Natural Selection and Innovation-Driven Growth

Chu, Angus and Cozzi, Guido and Fan, Haichao

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

June 2022

Natural Selection and Innovation-Driven Growth

Angus C. Chu Guido Cozzi Haichao Fan

January 2024

Abstract

How does the interplay between natural selection, household education choices, and R&D activities shape our macroeconomic trajectory? Delving deep into this question, we present a novel innovation-driven growth model that intricately connects household heterogeneity in education ability with fertility and R&D-driven technological progress. Our findings unravel a captivating paradox: while households with lower education abilities might amass less human capital and choose to have more offspring, they gain a fleeting evolutionary advantage. This advantage, however, exacts a significant toll, stifling R&D and curtailing long-term economic growth. Our model not only theoretically reveals this complex dynamic but validates it with cross-country data and an instrumental variable, suggesting that education disparities can hamper R&D output, education outcomes, and economic expansion in the long run. This research unveils crucial insights into the nuanced relationships between natural selection, household education choices, and R&D.

JEL classification: O30, O40

Keywords: natural selection, innovation, economic development

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, School of Economics, 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 a novel growth model with endogenous fertility 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, in an early stage of development, households that have a lower education ability accumulate less human capital but choose to have more children and enjoy an evolutionary advantage. In a later stage of development, households with a higher education ability choose to increase their number of children as their human capital rises over time because their higher level of human capital compensates for their lower fertility. 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.

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-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 than the case without natural selection because the more educated parents have fewer children. This 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

[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.

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.

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 indeed has adverse effects on education, innovation and economic growth in the long run. This result remains robust when we use heterogeneity in ability test scores as an instrumental variable for heterogeneity in education.

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 in early stages 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 substantial 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 into 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. An early 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;[4] 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 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 demo-

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

graphic transition and fosters technological progress,[5] but the selective advantage is reversed to favor the quantity type after the demographic transition.[6] Although the present study does not explore the interesting demographic transition in the pre-industrial era captured elegantly by unified growth theory, it contributes to 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 in the modern era.

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.[7] 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 and human capital accumulation.[8] Each individual lives for 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.

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[9]

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

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

[7] 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.

[8] 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.

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

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

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[10]

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

where the parameter $\delta \in (0,1)$ is the depreciation rate of human capital that a generation passes onto the next.[11] As for the ability parameter $\phi(i) > 0$ of household i,[12] it is heterogeneous across households and follows a general distribution with the following mean:[13]

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

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 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.[14]

The individual devotes her entire wage income to saving at time t and consumes the return at time $t+1$:[15]

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

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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 $e_t(i)w_th_t(i)$.

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

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[16]

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

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.

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

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

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.

[16] 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.

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.[17] 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.

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

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

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

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 $H_{Y,t} + H_{R,t} = \int_0^1 \left[1 - e_t(i) - \sigma n_t(i)\right] h_t(i) L_t(i) di$;
- total saving equals asset value such that $w_t \int_0^1 \left[1 - e_t(i) - \sigma n_t(i)\right] h_t(i) L_t(i) di = N_{t+1} v_t$.

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.[18] 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

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

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.

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

3.2 Stage 2: Innovation and human capital accumulation

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

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

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

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

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

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:

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

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.

We now 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[19]

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

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

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 equilibrium growth rate under homogeneous households (i.e., $\phi(i) = \overline{\phi}$ for all $i \in [0,1]$). 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 level 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.*

Proof. See Appendix A. ■

4.1 An example

In this section, we provide a simple parametric example to illustrate our results more clearly. 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 $\phi(i)$.

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

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

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

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

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

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.

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 1 *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

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.[20] 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 from 1955 to 2015, Figure 1 in Appendix B presents a well known negative relationship between fertility and education. 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

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

$h_t(i)$ and lower fertility $n_t(i)$, generating a negative relationship between these two endogenous variables.[21]

