I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11804

# Geography of Skills and Global Inequality

Michal Burzynski
Christoph Deuster
Frédéric Docquier

SEPTEMBER 2018

DISCUSSION PAPER SERIES

IZA DP No. 11804

# Geography of Skills and Global Inequality

**Michal Burzynski**
*CREA*

**Christoph Deuster**
*IRES and Universidade Nova de Lisboa*

**Frédéric Docquier**
*FNRS, IRES, Université Catholique de Louvain and IZA*

SEPTEMBER 2018



**IZA – Institute of Labor Economics**

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

# ABSTRACT

## Geography of Skills and Global Inequality*

This paper analyzes the factors underlying the evolution of the worldwide distribution of skills and their implications for global inequality. We develop and parameterize a two-sector, two-class, world economy model that endogenizes education and mobility decisions, population growth, and income disparities across and within countries. First, our static experiments reveal that the geography of skills matters for global inequality. Low access to education and sectoral misallocation of skills substantially impact income in poor countries. Second, we produce unified projections of population and income for the 21st century. Assuming the continuation of recent education and migration policies, we predict stable disparities in the world distribution of skills, slow-growing urbanization in developing countries and a rebound in income inequality. These prospects are sensitive to future education costs and to internal mobility frictions, which suggests that policies targeting access to all levels of education and sustainable urban development are vital to reduce demographic pressures and global inequality in the long term.

**JEL Classification:** E24, J24, O15

**Keywords:** human capital, migration, urbanization, growth, inequality

**Corresponding author:**
Frédéric Docquier
IRES, Université Catholique de Louvain
Office D.232
3, Place Montesquieu
B-1348 Louvain-La-Neuve
Belgium
E-mail: frederic.docquier@uclouvain.be

* This paper benefited from helpful comments from two anonymous referees. We also thank the participants of the First NOVAFRICA Workshop on Migration and Development (Universidade Nova de Lisboa, July 2016), participants of the OLG Days (University of Luxembourg, December 2016), seminar participants at the University of Western Australia (February 2017), participants of the CSAE Conference 2017: Economic Development in Africa (University of Oxford, March 2017), participants of the International Conference on Migration and Welfare (Sapienza Universit a di Roma, May 2017), and seminar participants at the University of Paris 1 Panthéon-Sorbonne (May 2017) for their suggestions.

# 1 Introduction

It is commonly accepted that human capital acts as a proximate cause of development. Recent studies show that highly educated workers, namely, those who have completed a tertiary/college education, exhibit the highest productivity levels, generate labor market complementarities with the less educated, and are instrumental in supporting democratization and in facilitating innovation and technology diffusion when knowledge becomes *economically useful*.[1] However, the factors governing the geography of skills, its long-term developments, and its interaction with the world distribution of income are quantitatively uncertain.

In this paper, we *quantitatively analyze the root drivers underlying the long-term trend in the worldwide distribution of skills* (i.e., domestic access to education, sector allocation of workers, and international migration) and *highlight the implications of these root drivers for economic convergence and global inequality.* To do so, we develop a two-sector, two-class, world economy model that endogenizes education and labor mobility decisions, population growth, and income disparities across countries and across regions/sectors. In our framework, each country has two sectors/regions (urban and rural or equivalently, nonagriculture and agriculture), which are populated by two types of adult workers (those who have completed a college education and the less educated) and by their offspring. Production and income depend on the size and structure of the domestic labor force. We parameterize the model to match the current structure of the world economy and the ongoing socio-demographic trends. We then carry out a set of static and dynamic numerical experiments. We first use the model to quantify the fraction of contemporaneous income inequality that is explained by the geographic allocation of skills. In particular, we shed light on the global inequality implications of disparities in education policies, for the allocation of labor across sectors and for international migration. We find that the heterogeneity with respect to the overall supply of tertiary educated workers and to their allocation across sectors matter. We then use dynamic simulations for the years 2010–2100 to gain an understanding of the main drivers of the geography of skills and of its interaction with global inequality. Again, we find that future global inequality is sensitive to future education costs and to internal mobility frictions. On the contrary, current and future income disparities are much less sensitive to international migration policies. We also assess the robustness of our results to the technological and preference assumptions of the model.

Figure 1 illustrates the importance of the subject matter. In many countries and regions, college graduates form a minority. Although the worldwide average proportion of college graduates increased from only 2.4% in 1970 to 8.8% in 2010, this share is currently smaller than 1% in fifteen developing countries, such as Niger, Malawi, Zambia, Zimbabwe, and Tanzania (Barro and Lee, 2013). Using our human capital estimates (see Section 4 below), Figure 1a shows the evolution of human capital inequality in ten-year intervals from 1970 to 2010. We use the Theil index of inequality and investigate its between-country compo-

