Human Capital Accumulation Across Space*

Klaus Desmet
SMU

Dávid Krisztián Nagy
CREI

Esteban Rossi-Hansberg
University of Chicago

September 10, 2025

Abstract

This paper studies how human capital shapes the economic geography of development. We develop a model in which the cost of acquiring human capital varies across space, and regions with higher human capital innovate more. Locations are spatially connected through migration and trade. There are localized agglomeration economies, and human-capital-augmenting technology diffuses across space. Using high-resolution data on income and schooling, we quantify and simulate the model at the $1^\circ \times 1^\circ$ resolution for the entire globe. Over the span of two centuries, the model predicts strong persistence in the spatial distribution of development — unlike spatial dynamic models without human capital, which predict convergence. Proportionally lowering the cost of education in sub-Saharan Africa or Central and South Asia raises local outcomes but reduces global welfare, whereas the same policy in Latin America improves global outcomes. An alternative policy equalizing educational costs across sub-Saharan Africa generates long-run local losses, as population reallocates within the region toward less productive areas. Central to these results is the estimated negative correlation between the education costs and local fundamentals, as well as inefficiencies in the spatial allocation due to externalities.

Keywords: human capital, geography, development, education policy, global inequality, dynamic spatial models
JEL Codes: E24, F10, I24, J24, O11, O18, O33; R12, R23

1 Introduction

Human capital is unevenly distributed across space. In 2000, people in the Netherlands had an average of 10.8 years of schooling, compared to 2.5 years in the Central African Republic. This range reflects the difference between countries at the 90^{th} and the 10^{th} percentiles. Looking at the same percentiles, but at the $1^\circ \times 1^\circ$ grid cell level, the range goes from 11.8 to 3.4 years.[1]

Where do these large differences come from? On the one hand, the cost of acquiring human capital varies widely across space. In countries or regions where access to education is relatively costly, the supply of human capital will be lower. On the other hand, the productivity of human capital also differs across locations. Where human capital is more productive, demand for it will be higher.

Migration, trade, and innovation interact with these geographic disparities in human capital to shape economic development across the globe. People may choose to move to areas where acquiring human capital is less costly

*Desmet: Department of Economics and Cox School of Business, Southern Methodist University, kdesmet@smu.edu, Nagy: CREI, Universitat Pompeu Fabra and Barcelona GSE, dnagy@crei.cat, Rossi-Hansberg: Department of Economics, University of Chicago, earossih@uchicago.edu. We thank Reigner Kane and Julian Tsang for outstanding research assistance.

[1] Data at the national and subnational level come from the Global Data Lab (Smits and Permanyer, 2019). As we explain in Section 3.2, we transformed these data to the $1^\circ \times 1^\circ$ grid cell level.

— particularly if those areas are also economically advanced. At the same time, residents of low human capital locations may benefit from trade by importing goods produced in more productive, high human capital regions. Because both migration and trade are costly, the strength of these forces depends on the degree of spatial frictions to moving people and goods. Dynamic forces can amplify spatial disparities, as locations with more human capital innovate more.

To study how human capital shapes the economic geography of development, we propose a dynamic spatial framework in the spirit of Desmet et al. (2018), but with an explicit role for human capital. The world economy is represented as a two-dimensional surface, which in the quantitative implementation is divided into 16,433 $1^\circ \times 1^\circ$ grid cells with positive land mass. Agents work and consume a basket of differentiated goods. They choose both where to live and how much human capital to acquire. Upgrading human capital is costly, as is moving between locations, and these costs vary across space. Firms produce using labor, human capital, and land. Each location produces a different variety, and trade across locations is subject to transport costs. A location's productivity depends on both TFP and human-capital-augmenting technology. TFP benefits from agglomeration economies: as density increases, so does TFP. Human-capital-augmenting productivity evolves over time, depending on diffusion from other locations and on firms' decisions on how much to locally innovate. This generates dynamic agglomeration economies: dense locations with more human capital innovate more, which further increases the demand for human capital, and attracts more migrants.

We quantify the model using high-resolution data on population, income, and schooling. We then simulate the model forward for 200 years, holding constant the location-specific costs of moving people, transporting goods, and acquiring human capital. The main finding is one of persistence — today's developed regions remain the most developed 200 years from now. The same persistence holds for population density — locations that are dense today continue to be dense two centuries from now. The transition to the balanced growth path is long and protracted; even after two centuries, the world economy remains far from steady-state growth.

Central to this persistence in the geography of development is the negative correlation between the local cost of education and economic fundamentals related to productivity and amenities. We identify local education costs from the quantified model, using schooling data in two different time periods. To externally validate these model-generated costs, we show that they are consistent with household spending per student as a share of income. Because highly developed regions tend to have low education costs, their human capital levels tend to be high, both in the short and the long run. This advantage is magnified by dynamic feedback loops over the transition path: current levels of human capital improve future productivity, because human capital is an input in the growth of human-capital-augmenting productivity. As a result, the low cost of acquiring human capital in the developed world keeps these regions ahead, generating the persistence we project.

This persistence stands in sharp contrast to the findings in Desmet et al. (2018). In their model, where there is no role for human capital, market size and local density are the strongest predictors of future development. Because across the world many dense regions tend to be poor, there is convergence as low-productivity, high-density locations catch up with the world's more productive regions. Although the model in this paper also includes agglomeration economies that benefit dense locations, these are not strong enough to offset the human capital disadvantages of densely populated regions that are less developed.

By exploring development policies through counterfactual experiments that affect the cost of schooling across the globe, our analysis yields two further insights. The first is that proportionally lowering the cost of human capital in different regions of the world generates local gains, but may in some cases cause global losses. Whether we implement such a policy in sub-Saharan Africa, Latin America, or Central and South Asia, the local economy benefits: human capital levels rise, both in the short and the long run. This enhances innovation in human-capital-augmenting productivity, generating positive dynamic effects. Higher local welfare retains a larger share of the global population in the target region, further reinforcing productivity through agglomeration economies.

While the target region gains in terms of development and welfare, the effect on the world as a whole is more complex. In the case of sub-Saharan Africa and Central and South Asia, the world becomes worse off over a 200-year horizon, whereas in the case of Latin America, the world improves. When the policy is implemented in a low-income region, such as sub-Saharan Africa, the local increase in population comes at the expense of regions that have better fundamentals. With fewer people residing in the world's most productive regions, global income and welfare decline. In contrast, when the policy is implemented in a middle-income region, such as Latin America, the local increase in population comes at the expense of regions that have relatively worse fundamentals. As a result, global income and welfare improve.

The finding that improving the education technology in certain regions may worsen global welfare reflects inefficiencies in the spatial distribution of population. If education in sub-Saharan Africa becomes more affordable, fewer people reside in the world's most productive places. The lower geographic concentration of population in those places weakens local agglomeration economies. In addition, by having smaller stocks of human capital, the more productive places generate less innovation — and ultimately less diffusion of these innovations to the less productive parts of the world. Because individual migrants do not take these effects into account when making location decisions, the decentralized spatial allocation need not be efficient. Hence, policies that improve local education technology may, in some cases, lead to lower global welfare.

The second insight is that equalizing opportunity of education across space may lead to unintended consequences by reallocating population from areas with better fundamentals to those with worse fundamentals. To illustrate this, we consider a policy experiment that sets schooling costs to a common lower level across sub-Saharan Africa. While equalizing costs yields short-run local gains in real income per capita, after about half a century, income starts to decline, relative to the baseline. Why might local real income per capita fall even as access to human capital becomes cheaper? The answer lies in the spatial reallocation of population within Africa. Because schooling costs are higher in less developed areas, equalizing them lowers costs more in those areas. This encourages migration toward areas with weaker fundamentals, so that an increasingly larger share of the sub-Saharan African population resides in less productive locations. This reallocation also increases the geographic dispersion of the population, weakening agglomeration economies and further reducing real income per capita. The somewhat counterintuitive effect underscores the importance of accounting for spatial disparities within broad regions such as sub-Saharan Africa when evaluating education policies.

This paper is most closely related to a small literature that uses spatial models to analyze educational choices and human capital formation. Burzynski et al. (2020) also propose a spatial model of the world economy that endogenizes skill acquisition. The spatial resolution is coarse, with each country divided into an urban and a rural region, and there is no trade. Agents live for two periods, with each adult choosing where to live and how much time to invest in improving the odds that their offspring become skilled. Assuming a gradual convergence in the access to education, they find that human capital inequality remains fairly persistent over the span of the 21^{st} century. While this resembles some of our findings, their results are not directly comparable because they do not report outcomes at the local level. Another important difference is that we contrast our findings with those from a model without human capital accumulation, highlighting the important differences between the two.

Another relevant paper in this area is Eckert and Kleineberg (2024) who use a spatial framework to examine the effects of equalizing school funding across U.S. counties. As expected, differences in educational outcomes decline, but this comes at the cost of reducing the supply of educated workers in highly productive counties that initially had high education funding. The negative impact in these counties is partly offset by the immigration of high-skilled workers from elsewhere. This paper differs from ours in two respects: it centers on the U.S., rather than the globe; and it is more narrowly focused on the interaction between school funding, educational outcomes, and the spatial distribution of skills, rather than on the broader impacts on the geography of growth and development.

Our paper is also related to the development accounting literature, which measures the contribution of human capital to economic development. Hall and Jones (1999) estimate a relatively modest contribution of about 10%. When accounting for cross-country differences in educational quality, Schoellman (2012) finds that this share rises to 20%. Other papers suggest an even larger role for human capital. Jones (2014) takes a broader view of human capital based on a model where workers provide differentiated services. In such a model, human capital can fully account for income differences between rich and poor countries. Similarly, Lucas (2015) develops a model of social learning where human capital is the only engine of growth, and shows that its predictions are consistent with the data. While related, our focus is on the interaction between human capital, economic growth, and spatial frictions.

Of interest as well is the literature on the effect of schooling on skilled-labor-augmenting technology. In a cross-country study of 28 manufacturing industries, Ciccone and Papaioannou (2009) find that countries with higher initial levels of schooling experience faster growth in schooling-intensive industries. Similarly, Caselli and Coleman (2006) show that higher-income countries use skilled labor more efficiently than lower-income countries. These results are consistent with the idea that human capital promotes the use of human-capital-augmenting technologies. This mechanism is present in our model too.

The rest of the paper is organized as follows. Section 2 develops a dynamic spatial model focusing on the geography of human capital accumulation. Section 3 shows how to solve the model, describes the data, and gives details about the quantification. Section 4 simulates the model forward to examine how human capital shapes the geography of development. Section 5 conducts counterfactual policy experiments that improve access to education in different parts of the world. Section 6 concludes.

2 Theory

In this section, we develop a spatial dynamic model that examines how human capital shapes the geography of development.

2.1 Endowments and Preferences

The world economy occupies a two-dimensional surface R, where a location is defined as a point $r \in R$. Location r has a land endowment $N(r)$. There are L agents in the world economy, each supplying one unit of labor and an amount of human capital.

Preferences. The period-t utility of agent i who resides at r_t with human capital h_t is given by:

$$U_t^i(r_{-t}, h_{-t}, r_t, h_t) = \frac{a_t(r_t) \cdot \left[\int_R c_t(r, r')^{\frac{\xi-1}{\xi}} dr'\right]^{\frac{\xi}{\xi-1}} \cdot \varepsilon_t^i(r_t)}{\prod_{t'=1}^{t} m(r_{t'-1}, r_{t'})\, s(h_{t'-1}, h_{t'}, r)}, \tag{1}$$

where $r_{-t} = (r_0, \ldots, r_{t-1})$ and $h_{-t} = (h_0, \ldots, h_{t-1})$ refer to the residences and the levels of human capital of the agent in past periods, $a_t(r)$ denotes local amenities, $c_t(r, r')$ is the consumption in location r of the variety produced in location r', $\xi > 1$ is the elasticity of substitution between the different varieties, $\varepsilon_t^j(r)$ is a location preference shock drawn from a Fréchet distribution with shape parameter $1/\Omega$, $m(r_{t'-1}, r_{t'})$ is the cost of moving from $r_{t'-1}$ in period $r' - 1$ to $r_{t'}$ in period t', and $s(h_{t'-1}, h_{t'}, r)$ is the cost in location r of changing human capital from $h_{t'-1}$ in period $t' - 1$ to $h_{t'}$ in period t'. Though (1) only displays the period-utility, the agent maximizes the present discounted sum of future utility applying a discount factor β.

There are two congestion forces in our model coming from preferences. A first relates to local amenities which take the form $a_t(r) = \bar{a}(r) L_t(r)^{-\lambda}$, where $\bar{a}(r)$ denotes a location's fundamental amenity, and $L_t(r)^{-\lambda}$ represents a dispersion or congestion force coming from local population density. Second, the existence of idiosyncratic preference shocks, $\varepsilon_t^j(r)$, implies that people are heterogeneous in their tastes for a particular location. The larger the value of Ω, the greater that heterogeneity, and the stronger the spatial dispersion force.

Migration costs. As in Desmet et al. (2018), the permanent flow utility cost of moving from r to r' is the product of an origin-specific cost, $m_1(r)$, and a destination-specific cost, $m_2(r')$, so that $m(r, r') = m_1(r) m_2(r')$. Remaining in the same place is costless, so $m(r, r) = m_1(r) m_2(r) = 1$. This implies that the cost of leaving a location is the inverse of the cost of entering that location, i.e., $m_2(r) = m_1(r)^{-1}$. As a result, the permanent utility flow cost paid by an immigrant when entering r' is compensated by a permanent utility flow benefit of the same magnitude when leaving r'. Hence, migrants only pay the flow utility cost while residing in the host location, making any decision to migrate fully reversible. This simplifies an agent's forward-looking migration decision to a sequence of static decisions.

Human capital accumulation costs. We make analogous assumptions to simplify the forward-looking human capital accumulation decision. Specifically, the permanent flow utility cost of upgrading human capital from h to h' in location r takes the the form $s(h, h', r) = s_1(h, r) s_2(h', r)$. This construct implies that maintaining human capital permanently at the higher level h' requires a sustained per-period utility cost, in line with the theoretical and empirical work on human capital depreciation (Ben-Porath, 1967; Aghion et al., 2024). In any period, agents may choose to raise, maintain, or lower their human capital. The choice of lowering human capital can be interpreted as letting it depreciate. By assuming that $s_1(h, r) = s_2(h, r)^{-1}$, the permanent flow utility cost of raising human capital from h to h' is exactly offset by a permanent flow utility benefit when human capital declines from h' back to h. As in the migration case, this makes the forward-looking human capital accumulation decision fully reversible, reducing it to a static choice.

Income and welfare. Residents of a location r in period t with human capital $h_t(r)$ have three sources of income. First, they earn income from the one unit of labor they supply, $w_t^L(r)$. Second, they receive additional income $w_t^H(r) h_t(r)$ from their human capital. Third, they get a proportional share of local land rents, $R_t(r)/L_t(r)$. We can define $u_t(r)$, the utility associated with local amenities, real income and the cost of acquiring human capital, as:

$$u_t(r) = \frac{a_t(r) \cdot \left(w_t^L(r) + w_t^H(r) h_t + \frac{R_t(r)}{L_t(r)} \right)}{P_t(r) s_2(h_t(r), r)}, \tag{2}$$

where P_t is the ideal price index. We use $u_t(r)$ as our measure of social welfare, though it does not include the idiosyncratic preferences of agents for a location nor mobility costs. As argued by Desmet et al. (2018), this is a good measure of the desirability of a location.

2.2 Technology

At each location r, a representative firm produces a location-specific good. Hence, goods are differentiated à la Armington. The market is perfectly competitive, and firms take prices as given. Shipping goods from r to r' is subject to iceberg costs, $\varsigma(r', r) \geq 1$.

Production. The production per unit of land of a firm in location r at time t is given by

$$Q_t(r) = T_t(r)\left(L_t(r)^{\frac{\rho-1}{\rho}} + \left(A_t(r)\, H_{\phi,t}(r)^{\gamma_1}\, H_t(r)^{1-\gamma_1}\right)^{\frac{\rho-1}{\rho}}\right)^{\mu\frac{\rho}{\rho-1}}, \tag{3}$$

where $T_t(r)$ is the location's TFP, $L_t(r)$ is labor, $H_t(r)$ is human capital used in production, $A_t(r)$ is the location's human-capital-augmenting productivity, $H_{\phi,t}(r)$ is human capital used for innovation,[2] $\rho > 1$ is the elasticity of substitution between labor and human capital in the production process, and $1-\mu \in (0,1)$ is the share of land in a firm's expenditures. Land use by firms, together with land being available in fixed supply, implies a congestion force on the technology side of the model.