In the previous section, we show that the negative effects of this natural selection can be captured by the heterogeneity in the ability to accumulate human capital, which in turn reduces the average level of education, innovation and economic growth in the long run. In this section, we use cross-country data to test this theoretical result. Specifically, we use the coefficient of variation in the level of education as a scale-invariant measure of heterogeneity in ability and estimate its effects on education, innovation and economic growth in the long run. Our regression equation is specified as follows:

$$y_{i,t+m} = \beta_0 + \beta_1 var_{i,t} + \beta_2 h_{i,t} + Z_{i,t} + \varphi_t + \epsilon_{i,t},$$

where $y_{i,t+m}$ is the dependent variable (i.e., education, innovation or economic growth) in country i at time $t+m$, $var_{i,t}$ is the variation in education and $h_{i,t}$ is the level of human capital in country i at time t. $Z_{i,t}$ is a vector of control variables including the log of population, the log of GDP per capita, trade as a share of GDP, gross capital formation as a share of GDP, and government expenditure as a share of GDP. In order to capture the long-run effect of heterogeneity in education, all explanatory variables are lagged 25 years (i.e., $m = 25$). φ_t denotes the year fixed effects, and $\epsilon_{i,t}$ is the error term. Our theory predicts that $\beta_1 < 0$ and $\beta_2 > 0$. In other words, upon controlling for the level of human capital, heterogeneity in education (reflecting heterogeneity in the ability to accumulate human capital) has a negative effect on education, innovation and economic growth in the next period.[22] Except for the variation in education, the human-capital index, the number of researchers in R&D and the number of patent applications, all other variables are from the Penn World Table. The human-capital index, the number of researchers in R&D and patent applications are from the World Bank. The variation in education is calculated from the Barro-Lee educational attainment dataset.[23] We provide the summary statistics in Appendix B.

Table 1 reports our main empirical results with control variables, but our results are also robust to excluding other control variables.[24] In the first two columns, the dependent variables are the share of the population with at least some primary education and the log of average years of education, respectively. These two variables capture the average level of education. In columns (3) and (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. These two variables capture the rate of innovation. Finally, in the last column, the dependent variable is the growth rate of GDP per capita, which captures economic growth. From Table 1, we see

[21] 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.

[22] For example, we use the annual growth rate of GDP per capita to capture economic growth. Then, in order to capture the long-run effects of heterogeneity in education on economic growth, we examine the impact of the variation of education on the annual growth rate of GDP per capita 25 years later.

[23] The Barro-Lee educational attainment dataset provides the fraction of each group completely or incompletely having attained primary, secondary and higher education. The duration for primary education and secondary education in each country is available from the UNESCO Statistical Yearbook. As in Barro and Lee (2013), we use a duration of four years for higher education and assign two years to persons who entered tertiary school but did not complete it. We compute the average years of education for each group and calculate their standard deviation in each country.

[24] See Table 4 in Appendix B, which shows that our findings are not driven by the inclusion of control variables.

that the coefficients of $var_{i,t}$ are all significantly negative, whereas the coefficients of $h_{i,t}$ are mostly significantly positive. This finding implies that upon controlling for the level of human capital, heterogeneity in education harms education, innovation and economic growth in the long run, as predicted by our theoretical model. Additionally, in order to address potential outliers, we depict a bin-scatter plot illustrating the residualized dependent variables plotted against the residualized heterogeneity in education, as outlined in Table 1. Figures 2 to 4 in Appendix B provide evidence suggesting that the observed negative relationship between any pair of the two variables is not likely to be influenced by outliers.

Table 1: Effects of heterogeneity in education

	Education		Innovation		Growth
	(1)	(2)	(3)	(4)	(5)
Heterogeneity in education	-20.187***	-0.381***	-0.679**	-0.847***	-0.992***
	(1.875)	(0.039)	(0.284)	(0.158)	(0.360)
Human capital	3.952	0.174***	0.809***	1.884***	1.311**
	(2.469)	(0.053)	(0.255)	(0.268)	(0.518)
log population	0.453	0.016**	0.227***	1.336***	0.278***
	(0.365)	(0.008)	(0.058)	(0.077)	(0.087)
log GDP per capita	2.843***	0.147***	1.047***	1.111***	-0.702***
	(0.849)	(0.022)	(0.133)	(0.153)	(0.191)
Trade share to GDP	1.566	0.005	-3.121***	-0.100	-0.251
	(1.812)	(0.055)	(0.752)	(0.400)	(0.457)
Capital formation share	8.206	0.267	0.934	1.977	0.793
	(5.570)	(0.167)	(1.117)	(1.265)	(1.602)
Government expenditure share	3.950	0.320*	1.948*	1.903*	0.275
	(5.869)	(0.175)	(1.065)	(1.143)	(1.697)
Year fixed effect	Yes	Yes	Yes	Yes	Yes
R-squared	0.826	0.827	0.747	0.825	0.090
Observations	954	954	244	624	954

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses are clustered by country. The dependent variables correspond to the share of population with schooling and log of average years of education respectively in the first two columns. The dependent variables correspond to the log of the number of researchers in R&D (per million people) and log of the number of patent applications respectively in the columns 3-4. The dependent variable is the growth rate of GDP per capita in the last column. In all columns, we control year fixed effects. All independent variables are lagged 25 years.