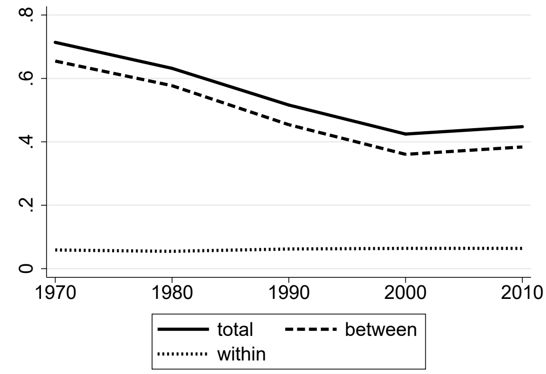
[1] This was the case during the Industrial Revolution (Mokyr, 2005; Squicciarini and Voigtländer, 2015) and it is still relevant in the modern world: see Castelló-Climent and Mukhopadhyay (2013), Jones (2014), Kerr et al. (2016) on productivity growth, or Castelló-Climent (2008), Bobba and Coviello (2007), Murtin and Wacziarg (2014) on democratization.

nent (capturing differences in the country average proportion of college graduates) and the within-country component (capturing differences between rural and urban regions). Human capital disparities are predominantly explained by the between-country component (as illustrated in Figure 1c). This means that between-country disparities are much greater than the within-country ones. Since 1970, the number of skilled workers has grown faster in poor countries. Hence, the Theil index has decreased, reflecting unconditional convergence in the share of college graduates (with a speed of approximately 0.7% per year). However, this process stalled after 2000, and large differences persist between the tails of the distribution. The latter is illustrated in Figure 1b, which depicts the density of the shares of college-educated workers in the year 2010 for a sample of 179 countries and 358 regions (i.e., rural and urban regions of the 179 countries). Figure 1d shows that the ratio of human capital between agriculture and nonagriculture reaches the lowest values for the developing countries. Hence, in poor countries, the share of college graduates is remarkably low in the rural areas (often smaller than 4%), in which a large fraction of the population lives.

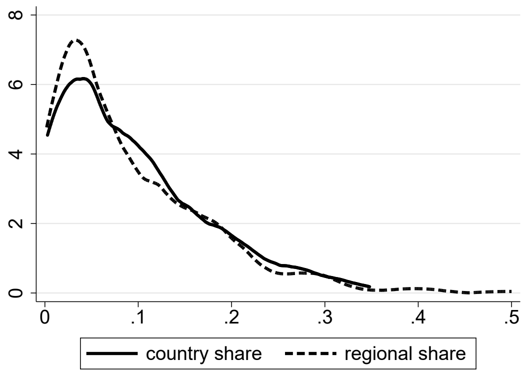
We study the drivers and implications of these geographic disparities in the world distribution of skills. The accumulation of human capital is clearly endogenous: higher-education investments are costly; returns to schooling depend on production technologies and labor market characteristics; and workers are mobile across nations and regions. To study interdependencies between the accumulation of skills and global income inequality, our model endogenizes the formation of human capital and the mobility decisions of workers. Adults decide how much to consume, how many of their children will be provided with higher education, and where to live. Internal and international migration decisions depend on geographic disparities in income and on moving costs. Accounting for international labor mobility helps to identify the effect of skill-biased migration flows on human capital and income disparities. Distinguishing between urban and rural regions allows us to model the differential in the access to education across regions (as in Lucas, 2009) and helps us to quantify the role of internal mobility frictions (as in Rodrik, 2013). The model is stylized and omits several features of the real world.[2] However, it does account for long-run interactions between human capital accumulation, migration and economic growth. Our quantitative theory is helpful for investigating how the geography of skills affects economic development and for identifying the key factors governing future demographic pressures and global inequality.

We first run static numerical experiments and use the technological block of the model to quantify the fraction of contemporaneous inequality that is explained by disparities in the share of college-educated workers. We show that the geography of skills matters for development, regardless of the size of technological externalities. In the absence of technological externality, transposing the US full educational structure (i.e., the US national share of college graduates and its allocation by sector/region) increases income per workers by a factor of 2.5 in the poorest countries (i.e., the bottom quartile of the income distribution). This is very much in line with Jones (2014); we obtain greater effects because in our two-sector model, transposing the US educational structure implies increasing the

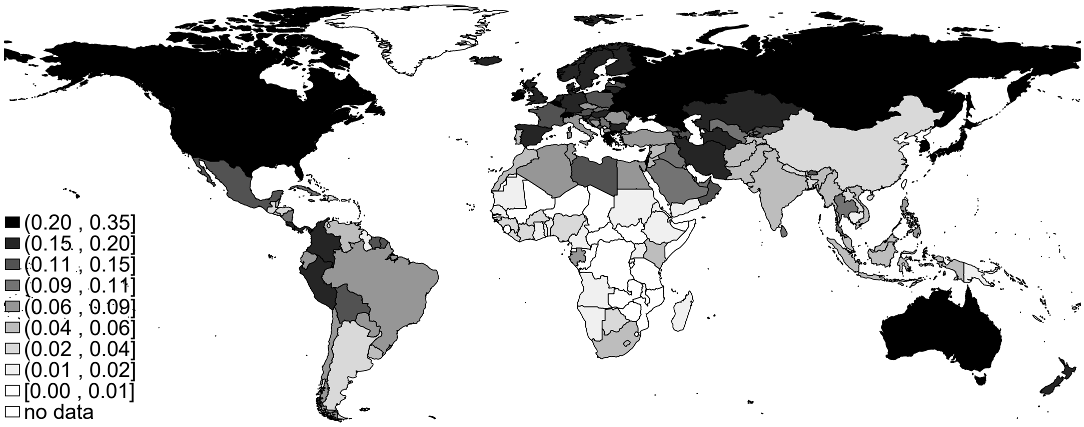
[2]The model does not account for all demographic variables (such as mortality or aging) and economic variables (such as trade, unemployment, or redistribution).



