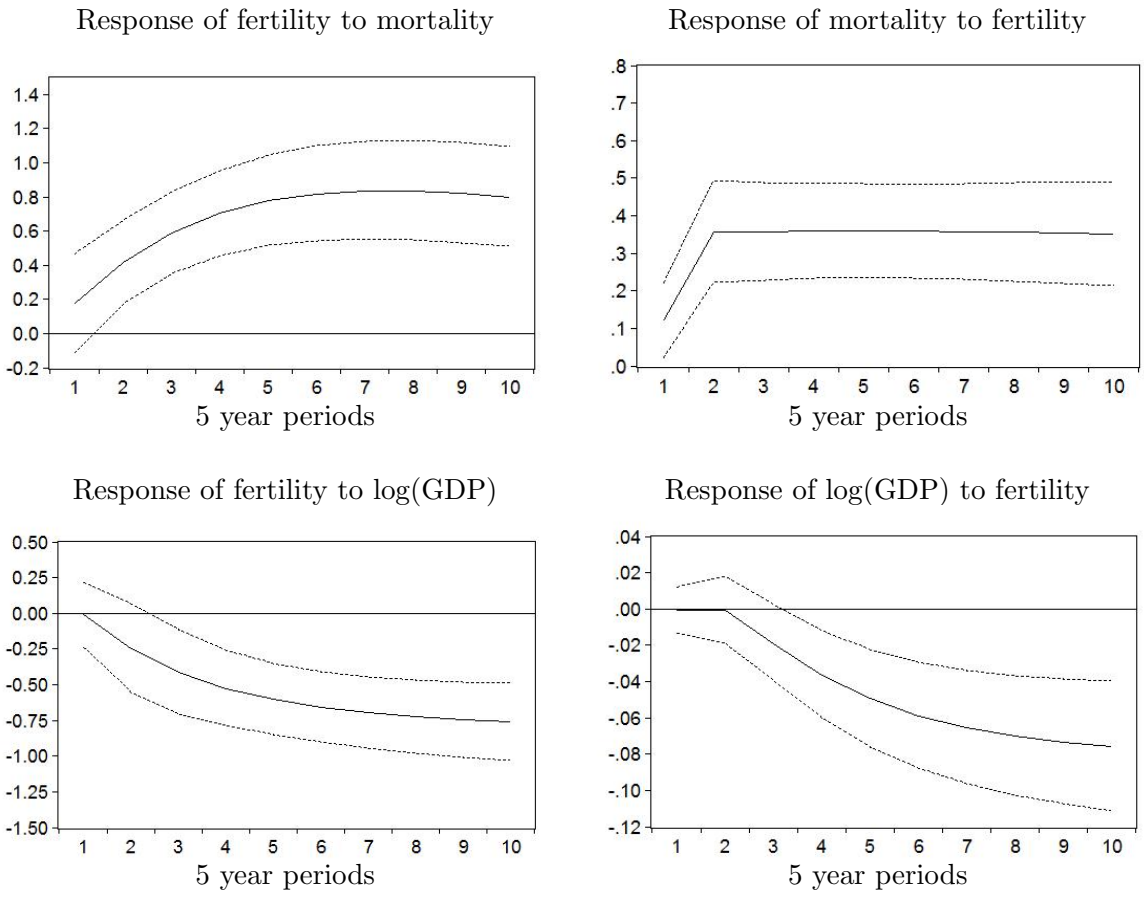
[10]The results (available on request) do not change substantially when alternative lag selection methods are used.
[11]We also experimented with VEC specifications of different lag orders, $k = 1, 3$, and found qualitatively similar results.

Table 5: Weak Exogeneity Tests / Long-run Causality Tests

	Weak exogeneity of...		
	$fert_{it}$ (significance of a_1)	$mort_{it}$ (significance of a_2)	$\log(gdp_{it})$ (significance of a_3)
χ^2 (1)	56.88	6.23	8.35
p-values	(0.000)	(0.013)	(0.004)

The number of degrees of freedom ν in the standard $\chi^2(\nu)$ tests correspond to the number of zero restrictions. The number of lags was determined by the general-to-specific procedure with a maximum of three lags.

Figure 2: Impulse-Responses



Lütkepohl and Reimers, 1992), which considers orthogonalized shocks based on the Cholesky decomposition, the generalized impulse response approach of Pesaran and Shin (1998) desirably yields unique impulse response functions that are invariant to the ordering of variables.

Figure 2 shows the responses of fertility to a one-standard-deviation innovation in mortality, the responses of mortality to one-standard-deviation impulse in fertility, the responses of fertility

to a one-standard-deviation innovation in log GDP per capita, and the responses of log GDP per capita to one-standard-deviation impulse in fertility over a 50-year horizon; the dashed lines mark plus and minus two standard errors obtained through Monte Carlo simulations using 1,000 replications.

The upper panels focus on the interaction between fertility and mortality. As the left panel shows, mortality has a gradual and permanent effect on fertility that reaches its full impact not before 6 periods (30 years) after the shock (i.e. after about one generation) and that is not statistically significant in the first period. This is consistent with the widespread belief in demography that fertility behavior reacts only gradually on declining mortality. The upper right panel confirms that there is also a significant positive feedback effect of fertility on mortality; it reaches its maximum in the second period.

The bottom panels in Figure 2 show the GDP-fertility interaction. The left panel indicates that fertility gradually and permanently decreases in response to a one-standard-deviation innovation in log GDP per capita and that the response becomes statistically different from zero after the second period. The lower right panel documents that there is also a (delayed) negative effect of fertility on GDP. It becomes statistically significant after about 15 years, i.e. at about the time when the individuals born 15 years ago begin to enter the workforce. This pattern of response of GDP per capita suggests that the dependency effect, which occurs immediately at birth when GDP is subdivided among more people, becomes only significant when it is amplified by the capital dilution effect, which occurs when the extra population enters the workforce. The observed response of GDP is inconsistent with the mechanism proposed by R&D-based growth theory (strictly interpreted). Our results do not support the view that more people cause income per capita to grow.

