Received: 6 May 2023 | Accepted: 23 February 2024

DOI: 10.1111/ecin.13216

ORIGINAL ARTICLE

Economic Inquiry

Fertility and long-term economic growth

Kaixing Huang

Peking University, Beijing, China

Correspondence
Kaixing Huang.
Email: kaixing.huang@pku.edu.cn

Funding information
National Natural Science Foundation of China

Abstract

Empirical studies have generally found that higher fertility has a negative or insignificant effect on economic growth. This article argues that this is because existing studies have failed to capture the long-term lagged effects of fertility. By estimating a long-term lagged panel model using data from 137 countries, I find that higher fertility first reduces and then increases economic growth, and the long-term average effect is significantly positive. This finding is robust when focusing on countries at different development levels, exploiting exogenous fertility shocks from global family planning campaigns, and capitalizing on within-country fertility variation resulting from China's one-child policy.

KEYWORDS

economic growth, fertility, long-term effect

JEL CLASSIFICATION

J11, J13, O47

1 | INTRODUCTION

Declining fertility is among the most salient features of global demography. Figure 1 shows the dramatic decline in the total fertility rate (TFR) in countries with different income levels over the past half-century. What is the impact of secular fertility declines on long-term economic growth?[1] The answer to this question is of great interest not only to demographers and economists but also to policymakers devising population policies. While many high-income countries have been adopting pro-natalist policies to increase their birth rates, most low-income countries have adopted anti-natalist policies to curb the birth rate (United Nations, 2015). The paper's main goal is to analyze the time-varying, long-term average effect of fertility rates on economic growth across countries. Additionally, the paper will examine the differential effects of fertility on economic growth among countries with varying development levels.

To answer these questions, one has to address two major identification issues. The first is endogenous fertility. Compelling evidence suggests the reverse causality, indicating that higher incomes reduce fertility rates (Chatterjee & Vogl, 2018; Herzer et al., 2012). It is also well-recognized that fertility decisions are jointly made with other choices influencing economic outcomes (Becker & Lewis, 1973; Galor & Weil, 1996). A decisive answer cannot be provided until the endogeneity bias from reverse causality and joint determination is sufficiently addressed. The second issue is estimating the long-term effects instead of the short-term ones. Individuals interact differently with the economy over different life-cycle stages; thus, the effect of fertility may last for decades and differ substantially in the short term versus the long term (Bloom et al., 2009; Simon, 1989).

Abbreviations: 2SLS, Two-Stage Least Squares; DID, difference-in-differences; GDP, Gross Domestic Product; IV, Instrumental Variable; OCP, one-child policy; R&D, Research and Development; TFR, total fertility rate; WTO, World Trade Organization.

Managing Editor: Dietrich Vollrath

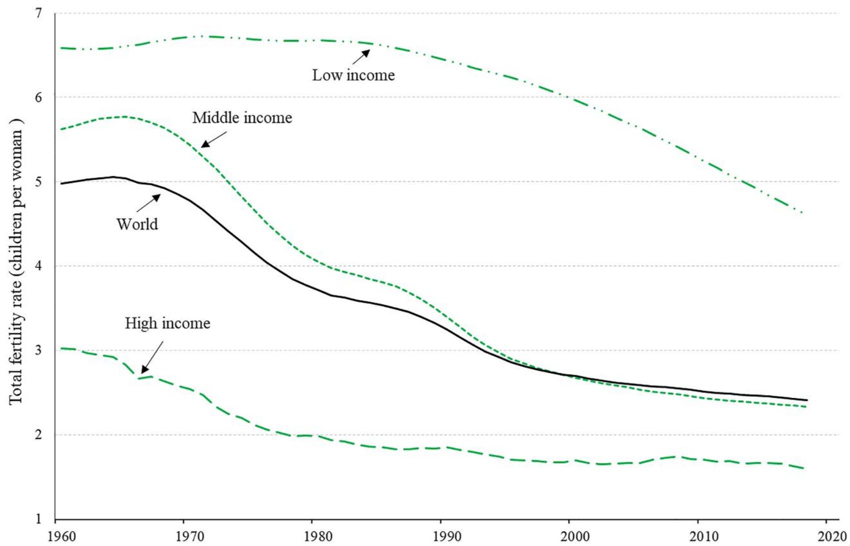

FIGURE 1 Fertility Trends. The figure classifies countries as high-, middle-, or low-income based on the 2016 gross national income per capita, following the World Bank's definition. The data are derived from the World Development Indicators.

However, while early empirical studies have failed to address the endogeneity bias, more recent studies have failed to capture the long-term effects of fertility. Early studies usually adopted a cross-sectional regression model comparing inter-country differences in fertility rates and growth and generally concluded that countries with lower fertility grow faster (see surveys from Kuznets (1967) and Kelley (1988)). Although they have the potential to capture the long-term effects by comparing long-term inter-country differences in fertility rates, the cross-sectional estimates could be primarily driven by endogeneity bias from omitted variables and reverse causality. To address the endogeneity bias, later studies have usually adopted a panel model with location-fixed effects and have generally found a smaller and sometimes statistically insignificant negative effect of fertility rates on growth (see surveys from Simon (1992), Kelley and Schmidt (1994), and Schultz (2008)). As the location-fixed effects eliminate most of the long-term cross-sectional differences in fertility rates, the panel fixed-effect estimate primarily captures the short-term effects of fertility rates (see Appendix C). Some studies have attempted to further address the endogeneity bias using instrumental variables in the fixed-effects panel model (e.g., Li & Zhang, 2007), but doing so does not improve the capability of capturing long-term effects.[2]

This study attempts to identify the long-term average effect of fertility rates while addressing the endogeneity bias. Based on data from 137 countries from 1960 to 2016, this study has estimated a series of panel models, each of which regresses the current income growth rate on the fertility rate lagged by different years, ranging from zero to decades. The endogeneity bias is addressed by fixed effects and by an instrument variable (IV) constructed from national birth control policies. The coefficients of the current and lagged fertility rates from the model reveal the dynamic effects of fertility rates on growth, and the average of these coefficients reflects the long-term average effect.

The estimates suggest that while the short-term effects of higher fertility rates on economic growth are significantly negative, the long-term lagged effects are significantly positive, and the long-term positive effects dominate the short-term negative effects, resulting in a significantly positive long-term average effect. This finding is robust to using an excluded IV constructed from the birth control policies, omitted variables, serial correlation, sub-samples, and measures of fertility and growth. The same effect pattern is found when focusing respectively on low-, middle-, and high-income countries and when exploiting within-country exogenous fertility variation from China's one-child policy (OCP). The baseline estimates suggest that a one-unit increase in the TFR would raise the GDP per capita by 5.0% for an average sample country over 4 decades, and that the secular fertility declines reduced global GDP per capita by 9.1% over the same period. This study also finds that the estimated long-term average effect of fertility is more positive in more developed countries.

The main findings of this study—a higher fertility rate first reduces and then increases economic growth, and the long-term average effect is significantly positive—have three major implications. First, the results imply that secular fertility declines represent a strong force driving down long-term economic growth. This implication is in sharp contrast with the conventional view that there is a virtuous cycle between fertility decline and economic growth. The observation that growth spurts are often associated with demographic transitions led to the hypothesis that fertility decline promotes economic growth.[3] Combining this hypothesis with the fact that higher income leads to lower fertility rates, a virtuous cycle emerges: growth of income per capita leads to reduced fertility rates, which in turn causes income growth to rise further, which leads to a further decline of fertility rates, and so forth. This virtuous cycle has been stressed in development economics. The finding of this article, however, does not support the existence of this cycle. Instead, it suggests the existence of a long-term equilibrium between fertility rates and growth: higher income leads to a lower fertility rate, which in turn reduces income growth.

Second, the finding of this study provides strong evidence supporting the scale effect prediction of the R&D-based growth models. First-generation models of R&D-based growth (Aghion & Howitt, 1992; Grossman & Helpman, 1991; Romer, 1990) predict a strong scale effect in which a larger population leads to faster economic growth. Later contributions eliminate the strong scale effect by assuming either lower inter-temporal knowledge spillovers (Jones, 1995), increasing difficulty of R&D (Kortum, 1997; Segerstrom, 1998), or a diluting effect of product proliferation (Howitt, 1999; Peretto, 1998). Nevertheless, even in these frameworks, a weak scale effect is still present—faster population growth increases economic growth. Although the scale effect prediction is naturally derived from the R&D-based growth models, empirical evidence supporting this prediction is scarce. This study shows that once an estimation method that can capture the long-term lagged causal effects of fertility rates is adopted, strong evidence for the scale effect can be found. Therefore, in line with the seminal work by Kremer (1993) and Jones (1995), this study contributes to establishing a positive link between population and economic growth in the long run.

Finally, the finding of this study is useful for evaluating the economic impact of family planning programs. The Malthusian fears that rapid population growth hinders income growth have led many developing countries to adopt family planning programs (Coale & Hoover, 1958; Ehrlich, 1968). The number of countries that adopted family planning programs reached 95 by 1976 and increased to 160 by 2013 (United Nations, 2015), with birth control efforts still ongoing in the developing world (Kuang & Brodsky, 2016; Stover & Sonneveldt, 2017). Many recent studies have evaluated the impact of family planning interventions on development and well-being (e.g., Ashraf et al., 2014; Cavalcanti et al., 2020). Existing evidence on the impact of family planning programs on economic growth largely comes from family-level studies, which generally find that families with fewer children have higher per capita incomes. However, micro-level studies often do not capture the positive spillovers associated with population growth because these effects are economy-wide (Dasgupta, 1995). This study presents macro-level evidence suggesting that, under some conditions,[4] the lower fertility caused by a family planning program may, nevertheless, reduce long-run economic growth. The finding of this study should not be interpreted as evidence denying the contribution of family planning programs. The primary target of family planning programs is not to promote long-run economic growth but to reduce unplanned pregnancies and births, support sexual and reproductive rights, or provide assistance to couples in realizing their family plans (Cleland et al., 2006).

2 | CONCEPTUAL FRAMEWORK

To comprehend the long-term dynamic effects of fertility on growth, this section briefly summarizes the theoretical literature on the fertility–growth nexus. Existing theoretical studies suggest that: (i) there are multiple channels for fertility to generate time-varying effects on economic growth throughout the life cycles of individuals; (ii) the net effect implied by these channels together is uncertain (could be positive or negative) both in the short and long term; and (iii) the implied long-term effect is likely to be more positive, or less negative, in more developed countries.