Table 1 uses the variation in education to represent differences in educational ability. However, it's important to note that unlike education, educational ability cannot be directly observed. Therefore, bias emerges when we employ heterogeneity in education as a proxy for heterogeneity in educational ability. In order to address this issue, we employ the global standardized tests of students' academic ability as an instrumental variable. These tests capture both individual intelligence and educational ability. There are two programs on tests of students' academic performance: the Trends in International Mathematics and Science Study (TIMSS) and the Progress in International Reading Literacy Study (PIRLS).[25] Angrist *et al.*

[25]The two programs evaluate students in both fourth and ninth grades. We focus on their tests of fourth-grade students. Assessing fourth-grade students covers a larger number of students and offers a more accurate

Table 2: Impacts of heterogeneity in education with an instrumental variable

	Education		Innovation		Growth
	(1)	(2)	(3)	(4)	(5)
Panel A: Total					
Heterogeneity in education	-25.740***	-0.470***	-2.103**	-1.207*	-1.888**
	(3.287)	(0.082)	(0.991)	(0.686)	(0.873)
Human capital	2.437	0.114	0.355	1.848***	1.221
	(4.116)	(0.074)	(0.343)	(0.292)	(0.772)
R-squared	0.746	0.814	0.715	0.840	0.069
Observations	567	567	180	429	567
Panel B: Read					
Heterogeneity in education	-24.663***	-0.452***	-1.859**	-1.208*	-1.432**
	(2.527)	(0.070)	(0.787)	(0.659)	(0.726)
Human capital	2.503	0.115	0.429	1.848***	1.249*
	(3.845)	(0.071)	(0.283)	(0.291)	(0.679)
R-squared	0.764	0.819	0.732	0.840	0.090
Observations	567	567	180	429	567
Panel C: Math					
Heterogeneity in education	-24.063***	-0.430***	-3.145*	-1.294**	-1.864**
	(2.635)	(0.065)	(1.701)	(0.612)	(0.837)
Human capital	2.540	0.116*	0.041	1.844***	1.223
	(3.744)	(0.068)	(0.529)	(0.303)	(0.765)
R-squared	0.773	0.824	0.593	0.839	0.070
Observations	567	567	180	429	567
Panel D: Science					
Heterogeneity in education	-26.624***	-0.504***	-1.756**	-1.111	-1.960**
	(3.885)	(0.098)	(0.885)	(0.717)	(0.934)
Human capital	2.383	0.111	0.460	1.852***	1.217
	(4.326)	(0.081)	(0.321)	(0.287)	(0.789)
R-squared	0.729	0.801	0.738	0.841	0.064
Observations	567	567	180	429	567
Country-level controls	Yes	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes	Yes

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses are clustered by country. The dependent variables correspond to the share of population with schooling and log of average years of education respectively in the first two columns. The dependent variables correspond to the log of the number of researchers in R&D (per million people) and log of the number of patent applications respectively in the columns 3-4. The dependent variable is the growth rate of GDP per capita in the last column. In all columns, we control year fixed effects.

(2021) demonstrate that the overall student performance across countries undergoes minimal changes over time, yet substantial differences exist at the national level. Similarly, the variations in educational ability exhibit small changes over time and considerable differences across countries.[26] Table 2 presents the empirical findings resulting from instrumenting the varia-

representation of nationwide differences compared to ninth-grade students. TIMSS conducted math and science tests every four years from 1995 to 2019, and PIRLS conducted tests in 2001, 2006, 2011, 2016 and 2021.

[26]The variation in science ability exhibited a modest change of only 1.79% from 2003 to 2019. Notably, at the 90th percentile, the variation reaches 542.14, in stark contrast to 302.93 at the 10th percentile. A consistent

tion in education with the corresponding variation in educational ability. The conducted tests are categorized into Reading, Math, and Science. From the obtained scores, we compute the variations in educational ability concerning Reading, Mathematics, and Science, as well as the overall ability (defined as the cumulative ability in Reading, Mathematics, and Science). Panel A uses the variation in total ability, Panel B uses the variation in reading ability, Panel C employs the variation in mathematical ability, and Panel D employs the variation in scientific ability as the instrumental variable. As seen in Table 2, the coefficients of educational variation are significantly negative across all columns, indicating a negative impact on innovation and economic growth. Table 5 in Appendix B presents the first stage of regression with instrumental variables. As shown in the table, we observe a significant positive correlation between the variation in ability and the variation in education.