TFP and human-capital-augmenting productivity. A location's TFP, $T_t(r)$, benefits from agglomeration economies coming from local employment density, so

$$T_t(r) = \tau(r)\, L_t(r)^{\alpha}, \tag{4}$$

where $\tau(r)$ is the location's fundamental productivity, and $\alpha \geq 0$ denotes the strength of agglomeration economies. A location's human-capital-augmenting productivity, $A_t(r)$, evolves according to

$$A_{t+1}(r) = H_{\phi,t}(r)^{\gamma_1}\left[\int_R \eta A_t(r')\, dr'\right]^{1-\gamma_2} A_t(r)^{\gamma_2}, \tag{5}$$

where γ_1 is the elasticity of human-capital-augmenting productivity to the human capital used for innovation, and $\gamma_2 \in (0,1]$ is a parameter driving the strength of global technology diffusion. If $\gamma_2 = 1$, there is no such technology diffusion and each location grows only as a result of local innovations. The model generates a dynamic agglomeration effect: more innovation today increases human-capital-augmenting productivity, attracting more people and thereby fueling more innovation tomorrow.

2.3 Utility Maximization

Following the argument in Desmet et al. (2018), in each period agents decide where to locate and how much human capital to accumulate, independently of past or future decisions. Specifically, conditional on the distribution of

[2]Note that human capital used for innovation at time t contributes to output at time t. It also serves as an input for human-capital-augmenting productivity at time $t+1$, as explained in (5).

current wages, prices and population, in period t an agent chooses the location where to reside and the level of human capital by maximizing her period utility:

$$\max_{r,h_t(r)} \frac{a_t(r) \cdot \left(w_t^L(r) + h_t(r) w_t^H(r) + R_t(r)/L_t(r)\right) \cdot \varepsilon_t^i(r)}{P_t(r) \cdot m_2(r) s_2(h_t(r), r)}. \tag{6}$$

The first-order condition with respect to human capital yields

$$\frac{\partial s_2(h_t(r), r)/\partial h_t(r)}{s_2(h_t(r), r)} = \frac{w_t^H(r)}{w_t^L(r) + h_t(r) w_t^H(r) + R_t(r)/L_t(r)}, \tag{7}$$

so that everyone in location r in period t chooses the same level of human capital, independently of where they resided before. Following Desmet et al. (2018), we can show that the number of people choosing to live in r in period t is given by

$$N(r) L_t(r) = \frac{u_t(r)^{1/\Omega} m_2(r)^{-1/\Omega} s_2(h_t(r), r)^{-1/\Omega}}{\int_R u_t(r')^{1/\Omega} m_2(r')^{-1/\Omega} s_2(h_t(r'), r')^{-1/\Omega} dr'} L. \tag{8}$$

2.4 Profit Maximization

Firms are perfectly competitive. Taking all prices as given, a firm in location r chooses its inputs $L_t(r)$, $H_{\phi,t}(r)$ and $H_t(r)$, subject to the technology in (3), to maximize its static profits per unit of land,

$$p_t(r,r) Q_t(r) - w_t(r) L_t(r) - w_t^H(r_t)(H_{\phi,t}(r) + H_t(r)) - R_t(r), \tag{9}$$

where $p_t(r,r)$ is the price of variety r sold in r. The first-order conditions of this problem yield

$$\frac{H_t(r)^{1-(1-\gamma_1)\frac{\rho-1}{\rho}}}{L_t(r)^{\frac{1}{\rho}}} = (1-\gamma_1) A_t(r)^{\frac{\rho-1}{\rho}} H_{\phi,t}(r)^{\gamma_1 \frac{\rho-1}{\rho}} \frac{w_t^L(r)}{w_t^H(r)}, \tag{10}$$

$$\frac{H_{\phi,t}(r)}{H_t(r)} = \frac{\gamma_1}{1-\gamma_1}, \tag{11}$$

and

$$p_t(r,r) Q_t(r) = \mu^{-1} W_t(r)^{1-\rho} w_t^L(r)^{\rho} L_t(r), \tag{12}$$

where

$$W_t(r) = \left[w_t^L(r)^{1-\rho} + \gamma_1^{\gamma_1(\rho-1)}(1-\gamma_1)^{(1-\gamma_1)(\rho-1)} A_t(r)^{\rho-1} w_t^H(r)^{1-\rho}\right]^{\frac{1}{1-\rho}} \tag{13}$$

is the CES price index for labor and human capital. The zero-profit condition implies that land rents are given by

$$R_t(r) = \frac{1-\mu}{\mu} W_t(r)^{1-\rho} w_t^L(r)^{\rho} L_t(r). \tag{14}$$

2.5 Equilibrium

Human capital market clearing. From (7), (12) and (14), we obtain that human capital supply is the solution to

$$\frac{\partial s_2\left(h_t\left(r\right),r\right)/\partial h_t\left(r\right)}{s_2\left(h_t\left(r\right),r\right)}=\frac{\mu w_t^H\left(r\right)}{W_t\left(r\right)^{1-\rho}w_t^L\left(r\right)^{\rho}} \tag{15}$$

Now assume that s_2 takes the following functional form:

$$s_2\left(h,r\right)=e^{\sigma\left(r\right)h}, \tag{16}$$

where $\sigma\left(r\right)>0$ is the exogenous cost of accumulating human capital. We allow $\sigma\left(r\right)$ to vary exogenously across locations. Under this specification, equation (15) reduces to

$$\frac{w_t^H\left(r\right)}{\mu^{-1}W_t\left(r\right)^{1-\rho}w_t^L\left(r\right)^{\rho}}=\sigma\left(r\right). \tag{17}$$

This is the inverse supply function of human capital — which is perfectly elastic.

From (10) and (11), we can express the demand for human capital at r as

$$h_t\left(r\right)=\frac{H_t\left(r\right)+H_{\phi,t}\left(r\right)}{L_t\left(r\right)}=\gamma_1^{\gamma_1\left(\rho-1\right)}\left(1-\gamma_1\right)^{\left(1-\gamma_1\right)\left(\rho-1\right)}A_t\left(r\right)^{\rho-1}\left[\frac{w_t^L\left(r\right)}{w_t^H\left(r\right)}\right]^{\rho} \tag{18}$$

We can rewrite this expression as

$$w_t^H\left(r\right)=\gamma_1^{\gamma_1\frac{\rho-1}{\rho}}\left(1-\gamma_1\right)^{\left(1-\gamma_1\right)\frac{\rho-1}{\rho}}A_t\left(r\right)^{\frac{\rho-1}{\rho}}h_t\left(r\right)^{-\frac{1}{\rho}}w_t^L\left(r\right) \tag{19}$$

Equalizing demand (18) and supply (17), and using (19), yields

$$\mu\gamma_1^{\gamma_1\frac{\rho-1}{\rho}}\left(1-\gamma_1\right)^{\left(1-\gamma_1\right)\frac{\rho-1}{\rho}}\frac{A_t\left(r\right)^{\frac{\rho-1}{\rho}}}{\sigma\left(r\right)}h_t\left(r\right)^{-\frac{1}{\rho}}=1+\gamma_1^{\gamma_1\frac{\rho-1}{\rho}}\left(1-\gamma_1\right)^{\left(1-\gamma_1\right)\frac{\rho-1}{\rho}}A_t\left(r\right)^{\frac{\rho-1}{\rho}}h_t\left(r\right)^{\frac{\rho-1}{\rho}} \tag{20}$$

from which we get that the equilibrium amount of human capital per capita is implicitly given by

$$\frac{\frac{\mu}{\sigma\left(r\right)}-h_t\left(r\right)}{h_t\left(r\right)^{\frac{1}{\rho}}}=\gamma_1^{-\gamma_1\frac{\rho-1}{\rho}}\left(1-\gamma_1\right)^{-\left(1-\gamma_1\right)\frac{\rho-1}{\rho}}A_t\left(r\right)^{-\frac{\rho-1}{\rho}}. \tag{21}$$

Combining (11) with $H_t\left(r\right)+H_{\phi,t}\left(r\right)=h_t\left(r\right)L_t\left(r\right)$ yields:

$$H_{\phi,t}\left(r\right)=\gamma_1 h_t\left(r\right)L_t\left(r\right). \tag{22}$$

Prices. Under perfect competition, the price of the good produced in location r is equal to its marginal cost of production, so

$$p_t\left(r,r\right)=mc_t\left(r\right)=\mu^{-1}T_t\left(r\right)^{-1}W_t\left(r\right)^{1-\left(1-\mu\right)\rho}w_t^L\left(r\right)^{\left(1-\mu\right)\rho}L_t\left(r\right)^{1-\mu}. \tag{23}$$

Because of iceberg shipping costs, the price of the variety produced in r and sold in r' is $p_t(r', r) = \varsigma(r', r)\, p_t(r, r)$. Given the utility function, the price index can then be written as

$$P_t(r)^{1-\xi} = \int_R p_t(r', r')^{1-\xi} \varsigma(r, r')^{1-\xi} dr'. \tag{24}$$

Goods market clearing. Goods market clearing implies that the revenues of firms producing the variety of location r equals total spending on this variety by all locations, namely,

$$W_t(r)^{1-\rho} w_t^L(r)^{\rho} N(r) L_t(r) = \int_R p_t(r, r)^{1-\xi} \varsigma(r', r)^{1-\xi} P_t(r')^{\xi-1} W_t(r')^{1-\rho} w_t^L(r')^{\rho} N(r') L_t(r') dr'. \tag{25}$$

Combining (23), (24) and (25) yields

$$P_t(r)^{1-\xi} = \mu^{\xi-1} \int_R T_t(r')^{\xi-1} W_t(r')^{-(1-(1-\mu)\rho)(\xi-1)} w_t^L(r')^{-(1-\mu)\rho(\xi-1)} L_t(r')^{-(1-\mu)(\xi-1)} \varsigma(r, r')^{1-\xi} dr' \tag{26}$$

and

$$\begin{aligned} &T_t(r)^{1-\xi} W_t(r)^{1-\rho+(1-(1-\mu)\rho)(\xi-1)} w_t^L(r)^{[1+(1-\mu)(\xi-1)]\rho} N(r) L_t(r)^{1+(1-\mu)(\xi-1)} \\ &= \mu^{\xi-1} \int_R P_t(r')^{\xi-1} W_t(r')^{1-\rho} w_t^L(r')^{\rho} N(r') L_t(r') \varsigma(r', r)^{1-\xi} dr'. \end{aligned} \tag{27}$$

We can rewrite $u_t(r)$ in equation (2) as

$$u_t(r) = \bar{a}(r) L_t(r)^{-\lambda} y_t(r), \tag{28}$$

where $y_t(r) = \left(w_t^L(r) + w_t^H(r) h_t(r) + R_t(r)/L_t(r)\right)/P_t(r)$. From (12), we can rewrite real income as $y_t(r) = \mu^{-1} W_t(r)^{1-\rho} w_t^L(r)^{\rho} P_t(r)^{-1}$. As a result, we can express the price index at r as

$$P_t(r) = \mu^{-1} \frac{\bar{a}(r)}{u_t(r)} W_t(r)^{1-\rho} w_t^L(r)^{\rho} L_t(r)^{-\lambda} \tag{29}$$

and use this relationship in equilibrium conditions (26) and (27). Assuming $\varsigma(r, r') = \varsigma(r', r)$, one can show that these two equilibrium conditions hold if and only if

$$T_t(r) = \frac{\bar{a}(r)}{u_t(r)} W_t(r)^{(1-\rho)\frac{\xi}{\xi-1}+1-(1-\mu)\rho} w_t^L(r)^{\frac{\xi}{\xi-1}\rho+(1-\mu)\rho} N(r)^{\frac{1}{\xi-1}} L_t(r)^{\frac{1}{\xi-1}+1-\mu-\lambda} \tag{30}$$

and

$$\begin{aligned} &\left[\frac{\bar{a}(r)}{u_t(r)}\right]^{1-\xi} W_t(r)^{-(1-\rho)(\xi-1)} w_t^L(r)^{-\rho(\xi-1)} L_t(r)^{\lambda(\xi-1)} \\ &= \int_R \left[\frac{\bar{a}(r')}{u_t(r')}\right]^{\xi-1} W_t(r')^{(1-\rho)\xi} w_t^L(r')^{\rho\xi} N(r') L_t(r')^{1-\lambda(\xi-1)} \varsigma(r', r)^{1-\xi} dr'. \end{aligned} \tag{31}$$

Dynamic competitive equilibrium. For a given period t, and a given distribution of fundamental amenities, $\bar{a}(\cdot)$, fundamental productivities, $\tau(\cdot)$, initial human-capital-augmenting productivities, $A_0(r)$, equations (8), (13), (17), (21), (30), and (31) pin down for each location the labor wage, $w_t^L(\cdot)$, the wage per unit of human capital,

$w_t^H(\cdot)$, the CES wage aggregator, $W_t(\cdot)$, population, $L_t(\cdot)$, utility, $u_t(\cdot)$, and human capital, $h_t(\cdot)$. These conditions determine the period-t equilibrium. Equations (22) and (5) then yield human-capital-augmenting productivity at time $t+1$.

2.6 Balanced Growth Path

The growth rate of productivity at location r can be expressed from equation (5) as

$$\frac{A_{t+1}(r)}{A_t(r)} = H_{\phi,t}(r)^{\gamma_1} \left[\int_R \eta \frac{A_t(r')}{A_t(r)} dr'\right]^{1-\gamma_2}. \tag{32}$$

On a balanced growth path, the growth rate of productivity must be constant and identical across locations. Dividing equation (32) by the corresponding equation for location r', and using (22), then yields:

$$\frac{A_t(r')}{A_t(r)} = \left[\frac{H_{\phi,t}(r')}{H_{\phi,t}(r)}\right]^{\frac{\gamma_1}{1-\gamma_2}} = \left[\frac{\bar{h}(r')}{\bar{h}(r)}\right]^{\frac{\gamma_1}{1-\gamma_2}} \left[\frac{\bar{L}(r')}{\bar{L}(r)}\right]^{\frac{\gamma_1}{1-\gamma_2}}, \tag{33}$$

where the bars above variables refer to their values in a balanced growth path. We now characterize a balanced growth path.

Proposition 1. *In a balanced growth path (BGP) (i) the allocation of people across space is constant, $L_t(r) = \bar{L}(r)$, (ii) the level of human capital across space is constant, $h_t(r) = \bar{h}(r) = \frac{\mu}{\sigma(r)}$, and (iii) productivity $A_t(r)$ and utility $u_t(r)$ grow at a constant rate.*

Proof. It is straightforward to derive an expression for the constant growth rate of productivity for an allocation of labor and human capital that does not change over time. If $L_t(r)$ and $h_t(r)$ do not change over time, we can rewrite (33) as

$$\frac{A_t(r')}{A_t(r)} = \left[\frac{\bar{L}(r')\bar{h}(r')}{\bar{L}(r)\bar{h}(r)}\right]^{\frac{\gamma_1}{1-\gamma_2}} \tag{34}$$

and therefore

$$A_t(r) = \kappa_t \left[\bar{L}(r)\bar{h}(r)\right]^{\frac{\gamma_1}{1-\gamma_2}}. \tag{35}$$

This implies that $A_t(r)$ grows at the same rate at every location. From (32) and (22), this growth rate can be expressed as

$$\begin{aligned} \frac{A_{t+1}(r)}{A_t(r)} &= [\gamma_1 h_t(r) L_t(r)]^{\gamma_1} \left[\int_R \eta \frac{A_t(r')}{A_t(r)} dr'\right]^{1-\gamma_2} \\ &= \left[\gamma_1 \bar{L}(r)\bar{h}(r)\right]^{\gamma_1} \left[\int_R \eta \left[\frac{\bar{L}(r')\bar{h}(r')}{\bar{L}(r)\bar{h}(r)}\right]^{\frac{\gamma_1}{1-\gamma_2}} dr'\right]^{1-\gamma_2} \\ &= \gamma_1^{\gamma_1} \left[\int_R \eta \left[\bar{L}(r')\bar{h}(r')\right]^{\frac{\gamma_1}{1-\gamma_2}} dr'\right]^{1-\gamma_2}. \end{aligned} \tag{36}$$

Similarly, we can derive the growth rate of utility and real consumption at any location as

$$\frac{u_{t+1}(r)}{u_t(r)} = \gamma_1^{\frac{\gamma_1}{\mu}\frac{\xi-1}{2\xi-1}} \left[\int_R \eta \left[\bar{L}(r')\bar{h}(r')\right]^{\frac{\gamma_1}{1-\gamma_2}} dr'\right]^{\frac{1-\gamma_2}{\mu}\frac{\xi-1}{2\xi-1}}. \tag{37}$$

From (21), if $h(r)$ does not change over time and $A_t(r)$ grows at a constant rate, in steady state we have

$$\bar{h}(r) = \frac{\mu}{\sigma(r)}. \tag{38}$$

□

3 Solution, Data and Quantification

This section discusses how we solve the model, which data we use, and how we quantify the different parameter values.

