Natural selection and innovation-driven growth

Chu, Angus and Cozzi, Guido and Fan, Haichao and Hu, Dongmin

University of Macau, University of St. Gallen, Fudan University, Fudan University

June 2022

Natural Selection and Innovation-Driven Growth

Angus C. Chu Guido Cozzi Haichao Fan Dongmin Hu

June 2024

Abstract

How does the interplay between natural selection, household education choices, and R&D activities shape our macroeconomic trajectory? To explore this question, we develop an innovation-driven growth model that connects household heterogeneity in education ability with endogenous fertility and endogenous takeoff. Because households with lower education abilities accumulate less human capital and choose to have more offspring, they gain a temporary evolutionary advantage over households with higher education abilities, which in turn has novel implications on innovation-driven growth. Initially, the heterogeneity of households makes it more likely for an endogenous takeoff to occur; however, the temporary evolutionary disadvantage of high-ability households stifles R&D and long-run growth. We also validate this theoretical result with cross-country data and instrumental variables, suggesting that education ability disparities can hamper R&D, education, and economic growth in the long run. This research unveils useful insights into the nuanced relationships between natural selection, household education choices, and innovation.

JEL classification: O30, O40
Keywords: natural selection, education, innovation, endogenous takeoff

Chu: anguscccc@gmail.com. Department of Economics, University of Macau, Macau, China.
Cozzi: guido.cozzi@unisg.ch. Department of Economics, University of St. Gallen, St. Gallen, Switzerland.
Fan: fan_haichao@fudan.edu.cn. Institute of World Economy, Fudan University, Shanghai, China.
Hu: dmhu21@m.fudan.edu.cn. Institute of World Economy, Fudan University, Shanghai, China
The authors are most grateful to Oded Galor for insightful comments. The usual disclaimer applies.

"Britons are becoming less educated and poorer because smart rich people are having fewer children." The Telegraph (2022)[1]

1 Introduction

Modern macroeconomic models often feature a representative household or a fixed composition of heterogeneous households. However, when heterogeneous households choose to have different fertility rates, their composition in the economy changes over time. This differential reproduction of individuals is famously known as natural selection. In this study, we explore how the heterogeneity of households and natural selection of heterogeneous households affect the macroeconomy. Family attitudes toward the education of their children last long, and the intra-family educational attitudes and human capital transmission abilities matter.[2] Unfortunately, not all households are equally endowed, so heterogeneity matters for human capital accumulation. How does this heterogeneity affect fertility? And how would the resulting natural selection influence technological progress and economic growth? To explore these questions, we develop an innovation-driven growth model with fertility, an endogenous activation of innovation and natural selection of heterogeneous households, which persistently differ in their propensity to educate their children.

Following the seminal unified growth theory of Galor (2005, 2011, 2022), we assume that households differ in their ability to accumulate human capital. In this case, families that are more able to provide high-quality learning focus on child quality and have fewer children than less able families. This negative relationship between child quantity and quality during the demographic transition is consistent with the empirical evidence in Becker *et al.* (2010), Fernihough (2017) and Klemp and Weisdorf (2019).[3] Naturally, this quality-quantity tradeoff magnifies the share of less able families in the economy, at least temporarily. Therefore, initially, households that have a lower education ability accumulate less human capital but choose to have more children and enjoy a temporary evolutionary advantage. As their human capital rises over time, households with a higher education ability choose to increase their number of children because their higher level of income enables them to have more children.[4] In the long run, households with a higher education ability end up having a higher level of human capital, and all households choose the same steady-state fertility rate. Therefore, households' population share and human capital converge to stationary distributions.

Initially, the heterogeneity of households makes it more likely for the activation of innovation to occur. The presence of heterogeneous households implies that some households supply more human capital for production and innovation whereas some households supply less. In our model, the former effect dominates the latter effect such that the initial amount of human capital available for production and innovation increases as a result of this heterogeneity. However, the evolutionary disadvantage of high-ability households during the transitional dynamics implies that the population share of high-ability households decreases and the population share of low-

[1] https://www.telegraph.co.uk/news/2022/07/06/britons-evolving-poorer-less-well-educated/

[2] For example, Alesina *et al.* (2021) find that differences in family attitudes toward education persist and rebound after even some of the most forceful attempts to eliminate differences in the population.

[3] See also Shiue (2017) and Bai *et al.* (2023) for evidence in pre-industrial China.

[4] See Strulik (2024) for evidence on this positive relationship between income and fertility.

ability households increases towards the steady state. Although the level of human capital rises over time due to all households accumulating human capital, the population becomes less educated (relative to the case without natural selection) because the more educated parents have fewer children during the transition dynamics.

The above finding resonates with the opening quote and shows that this phenomenon may not be specific to Britain. The lower long-run share of high-ability households is due to a well-known property that a temporary growth effect has a permanent level effect. Suppose two variables start at an equal level. Then, one of them grows at a slower rate temporarily before growing at the same rate as the other variable. In this case, the temporary disadvantage of the former will endure forever. So, despite population trends being similar in the long run, a temporarily lower population growth rate of the higher-ability households will never be compensated, which in turn has novel implications on R&D and innovation-driven growth.

The scale-invariant property of our model then implies that economic growth depends on the average level of human capital in the economy and that the lower share of high-ability households (relative to the case without natural selection) in the long run gives rise to a lower steady-state equilibrium growth rate as a result of natural selection of heterogeneous households. Finally, we show that the negative effects of this natural selection can be captured by the heterogeneity in the ability to accumulate human capital and provide evidence that heterogeneity in education ability indeed has adverse effects on education, innovation and economic growth in the long run. This result remains robust when we use ancestral population diversity and prehistoric migratory distance in Ashraf and Galor (2013) as instrumental variables for heterogeneity in education ability.

Central to this exploration is the link between natural selection, innovation, and R&D. The interplay between household heterogeneity in education, fertility choices, and the effect on R&D activities is a focal point of our study. This dynamic plays a significant role in shaping technological progress, where the temporary disadvantage of high-ability households during the transitional dynamics can have lasting effects on R&D efforts and, consequently, long-term economic growth. Our model sheds light on these complex interactions and how natural selection can influence the broader R&D landscape, innovation, and growth trajectories. Such insights hold useful implications for both economic policy and the theoretical understanding of growth mechanisms.

This study relates to the literature on innovation and economic growth. The pioneering study by Romer (1990) develops the seminal innovation-driven growth model; see also Aghion and Howitt (1992), Grossman and Helpman (1991) and Segerstrom *et al.* (1990) for other early studies. Some subsequent studies introduce endogenous fertility to variants of the innovation-driven growth model to explore the relationship between economic growth and endogenous population growth; see, for example, Jones (2001), Connolly and Peretto (2003), Chu *et al.* (2013), Peretto and Valente (2015) and Brunnschweiler *et al.* (2021). This study contributes to this literature by exploring the endogenous fertility decisions of heterogeneous households and their evolutionary differences in an innovation-driven growth model.

This study also relates to the literature on endogenous takeoff and economic growth. The seminal study by Galor and Weil (2000) develops the unified growth theory that explores the endogenous transition of an economy from pre-industrial stagnation to modern economic growth;[5] see Galor (2005) for a comprehensive review of unified growth theory and also Galor and Mountford (2008), Galor, Moav and Vollrath (2009) and Ashraf and Galor (2011) for subsequent studies and empirical evidence that supports unified growth theory. Galor and Moav (2002), Galor and Michalopoulos (2012) and Carillo *et al.* (2019) explore how natural selection of different traits, such as the quality preference of fertility, the degree of entrepreneurs' risk aversion and the level of family-specific human capital, affects the transition from stagnation to growth. Specifically, Galor and Moav (2002) show that natural selection favors the quality type during the demographic transition and fosters technological progress,[6] but the selective advantage is reversed to favor the quantity type after the demographic transition.[7] Galor and Michalopoulos (2012) show that as an economy develops, risk-tolerant individuals allocate less resources towards fertility, which in turn leads to a decline in entrepreneurial spirit over time. Although the present study does not explore the interesting demographic transition in the pre-industrial era captured elegantly by unified growth theory, it complements the interesting studies in this literature by showing how natural selection of heterogeneous households with different ability to accumulate human capital affects the transition of an economy from human capital accumulation to innovation-driven growth, which is the novel contribution of this study.

Therefore, this study also relates to a recent branch of this literature on the endogenous transition from pre-industrial stagnation to innovation-driven growth. For example, Funke and Strulik (2000) develop a growth model in which the economy experiences capital accumulation and variety-expanding innovation in different stages of economic development. A more recent study by Peretto (2015) develops a Schumpeterian growth model with the endogenous activations of variety-expanding innovation and quality-improving innovation. Subsequent studies extend the model in Peretto (2015) to explore different mechanisms that trigger an endogenous takeoff.[8] This study contributes to this branch of the literature by introducing natural selection of heterogeneous households to a tractable innovation-driven growth model with different stages of economic development and an endogenous activation of innovation.

The rest of this study is organized as follows. Section 2 sets up the model. Section 3 presents the two stages of economic development. Section 4 explores the implications of heterogeneous households and natural selection. Section 5 provides empirical evidence. Section 6 concludes.

2 An R&D-based growth model with natural selection

To model natural selection, we introduce heterogeneous households and endogenous fertility to the seminal Romer model. To keep the model tractable, we consider a simple structure of overlapping generations (OLG) and human capital accumulation.[9] Each individual lives for

[5] Other early studies on endogenous takeoff and economic growth include Hansen and Prescott (2002), Jones (2001) and Kalemli-Ozcan (2002).

[6] See Galor and Klemp (2019) for empirical evidence.

[7] Galor and Maov (2001) argue that this result is generalizable to the case of heterogeneity in ability.

[8] See for example, Chu, Fan and Wang (2020) on status-seeking culture, Chu, Kou and Wang (2020) on intellectual property rights, Iacopetta and Peretto (2021) on corporate governance, Chu, Furukawa and Wang (2022) on rent-seeking government, Chu, Peretto and Wang (2022) on agricultural revolution, and Chu, Peretto and Xu (2023) on international trade.

[9] The formulation is based on Chu, Furukawa and Zhu (2016) and Chu, Kou and Wang (2022), who however focus on homogeneous households and exogenous fertilty.

three periods. In the young age, the individual accumulates human capital. In the working age, the individual allocates her time between work, fertility and education of the next generation. In the old age, the individual consumes her saving. Saving is required in the OLG model of innovation-driven growth because inventions are owned by agents as assets.

2.1 Heterogeneous households

There is a unit continuum of households indexed by $i \in [0,1]$. Within household i, the utility of an individual who works at time t is given by[10]

$$U^t(i) = u\left[n_t(i), h_{t+1}(i), c_{t+1}(i)\right] = \eta \ln n_t(i) + \gamma \ln h_{t+1}(i) + \ln c_{t+1}(i), \tag{1}$$

where $c_{t+1}(i)$ is the individual's consumption at time $t+1$, $n_t(i)$ denotes the number of children the individual has at time t, $\eta > 0$ is the fertility preference parameter, $h_{t+1}(i)$ denotes the level of human capital that the individual passes onto each child, and γ is the quality preference parameter. We assume that all individuals within the same household i have the same level of human capital at time 0. Then, they will also have the same level of human capital for all t as an endogenous outcome.