We also calculate the coefficients of variation in education for the male and female populations, respectively. As shown in Table 6 in Appendix B, all main results still hold for both samples.[27] The coefficients of variation in education are all significantly negative (except for the impact of heterogeneity in education on the number of researchers in R&D). Comparing panel A and B, the negative impact of heterogeneity in the ability to accumulate human capital is more significant for the male population with a larger magnitude of coefficients. This result is possibly due to the advantages enjoyed by men in the economy (i.e., the employment gender gap, women's discrimination in senior positions, and their lower labor force participation).

Finally, to account for differences in countries' education systems, as depicted in Table 7, we incorporate the proportion of public education expenditure and female enrollment rate into the regression. This allows us to isolate a portion of the impact attributed to differences in national education systems. Even after further controlling for these factors, our predictions remain robust.

6 Conclusion

In this study, we have constructed an innovation-driven growth model with endogenous takeoff, 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 an unanticipated 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 long-term economic growth.

pattern is observed in both math and reading abilities. In math, the 90th percentile exhibits a variation of 454.05, while the 10th percentile shows 308.75. In reading, the 90th percentile has a variation of 481.33, significantly differing from 302.56 at the 10th percentile. Considering that both TIMSS and PIRLS were conducted in 2011, we calculated the standard deviation of scores for each country in that year to serve as an instrumental variable. Employing scores from other years for calculating ability variation would not impact our results.

[27] In Table 6 and 7, we present the second-stage regression results from instrumenting the heterogeneity in education using the corresponding variation in ability test scores. Here, we use the variation in overall ability (defined as the cumulative ability in Reading, Mathematics, and Science).

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. A caveat is that heterogeneity in education in the data may be also driven by heterogeneity in preferences and other considerations. To partly mitigate this problem, we adopt variation in ability test scores as an instrumental variable, 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 in innovation-driven growth models, 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 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 and endogenous takeoff but also raises thought-provoking 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. Persistance 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] Ashraf, Q., and Galor, O., 2011. Dynamics and stagnation in the Malthusian epoch. *American Economic Review*, 101, 2003-2041.

[5] 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.

[6] 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.

[7] Bisin, A., and Verdier, T., 1998. On the cultural transmission of preferences for social status. *Journal of Public Economics*, 70(1), 75–97.

[8] 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(3), 955–988.

[9] Bisin, A., and Verdier, T., 2001. The economics of cultural transmission and the dynamics of preferences. *Journal of Economic Theory*, 97(2), 298–319.

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

[11] 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.

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

[13] 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.

[14] 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.

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

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

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

[18] 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.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

[34] 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.

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

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

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

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

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

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

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

[42] 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.

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

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

[45] 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.

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

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

[48] 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.