3.1 Solving the Model

For a given initial distribution of human-capital-augmenting productivity $A_0(r)$, equations (8), (16), (17), (19), (21), (30) and (31) of the model allow us to solve for the equilibrium in period 0. Equation (5) can then be used to update human-capital-augmenting productivity at every location, which then allows us to solve for the equilibrium in period 1, and so on. In what follows, we explain how we can solve for the equilibrium distribution of economic activity in any period t as a function of $A_t(r)$ and fundamentals. The procedure consists of two steps. In the first step, we use equation (21) to solve for human capital per capita at every location, $h_t(r)$. In the second step, we use $h_t(r)$, $A_t(r)$ and fundamentals in the remaining equations to obtain the equilibrium distribution of $u_t(r)$, $L_t(r)$, $w_t^L(r)$, $W_t(r)$ and $w_t^H(r)$.

Although equation (21) cannot be solved for $h_t(r)$ in closed form, the following proposition shows that it has a unique solution for $h_t(r)$.

Proposition 2. *For any $A_t(r) > 0$ and $0 < \sigma(r) < \mu$, equation (21) has exactly one positive solution. Moreover, the solution satisfies $0 < h_t(r) < \frac{\mu}{\sigma(r)}$.*

Proof. See Appendix A.1. □

The next proposition shows how to solve for the remaining period-t endogenous variables.

Proposition 3. *Given $h_t(r)$, equations (8), (16), (17), (19), (30) and (31) pin down the remaining period-t endogenous variables as a function of $A_t(r)$ and fundamentals.*

Proof. See Appendix A.2. □

We can also show that an equilibrium exists and is unique if static agglomeration economies are not too strong compared to congestion forces.

Proposition 4. *An equilibrium exists and is unique if*

$$\alpha < \lambda + 1 - \mu + \Omega$$

and the equilibrium can be solved by iteration.

Proof. See Appendix A.3. □

3.2 Identifying the Cost of Human Capital and Other Fundamentals

To solve the model, we need values for exogenous location fundamentals $\sigma(r)$, $\bar{a}(r)$, $\tau(r)$, $A_0(r)$ and $m_2(r)$. We start by showing how we recover the exogenous cost of human capital, $\sigma(r)$. We write equation (21) in period 0:

$$\frac{\frac{\mu}{\sigma(r)} - h_0(r)}{h_0(r)^{\frac{1}{\rho}}} = \gamma_1^{-\gamma_1 \frac{\rho-1}{\rho}} (1-\gamma_1)^{-(1-\gamma_1)\frac{\rho-1}{\rho}} A_0(r)^{-\frac{\rho-1}{\rho}}$$

and divide this by the same equation in period 1. This yields

$$\frac{\hat{\sigma}_0(r)}{\hat{\sigma}_0(r) + h_0(r) - h_1(r)} = \left[\frac{h_0(r)}{h_1(r)}\right]^{\frac{1}{\rho}} \left[\frac{A_1(r)}{A_0(r)}\right]^{\frac{\rho-1}{\rho}}$$

where $\hat{\sigma}_0(r) = \frac{\mu}{\sigma(r)} - h_0(r)$. Next, using (5), (21) and (22), we can write this equation as

$$\tilde{\gamma}\tilde{A}_0^{-(1-\gamma_2)\frac{\rho-1}{\rho}} \frac{h_1(r)^{\frac{1}{\rho}}}{h_0(r)^{\gamma_1 \frac{\rho-1}{\rho} + \gamma_2 \frac{1}{\rho}}} L_0(r)^{-\gamma_1 \frac{\rho-1}{\rho}} \hat{\sigma}_0(r)^{\gamma_2} = \hat{\sigma}_0(r) + h_0(r) - h_1(r) \tag{39}$$

where $\tilde{\gamma} = \gamma_1^{-\gamma_1(2-\gamma_2)\frac{\rho-1}{\rho}} (1-\gamma_1)^{-(1-\gamma_1)(1-\gamma_2)\frac{\rho-1}{\rho}}$, and

$$\tilde{A}_0 = \gamma_1^{-\gamma_1} (1-\gamma_1)^{-(1-\gamma_1)} \int_R \eta \frac{h_0(r')^{\frac{1}{\rho-1}}}{\hat{\sigma}_0(r')^{\frac{\rho}{\rho-1}}} dr'. \tag{40}$$

Assume $h_1(r) > h_0(r)$ for every r.[3] For a given $\tilde{A}_0$, equation (39) then has a unique solution for $\hat{\sigma}_0(r)$. Once $\hat{\sigma}_0(r)$ is known, $\sigma(r)$ can be obtained as $\sigma(r) = \frac{\mu}{\hat{\sigma}_0(r) + h_0(r)}$. This suggests the following algorithm to back out the costs of human capital accumulation: (i) guess an $\tilde{A}_0$; (ii) solve equation (39) for $\hat{\sigma}_0(r)$; (iii) check if (40) holds; if it does not, adjust $\tilde{A}_0$ and go back to step (ii); (iv) if (40) holds, calculate $\sigma(r)$ for every r.

Armed with $\sigma(r)$, the next proposition states that we can uniquely recover other location fundamentals (except for moving costs) for given values of structural parameters.

Proposition 5. *Using data on land $N(r)$, population density $L_0(r)$, nominal income per capita $\mu^{-1} W_0(r)^{1-\rho} w_0^L(r)^{\rho}$, human capital costs $\sigma(r)$, and human capital per capita $h_0(r)$, at every location r, the equilibrium conditions uniquely identify $\frac{\bar{a}(\cdot)}{u_0(\cdot)}$, $\tau(\cdot)$ and $A_0(\cdot)$.*

Proof. See Appendix A.4. □

Once $\frac{\bar{a}(r)}{u_0(r)}$ has been obtained, subjective wellbeing data on $u_0(r)$ can be used to identify fundamental amenities $\bar{a}(r)$ at each location, as in Desmet et al. (2018).

Finally, observing the population distribution in period 1, $L_1(r)$, we can back out the cost of moving to a location, $m_2(r)$, using a procedure analogous to the one in Desmet et al. (2018). The following proposition presents this result.

Proposition 6. *Using data on $L_1(r)$, the equilibrium conditions identify moving costs $m_2(r)$ up to a scale.*

[3] In the quantitative implementation, $h_0(r)$ is based on the years of schooling in 2000, while $h_1(r)$ is based on an estimate of the years of schooling for 2001. This estimate is obtained by annualizing the growth rate in the years of schooling between 2000 and the maximum observed in 2005, 2010, 2015, and 2018, and then applying it to the 2000 level.

Proof. See Appendix A.5. □

Table 1: Parameter Values

Parameter	Target/Comment
1. Preferences	
$\xi = 7.5$	Elasticity of substitution equal to trade elasticity of 6.5 plus one[1]
$\lambda = 0.32$	Relation between amenities and population[1]
$\Omega = 0.5$	Inverse elasticity of migration flows with respect to income[1]
$\psi = 1.8$	Subjective well-being parameter[1]
2. Technology	
$\rho = 4$	Elasticity of substitution between labor and human capital
$\mu = 0.8$	Non-land share in production[1]
$\alpha = 0.06$	Agglomeration externalities[1]
$\gamma_1 = 0.005655253$	Share of human capital used for innovation
$\gamma_2 = 0.994344747$	Elasticity of A_{t+1} with respect to A_t
$\eta = 0.04388$	Scale of technology diffusion
3. Human Capital	
$\phi = 0.1$	Semi-elasticity of human capital with respect to years of schooling

Note: The parameter values with a subscript [1] come from Desmet et al. (2018). The choices for all other parameter values are further explained in the text.

3.3 Data and Quantification

Data. We discretize the world into 64,800 $1° \times 1°$ cells. At that level of resolution, we require data on the initial population density, $L_0(r)$, total output per unit of land, $\mu^{-1} W_0\left(r\right)^{1-\rho} w_0^L\left(r\right)^{\rho} L_0\left(r\right)$, land, $N(r)$, and human capital per capita, $h_0\left(r\right)$. We take the initial year to be the year 2000. Data on population, output, and land come from the G-Econ 4.0 database of Nordhaus et al. (2006). With the exception of a few countries, this dataset has global coverage.

Data on human capital comes from the Global Data Lab, which provides information on the average years of schooling at the level of subnational administrative regions (Smits and Permanyer, 2019).[4] We rasterize these data at the $30'' \times 30''$ resolution. Combining them with population data from the Gridded Population of the World Version 4 at the same resolution (CIESIN, 2018), we can compute average years of schooling at the $1° \times 1°$ resolution (Figure 1 panel (a)). We then transform the average years of schooling, denoted by $b\left(r\right)$, into human capital, using the transformation $h\left(r\right) = e^{\phi \cdot b(r)}$.

We also need data on subjective wellbeing and bilateral trade costs. For $u_0\left(r\right)$, we use the same data as Desmet et al. (2018), and for $\varsigma\left(r, r'\right)$, we take the estimates of that same paper. This allows for easy comparison of results.

Parameter values. Table 1 reports the parameter values we use to simulate the model. Many of these values are taken from Desmet et al. (2018), with two exceptions: some of the technology parameters, and those related to human capital, which are new to this paper.

To estimate the technology parameter values γ_1 and γ_2, recall that the growth rate of utility along the balanced growth path in equation (37) is the same as the growth rate of real income. Discretizing and taking logs, we can

[4] See https://globaldatalab.org/shdi/. Data accessed on April 29, 2021.

therefore rewrite (37) as

$$\log y_{t+1} - \log y_t = \alpha_1 + \alpha_2 \log \sum_R \left[L\left(r'\right) h\left(r'\right)\right]^{\alpha_3} \tag{41}$$

where α_1 is a constant and

$$\begin{aligned} \alpha_2 &= \frac{1-\gamma_2}{\mu} \frac{\xi}{2\xi - 1} \\ \alpha_3 &= \frac{\gamma_1}{1-\gamma_2}. \end{aligned}$$

The theory therefore gives us the following measure of effect of the spatial distribution of population on growth:

$$\sum_R \left[L\left(r'\right) h\left(r'\right)\right]^{\alpha_3}. \tag{42}$$

To ensure that our measure does not depend on a country's number of cells, we normalize (42) to

$$\frac{2}{R} \sum_R \left[L\left(r'\right) h\left(r'\right)\right]^{\alpha_3}, \tag{43}$$

where S is the number of cells in a country.

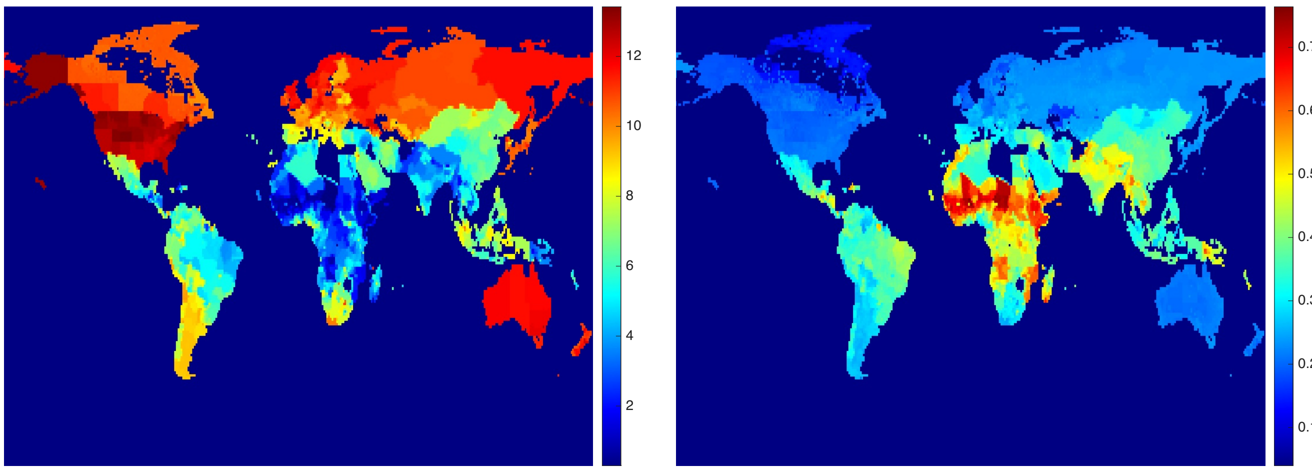
Using income per capita and population data from G-Econ and schooling data from the Global Data Lab, we compute measures of (43) and then estimate (41). For the estimation, we focus on countries with at least 10 grid cells, we use four years of data, 1990, 1995, 2000 and 2005, and use the between-estimator. Since (41) is a steady-state relationship, it makes sense to focus on the 5-year growth rates. Our estimates give values of $\alpha_2 = 0.003787$ and $\alpha_3 = 1.0$, which correspond $\gamma_1 = 0.005655253$ and $\gamma_2 = 0.994344747$. The parameter η which governs technological diffusion is then chosen to target a growth rate in nominal GDP per capita in 2000 of 3.4%.

As for the parameters related to human capital, we set ρ, the elasticity between labor and human capital, to 4. In a survey article of fifteen years ago, Acemoglu and Autor (2011) conclude that the elasticity of substitution between skilled and unskilled labor lies between 1.4 and 2. Jerzmanowski and Tamura (2023) report slightly higher estimates, between 1.8 and 2.6, based on data from 32 developed countries. Havranek et al. (2024) conduct a meta-analysis of 682 estimates from 77 different studies. After correcting for publication and attenuation bias, they find an implied mean elasticity of 4. We adopt this value for ρ in our model.

For the value of ϕ, we can write the semi-elasticity of human capital income, $w_H(r)\,h(r)$, with respect to years of schooling, $b(r)$, as $\frac{1}{w_H(r)h(r)} \frac{\partial w_H(r)h(r)}{\partial b(r)} = \phi$. This relative increase in income for an additional year of schooling can be interpreted as the skill premium for an additional year of schooling. Using data from Montenegro and Patrinos (2014), the average return to one more year of schooling is 10.0% when using data on 119 countries for the period 1995-2005. We therefore set ϕ to 0.1.

Cost of human capital. The location-specific parameters that measure the cost of acquiring human capital, $\sigma(r)$, are identified through the model. These costs are relatively high in regions with low initial schooling levels (Figure 1). Human capital acquisition costs are also lower in countries and cells with better fundamentals. To quantify a location's fundamentals, we regress log GDP per capita at $t = 1$ on fundamental amenities, fundamental productivity, human-capital-augmenting productivity, and migration costs. The predicted log GDP from this regression serves as our measure of a location's fundamentals. Figure 2, panels (a) and (c), shows at the country and grid-cell level the negative relationship between σ and fundamentals. Panels (b) and (d) show a similar negative relationship between σ and income per capita. At the country level, the correlation between log income per capita and $\sigma(r)$ is -0.79, and at the grid-cell level it is -0.83.