The individual allocates $e_t(i)$ units of time to her children's education. The accumulation equation of human capital is given by[11]

$$h_{t+1}(i) = \phi(i) e_t(i) + (1-\delta) h_t(i), \tag{2}$$

where the parameter $\delta \in (0,1)$ is the depreciation rate of human capital that a generation passes onto the next.[12,13] As for the ability parameter $\phi(i) > 0$ of household i,[14] it is heterogeneous across households and follows a general distribution with the following mean:[15]

$$\overline{\phi} \equiv \int_0^1 \phi(i) di.$$

The heterogeneity of households is captured by their differences in $\phi(i)$, which in turn give rise to an endogenous distribution of human capital. We focus on heterogeneity in $\phi(i)$ because it

[10] de la Croix and Doepke (2003) consider a similar utility function by assuming $\eta = \gamma$, such that utility depends on $\gamma \ln[n_t(i) h_{t+1}(i)]$.

[11] Our specification differs from de la Croix and Doepke (2003), which in turn is based on Lucas (1988). In the seminal Lucas model, human capital accumulation alone gives rise to long-run growth, so the addition of technological progress causes exploding growth. In our model, human capital accumulation alone gives rise to a higher level of output in the steady state, whereas long-run growth requires endogenous technological progress driven by innovation.

[12] In an OLG setting, Becker *et al.* (1990) and Blankenau and Simpson (2004) also assume intergenerational transmission of human capital, which is supported by empirical evidence; see Solon (1999) and Black and Devereux (2010) for comprehensive surveys.

[13] The quality-quantity tradeoff would still be present if (2) is replaced by $h_{t+1}(i) = \phi(i) e_t(i) + (1-\delta) h_t$, where h_t is the average level of human capital in the society. However, the population would converge to a degenerate distribution, in which households with the lowest $\phi(i)$ would dominate in the long run.

[14] Black, Devereux and Salvanes (2009) provide empirical evidence for a significant intergenerational transmission of IQ scores. See also Jones and Schneider (2006) for data on the variation of average IQ across countries.

[15] It is useful to note that $\overline{\phi}$ is the unweighted mean which is exogenous, whereas the weighted mean changes endogenously as the population share of households evolves over time.

allows for a stationary distribution of the population share of different households in the long run, whereas heterogeneity in other parameters, such as η or γ, imply that households with the largest η or smallest γ would dominate the population in the long run.

An individual in household i allocates $1 - e_t(i) - \sigma n_t(i)$ units of time to work and earns $w_t \left[1 - e_t(i) - \sigma n_t(i)\right] h_t(i)$ as real wage income, where the parameter $\sigma \in (0, 1)$ determines the time cost of fertility. For simplicity, we assume that there are economies of scale in the time spent in educating children within a family, and the cost of having more children is reflected in the time cost of childrearing.[16]

The individual devotes her entire wage income to saving at time t and consumes the return at time $t + 1$:[17]

$$c_{t+1}(i) = (1 + r_{t+1})w_t \left[1 - e_t(i) - \sigma n_t(i)\right] h_t(i), \tag{3}$$

where r_{t+1} is the real interest rate. Substituting (2) and (3) into (1), the individual maximizes

$$\max_{e_t(i),\, n_t(i)} U^t(i) = \eta \ln n_t(i) + \gamma \ln\left[\phi(i)e_t(i) + (1-\delta)h_t(i)\right] + \ln\left\{(1 + r_{t+1})w_t \left[1 - e_t(i) - \sigma n_t(i)\right] h_t(i)\right\},$$

taking $\{r_{t+1}, w_t, h_t(i)\}$ as given. The utility-maximizing level of fertility $n_t(i)$ is

$$n_t(i) = \frac{\eta}{\sigma(1 + \eta + \gamma)} \left[1 + (1 - \delta)\frac{h_t(i)}{\phi(i)}\right], \tag{4}$$

which is decreasing in $\phi(i)$ but increasing in $h_t(i)$. In other words, households with a lower ability to accumulate human capital and a higher level of human capital choose to have more children. In (4), fertility $n_t(i)$ is decreasing in $\phi(i)/h_t(i)$. As we will show, households with higher $\phi(i)$ have higher $h_t(i)$ and also higher $\phi(i)/h_t(i)$ before the level of human capital reaches the steady state, at which point all households share the same $\phi(i)/h_t(i)$. Therefore, households with higher ability $\phi(i)$ generally have higher human capital $h_t(i)$ and lower fertility $n_t(i)$, generating a negative relationship between these two variables. To understand this negative relationship, we also derive the utility-maximizing level of education $e_t(i)$ as[18]

$$e_t(i) = \frac{1}{1 + \eta + \gamma} \left[\gamma - (1 + \eta)(1 - \delta)\frac{h_t(i)}{\phi(i)}\right], \tag{5}$$

which is increasing in $\phi(i)$ but decreasing in $h_t(i)$. In summary, for a given $h_t(i)$, households with a larger $\phi(i)$ choose a higher level of education $e_t(i)$ but a smaller number $n_t(i)$ of children, reflecting the quality-quantity tradeoff. Given the same initial human capital $h_0(i) = h_0$, differences in education ability $\phi(i)$ give rise to differences in education level $e_t(i)$.[19]

Substituting (5) into (2) yields the autonomous and stable dynamics of human capital as

$$h_{t+1}(i) = \frac{\gamma}{1 + \eta + \gamma} \left[\phi(i) + (1 - \delta)h_t(i)\right], \tag{6}$$

[16] In de la Croix and Doepke (2003), childrearing also requires time as an input, but education costs income instead. Our education time cost $e_t(i)$ is equivalent to a reduction in income of $e_t(i)w_t h_t(i)$.

[17] Our results are robust to individuals consuming also in the working age; derivations available upon request.

[18] In (5), $e_0(i) = 0$ if $\phi(i) < (1 + \eta)(1 - \delta)h_0(i)/\gamma$, and $e_t(i) = 0$ until $h_t(i)$ depreciates to a level that reverses this inequality. Then, $e_t(i)$ becomes positive and remains to be so even at the steady state.

[19] These differences persist until $h_t(i)$ reaches the steady state.

where $h_{t+1}(i)$ is increasing in $\phi(i)$ and $h_t(i)$. The total amount of human capital in the economy at time t is

$$H_t = \int_0^1 h_t(i) L_t(i) di,$$

where $L_t(i)$ is the working-age population size of household i. The law of motion for $L_t(i)$ is

$$L_{t+1}(i) = n_t(i) L_t(i) = \frac{\eta}{\sigma(1+\eta+\gamma)} \left[1 + (1-\delta)\frac{h_t(i)}{\phi(i)}\right] L_t(i), \tag{7}$$

and the size of the aggregate labor force in the economy at time t is

$$L_t = \int_0^1 L_t(i) di.$$

Let's define $s_t(i) \equiv L_t(i)/L_t$ as the working-age-population (i.e., labor) share of household i.

Lemma 1 *The labor share $s_t(i)$ of household i at time $t \geq 1$ is given by*

$$s_t(i) = \frac{\prod_{\tau=0}^{t-1} n_\tau(i) L_0(i)}{\int_0^1 \prod_{\tau=0}^{t-1} n_\tau(i) L_0(i) di},$$

where the fertility decision $n_t(i)$ of household i at time $t \geq 1$ is given by

$$n_t(i) = \frac{\eta}{\sigma(1+\eta+\gamma)} \left\{ \sum_{\tau=0}^{t-1} \left[\frac{\gamma(1-\delta)}{1+\eta+\gamma}\right]^\tau + \left[\frac{\gamma(1-\delta)}{1+\eta+\gamma}\right]^t \left[1 + (1-\delta)\frac{h_0(i)}{\phi(i)}\right] \right\},$$

which is a decreasing function of $\phi(i)/h_0(i)$.
Proof. See Appendix A. ■

Notice that changes to $n_\tau(i)$ in any one period will affect $s_t(i)$ in all future generations. The reason is general and does not depend on the specific assumptions of this model: a temporary growth effect has a permanent level effect. Therefore, if the fertility rate of an ability group drops temporarily, this group would *ceteris paribus* forever have a lower population share than it would otherwise have had. As we will later see, if the high-ability household experiences a temporary reproduction loss, the economy will have a lower share of high-ability people forever. We will also show that this loss will permanently lower human capital, innovation and economic growth.

2.2 Final good

Perfectly competitive firms use the following production function to produce final good Y_t, which is chosen as the numeraire:

$$Y_t = H_{Y,t}^{1-\alpha} \int_0^{N_t} X_t^{\alpha}(j)dj, \tag{8}$$

where the parameter $\alpha \in (0,1)$ determines production labor intensity $1-\alpha$, and $H_{Y,t}$ denotes human-capital-embodied production labor. $X_t(j)$ denotes a continuum of differentiated intermediate goods indexed by $j \in [0, N_t]$. Firms maximize profit, and the conditional demand functions for $H_{Y,t}$ and $X_t(j)$ are given by

$$w_t = (1-\alpha)\frac{Y_t}{H_{Y,t}}, \tag{9}$$

$$p_t(j) = \alpha \left[\frac{H_{Y,t}}{X_t(j)}\right]^{1-\alpha}. \tag{10}$$

2.3 Intermediate goods

Each intermediate good j is produced by a monopolistic firm, which uses a one-to-one linear production function that transforms $X_t(j)$ units of final good into $X_t(j)$ units of intermediate good $j \in [0, N_t]$. The profit function is

$$\pi_t(j) = p_t(j)X_t(j) - X_t(j), \tag{11}$$

where the marginal cost of production is constant and equal to one (recall that final good is the numeraire). The monopolist maximizes (11) subject to (10) to derive the monopolistic price as

$$p_t(j) = \frac{1}{\alpha} > 1, \tag{12}$$

where $1/\alpha$ is the markup ratio. One can show that $X_t(j) = X_t$ for all $j \in [0, N_t]$ by substituting (12) into (10). Then, we substitute (10) and (12) into (11) to derive the equilibrium amount of monopolistic profit as

$$\pi_t = \left(\frac{1}{\alpha} - 1\right) X_t = (1-\alpha)\alpha^{(1+\alpha)/(1-\alpha)} H_{Y,t}. \tag{13}$$

2.4 R&D

We denote v_t as the value of a newly invented intermediate good at the end of time t. The value of v_t is given by the present value of future profits from time $t+1$ onwards:

$$v_t = \sum_{s=t+1}^{\infty} \left[\pi_s / \prod_{\tau=t+1}^{s} (1+r_\tau)\right]. \tag{14}$$

Competitive R&D entrepreneurs invent new products by employing $H_{R,t}$ units of human-capital-embodied labor. We specify the following innovation process:

$$\Delta N_t = \frac{\theta N_t H_{R,t}}{L_t}, \tag{15}$$

where $\Delta N_t \equiv N_{t+1} - N_t$. The parameter $\theta > 0$ determines R&D productivity $\theta N_t / L_t$, where N_t captures intertemporal knowledge spillovers as in Romer (1990) and $1/L_t$ captures a dilution effect that removes the scale effect.[20] If the following free-entry condition holds:

$$\Delta N_t v_t = w_t H_{R,t} \Leftrightarrow \frac{\theta N_t v_t}{L_t} = w_t, \tag{16}$$

then R&D $H_{R,t}$ would be positive at time t. If $\theta N_t v_t / L_t < w_t$, then R&D does not take place at time t (i.e., $H_{R,t} = 0$). Lemma 2 provides the condition for $H_{R,t} > 0$, which requires R&D productivity θ to be sufficiently high in order for innovation to take place.