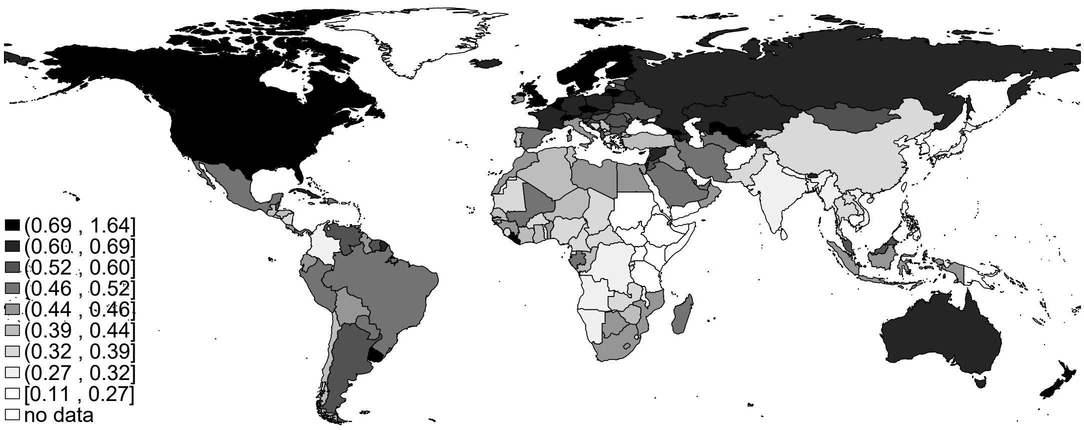

(a) Theil index of inequality in the share of college graduates 1970-2010




(b) Kernel density of the share of college graduates in 2010




(c) Share of college graduates by country in 2010




(d) Agriculture-to-nonagriculture ratio in the share of college graduates by country in 2010

Figure 1: Worldwide distribution of skills

share of the labor force employed in the urban sector, in which productivity is greater. Our baseline scenario is even more optimistic; it assumes that half the correlation between productivity (aggregate or skill bias) and the share of college-educated workers is due to technological externalities. In this context, the growth factor increases from 2.5 to 5 in the poorest countries.[3] Interestingly, we show that keeping the share of college-educated workers constant but transposing the US sector allocation explains one third of the total effect above. This suggests that internal mobility frictions (such as liquidity constraints, imperfect information, or congestion effects) generate a misallocation of workers in poor countries and shows the relevance of a two-sector approach (see Bryan et al., 2014; Hsieh and Klenow, 2009). In contrast, with the exception of small island developing states, the effect of international migration on economic development is small.

Second, we use the model to predict the future geography of skills (i.e., the evolution of human capital and urbanization), population and income during the 21st century. Accounting for interdependencies among demographic, economic and educational variables has rarely been done in projection exercises.[4] In contrast, our micro-founded structure enables us to produce consistent projections and to identify the key factors that will govern the future geography of skills and income. Our baseline scenario assumes a continuation of the ongoing convergence trends in the access to education (possibly initiated by the Millennium Development Goals). In terms of education and urbanization, our baseline prospects are less optimistic than official projections. In line with the evolution of the last decade (see Figure 1a), the baseline predicts fairly stable disparities in the world distribution of skills. We also envisage slower urbanization in developing countries, due to persistent mobility frictions. When extrapolating ongoing trends, the dynamics of the geography of skills *per se* does not translate into drastic changes in global income inequality. These socio-demographic and inequality prospects are highly robust to the size of technological externalities, to the preference structure, and to future international migration policies.

Within the context of the convergence literature,[5] this means that the current convergence in the access to education is too slow to drastically reduce income inequality. The recent decline in inequality is due to the success of some

[3] In a maximalist scenario in which the sizes of externality are proxied by the correlations, human capital almost becomes the single determining factor for global inequality.