Figure 1: Years of Schooling and Cost of Education



(a) Average years of schooling 2000

(b) Education cost (σ)

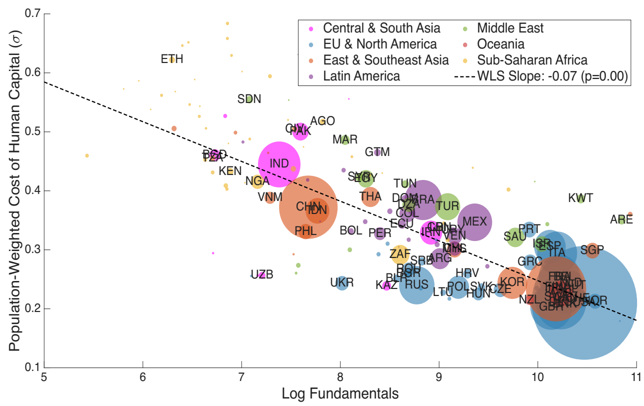
Note: Panel (a) shows the average years of schooling in year 2000. Panel (b) shows the education cost, as recovered from the model.

As we will see, this negative relationship between the model-implied cost of schooling and fundamentals plays a central role in the persistence of the geography of development. In particular, the relatively low cost of human capital acquisition in developed regions ensures higher levels of human capital, both in the short and the long run. These regions therefore remain ahead, generating persistence.

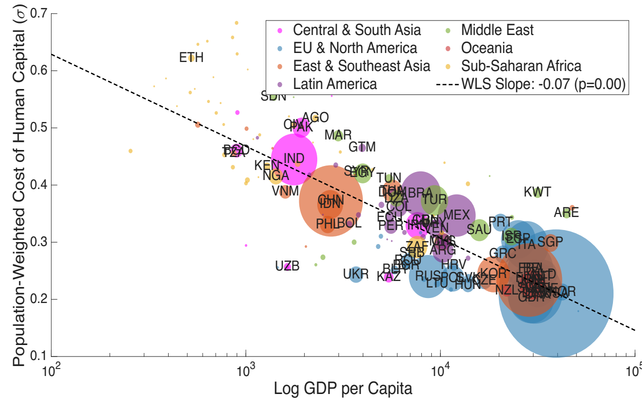
External validity test. As an external validity test, we examine whether the model-generated costs of education correlate well with actual data. Given that in the model $\sigma(r)$ determines an individual's demand for education, we focus on the private cost faced by the individual. In other words, we ignore government spending, since the demand for schooling is sensitive to the private, rather than the public, component of the cost.[5] Using data from the UNESCO Institute of Statistics, we take household spending per student as a share of income per capita. These data are available at the country level, and we focus on primary and secondary education. We then compare these data with the model-generated population-weighted values of σ at the country level. The scatterplot in Figure 3 shows a strong positive relationship between household spending on education and σ, with a correlation of 0.70. This suggests that our modeling choices and our quantification strategy yield reasonable estimates for the cross-country variation in σ.

[5] Of course, for a given level of private spending, an increase in public spending might improve the quality of a year of schooling. Hence, if our model were to incorporate quality concerns, public spending might also determine the incentives to acquire human capital.

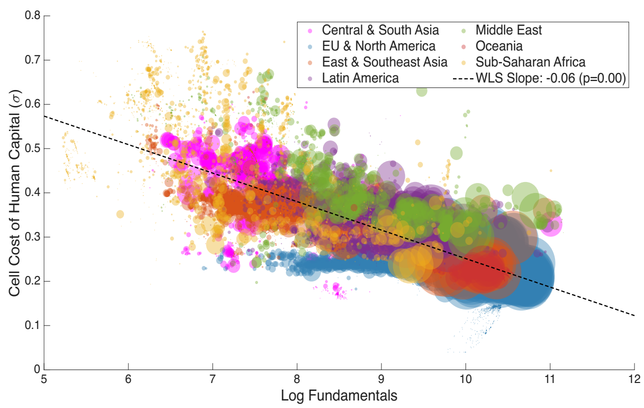
Figure 2: Cost of Education and GDP per Capita



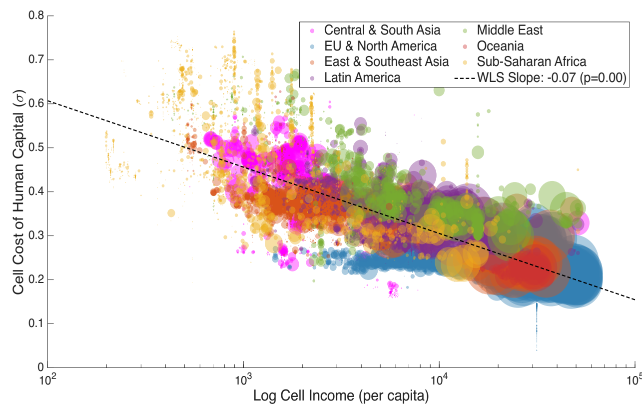
(a) σ vs fundamentals (country)



(b) σ vs log income per capita (country)



(a) σ vs fundamentals (cell)



(b) σ vs log income per capita (cell)

Note: This figure displays the relationships between σ and fundamentals (panel (a) for countries and panel (c) for grid cells), and between σ and income per capita in 2000 (panel (b) for countries and panel (d) for grid cells). The sizes of the circles are proportional to the size of GDP, and the colors represent different regions of the world.

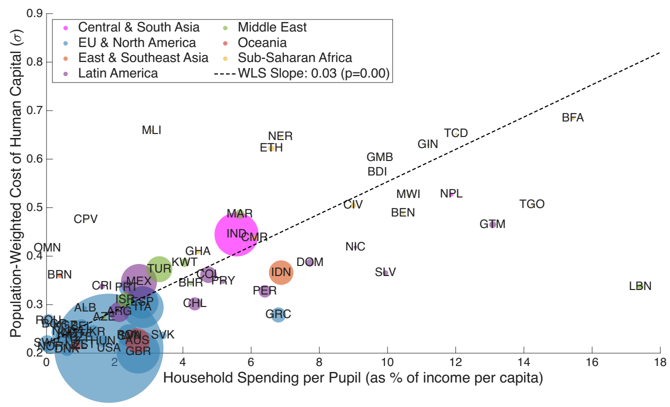
4 Human Capital Accumulation and the Geography of Development

This section examines how human capital shapes the geography of development. To explore how today's spatial variation in the cost of human capital affects the future geographic distribution of population and development, we simulate the model forward for 200 years, until the year 2200. We then conduct a series of counterfactual policy experiments that improve the access to schooling in different regions of the world.

By simulating our model forward for 200 years, we analyze how the geography of human capital and economic development evolves over time. As shown in Figure 4(a), some of today's poorest regions, particularly in sub-Saharan Africa and South Asia, are projected to experience relatively large gains in years of schooling. Yet, despite these improvements, in the long run their levels of schooling remain lower than those in the rest of the world (Figure 4(b)).

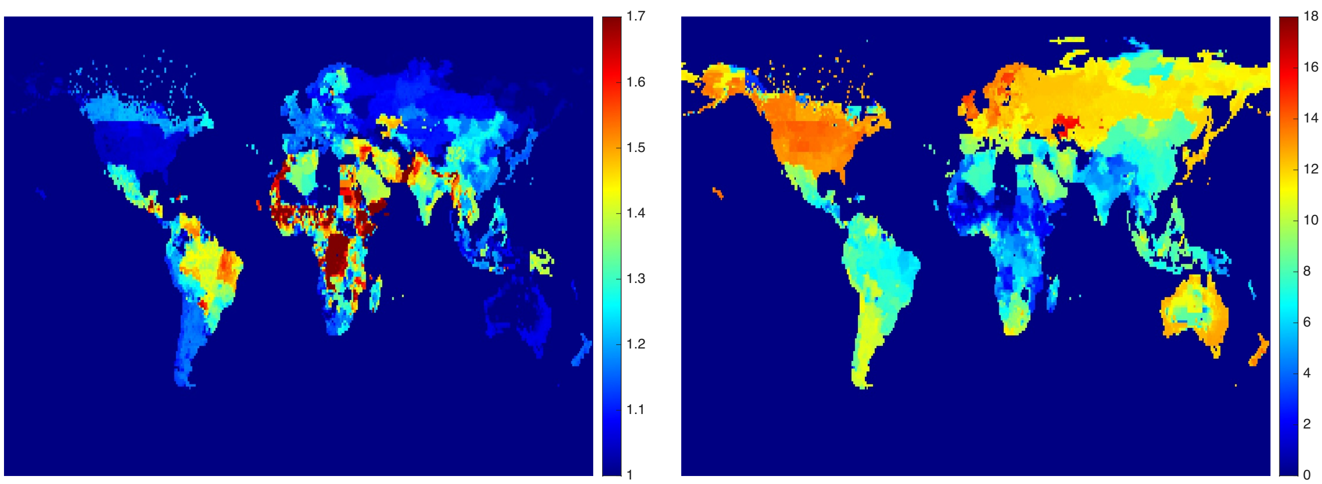
As a result, Figure 5 shows that in terms of development, the overall picture is one of persistence. Today's most developed regions — coastal Australia, Western Europe, Japan, and the U.S. — remain the most developed two centuries later. The geography of population density also changes little: today's densest places — parts of

Figure 3: Cost of Education and GDP per Capita



Note: This figure displays the relationship between σ and household spending per pupil as a share of income per capita.

Figure 4: Years of Schooling: 2200 vs 2000



(a) Years of schooling 2200 / years of schooling 2000

(b) Years of schooling 2200

Note: Panel (a) displays the ratio of the years of schooling in 2200, as predicted by the model, and the years of schooling in 2000, as observed in the data. Panel (b) displays the years of schooling in 2200, as predicted by the model.

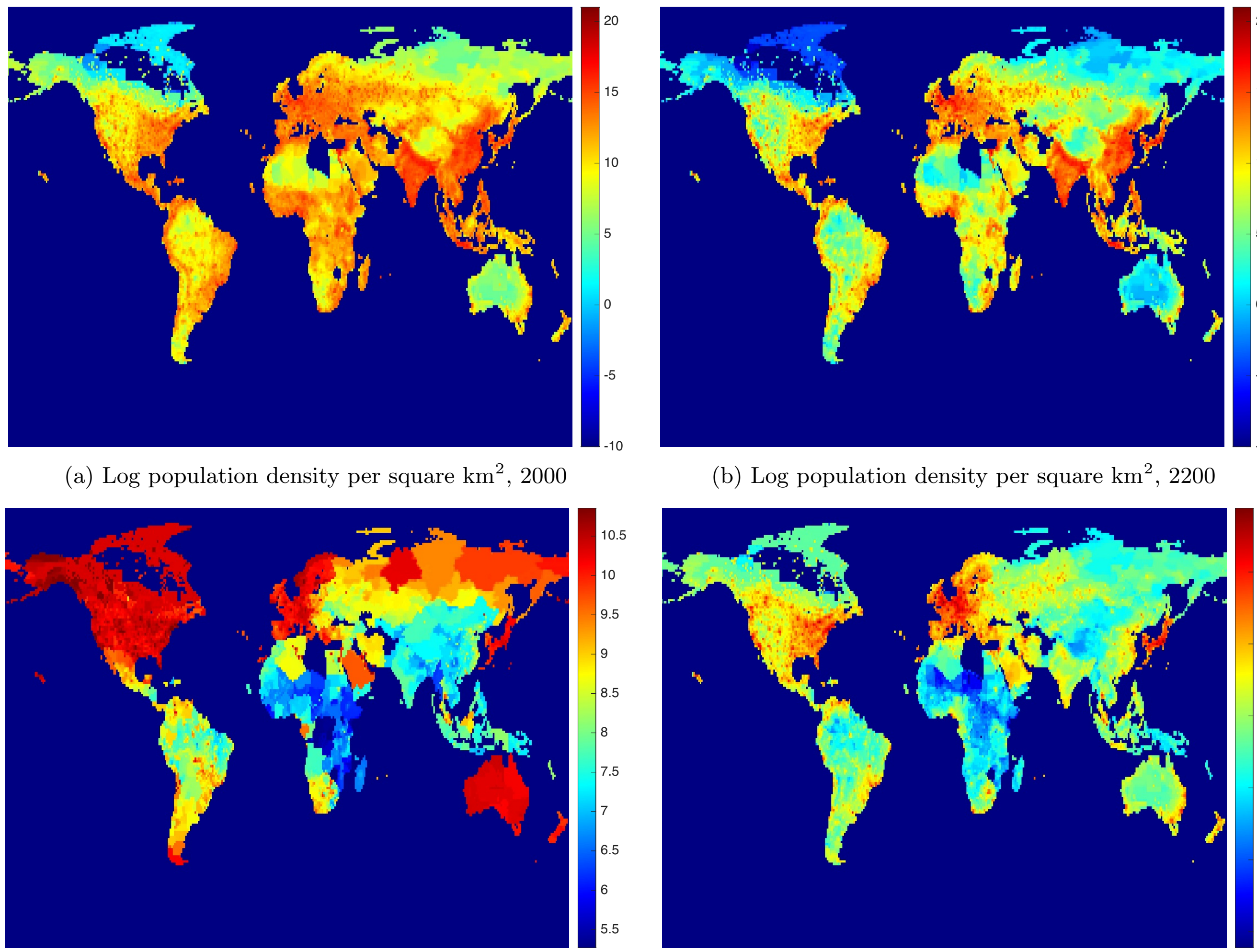
sub-Saharan Africa, eastern China, northwest Europe, India, and the US East Coast — are projected to still be the densest in 200 years.

At the core of this persistence in the geography of development is the negative correlation between local education costs and economic development. Since highly developed regions tend to face low education costs, their human capital levels remain consistently high, which in turn fosters innovation in human-capital-augmenting productivity. As a result, the low cost of acquiring human capital in the developed world allows it to maintain its lead, generating the persistence predicted by the model.

This remarkable persistence in the geography of development contrasts with models that omit human capital. In Desmet et al. (2018), where human capital accumulation plays no role, today's poorest places are projected to grow faster over the transition path. As a result, many low-productivity, high-density regions in sub-Saharan

Africa, South Asia and East Asia catch up with the world's most productive areas. This contrast is illustrated in Figure 6 which plots the long-run evolution of GDP per capita growth by income quintile: in the current model, more developed regions exhibit higher growth rates over the transition path, whereas the opposite holds in Desmet et al. (2018). Consistent with this, Figure 7 shows that the current model, which incorporates human capital accumulation, predicts higher future population density and GDP per capita in today's developed regions, compared to a model without human capital.

Figure 5: Population and GDP per Capita: 2200 vs 2000



(a) Log population density per square km^2, 2000

(b) Log population density per square km^2, 2200

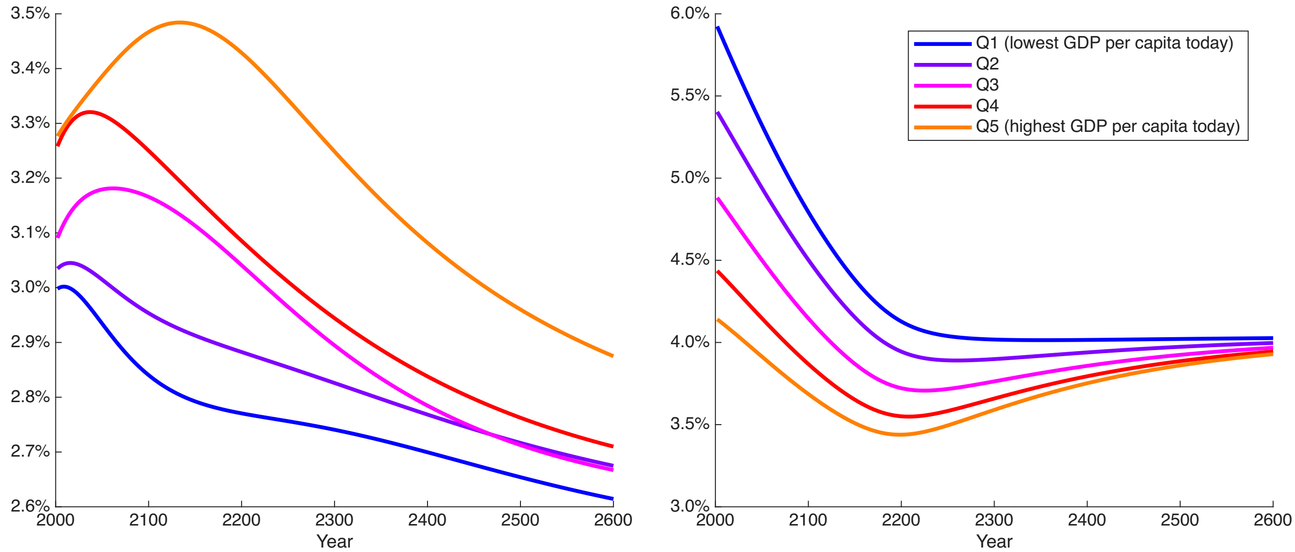
(c) Log GDP per capita (US$, PPP), 2000

(d) Log GDP per capita (US$, PPP), 2200

Note: Panel (a) displays population density in 2000, as observed in the data. Panel (b) displays population density in 2200, as predicted by the model. Panels (c) and (d) do the same for GDP per capita.