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

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

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

[52] 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

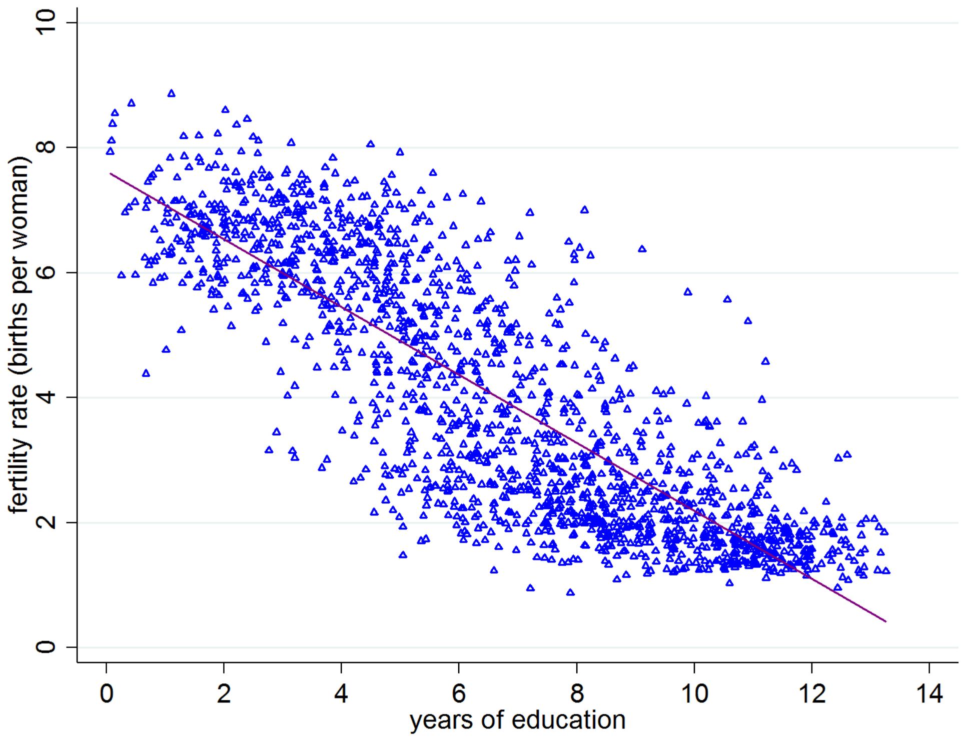

Figure 1: The relationship between fertility and education

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

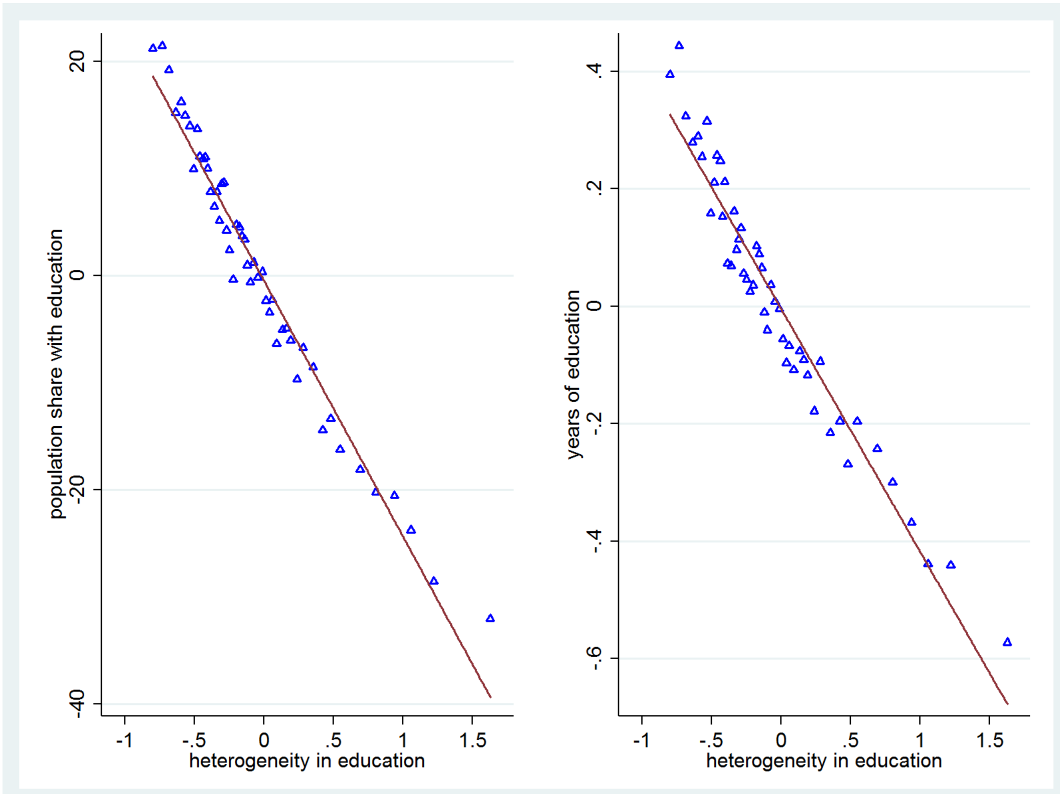

Figure 2: The relationship between heterogeneity in education and the average level of education

Notes: This figure illustrates a negative association between the heterogeneity in education and the average education level. In Panel A, the vertical axis represents the percentage of the population with at least some primary education, whereas in Panel B, the vertical axis corresponds to the log of the average years of education.