[4] For example, the demographic projections of the United Nations do not anticipate the economic forces and policy reforms that shape demography (see Mountford and Rapoport, 2016). The recent projections by International Institute for Applied Systems Analysis (IIASA) include the educational dimension (see Samir et al., 2010), predicting the population of 120 countries by level of educational attainment and accounting for differentials in fertility, mortality and migration by education. However, assumptions about future educational development (e.g., partial convergence in enrollment rates) are also deterministic and seemingly disconnected from changes in the economic environment. Given the high correlation between economic and socio-demographic variables, assuming cross-country convergence in demographic indicators implicitly suggests that economic variables should also converge in the long run. This is not what historical data reveal (see Bourguignon and Morrisson, 2002; Sala-i Martin, 2006).

[5] The convergence literature studies the evolution of inequality between people and between countries. Absolute divergence in income per capita is obtained when countries are not weighted by their size (Pritchett, 1997). When country size is accounted for, global inequality continuously increased between the Industrial Revolution and the 1970s (Bourguignon and Morrisson, 2002) but has decreased since then (Sala-i Martin, 2006).

of the largest countries in the planet (for example, China, India and the rest of Asia), which offsets the divergent incomes of the poorest countries (for example, the African continent). Demographic imbalances are such that the weight of the poorest countries will continuously increase. Without drastic changes in the ongoing productivity and socio-demographic trends, our baseline shows that world income inequality should start rising again. In addition, the future geography of skills and income is sensitive to education policies and to internal mobility frictions. Attenuating or eliminating the convergence in education costs induces dramatic effects on population growth, urbanization and income inequality. In the same vein, obstructing internal mobility generates huge misallocation costs. In line with the Sustainable Development Agenda, our analysis clearly suggests that policies targeting access to all levels of education (what is needed to promote higher education), education quality and sustainable urban development are vital to limit demographic pressures and global inequality.

The rest of this paper is organized as follows. Section 2 provides a summary of the related literature. Section 3 describes our model. In Section 4, we parameterize this model to match historical data over the period 1980-2010 and the socio-demographic prospects for 2040. Section 5 discusses our simulation results, distinguishing between the contemporaneous implications of human capital inequality, the projections for the 21st century, and a sensitivity analysis. Finally, Section 6 concludes.

## 2 Related Literature

Our paper speaks to the literature on the links between human capital accumulation and productivity growth and the literature on the determinants of labor mobility and its effect on economic development. In this section, we review the body of literature that helps contextualizing our approach.

Although the role of human capital as a determinant of productivity growth has been debated, its importance as a proximate cause of development is much less disputed (Acemoglu et al., 2014; Glaeser et al., 2004; Jones, 2014). Our technological specification distinguishes between college and non-college educated workers. This is consistent with Goldin and Katz (2007), Card (2009) and Ottaviano and Peri (2012), who find high substitutability between workers with no schooling and those with a high school degree but small substitutability between those with no schooling and workers with a college education. In this context, increasing the share of college-educated workers not only affects their average skill level and cognitive ability but also generates positive labor market complementarities for the less educated. Jones (2014) builds a generalized development accounting framework that includes such complementarities; he shows that for a reasonable level of the elasticity of substitution (e.g., equal to 2), human capital explains approximately 50% of the ratio of income per worker between the richest and poorest countries. Although such a success rate is still limited, it is greater than what was found in earlier studies that assumed perfect substitution between all categories of workers.[6]

[6] Assuming the income per worker equals \$100,000 in the richest countries and \$5,000 in the poorest countries, a success rate of 50% means that income per capita would reach \$10,000 in

Furthermore, greater contributions of human capital to growth can be obtained by assuming technological externalities. These externalities have been the focus of many recent articles and have generated a certain level of debate. Using data from US cities (Moretti, 2004) or US states (Acemoglu and Angrist, 2000; Iranzo and Peri, 2009), some instrumental-variable approaches show substantial externalities (Moretti, 2004), while others do not (Acemoglu and Angrist, 2000). In the cross-country literature, there is evidence of a positive effect of schooling on innovation and technology diffusion (see Benhabib and Spiegel, 1994; Caselli and Coleman, 2006; Ciccone and Papaioannou, 2009). Other studies identify skill-biased technical changes: when the supply of human capital increases, firms invest in skill-intensive technologies (Acemoglu, 2002; Autor et al., 2003; Restuccia and Vandenbroucke, 2013). Finally, another set of contributions highlights the effect of human capital on the quality of institutions (Bobba and Coviello, 2007; Castelló-Climent, 2008; Murtin and Wacziarg, 2014). Comparative development studies suggest that focusing on highly skilled workers is more appropriate for accounting for such externalities.[7] Squicciarini and Voigtländer (2015) show that upper-tail human capital was instrumental in explaining the process of technology diffusion during the French Industrial Revolution. However, they assert that mass education (proxied by the average level of literacy) was positively associated with development at the onset of the Industrial Revolution but did not explain growth. Confirming Mokyr's findings for the British Revolution, they conclude that the effect of "the educated elite" on local development becomes stronger when the aggregate technology frontier expands more rapidly. It can be argued that this situation also characterizes the modern globalized world, in which most rich countries use advanced technologies, while poor countries struggle to adopt them. The contemporaneous contributions of human capital in poor countries are studied in Castelló-Climent and Mukhopadhyay (2013). They use data on Indian states over the period 1961-2001 and show that a one percent change in the proportion of tertiary-educated workers has the same effect on growth as a 13% decrease in illiteracy rates (equivalently, a one standard deviation in the share of college graduates has the same effect as three standard deviations in literacy). Aggregate and skill-biased externalities cannot be ignored when dealing with long-run growth and inequality. However, given the uncertainty about their levels, our analyses and projections cover several plausible scenarios.