Taken together and keeping in mind that GDP is measured in logs, the impulse-response pattern confirms empirically – and to our best knowledge for the first time – that the virtuous cycle, which has been stressed so much in development economics and in unified growth theory, does indeed exist: Growth of income per capita leads to reduced fertility, which in turn causes income growth to rise further, which leads to a further decline of fertility etc. Low fertility is both a cause and consequence of successful economic development.

5. Concluding Remarks

Given the available data from the last century our analysis has shown that (1) declining mortality leads to declining fertility, that (2) growth of income per capita leads to declining fertility, that (3) declining mortality per se is insufficient to explain the secular decline of population growth over the last century, and that (4) fertility changes are both cause and consequence of economic development such that the income-fertility interaction provides a virtuous cycle of demo-economic development. We have furthermore shown that these conclusions are robust against alternative estimation methods, potential outliers, sample selection, different measures of mortality, and the sample period.

Under the prospect of perpetual income growth our result that there exists a linear negative relationship between income and fertility (and thus between fertility change and income growth) may appear to be puzzling. After all, fertility and mortality are bounded to be non-negative and cannot continue to fall infinitely with forever rising income. The evidence derived from historical data, however, does not mean that the empirical model predicts a persistence of this association for the (infinite) future.

With income growing further, the association between income growth and fertility change has to become non-linear sooner or later and eventually it must disappear. The correct assessment of our results is that so far (i.e. over the last century) a linear model describes the data adequately, a fact that we have proven with extensive tests. The implied conclusion is thus that the leveling-off of fertility's reaction on income growth is not yet visible in the data. This conclusion is in line with Strulik and Vollmer (2010) who investigate convergence behavior of fertility across countries and find that the end of the fertility transition is not reached even by very rich countries where fertility is considerably below replacement level.

Appendix A1. Key variables by Country over the Sample Period

Figure A.1: The 20 Countries of our Main Sample on A Map

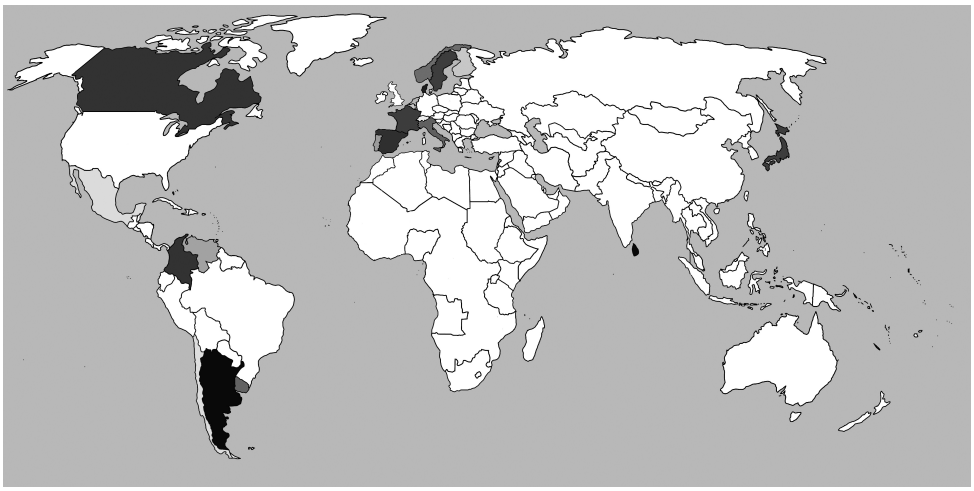
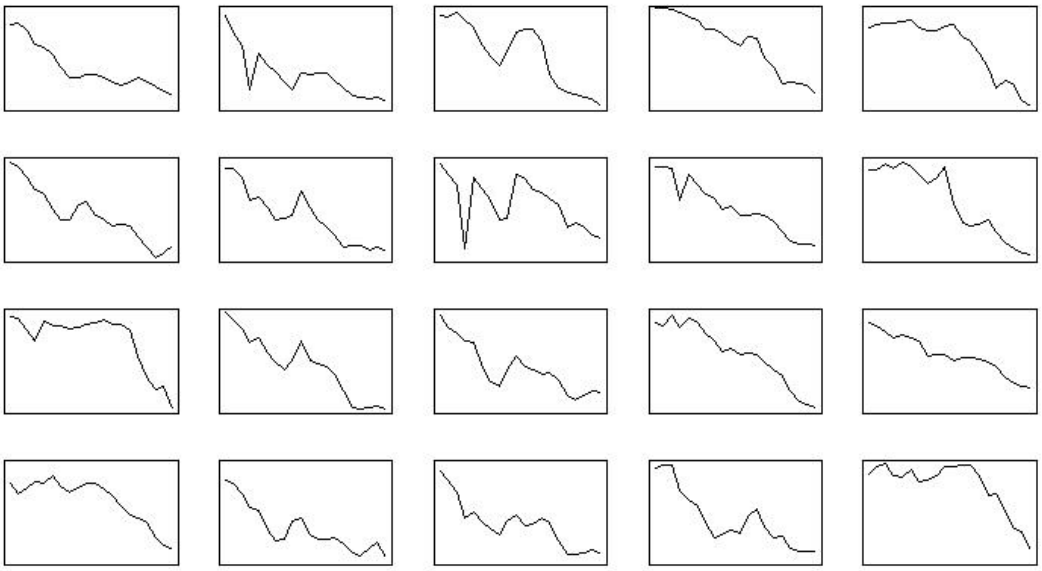
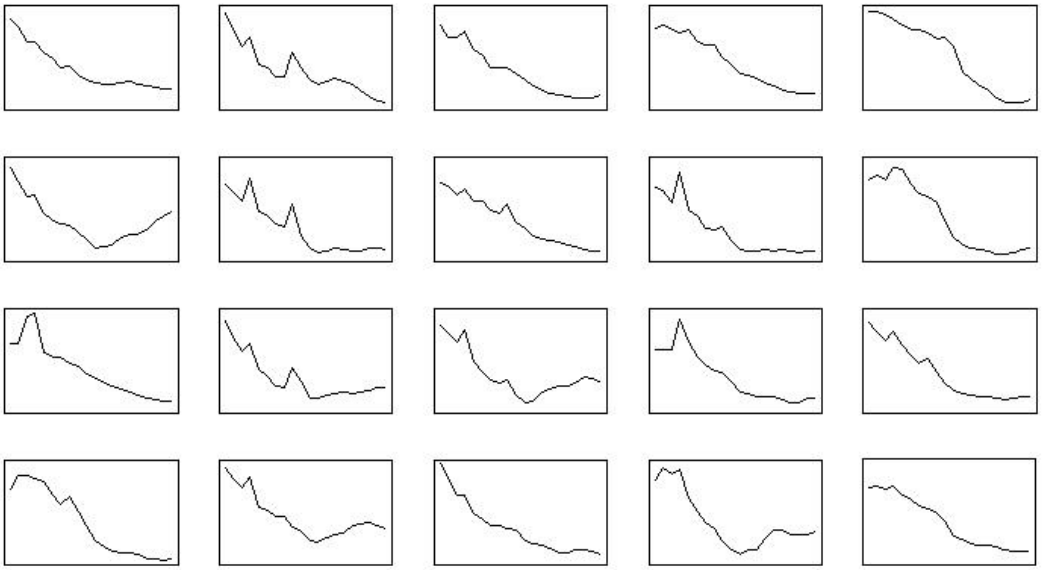