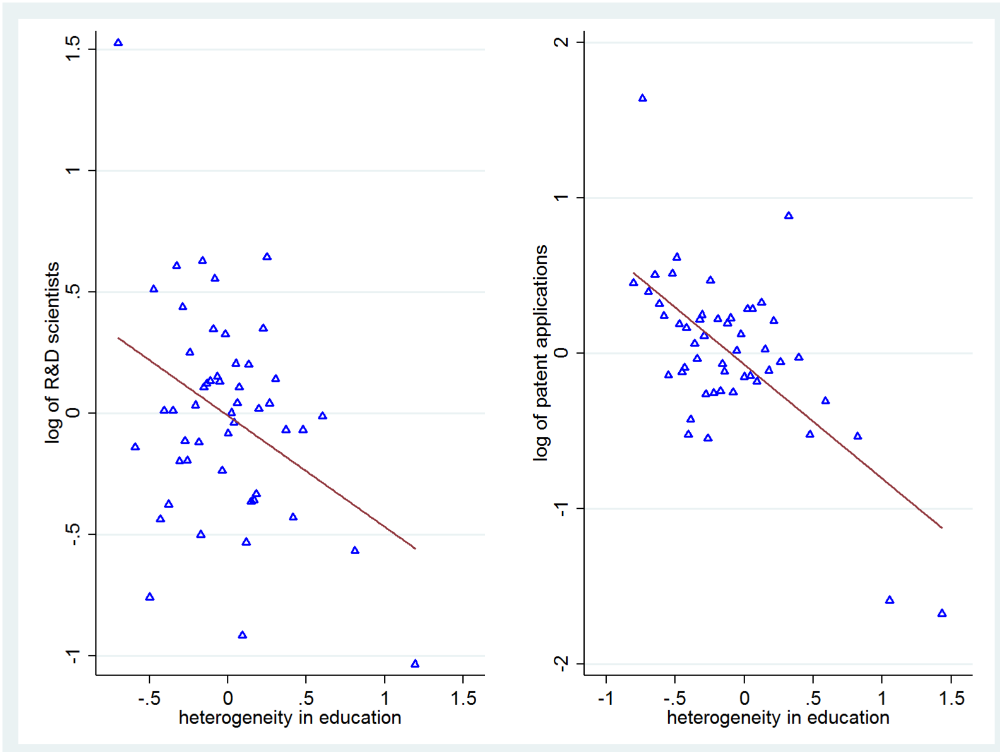

Figure 3: The relationship between heterogeneity in education and the rate of innovation

Notes: This figure illustrates a negative relationship between heterogeneity in education and the rate of innovation. In Panel A, the vertical axis represents the log of the number of researchers in R&D (per million people), whereas in Panel B, the vertical axis corresponds to the log of the number of patent applications.