Lemma 2 *R&D $H_{R,t}$ is positive at time t if and only if the following inequality holds:*

$$\int_0^1 \left[1 - e_t(i) - \sigma n_t(i)\right] h_t(i) s_t(i) di > \frac{1}{\theta}. \tag{17}$$

Proof. See Appendix A. ■

2.5 Aggregation

Imposing symmetry on (8) yields $Y_t = H_{Y,t}^{1-\alpha} N_t X_t^{\alpha}$. Then, we substitute (10) and (12) into this equation to derive the aggregate production function as

$$Y_t = \alpha^{2\alpha/(1-\alpha)} N_t H_{Y,t}. \tag{18}$$

Using $N_t X_t = \alpha^2 Y_t$, we obtain the following resource constraint on final good:

$$C_t = Y_t - N_t X_t = (1 - \alpha^2) Y_t, \tag{19}$$

where C_t denotes aggregate consumption. Finally, the resource constraint on human-capital-embodied labor is

$$\int_0^1 \left[1 - e_t(i) - \sigma n_t(i)\right] h_t(i) L_t(i) di = H_{Y,t} + H_{R,t}. \tag{20}$$

[20] See Laincz and Peretto (2006) for a discussion of the scale effect.

2.6 Equilibrium

The equilibrium is a sequence of allocations $\{X_t(j), Y_t, e_t(i), n_t(i), c_t(i), C_t, h_t(i), H_t, H_{Y,t}, H_{R,t}, L_t\}$ and prices $\{p_t(j), w_t, r_t, v_t\}$ that satisfy the following conditions:

- individuals choose $\{e_t(i), n_t(i), c_t(i)\}$ to maximize utility taking $\{r_{t+1}, w_t, h_t(i)\}$ as given;
- competitive firms produce Y_t to maximize profit taking $\{p_t(j), w_t\}$ as given;
- a monopolistic firm produces $X_t(j)$ and chooses $p_t(j)$ to maximize profit;
- competitive entrepreneurs perform R&D to maximize profit taking $\{w_t, v_t\}$ as given;
- the market-clearing condition for the final good holds such that $Y_t = N_t X_t + C_t$;
- the resource constraint on human-capital-embodied labor holds such that $H_{Y,t} + H_{R,t} = \int_0^1 [1 - e_t(i) - \sigma n_t(i)] \, h_t(i) L_t(i) di$;
- total saving equals asset value such that $w_t \int_0^1 [1 - e_t(i) - \sigma n_t(i)] \, h_t(i) L_t(i) di = N_{t+1} v_t$.

3 Stages of economic development

Our model features two stages of economic development. The first stage features only human capital accumulation. The second stage features both human capital accumulation and innovation.[21] The activation of innovation and the resulting transition from the first stage to the second stage are endogenous and do not always occur.

3.1 Stage 1: Human capital accumulation only

The initial level of human capital for each individual in household i is $h_0(i)$. Suppose the following inequality holds at time 0:

$$\int_0^1 [1 - e_0(i) - \sigma n_0(i)] \, h_0(i) s_0(i) di = \frac{1}{1+\eta+\gamma} \int_0^1 \left[1 + (1-\delta) \frac{h_0(i)}{\phi(i)} \right] h_0(i) s_0(i) di < \frac{1}{\theta}, \quad (21)$$

which uses (4) and (5). In (21), both the initial labor share $s_0(i) \equiv L_0(i)/L_0$ and initial human capital $h_0(i)$ are exogenously given. Then, Lemma 2 implies that $H_{R,0} = 0$ and

$$H_{Y,0} = \frac{1}{1+\eta+\gamma} \int_0^1 \left[1 + (1-\delta) \frac{h_0(i)}{\phi(i)} \right] h_0(i) L_0(i) di. \quad (22)$$

[21] See Iacopetta (2010) who considers a model in which innovation occurs before human capital accumulation.

In this stage of development, the economy features only human capital accumulation. Human capital $h_t(i)$ accumulates according to the autonomous and stable dynamics in (6), and $s_t(i)$ evolves according to Lemma 1. However, so long as the following inequality holds at time t:

$$\int_0^1 [1 - e_t(i) - \sigma n_t(i)] h_t(i) s_t(i) di = \frac{1}{1+\eta+\gamma} \int_0^1 \left[1 + (1-\delta)\frac{h_t(i)}{\phi(i)}\right] h_t(i) s_t(i) di < \frac{1}{\theta}, \quad (23)$$

we continue to have $H_{R,t} = 0$ and

$$H_{Y,t} = \frac{1}{1+\eta+\gamma} \int_0^1 \left[1 + (1-\delta)\frac{h_t(i)}{\phi(i)}\right] h_t(i) L_t(i) di. \quad (24)$$

Substituting (24) into (18) yields the level of output per worker as

$$y_t \equiv \frac{Y_t}{L_t} = \alpha^{2\alpha/(1-\alpha)} N_0 \frac{H_{Y,t}}{L_t} = \frac{\alpha^{2\alpha/(1-\alpha)} N_0}{1+\eta+\gamma} \int_0^1 \left[1 + (1-\delta)\frac{h_t(i)}{\phi(i)}\right] h_t(i) s_t(i) di, \quad (25)$$

where N_0 remains at the initial level and output increases as human capital accumulates.

3.2 Does innovation occur?

Equation (6) shows that human capital $h_t(i)$ converges to a steady state given by

$$h^*(i) = \frac{\gamma \phi(i)}{1+\eta+\gamma\delta}, \quad (26)$$

which is increasing in household i's ability $\phi(i)$. Substituting (26) into (4) and (5) yields the steady-state levels of education and fertility given by

$$e^*(i) = e^* = \frac{\gamma\delta}{1+\eta+\gamma\delta}, \quad (27)$$

$$n^*(i) = n^* = \frac{\eta}{\sigma(1+\eta+\gamma\delta)}, \quad (28)$$

where we assume positive population growth (i.e., $n^* > 1$) by imposing $\eta > (1+\gamma\delta)\sigma/(1-\sigma)$. Also, n^* is the same across all households because they are independent of $\phi(i)$. In other words, the negative effect of $\phi(i)$ and the positive effect of $h^*(i)$ on $n^*(i)$ cancel each other. As a result, the distribution of the population share of different households is stationary in the long run.

In this case, Lemma 2 implies that if the following inequality holds:

$$(1 - e^* - \sigma n^*) \int_0^1 h^*(i) s^*(i) di = \frac{\gamma}{(1+\eta+\gamma\delta)^2} \int_0^1 \phi(i) s^*(i) di > \frac{1}{\theta}, \quad (29)$$

then human capital accumulation eventually triggers the activation of innovation, under which the R&D condition in (16) holds and R&D $H_{R,t}$ becomes positive. Therefore, the endogenous activation of innovation requires a sufficiently large R&D productivity parameter θ, such that (29) holds before human capital converges to a steady state. If innovation does not occur, then the economy features only human capital accumulation and converges to the following steady-state level of output per worker:

$$y^* = \frac{\gamma \alpha^{2\alpha/(1-\alpha)} N_0}{(1+\eta+\gamma\delta)^2} \int_0^1 \phi(i) s^*(i) di,$$

which uses (26) in (25).

3.3 Stage 2: Innovation and human capital accumulation

We now consider the case in which the activation of innovation has occurred and derive the equilibrium growth rate in the presence of innovation. Substituting (18) into (9) yields the equilibrium wage rate as

$$w_t = (1-\alpha)\alpha^{2\alpha/(1-\alpha)} N_t. \tag{30}$$

Then, substituting (30) into (16) yields the equilibrium invention value as

$$\frac{v_t}{L_t} = \frac{(1-\alpha)\alpha^{2\alpha/(1-\alpha)}}{\theta}. \tag{31}$$

The structure of overlapping generations implies that the value of assets at the end of time t must equal the amount of saving at time t given by wage income at time t:

$$N_{t+1} v_t = w_t \int_0^1 \left[1 - e_t(i) - \sigma n_t(i)\right] h_t(i) L_t(i) di = w_t (H_{Y,t} + H_{R,t}), \tag{32}$$

where the second equality uses (20). Substituting (30) and (31) into (32) yields

$$N_{t+1} = \frac{\theta N_t}{L_t}(H_{Y,t} + H_{R,t}). \tag{33}$$

Combining (15) and (33) yields the equilibrium level of $H_{Y,t}$ as

$$\frac{H_{Y,t}}{L_t} = \frac{1}{\theta} \tag{34}$$

for all t. Substituting (4), (5) and (34) into (20) yields the equilibrium level of $H_{R,t}$ as

$$\frac{H_{R,t}}{L_t} = \int_0^1 \left[1 - e_t(i) - \sigma n_t(i)\right] h_t(i) s_t(i) di - \frac{H_{Y,t}}{L_t} = \frac{1}{1+\eta+\gamma} \int_0^1 \left[1 + (1-\delta)\frac{h_t(i)}{\phi(i)}\right] h_t(i) s_t(i) di - \frac{1}{\theta}. \tag{35}$$

We can now substitute (35) into (15) to derive the equilibrium growth rate of N_t as

$$g_t \equiv \frac{\Delta N_t}{N_t} = \frac{\theta H_{R,t}}{L_t} = \frac{\theta}{1+\eta+\gamma} \int_0^1 \left[1 + (1-\delta)\frac{h_t(i)}{\phi(i)}\right] h_t(i) s_t(i) di - 1, \tag{36}$$

which is also the equilibrium growth rate of output per worker $y_t = \alpha^{2\alpha/(1-\alpha)} N_t/\theta$. Finally, the steady-state equilibrium growth rate of N_t and y_t is

$$g^* = \frac{\theta\gamma}{(1+\eta+\gamma\delta)^2} \int_0^1 \phi(i) s^*(i) di - 1. \tag{37}$$

In the steady state, $s^*(i)$ is also the population share of household i and still depends on the initial distribution of $h_0(i)$ and the exogenous distribution of $\phi(i)$ as shown in Lemma 1.

4 Heterogeneous households and evolutionary differences

Equation (21) shows that the activation of innovation-driven growth occurs at time 0 if and only if the following inequality holds:

$$\frac{1}{1+\eta+\gamma}\int_0^1 \left[1+(1-\delta)\frac{h_0(i)}{\phi(i)}\right] h_0(i)s_0(i)di > \frac{1}{\theta}. \tag{38}$$

Suppose we consider a useful benchmark of an equal initial labor share $s_0(i) = 1$ and an equal initial level of human capital $h_0(i) = h_0$ for all $i \in [0, 1]$. Then, the left-hand side of (38) simplifies to

$$\frac{h_0}{1+\eta+\gamma}\left[1+(1-\delta)h_0\int_0^1 \frac{1}{\phi(i)}di\right] > \frac{h_0}{1+\eta+\gamma}\left[1+\frac{(1-\delta)h_0}{\overline{\phi}}\right], \tag{39}$$

where $\int_0^1[1/\phi(i)]di > 1/\overline{\phi}$ due to Jensen's inequality. In other words, the presence of heterogeneity in $\phi(i)$ makes the activation of innovation-driven growth more likely to occur at time 0 than the absence of heterogeneity (i.e., $\phi(i) = \overline{\phi}$ for all $i \in [0, 1]$) does. Due to heterogeneity, some households supply more human capital for production and innovation while others supply less. Equation (39) implies that the former effect dominates the latter effect such that the initial amount of human capital available for production and innovation increases as a result of heterogeneity. The intuition can be explained as follows.