To explain these contrasting predictions, we must understand the two key forces at work in the current model. On the one hand, a location's TFP is increasing in local employment density. These agglomeration economies imply that denser locations tend to be more productive. Over time, this attracts more people, further increasing density and reinforcing productivity. On the other hand, over time each location's human capital converges to a level that depends on the cost of acquiring human capital. Locations with a lower $\sigma(r)$ achieve higher long-run

Figure 6: GDP per Capita Growth by GDP Quintile: Current Model vs Desmet et al. (2018)

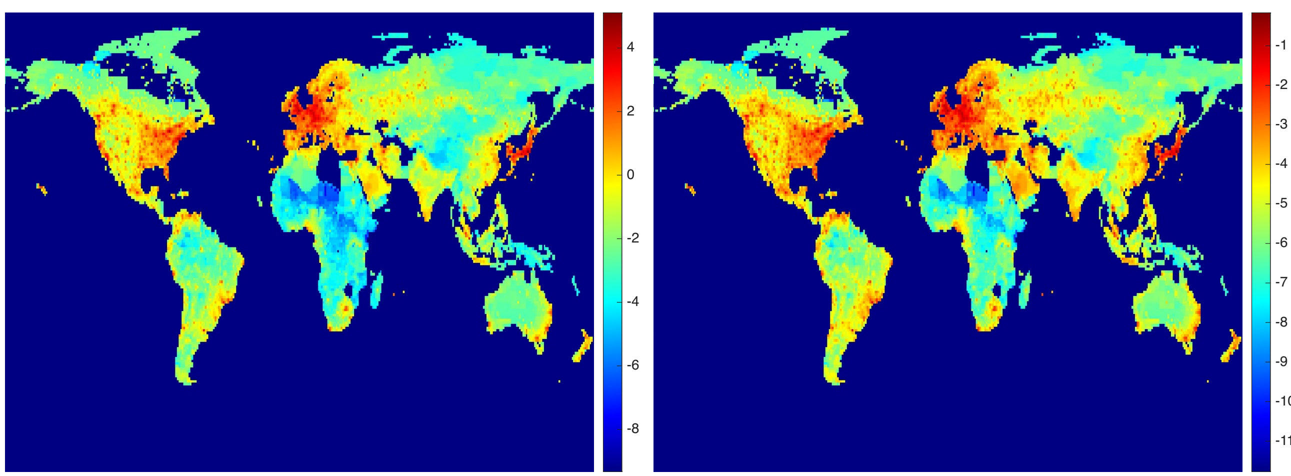


(a) GDP per capita growth by GDP quintile current model

(b) GDP per capita growth by GDP quintile (Desmet et al., 2018)

Note: Panel (a) displays GDP per capita growth by quintile based on GDP per capita in 2000, as predicted by the model. Panel (b) displays the same growth rates, as predicted by Desmet et al. (2018).

Figure 7: Population and GDP per Capita in 2200: Current Model vs Desmet et al. (2018)



(a) Log population density per square km^2, 2200, current model – Desmet et al. (2018)

(b) Log GDP per capita (US$, PPP), 2200, current model – Desmet et al. (2018)

Note: This figure displays the difference in log population density in 2200 (panel (a)) and log GDP per capita in 2200 (panel (b)) between the current model and Desmet et al. (2018).

levels of human capital. Since human capital affects TFP, these locations tend to be more productive.

The first force, linking density to productivity, tends to weaken persistence in the geography of development. Because some of today's densest regions are relatively poor, they are likely to grow and develop in the future, leading to some degree of convergence. The second force, linking the cost of human capital accumulation to long-run productivity, strengthens persistence. Because today's richest regions also tend to face lower costs of acquiring human capital, they are likely to remain ahead in terms of future development. In a model without human capital, the first force is present, but not the second. This explains why in Desmet et al. (2018) the densest regions in sub-Saharan Africa, South Asia, and East Asia become highly developed in the future. In contrast, in the current model with human capital, the second force dominates, leading to persistence in the geography of development.

Over time, both models converge to a balanced growth path, where the spatial distribution of population no longer changes and all regions grow at the same rate in terms of income per capita. Another difference between both models is the time it takes to reach the balanced growth path. While in Desmet et al. (2018) this transition is already long and protracted, with growth rates becoming similar only after 500 years, in the current model it appears to take even longer. Appendix Figure B1 confirms that, if run far enough into the future, the model does ultimately converge to balanced growth.

5 Improving Human Capital Accumulation around the World

In this section, we explore the implications of changing the cost of human capital accumulation in different regions of the world.

5.1 Proportionally Improving Access to Education in Sub-Saharan Africa

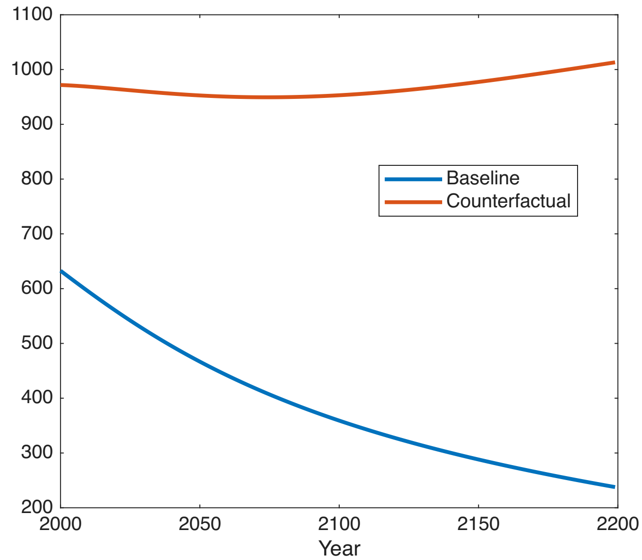
Our first policy experiment reduces the cost of human capital acquisition, $\sigma(r)$, by the same proportion across all grid cells of sub-Saharan Africa. To enable easy comparison with subsequent counterfactuals, we set this proportional reduction to 49.5%, equivalent to the population-weighted average decline that would occur if costs were lowered to the world median throughout the region. We then simulate the model forward for 200 years, and compare the resulting outcomes to those of the baseline.

Lowering the cost of human capital acquisition in sub-Saharan Africa reduces the incentives to emigrate, so the region's population remains larger (Figure 8). Part of this adjustment shows up as an upward discrete jump at the time of the policy change, reflecting the immediate increase in human capital when the policy gets implemented. Higher levels of human capital have two positive effects on output: a direct effect, because human capital is a factor of production, and an indirect effect, by raising human-capital-augmenting productivity. As a result, sub-Saharan Africa benefits, in terms of both income per capita and welfare levels and growth rates (Figure 9, panels (a) and (c)).

The policy of proportionally improving access to education across sub-Saharan Africa results in global losses. While sub-Saharan Africa gains, world GDP per capita and welfare decline following the policy change (Figure 9, panels (b) and (d)). The main reason is a composition effect: a smaller share of the world's population resides in its most productive regions.

The finding that improving access to education in sub-Saharan Africa may worsen global welfare reflects inefficiencies in the spatial allocation. By keeping a larger share of the global population in sub-Saharan Africa, the policy reduces agglomeration economies in some of the world's most productive places. This effect is magnified by weaker dynamic agglomeration effects: with smaller stocks of human capital, these productive areas innovate less, lowering the diffusion of technology to the rest of the world. Because of the inefficiencies that arise from these externalities, lowering schooling costs in sub-Saharan Africa may reduce global welfare.

Figure 8: Proportionately Lowering Human Capital Cost in Sub-Saharan Africa: Effect on Population



Note: Figure displays the population in sub-Saharan Africa (in millions) in a counterfactual where the cost of education is lowered by 49.5% across all sub-Saharan African grid cells, compared to the baseline. The time period goes from 2000 to 2200.

5.2 Proportionally Improving Access to Education in Other Regions

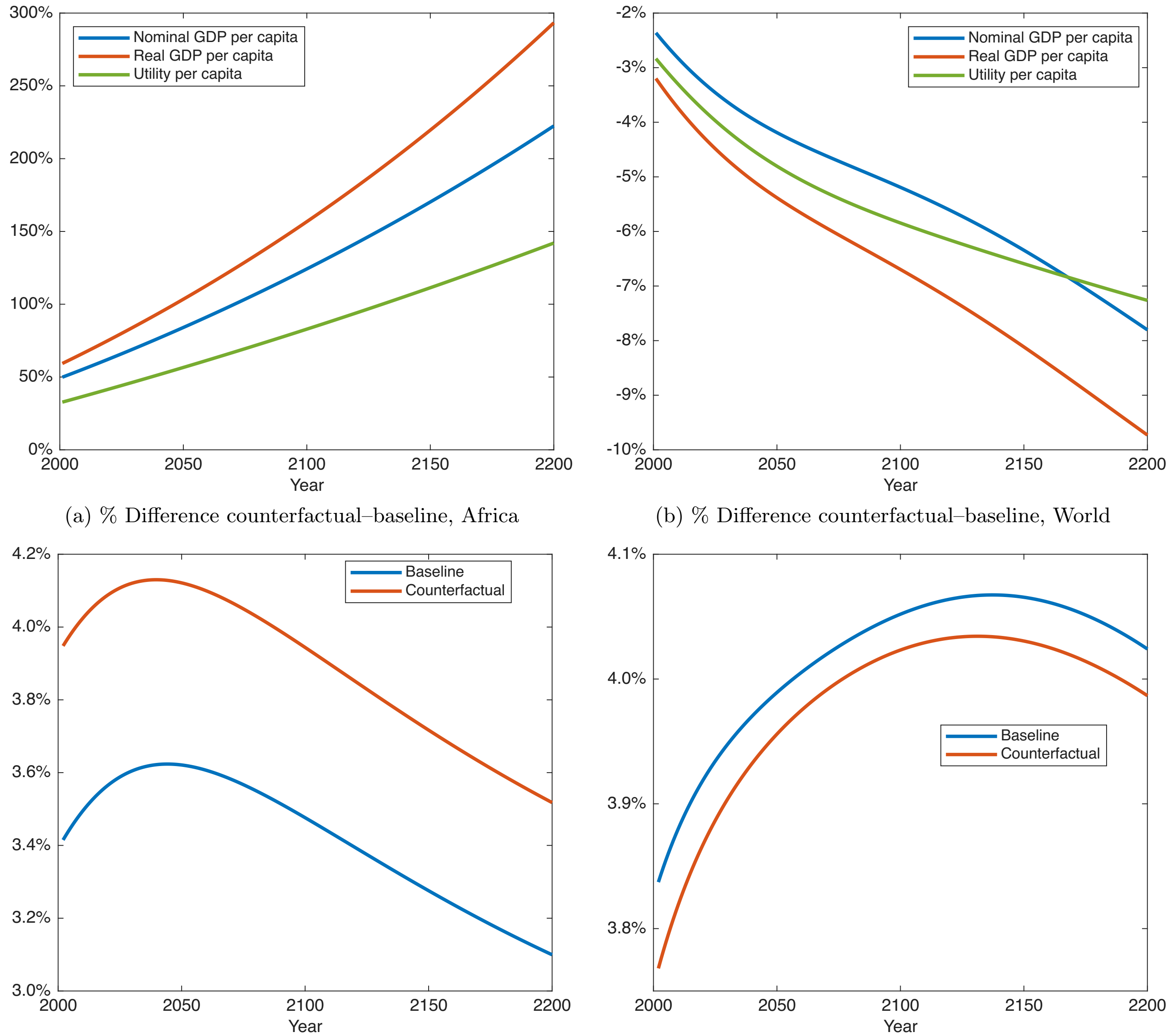
Next, we explore the counterfactual scenario of proportionally reducing the cost of human capital in other parts of the world. As before, the magnitude of the proportional drop in $\sigma(r)$ in the target region equals the population-weighted average decline that would result from setting the cost equal to the world median in that region. This corresponds to a 44.6% reduction in Central and South Asia and a 31.9% reduction in Latin America.

Proportionally lowering the cost in Central and South Asia produces effects similar to those observed in sub-Saharan Africa: locally, GDP per capita and welfare improve, but globally, outcomes worsen (Figure 10, panels (a) and (b)). By contrast, the same counterfactual policy in Latin America generates gains, both locally and globally (panels (c) and (d)).

This contrasting global impact arises from how such policies reshape the global distribution of population. While the target region consistently gains population, this does not always come at the expense of the same areas of the world. In sub-Saharan Africa and Central and South Asia, the policy disproportionately draws people away from regions with relatively strong fundamentals, resulting in global losses. In contrast, in Latin America, the reallocation comes mainly from areas with relatively weak fundamentals, generating global gains. To illustrate this, we regress log GDP per capita at $t = 1$ on fundamental amenities, fundamental productivity, human-capital-augmenting productivity, and migration costs. We take the predicted log GDP from this regression as our measure of a cell's fundamentals. Figure 11 shows that when lowering education costs in sub-Saharan Africa, the population losses in the rest of the world occur mostly in areas with stronger fundamentals than sub-Saharan Africa (panel (a)), whereas for Latin America, they are concentrated in areas with relatively weaker fundamentals (panel (b)).

Intuitively, whether the world gains or loses depends on how the fundamentals of the target region compare with those of the rest of the world. As shown in Figure 12, in sub-Saharan Africa the distribution of fundamentals lies mostly to the left of the rest of the world. By retaining a larger share of the global population in sub-Saharan

Figure 9: Proportionately Lowering Human Capital Cost in Sub-Saharan Africa: Effects on Region and World



(a) % Difference counterfactual–baseline, Africa

(b) % Difference counterfactual–baseline, World

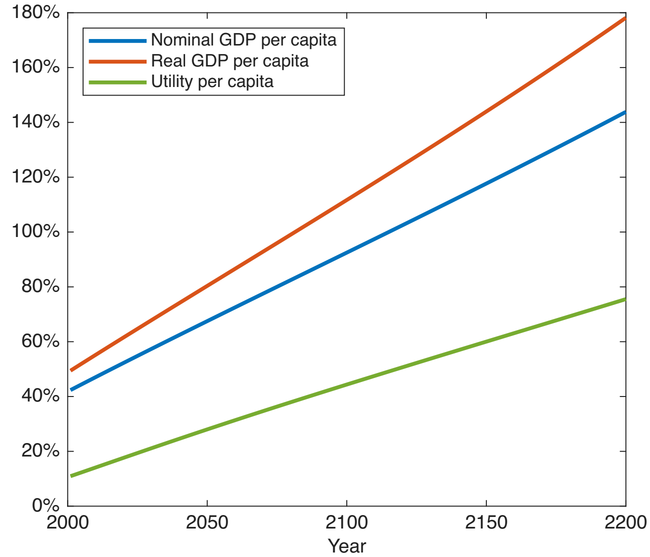
(c) Real GDP per capita growth, Africa

(d) Real GDP per capita growth, World

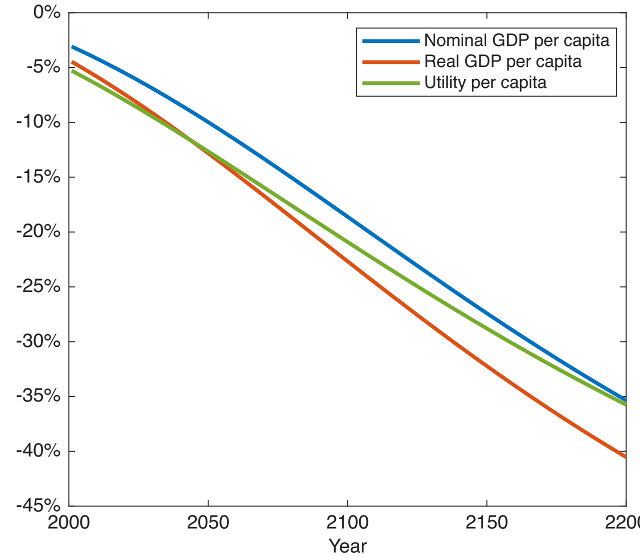
Note: Figure compares a counterfactual where the cost of education is lowered by 49.5% across all sub-Saharan African grid cells to the baseline. The top row displays the percentage difference between the counterfactual and the baseline in nominal GDP per capita, real GDP per capita and utility per capita in sub-Saharan Africa (panel (a)) and in the World as a whole (panel (b)). The bottom row displays the real GDP per capita growth rates in the counterfactual and the baseline for sub-Saharan Africa (panel (c)) and the World as a whole (panel (d)). The time period goes from 2000 to 2200.