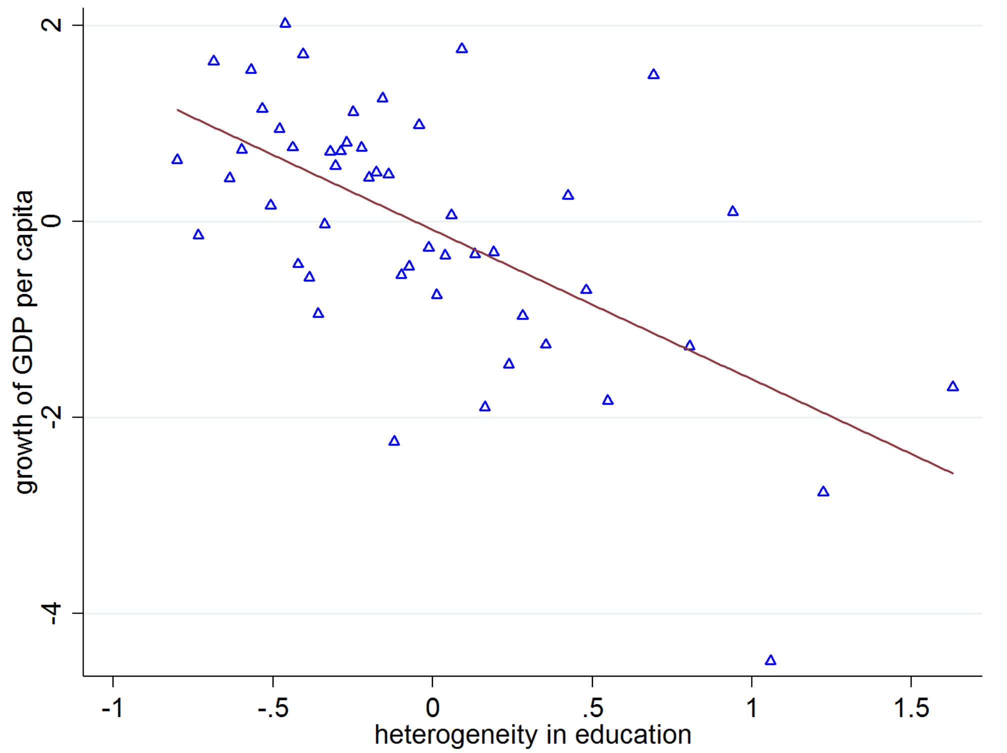

Figure 4: The relationship between heterogeneity in education and the growth rate of GDP per capita

Notes: This figure displays a negative relationship between heterogeneity in education and the growth rate of GDP per capita.

Table 3: Summary statistics

Variable	Obs	Mean	S.D.	Min	Max
Share of population with schooling (%)	954	81.15	20.86	13.72	100
Years of education (logarithm)	954	1.895	0.509	-0.191	2.586
Number of researchers (logarithm)	244	6.728	1.665	2.003	8.952
Number of patent applications (logarithm)	624	5.313	2.829	0	13.78
Growth of GDP per capita (%)	954	1.949	4.889	-50.23	35.26
Coefficient of variation in education	954	1.140	0.822	0.220	8.075
Variation in education (male)	954	0.991	0.638	0.228	5.984
Variation in education (female)	954	1.432	1.409	0.209	17.71
Variation in educational ability (total)	567	1201	230.9	873.5	1669
Variation in educational ability (read)	567	418.9	71.72	315.0	578.2
Variation in educational ability (math)	567	398.5	64.68	291.5	521.8
Variation in educational ability (science)	567	436.6	110.3	292.7	636.2
Human capital	954	1.808	0.613	1.009	3.463
Log of population	954	1.883	1.686	-2.212	7.067
Log of GDP per capita	954	8.587	1.175	5.683	12.38
Trade share to GDP	954	-0.050	0.332	-8.188	0.860
Capital formation share	954	0.211	0.133	0.002	2.000
Government expenditure share	954	0.178	0.106	0.012	1.122
Education expenditure share	856	0.042	0.023	0.007	0.417
Gender disparities	954	1.134	0.223	0.145	1.745

Note: The coefficient of variation in education is calculated from the Barro-Lee educational attainment dataset. The variation in educational ability is calculated from the TIMSS and PIRLS for the Fourth Grade Combined International Database. The human-capital index, the number of researchers in R&D and the number of patent applications are from the World Bank. All other variables are from the Penn World Table.

Table 4: Effects of heterogeneity in education (without controls)

	Education		Innovation		Growth
	(1)	(2)	(3)	(4)	(5)
Heterogeneity in education	-22.744***	-0.504***	-2.727***	-1.451***	-0.503
	(1.858)	(0.047)	(0.290)	(0.337)	(0.375)
Year fixed effect	Yes	Yes	Yes	Yes	Yes
R-squared	0.798	0.721	0.467	0.119	0.049
Observations	954	954	244	624	954

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses are clustered by country. The dependent variables correspond to the share of population with schooling and log of average years of education, respectively, in the first two columns. The dependent variables correspond to the log of the number of researchers in R&D (per million people) and log of the number of patent applications, respectively, in columns 3 and 4. The dependent variable is the growth rate of GDP per capita in the last column. In all columns, we control year fixed effects.

Table 5: Impacts of heterogeneity in education with an instrumental variable (First stage)

	Education		Innovation		Growth
	(1)	(2)	(3)	(4)	(5)
Panel A: Total					
Variation in scores	0.001***	0.001***	0.001***	0.001**	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R-squared	0.412	0.412	0.592	0.372	0.412
Observations	567	567	180	429	567
Panel B: Read					
Variation in scores	0.005***	0.005***	0.002***	0.003**	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
R-squared	0.452	0.452	0.617	0.387	0.452
Observations	567	567	180	429	567
Panel C: Math					
Variation in scores	0.005***	0.005***	0.001*	0.003**	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
R-squared	0.415	0.415	0.558	0.373	0.415
Observations	567	567	180	429	567
Panel D: Science					
Variation in scores	0.003***	0.003***	0.002***	0.002**	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
R-squared	0.400	0.400	0.613	0.369	0.400
Observations	567	567	180	429	567
Country-level controls	Yes	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes	Yes

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses are clustered by country. In the first stage, the dependent variable is heterogeneity in education in all columns. In the second stage, the dependent variables correspond to the share of population with schooling and log of average years of education respectively in the first two columns. In the second stage, the dependent variables correspond to the log of the number of researchers in R&D (per million people) and log of the number of patent applications respectively in the columns 3-4. In the second stage, the dependent variable is the growth rate of GDP per capita in the last column. In all columns, we control year fixed effects.