Economic growth theories have suggested at least six channels for fertility to affect economic growth. First, the classical growth theory of Thomas Malthus (1798) highlighted that per capita income is simply the ratio of output to population, meaning that a larger newborn cohort corresponds to a lower per capita income. Second, while recognizing the income-dilution effect of increased population growth, neoclassical growth theories (Cass, 1965; Solow, 1956) also suggested that higher demand arising from population growth might promote income growth by inducing physical capital accumulation. Third, most of the R&D-based growth models have predicted that a larger population, which means more potential innovators, can improve economic growth by promoting technological progress (see Jones (1995) for a review). Fourth, the theory of quality and quantity trade-off (Becker et al., 1990) implies that lower fertility increases human capital formation and thus future economic growth. Fifth, higher fertility rates could also affect economic growth by initially reducing parental labor-market participation and subsequently increasing the number of workers (Bloom et al., 2009; Galor & Weil, 1996). Finally, a larger newborn cohort naturally leads to a larger aging cohort in the future, which may hinder (Gordon, 2016) or improve (Acemoglu & Restrepo, 2017) long-term economic growth.

These channels together suggest a time-varying effect of fertility rates on growth, and the net effect implied by these channels could be positive or negative, both in the short and long term. On the one hand, a larger newborn cohort could reduce short-term economic growth by reducing parental labor supply (Galor & Weil, 1996) and reduce long-term economic growth by reducing human capital investment (Becker et al., 1990) and diluting per capita physical capital (Malthus, 1798). On the other hand, a larger newborn cohort could increase short-term economic growth by generating higher demands (Solow, 1956) and increase long-term economic growth by providing a larger working age cohort (Bloom et al., 2009) and more potential innovators (Romer, 1990). The theoretical uncertainty implies that the effect of fertility on economic growth is more of an empirical question.

The theoretical literature also implies that the net effect of fertility depends on the development level of the economy being considered. Growth models with endogenous savings (e.g., the Ramsey–Cass–Koopmans models) generally predict that the rate of physical capital accumulation increases when an economy emerges from a relatively low-income level. Therefore, in more developed countries, the effects of higher fertility through the channels of physical capital dilution (Malthus, 1798) could be less negative and through physical capital accumulation (Solow, 1956) could be more positive. Similarly, growth models with endogenous human capital accumulation (Becker et al., 1990) predict that the accumulation of human capital accelerates only after the economy develops to a level with sufficiently high returns on human capital, which implies that the long-term positive effects of higher fertility from the channels of human capital formation are likely to be larger in more developed countries.

To express these theoretical predictions in a concise way, the current study develops a simple theoretical model that characterizes the major channels through which fertility affects growth, shown in Appendix A. To characterize multiple channels working over different time horizons, the model has to make strong assumptions. Nevertheless, the simple model still facilitates understanding the complex and time-varying nature of the effect of fertility on economic growth. The model demonstrates that the effects of fertility on growth over the life cycle of individuals depend not only on constant elasticity coefficients but also on time-varying elasticity coefficients, the first derivative of the time-varying elasticity coefficients, and the speed of convergence.

The finding of this conceptual framework is somewhat inconsistent with the common understanding that lower fertility improves economic growth. The formation of this common understanding is mainly due to the influences of two major strands of literature. The first strand is the large body of research on the demographic dividend, which examines the observed positive association between fertility declines and economic growth, particularly in Asian developing countries during the latter half of the 20th century (e.g., Bloom & Williamson, 1998; Kotschy et al., 2020; Navaneetham & Dharmalingam, 2012). The second strand is the influential literature on the trade-off between the quality and quantity of children, which provides rigorous theories and strong empirical evidence showing the positive effect of lower fertility on education (e.g., Ashraf et al., 2013; Becker et al., 1990; Canning & Schultz, 2012; Navaneetham & Dharmalingam, 2012). Higher education is expected to increase long-run economic growth by enhancing employment, especially for women, and productivity. However, these are only two major channels through which fertility can affect economic growth. When we combine these channels with other channels highlighted in this conceptual framework, the net impact of fertility on economic growth becomes theoretically uncertain.

3 | DATA AND EMPIRICAL STRATEGY

3.1 | Data

The main analysis of this article focuses on 137 countries (as listed in Table A1) for which data on GDP per capita and TFR are available for each year from 1960 to 2016. The GDP per capita data is sourced from the Maddison Project Database 2018, while the TFR data is obtained from the World Development Indicators. The analysis also incorporates various control variables derived from different sources. Table 1 presents the data sources, definitions, and summary statistics for all variables involved in the primary analysis. In robustness checks, TFR data before 1960 from Roser (2014), national family planning program data from De Silva and Tenreyro (2017), province-level data in China

TABLE 1 Data sources and summary statistics of variables.

Variable name	Definition	Source	Mean
Growth rate of GDP per capita	Annual growth rate of real GDP per capita in 2011 USD	A	0.019
Total fertility rate	The average number of children a woman would have over childbearing years	B, E	4.1
Five-year lagged GDP per capita	Five-year lagged GDP per capita in natural log, 2011 USD	A	8.5
Infant mortality rate	The number of deaths per 1000 live births	B	58.5
Life expectancy	Life expectancy at birth, year	B	62.6
Years of schooling	Years of total schooling for individuals aged 25 and over	C	5.1
Income ratio to US	GDP per capita as a ratio of US GDP per capita	A	0.3
Share of urban population	Urban population as a percentage of total population	B	47.7
Migration	Net international migration, 1000 person	B	801
Resource rents	The share of natural resource rents in GDP, %	B	6.6
Landlocked	The dummy of landlocked	F	0.2
Official language	The first official language	F	–
Political stability	Average of the political stability index	B	−0.2
Program staring year	Starting year of national family planning program	D	1966
Women's labor force participation	Percentage of female labor force	B	39.1

Note: 1. Data sources A: The Maddison Project Database 2018; B: World Development Indicators, the World Bank; C: Barro & Lee (2013); D: De Silva and Tenreyro (2017); E: Roser (2014), for data before 1960; F: Mayer & Zignago (2011). 2. All data pertain to the 137 sample countries over the period of 1960–2016. The growth rate of GDP per capita and fertility rates is calculated as five-year moving averages to minimize the confounding effects of short-term fluctuations; similar results are obtained if the moving averages are not taken into consideration.

from various sources (summarized in Appendix Table A2), and OCP violation fine data in China from Ebenstein (2010) are also utilized. More details will be provided when these data are used in the analysis.

3.2 | Fixed effects panel model and the short-term effect

To address the omitted variable bias, existing studies usually depend on a panel model with location fixed effects to estimate the effect of fertility on economic growth, such as:

$$y_{i,t} = \nu_i + \tau_t + \psi TFR_{i,t} + Z_{i,t}\mu + \varepsilon_{i,t}, \tag{1}$$

where $y_{i,t}$ is the growth rate of GDP per capita (or log GDP per capita) in country i and year t, ν_i denotes the country fixed effects, τ_t denotes the year fixed effects, $TFR_{i,t}$ is the TFR (or the crude birth rate (CBR)) in country i and year t, $Z_{i,t}$ is a set of control variables, and $\varepsilon_{i,t}$ is an error term. The coefficient ψ is expected to capture the effect of fertility on economic growth.

The panel model with location fixed effects has the advantage of addressing the bias from omitted time-invariant factors. It also serves as a starting point for further addressing the endogeneity bias using IVs: conditioning on location fixed effects is generally necessary to ensure the validity of the IV, which is most likely correlated with time-invariant, unobservable determinants of economic growth. Therefore, the fixed effects panel model has been most frequently used when estimating the effect of fertility on economic growth (see Kelley and Schmidt (1994) and Headey and Hodge (2009) for surveys).

However, the coefficient ψ from model (1) mainly captures the short-term effect of fertility. The country fixed effects included in the model eliminate long-term, cross-country fertility differences, and thus, the identification depends primarily on short-term, inter-annual fertility variation. As illustrated in Appendix C, even for a panel of data covering 5 decades, the fixed-effect estimate of ψ still primarily captures short-term effects. Note that existing studies that lag fertility by several years (usually one or 5 years) could not address this concern because the lags applied are too short relative to the time span of the long-term effect of fertility.

Appendix C also shows that there are two ways of extending the fixed effects panel model to capture the long-term lagged effects of fertility rates while keeping the model's advantage of addressing omitted variable bias. The first is to use long-term lagged fertility as the explanatory variable, which is the primary identification strategy used in this article and will be detailed in the next subsection. The second way is to estimate the fixed effects panel model based on fertility shocks that occurred in the early stages of the sample period, following the logic of a standard difference-in-differences (DID) model. The second strategy will be used for robustness checks.

3.3 | Local projections and long-term dynamic effects

A natural way to estimate the long-term dynamic effects of fertility is by extending model (1) to a distributed-lag model that includes long-term lags of fertility:

$$y_{i,t} = \nu_i + \tau_t + \sum_{h=0}^{H} \psi_h TFR_{i,t-h} + Z_{i,t}\mu + \varepsilon_{i,t}, \tag{2}$$

where y_{it} is the annual growth rate of GDP per capita in country i and year t, $TFR_{i,t-h}$ is the total fertility rate h years prior to the current year t, and $h = 0, 1, 2, \cdots H$. The coefficient ψ_h captures the s-year lagged effect of fertility rate on economic growth, and the average of ψ_h over $0 \le h \le H$ reflects the long-term average effect of fertility rate.

However, directly estimating model (2) is infeasible due to serious collinearity issues: fertility rates in successive years are highly correlated with one another. Unless collinearity is adequately addressed, the estimate of ψ_h could be imprecise and have incorrect signs. A standard method for addressing collinearity is to use a restricted least squares estimator that depends on a polynomial distributed lag, which was first explored by Almon (1965). To do this, however, one must first know the pattern of the time effects, which can then be translated into parameter restrictions. Unfortunately, the time evolution of the effects of fertility is too complicated to be characterized by a tractable functional form.[5] Imposing incorrect restrictions on parameters can lead to additional biases, so I do not seek to solve the collinearity problem by using restricted least squares estimators.