Africa, this policy shifts people away from areas with relatively good fundamentals (panel (a)). In contrast, Latin America's fundamentals reflects those of a middle-income region, with some parts of the rest of the world better off and others worse off. A policy that keeps more people in Latin America therefore does not necessarily draw people away from areas with relatively strong fundamentals.

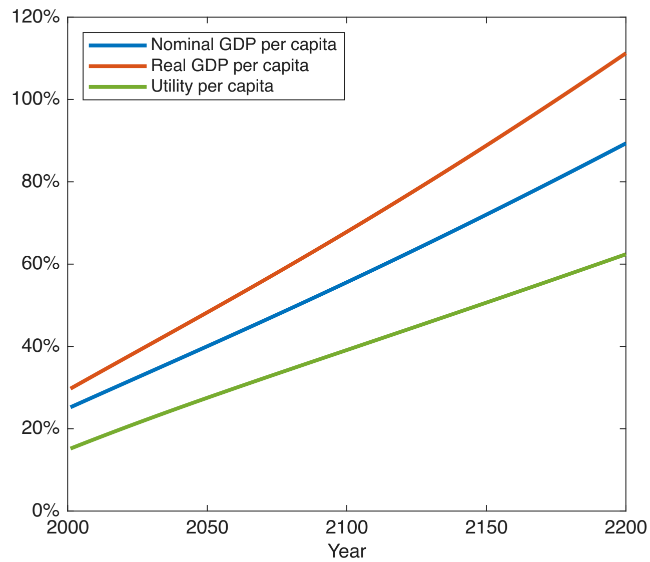
Figure 10: Proportionately Lowering Human Capital Cost in Other Regions of the World



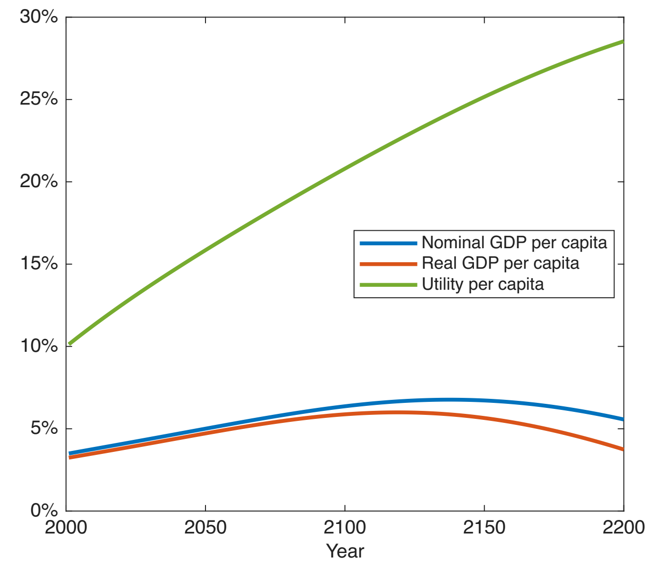
(a) % Difference counterfactual–baseline, Central and South Asia



(b) % Central-South Asia counterfactual–baseline, World



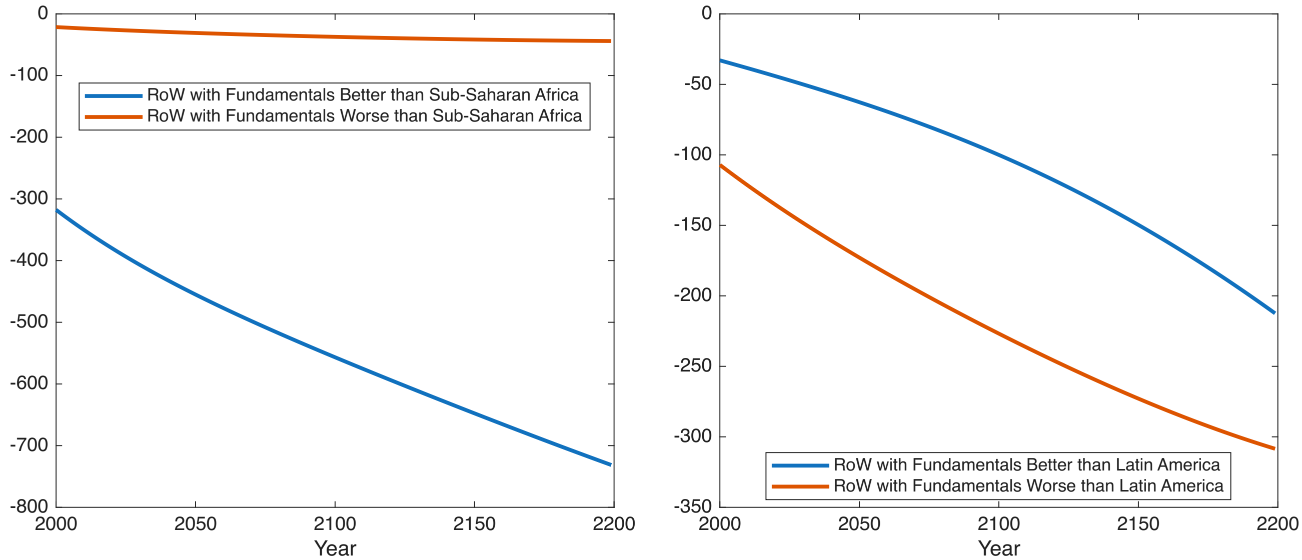
(c) % Difference counterfactual–baseline, Latin America



(d) % LatAm counterfactual–baseline, World

Note: The top row compares a counterfactual where the cost of education is lowered by 44.6% across all Central and South Asian grid cells to the baseline. The percentage difference between the counterfactual and the baseline in nominal GDP per capita, real GDP per capita, and utility per capita is displayed for Central and South Asia (panel (a)) and for the World as a whole (panel (b)). The bottom row compares a counterfactual where the cost of education is lowered by 31.9% across all Latin American grid cells to the baseline. The percentage difference between the counterfactual and the baseline in nominal GDP per capita, real GDP per capita, and utility per capita is displayed for Latin America (panel (c)) and for the World as a whole (panel (d)). The time period goes from 2000 to 2200.

Figure 11: Proportionately Lowering Human Capital Cost: Reallocation in Rest of the World

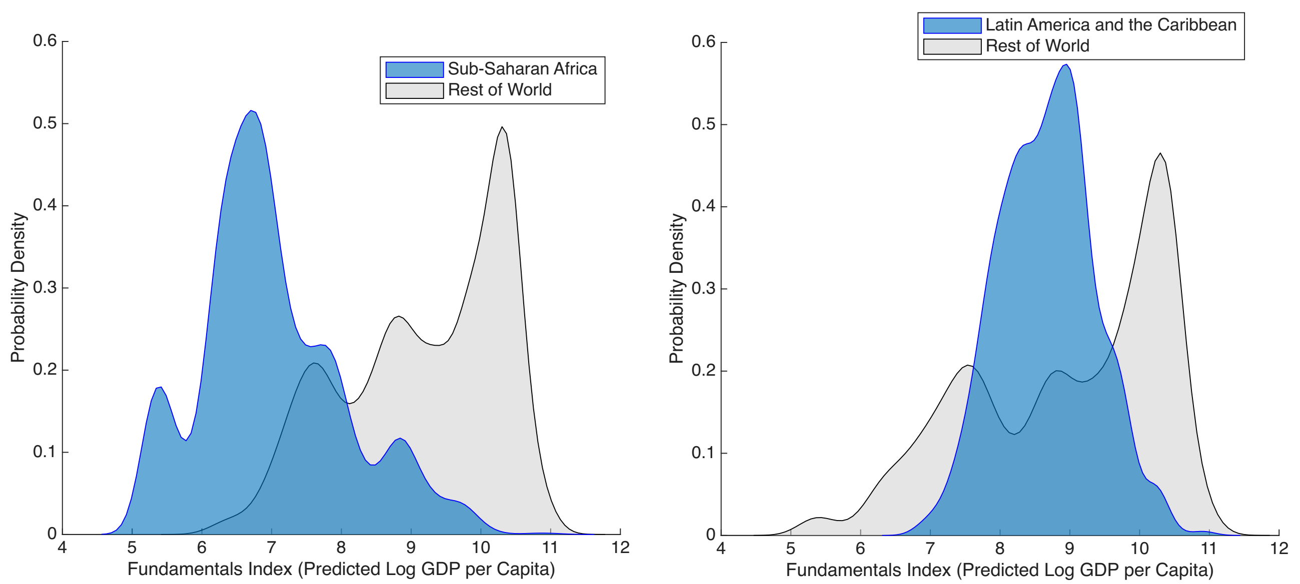


(a) Effect on population (millions) in the rest of world from lowering cost in sub-Saharan Africa

(b) Effect on population (millions) in the rest of the world from lowering cost in Latin America

Note: Panel (a) displays the effect on population (millions) in the rest of the world from proportionally lowering the cost of education by 49.5% across sub-Saharan Africa, distinguishing between cells with better and worse fundamentals than the average of sub-Saharan Africa. Panel (b) displays the same effect from a policy that proportionally lowers the cost of education by 31.9% across Latin America.

Figure 12: Distribution of Fundamentals



(a) Sub-Saharan Africa vs rest of the world

(b) Latin America vs rest of the world

Note: Panel (a) displays the distribution of fundamentals in sub-Saharan Africa vs those in the rest of the world. Panel (b) displays the distribution of fundamentals in Latin America vs those of the rest of the world.

5.3 Equality of Opportunity

As an alternative counterfactual experiment, we implement an equality of opportunity policy. Specifically, we lower the cost of acquiring human capital, $\sigma(r)$, in all sub-Saharan African grid cells to the world median, giving everyone in the region equal access to education. To facilitate comparison, the population-weighted average decline in $\sigma(r)$ matches that of the previous counterfactual, where the cost was reduced proportionally. As shown in Figure 13 panel (c), under this alternative policy sub-Saharan Africa still experiences welfare gains, but the effect on real income per capita is now more nuanced: after an initial improvement, it starts to decline relative to the baseline after about half a century.

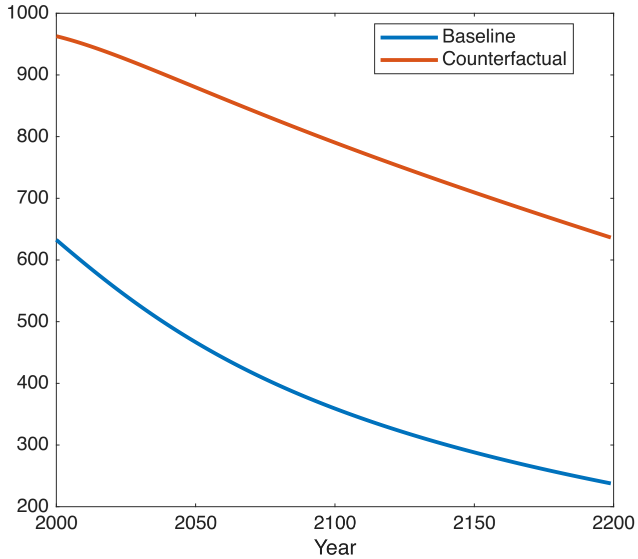
Why might real income per capita in sub-Saharan Africa fall when introducing equal access to education, when it did not when lowering costs proportionally? The answer lies in the regional reallocation of population within Africa. Equalizing access across all grid cells lowers the cost of education more in less developed cells than in more developed ones. As a result, population shifts away from cells with better fundamentals towards cells with worse fundamentals. To show this pattern, we plot the population change in sub-Saharan Africa by quintile of fundamentals. As can be seen in Figure 14, panel (a), following the change in the cost of education, locations in the bottom two quintiles experience the fastest population growth, whereas those in the highest quintile see almost no change. As a result, a rising share of the sub-Saharan African population resides in locations with worse fundamentals. This reallocation also implies a greater geographic dispersion of the population. As shown in panel (b), the Herfindahl–Hirschman Index (HHI) of population increases much more slowly in the counterfactual than in the baseline. In relative terms, this greater dispersion weakens agglomeration economies, further reducing real income per capita.

In contrast, in our first counterfactual, no such within-region population reallocation occurs. Since in that case access to education improves proportionally across all locations, there is no incentive for people to move away from the more productive areas within sub-Saharan Africa. Unlike under the equal opportunity policy, no force counteracts the local gains from lowering the cost of education. As a result, in our first counterfactual the region experiences consistent gains throughout the 200-year period.

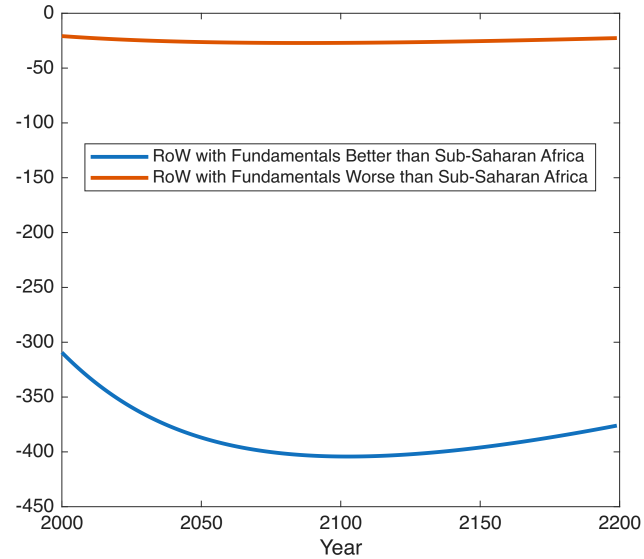
As for the effect on the world as a whole, introducing a policy of equal opportunity in sub-Saharan Africa leads to global losses (Figure 13, panel (d)). The main reason is, once again, a composition effect: a smaller share of the global population resides in its most productive regions. Compared with the first counterfactual, however, these losses diminish somewhat over time. The reason for this is two-fold. First, the population size of sub-Saharan Africa grows less over time under the equal opportunity policy, so the composition effect is weaker (compare Figure 13(a) with Figure 8). Second, outside sub-Saharan Africa, there is over time some reallocation of population towards more developed regions — a pattern absent when lowering costs proportionally (compare Figure 13(b) with Figure 11(a)).

When the same equal opportunity policy is applied in Central and South Asia, we do not find losses for the target region over a 200-year period, though the gains taper off over time (Figure 15(a)). When implemented in Latin America, the region gains consistently throughout, with the benefits growing over time (panel (b)). While the within-region reallocation towards less productive areas occurs here as well — just as in sub-Saharan Africa — this negative effect is not strong enough to offset the direct gains from improved access to schooling.