Table 6: Impacts of heterogeneity in education with an instrumental variable (male vs female)

	Education		Innovation		Growth
	(1)	(2)	(3)	(4)	(5)
Panel A: Male					
Heterogeneity in education	-28.871***	-0.474***	-2.672**	-1.563	-2.343**
	(4.773)	(0.101)	(1.166)	(0.963)	(1.141)
Human capital	-0.086	0.116**	0.314	1.868***	1.331*
	(3.396)	(0.056)	(0.324)	(0.289)	(0.793)
R-squared	0.662	0.777	0.735	0.839	0.065
Observations	567	567	180	429	567
Panel B: Female					
Heterogeneity in education	-17.917***	-0.381***	-1.525*	-0.695	-1.235**
	(2.574)	(0.056)	(0.843)	(0.487)	(0.555)
Human capital	10.853	0.204	0.388	1.982***	1.258*
	(6.658)	(0.134)	(0.387)	(0.360)	(0.741)
R-squared	0.622	0.774	0.675	0.837	0.047
Observations	567	567	180	429	567
Country-level controls	Yes	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes	Yes

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses are clustered by country. In the first-stage regression (available upon request), we use the variation in overall test scores as the IV for heterogeneity in education. Here, we only report the second-stage regression results. The dependent variables correspond to the share of population with schooling and log of average years of education respectively in the first two columns. The dependent variables correspond to the log of the number of researchers in R&D (per million people) and log of the number of patent applications respectively in the columns 3-4. The dependent variable is the growth rate of GDP per capita in the last column. In all columns, we control year fixed effects.

Table 7: Impacts of heterogeneity in education with an instrumental variable (More controls)

	Education		Innovation		Growth
	(1)	(2)	(3)	(4)	(5)
Heterogeneity in education	-25.285***	-0.462***	-1.956**	-1.098*	-1.355*
	(2.635)	(0.065)	(0.792)	(0.591)	(0.707)
Human capital	2.815	0.148*	0.447	1.750***	1.333**
	(4.371)	(0.079)	(0.296)	(0.331)	(0.674)
Log of population	0.043	0.011	0.144**	1.288***	0.157*
	(0.376)	(0.008)	(0.070)	(0.078)	(0.092)
Log of GDP per capita	0.422	0.123***	0.587**	0.800***	-0.978***
	(0.994)	(0.024)	(0.272)	(0.255)	(0.292)
Trade share to GDP	4.449	0.065	-1.751*	0.273	0.701
	(3.038)	(0.082)	(1.011)	(0.426)	(0.636)
Capital formation share	5.835	0.262	2.710**	2.263	3.482*
	(6.929)	(0.192)	(1.243)	(1.654)	(2.021)
Government expenditure share	15.966	0.685**	3.958***	2.139	2.562
	(11.679)	(0.321)	(1.344)	(1.339)	(2.432)
Education expenditure share	-0.419	-0.015	-0.071	0.021	-0.408***
	(0.838)	(0.019)	(0.104)	(0.086)	(0.137)
Gender disparities	0.940	0.154**	-0.026	-0.555	-0.548
	(2.251)	(0.072)	(0.488)	(0.519)	(0.708)
R-squared	0.755	0.822	0.729	0.843	0.103
Observations	567	567	180	429	567

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses are clustered by country. In the first-stage regression (available upon request), we use the variation in overall test scores as the IV for heterogeneity in education. Here, we only report the second-stage regression results. The dependent variables correspond to the share of population with schooling and logarithm of average years of education respectively in the first two columns. The dependent variables correspond to logarithm of the number of researchers in R&D (per million people) and logarithm of the number of patent applications respectively in the columns 3-4. The dependent variable is the growth rate of GDP per capita in the last column. In all columns, we control year fixed effects.