Instead, model (2) was transformed into a series of $(H+1)$ estimating equations, each of which only includes one of the lag terms of TFR, with a lag length ranged from 0 to H years:

$$y_{i,t} = \nu_i^h + \tau_t^h + \psi_h TFR_{i,t-h} + Z_{i,t-h}\mu^h + \varepsilon_{i,t}^h, \qquad h = 0, 1, 2, \cdots H. \tag{3}$$

To the extent that nearby lags are correlated with one another, the coefficient of the included lag, ψ_h, captures the effect of the "omitted" nearby lags. The estimate of ψ_h in model (3) can be seen as the weighted average of the effects of the included lag and the omitted nearby lags, and the weighting is the strength of the correlation. It is well-known that this kind of pure serial correlation does not cause bias in the regression coefficient estimates (Greene, 2010, p. 903), but it tends to bias the estimated variances of the regression coefficients. For this reason, this study reports autocorrelation-consistent standard errors (i.e., Newey–West standard errors (Newey & West, 1987)). The simple average of ψ_h (over $h = 0, 1, 2, ..., H$) from model (3) captures the long-term average effect of fertility rates on economic growth. This study depends mainly on model (3) to estimate the long-term dynamic effects of fertility rates.

Model (3) is identical to the local projections model of Jordà (2005), although the latter usually regresses future outcomes on current explanatory variables. Specifically, the local projections model in our case is

$$y_{i,t+h} = \nu_i^h + \tau_t^h + \psi_h TFR_{i,t} + Z_{i,t}\mu^h + \varepsilon_{i,t+h}^h, \qquad h = 0, 1, 2, \cdots H, \tag{4}$$

where $y_{i,t+h}$ is the growth rate of GDP per capita at time $t+h$, and TFR_{it} is the fertility rate at time t. Equation (4) can be obtained by simply replacing the time t of Equation (3) with $t+h$. The local projection approach is a well-established methodology for the estimation of dynamic effects in the macroeconomic literature (Angrist et al., 2018; Nakamura & Steinsson, 2018; Stock & Watson, 2018). Ramey and Zubairy (2018) assessed that the only complication associated with the Jordà method is the serial correlation in the error terms, which can be addressed by using the Newey–West correction for the standard errors. Importantly, Montiel Olea and Plagborg-Møller (2021) showed that after correcting standard errors for serial correlation, local projection inference is robust to using highly persistent data and estimating the effect over long horizons.

3.4 | Addressing the remaining endogeneity bias

The lagged panel model (3) can significantly reduce the endogeneity bias from omitted variables and reverse causality. Country fixed effects account for any confounding factors that are time-invariant within the sample period. The lagged model setting is effective in substantially reducing the bias from reverse causality, especially when the lag length is long. The model will also include various time-varying control variables, as well as year-fixed effects, to account for time-varying confounding factors.

However, it is reasonable to argue that the endogeneity bias may not have been fully addressed. The bias from omitted variables still exists as long as there are omitted time-varying factors that are correlated with both lagged fertility and current growth. The bias from reverse causality still exists to the extent that individuals could make fertility decisions based on expectations of future income growth. As economic growth generally has a negative effect on fertility in modern times (Chatterjee & Vogl, 2018; Eckstein et al., 1999; Herzer et al., 2012), the estimate of ψ_h is expected to be downwardly biased by the remaining reverse causality.

Two approaches have been adopted to address the remaining endogeneity bias. The first is to construct an excluded IV for current and lagged fertility. This approach was only applied when using China provincial data in robustness checks because a time-varying IV is not available at the national level. The second approach is to follow the logic of a standard DID model to estimate the long-term lagged causal effect of an exogenous fertility shock. This approach can be applied when using country-level data by exploiting exogenous fertility shocks from the global family planning campaigns starting around the mid-1960s. Details of these approaches will be introduced later.

4 | EMPIRICAL RESULTS

4.1 | Baseline estimates

Figure 2 presents the estimate of ψ_h and the corresponding 95% confidence intervals from model (3), which regresses the annual growth rate of GDP per capita on the TFR lagged by different years. The confidence intervals are calculated based on the Newey–West standard errors that are robust to autocorrelation and heteroskedasticity. The estimation depends on annual data from 1960 to 2016 for the 137 countries. The baseline estimation controls for four important determinants of fertility and growth: the 5-year lagged log GDP per capita (capturing the effect of economic convergence), years of total schooling (capturing the effect of human capital), infant mortality, and life expectancy. Various other control variables are set aside for robustness checks. All control variables are lagged by the same years as the TFR in each regression to avoid over-control bias.[6] The coefficients of the current TFR and the TFR lagged up to 50 years are estimated, and the figure only reports the coefficients up to 40 lagged years because the remaining estimates have wide confidence intervals (complete results are reported in Appendix Table A3).

The figure shows that a higher fertility rate initially reduces and then increases the growth rate of GDP per capita, and these effects persist for more than 3 decades. Specifically, the initially negative effects turn positive after 15 lagged years, and the positive effects peak at 25 lagged years, after which they decline and become statistically insignificant after 33 lagged years. This finding supports the main argument of this study: it is important to estimate the long-term average effects of fertility. Studies focusing only on the short-term effects capture only the initial negative effects and omit the subsequent positive effects. This finding explains why existing fixed effects panel studies, which primarily capture short-term effects, generally find that fertility has a negative or insignificant effect on growth.

The average of the estimate of ψ_h over 0 to 40 lagged years is 0.12%, with a 95% confidence interval (0.03, 0.21). This implies that a one-unit increase in the TFR would significantly raise the long-term average growth rate of GDP per capita by 0.12% over 4 decades. The accumulated effect calculated based on the estimates (according to $\prod_{0 \le h \le 40}(1 + \psi_h) - 1$) suggests that a one-unit increase in the TFR would raise the GDP per capita by 5.0% for an average sample country over 4 decades (the estimated accumulated effect is comparable when using log GDP per capita as the dependent variable, see Appendix Figure A1). Considering that the TFR had declined by 2.53 from 1960 to 2016 for an

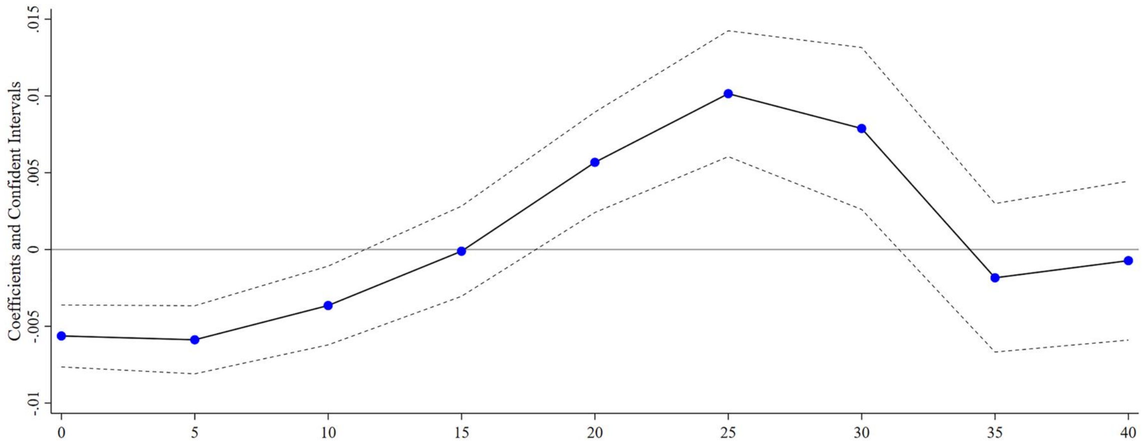

FIGURE 2 Current and lagged effects of the total fertility rate (TFR) on the annual growth rate of GDP per capita. The figure presents the estimated TFR coefficients of model (3), using data from 137 countries from 1960 to 2016. Each dot on the solid line is the point estimate of the coefficient of the TFR lagged by the year indicated by the x-axis, and the dotted lines indicate the corresponding 95% confidence intervals calculated based on the Newey–West standard errors.

average sample country, this estimate implies that secular fertility declines reduced global GDP per capita by 9.1% ($5.0 * 2.53 * 40/56$) over 4 decades.

It is not difficult to explain the dynamic effects presented in Figure 2 by the rich mechanisms summarized in Section 2. First, the initial negative effects of higher fertility on growth could be primarily driven by the negative effect of parental time cost of childrearing. Second, the negative effect gradually disappears and becomes positive as the time cost of childrearing declines with the age of children, while the stimulating effect of higher demands on physical capital formation increases with the age of children. Third, the positive effect increases further when the new cohort directly contributes to the labor force and innovations. Finally, the effect peaks at 25 lagged years and then gradually declines, possibly because 25 years old is approximately the reproductive age of individuals (and thus a new cycle begins).

4.2 | Robustness checks

This subsection presents the nine robustness checks for the baseline estimates presented in Figure 2. All robustness checks are based on model (3) and have the same model setting as the baseline estimation, except for the one specified in each check. All these robustness checks find the same effect pattern as that from the baseline estimation.

Panels A and B of Figure 3 show the robustness of omitted variables. Panel A excludes the four control variables (i.e., 5-year lagged log GDP per capita, years of total schooling, infant mortality, and life expectancy). The estimates suggest smaller initial negative effects, but the effect pattern remains the same. Panel B additionally controls for nine other variables that have been frequently used in the growth literature: the share of the urban population, net international migration, share of natural resource rents in GDP, income ratio to the US, trade share in GDP, the share of the female labor force,[7] pasture to cropland ratio, index of political stability, first official language, and the dummy of landlocked.[8] The definitions and data sources of these control variables are presented in Table 1. The resulting estimates are virtually identical to the baseline estimates. I have also tried to control for subsets of these variables or other relevant variables (such as temperature, inflation, and foreign direct investment) and found very similar results. These robustness checks suggest that the effect pattern estimated is robust to omitted variables.

Panels C and D of the figure examine the robustness of serial correlation. The baseline estimation follows Ramey and Zubairy (2018) to address the potential bias in the standard error from serial correlation by using the Newey–West standard error. Instead, Panel C clusters the error terms at the country level to adjust for within-country correlation of the error terms caused for any reason, not only by serial correlation (Abadie et al., 2017). This approach of addressing error correlation is not preferred in a lagged model setting because the clustering also accounts for error correlations caused by omitted channel variables and thus tends to overestimate the standard error.[9] The resulting confidence intervals are wider, but the corresponding estimates are still statistically significant at least at the 10% level. Panel D directly accounts for serial correlation by controlling for linear and quadratic country-specific time trends and finds comparable estimates.