As far as the source of human capital disparities is concerned, the geography of skills is clearly endogenous. Investments in higher education depend on access to education—which varies across income groups (e.g. Galor and Zeira, 1993; Mookherjee and Ray, 2003) and regions (e.g. Lucas, 2009)—as well as on the quality of education (e.g. Castelló-Climent and Hidalgo-Cabrillana, 2012). Human capital disparities are also affected by international and internal labor mobility. International migration affects knowledge accumulation, as well-educated people exhibit much greater propensity to emigrate than do the less educated and tend

poor countries after transferring the human capital level of the richest countries to the poorest countries (i.e., the income ratio would decrease from 20 to 10).

[7] Meisenzahl and Mokyr (2011) argue that the British Industrial Revolution is not so much due to the few dozens of "great inventors" (scientists, PhD holders) nor to the mass of literate factory workers. Instead, in terms of skills, they highlight the role of the top 3-5% of the labor force, including artisans, entrepreneurs and employees.

to agglomerate in countries/regions with high rewards to skill (Belot and Hatton, 2012; Docquier and Rapoport, 2012; Grogger and Hanson, 2011; Kerr et al., 2016). This predominating high-skilled bias in international migration is due to migrants' self-selection (high-skilled people being more responsive to economic opportunities and political conditions abroad, having more transferable skills, having greater ability to gather information or finance emigration costs, etc.) and to the skill-selective immigration policies conducted in the major destination countries (Docquier et al., 2009).

Internal mobility frictions can also be responsible for development inequality. Rodrik (2013) demonstrates that manufacturing industries exhibit unconditional convergence in productivity, while the whole-economy income per worker does not converge across countries. The reason is that a fraction of workers is stuck in the wrong sectors and that these sectoral and/or regional misallocations are likely to be important in poor countries. Such misallocations can be driven by the existence of liquidity constraints, imperfect information, or congestion effects (Bryan et al., 2014; Hsieh and Klenow, 2009). In the same vein, our analysis sheds light on the effect of international migration on global inequality, on the fraction of income disparities explained by internal mobility frictions, and on the implications of labor mobility for future development.

# 3 Model

Our model sheds light on the interactions between the geography of skills and the distribution of income. It endogenizes the accumulation of skills and its implications for economic development.[8] We depict a set of economies with two sectors/regions, $r = (a, n)$, denoting agriculture ($a$) and nonagriculture ($n$), and two types of workers, $s = (h, l)$, denoting college-educated workers ($h$) and the less educated workers ($l$). We assume that agents live for two periods (childhood and adulthood). The number of adults of type $s$ living in region $r$ at time $t$ is denoted by $L_{r,s,t}$. Time is discrete, and one period is meant to represent the active life of one generation (30 years). The retirement period is ignored. In the benchmark version of the model, goods produced in the two sectors are assumed to be perfectly substitutable from the point of view of consumers; their price is normalized to unity. In the robustness checks, we consider an alternative specification with imperfectly substitutable goods entering into a non-homothetic preference structure, as in Boppart (2014). Adults are the only decision makers. They maximize their well-being and decide where to live, how much to consume, and how much to invest in their children's quantity and quality. The latter decisions are governed by a warm-glow motive; adults directly value investments in the quality and quantity of their children, but they do not anticipate the future income and utility of their children (as in De La Croix and Doepke, 2003; De la Croix and Doepke, 2004; Galor, 2011; Galor and Weil, 2000). The dynamic structure of the model is thus totally recursive. The model endogenizes the levels of productivity of both sectors/regions (and the resulting productivity gap), human capital accu-

[8]Our model is similar to Delogu et al. (2018) but relies on a different training technology, accounts for richer technological externalities, includes two sectors per country, and jointly endogenizes internal and international migration flows.

mulation, fertility decisions, and internal and international labor mobility. This section describes our assumptions and defines the intertemporal equilibrium.