Figure A.2: Fertility by Country over the Period 1900-1999



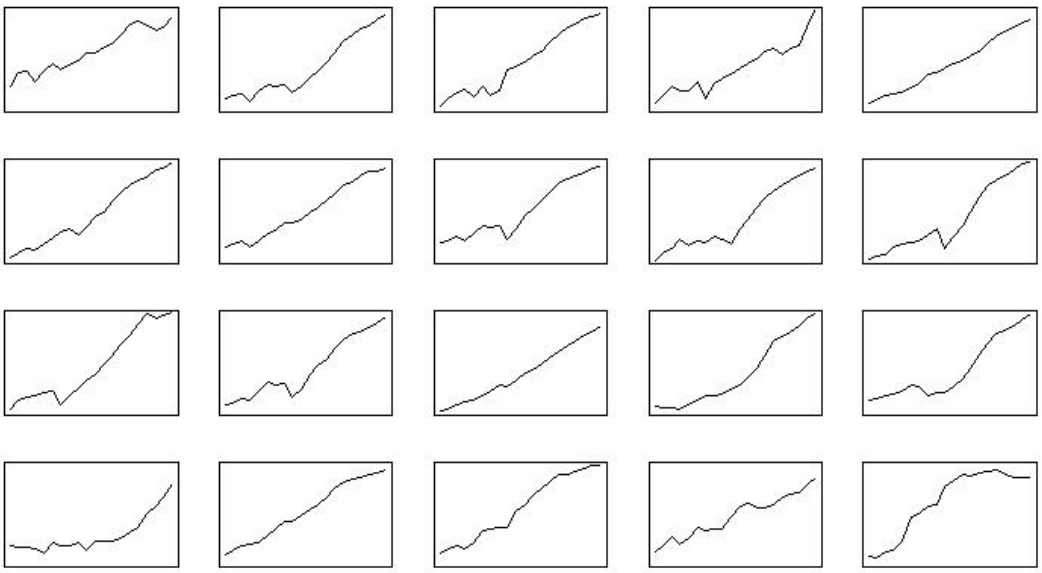
The countries from the left to the right are: Argentina, Belgium, Canada, Chile, Colombia, Denmark, Finland, France, Italy, Japan, Mexico, Netherlands, Norway, Portugal, Spain, Sri Lanka, Sweden, Switzerland, Uruguay, and Venezuela.

Figure A.3: Mortality by Country over the Period 1900-1999



The countries from the left to the right are: Argentina, Belgium, Canada, Chile, Colombia, Denmark, Finland, France, Italy, Japan, Mexico, Netherlands, Norway, Portugal, Spain, Sri Lanka, Sweden, Switzerland, Uruguay, and Venezuela.

Figure A.4: Log GDP per Capita by Country over the Period 1900-1999



The countries from the left to the right are: Argentina, Belgium, Canada, Chile, Colombia, Denmark, Finland, France, Italy, Japan, Mexico, Netherlands, Norway, Portugal, Spain, Sri Lanka, Sweden, Switzerland, Uruguay, and Venezuela.

Appendix A2. Panel Unit-Root Tests

In order to investigate the time-series properties of the data, we use the Levin, Lin, and Chu (2002) (LLC), the Im, Pesaran, and Shin (2003) (IPS), and the cross-sectionally augmented IPS

test of Pesaran (2007). All these tests are based on an augmented Dickey-Fuller (ADF) regression where the variable of interest is observed for $N(= 20)$ cross-sectional units and $T(= 20)$ time periods:

$$\Delta x_{it} = z_{it}\gamma_i + z_{it}\rho_i x_{it-1} + \sum_{j=1}^{k_i} \varphi_{ij}\Delta x_{it-1} + \epsilon_{it}, \quad i = 1, 2, \ldots, N, \quad t = 1, 2, \ldots, T \tag{A.1}$$

where k_i is the lag length, z_{it} is a vector of deterministic terms, such as fixed effects or fixed effects plus individual trends, and γ_i is the corresponding vector of coefficients.