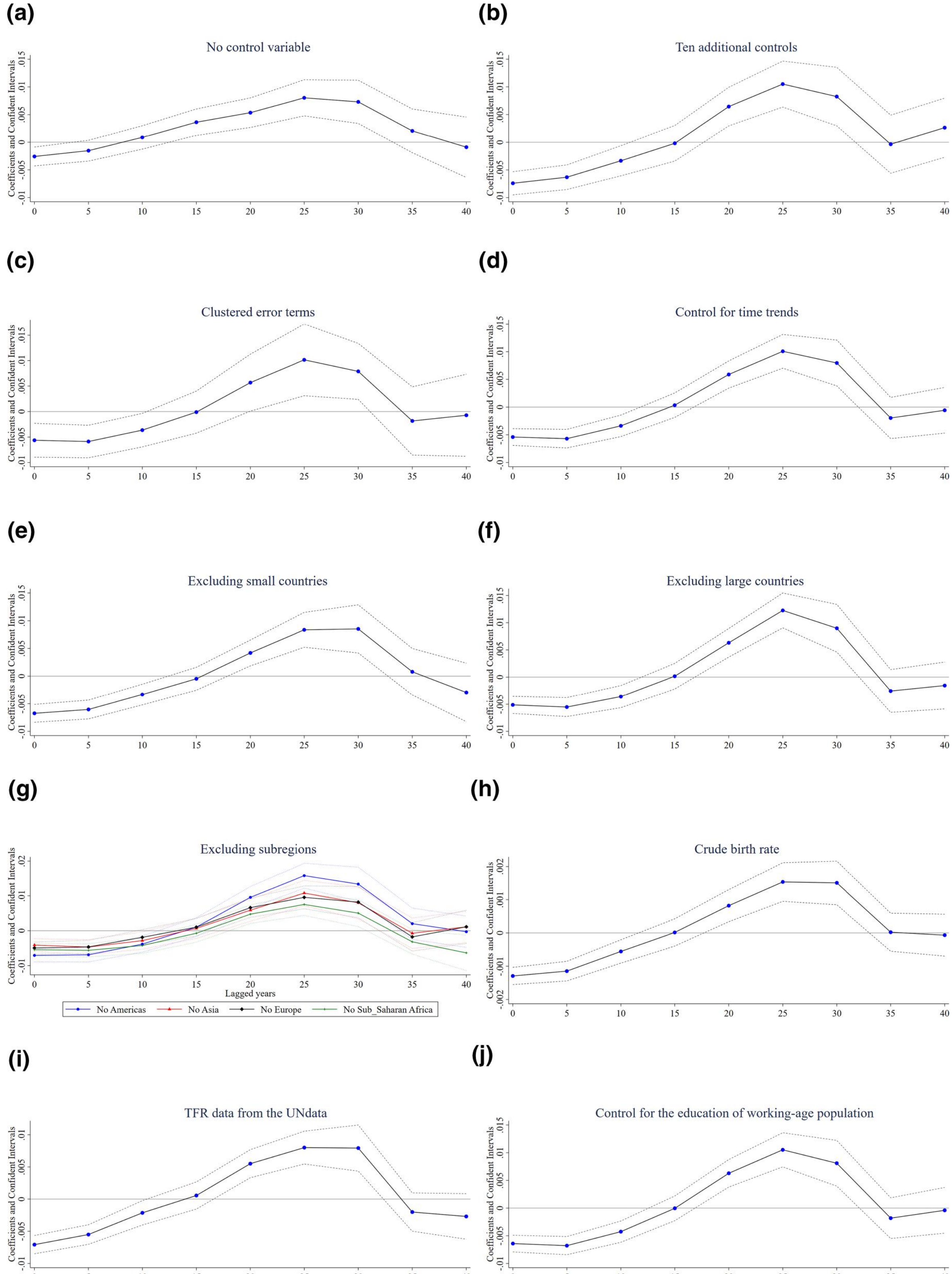

FIGURE 3 Various robustness checks. Each panel of the figure presents a robustness check for the baseline estimates presented in Figure 2. Specifically, Panel (a) excludes the four time-varying control variables, Panel (b) includes 10 additional control variables, Panel (c) clusters the error terms at the country level, Panel (d) controls for linear and quadratic country-specific time trends, Panel (e) excludes 40 countries with a 1960 population smaller than 2 million, Panel (f) excludes 12 countries with a 1960 population larger than 50 million, Panel (g) presents four groups of estimates that respectively exclude countries from one region (America, Asia, Europe, and Sub-Saharan Africa), Panel (h) measures fertility by the crude birth rate (CBR), Panel (i) utilizes the total fertility rate (TFR) data obtained from UNdata, and Panel (j) replaces the control variable of years of total schooling for individuals aged 25 and over with the years of schooling for the working-age population, specifically those aged 15–64.

Panels E, F, and G test the robustness of sub-samples. To determine whether the findings are primarily driven by small or large countries, Panel E excludes 40 countries with a 1960 population smaller than 2 million, and Panel F excludes 12 countries with a 1960 population larger than 50 million. The resulting estimates are almost identical to the baseline estimates. To check if the finding is sensitive to countries from a specific region, Panel G presents four estimations that respectively exclude countries from America, Asia, Europe, and Sub-Saharan Africa. The resulting estimates are comparable, and the indicated effect pattern is the same. Panel H adopts an alternative fertility measure—the CBR—to replace the TFR and finds the same effect pattern, although in different units; the main analysis does not measure fertility by the CBR because it is more sensitive to the age distribution of the population (see Footnote 14 for details). Panel I employs the TFR data from the United Nations' World Population Prospects to assess the robustness of the results concerning different data sources. The estimated results are found to be very similar to the baseline estimates. Panel J substitutes the control variable of years of schooling for the population aged 25 or over with the years of schooling for the working-age population, specifically those aged 16–54. This is done to address the concern that the original education measure might change slowly due to lower fertility, even 25 years after the onset of fertility decline. The resulting estimates are fully comparable, suggesting that the results are not sensitive to the education controls.

Finally, Appendix Figure A1 examines the effect on income levels. The main analysis of this study uses the annual growth rate of GDP per capita (instead of its level) as the dependent variable for two reasons. First, the effect on the growth rate is relevant, as the modified R&D-based growth models (e.g., Howitt, 1999; Jones, 1995; Kortum, 1997) generally predict a weak scale effect in the sense that faster population growth (measured here by fertility rate) increases the long-term economic growth rate. Second, the dynamic effect pattern could be obscured when using the income level as the dependent variable: the resulting estimates reflect the accumulated effect on income level instead of the dynamic effect on the growth rate. Nevertheless, when using the log of real GDP per capita as the dependent variable, Figure a1 shows the same effect pattern: higher fertility first reduces and then increases GDP per capita, and this pattern applies to countries of different development levels. In addition, consistent with the baseline estimation, the estimated long-term accumulated effect on GDP per capita is around 5%.

4.3 | Heterogeneity

Figure 4 presents the estimates of model (3) for low-, middle-, and high-income countries. To avoid using a subjective definition of income levels, the 137 countries were classified into three (approximately) equal-sized groups based on the ranking of their 1960 GDP per capita. I have also tried the development level classification of the World Bank and found a comparable result. These sub-sample regressions show the same effect pattern as the baseline estimation: a higher fertility rate first reduces and then increases income growth, and the long-term average effect is significantly positive. The estimated average effects of a one-unit increase in the TFR on the growth rate over 4 decades are 0.33%, 0.19%, and 0.96%, for low-, middle-, and high-income countries, respectively, with the corresponding 95% confidence intervals (0.20, 0.45), (0.07, 0.31), and (0.76, 1.17). Therefore, a 10% increase in the TFR from the corresponding means (which are 5.4, 3.9, and 2.7 over 1960–2016) would increase the average growth rate of GDP per capita by 0.18%, 0.07%, and 0.26%, in the low-, middle-, and high-income groups of countries, respectively. Indirectly, these results also suggest a

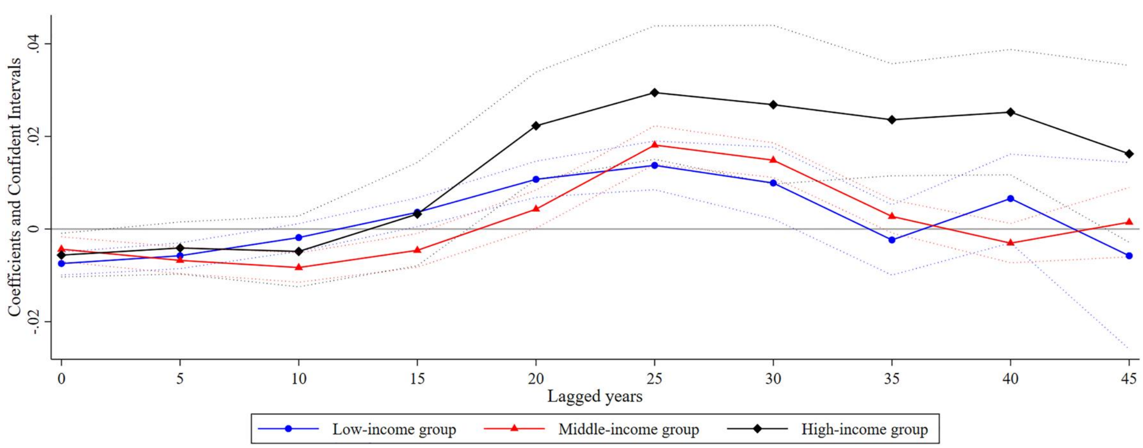

FIGURE 4 Robust to the development level of sample countries. The figure examines the robustness of the baseline estimates to development levels by estimating model (3) separately for countries in the low-, middle-, and high-income groups.

different impact of fertility on economic growth in countries with different fertility levels, as a country's level of development was closely linked to its fertility in the past.

These sub-sample regressions serve two purposes. The first is to test the theoretical prediction that the long-term positive effect of higher fertility is larger in more developed countries (Section 2). Consistent with this prediction, the figure shows that the lagged positive effects are substantially larger and last longer in the high-income group than in the middle- and low-income groups. Additionally, the figure also shows that the lagged positive effects are generally larger in the middle-income group than in the low-income group after 23 lagged years, though the effects in the low-income group become positive earlier.[10]

The second purpose is to verify the existence of downward bias from reverse causality. As previously discussed, the lagged model setting may not fully address reverse causality, which tends to bias the estimated effect downward if economic growth has a negative effect on fertility. These sub-sample regressions provide a way to examine the bias from reverse causality: If the bias exists, the effect estimated separately for different income groups should be more positive than that based on the pooled data. This is because when using pooled data, more variation in the estimation comes from cross-country differences in fertility declines and economic growth, which tends to amplify the downward bias from reverse causality. Figures 2 and 4 confirm the existence of a downward bias from reverse causality: the sub-sample estimates presented in Figure 4 are all larger than the baseline estimates presented in Figure 2. The bias from reverse causality will be addressed further in the next section. Note that the existence of downward bias from reverse causality does not undermine the main conclusion of this study that lower fertility reduces economic growth in the long run.