## 3.1 Technology

Total output in period $t$ is a sum of the production in agriculture and nonagriculture, $Y_t = Y_{a,t} + Y_{n,t}$. In each sector, production is proportional to labor in efficiency units. Such a model without physical capital features a globalized economy with a common international interest rate. This hypothesis is in line with Kennan (2013) or Klein and Ventura (2009), who assume that capital "chases" labor.[9] In line with Gollin et al. (2014) or Vollrath (2009), each country is characterized by a pair of production functions with two types of labor, college-educated and low-skilled labor ($\ell_{r,s,t}$ $\forall r, s$). We generalize their work by assuming CES (constant elasticity of substitution) specifications with sector-specific elasticities of substitution.[10] The supply of labor, $\ell_{r,s,t}$, differs from the adult population size, $L_{r,s,t}$, because participation rates are smaller than one: as explained below, raising children induces a time cost and decreases labor market participation. Output levels at time $t$ are given by the following CES function:

$$Y_{r,t} = A_{r,t} \left( \sum\nolimits_s \varpi_{r,s,t} \ell_{r,s,t}^{\frac{\sigma_r - 1}{\sigma_r}} \right)^{\frac{\sigma_r}{\sigma_r - 1}} \qquad \forall r, t, \tag{1}$$

where $A_{r,t}$ denotes the productivity scale in sector $r$ at time $t$, $\varpi_{r,s,t}$ is a sector-specific variable governing the relative productivity of workers of type $s$ (such that $\varpi_{r,h,t} + \varpi_{r,l,t} = 1$) and $\sigma_r \in \mathbb{R}_+$ is the sector-specific elasticity of substitution between the two types of workers employed in sector $r$.

The CES specification is flexible enough to account for substitutability differences across sectors. In particular, we consider a greater elasticity of substitution in the agricultural sector ($\sigma_a > \sigma_n$). Wage rates are determined by the marginal productivity of labor and there is no unemployment. This yields:

$$w_{r,s,t} = A_{r,t} \left( \sum\nolimits_s \varpi_{r,s,t} \ell_{r,s,t}^{\frac{\sigma_r - 1}{\sigma_r}} \right)^{\frac{1}{\sigma_r - 1}} \varpi_{r,s,t} \ell_{r,s,t}^{\frac{-1}{\sigma_r}} \qquad \forall r, s, t. \tag{2}$$

It follows that the wage ratio between high-skilled and low-skilled workers in region $r$ is given by the following:

$$R^w_{r,t} \equiv \frac{w_{r,h,t}}{w_{r,l,t}} = R^\varpi_{r,t} \left( R^\ell_{r,t} \right)^{\frac{-1}{\sigma_r}} \qquad \forall r, t, \tag{3}$$

where $R^\ell_{r,t} \equiv \frac{\ell_{r,h,t}}{\ell_{r,l,t}}$ is the skill ratio in the labor force of region $r$ at time $t$ and $R^\varpi_{r,t} \equiv \frac{\varpi_{r,h,t}}{\varpi_{r,l,t}}$ measures the skill bias in relative productivity. Although human

[9] Ortega and Peri (2014) find that capital adjustments are rapid in open economies: an inflow of immigrants increases one-for-one employment and capital stocks in the short term (i.e. within one year), leaving the capital/labor ratio unchanged. In the medium term, demographic change may affect the worldwide capital/labor ratio. Nevertheless, in a closed setting in the vein of Ramsey (1928) or Solow (1956), the interest rate is totally determined by the inter-temporal discount rate of individuals (or by the savings rate) on the long-run balanced growth path. In this paper, we abstract from potential variations in the international interest rate and its impact on within- and between-country inequality.

[10] This elasticity plays a key role in development accounting and is shown to vary across sectors (Caselli and Ciccone, 2013; Jones, 2014; Lucas, 2009).

capital is used in agriculture, the literature has emphasized that the marginal product of human capital is greater in the nonagricultural sector (see Gollin et al., 2014; Lucas, 2009; Vollrath, 2009).

Two types of technological externality are factored in. First, we consider a simple Lucas-type, aggregate externality (see Lucas, 1988) and assume that the scale of the total factor productivity (TFP) in each sector is a concave function of the skill ratio in the resident labor force. This specification captures the fact that college-educated workers facilitate democratization, innovation and the adoption of advanced technologies. We assume that the region-specific TFP equals to the following:

$$A_{r,t} = \gamma^t \overline{A}_{r,t} \left(R^{\ell}_{r,t}\right)^{\epsilon_r} \qquad \forall r, t, \tag{4}$$

where $\gamma^t$ is a time trend in productivity that is common to all countries ($\gamma > 1$), $\overline{A}_{r,t}$ is the exogenous component of TFP in region $r$ (reflecting exogenous factors such as the proportion of arable land, climatic factors, soil fertility, ruggedness, etc.), and $\epsilon_r \in (0,1)$ is a pair of elasticities of TFP to the skill-ratio in the sector. The TFP gap between the two sectors is thus given by the following:

$$\Gamma_t \equiv \frac{A_{n,t}}{A_{a,t}} = \frac{\overline{A}_{n,t}\left(R^{\ell}_{n,t}\right)^{\epsilon_n}}{\overline{A}_{a,t}\left(R^{\ell}_{a,t}\right)^{\epsilon_a}}. \tag{5}$$

In Gollin et al. (2014), the "nonagriculture/agriculture" ratio of value added per worker decreases with development; it amounts to 5.6 in poor countries (bottom 25%) and 2.0 in rich countries (top 25%). After adjusting for hours worked and human capital, the ratio falls to 3.0 in poor countries and 1.7 in rich countries. In our model the findings of Gollin et al. (2014) can then be driven by the correlation between economic development and three country-specific characteristics: (i) the exogenous productivity gap between sectors, $\overline{A}_{n,t} \neq \overline{A}_{a,t}$, (ii) the differences in the elasticity of TFP to human capital, $\epsilon_n \neq \epsilon_a$, or (iii) the disparities in human capital across sectors, $R^{\ell}_{n,t} \neq R^{\ell}_{a,t}$. The latter operate through the ratio of TFP (as shown in Eq. (5)) and through labor market complementarities (captured by the CES transformation function in Eq. (1)).