The within-dimension-based LLC panel unit-root test pools the autoregressive coefficient across the countries during the unit-root test and thus restrict the first-order autoregressive parameter to be the same for all countries, $\rho_i = \rho$. Thus, the null hypothesis is that all series contain a unit root, $H_0 : \rho = 0$, while the alternative hypothesis is that no series contains a unit root, $H_1 : \rho = \rho_i < 0$, that is, all are (trend) stationary.

To conduct the LLC-test statistic, the following steps are performed. The first step is to obtain the residuals, $\hat{e}_{it}$, from individual regressions of Δx_{it} on its lagged values (and on z_{it}), $\Delta x_{it} = \sum_{j=1}^{k_i} \theta_{1ij}\Delta x_{it-j} + z_{ij}\gamma_i + e_{it}$. Second, x_{it-1} is regressed on the lagged values of Δx_{it} (and on z_{it}) to obtain $\hat{\nu}_{it-1}$, that is, the (lagged) residuals of this regression,$x_{it} = \sum_{j=1}^{k_i} \theta_{2ij}\Delta x_{it-j} + z_{ij}\gamma_i + \nu_{it}$. In the third step, $\hat{e}_{it}$ is regressed on $\hat{\nu}_{it-1}$, $\hat{e}_{it} = \delta\hat{\nu}_{it-1} + \xi_{it}$. The standard error, $\hat{\sigma}^2_{ei}$, of this regression is then used to normalize the residuals $\hat{e}_{it}$ and $\hat{\nu}_{it-1}$ (to control for heterogeneity in the variances of the series), $\tilde{e}_{it} = \hat{e}_{it}/\hat{\sigma}^2_{ei}$, $\tilde{\nu}_{it-1} = \hat{\nu}_{it-1}/\hat{\sigma}^2_{ei}$. Finally, ρ is estimated from a regression of $\tilde{e}_{it}$ on $\tilde{\nu}_{it-1}$, $\tilde{e}_{it} = \rho\tilde{\nu}_{it-1} + \xi_{it}$. The conventional t-statistic for the autoregressive coefficientρ has a standard normal limiting distribution if the underlying model does not include fixed effects and individual time trends (z_{it}). Otherwise, this statistic has to be corrected using the first and second moments tabulated by Levin et al. (2002) and the ratio of the long-run variance to the short-run variance, which accounts for the nuisance parameters present in the specification. The limiting distribution of this corrected statistic is normal as $N \to \infty$ and $T \to \infty$.

In contrast to the LLC test, the between-dimension-based IPS panel unit-root test allows the first-order autoregressive parameter to vary across countries by estimating the ADF equation separately for each country. Thus, the null hypothesis is that each series contains a unit-root, $H_0 : \rho_i = 0$ for all i, while the alternative hypothesis is that at least one of the individual series in the panel is (trend) stationary, $H_1 : \rho_i < 0$ for at least one i. H_0 is tested against H_1 using

the standardized t-bar test statistic

$$\Gamma_i = \frac{\sqrt{N}\left[\bar{t}_{NT} - \mu\right]}{\sqrt{\nu}} \tag{A.2}$$

where $\bar{t}_{NT}$ is the average of the N cross-section ADF t statistics, and μ and ν are, respectively, the mean and variance of the average of the individual t-statistics, tabulated by Im et al. (2003). The standardized t-bar statistic converges to a standard normal distribution as N and $T \to \infty$.

However, both the LLC and the IPS test procedures assume cross-sectional independence and thus may lead to spurious inference if the errors, ϵ_{it}, are not independent across i. Therefore, we also use the cross-sectionally augmented IPS test, which allows for cross-sectional dependence by augmenting the ADF regression with the cross-section averages of lagged levels and first-differences of the individual series. An attractive feature of this test is that it permits the individual countries to respond differently to the common time effects as reflected by the country-specific coefficients on the cross-section averages of the variables. The cross-section augmented ADF (CADF) regression, carried out separately for each country, is given by

$$\Delta x_{it} = z_{it}\gamma_i + \rho x_{it-1} + \sum_{j=1}^{k_i} \varphi_{ij}\Delta x_{it-j} + \alpha_i \bar{x}_{t-1} + \sum_{j=0}^{k_i} \eta_{ij}\Delta \bar{x}_{t-j} + \nu_{it} \tag{A.3}$$

where $\bar{x}_t$ is the cross-section mean of x_{it}, $\bar{x}_t = N^{-1}\sum_{i=1}^{N} x_{it}$. The cross-section augmented IPS statistic is a simple average of t_i defined by

$$CIPS = N^{-1}\sum_{i=1}^{N} t_i \tag{A.4}$$

where t_i is the OLS t ratio of ρ_i in the above CADF regression. Critical values are tabulated by Pesaran (2007).

Table A1 reports the results of these tests for the variables in levels and in first differences. As can be seen, all three test statistics are unable to reject the null hypothesis that $fert_{it}$, $mort_{it}$, and $\log(gdp_{it})$ have a unit-root in levels. Since the unit-root hypothesis can be rejected for the first differences, it can be concluded that all series are integrated of order one, $I(1)$.