Another interesting observation from the figure is that the effect size is still large in the high-income group when the TFR is lagged by more than 40 years, though statistically insignificant as indicated by the wide confidence intervals. This observation suggests the possibility that significantly positive effects may last beyond 40 lagged years in high-income countries in a longer time series. However, Appendix Figure A2 rejects this possibility when using data from 1900 to 2016 for 20 high-income countries (marked by red in Table A1). It shows that while the baseline effect pattern remains, the effects after 40 lagged years are all close to zero and statistically insignificant.

4.4 | A discussion of the main finding

Up to this point, the main finding of this article—a higher fertility rate first reduces and then increases economic growth, and the long-term average effect is significantly positive—has been well supported. This finding is robust to omitted variables, serial correlation, sub-samples, measures of fertility and economic growth, and the development level of the country examined. Reverse causality has not been fully addressed, but it tends to bias the estimated effect downward and, therefore, does not undermine this finding; we expect to see a more positive long-term average effect when the bias from reverse causality is further addressed, as presented in the next section.

This finding has four important implications. First, studies exploiting short-term fertility variation tend to overestimate the negative effect of fertility. Second, secular fertility declines represent a strong force driving down long-term economic growth. Third, the scale effect prediction of the R&D-based growth models is well supported when taking into account the long-term lagged effects of fertility. Fourth, the family planning programs prevalent in developing countries significantly hinder long-term economic growth.

The remainder of this article further supports this finding by exploiting plausibly exogenous fertility variations from the global family planning campaigns and China's OCP. The resulting estimates indicate a more positive long-term average effect of fertility on economic growth. The resulting causal effect estimates are useful for evaluating the economic impact of exogenous fertility changes, such as those from birth control policies, medical advancements, and wars. However, one should not directly use the following causal effect estimates to predict the impact of the observed secular fertility declines, which could be driven primarily by economic growth; the causal effect estimates can only be used to credibly predict the economic impact of fertility changes that are exogenous to economic growth.

5 | FURTHER ROBUSTNESS CHECKS

5.1 | Evidence from the global family planning campaigns

Starting around the mid-1960s, increasing concerns over the unprecedented levels of population growth in the developing world led many developing countries to adopt family planning programs (Robinson & Ross, 2007). De Silva and Tenreyro (2017) presented strong evidence that national family planning programs significantly reduced the fertility rate in developing countries. The programs provide exogenous fertility shocks that can be used to identify the long-term effects of fertility, following the logic of a standard DID model.

While there are several intensity measures of the national family planning programs available (i.e., the effort score and the funds for family planning), the identification here does not depend on any of them because they are most likely endogenous to a country's economic performance (Bloom et al., 2009). Instead, I use the variation in the starting year of the programs across countries, which is much less likely to be endogenous. As compiled by De Silva and Tenreyro (2017), the exact starting year of state-led family planning programs is available for 31 developing countries. Figure A3 presents the program starting year for each country.

The program starting year dummy (equal to 1 after the starting year) is used as the IV for fertility for the two-stage least squares (2SLS) estimation of a version of model (1) that uses log GDP per capita as the dependent variable. The 2SLS estimate of the fertility coefficient, ψ, captures the accumulated effect of the fertility shock from the program on GDP per capita over the post-shock period. As the post-shock period is long enough, the estimated effect can be interpreted as the long-term average effect of fertility. Deviating from the baseline estimation, here I use the level of GDP per capita (instead of growth rate) as the dependent variable because the DID model setting is more suitable for capturing the accumulated effect.

The identification is based on the assumption that—conditional on country and year fixed effects—the program starting year is exogenous to a country's economic growth. I provide three pieces of evidence to support this assumption. First, Figure A3 shows that the program starting year is neither obviously correlated with the initial GDP per capita nor with the initial TFR. This finding reduces the concern that countries with better economic performances or higher fertility rates are more likely to start a program earlier. Second, columns 1–3 of Table A4 show that the starting year is not determined by the average growth rate of GDP per capita that is 3, 5, or 10 years prior to the starting year. This finding reduces the endogeneity concern that countries which grow faster may adopt a program earlier. Third, as presented in column 4 of Table A4, conditional on fixed effects and the starting year, the 1-, 5-, and 10-year leads of the starting year have no effect on TFR. This finding suggests that the effect of the starting year on fertility indeed comes from family planning programs instead of from any preexisting omitted factors correlated with the starting year.

Table 2 presents the 2SLS estimates of model (1), using log GDP per capita as the dependent variable and the program starting year as the IV, based on annual data from 1960 to 2016 for the 31 countries. The estimation includes 11 control variables as well as the country and year fixed effects.[11] The first-stage estimates presented in column 1 of Panel B suggest that the adoption of family planning programs significantly reduced TFR by 0.25.[12] The reported Kleibergen–Paap first-stage F-statistic is 16.4, indicating a strong first stage.

TABLE 2 Causal evidence from global family planning campaigns.

	(1) Baseline	(2) Excluding control variables	(3) Excluding samples after 2000	(4) Excluding samples after 1990	(5) Excluding samples after 1980
Panel A. Second stage (dependent variable: Log GDP per capita)					
TFR (IV=Program starting year dummy)	0.473***	0.427***	0.353**	0.094	−0.188
	[0.127]	[0.111]	[0.151]	[0.108]	[0.146]
Panel B. First stage (dependent variable: TFR)					
Program starting year dummy	−0.253***	−0.271***	−0.180***	−0.139**	−0.098*
	[0.057]	[0.057]	[0.059]	[0.055]	[0.055]
Country and year fixed effects	Yes	Yes	Yes	Yes	Yes
11 control variables	Yes	No	Yes	Yes	Yes
First-stage F-statistics	16.4	22.4	11.1	7.8	3.8
Observations	1683	1683	1218	908	598
Countries	31	31	31	31	31
R-squared	0.920	0.920	0.936	0.980	0.983

Note: The table presents the 2SLS estimates of model (1) that uses log GDP per capita as the dependent variable and uses the dummy of the family planning programs starting year as the IV. Heteroskedasticity-consistent standard errors are reported in square brackets. Significance levels are ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

The second-stage estimates reported in column 1 of Panel A indicate that a one-unit increase in TFR significantly increases the long-term GDP per capita (over the post-program period) by 0.47 log points. If the effect of a fertility shock on income growth lasts for about 4 decades (as suggested by the baseline estimation), this accumulated effect estimate implies an average effect on the annual growth rate of GDP per capita of 0.97%. This estimate is larger than the baseline estimates for middle- and low-income countries (presented in Figure 4), confirming that further addressing the endogeneity bias from reverse causality (and other sources) indeed increases the estimated positive effect of fertility. Combining with the first-stage estimate, I calculate that the adoption of the national family planning program reduced an average country's GDP per capita by 11.9% (0.473 × 0.253) or reduced the average annual growth rate of GDP per capita over 4 decades by 0.24%. Column 2 provides a robustness check that excludes all 11 control variables and finds a comparable result.

The remaining columns of the table confirm the baseline finding that while the initial effect of higher fertility is negative, the long-term lagged effect is positive. This is done by excluding the samples from later years and re-estimating the model. If the lagged effect is more positive than the initial effect, excluding later sample years should find a less positive (more negative) accumulated effect estimate. Column 3 excludes sample years after 2000 and shows a smaller but still significantly positive effect of fertility. Column 4 excludes sample years after 1990 and shows that the effect is much smaller and statistically insignificant. Column 5 excludes sample years after 1980 and shows that the effect becomes negative, though not statistically significant at conventional levels. Therefore, the indicated effect pattern is the same as that from the baseline estimation.

5.2 | Evidence from China's one-child policy

China enacted the OCP in 1979 to curb its population explosion (Coale, 1981). The OCP lasted for 3 decades and was significantly modified in 2011. From 1979 to 2010, the OCP generally allowed each couple to only have one child but had several exemption rules.[13] Residents who violated the OCP faced a stiff fine. Ebenstein (2010) collected province-level OCP violation fine rates (measured in times of local yearly household income) in China from 1979 to 2000, finding substantial cross-province and temporal variations in the fine rate (see Figure A4).

By using the time-varying fine rate as an IV, this subsection is able to estimate the dynamic cause effects of fertility on growth. Existing studies have used the fine rate as an IV or proxy variable for fertility when examining the effect on the sex ratio (Ebenstein, 2010), saving rates (Wei & Zhang, 2011), man-made twins (Huang et al., 2016), and various micro-level individual outcomes (Huang et al., 2020). To the best of my knowledge, this study is the first to use the fine rate to examine the dynamic effects of fertility on growth.

Specifically, I use the policy fine rate lagged by h years ($Fine_{p,t-h}$) as the IV for the CBR lagged by the same years ($CBR_{p,t-h}$) in the 2SLS estimation of a version of model (3):

$$y_{p,t} = v_p^h + \tau_t^h + \psi_h CBR_{p,t-h} + Z_{p,t-h}\lambda^h + \varepsilon_{p,t}^h, \qquad h = 0, 1, 2, ..., T. \tag{5}$$

The key explanatory variable, $CBR_{p,t-h}$, is the CBR in province p lagged by h years. Here, fertility is measured by CBR instead of TFR because province-level TFR data are not available for China.[14] All other variables are the same as defined before but at the province level. The estimation depends on annual data from 1980 to 2010 for 27 of the 31 mainland Chinese provincial districts that enforced the OCP.[15] The data sources and summary statistics of all variables constructed from China data are presented in Table A2. The first-stage estimates presented in Appendix Table A6 indicate that the fine rate had a significantly negative and robust effect on fertility.[16]

The identification assumption is that, conditional on the province and year fixed effects, the spatial and temporal variations in the fine rate are exogenous. Studies using the fine rate as an IV have usually argued that changes in the province-level policy fine rate are determined by local-specific factors, such as the new provincial governors' preferences, which are not systematically correlated with the outcomes of interest. This argument is supported by Figure A4: no obvious common patterns are found for the timing, magnitude, or direction of the changes in the fine rate. A series of tests presented in Appendix E.3 also supports the exogeneity assumption. First, Table A7 shows that prior income levels or growth rates have no predictive power on the current policy fine rate. Second, Table A8 shows that the policy fine rate is uncorrelated with various important determinants of economic growth. Third, Table A9 shows that the lead of the policy fine rate is not correlated with current fertility and economic growth.