Although some low-ability households may devote almost no time to education and most of their time to work (and fertility), high-ability households always spend some time to work, as the following shows:

$$1-e_0(i)-\sigma n_0(i) = \frac{1}{1+\eta+\gamma}\left[1+\frac{(1-\delta)h_0}{\phi(i)}\right] > \frac{1}{1+\eta+\gamma} > 0.$$

The convexity of $1/\phi(i)$ in $1 - e_0(i) - \sigma n_0(i)$ gives rise to the positive effect of heterogeneity on the amount of human capital available for production and innovation. To put it differently, the low-ability households being less willing to educate their children contribute to a larger workforce, which in turn rewards the innovation pioneers with more profits extracted from a larger market size of the economy. We summarize this result in the following lemma.

Lemma 3 *Heterogeneity makes it more likely for innovation to be activated at time 0.*

Proof. If the following inequality holds:

$$\frac{h_0}{1+\eta+\gamma}\left[1+(1-\delta)h_0\int_0^1 \frac{1}{\phi(i)}di\right] > \frac{1}{\theta} > \frac{h_0}{1+\eta+\gamma}\left[1+\frac{(1-\delta)h_0}{\overline{\phi}}\right], \tag{40}$$

which is a nonempty parameter space due to $\int_0^1[1/\phi(i)]di > 1/\overline{\phi}$, then the takeoff of the economy occurs at time 0 under heterogeneous households but not under homogeneous households. ■

Next we examine how the labor share of households evolves over time. Given the benchmark of an equal initial labor share $s_0(i) = 1$ and an equal initial level of human capital $h_0(i) = h_0$ for all $i \in [0, 1]$, the fertility of household i at time 0 is

$$n_0(i) = \frac{\eta}{\sigma(1 + \eta + \gamma)} \left[1 + (1 - \delta)\frac{h_0}{\phi(i)}\right],$$

which is decreasing in $\phi(i)$. For households with $\phi(i) > \overline{\phi}$, their growth rate $n_0(i)$ would be lower than $n_0(\overline{\phi})$. However, they will have a higher level of human capital in the next period:

$$h_1(i) = \gamma\frac{\phi(i) + (1 - \delta)h_0}{1 + \eta + \gamma} > \gamma\frac{\overline{\phi} + (1 - \delta)h_0}{1 + \eta + \gamma}.$$

This higher level of human capital gives rise to a higher growth rate $n_1(i)$ and reduces the difference between $n_1(i)$ and $n_1(\overline{\phi})$. However, as shown in Lemma 1, $n_t(i)$ remains lower than $n_t(\overline{\phi})$ for $\phi(i) > \overline{\phi}$ until $h_t(i)$ converges to its steady-state level in (26) at which point the population growth rate of all households $i \in [0, 1]$ converges to n^* in (28). Therefore, the population growth rates of households with $\phi(i) > \overline{\phi}$ are lower than the population growth rates of households with $\phi(i) < \overline{\phi}$ until $h_t(i)$ converges to its steady-state level in (26). This temporary evolutionary disadvantage of high-ability households will never be compensated despite population trends being equal across households in the long run.

The above analysis implies that there exists a threshold for $\phi(i)$ above (below) which $s^*(i) < 1$ ($s^*(i) > 1$). This in turn implies that[22]

$$\int_0^1 \phi(i)s^*(i)di < \int_0^1 \phi(i)di = \overline{\phi}, \tag{41}$$

because the households with larger $\phi(i)$ end up having a lower steady-state population share $s^*(i)$. Therefore, we also have the following inequality:

$$g^* = \frac{\theta\gamma}{(1 + \eta + \gamma\delta)^2} \int_0^1 \phi(i)s^*(i)di - 1 < \frac{\theta\gamma}{(1 + \eta + \gamma\delta)^2}\overline{\phi} - 1, \tag{42}$$

where the right-hand side of the inequality is the steady-state innovation-driven growth rate under homogeneous households (i.e., $\phi(i) = \overline{\phi}$ for all $i \in [0, 1]$) in an economy that has experienced the transition to innovation. In other words, the steady-state growth rate g^* becomes lower because the heterogeneity in households and the temporary evolutionary disadvantage of the high-ability households reduce the average level of human capital and consequently the rate of innovation (recall that $g_t = \theta H_{R,t}/L_t$) in the long run. We summarize the above result in the following proposition.

Proposition 1 *The temporary evolutionary disadvantage of the high-ability households causes a lower steady-state equilibrium growth rate g^* than the case of homogeneous households without natural selection.*

Proof. See Appendix A. ■

[22] See the proof of Proposition 1 in Appendix A.

4.1 A parametric example

In this section, we provide a simple parametric example to illustrate our results more clearly and prepare for the empirical analysis in Section 5. We consider two types of households. Specifically, $\phi(i) = \overline{\phi} + \varsigma$ for $i \in [0, 0.5]$ and $\phi(j) = \overline{\phi} - \varsigma$ for $j \in [0.5, 1]$. As before, the households own the same initial amount of human capital (i.e., $h_0(i) = h_0$ for $i \in [0, 1]$). Their initial population shares are also the same (i.e., $s_0(i) = 1$ for $i \in [0, 1]$); in this case, the mean of $\phi(i)$ is simply $\overline{\phi}$ and the coefficient of variation in $\phi(i)$ is $\varsigma/\overline{\phi}$. Therefore, for a given $\overline{\phi}$, an increase in ς raises the coefficient of variation in $\phi(i)$ and also makes (40) more likely to hold by raising $\int_0^1 [1/\phi(i)]di = 1/(\overline{\phi} - \varsigma^2/\overline{\phi}) > 1/\overline{\phi}$.

From (26), their steady-state levels of human capital are different and given by $h^*(i) = \gamma(\overline{\phi} + \varsigma)/(1 + \eta + \gamma\delta)$ for $i \in [0, 0.5]$ and $h^*(j) = \gamma(\overline{\phi} - \varsigma)/(1 + \eta + \gamma\delta)$ for $j \in [0.5, 1]$. From (42), the steady-state growth rate g^* is given by

$$g^* = \frac{\theta\gamma}{(1+\eta+\gamma\delta)^2}\left[(\overline{\phi}+\varsigma)s_H^* + (\overline{\phi}-\varsigma)s_L^*\right] - 1 = \frac{\theta\gamma}{(1+\eta+\gamma\delta)^2}\left\{\overline{\phi} + \varsigma\left[\underset{-}{s_H^*(\varsigma)} - \underset{+}{s_L^*(\varsigma)}\right]\right\} - 1, \tag{43}$$

where $s_L^* \equiv \int_{0.5}^1 s^*(j)dj = s^*(j)/2$ is the steady-state population share of household $j \in [0.5, 1]$ with low ability $\phi(j) = \overline{\phi} - \varsigma$ whereas $s_H^* \equiv \int_0^{0.5} s^*(i)di = s^*(i)/2$ is the steady-state population share of household $i \in [0, 0.5]$ with high ability $\phi(i) = \overline{\phi} + \varsigma$. We note that $s_H^* + s_L^* = 1$. Then, from Lemma 1, we have

$$\frac{s_L^*}{s_H^*} = \frac{\prod_{t=0}^{\infty} n_t(j)}{\prod_{t=0}^{\infty} n_t(i)} > 1, \tag{44}$$

where

$$\begin{aligned} n_t(j) &= \frac{\eta}{\sigma(1+\eta+\gamma)}\left\{\sum_{\tau=0}^{t-1}\left[\frac{\gamma(1-\delta)}{1+\eta+\gamma}\right]^{\tau} + \left[\frac{\gamma(1-\delta)}{1+\eta+\gamma}\right]^{t}\left[1 + (1-\delta)\frac{h_0}{\overline{\phi}-\varsigma}\right]\right\}, \\ n_t(i) &= \frac{\eta}{\sigma(1+\eta+\gamma)}\left\{\sum_{\tau=0}^{t-1}\left[\frac{\gamma(1-\delta)}{1+\eta+\gamma}\right]^{\tau} + \left[\frac{\gamma(1-\delta)}{1+\eta+\gamma}\right]^{t}\left[1 + (1-\delta)\frac{h_0}{\overline{\phi}+\varsigma}\right]\right\}. \end{aligned}$$

Therefore, s_L^*/s_H^* is increasing in ς, which together with $s_H^* + s_L^* = 1$ implies that s_L^* is increasing in ς and s_H^* is decreasing in ς as stated in (43).

In summary, an increase in ς leads to an immediate increase in the coefficient of variation in $\phi(i)$ given by $\varsigma/\overline{\phi}$ and a subsequent decrease in the steady-state growth rate g^* given by (43) by reducing the average level of human capital and the level of innovation in the long run due to the temporary evolutionary disadvantage of the high-ability households. In the next section, we will test this theoretical prediction using cross-country data.

Corollary 1 *Raising ς causes a larger coefficient of variation in $\phi(i)$ and a lower steady-state growth rate g^*.*

4.1.1 Education policy

We now consider a simple policy experiment. Any proportional shock $\lambda_e > 1$ to the household's education abilities will scale up all $\phi(i)$, but it will also emphasize differences. High-ability households' ability will become $\lambda_e\left(\overline{\phi}+\varsigma\right)$, while low-ability households' ability will become $\lambda_e\left(\overline{\phi}-\varsigma\right)$. Since $\overline{\phi}-\varsigma > 0$, the effects on fertility and on $n_t(j)$ and $n_t(i)$ are both negative. This result means that education facilities and support will reduce population growth by increasing the family's potential to educate. For example, after decades of education policies, China's fertility rate has dropped despite the 2016 abandonment of the single-child policy. Our model allows arguing that China's recent population decline is not easily revertible because the country's fertility transition to quality children is a by-product of its inclusive and meritocratic education tradition. Will it hamper economic growth? According to our model, it will not. The reader can easily prove that

$$\frac{1+(1-\delta)\frac{h_0}{\lambda_e\left(\overline{\phi}-\varsigma\right)}}{1+(1-\delta)\frac{h_0}{\lambda_e\left(\overline{\phi}+\varsigma\right)}}$$

decreases in λ_e, which implies - by (44) - that s_L^*/s_H^* decreases as well, thereby leading to an increase in g^*. Therefore, we can state that:

Corollary 2 *A policy that proportionally raises all education abilities will lead to a decrease in fertility and an increase in long-term economic growth.*

5 Empirical evidence

5.1 Empirical strategy

The main result in this study is driven by the quality-quantity tradeoff in fertility transition highlighted by Galor (2005, 2011, 2022) and others in the related literature.[23] The core of this transition is the parents' decision to educate their children: education takes time and resources, and hence, it cannot be effective on too many children. Using data from a sample of 137 countries, Figure 1 in Appendix B presents a well known negative relationship between fertility and education and shows that this relationship appears even in cross-country panel data. This well-documented quality-quantity tradeoff implies that households with higher education experience an evolutionary disadvantage and represent a smaller share of the population over time. This stylized fact is consistent with (4) in our model, in which households with higher ability $\phi(i)$ generally have higher human capital $h_t(i)$ and lower fertility $n_t(i)$, generating a negative relationship between these two endogenous variables.[24]

[23] See for example, Becker *et al.* (2010), Fernihough (2017) and Klemp and Weisdorf (2019).