Appendix A3. Panel Cointegration Tests

We use several panel cointegration test procedures to determine whether there is a long-run relationship between fertility, mortality, and economic development. The first is the two-step

Table A.1: Panel Unit Root Tests

Variables	Deterministic terms	LLC statistics	IPS statistics	CIPS statistics
Levels				
$fert_{it}$	c, t	0.46	-0.72	-2.08
$mort_{it}$	c, t	2.1	4.06	-2.3
$\log(gdp_{it})$	c, t	-0.46	0.4	-2.32
First differences				
$\Delta fert_{it}$	c	-6.38^{**}	-6.17^{**}	-2.55^{**}
$\Delta mort_{it}$	c	-1.75^{*}	-3.49^{**}	-2.41^{**}
$\Delta \log(gdp_{it)}$	c	-2.71^{**}	-3.31^{**}	-2.44^{**}

c (t) indicates that we allow for different intercepts (and time trends) for each country. Two lags were selected to adjust for autocorrelation. The relevant 1% (5%) critical value for the CIPS statistics is -2.92 (-2.73) with an intercept and a linear trend, and -2.40 (-2.21) with an intercept. $**$ ($*$) denote significance at the 1% (5%) level.

residual-based procedure suggested by Pedroni (1999, 2004), which can be intuitively described as follows. In the first step, the hypothesized cointegrating regression

$$fert_{it} = a_i + \beta_{1i} mort_{it} + \beta_{2i} \log(gdp_{it}) + \epsilon_{it} \tag{A.5}$$

is estimated separately for each country, thus allowing for heterogeneous cointegrating vectors. In the second step, the residuals, $\hat{\epsilon}_{it}$, from these regressions are tested for stationarity. To test the null hypothesis of non-stationarity (or no cointegration) Pedroni proposes seven statistics. Here, we employ the two statistics with the highest power for small T-panels like ours: the panel ADF and group ADF statistics (see, e.g., Pedroni, 2004, Wagner and Hlouskova, 2010). The former is analogous to the LLC (2002) panel unit root test, while the latter is analogous to the IPS (2003) panel unit root test (both discussed above). The standardized distributions for the test statistics are given by

$$\kappa = \frac{\varphi - \mu\sqrt{N}}{\sqrt{\nu}} \quad \Rightarrow N(0,1). \tag{A.6}$$

where φ is the respective ADF panel or group ADF statistic, and μ and ν are the expected mean and variance of the corresponding statistic, tabulated by Pedroni (1999).

In addition, we use the panel cointegration tests developed by Kao (1999). Kao follows basically the same approach as Pedroni (1999, 2004), but constrains the cointegrating coefficients to be homogeneous across countries by employing a within regression of the form

$$fert_{it} = a_i + \beta_1 mort_{it} + \beta_2 \log(gdp_{it}) + e_{it}. \tag{A.7}$$

To test the stationarity of the residuals, $\hat{e}_{it}$, from this regression Kao presents four within-dimension-based DF test statistics and one within-dimension-based ADF statistic: The first two DF statistics, DF_ρ and DF_t, as well as the ADF statistic, assume strict exogeneity of the regressors, while the other two DF-type tests, DF_ρ^* and DF_t^*, do not require this assumption. DF_ρ and DF_ρ^* are calculated based on the estimated first-order autoregressive coefficient in the panel DF regression; the associated t-statistic is used in calculating DF_t and DF_t^*.

The problem with these two approaches is that they do not take into account potential error cross-sectional dependence, which could bias the results. To test for cointegration in the presence of possible cross-sectional dependence we use the two-step residual-based procedure suggested by Holly et al. (2010), who apply the common correlated effects (CCE) estimator of Pesaran (2006) in the first-step regression. Like the cross-sectionally augmented IPS test, the CCE estimator allows for cross-sectional dependencies that potentially arise from multiple unobserved common factors and permits the individual responses to these factors to differ across countries. In our case, the cross-section augmented cointegrating regression (for the ith cross-section) is given by

$$fert_{it} = a_i + \beta_{1i} mort_{it} + \beta_{2i} \log(gdp_{it}) + g_{1i}\overline{fert_t} + g_{2i}\overline{mort_t} + g_{3i}\overline{\log(gdp_t)} + \xi_{it} \qquad \text{(A.8)}$$

where the cross-section averages $\overline{fert_t} = N^{-1}\sum_{i=1}^{N} fert_{it}$, $\overline{mort_t} = N^{-1}\sum_{i=1}^{N} mort_{it}$ and $\overline{\log(gdp_t)} = N^{-1}\sum_{i=1}^{N}\log(gdp_{it})$ serve as proxies for the unobserved factors. In the second step, we compute the cross-section augmented IPS statistic for the residuals from the individual CCE long-run relations $\hat{\mu} = fert_{it} - \hat{\beta}_{1i} mort_{it} - \hat{\beta}_{2i}\log(gdp_{it})$, including an intercept. In doing so, we account for unobserved common factors that could be correlated with the observed regressors in both steps.

However, residual-based (panel) cointegration tests restrict the long-run elasticities to be equal to the short-run elasticites. If this restriction is invalid, residual-based (panel) cointegration tests may suffer from low power (see, e.g., Westerlund, 2007). Another drawback of single-equation, residual-based (panel) cointegration tests is that they are generally not invariant to the normalization of the cointegrating regression, and, moreover, such tests are unable to identify more than one cointegrating relationship in systems with more than two variables. Therefore, we also use the Larsson et al. (2001) procedure, which is based on Johansen's (1988) system approach. Like the Johansen time-series cointegration test, the Larsson et al. panel test treats all variables as potentially endogenous, thus avoiding normalization problems inherent in

residual-based cointegration tests. In addition, the Larsson et al. procedure allows the long-run elasticities to differ from the short-run elasticities and hence does not impose a possibly invalid common factor restriction. Finally, an important feature of the Larsson et al. approach is that it allows the determination of the number of cointegrating vectors.