The estimation further addresses endogeneity concerns by including various control variables. First, it controls for five time-varying determinants of fertility and growth: 5-year lagged log GDP per capita, share of labor with secondary education, crude death rate, net migration rate, and share of the urban population. All these controls are lagged by the same years as the CBR to avoid over-control bias. Second, the estimation controls for the indicators of two important events that may confound the estimated effect: the tax system reform in 1994 and joining the World Trade Organization (WTO) in 2001. Specifically, the estimation controls for the interactions between the timing and intensity of each event.[17] China reformed its tax system in 1994 to triple the central government's share of revenues in GDP from 3% to 9% (Brandt & Rawski, 2008, pp. 431–440). I control for this event using the interaction between the 1994 dummy and government spending share of GDP. China joined the WTO in 2001, which dramatically increased its international trade and liberalized its service sectors (Brandt & Rawski, 2008, pp. 657–659). I control for this event using the interactions between the 2001 dummy and trade-to-GDP ratio and the contribution of services to the GDP, respectively. Finally, the model also controls for province-specific time trends to address serial correlation.[18]

The second-stage estimates of the 2SLS estimation of model (5) are reported in Table 3 (the first-stage estimates are reported in Table A6). To make a comparison with the baseline ordinary least squares (OLS) estimates based on country-level data, column 1 of the table presents the OLS estimates based on China's provincial data. Column 2 presents the 2SLS estimates using the fine rate as the IV. Column 3 tests the robustness of the 2SLS estimates to omitted variables by excluding all control variables (except the fixed effects). Column 4 uses the one-additional-year lagged fine rate $Fine_{p,t-h-1}$ as the IV for $CBR_{p,t-h}$ to allow a lag for the translation of the fine rate change to fertility change. The 2SLS estimates from columns 3 and 4 are comparable to those from column 2.

Both the OLS and 2SLS estimates suggest the same effect pattern as the baseline estimation: higher fertility first reduces and then increases economic growth, and the long-term average effect is significantly positive.[19] In addition, as expected, the 2SLS estimates are more positive than the OLS estimates. The average of OLS estimates of ψ_h over 0 to 20 lagged years (the last year with a significant effect) is 0.07%, while the average of 2SLS estimates (column 2) over the same lagged years is 0.73%. When calculating the accumulated effect over 20 years, the OLS estimates suggest that a one-unit increase in CBR raises GDP per capita by 1.5%, while the 2SLS estimates suggest a corresponding effect of 15.6%. The 2SLS estimates can be used to roughly calculate the economic impact of the OCP violation fine. The mean of the fine rate was 1.74 (Table A2), and the marginal effect of the fine rate on CBR is −0.44 (column 4 of Table A6). Therefore, removing the policy violation fine could have increased China's GDP per capita by 12.0% (i.e., $-0.44 \times 1.74 \times 15.6$).

6 | CONCLUDING REMARKS

This study revisits the causal effect of fertility on economic growth. Existing empirical studies have generally found fertility rates to have a negative or statistically insignificant effect on growth. This article highlights that failing to capture the long-term lagged effects of fertility rates is a crucial reason for this finding in the literature. When adopting estimation methods that capture the long-term lagged effects of fertility rates, both the OLS and 2SLS estimates indicate a significantly positive and economically important long-term average effect of fertility rates on growth. This result can be found separately for low-, middle-, and high-income countries and when using province-level data from China.

The finding of this article has important implications for economists and policymakers. Although the R&D-based growth models unambiguously predict the scale effect, indicating that faster population growth raises long-term economic growth, existing empirical studies have found little evidence supporting this prediction. At the same time, development economists usually hypothesize a virtuous cycle between fertility decline and economic growth based on the assumption that lower fertility rates promote economic growth (and the fact that higher income reduces fertility rates). This study provides strong evidence supporting the scale effect prediction, thus questioning the existence of the virtuous cycle between fertility rate decline and economic growth.

Although fertility rates have dramatically declined in almost all countries over the past few decades and many high-income countries adopting pro-natalist policies to increase their birth rates, many developing countries still adopt family planning programs to curb their population growth, partly motivated by the conventional view that fast population growth hinders long-term economic growth. The findings of this study, however, suggest that under some conditions, fertility declines caused by family planning programs may reduce these countries' long-term economic growth.

I conclude by highlighting three limitations of this study. First, due to data limitations, this study focuses mainly on the effect of fertility on economic growth over 4 decades, although Appendix D3 presents evidence over 70 years when

TABLE 3 Current and lagged effects of the crude birth rate (CBR) on the growth rate of GDP per capita in China.

	OLS		2SLS		2SLS, no controls		2SLS, lagged fine rate	
Lags	(1a) Coefficient	(1b) Se	(2a) Coefficient	(2b) Se	(3a) Coefficient	(3b) Se	(4a) Coefficient	(4b) Se
0	−0.0033***	[0.0007]	−0.0015	[0.0041]	0.0015	[0.0041]	0.0032	[0.0045]
1	−0.0025***	[0.0007]	0.0042	[0.0044]	0.0048	[0.0045]	0.0069	[0.0048]
2	−0.0019***	[0.0006]	0.0074	[0.0047]	0.0077	[0.0047]	0.0061	[0.0047]
3	−0.0007	[0.0006]	0.0040	[0.0042]	0.0063	[0.0044]	0.0010	[0.0044]
4	0.0010	[0.0007]	0.0014	[0.0040]	0.0049	[0.0041]	−0.0009	[0.0042]
5	0.0013*	[0.0007]	−0.0011	[0.0038]	0.0035	[0.0039]	−0.0015	[0.0040]
6	0.0007	[0.0007]	−0.0018	[0.0038]	0.0028	[0.0040]	−0.0001	[0.0042]
7	0.0011	[0.0007]	−0.0001	[0.0039]	0.0038	[0.0041]	0.0041	[0.0043]
8	0.0004	[0.0007]	0.0053	[0.0039]	0.0075*	[0.0043]	0.0083*	[0.0045]
9	0.0013*	[0.0007]	0.0095**	[0.0040]	0.0093**	[0.0042]	0.0126***	[0.0043]
10	0.0021***	[0.0007]	0.0112***	[0.0037]	0.0110***	[0.0039]	0.0123***	[0.0041]
11	0.0025***	[0.0007]	0.0135***	[0.0038]	0.0124***	[0.0038]	0.0127***	[0.0041]
12	0.0027***	[0.0007]	0.0127***	[0.0038]	0.0116***	[0.0035]	0.0081**	[0.0037]
13	0.0019**	[0.0007]	0.0079**	[0.0036]	0.0082**	[0.0033]	0.0084*	[0.0044]
14	0.0007	[0.0007]	0.0066*	[0.0040]	0.0080**	[0.0037]	0.0078	[0.0055]
15	0.0009	[0.0006]	0.0082*	[0.0046]	0.0093**	[0.0044]	0.0021	[0.0066]
16	0.0006	[0.0006]	0.0061	[0.0052]	0.0053	[0.0053]	0.0138	[0.0106]
17	0.0017**	[0.0007]	0.0145**	[0.0073]	0.0111*	[0.0065]	0.0071	[0.0120]
18	0.0012	[0.0007]	0.0103	[0.0089]	0.0060	[0.0083]	0.0247	[0.0571]
19	0.0019***	[0.0007]	0.0206	[0.0192]	0.0125	[0.0170]	−0.0038	[0.0573]
20	0.0017**	[0.0008]	0.0128	[0.0300]	−0.0074	[0.1177]	−0.0108	[0.0169]
21	0.0014	[0.0008]	−0.0190	[0.0243]	−0.0354	[0.0972]	−0.0259	[0.0322]
22	0.0000	[0.0008]	−0.0183	[0.0133]	−0.0184	[0.0211]	−0.0233	[0.0297]
23	−0.0018	[0.0011]	−0.0122	[0.0086]	−0.0062	[0.0107]	−0.0271	[0.0272]
24	−0.0028	[0.0013]	−0.0108	[0.0057]	−0.0092	[0.0068]	−0.0125	[0.0130]
25	−0.0006	[0.0011]	−0.0094	[0.0082]	−0.0118	[0.0119]	0.0004	[0.0178]
26	−0.0003	[0.0012]	0.0025	[0.0078]	−0.0014	[0.0142]	0.0186	[0.0201]
27	0.0011	[0.0022]	0.0054	[0.0045]	0.0096	[0.0077]	0.0011	[0.0107]
28	0.0004	[0.0023]	0.0012	[0.0044]	0.0048	[0.0070]	−0.0036	[0.0108]

Note: All estimates are based on data from 1980 to 2010 for 27 Chinese provinces. All columns (except column 3) control for the province and year fixed effects, the five time-varying control variables, the indicators of the tax system reform in 1994 and joining the WTO in 2001, and province-specific time trends. Column 1 presents the OLS estimates of model (5). Column 2 presents the 2SLS estimates of model (5) by using the policy fine rate as the IV. Column 3 presents the 2SLS estimates that exclude all control variables (except the province and year fixed effects). Column 4 lags the policy fine rate (the IV) by one additional year. Heteroskedasticity-consistent standard errors are reported in square brackets. Significance levels are ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.
Abbreviation: OLS, ordinary least squares.

examining 20 high-income countries with earlier data available. Theoretically, the impact of a fertility change can extend beyond 40 years and even 70 years. For example, a lower fertility rate could lead to higher education levels, which could affect the productivity of both current and future generations. Therefore, the main finding of this study should be interpreted as higher fertility increasing average economic growth over the 40 years following the fertility change.

Second, although this study finds that higher fertility has a net positive effect on economic growth in both developed and developing countries, the possibility cannot be excluded that the net effect could be negative for countries with very high fertility. Very high population growth could accelerate resource pressures, contribute to environmental degradation and food insecurity, lead to conflict, and ultimately result in governance collapse in some countries (e.g., Bongaarts, 2016). Given that this study relies mainly on country-level data, the sample size is not sufficient to identify the threshold fertility rate beyond which the effect may turn negative. Determining such a threshold requires comparing small subgroups of countries within the context of this study. Future studies that explore long-run micro-level data to identify this threshold fertility rate would make significant contributions to the literature.