[24] Note that $\phi(i)/h_t(i)$ is increasing in $\phi(i)$ until $h_t(i)$ reaches its steady state $h^*(i)$ in (26), at which point $\phi(i)/h^*(i) = (1+\eta+\gamma\delta)/\gamma$ for all i.

Corollary 1 in the theoretical analysis demonstrates that the negative relationship between fertility and education further implies a negative impact of heterogeneity in the ability to accumulate human capital on economic growth. In this section, we use cross-country data to test this theoretical prediction. Specifically, we use global standardized tests of students' academic ability as our proxy variable for educational ability. We focus on the Trends in International Mathematics and Science Study (TIMSS), a renowned assessment that evaluates the mathematics and science performance of students around the world. From this data, we calculate the coefficient of variation of scores as a measure of the heterogeneity of educational ability within a country.

As a standardized test, the indicators constructed from TIMSS can be directly compared across countries.[25] It is less influenced by institutional factors (e.g., educational systems) compared to measures based on years of education, which more precisely focus on the abilities and endowments of the population. Angrist *et al.* (2021) demonstrate that overall student performance across countries undergoes minimal changes over time, yet substantial differences exist at the national level. Similarly, the coefficient of variation of TIMSS scores exhibits small changes over time but considerable differences across countries, making it an effective tool for analyzing the impact of heterogeneity of ability at the national level.[26]

Given that the heterogeneity of educational ability may correlate with other determinants of economic growth, our analysis could be subject to omitted variable bias. To address this, we account for cultural characteristics like time preference, which could influence both our dependent and outcome variables (Galor and Özak, 2016).[27] Additionally, we include a comprehensive set of geographic variables as suggested by Arbatlı *et al.* (2020).[28] Furthermore, recognizing the impact of continent-specific factors (e.g., land suitability for agriculture, disease environment, and climatic conditions) on economic growth, we control for continent-specific fixed effects.

We focus on 67 countries that participated in the TIMSS test and study the impacts of heterogeneity of educational ability on economic growth from 1951 to 2017. Our regression equation is as follows:

$$y_{i,t} = \beta_0 + \beta_1 var_i + \gamma Z_i + \varphi_t + \varphi_c + \epsilon_{i,t},$$

[25] The program evaluates both fourth and ninth-grade students. Since fourth-grade assessments cover a larger and more representative sample of students, they offer a more accurate depiction of nationwide differences compared to assessments focused solely on ninth-grade students. Therefore, our analysis focuses on fourth-grade students.

[26] The coefficient of variation in overall TIMSS scores (math and science combined) exhibited a modest 1.28% change from 2003 to 2019. A similar pattern emerges when analyzing mathematical and scientific ability separately. The coefficient of variation in mathematical ability increased by only 1.6% during this period. For scientific ability, the coefficient of variation increased by a modest 1.91% from 2003 to 2019.

[27] We follow Hofstede (1991) and use the average level of long-term orientation among individuals in a country as a proxy for the country's rate of time preference. As highlighted by Galor and Özak (2016), long-term orientation significantly impacts the formation of human and physical capital, technological advancement, and economic growth.

[28] Inspired by Arbatlı *et al.* (2020), who use cross-sectional data to investigate the impact of population diversity on the annual frequency of new civil conflict onsets, we introduce similar geographic variables due to the following considerations: (i) absolute distance from the equator and proximity to the nearest waterway, which influence economic development through climatological, institutional, and trade-related mechanisms; (ii) geographical isolation, which provides relative immunity from cross-border spillovers; and (iii) variability in land suitability for agriculture and elevation, which has been shown to foster ethnic diversity (Michalopoulos, 2012) and also impacts economic growth through various mechanisms, such as productivity.

where (i) the baseline dependent variable $y_{i,t}$ is economic growth of country i at time t, measured by the annual growth rate of GDP per capita. (ii) the independent variable var_i is the heterogeneity of educational ability in country i, which we capture by using the coefficient of variation of TIMSS scores.[29] (iii) Z_i includes time preference and geographical characteristics (e.g., distance to the nearest waterway, absolute latitude, mean elevation and agricultural land suitability, standard deviation in elevation and agricultural land suitability, and an island dummy),[30] and (iv) φ_t and φ_c are year fixed-effects and continent fixed-effects, respectively. We provide the summary statistics in Appendix B. In Corollary 1, our theoretical framework predicts that a larger coefficient of variation in ability induces a lower steady-state growth rate, which implies $\beta_1 < 0$.

5.2 Empirical results

Table 1 presents the results from our baseline cross-country analysis. The analysis begins with a bivariate regression in Column 1, which shows that the heterogeneity of education is a negative and highly significant correlate of the growth rate of GDP per capita. To provide a more visual representation of the relationship between educational ability heterogeneity and economic growth, we have created a scatter plot, which is presented in Figure 2 in Appendix B. For each country in our sample, we calculated the average annual economic growth rate over the study period. We then plotted these average growth rates against the country's heterogeneity of educational ability. The scatter plot reveals a clear negative relationship between the two variables. Notably, this inverse pattern is observed across countries from different continents, as represented by the distinct color-coded data points.

To ensure our results are not influenced by other factors, we gradually introduce controls for various variables. Column 2 includes year fixed effects to address time-varying factors, while Column 3 adds continent fixed effects to control for regional differences. In our baseline results, shown in Column 4, we include additional controls for time preference and exogenous geographical characteristics. These characteristics include distance to the nearest waterway, absolute latitude, mean elevation, mean agricultural land suitability, standard deviation in elevation and agricultural land suitability, and an island dummy variable. These results demonstrate the robustness of our core finding, regardless of the extent of controls included in the regression analysis.

However, this association may be influenced by endogeneity bias. The heterogeneity of educational ability and the spatial pattern of economic growth in the modern era could be jointly determined by unobserved cultural, institutional, and human characteristics that are not fully

[29] The TIMSS test has been conducted every four years from 1995 to 2023, with continuous testing of 4th-grade students starting in 2003. Since data for 2023 is not yet available, we utilize test data from 2003, 2007, 2011, 2015, and 2019. To quantify the heterogeneity of educational ability, we compute the coefficient of variation for each country using student scores from each test cycle. Subsequently, we derive the average heterogeneity of educational ability for each country by computing the mean across these test years.

[30] The data on time preference is sourced from Galor and Özak (2016). Data for absolute latitude, mean elevation, standard deviation in elevation, and an island dummy variable are obtained from the Geographically Based Economic Data (G-ECON) project (Nordhaus, 2006). Data for distance to the nearest waterway is from Arbatlı *et al.* (2020). The land agricultural suitability data is sourced from Michalopoulos (2012), with the mean or standard deviation at the country level reflecting the average or standard deviation value of the index across the grid cells located within a country's national borders.

Table 1: Heterogeneity of Educational Ability and Economic Growth

	Economic Growth			
	(1)	(2)	(3)	(4)
Heterogeneity of educational ability	-8.156***	-7.952***	-13.269***	-9.234***
	(2.484)	(2.279)	(2.787)	(3.366)
Time preference				0.012
				(0.009)
Distance to nearest waterway				0.134
				(0.458)
Absolute latitude				-0.023
				(0.016)
Mean elevation				0.004*
				(0.002)
Mean land suitability				0.782
				(0.561)
Standard deviation of elevation				-0.004**
				(0.002)
Standard deviation of land suitability				0.436
				(1.593)
Island nation dummy				0.114
				(0.573)
Year FE	N	Y	Y	Y
Continent FE	N	N	Y	Y
R-square	0.009	0.118	0.128	0.136
Observations	3476	3476	3476	3356

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses are clustered by country. The dependent variable is the growth rate of GDP per capita.

captured by covariates. For example, in a society that highly values knowledge and education, individuals might be strongly self-motivated to strive for academic excellence, leading to smaller differences in test scores and rapid economic growth. If we ignore these unobserved factors in our analysis, our results may be biased. To address this issue, we employ two instrumental variables. The first instrumental variable we employ is ancestry-adjusted population diversity, which has been discussed in the related literature (Ashraf and Galor, 2013; Arbatlı *et al.*, 2020). We believe this measure effectively relates to differences in educational ability and exhibits a high degree of exogeneity.[31] The second instrumental variable is migratory distance from East Africa, as utilized in previous studies (Ashraf and Galor, 2013; Arbatlı *et al.*, 2020; Ashraf *et al.*, 2021 and Galor *et al.*, 2023). As explained by Galor *et al.* (2023), this migration entailed a gradual expansion. Subgroups of individuals left their ancestral settlement in each step to establish new colonies farther away, carrying only a subset of their ancestral traits. This process, known as the Serial Founder Effect, has resulted in a negative correlation between migratory

[31] The interpersonal population diversity of each country is captured by a measure of predicted genetic diversity developed by Ashraf and Galor (2013). This measure takes into account the proportional representation of each ancestral population within a contemporary nation, the genetic diversity of each ancestral population based on its migratory distance from Africa, and the pairwise genetic distances between ancestral populations derived from their migratory distances from one another.

distance from Africa and genetic, phenotypic, and potentially phonemic heterogeneity.[32]

Column (1) of Table 2 presents the results of the 2SLS regressions using ancestry-adjusted population diversity as an instrument for the heterogeneity of educational ability. The second stage of the regression shows that the coefficient for the heterogeneity of educational ability is significantly negative, indicating a detrimental impact on economic growth. Column (2) further supports our analysis by using migratory distance from East Africa as an alternative instrumental variable. The second-stage regression results in this column also show that the coefficient for the heterogeneity of educational ability remains significantly negative.

In last two columns of Table 2 , we present the first-stage regression results with the instrumental variables. The results show that there is a significant positive correlation between the heterogeneity of educational ability and ancestry-adjusted population diversity, as shown in Column (3). Furthermore, Column (4) shows a significant negative correlation between the heterogeneity of educational ability and migratory distance from East Africa. Specifically, as migratory distance increases, the heterogeneity of educational ability decreases. These results support the validity of our instrumental variables.

Table 2: Impacts of Heterogeneity of Educational Ability Using Instrumental Variables

	Second stage		First stage	
	Economic Growth		Heterogeneity of educational ability	
	(1)	(2)	(3)	(4)
Heterogeneity of educational ability	-27.314***	-32.498***		
	(9.099)	(10.165)		
Population diversity (ancestry adjusted)			1.432***	
			(0.389)	
Migratory distance from East Africa				-0.011***
				(0.003)
Controls	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Continent FE	Y	Y	Y	Y
R-square	0.111	0.101	0.738	0.740
Observations	3065	3065	3065	3065
First-stage F statistic			13.064	10.099

Note: p < 0.01, ** p < 0.05, * p < 0.1. Standard errors in parentheses are clustered by country. The dependent variables in columns (1) and (2) are the growth rate of GDP per capita. Column (1) shows the second-stage results using population diversity (ancestry adjusted) as the instrumental variable, with the corresponding first-stage results presented in Column (3). Column (2) presents the second-stage results using prehistoric migratory distance from East Africa as the instrumental variable, with the corresponding first-stage results shown in Column (4). Controls include time preference, distance to the nearest waterway, absolute latitude, mean elevation, mean agricultural land suitability, standard deviation in elevation, standard deviation in agricultural land suitability, and an island dummy.