Second, we assume a skill-biased technical change. As the technology improves, the relative productivity of college-educated workers increases, and this is particularly the case in the nonagricultural sector (Acemoglu, 2002; Restuccia and Vandenbroucke, 2013). For example, Autor et al. (2003) show that computerization is associated with a declining relative industry demand for routine manual and non-cognitive tasks and an increased relative demand for non-routine cognitive tasks. The observed relative demand shift favors college versus non-college labor. We write:

$$R^{\varpi}_{r,t} = \overline{R}^{\varpi}_{r} \left(R^{\ell}_{r,t}\right)^{\kappa_r} \qquad \forall r, t, \tag{6}$$

where $\overline{R}^{\varpi}_{r}$ is an exogenous term, and $\kappa_r \in (0,1)$ is a pair of elasticities of the skill-bias to the skill-ratio in the sector.

## 3.2 Preferences

We now model the process of skill accumulation as the outcome of education and mobility decisions. First, individual decisions to emigrate result from the comparison of discrete alternatives: staying in the region of birth, emigrating to the other region, or emigrating to a foreign country. To model these decisions, we use a logarithmic *outer utility function* with a deterministic and a random component. The utility of an adult of type $s$, who is born in region $r^*$ and is moving to region/country $r$, is given by:

$$U_{r^*r,s,t} = \ln v_{r,s,t} + \ln(1 - x_{r^*r,s,t}) + \xi_{r^*r,s,t} \quad \forall r^*, r, s, t, \tag{7}$$

where $v_{r,s,t} \in \mathbb{R}$ is the deterministic level of utility that can be reached in the location $r$ at period $t$ (governed by the inner utility function described below) and $x_{r^*r,s,t} \leq 1$ captures the effort required to migrate from region $r^*$ to location $r$ (such that $x_{r^*r^*,s,t} = 0$). Migration costs are exogenous; they vary across location pairs, across education levels, and over time. The individual-specific random taste shock for moving from country $r^*$ to $r$ is denoted by $\xi_{r^*r,s,t} \in \mathbb{R}$ and follows an *iid* Type-I Extreme Value distribution:

$$F(\xi) = \exp\left[-\exp\left(-\frac{\xi}{\mu} - \vartheta\right)\right],$$

where $\mu > 0$ is a common scale parameter governing the responsiveness of migration decisions to changes in $v_{r,s,t}$ and $x_{r^*r,s,t}$ and $\vartheta \approx 0.577$ is the Euler's constant. Although $\xi_{r^*r,s,t}$ is individual-specific, we omit individual subscripts for notational convenience.

Second, we model education decisions as in Galor and Weil (2000), Galor (2011), De La Croix and Doepke (2003), De la Croix and Doepke (2004), Delogu et al. (2018). We assume that the *inner utility* $\ln v_{r,s,t}$ is a function of consumption ($c_{r,s,t}$), fertility ($n_{r,s,t}$) and the probability that each child becomes highly skilled ($p_{r,s,t}$):

$$\ln v_{r,s,t} = \ln c_{r,s,t} + \theta \ln\left(n_{r,s,t} p_{r,s,t}\right) \quad \forall r, s, \tag{8}$$

where $\theta \in (0, 1)$ is a preference parameter for the quantity and quality of children.

The probability that a child becomes high skilled increases with the share of time that is spent in education ($q_{r,s,t}$):

$$p_{r,s,t} = (\pi_r + q_{r,s,t})^\lambda \quad \forall r, s, \tag{9}$$

where $\pi_r$ is an exogenous parameter that is region-specific and $\lambda$ governs the elasticity of knowledge acquisition to the education investment.

A type-$s$ adult in region $r$ receives a wage rate $w_{r,s,t}$ per unit of time worked. Raising a child requires a time cost $\phi$ (thereby reducing the labor market participation rate), and each unit of time spent by a child in education incurs a cost equal to $E_{r,t}$. The budget constraint is written as follows:

$$c_{r,s,t} = w_{r,s,t}(1 - \phi n_{r,s,t}) - n_{r,s,t} q_{r,s,t} E_{r,t}. \tag{10}$$

It follows that the labor supply of type-$s$ adults in region $r$ at time $t$ is given by the following:

$$\ell_{r,s,t} = L_{r,s,t}(1 - \phi n_{r,s,t}). \tag{11}$$

In the following sub-sections, we solve the optimization problem backwards. We first determine the optimal fertility rate and investment in education in a given location $r$, which characterizes the optimal level of utility, $v_{r,s,t}$, that can be reached in any location. We then characterize the choice of the optimal location.