Finally, one should be careful when applying the findings of this study to evaluate the merits of family planning programs. This study does not consider population pressure on resources and the environment and potential resource scarcity in many countries with lower levels of development and high levels of fertility. Family planning programs are helpful for addressing these resource and environmental issues in developing countries. In addition, from an individual perspective, the basic provision of family planning and reproductive health services can be seen as a human and reproductive rights issue that should not be questioned from a purely economic cost-benefit analysis.

ACKNOWLEDGMENTS

I am grateful to the editor and three anonymous referees for the insightful comments and suggestions. The work was funded by National Natural Science Foundation of China (NSFC72273001).

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in openICPSR at https://doi.org/10.3886/E197782V4 (Huang, 2024).

ORCID

Kaixing Huang https://orcid.org/0000-0002-7761-4153

ENDNOTES

[1] The long-term effects investigated in this study refer to the effects over decades (within the life cycle of individuals) instead of the cross-generation effect usually examined in theoretical growth models.

[2] The argument that existing studies are weak either in addressing the endogeneity bias or in capturing the long-term effect of fertility rates also generally applies to those studies adopting micro approaches, using pure time series methods, or focusing on pre-modern times. There are excellent studies examining the fertility–income nexus by using family-level data (e.g., Ashraf et al., 2014; Black et al., 2005), adopting time series methods (e.g., Hafner & Mayer-Foulkes, 2013; Herzer et al., 2012), and focusing on pre-modern times (e.g., Ashraf & Galor, 2011; Lee & Anderson, 2002).

[3] It should be noted that there are good reasons to believe that this association may actually reflect the fact that income growth reduces fertility rates. For example, it has been proposed that rising income increases the opportunity cost of fertility for women (Galor & Weil, 1996), that technological progress raises the importance of human capital relative to raw labor (Galor & Weil, 2000), and that higher income reduces the needs of old-age security from children (Strulik, 2003).

[4] In some lower-income countries with very high fertility, fertility might have a negative impact on economic growth. As detailed in the Concluding Remarks, a more nuanced analysis than presented in this study would be needed to prove that the positive effect of fertility on economic growth also prevails in countries with high or very high fertility levels.

[5] Because an individual's interaction with the economy lasts for decades and changes in a nonlinear manner over time.

[6] If the control variables are not lagged, which means that they are in the same time period as the dependent variable, the control variables could partially account for the true effect of the lagged fertility. Specifically, if there are any correlations between the future values of the control variables and the past TFR, it is most likely that the past TFR is the cause of the correlation. In this case, these (current) control variables are the channel variables through which the lagged TFR affects income growth, and controlling for these channel variables would account for the true effect of the lagged fertility.

[7] Controlling for the share of the female labor force is important in addressing concerns regarding endogeneity bias stemming from the evolution of women's fertility choices. As emphasized by Doepke et al. (2023), the role of women in fertility choices has changed over time and can have a significant impact on economic outcomes. It should be noted that data on the share of the female labor force was not available for most countries before 1990, and I have filled in the missing values through extrapolation.

[8] Each of the last three time-invariant control variables is interacted with a full set of year dummies; the index of political stability is time-varying, but here I control for the country mean of it (and interacted with year dummies) because the data are only sparsely available after the 1990s.

[9] For example, lagged TFR could affect future growth by affecting future education, so the model intentionally excludes future (non-lagged) education as a control variable to avoid over-control bias (see Footnote 6 for details). The "omitted" future education naturally leads to within-country error correlation given the lagged setting of model (3). Adjusting for this kind of error correlation could lead to upward-biased standard errors.

[10] A potential explanation for the smaller initial negative effect in the low-income group is that parents may invest less time in childrearing (i.e., human capital investment) in low-income countries than in middle-income countries.

[11] The control variables are years of total schooling, infant mortality, life expectancy, the share of the urban population, net international migration, share of natural resource rents in GDP, income ratio to the US, trade share in GDP, the index of political stability, first official language, and the dummy for landlocked. The last three time-invariant variables are interacted with a full set of year dummies. The 5-year lagged GDP per capita (a control variable in the baseline analysis) is excluded here because the dependent variable is income level. I have also tried to control for subsets of these variables and found comparable results.

[12] This effect could be permanent if the family planning programs successfully and permanently altered the fertility preferences of individuals or couples.

[13] The three most important exemptions were (1) couples with an agricultural hukou (a system of household registration) were allowed to have a second child if their first child was a girl; (2) residents who belonged to an ethnic minority group were allowed to have more than one child; and (3) residents in Xinjiang and Tibet were not subject to the OCP until the early 1990s (Baochang et al., 2007).

[14] The CBR is defined as the annual number of births per thousand populations, while TFR is defined as the average number of children that a woman would have over her childbearing years. Therefore, TFR is a better measure of current fertility than CBR because it is less affected by the age distribution of the population. Note that using CBR as the fertility measure in the global analysis yielded a comparable result (see Figure 3 Panel H).

[15] The provinces of Xinjiang and Tibet are excluded because they were not subjected to the OCP until the early 1990s, and Hainan and Chongqing are excluded because they were separated from Guangdong and Sichuan in 1988 and 1997, respectively. The year of 1979 is excluded because the OCP was implemented at the end of 1979, thus having no effect on fertility in 1979 (due to the nine-and-a-half-month length of a pregnancy). Data after 2010 are excluded because the OCP was significantly modified in 2011.

[16] A potential concern associated with using this dataset is the possibility of under-reporting or incorrect reporting of crude birth rates in certain periods and provinces (Yang et al., 2022). This could potentially result in a biased OLS estimate of the effect of fertility on economic growth. However, as long as the measurement error in the crude birth rate is uncorrelated with the policy fine rate, the 2SLS estimates will be unbiased.

[17] Another important event, the reform and opening up in 1978, is not controlled for because it occurred before the OCP. I have attempted to control for this event using the interactions between a full set of year dummies and two intensity measures of this event (the trade-to-GDP ratio and distance to the nearest port) and found a similar result.

[18] Recall that a key assumption in using the fine rate as the IV is that it is not correlated with omitted determinants of economic growth. One important concern regarding this assumption is that the OCP and its enforcement varied significantly between provinces and over time (Baochang et al., 2007). If these variations are correlated with both the IV and economic growth, it would violate the exogeneity assumption of the IV. To address this concern, I control for year- and province-fixed effects, include various other control variables, and demonstrate that the policy fine rate is not correlated with prior income levels, prior fertility rates, and other important determinants of economic growth.

[19] The OLS estimates based on China data (column 1 of Table 3) indicate shorter initial negative effects than the OLS estimates based on global data (Figure 2). This may reflect the fact that China implemented the most coercive birth control policy in the world: when more fertility declines are exogenous, the OLS estimates are less likely biased downward by reverse causality.

REFERENCES

Abadie, A., Athey, S., Imbens, G.W. & Wooldridge, J. (2017) *When should you adjust standard errors for clustering?* National Bureau of Economic Research Working Paper.

Acemoglu, D. & Restrepo, P. (2017) Secular stagnation? The effect of aging on economic growth in the age of automation. *The American Economic Review*, 107(5), 174–179. Available from: https://doi.org/10.1257/aer.p20171101

Aghion, P. & Howitt, P. (1992) A model of growth through creative destruction. *Econometrica*, 323–351.

Almon, S. (1965) The distributed lag between capital appropriations and expenditures. *Econometrica*, 33(1), 178–196. Available from: https://doi.org/10.2307/1911894

Angrist, J.D., Jordà, Ò. & Kuersteiner, G.M. (2018) Semiparametric estimates of monetary policy effects: string theory revisited. *Journal of Business and Economic Statistics*, 36(3), 371–387. Available from: https://doi.org/10.1080/07350015.2016.1204919

Ashraf, N., Field, E. & Lee, J. (2014) Household bargaining and excess fertility: an experimental study in Zambia. *The American Economic Review*, 104(7), 2210–2237. Available from: https://doi.org/10.1257/aer.104.7.2210

Ashraf, Q. & Galor, O. (2011) Dynamics and stagnation in the Malthusian epoch. *The American Economic Review*, 101(5), 2003–2041. Available from: https://doi.org/10.1257/aer.101.5.2003

Ashraf, Q.H., Weil, D.N. & Wilde, J. (2013) The effect of fertility reduction on economic growth. *Population and Development Review*, 39(1), 97–130. Available from: https://doi.org/10.1111/j.1728-4457.2013.00575.x

Baochang, G., Feng, W., Zhigang, G. & Erli, Z. (2007) China's local and national fertility policies at the end of the twentieth century. *Population and Development Review*, 33(1), 129–148. Available from: https://doi.org/10.1111/j.1728-4457.2007.00161.x

Barro, R.J. & Lee, J.W. (2013) A new data set of educational attainment in the world, 1950–2010. *Journal of Development Economics*, 104, 184–198. Available from: https://doi.org/10.1016/j.jdeveco.2012.10.001

Becker, G.S. & Lewis, H.G. (1973) On the interaction between the quantity and quality of children. *Journal of Political Economy*, 81(2), S279–S288. Available from: https://doi.org/10.1086/260166

Becker, G.S., Murphy, K.M. & Tamura, R. (1990) Human capital, fertility, and economic growth. *Journal of Political Economy*, 98(5), S12–S37. Available from: https://doi.org/10.1086/261723

Black, S.E., Devereux, P.J. & Salvanes, K.G. (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. Available from: https://doi.org/10.1093/qje/120.2.669

Bloom, D.E., Canning, D., Fink, G. & Finlay, J.E. (2009) Fertility, female labor force participation, and the demographic dividend. *Journal of Economic Growth*, 14(2), 79–101. Available from: https://doi.org/10.1007/s10887-009-9039-9

Bloom, D.E. & Williamson, J.G. (1998) Demographic transitions and economic miracles in emerging Asia. *The World Bank Economic Review*, 12(3), 419–455. Available from: https://doi.org/10.1093/wber/12.3.419

Bongaarts, J. (2016) Development: slow down population growth. *Nature*, 530(7591), 409–412. Available from: https://doi.org/10.1038/530409a

Brandt, L. & Rawski, T.G. (2008) *China's great economic transformation*. Cambridge: Cambridge University Press.