Next, we use education and innovation as alternative proxies for economic growth. In the first two columns of Table 3, the dependent variables are the share of the population with at least some primary education and the logarithm of the average years of education, respectively. These variables capture the average level of education. In Columns (3) and (4),

[32] The data on ancestry-adjusted population diversity and migratory distance from East Africa are sourced from Arbatlı *et al.* (2020).

the dependent variables are the logarithm of the number of researchers in R&D (per million people) and the logarithm of the number of patent applications, respectively. These variables capture the rate of innovation. The coefficients of the heterogeneity of educational ability in Table 3 are all significantly negative, which implies that heterogeneity of educational ability negatively impacts education and innovation as well, which further support the predictions of our theoretical model.[33]

Table 3: Impacts of Heterogeneity of Educational Ability on Education and Innovation

	(1) Share of schooling	(2) Average education	(3) R&D	(4) Patent
Heterogeneity of educational ability	-109.662**	-2.080**	-12.572***	-12.686***
	(50.193)	(0.986)	(1.360)	(4.417)
Controls	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Continent FE	Y	Y	Y	Y
R-square	0.639	0.737	0.761	0.540
Observations	603	603	855	1735

Note: $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses are clustered by country. In columns 1-2, the dependent variables are the share of the population with primary schooling and the log of average years of education in the first two columns, respectively. In columns 3-4, the dependent variables are the log of the number of researchers in R&D (per million people) and the log of the number of patent applications, respectively. Controls include time preference, distance to the nearest waterway, absolute latitude, mean elevation, mean agricultural land suitability, standard deviation in elevation and agricultural land suitability, and an island dummy.

In the subsequent analysis, we will conduct a series of robustness checks to affirm the reliability of our core findings. First, we use other indicators to estimate the heterogeneity of educational ability. The TIMSS tests are segmented into Mathematics and Science categories. These tests yield scores for both Mathematics and Science, from which we derive the coefficient of variation for educational ability in each subject. This coefficient reflects the diversity of mathematical and scientific skills, respectively. As depicted in columns (1) and (2) of Table 5 in Appendix B, both of them have a negative impact on economic growth. Lastly, we utilize years of education as a proxy for educational ability and compute its coefficient of variation. The coefficient remains notably negative, providing additional backing for our theoretical forecasts. Second, we introduce additional variables to demonstrate the robustness of our main findings. Specifically, in column (1) of Table 6 in Appendix B, we include institutional controls such as legal origins. In column (2), we add diversity measures such as ethnic fractionalization, ethnolinguistic polarization, and linguistic fractionalization.[34] In column (3), We include the lagged logarithm of GDP per capita and the lagged logarithm of population. Table 6 in Appendix B confirms that our baseline findings remain qualitatively intact after accounting for these additional controls. Additionally, even after controlling for the lagged growth rate of GDP per

[33] The share of the population with at least some primary education and the logarithm of the average years of education are calculated from the Barro-Lee educational attainment dataset. The number of researchers in R&D, and the number of patent applications are sourced from the World Bank.

[34] The data on legal origins, ethnic fractionalization, ethnolinguistic polarization, and linguistic fractionalization come from Arbatlı et al.(2020).

capita to account for unobserved economic growth potential, our results still hold. Third, we use the subsample to confirm that the impact of educational ability heterogeneity is prevalent across all continents. Specifically, we group the data by continent: Africa, Europe, Asia, Oceania, and the Americas. As shown in Table 7 in Appendix B, despite the significant cultural and institutional differences among these continents, our empirical results hold for each of them. This indicates that our main finding is not driven by outliers from any specific continent.

6 Conclusion

In this study, we have constructed an innovation-driven growth model with endogenous takeoff and endogenous fertility, elucidating the natural selection of heterogeneous households, differentiated by their ability to accumulate human capital. The followings are the core findings of our research. In terms of short-run dynamics, we show that a survival-of-the-weakest scenario emerges in the short run, where high-ability households experience a temporary evolutionary disadvantage, later offset by human capital accumulation. In terms of long-run implications, the temporary disadvantage of high-ability households has a lingering negative impact on R&D, technological progress and long-term economic growth.

As for empirical evidence, our cross-country data analysis affirms the model's predictions, highlighting the adverse effects of educational heterogeneity on long-run education, innovation, and economic growth. To mitigate any potential endogeneity issue, we adopt instrumental variables, but we acknowledge that this may not completely resolve the issue. In terms of theoretical contributions, this work introduces the novel concept of natural selection of heterogeneous households to the innovation-driven growth model, thereby enriching the existing literature on economic growth. While our model yields several noteworthy insights, it also opens up intriguing avenues for future research. One potential extension could involve a more granular examination of the policy implications. Understanding how government interventions or institutional reforms might affect the complex interplay between natural selection, human capital accumulation and economic growth could further refine the applicability of our model to real-world scenarios.

Additionally, the integration of other socio-economic factors, such as cultural attitudes, as in Cozzi (1998) and Tabellini (2010), and their dynamics, as in Bisin and Verdier (1998, 2000, 2001 and 2017), could add layers of realism and relevance to the theory. The exploration of these and other extensions could provide valuable insights into how the subtle dynamics of natural selection within heterogeneous households shape macroeconomic outcomes. In conclusion, our study not only contributes to the understanding of innovation-driven growth, endogenous takeoff and natural selection but also raises questions for subsequent research, emphasizing the multifaceted nature of human capital, fertility choices, and economic development.

References

[1] Aghion, P., and Howitt, P., 1992. A model of growth through creative destruction. *Econometrica*, 60, 323-351.

[2] Alesina, A., Seror, M., Yang, D. Y., You, Y., and Zeng, W., 2021. Persistence through revolutions. NBER Working Paper 27053.

[3] Angrist, N., Djankov, S., Goldberg, P. K., and Patrinos, H. A., 2021. Measuring human capital using global learning data. *Nature*, 592, 403-408.

[4] Arbatlı C. E., Ashraf, Q. H., Galor, O., and Klemp, M., 2020, Diversity and conflict. *Econometrica*, 88, 727-797.

[5] Ashraf, Q., and Galor, O., 2011. Dynamics and stagnation in the Malthusian epoch. *American Economic Review*, 101, 2003-2041.

[6] Ashraf Q, Galor O., 2013, The "Out of Africa" hypothesis, human genetic diversity, and comparative economic development, *American Economic Review*, 103, 1-46.

[7] Ashraf, Q., Galor, O., Klemp, M., 2021, The ancient origins of the wealth of nations, *The Handbook of Historical Economics*, 675-717.

[8] Bai, Y., Li, Y. and Lam, P. H., 2023. Quantity-quality trade-off in Northeast China during the Qing dynasty. *Journal of Population Economics*, 36, 1657-1694.

[9] Becker, G., Murphy, K., and Tamura, R., 1990. Human capital, fertility, and economic growth. *Journal of Political Economy*, 98, S12-S37.

[10] Becker, S. O., Cinnirella, F., and Woessmann, L., 2020. The trade-off between fertility and education: Evidence from before the demographic transition. *Journal of Economic Growth*, 15, 177-204.

[11] Bisin, A., and Verdier, T., 1998. On the cultural transmission of preferences for social status. *Journal of Public Economics*, 70, 75-97.

[12] Bisin, A., and Verdier, T., 2000. Beyond the melting pot: Cultural transmission, marriage, and the evolution of ethnic and religious traits. *Quarterly Journal of Economics*, 115, 955-988.

[13] Bisin, A., and Verdier, T., 2001. The economics of cultural transmission and the dynamics of preferences. *Journal of Economic Theory*, 97, 298-319.

[14] Bisin, A., and Verdier, T., 2017. On the joint evolution of culture and institutions. *NBER* Working Paper 23375.

[15] Black, S. E., and Devereux, P. J., 2010. Recent developments in intergenerational mobility. In: O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, vol. 4B, ch. 16, 1487-1541, Amsterdam: Elsevier.

[16] Black, S. E., Devereux, P. J., and Salvanes, K. G., 2009. Like father, like son? A note on the intergenerational transmission of IQ scores. *Economics Letters*, 105, 138-140.

[17] Blankenau, W., and Simpson, N., 2004. Public education expenditures and growth. *Journal of Development Economics*, 73, 583-605.

[18] Brunnschweiler, C., Peretto, P., and Valente, S., 2021. Wealth creation, wealth dilution and demography. *Journal of Monetary Economics*, 117, 441-459.

[19] Carillo, M. R., Lombardo, V., and Zazzaro, A., 2019. The rise and fall of family firms in the process of development. *Journal of Economic Growth*, 24, 43-78.

[20] Chu, A., Cozzi, G., and Liao, C., 2013. Endogenous fertility and human capital in a Schumpeterian growth model. *Journal of Population Economics*, 26, 181-202.

[21] Chu, A., Fan, H., and Wang, X., 2020. Status-seeking culture and development of capitalism. *Journal of Economic Behavior and Organization*, 180, 275-290.

[22] Chu, A., Furukawa, Y., and Wang, X., 2022. Rent-seeking government and endogenous takeoff in a Schumpeterian economy. *Journal of Macroeconomics*, 72, 103399.

[23] Chu, A., Furukawa, Y., and Zhu, D., 2016. Growth and parental preference for education in China. *Journal of Macroeconomics*, 49, 192-202.

[24] Chu, A., Kou, Z., and Wang, X., 2020. Effects of patents on the transition from stagnation to growth. *Journal of Population Economics* , 33, 395-411.

[25] Chu, A., Kou, Z., and Wang, X., 2022. Culture and stages of economic development. *Economics Letters*, 210, 110213.

[26] Chu, A., Peretto, P., and Wang, X., 2022. Agricultural revolution and industrialization. *Journal of Development Economics*, 158, 102887.

[27] Chu, A., Peretto, P., and Xu, R., 2023. Export-led takeoff in a Schumpeterian economy. *Journal of International Economics*, 145, 103798.

[28] Connolly, M., and Peretto, P., 2003. Industry and the family: Two engines of growth. *Journal of Economic Growth*, 8, 115-148.

[29] Cozzi, G., 1998. Culture as a bubble. *Journal of Political Economy*, 106, 376-394.

[30] de la Croix, D., and Doepke, M., 2003. Inequality and growth: Why differential fertility matters. *American Economic Review*, 93, 1091-1113.

[31] Fernihough, A., 2017. Human capital and the quantity-quality trade-off during the demographic transition. *Journal of Economic Growth*, 22, 35-65.

[32] Funke, M., and Strulik, H., 2000. On endogenous growth with physical capital, human capital and product variety. *European Economic Review*, 44, 491-515.

[33] Galor, O., 2005. From stagnation to growth: Unified growth theory. *Handbook of Economic Growth*, 1, 171-293.

[34] Galor, O., 2011. *Unified Growth Theory.* Princeton University Press.

[35] Galor, O., 2022. *The Journey of Humanity: The Origins of Wealth and Inequality.* Penguin Random House.

[36] Galor, O., and Klemp, M., 2019. Human genealogy reveals a selective advantage to moderate fecundity. *Nature Ecology & Evolution*, 3, 853-857.

[37] Galor, O., Klemp, M., Wainstock, D. C., 2023. The Impact of the Prehistoric Out of Africa Migration on Cultural Diversity, *National Bureau of Economic Research.*

[38] Galor, O., and Michalopoulos, S., 2012. Evolution and the growth process: Natural selection of entrepreneurial traits. *Journal of Economic Theory*, 147, 759-780.

[39] Galor, O., and Moav, O., 2001. Evolution and growth. *European Economic Review*, 45, 718-729.