TABLE A.2 COINTEGRATION TESTS

Pedroni (1999, 2004)			
Panel ADF t-statistic		-3.82^{**}	
Group ADF t-statistic		-3.24^{**}	
Kao (1999)			
DF_ρ statistic		-3.38^{**}	
DF_t statistic		-2.45^{**}	
ADF_t -statistic		-3.38^{**}	
DF_ρ^* statistic		-3.69^{**}	
DF_t^* statistic		-2.67^{**}	
Holly et al. (2010)			
CIPS statistic		-2.47^{**}	
Larsson et al. (2001)		Cointegration rank	
	r = 0	r = 1	r = 2
Standardized panel trace statistics	4.70^{**}	0.01	1.51

$**$ indicate a rejection of the null of no cointegration at the one percent level. The relevant 1% critical value for the CIPS statistic is -2.40. All other test statistics are asymptotically normally distributed. The right tail of the normal distribution is used to reject the null hypothesis in the standardized panel trace statistics as recommended by Wagner and Hlouskova (2010), while the left tail is used for the other statistics. The number of lags in the ADF tests was determined by the Schwarz criterion with a maximum number of four lags. For the Larsson et al. (2001) technique we used one lag.

The Larsson et al. approach involves estimating the Johansen vector error-correction model for each country separately:

$$\Delta y_{it} = \Pi_i y_{it-1} + \sum_{i=1}^{k_i} \Gamma_{ik} \Delta y_{it-k} + z_{it}\gamma_i + \epsilon_{it} \tag{A.9}$$

where y_{it} is a $p \times 1$ vector of endogenous variables ($y_{it} = [fert_{it}, mort_{it}, \log(gdp_{it}]'$); p is the number of variables) and Π_t is the long-run matrix of order $p \times p$. If Π_i is of reduced rank, $r_i < p$, it is possible to let $\Pi_i = \alpha_i \beta_i$, where β_i is a $p \times r_i$ matrix, the r_i columns of which represent the cointegrating vectors, and α_i is a $p \times r_i$ matrix whose p rows represent the error correction coefficients. The null hypothesis is that all of the N countries in the panel have a common cointegrating rank, i.e. at most r (possibly heterogeneous) cointegrating relationships among the p variables: $H_0 : \text{rank}(\Pi_i) = r_i \leq r$ for all $i = 1, \ldots, N$, whereas the alternative hypothesis is that all the cross-sections have a higher rank: $H_1 : \text{rank}(\Pi_i) = p$ for all $i = 1, \ldots, N$. To test

H_0 against H_1, a panel cointegration rank trace-test statistic is computed by calculating the average of the individual trace statistics, $LR_{iT}\{H(r)|H(p)\}$.

$$\overline{LR}_{NT}\{H(r)|H(p)\} = \frac{1}{N}\sum_{i=1}^{N} LR_{iT}\{H(r)|H(p)\} \tag{A.10}$$

and then standardizing it as follows:

$$\Psi_{\overline{LR}}\{H(r)|H(p)\} = \frac{\sqrt{N}\left[\overline{LR}_{NT}\{H(r)|H(p)\} - E(Z_k)\right]}{\sqrt{Var(Z_k)}} \quad \Rightarrow N(0,1). \tag{A.11}$$

The mean $E(Z_k)$ and variance $Var(Z_k)$ of the asymptotic trace statistic are tabulated by Breitung (2005) for the model we use (the model with a constant in the cointegrating vector and a linear trend in the data). However, a well-known problem is that the Johansen trace statistics tend to over-reject the null in small samples. To avoid the Larsson et al. test also overestimating the cointegrating rank, we compute the standardized panel trace statistics based on small-sample corrected country-specific trace statistics. More specifically, to adjust the individual trace statistics we use the small-sample correction factor suggested by Reinsel and Ahn (1992):

$$LR_{iT}\{H(r)|H(p)\} \times \left[\frac{T - k_i \times p}{T}\right]. \tag{A.12}$$

The results of these tests are presented in Table A2. As can be seen, all tests strongly suggest that $fert_{it}$, $mort_{it}$, and $\log(gdp_{it})$ are cointegrated. The standardized trace statistics clearly supports the presence of one cointegrating vector. Also, the CIPS, the Kao, and the Pedroni statistics reject the null hypothesis of no cointegration at the 1 level, implying that there exists a single long-run relationship between fertility, mortality, and economic development.

Appendix A4. Countries in the sample for Table 3

Algeria, Angola, Argentina, Austria, Bahrain, Bangladesh, Belgium, Benin, Bolivia, Botswana, Brazil, Bulgaria, Burkina Faso, Burundi, Cambodia, Cameroon, Canada, Central African Republic, Chad, Chile, China, Colombia, Comoros, Congo (Dem. Rep.), Congo (Rep.), Costa Rica, Cote d'Ivoire, Cuba, Denmark, Djibouti, Dominican Republic, Ecuador, Egypt, El Salvador, Finland, France, Gabon, Gambia, Germany, Ghana, Guatemala, Guinea, Guinea Bissau, Haiti, Honduras, Hungary, India, Indonesia, Iran, Iraq, Israel, Italy, Jamaica, Japan,

Jordan, Kenya, Korea (Republic), Kuwait, Laos, Lebanon, Lesotho, Liberia, Libya, Madagascar, Malawi, Malaysia, Mali, Mauritania, Mauritius, Mexico, Mongolia, Morocco, Mozambique, Myanmar, Namibia, Nepal, Netherlands, Nicaragua, Niger, Nigeria, North Korea , Norway, Oman, Panama, Paraguay, Peru, Philippines, Portugal, Puerto Rico, Qatar, Romania, Rwanda, Saudi Arabia, Senegal, Seychelles, Sierra Leone, Singapore, Somalia, South Africa, Spain, Sri Lanka, Sudan, Swaziland, Sweden, Switzerland, Syria, Tanzania, Thailand, Togo, Trinidad and Tobago, Tunisia, Uganda, United States, Uruguay, Venezuela, Vietnam, Yemen, Zambia, and Zimbabwe.