Canning, D. & Schultz, T.P. (2012) The economic consequences of reproductive health and family planning. *The Lancet*, 380(9837), 165–171. Available from: https://doi.org/10.1016/s0140-6736(12)60827-7

Cass, D. (1965) Optimum growth in an aggregative model of capital accumulation. *The Review of Economic Studies*, 32(3), 233–240. Available from: https://doi.org/10.2307/2295827

Cavalcanti, T., Kocharkov, G. & Santos, C. (2020) Family planning and development: aggregate effects of contraceptive use. *The Economic Journal*, 131(634), 624–657. Available from: https://doi.org/10.1093/ej/ueaa070

Chatterjee, S. & Vogl, T. (2018) Escaping Malthus: economic growth and fertility change in the developing world. *The American Economic Review*, 108(6), 1440–1467. Available from: https://doi.org/10.1257/aer.20170748

Cleland, J., Bernstein, S., Ezeh, A., Faundes, A., Glasier, A. & Innis, J. (2006) Family planning: the unfinished agenda. *The Lancet*, 368(9549), 1810–1827. Available from: https://doi.org/10.1016/s0140-6736(06)69480-4

Coale, A.J. (1981) Population trends, population policy, and population studies in China. *Population and Development Review*, 7(1), 85–97. Available from: https://doi.org/10.2307/1972766

Coale, A.J. & Hoover, E.M. (1958) *Population growth and economic development in low income countries: a case study of India's prospects*. Princeton: Princeton University Press.

Dasgupta, P. (1995) The population problem: theory and evidence. *Journal of Economic Literature*, 33(4), 1879–1902.

De Silva, T. & Tenreyro, S. (2017) Population control policies and fertility convergence. *The Journal of Economic Perspectives*, 31(4), 205–228. Available from: https://doi.org/10.1257/jep.31.4.205

Doepke, M., Hannusch, A., Kindermann, F. & Tertilt, M. (2023) The economics of fertility: a new era. In: *Handbook of the economics of the family*. Elsevier, 1, pp. 151–254.

Ebenstein, A. (2010) The "missing girls" of China and the unintended consequences of the one child policy. *Journal of Human Resources*, 45(1), 87–115. Available from: https://doi.org/10.3368/jhr.45.1.87

Eckstein, Z., Mira, P. & Wolpin, K.I. (1999) A quantitative analysis of Swedish fertility dynamics: 1751–1990. *Review of Economic Dynamics*, 2(1), 137–165. Available from: https://doi.org/10.1006/redy.1998.0041

Ehrlich, P.R. (1968) *The population bomb*. New York: Ballentine.

Galor, O. & Weil, D.N. (1996) The gender gap, fertility, and growth. *The American Economic Review*, 374–387.

Galor, O. & Weil, D.N. (2000) Population, technology, and growth: from Malthusian stagnation to the demographic transition and beyond. *The American Economic Review*, 90(4), 806–828. Available from: https://doi.org/10.1257/aer.90.4.806

Gordon, R.J. (2016) Perspectives on the rise and fall of American growth. *The American Economic Review*, 106(5), 72–76. Available from: https://doi.org/10.1257/aer.p20161126

Greene, W.H. (2010) *Econometric analysis*, 7 edn. NJ: Prentice Hall.

Grossman, G.M. & Helpman, E. (1991) Quality ladders in the theory of growth. *The Review of Economic Studies*, 58(1), 43–61. Available from: https://doi.org/10.2307/2298044

Hafner, K.A. & Mayer-Foulkes, D. (2013) Fertility, economic growth, and human development causal determinants of the developed lifestyle. *Journal of Macroeconomics*, 38, 107–120. Available from: https://doi.org/10.1016/j.jmacro.2013.04.001

Headey, D.D. & Hodge, A. (2009) The effect of population growth on economic growth: a meta-regression analysis of the macroeconomic literature. *Population and Development Review*, 35(2), 221–248. Available from: https://doi.org/10.1111/j.1728-4457.2009.00274.x

Herzer, D., Strulik, H. & Vollmer, S. (2012) The long-run determinants of fertility: one century of demographic change 1900–1999. *Journal of Economic Growth*, 17(4), 1–29. Available from: https://doi.org/10.1007/s10887-012-9085-6

Howitt, P. (1999) Steady endogenous growth with population and R&D inputs growing. *Journal of Political Economy*, 107(4), 715–730. Available from: https://doi.org/10.1086/250076

Huang, K. (2024) *ECIN replication package for "Fertility and long-term economic growth"*. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2024-02-21. Available from: https://doi.org/10.3886/E197782V4

Huang, W., Lei, X. & Sun, A. (2020) Fertility restrictions and life-cycle outcomes: evidence from the one-child policy in China. *The Review of Economics and Statistics*, 1–41. Available from: https://doi.org/10.1162/rest_a_00921

Huang, W., Lei, X. & Zhao, Y. (2016) One-child policy and the rise of man-made twins. *The Review of Economics and Statistics*, 98(3), 467–476. Available from: https://doi.org/10.1162/rest_a_00567

Jones, C.I. (1995) R&D-based models of economic growth. *Journal of Political Economy*, 103(4), 759–784. Available from: https://doi.org/10.1086/262002

Jordà, s. (2005) Estimation and inference of impulse responses by local projections. *The American Economic Review*, 95(1), 161–182. Available from: https://doi.org/10.1257/0002828053828518

Kelley, A.C. (1988) Economic consequences of population change in the Third World. *Journal of Economic Literature*, 26(4), 1685–1728.

Kelley, A.C. & Schmidt, R.M. (1994) *Population and income change: recent evidence*. Washington: The World Bank.

Kortum, S.S. (1997) Research, patenting, and technological change. *Econometrica*, 65(6), 1389–1419. Available from: https://doi.org/10.2307/2171741

Kotschy, R., Suarez Urtaza, P. & Sunde, U. (2020) The demographic dividend is more than an education dividend. *Proceedings of the National Academy of Sciences*, 117(42), 25982–25984. Available from: https://doi.org/10.1073/pnas.2012286117

Kremer, M. (1993) Population growth and technological change: one million B.C. to 1990. *Quarterly Journal of Economics*, 108(3), 681–716. Available from: https://doi.org/10.2307/2118405

Kuang, B. & Brodsky, I. (2016) Global trends in family planning programs, 1999–2014. *International Perspectives on Sexual and Reproductive Health*, 42(1), 33–44. Available from: https://doi.org/10.1363/42e0316

Kuznets, S. (1967) Population and economic growth. *Proceedings of the American Philosophical Society*, 111(3), 170–193.

Lee, R. & Anderson, M. (2002) Malthus in state space: macro economic-demographic relations in English history, 1540 to 1870. *Journal of Population Economics*, 15(2), 195–220. Available from: https://doi.org/10.1007/s001480100091

Li, B. & Zhang, J. (2007) Do high birth rates hamper economic growth? *The Review of Economics and Statistics*, 89(1), 110–117. Available from: https://doi.org/10.1162/rest.89.1.110

Malthus, T.R. (1798) *An essay on the principle of population*. Cambridge: W. Pickering, 1986.

Mayer, T. & Zignago, S. (2011) Notes on CEPII's distances measures: the geodist database. *CEPII Working Paper 2011-25*.

Montiel Olea, J.L. & Plagborg-Møller, M. (2021) Local projection inference is simpler and more robust than you think. *Econometrica*, 89(4), 1789–1823. Available from: https://doi.org/10.3982/ecta18756

Nakamura, E. & Steinsson, J. (2018) Identification in macroeconomics. *The Journal of Economic Perspectives*, 32(3), 59–86. Available from: https://doi.org/10.1257/jep.32.3.59

Navaneetham, K. & Dharmalingam, A. (2012) A review of age structural transition and demographic dividend in South Asia: opportunities and challenges. *Journal of Population Ageing*, 5(4), 281–298. Available from: https://doi.org/10.1007/s12062-012-9071-y

Newey, W.K. & West, K.D. (1987) Hypothesis testing with efficient method of moments estimation. *International Economic Review*, 28(3), 777–787. Available from: https://doi.org/10.2307/2526578

Peretto, P.F. (1998) Technological change and population growth. *Journal of Economic Growth*, 3(4), 283–311. Available from: https://doi.org/10.1023/a:1009799405456

Ramey, V.A. & Zubairy, S. (2018) Government spending multipliers in good times and in bad: evidence from US historical data. *Journal of Political Economy*, 126(2), 850–901. Available from: https://doi.org/10.1086/696277

Robinson, W.C. & Ross, J.A. (2007) *The global family planning revolution: three decades of population policies and programs*. Washington: World Bank.

Romer, P.M. (1990) Endogenous technological change. *Journal of Political Economy*, 98(5, Part 2), S71–S102

Roser, M. (2014) Fertility rate.

Schultz, P. (2008) *Population policies, fertility, women's human capital, and child quality, Vol. 4 of Handbook of Development Economics*. Elsevier, pp. 3249–3303.

Segerstrom, P.S. (1998) *Endogenous growth without scale effects*. American Economic Review, pp. 1290–1310.

Simon, J. (1992) *Population and development in poor countries: selected essays*. Princeton: Princeton University Press.

Simon, J.L. (1989) On aggregate empirical studies relating population variables to economic development. *Population and Development Review*, 323–332.

Solow, R.M. (1956) A contribution to the theory of economic growth. *Quarterly Journal of Economics*, 70(1), 65–94. Available from: https://doi.org/10.2307/1884513

Stock, J.H. & Watson, M.W. (2018) Identification and estimation of dynamic causal effects in macroeconomics using external instruments. *The Economic Journal*, 128(610), 917–948. Available from: https://doi.org/10.1111/ecoj.12593

Stover, J. & Sonneveldt, E. (2017) Progress toward the goals of FP2020. *Studies in Family Planning*, 48(1), 83–88. Available from: https://doi.org/10.1111/sifp.12014

Strulik, H. (2003) Mortality, the trade-off between child quality and quantity, and demo-economic development. *Metroeconomica*, 54(4), 499–520. Available from: https://doi.org/10.1111/1467-999x.00177

United Nations (2015) World population policies database. Data retrieved from http://data.worldbank.org/indicator/SP.DYN.LE00.FE

Wei, S.-J. & Zhang, X. (2011) The competitive saving motive: evidence from rising sex ratios and savings rates in China. *Journal of Political Economy*, 119(3), 511–564. Available from: https://doi.org/10.1086/660887

Yang, S., Jiang, Q. & Sánchez-Barricarte, J.J. (2022) China's fertility change: an analysis with multiple measures. *Population Health Metrics*, 20(1), 1–14. Available from: https://doi.org/10.1186/s12963-022-00290-7

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Huang, K. (2024) Fertility and long-term economic growth. *Economic Inquiry*, 62(3), 1152–1171. Available from: https://doi.org/10.1111/ecin.13216