[40] Galor, O., and Moav, O., 2002. Natural selection and the origin of economic growth. *Quarterly Journal of Economics*, 117, 1133-1191.

[41] Galor, O., Moav, O., and Vollrath, D., 2009. Inequality in landownership, the emergence of human-capital promoting institutions, and the great divergence. *Review of Economic Studies*, 76, 143-179.

[42] Galor, O., and Mountford, A., 2008. Trading population for productivity: Theory and evidence. *Review of Economic Studies*, 75, 1143-1179.

[43] Galor, O., and Özak, Ö., 2016. The Agricultural Origins of Time Preference. *American Economic Review*, 106, 3064-3103.

[44] Galor, O., and Weil, D., 2000. Population, technology and growth: From the Malthusian regime to the demographic transition. *American Economic Review*, 110, 806-828.

[45] Grossman, G., and Helpman, E., 1991. Quality ladders in the theory of growth. *Review of Economic Studies*, 58, 43-61.

[46] Hansen, G., and Prescott, E., 2002. Malthus to Solow. *American Economic Review*, 92, 1205-1217.

[47] Hofstede, G. H., 1991. *Cultures and organizations: software of the mind*, McGraw-Hill, London.

[48] Iacopetta, M., 2010. Phases of economic development and the transitional dynamics of an innovation-education growth model. *European Economic Review*, 54, 317-330.

[49] Iacopetta, M., and Peretto, P., 2021. Corporate governance and industrialization. *European Economic Review*, 135, 103718.

[50] Jones, C., 2001. Was an industrial revolution inevitable? Economic growth over the very long run. *The B.E. Journal of Macroeconomics*, 1, 1-45.

[51] Jones, G., and Schneider, W., 2006. Intelligence, human capital, and economic growth: A Bayesian averaging of classical estimates (BACE) approach. *Journal of Economic Growth*, 11, 71-93.

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

[53] Klemp, M., and Weisdorf, J., 2019. Fecundity, fertility and the formation of human capital. *Economic Journal*, 129, 925-960.

[54] Laincz, C., and Peretto, P., 2006. Scale effects in endogenous growth theory: An error of aggregation not specification. *Journal of Economic Growth*, 11, 263-288.

[55] Lucas, R., 1988. On the mechanics of economic development. *Journal of Monetary Economics*, 22, 3-42.

[56] Michalopoulos, S., 2012, The Origins of Ethnolinguistic Diversity, *American Economic Review*, 102, 1508–1539.

[57] Peretto, P., 2015. From Smith to Schumpeter: A theory of take-off and convergence to sustained growth. *European Economic Review*, 78, 1-26.

[58] Peretto, P., and Valente, S., 2015. Growth on a finite planet: resources, technology and population in the long run. *Journal of Economic Growth*, 20, 305-331.

[59] Romer, P., 1990. Endogenous technological progress. *Journal of Political Economy*, 98, S71-S102.

[60] Segerstrom, P., Anant, T., and Dinopoulos, E., 1990. A Schumpeterian model of the product life cycle. *American Economic Review*, 80, 1077-91.

[61] Shiue, C. 2017. Human capital and fertility in Chinese clans before modern growth. *Journal of Economic Growth*, 22, 351-396.

[62] Solon, G., 1999. Intergenerational mobility in the labor market. In: O. Ashenfelter & D. Card (ed.), *Handbook of Labor Economics*, vol. 3, ch. 29, 1761-1800, Amsterdam: Elsevier.

[63] Strulik, H., 2024. Higher education and the income-fertility nexus. *Journal of Population Economics*, 37, article 35.

[64] Tabellini, G., 2010. Culture and institutions: Economic development in the regions of Europe. *Journal of the European Economic Association*, 8, 677-716.

Appendix A: Proofs

Proof of Lemma 1. The labor share of household i is $s_t(i) \equiv L_t(i)/L_t$, where

$$L_t(i) = n_{t-1}(i)L_{t-1}(i) = n_{t-1}(i)n_{t-2}(i)L_{t-2}(i) = ... = \prod_{\tau=0}^{t-1} n_\tau(i)L_0(i). \tag{A1}$$

From (4), the fertility choice at time 0 is given by

$$n_0(i) = \frac{\eta}{\sigma(1+\eta+\gamma)}\left[1 + (1-\delta)\frac{h_0(i)}{\phi(i)}\right]. \tag{A2}$$

From (6), the level of human capital at time 1 is given by

$$h_1(i) = \frac{\gamma\phi(i)}{1+\eta+\gamma}\left[1 + (1-\delta)\frac{h_0(i)}{\phi(i)}\right]. \tag{A3}$$

Substituting (A3) into (4) yields the fertility choice at time 1 as

$$n_1(i) = \frac{\eta}{\sigma(1+\eta+\gamma)}\left\{1 + \frac{\gamma(1-\delta)}{1+\eta+\gamma}\left[1 + (1-\delta)\frac{h_0(i)}{\phi(i)}\right]\right\}. \tag{A4}$$

Substituting (A3) into (6) yields the level of human capital at time 2 as

$$h_2(i) = \frac{\gamma\phi(i)}{1+\eta+\gamma}\left\{1 + \frac{\gamma(1-\delta)}{1+\eta+\gamma}\left[1 + (1-\delta)\frac{h_0(i)}{\phi(i)}\right]\right\}. \tag{A5}$$

Substituting (A5) into (4) yields the fertility choice at time 2 as

$$n_2(i) = \frac{\eta}{\sigma(1+\eta+\gamma)}\left\{1 + \frac{\gamma(1-\delta)}{1+\eta+\gamma} + \left[\frac{\gamma(1-\delta)}{1+\eta+\gamma}\right]^2\left[1 + (1-\delta)\frac{h_0(i)}{\phi(i)}\right]\right\}. \tag{A6}$$

Then, we can continue the process to derive the fertility choice at time $t \geq 3$ as

$$n_t(i) = \frac{\eta}{\sigma(1+\eta+\gamma)}\left\{1 + \frac{\gamma(1-\delta)}{1+\eta+\gamma} + ... + \left[\frac{\gamma(1-\delta)}{1+\eta+\gamma}\right]^{t-1} + \left[\frac{\gamma(1-\delta)}{1+\eta+\gamma}\right]^{t}\left[1 + (1-\delta)\frac{h_0(i)}{\phi(i)}\right]\right\}, \tag{A7}$$

which can then be re-expressed using a summation sign as in Lemma 1. ■

Proof of Lemma 2. If (17) holds, then (35) shows that $H_{R,t} > 0$. Now, let's consider the case in which

$$\int_0^1 \left[1 - e_t(i) - \sigma n_t(i)\right] h_t(i)\frac{L_t(i)}{L_t}di < \frac{1}{\theta}. \tag{A8}$$

Recall that the value of assets at the end of time t must equal the amount of saving at time t given by wage income at time t such that

$$N_{t+1}v_t = w_t \int_0^1 \left[1 - e_t(i) - \sigma n_t(i)\right] h_t(i)L_t(i)di. \tag{A9}$$

Substituting (A9) into (A8) yields

$$w_t > \frac{\theta N_{t+1} v_t}{L_t} \geq \frac{\theta N_t v_t}{L_t}, \tag{A10}$$

where the second inequality uses $N_{t+1} \geq N_t$. Equation (A10) implies that $\Delta N_t v_t = w_t H_{R,t}$ in (16) cannot hold unless $H_{R,t} = 0$. ■

Proof of Proposition 1. From Lemma 1, the steady-state population share of household i is given by

$$s^*(i) = \frac{\prod_{t=0}^{\infty} n_t(i)}{\int_0^1 \prod_{t=0}^{\infty} n_t(i) di},$$

where we have used $L_0(i) = L_0$ for all i. Lemma 1 shows that $n_t(i)$ is monotonically decreasing in $\phi(i)$ before reaching the steady state n^* in (28), which then becomes independent of $\phi(i)$. Therefore, it must be the case that

$$s^*(i) < s^*(j) \Leftrightarrow \phi(i) > \phi(j).$$

Given that $\int_0^1 s^*(i)di = 1$, there must exist a threshold for $\phi(i)$ above (below) which $s^*(i) < 1$ ($s^*(i) > 1$). Let's define:

$$\Delta \equiv \int_0^1 \phi(i)s^*(i)di - \overline{\phi} = \int_0^1 \phi(i)s^*(i)di - \int_0^1 \phi(i)di = \int_0^1 \phi(i)[s^*(i) - 1]di.$$

We order the households such that $\phi(i) > \phi(j)$ for any $i < j$. In this case, $s^*(i) < 1$ for $i \in [0, \varepsilon]$ and $s^*(i) > 1$ for $i \in [\varepsilon, 1]$. Therefore, we can re-express Δ as

$$\Delta = \underbrace{\int_0^{\varepsilon} \phi(i)[s^*(i) - 1]di}_{<0} + \underbrace{\int_{\varepsilon}^1 \phi(i)[s^*(i) - 1]di}_{>0}.$$

If $\phi(i) = \phi(j) = \phi(\varepsilon)$ for all $i \in [0, \varepsilon]$ and $j \in [\varepsilon, 1]$, then $\Delta = 0$ because

$$\phi(\varepsilon) \int_0^{\varepsilon} [s^*(i) - 1]di + \phi(\varepsilon) \int_{\varepsilon}^1 [s^*(i) - 1]di = \phi(\varepsilon) \int_0^1 [s^*(i) - 1]di = 0.$$

Otherwise, $\Delta < 0$ because $\phi(i) > \phi(\varepsilon) > \phi(j)$ for any $i \in [0, \varepsilon)$ and $j \in (\varepsilon, 1]$ such that

$$\int_0^{\varepsilon} \phi(i)[s^*(i) - 1]di < \phi(\varepsilon) \int_0^{\varepsilon} [s^*(i) - 1]di < 0,$$

$$\phi(\varepsilon) \int_{\varepsilon}^1 [s^*(i) - 1]di > \int_{\varepsilon}^1 \phi(i)[s^*(i) - 1]di > 0,$$

implying $\Delta < \phi(\varepsilon) \int_0^1 [s^*(i) - 1]di = 0$. Therefore, (41) and (42) hold. ■