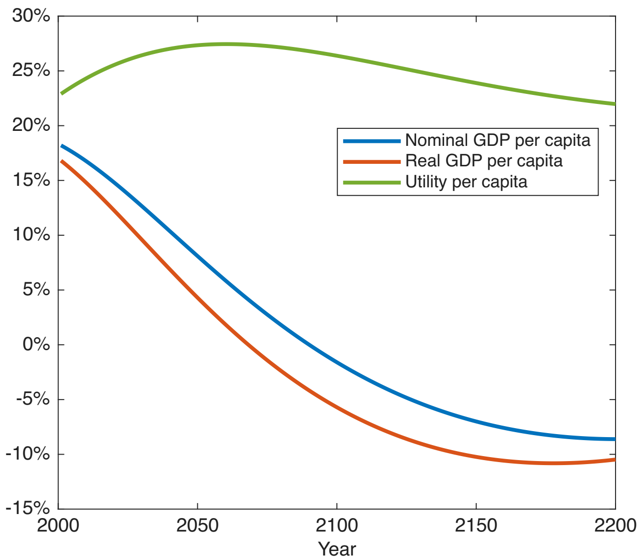
Figure 13: Setting Human Capital Cost in Sub-Saharan Africa to World Median: Effects on Region and World



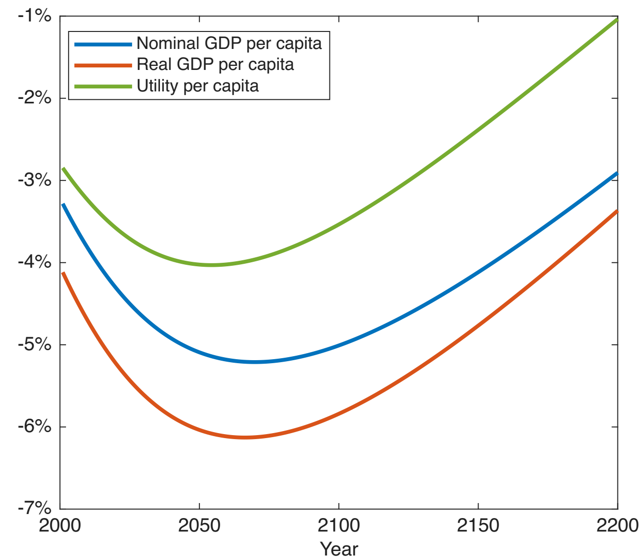
(a) Population in sub-Saharan Africa (millions)



(b) Population reallocation in rest of the world (millions)



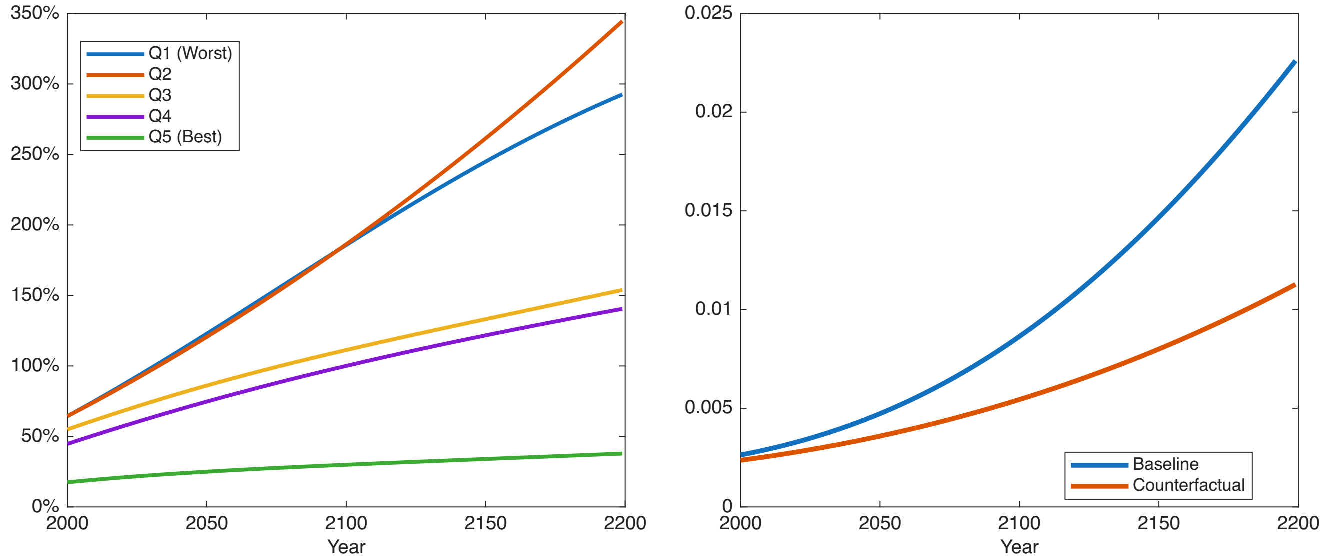
(c) % Difference counterfactual–baseline, Africa



(d) % Difference counterfactual–baseline, World

Note: Figure compares a counterfactual where the cost of education is lowered to the world median across all sub-Saharan African grid cells to the baseline. Panel (a) displays the population in sub-Saharan Africa (in millions) in the counterfactual and the baseline. Panel (b) displays the change in population in the rest of the world (in millions) in the counterfactual, comparing cells with better and worse fundamentals than sub-Saharan Africa. The bottom row displays the percentage difference between the counterfactual and the baseline in nominal GDP per capita, real GDP per capita and utility per capita in sub-Saharan Africa (panel (c)) and in the world as a whole (panel (d)). The time period goes from 2000 to 2200.

Figure 14: Setting Human Capital Cost in Sub-Saharan Africa to World Median: Within-Region Changes

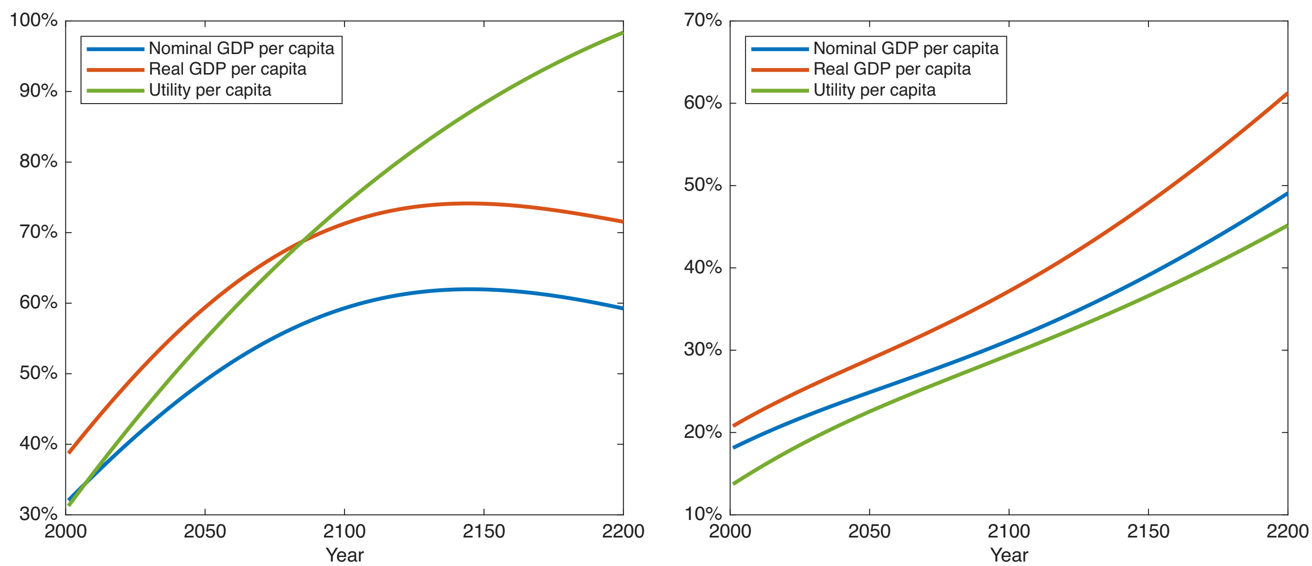


(a) Population change by fundamentals quintile in sub-Saharan Africa

(b) Population concentration (HHI) within sub-Saharan Africa

Note: Panel (a) displays the percentage population change within sub-Saharan Africa by fundamentals quintile under a counterfactual policy that lowers the cost of education to the world median across all cells of sub-Saharan Africa. Panel (b) plots the Herfindahl-Hirschman Index for sub-Saharan Africa, comparing that same counterfactual to the baseline. The time period goes from 2000 to 2200.

Figure 15: Setting Human Capital Cost in Other Regions to World Median



(a) % Difference counterfactual–baseline Central and South Asia

(b) % Difference counterfactual–baseline Latin America

Note: Panel (a) shows the effect on Central and South Asia of a counterfactual policy that sets education costs to the world median across all cells of Central and South Asia. The effect is displayed as the percentage difference between the counterfactual and the policy. Panel (b) shows the effect on Latin America of a similar policy applied to all cells of Latin America.

6 Conclusion

This paper has examined the role of human capital in shaping the economic geography of development. The cost of acquiring human capital varies across space, and so does its productivity. This affects how the supply of and the demand for human capital evolve over time and space. We embed these insights into a spatial dynamic model, with migration, trade, and endogenous human-capital-augmenting productivity growth. When simulating a quantified high-resolution version of the model for the entire globe, we find a high degree of persistence in the geography of development over the next centuries. This stands in contrast to some other spatial dynamic models that ignore human capital and predict convergence in economic development.

At the core of this persistence is the negative correlation between the cost of education and fundamentals. This relationship ensures that today's developed regions continue to have high levels of human capital in the future, keeping them ahead of the rest of the world. This mechanism is strong enough to offset other forces that might otherwise benefit some of the world's low-productivity high-density areas through market size effects and agglomeration economies. To summarize, two key drivers shape the geography of development. First, local market size, partly determined by density, incentivizes firms to innovate and grow. Second, the local stock of human capital, partly dependent on education costs, determines output and human-capital-augmenting innovation. Hence, once human capital is introduced, heterogeneity in education costs becomes central to the geography of development.

We then use the quantified model to conduct counterfactual policy experiments aimed at improving the access to education in different parts of the world. When proportionally lowering the cost of acquiring human capital in sub-Saharan Africa, we find local gains but global losses. The local gains come from higher levels of human capital that stimulate innovation. This reduces out-migration, further improving productivity through agglomeration economies. However, these local benefits come at the cost of fewer people residing in some of the world's most productive regions. This undermines global innovation and technology diffusion, and leads to global losses as the spatial equilibrium allocation is inefficient. When implementing a similar policy in other low-income regions, such as Central and South Asia, the world also loses, because the local population gains come at the expense of regions with better fundamentals. By contrast, introducing the same policy in middle-income regions, such as Latin America, generates global gains, as the local increase in population comes at the expense of regions with worse fundamentals.

In another policy counterfactual experiment, we analyze the impact of an equal opportunity policy that lowers education costs across sub-Saharan Africa to a common level. While this also generates short-run local gains, the sub-Saharan African economy experiences long-run losses due to population shifting within the region toward less productive areas. This within-region spatial reallocation offsets the benefits of cheaper access to education. A similar equal opportunity policy, when implemented in Central and South Asia or Latin America, does not generate these unintended negative consequences on the local economy.

The analysis of the relationship between human capital and spatial growth leaves several open questions. First, our model treats local education costs as exogenous. One could, instead, endogenize how these costs depend on the degree of development. In such a scenario, the model may generate less persistence. Second, we abstract from spatial differences in the quality of education, an issue that has been shown to be important (Schoellman, 2012). That said, the high correlation between the local cost of education as identified by the model and observed educational spending by households suggests that our approach captures the essential spatial differences in the cost of acquiring human capital. Third, in our framework all individuals in a given location choose the same level of human capital. Allowing for within-location heterogeneity between individuals would be a valuable extension, as it would generate differential migration flows by human capital, as in Burzynski et al. (2020). Fourth, we simplify the forward-looking capital accumulation decision to a sequence of static decisions. As discussed in Desmet and Parro (2025), spatial endogenous growth models face a computational tradeoff between incorporating anticipation effects and maintaining high spatial resolution. In this paper, we chose to prioritize spatial heterogeneity, which we have shown to be crucial for evaluating the impact of certain educational policies, such as equalizing opportunity across locations.

References

Acemoglu, D. and D. Autor (2011). Skills, Tasks and Technologies: Implications for Employment and Earnings. In D. Card and O. Ashenfelter (Eds.), *Handbook of Labor Economics*, Volume 4, pp. 1043–1171. Elsevier.

Aghion, P., U. Akcigit, A. Hyytinen, and O. Toivanen (2024). A Year Older, a Year Wiser (and Farther from Frontier): Invention Rents and Human Capital Depreciation. *Review of Economics and Statistics 106*(4), 974–982.

Ben-Porath, Y. (1967). The Production of Human Capital and the Life Cycle of Earnings. *Journal of political economy 75*(4, Part 1), 352–365.

Burzynski, M., C. Deuster, and F. Docquier (2020). Geography of Skills and Global Inequality. *Journal of Development Economics 142*, 102333.

Caselli, F. and W. J. Coleman (2006). The World Technology Frontier. *American Economic Review 96*(3), 499–522.

Ciccone, A. and E. Papaioannou (2009). Human Capital, the Structure of Production, and Growth. *The Review of Economics and Statistics 91*(1), 66–82.

CIESIN (2018, Revision 11). Gridded Population of the World, Version 4 (GPWv4): Basic Demographic Characteristics.

Desmet, K., D. K. Nagy, and E. Rossi-Hansberg (2018). The Geography of Development. *Journal of Political Economy 126*(3), 903 – 983.

Desmet, K. and F. Parro (2025). Spatial Dynamics. In D. Donaldson and S. J. Redding (Eds.), *Handbook of Regional and Urban Economics*, Volume 6 of *Handbook of Regional and Urban Economics*, pp. 225–286. Elsevier.

Eckert, F. and T. Kleineberg (2024). The Geography of Opportunity: Education, Work, and Intergenerational Mobility Across US Counties. unpublished manuscript.

Hall, R. E. and C. I. Jones (1999). Why Do Some Countries Produce So Much More Output per Worker Than Others? *The Quarterly Journal of Economics 114*(1), 83–116.

Havranek, T., Z. Irsova, L. Laslopova, and O. Zeynalova (2024). Publication and Attenuation Biases in Measuring Skill Substitution. *Review of Economics and Statistics 106*(5), 1187–1200.

Jerzmanowski, M. and R. Tamura (2023). Aggregate Elasticity of Substitution between Skills: Estimates from a Macroeconomic Approach. *Macroeconomic Dynamics 27*(6), 1597–1627.

Jones, B. F. (2014). The Human Capital Stock: A Generalized Approach. *American Economic Review 104*(11), 3752–3777.

Lucas, R. E. (2015). Human Capital and Growth. *American Economic Review 105*(5), 85–88.

Montenegro, C. E. and H. A. Patrinos (2014). Comparable Estimates of Returns to Schooling Around the World. World Bank Policy Research Working Paper 7020.

Nordhaus, W., Q. Azam, D. Corderi Novoa, K. Hood, N. Victor, M. Mohammed, A. Miltner, and J. Weiss (2006). The G-Econ Database on Gridded Output: Methods and Data. Working Paper, Yale University.

Schoellman, T. (2012). Education Quality and Development Accounting. *The Review of Economic Studies 79*(1), 388–417.

Smits, J. and I. Permanyer (2019). The Subnational Human Development Database. *Scientific Data 6*(1), 1–15.

Zabreyko, P. P., A. I. Koshelev, M. A. Rasnosel'skii, S. G. Mikhlin, L. S. Rakovshchik, and V. Y. Stet'senko (1975). *Integral Equations: A Reference Text*. Leyden: Noordhoff.