References

Acemoglu, D., and S. Johnson, 2007, Disease and development: The effect of life-expectancy on economic growth, *Journal of Political Economy* 115, 925-985.

Ahituv, A., 2001, Be fruitful or multiply: On the interplay between fertility and economic development, *Journal of Population Economics* 14, 51-71.

Angeles, L., 2010, Demographic transitions: analyzing the effects of mortality on fertility, *Journal of Population Economics* 23, 99-120.

Ashraf, Q. and Galor, O., 2011, Dynamics and stagnation in the Malthusian epoch, *American Economic Review*, forthcoming.

Arellano, M., Bond, S.R., 199, Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations *Review of Economic Studies* 58, 277-297.

Angeles, L., 2010, Demographic transitions: analyzing the effects of mortality on fertility ,*Journal of Population Economics* 23, 99-120.

Baltagi, B.H., Griffin, J.M., 1997, Pooled estimators vs. their heterogeneous counterparts in the context of dynamic demand for gasoline, *Journal of Econometrics* 77, 303-327.

Baltagi, B., Kao, C., 2000, Nonstationary panels, cointegration in panels, and dynamic panels: a survey. *Advances in Econometrics* 15, 7-52.

Banerjee, A., 1999, Panel data unit roots and cointegration: an overview, *Oxford Bulletin of Economics and Statistics* 61, 607-29.

Becker, G.S., 1960, An economic analysis of fertility. In: National Bureau of Economic Research (ed), *Demographic and Economic Change in Developed Countries*, Princeton University Press, Princeton, 209-231.

Becker, G.S., 1965, A theory of the allocation of time, *Economic Journal* 75, 493-517.

Becker, G.S. and Lewis, H.G., 1973, On the Interaction between the Quantity and Quality of Children, *Journal of Political Economy* 8, 279-288.

Becker, G.S. and R.J. Barro, 1988, A Reformulation of the theory of fertility, *Quarterly Journal of Economics* 103, 1-25.

Becker, G., K. Murphy and R. Tamura, 1990, Human Capital, fertility and growth, *Journal of Political Economy* 98, S12-37.

Brander, J.A. and Dowrick, S., 1994, The role of fertility and population in economic growth, *Journal of Population Economics* 7, 1-25.

Breitung, J., 2005, A parametric approach to the estimation of cointegrating vectors in panel data, *Econometric Reviews* 24, 151-173.

Cervellati, M., and Sunde, U., 2007, Human capital, mortality and fertility: A unified theory of the economic and demographic transition, IZA Discussion Paper 2905.

Cervellati, M. and Sunde, U., 2009, Life Expectancy and Economic Growth: The Role of the Demographic Transition, Institute for the Study of Labor (IZA), Discussion paper DP 4016.

Clark, G., 2005, Human capital, fertility and the industrial revolution, *Journal of the European Economic Association* 3, 505-515.

Doepke, M., 2005, Child mortality and fertility decline: does the Barro-Becker model fit the facts?, *Journal of Population Economics* 18, 337-366.

Eckstein, Z. and Schultz, T.P. and Wolpin, K.I., 1984, Short-run fluctuations in fertility and mortality in pre-industrial Sweden, *European Economic Review* 26, 295-317.

Eckstein, Z. Mira, P., and Wolpin, K., 1999, A quantitative analysis of swedisch fertility dynamics, *Review of Economic Dynamics* 2, 137-165.

Galloway, P.R., 1988, Basic patterns in annual variations in fertility, nuptiality, mortality, and prices in pre-industrial Europe, *Population Studies* 42, 275-303.

Galor, O., 2005, From stagnation to growth: unified growth theory, in: *Handbook of Economic Growth*, Amsterdam: North-Holland.

Galor, O. and Weil, D.N., 1996, The gender gap, fertility, and growth, *American Economic Review* 86, 374-387.

Galor, O. and D. Weil, 2000, Population, technology and growth: From Malthusian stagnation to the demographic transition and neyond, *American Economic Review* 90, 806-828.

Galor, O., and O. Moav, 2002. Natural selection and the origin of economic growth. *Quarterly Journal of Economics*, 117, 11331191

Greenwood, J. and Seshadri, A. and Vandenbroucke, G., 2005, The baby boom and baby bust, *American Economic Review* 95, 183-207.

Hall, S.G., Milne. A., 1994, The relevance of p-star analysis to UK monetary policy. *Economic Journal* 104, 597-604.

Hendry, D.F., 1995, *Dynamic Econometrics*, Oxford University Press, Oxford.

Herzer, D., 2008, The long-run relationship between outward FDI and domestic output: evidence from panel data, *Economics Letters* 100, 146-149.

Holly, S., Pesaran, M.H., Yamagata, T., 2010, A spatio-temporal model of house prices in the US. *Journal of Econometrics* 158, 160-173.

Im, K.S., Pesaran, M.H., Shin, Y., 2003, Testing for unit roots in heterogeneous panels, *Journal of Econometrics* 115, 53-74.

Johansen, S., 1988, Statistical analysis of cointegrating vectors, *Journal of Economic Dynamics and Control* 12, 231-254.

Johansen, S., 2000, Modelling of cointegration in the vector autoregressive model, *Economic Modeling* 17, 359-73.

Jones, C.I., 1995, R&D-based models of economic growth, *Journal of Political Economy* 103, 759-784.