Appendix B: Figures and Tables

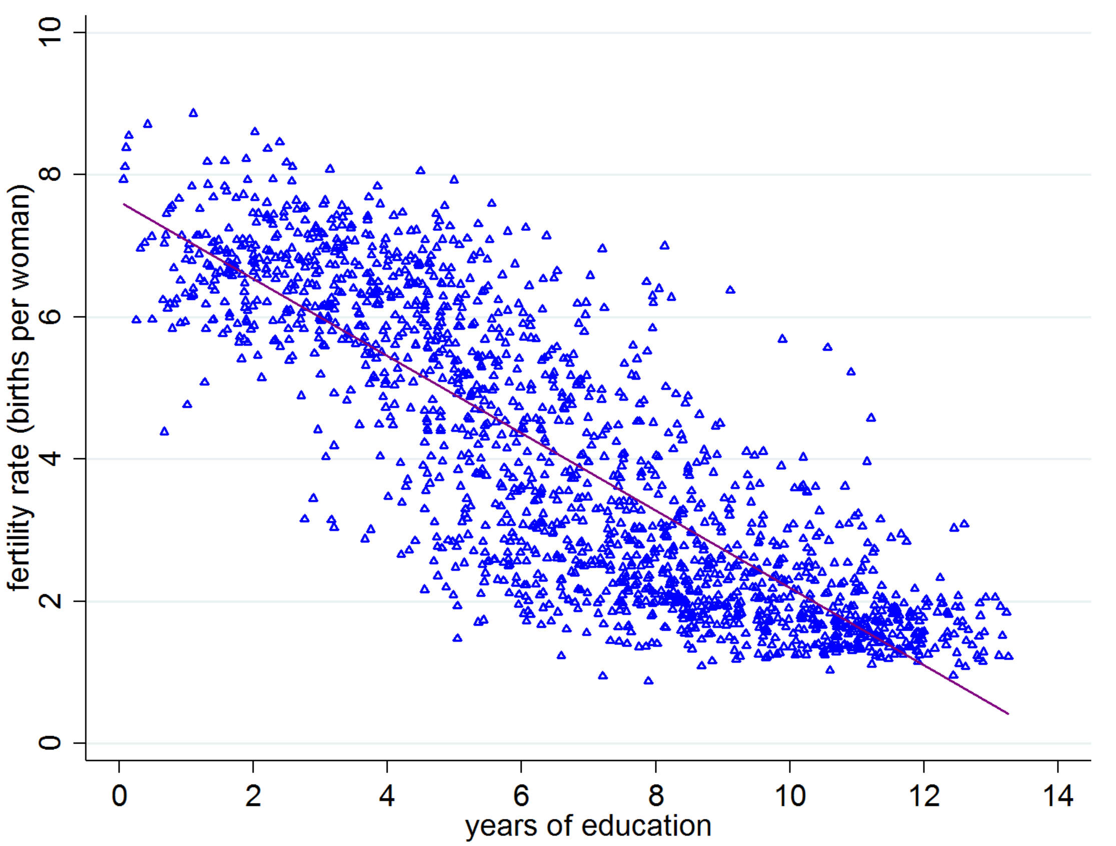

Figure 1: The relationship between fertility and education

Notes: This figure depicts the negative correlation between fertility and education. The vertical axis represents the fertility rate, whereas the horizontal axis denotes the number of years of education.

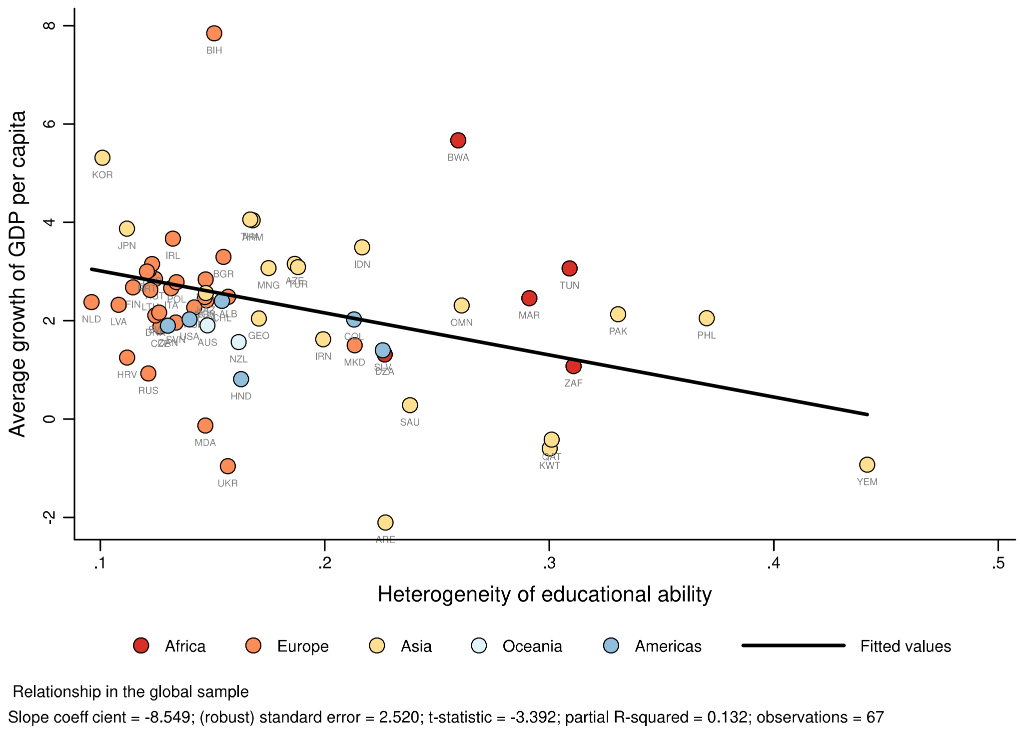

Figure 2: The relationship between heterogeneity of educational ability and economic growth

Notes: This figure illustrates a negative relationship between heterogeneity in educational ability against the country's average economic growth rate from 1951 to 2017.

Table 4: Summary statistics

Variable	Obs	Mean	S.D.	Min	Max
Growth rate of GDP per capita (%)	3,476	2.481	5.968	-43.743	94.138
Heterogeneity of educational ability (Total)	3,476	0.177	0.070	0.096	0.442
Heterogeneity of educational ability (Mathematics)	3,476	0.174	0.057	0.100	0.397
Heterogeneity of educational ability (Science)	3,476	0.192	0.091	0.102	0.541
Heterogeneity of educational years	627	0.667	0.511	0.166	4.902
Time preference (%)	3,476	39.676	27.739	13.000	100
Absolute latitude	3,476	36.374	15.163	1.300	64.000
Distance to nearest waterway (km)	3,476	264.034	425.176	14.176	2385.580
Mean elevation (km)	3,476	141.661	226.068	0.003	1096.503
Mean land suitability	3,413	0.436	0.285	0.003	0.954
Standard deviation of elevation	3,476	106.871	171.805	0.000	965.317
Standard deviation of land suitability	3,356	0.185	0.099	0.001	0.387
Island nation dummy	3,476	0.161	0.367	0	1
Population diversity (ancestry adjusted)	3,065	0.723	0.022	0.671	0.746
Migratory distance from East Africa (in 1,000 km)	3,065	8.520	5.915	3.571	25.898
Share of schooling (%)	627	84.422	21.136	5.009	100
Average education	627	1.944	0.550	-0.934	2.586
Number of researchers (logarithm)	893	7.414	1.078	3.108	8.978
Patent application (logarithm)	1,794	6.293	2.532	0	12.859
Legor_uk	3,476	0.281	0.450	0	1
Legor_fr	3,476	0.347	0.476	0	1
Legor_so	3,476	0.184	0.388	0	1
Legor_ge	3,476	0.076	0.265	0	1
Legor_sc	3,476	0.093	0.290	0	1
Ethnic fractionalization	3,037	0.321	0.233	0.002	0.752
Ethnolinguistic polarization	3,065	0.441	0.264	0.006	0.958
Linguistic fractionalization	2,998	0.303	0.249	0.002	0.865
L1.log GDP per capita	3,409	9.491	1.000	6.645	12.349
L1.log Population	3,409	2.344	1.543	-2.125	5.775

Note: The heterogeneity of educational ability is calculated using the coefficient of variation of academic scores from the TIMSS database. The heterogeneity of educational years, share of schooling, and average education data are sourced from the Barro-Lee educational attainment dataset. The data on time preference is sourced from Galor and Ozak (2016). The land agricultural suitability data is sourced from Michalopoulos (2012). Data for absolute latitude, mean elevation, standard deviation in elevation, and an island dummy variable are obtained from the Geographically Based Economic Data (G-ECON) project (Nordhaus, 2006). Data for distance to the nearest waterway, legal origins, ethnic fractionalization, ethnolinguistic polarization, and linguistic fractionalization are from Arbath et al. (2020). All other variables are from the Penn World Table.

Table 5: Robustness to Alternative Explanatory Variables

	Economic Growth		
	(1)	(2)	(3)
Heterogeneity of educational ability (Mathematics)	-12.753***		
	(4.123)		
Heterogeneity of educational ability (Science)		-5.961**	
		(2.467)	
Heterogeneity of educational years			-1.606*
			(0.883)
Controls	Y	Y	Y
Year FE	Y	Y	Y
Continent FE	Y	Y	Y
R-square	0.137	0.134	0.136
Observations	3356	3356	603

Note: $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses are clustered by country. The dependent variable is the growth rate of GDP per capita. Controls include time preference, distance to the nearest waterway, absolute latitude, mean elevation, mean agricultural land suitability, standard deviation in elevation and agricultural land suitability, and an island dummy.

Table 6: Robustness with Additional Control Variables

	Economic Growth			
	(1)	(2)	(3)	(4)
Heterogeneity of educational ability	-8.146***	-8.181**	-17.057***	-6.230***
	(2.993)	(3.922)	(3.084)	(2.331)
Time preference	0.008	0.009	0.001	0.008
	(0.010)	(0.009)	(0.007)	(0.006)
Distance to nearest waterway	0.238	-0.020	-0.636	0.232
	(0.436)	(0.622)	(0.620)	(0.301)
Absolute latitude	-0.026	-0.022	0.002	-0.012
	(0.018)	(0.018)	(0.015)	(0.011)
Mean elevation	0.004**	0.003	-0.000	0.002*
	(0.002)	(0.002)	(0.002)	(0.001)
Mean land suitability	1.154*	0.461	-0.743	0.607*
	(0.682)	(0.633)	(0.572)	(0.346)
Standard deviation of elevation	-0.005**	-0.003	0.001	-0.003**
	(0.002)	(0.002)	(0.002)	(0.001)
Standard deviation of land suitability	0.446	1.195	0.789	0.407
	(1.564)	(2.033)	(1.825)	(1.011)
Island nation dummy	0.070	0.224	0.653	-0.023
	(0.560)	(0.611)	(0.567)	(0.392)
Ethnic fractionalization		-0.357		
		(1.044)		
Ethnolinguistic polarization		-0.621		
		(0.640)		
Linguistic fractionalization		-0.039		
		(0.850)		
L1.log GDP per capita			-1.251***	
			(0.253)	
L1.log Population			-0.097	
			(0.132)	
L1.Growth rate of GDP per capita				0.368***
				(0.046)
Legal Origin Dummies	Y	N	N	N
Controls	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Continent FE	Y	Y	Y	Y
R-square	0.137	0.126	0.146	0.249
Observations	3356	2970	3291	3243

Note: $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses are clustered by country. The dependent variable is the growth rate of GDP per capita. Legal origin dummies consist of a set of variables that identify the legal origin of a country's Company Law or Commercial Code. The five possible legal origins are: (i) English Common Law, (ii) French Commercial Code, (iii) German Commercial Code, (iv) Scandinavian Commercial Code, and (v) Socialist or Communist Laws.

Table 7: Robustness to Sub-Sampling

	Economic Growth				
	(1)	(2)	(3)	(4)	(5)
	Africa	Europe	Asia	Oceania	Americas
Heterogeneity of educational ability	-11.666***	-4.029	-5.814*	-24.504***	-6.991***
	(0.641)	(8.476)	(2.825)	(0.000)	(0.509)
Controls	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y
R-square	0.315	0.296	0.135	0.585	0.372
Observations	305	1410	1106	134	401

Note: $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses are clustered by country. The dependent variable is the growth rate of GDP per capita. Controls include time preference, distance to the nearest waterway, absolute latitude, mean elevation, mean agricultural land suitability, standard deviation in elevation and agricultural land suitability, and an island dummy.