A Proofs of Propositions

A.1 Proof of Proposition 2

Proposition 2. *For any $A_t(r) > 0$ and $0 < \sigma(r) < \mu$, equation (21) has exactly one positive solution. Moreover, the solution satisfies $0 < h_t(r) < \frac{\mu}{\sigma(r)}$.*

Proof. Denote the left-hand side of equation (21) by $F(h_t(r))$. It is clear that $F(\cdot)$ is continuous and strictly decreasing over the interval $(0, \mu/\sigma(r))$. Moreover, given that $0 < \sigma(r) < \mu$, $F(\cdot)$ takes on any value between 0 and $+\infty$: as $h_t(r) \to \mu/\sigma(r)$, $F(h_t(r)) \to 0$, while as $h_t(r) \to 0$, $F(h_t(r)) \to \infty$. This implies that equation (21) has exactly one solution over the interval $(0, \mu/\sigma(r))$. □

A.2 Proof of Proposition 3

Proposition 3. *Given $h_t(r)$, equations (8), (16), (17), (19), (30) and (31) pin down the remaining period-t endogenous variables as a function of $A_t(r)$ and fundamentals.*

Proof. Combining (17) and (19) yields

$$W_t(r) = \mu^{-\frac{1}{\rho-1}} \gamma_1^{-\frac{\gamma_1}{\rho}} (1-\gamma_1)^{-\frac{1-\gamma_1}{\rho}} A_t(r)^{-\frac{1}{\rho}} \sigma(r)^{\frac{1}{\rho-1}} h_t(r)^{\frac{1}{\rho(\rho-1)}} w_t^L(r). \tag{A1}$$

Plugging this into equations (30) and (31) implies

$$\begin{aligned} T_t(r) =& \tilde{\kappa}_1 \frac{\bar{a}(r)}{u_t(r)} A_t(r)^{1-\mu-\frac{1}{\rho}-\frac{1-\rho}{\rho}\frac{\xi}{\xi-1}} \sigma(r)^{\frac{1}{\rho-1}-(1-\mu)\frac{\rho}{\rho-1}-\frac{\xi}{\xi-1}} \\ & h_t(r)^{\frac{1}{\rho(\rho-1)}\left[(1-\rho)\frac{\xi}{\xi-1}+1-(1-\mu)\rho\right]} w_t^L(r)^{\frac{2\xi-1}{\xi-1}} N(r)^{\frac{1}{\xi-1}} L_t(r)^{\frac{1}{\xi-1}+1-\mu-\lambda} \end{aligned} \tag{A2}$$

and

$$\begin{aligned} & \left[\frac{\bar{a}(r)}{u_t(r)}\right]^{1-\xi} A_t(r)^{\frac{1-\rho}{\rho}(\xi-1)} \sigma(r)^{\xi-1} h_t(r)^{\frac{1}{\rho}(\xi-1)} w_t^L(r)^{1-\xi} L_t(r)^{\lambda(\xi-1)} \\ =& \tilde{\kappa}_2 \int_R \left[\frac{\bar{a}(r')}{u_t(r')}\right]^{\xi-1} A_t(r')^{\frac{\rho-1}{\rho}\xi} \sigma(r')^{-\xi} h_t(r')^{-\frac{1}{\rho}\xi} w_t^L(r')^{\xi} N(r') L_t(r')^{1-\lambda(\xi-1)} \varsigma(r',r)^{1-\xi} dr' \end{aligned} \tag{A3}$$

where

$$\tilde{\kappa}_1 = \mu^{\frac{\xi}{\xi-1}-\frac{1}{\rho-1}+(1-\mu)\frac{\rho}{\rho-1}} \gamma_1^{\gamma_1\frac{\rho-1}{\rho}\frac{\xi}{\xi-1}-\frac{\gamma_1}{\rho}+\gamma_1(1-\mu)} (1-\gamma_1)^{(1-\gamma_1)\frac{\rho-1}{\rho}\frac{\xi}{\xi-1}-\frac{1-\gamma_1}{\rho}+(1-\gamma_1)(1-\mu)}$$

and

$$\tilde{\kappa}_2 = \mu^{2\xi-1} \gamma_1^{\gamma_1\frac{\rho-1}{\rho}(2\xi-1)} (1-\gamma_1)^{(1-\gamma_1)\frac{\rho-1}{\rho}(2\xi-1)}.$$

Expressing $w_t^L(r)$ from (A2) and plugging it into (A3) yields

$$\begin{aligned} & \left[\frac{\bar{a}(r)}{u_t(r)}\right]^{-\frac{\xi(\xi-1)}{2\xi-1}} A_t(r)^{-\mu\frac{(\xi-1)^2}{2\xi-1}} \sigma(r)^{\mu\frac{\rho}{\rho-1}\frac{(\xi-1)^2}{2\xi-1}} h_t(r)^{\mu\frac{1}{\rho-1}\frac{(\xi-1)^2}{2\xi-1}} \tau(r)^{-\frac{(\xi-1)^2}{2\xi-1}} N(r)^{\frac{\xi-1}{2\xi-1}} \cdot \\ & L_t(r)^{\lambda(\xi-1)-\frac{\xi-1}{2\xi-1}[(\lambda+\alpha-(1-\mu))(\xi-1)-1]} = \tilde{\kappa}_3 \int_R \left[\frac{\bar{a}(r')}{u_t(r')}\right]^{\frac{(\xi-1)^2}{2\xi-1}} A_t(r')^{\mu\frac{\xi(\xi-1)}{2\xi-1}} \sigma(r')^{-\mu\frac{\rho}{\rho-1}\frac{\xi(\xi-1)}{2\xi-1}} \cdot \\ & h_t(r')^{-\mu\frac{1}{\rho-1}\frac{\xi(\xi-1)}{2\xi-1}} \tau(r')^{\frac{\xi(\xi-1)}{2\xi-1}} N(r')^{\frac{\xi-1}{2\xi-1}} L_t(r')^{1-\lambda(\xi-1)+\frac{\xi}{2\xi-1}[(\lambda+\alpha-(1-\mu))(\xi-1)-1]} \varsigma(r',r)^{1-\xi} dr' \end{aligned} \tag{A4}$$

where $\tilde{\kappa}_3 = \tilde{\kappa}_1^{1-\xi}\tilde{\kappa}_2$, and we have used $T_t(r) = \tau(r) L_t(r)^{\alpha}$. Finally, combining (A4) with equations (8) and (16) yields

$$\begin{aligned} & B_{1t}(r)\,\hat{u}_t(r)^{\frac{1}{\Omega}\left[\lambda(\xi-1)-\frac{\xi-1}{2\xi-1}[(\lambda+\alpha-(1-\mu))(\xi-1)-1]\right]+\frac{\xi(\xi-1)}{2\xi-1}} = \\ & \tilde{\kappa}_3 \int_R \hat{u}_t(r')^{\frac{1}{\Omega}\left[1-\lambda(\xi-1)+\frac{\xi}{2\xi-1}[(\lambda+\alpha-(1-\mu))(\xi-1)-1]\right]-\frac{(\xi-1)^2}{2\xi-1}} B_{2t}(r')\,\varsigma(r',r)^{1-\xi}\,dr' \end{aligned} \tag{A5}$$

where

$$\begin{aligned} B_{1t}(r) =& \bar{a}(r)^{-\frac{\xi(\xi-1)}{2\xi-1}} A_t(r)^{-\mu\frac{(\xi-1)^2}{2\xi-1}} \sigma(r)^{\mu\frac{\rho}{\rho-1}\frac{(\xi-1)^2}{2\xi-1}} h_t(r)^{\mu\frac{1}{\rho-1}\frac{(\xi-1)^2}{2\xi-1}} \tau(r)^{-\frac{(\xi-1)^2}{2\xi-1}} \cdot \\ & N(r)^{[\lambda+\alpha-(1-\mu)]\frac{(\xi-1)^2}{2\xi-1}-\lambda(\xi-1)} m_2(r)^{-\frac{1}{\Omega}\left[\lambda(\xi-1)-\frac{\xi-1}{2\xi-1}[(\lambda+\alpha-(1-\mu))(\xi-1)-1]\right]} \cdot \\ & e^{-\frac{\sigma(r)h_t(r)}{\Omega}\left[\lambda(\xi-1)-\frac{\xi-1}{2\xi-1}[(\lambda+\alpha-(1-\mu))(\xi-1)-1]\right]} \end{aligned}$$

$$\begin{aligned} B_{2t}(r') =& \bar{a}(r')^{\frac{(\xi-1)^2}{2\xi-1}} A_t(r')^{\mu\frac{\xi(\xi-1)}{2\xi-1}} \sigma(r')^{-\mu\frac{\rho}{\rho-1}\frac{\xi(\xi-1)}{2\xi-1}} h_t(r')^{-\mu\frac{1}{\rho-1}\frac{\xi(\xi-1)}{2\xi-1}} \tau(r')^{\frac{\xi(\xi-1)}{2\xi-1}} \cdot \\ & N(r')^{\frac{\xi-1}{2\xi-1}-1+\lambda(\xi-1)-\frac{\xi}{2\xi-1}[(\lambda+\alpha-(1-\mu))(\xi-1)-1]} \cdot \\ & m_2(r')^{-\frac{1}{\Omega}\left[1-\lambda(\xi-1)+\frac{\xi}{2\xi-1}[(\lambda+\alpha-(1-\mu))(\xi-1)-1]\right]} e^{-\frac{\sigma(r')h_t(r')}{\Omega}\left[1-\lambda(\xi-1)+\frac{\xi}{2\xi-1}[(\lambda+\alpha-(1-\mu))(\xi-1)-1]\right]} \end{aligned}$$

and

$$\hat{u}_t(r) = u_t(r)\left[\frac{L}{\int_R u_t(r')^{1/\Omega} m_2(r')^{-1/\Omega} e^{-\sigma(r')\frac{h_t(r')}{\Omega}} dr'}\right]^{\Omega\left[1-\frac{1}{\frac{1}{\Omega}(\lambda+(1-\mu)-\alpha)+1}\right]}. \tag{A6}$$

Equations (A5) and (A6) can be solved for $u_t(r)$ by following the same procedure as in Desmet et al. (2018). Next, $L_t(r)$ can be obtained from (8) and (16); $w_t^L(r)$ from (A2); $W_t(r)$ from (A1); and $w_t^H(r)$ from (19). □

A.3 Proof of Proposition 4

Proposition 4. *An equilibrium exists and is unique if*

$$\alpha < \lambda + 1 - \mu + \Omega$$

and the equilibrium can be solved by iteration.

Proof. The proof follows the same steps as the proof of Lemma 3 in Desmet et al. (2018). In particular, it follows from Zabreyko et al. (1975) that the solution to equation (A5) exists, is unique, and can be solved by iterating on the equation if the ratio of exponents on $\hat{u}_t(r)$ on the right-hand side and on the left-hand side is below one in absolute value. This means that the condition

$$\left|\frac{\frac{1}{\Omega}\left[\lambda(\xi-1)-\frac{\xi-1}{2\xi-1}\left[(\lambda+\alpha-(1-\mu))(\xi-1)-1\right]\right]+\frac{\xi(\xi-1)}{2\xi-1}}{\frac{1}{\Omega}\left[1-\lambda(\xi-1)+\frac{\xi}{2\xi-1}\left[(\lambda+\alpha-(1-\mu))(\xi-1)-1\right]\right]-\frac{(\xi-1)^2}{2\xi-1}}\right| < 1$$

guarantees existence, uniqueness, and solvability by iteration. Rearranging and simplifying this inequality condition yields

$$\alpha < \lambda + 1 - \mu + \Omega.$$

□

A.4 Proof of Proposition 5

Proposition 5. *Using data on land $N(r)$, population density $L_0(r)$, nominal income per capita $\mu^{-1}W_0(r)^{1-\rho}\, w_0^L(r)^{\rho}$, human capital costs $\sigma(r)$, and human capital per capita $h_0(r)$, at every location r, the equilibrium conditions uniquely identify $\frac{\bar{a}(\cdot)}{u_0(\cdot)}$, $\tau(\cdot)$ and $A_0(\cdot)$.*

Proof. Rearrange equation (21) to express $A_0(r)$ as a function of $h_0(r)$ and $\sigma(r)$:

$$A_0(r) = \gamma_1^{-\gamma_1}(1-\gamma_1)^{-(1-\gamma_1)}\left[\frac{h_0(r)^{\frac{1}{\rho}}}{\frac{\mu}{\sigma(r)} - h_0(r)}\right]^{\frac{\rho}{\rho-1}}$$

Next, rearrange equation (17) to express $w_0^H(r)$ as a function of income and $\sigma(r)$:

$$w_0^H(r) = \sigma(r)\,\mu^{-1}W_0(r)^{1-\rho}\, w_0^L(r)^{\rho}$$

Once we know $A_0(r)$ and $w_0^H(r)$, we can express $w_0^L(r)$ from equation (19) as

$$w_0^L(r) = \gamma_1^{\gamma_1\frac{\rho-1}{\rho}}(1-\gamma_1)^{(1-\gamma_1)\frac{\rho-1}{\rho}}\, A_0(r)^{\frac{\rho-1}{\rho}}\, h_0(r)^{-\frac{1}{\rho}}\, w_0^H(r)$$

and then equation (13) gives us $W_0(r)$. Plugging $W_0(r)$ and $w_0^L(r)$ into equations (30) and (31), we can solve for $T_0(r)$ and $\frac{\bar{a}(r)}{u_0(r)}$ using the standard inversion procedure in Desmet et al. (2018). Finally, fundamental productivity can be obtained as

$$\tau(r) = T_0(r)\, L_0(r)^{-\alpha}.$$

□

A.5 Proof of Proposition 6

Proposition 6. *Using data on $L_1(r)$, the equilibrium conditions identify moving costs $m_2(r)$ up to a scale.*

Proof. Write equation (8) for period 1 after plugging in (16). Solving out the resulting equation for $u_1(r)$, and plugging it into equation (A4), yield

$$\begin{aligned}
&\left[\frac{\bar{a}(r)}{\hat{m}_2(r)}\right]^{-\frac{\xi(\xi-1)}{2\xi-1}} A_1(r)^{-\mu\frac{(\xi-1)^2}{2\xi-1}}\,\sigma(r)^{\mu\frac{\rho}{\rho-1}\frac{(\xi-1)^2}{2\xi-1}}\, h_1(r)^{\mu\frac{1}{\rho-1}\frac{(\xi-1)^2}{2\xi-1}}\, e^{\sigma(r)h_1(r)\frac{\xi(\xi-1)}{2\xi-1}}\,\tau(r)^{-\frac{(\xi-1)^2}{2\xi-1}}\cdot \\
&N(r)^{\frac{\xi-1}{2\xi-1}(1+\xi\Omega)}\, L_1(r)^{\lambda(\xi-1)-\frac{\xi-1}{2\xi-1}[(\lambda+\alpha-(1-\mu))(\xi-1)-1]+\frac{\xi(\xi-1)}{2\xi-1}\Omega} = \tilde{\kappa}_3\int_R\left[\frac{\bar{a}(r')}{\hat{m}_2(r')}\right]^{\frac{(\xi-1)^2}{2\xi-1}}\cdot \qquad \text{(A7)}\\
&A_t(r')^{\mu\frac{\xi(\xi-1)}{2\xi-1}}\,\sigma(r')^{-\mu\frac{\rho}{\rho-1}\frac{\xi(\xi-1)}{2\xi-1}}\, h_t(r')^{-\mu\frac{1}{\rho-1}\frac{\xi(\xi-1)}{2\xi-1}}\, e^{-\frac{(\xi-1)^2}{2\xi-1}\sigma(r')h_1(r')}\tau(r')^{\frac{\xi(\xi-1)}{2\xi-1}}\cdot \\
&N(r')^{\frac{\xi-1}{2\xi-1}-\Omega\frac{(\xi-1)^2}{2\xi-1}}\, L_1(r')^{1-\lambda(\xi-1)+\frac{\xi}{2\xi-1}[(\lambda+\alpha-(1-\mu))(\xi-1)-1]-\frac{(\xi-1)^2}{2\xi-1}\Omega}\,\varsigma(r',r)^{1-\xi}\, dr'
\end{aligned}$$

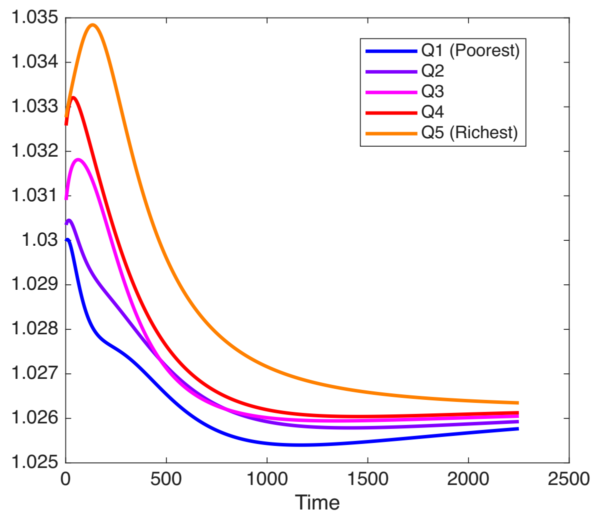
where

$$\hat{m}_2(r) = \frac{m_2(r)}{L^{\Omega}\left[\int_R u_t(r')^{1/\Omega} m_2(r')^{-1/\Omega} e^{-\sigma(r')\frac{h_t(r')}{\Omega}} dr'\right]^{-\Omega}}. \tag{A8}$$

Solving equation (A7) identifies $\hat{m}_2(r)$ for each location r. Per equation (A8), this is identical to identifying moving costs up to a scale, according to the same argument as in Desmet et al. (2018). □

B Additional Figures

Figure B1: GDP per Capita Growth By Quintile: Transition to Balanced Growth



Note: Figure displays the GDP per capita growth rates by GDP per capita quintiles over a 2500-year period, showing the protracted transition to a balanced growth path.