Jones, C.I., 2003, Population and ideas: a theory of endogenous growth, in: Aghion, P., R. Frydman; J. Stiglitz and M. Woodford (eds.) *Knowledge, Information and Expectations in Modern Macroeconomics*, Princeton, New Jersey: Princeton University Press. 498-521

Juselius, K., 2001, European integration and monetary transmission mechanisms: the case of Italy, *Journal of Applied Econometrics* 16, 314-358.

Kalemli-Ozcan, S., 2002, Does the mortality decline promote economic growth, *Journal of Economic Growth* 7, 411-439.

Kao, C., 1999, Spurious regression and residual-based tests for cointegration in panel data, *Journal of Econometrics* 90, 1-44.

Kao, C., Chiang, M.H., 2000, On the estimation and inference of a cointegrated regression in panel data, *Advances in Econometrics* 15, 179 -222.

Kögel, T., and A. Prskawetz, 2001, Agricultural productivity growth and escape from the Malthusian trap, *Journal of Economic Growth* 6, 337-357.

Larsson, R., Lyhagen, J., Lthegren, M., 2001. Likelihood-based cointegration tests in heterogeneous panels. Econometrics Journal 4, 109-142.

Lee, R. and M. Anderson, 2002, Malthus in State Space: Macro Economic-demographic Relations in English History, 1540 to 1840, *Journal of Population Economics* 15, 195-220.

Levin, A., Lin, C.-F., Chu, C.-S.J., 2002, Unit root test in panel data: asymptotic and finite-sample properties, *Journal of Econometrics* 108, 1-24.

Lütkepohl, H., 2007, General-to-specific or specific-to-general modelling? An opinion on current econometric terminology, *Journal of Econometrics* 136, 319-324.

Lütkepohl, H., and Reimers, H.E., 1992, Impulse response analysis of cointegrated systems, *Journal of Economic Dynamics and Control* 16, 53-78.

Lütkepohl, H., and Wolters, J., 1998, A money demand system for German M3, *Empirical Economics* 23, 371-386.

Lütkepohl, H., and Wolters, J., 2003, Transmission of German monetary policy in the pre-Euro period, *Macroeconomic Dynamics* 7, 711-733.

Maddison, A., 2003, *The World Economy: Historical statistics*, OECD, Paris.

Mankiw, N.G., Romer, D., and Weil, D.N., 1992, A contribution to the empirics of economic growth, *Quarterly Journal of Economics* 107, 407-437.

Nicolini, E., 2007, Was Malthus right? A VAR analysis of economic and demographic interactions in pre-industrial England, *European Review of Economic History* 11, 99-121.

Pedroni, P., 1999, Critical values for cointegration tests in heterogeneous panels with multiple regressors, *Oxford Bulletin of Economics and Statistics* 61, 653-670.

Pedroni, P., 2001, Purchasing power parity tests in cointegrated panels, *Review of Economics and Statistics* 83, 727-731.

Pedroni, P., 2004, Panel cointegration: Asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis, *Econometric Theory* 20, 597-625.

Pedroni, P., 2007, Social capital, barriers to production and capital shares: implications for the importance of parameter heterogeneity from a nonstationary panel approach, *Journal of Applied Econometrics* 22, 429-451.

Pesaran, M.H., 2006, Estimation and inference in large heterogeneous panels with a multifactor error structure, *Econometrica* 74, 967-1012.

Pesaran, M.H., 2007, A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics* 22, 265-312.

Pesaran, M.H., and Shin, Y., 1998. Generalized impulse response analysis in linear multivariate models, *Economics Letters* 58, 17-29.

Phillips, P.C.B, and Hansen, B.E., 1990, Statistical inference in instrumental variables regression with I(1) processes, *Review of Economic Studies* 57, 99-125.

Reher, D.S., 2004, The demographic transition revisited as a global process, *Population Space and Place* 10, 19-41.

Reinsel, G.C., Ahn, S.K., 1992, Vector autoregressive models with unit roots and reduced rank structure: estimation, likelihood ratio test and forecasting, *Journal of Time Series Analysis* 13, 353-375.

Romer, P.M., 1990, Endogenous technological change, *Journal of Political Economy* 98, S71-S102.

Sah, R.K., 1991, The effects of child mortality changes on fertility choice and parental welfare, *Journal of Political Economy* 99, 582-606.

Schultz, T.P., 1997, Demand for children in low income countries, in: M.R. Rosenzweig and O. Stark (eds.) *Handbook of Population and Family Economics* 1, 349-430.

Soares, R.R., 2005, Mortality reductions, educational attainment, and fertility choice, *American Economic Review* 95, 580-601.

Solow, R.M., 1956, A Contribution to the Theory of Economic Growth, Quarterly Journal of Economics 70, 65-94.

Stock, J.H., Watson, M.W., 1993, A simple estimator of cointegrating vectors in higher-order integrated systems, *Econometrica* 61, 783-820.

Strulik, H., 2008, Geography, health, and the pace of demo-economic development, *Journal of Development Economics* 86, 61-75.

Strulik, H. and Vollmer, S., 2010, The fertility transition around the world – 1950-2005, Discussion Paper, Harvard University.

Strulik, H. and Weisdorf, J., 2008, Population, food, and knowledge: A simple unified growth theory, *Journal of Economic Growth* 13, 169-194.

Urbain, J.-P., 1995, Partial versus full system modelling of cointegrated systems: an empirical illustration, *Journal of Econometrics* 69, 177-210.

Wagner, M., and Hlouskova, K., 2010, The performance of panel cointegration methods: results from a large scale simulation study, *Econometric Reviews* 29, 182-223.

Wang, P., Yip, C.K. and Scotese, C.A., 1994, Fertility choice and economic growth: Theory and evidence, *The Review of Economics and Statistics* 76, 255-266.

Westerlund, J., 2007, Testing for error correction in panel data, *Oxford Bulletin of Economics and Statistics* 69, 709-748.