### 3.2.1 Education and fertility

Each adult in region $r$ maximizes her utility (8) subject to the constraints (9) and (10). The first-order conditions for an interior solution are as follows:

$$\begin{aligned}\frac{\phi w_{r,s,t} + q_{r,s,t}E_{r,t}}{w_{r,s,t}(1-\phi n_{r,s,t}) - n_{r,s,t}q_{r,s,t}E_{r,t}} &= \frac{\theta}{n_{r,s,t}},\\ \frac{n_{r,s,t}E_{r,t}}{w_{r,s,t}(1-\phi n_{r,s,t}) - n_{r,s,t}q_{r,s,t}E_{r,t}} &= \frac{\theta\lambda}{\pi_r + q_{r,s,t}}.\end{aligned}$$

Solving this system gives the following:

$$\left\{\begin{array}{l} q_{r,s,t} = \frac{\lambda\phi w_{r,s,t} - \pi_r E_{r,t}}{(1-\lambda)E_{r,t}} \\ n_{r,s,t} = \frac{\theta(1-\lambda)}{1+\theta}\cdot\frac{w_{r,s,t}}{\phi w_{r,s,t} - \pi_r E_{r,t}} \end{array}\right. \quad \forall r, s.$$

The cost of education is assumed to be proportional to the wage of high-skilled workers in the region, multiplied by a fixed, region-specific factor $\psi_{r,t}$ (capturing education policy/quality, population density, average distance to schools, etc.):

$$E_{r,t} = \psi_{r,t} w_{r,h,t} \quad \forall r, s. \tag{12}$$

Factoring Eq. (12) into the first-order conditions gives the following:

$$\left\{\begin{array}{l} q_{r,h,t} = \frac{\lambda\phi}{(1-\lambda)\psi_{r,t}} - \frac{\pi_r}{1-\lambda} \\ q_{r,l,t} = \frac{\lambda\phi}{(1-\lambda)\psi_r R^w_{r,t}} - \frac{\pi_r}{1-\lambda} \end{array}\right. \quad \text{and} \quad \left\{\begin{array}{l} n_{r,h,t} = \frac{\theta(1-\lambda)}{1+\theta}\frac{1}{\phi - \pi_r\psi_r} \\ n_{r,l,t} = \frac{\theta(1-\lambda)}{1+\theta}\frac{1}{\phi - \pi_r\psi_r R^w_{r,t}} \end{array}\right. \tag{13}$$

Note that $R^w_{r,t} > 1$ implies that college-educated workers have fewer and more educated children in all regions ($q_{r,h,t} > q_{r,l,t}$ and $n_{r,h,t} < n_{r,l,t}$). The model also predicts that investments in education vary across regions, and are likely to be greater in the nonagricultural region. Under the plausible condition $\psi_{a,t}/\psi_{n,t} > 1$, college-educated workers living in urban areas have fewer and more educated children ($q_{n,h,t} > q_{a,h,t}$ and $n_{n,h,t} < n_{a,h,t}$). Finally, when $(\psi_{a,t}R^w_{a,t})/(\psi_{n,t}R^w_{n,t}) > 1$, this is also the case for the low skilled ($q_{n,l,t} > q_{a,l,t}$ and $n_{n,l,t} < n_{a,l,t}$). These results are in line with Lucas (2009), who assumes that human capital accumulation increases with the fraction of people living in cities (seen as *centers of intellectual interchange and recipients of technological inflows*).

The deterministic indirect utility function can be obtained by substituting Eq. (13) into Eq. (8):

$$\left\{\begin{array}{l} \ln v_{r,h,t} = \chi + \ln(w_{r,h,t}) + \theta\lambda\ln\left(\frac{1}{\psi_{r,t}}\right) - \theta(1-\lambda)\ln(\phi - \pi_r\psi_{r,t}) \\ \ln v_{r,l,t} = \chi + \ln(w_{r,l,t}) + \theta\lambda\ln\left(\frac{1}{\psi_{r,t}}\right) - \theta(1-\lambda)\ln\left(\phi - \pi_r\psi_{r,t}R^w_{r,t}\right) \\ \qquad\qquad + \ln\left(\frac{\phi\left(1+\theta\lambda\left(1-1/R^w_{r,t}\right)\right) - \pi_r\psi_{r,t}R_{r,t}\left(1+\theta\left(1-1/R^w_{r,t}\right)\right)}{\phi - \pi_r\psi_{r,t}R^w_{r,t}}\right) \end{array}\right. \tag{14}$$

where $\chi = \theta\ln\left(\frac{\theta}{1+\theta}(1-\lambda)^{1-\lambda}\lambda^\lambda\right) - \ln(1+\theta)$ is a constant.

Together with the number and structure of the resident population at time $t$ ($L_{r,s,t}\ \forall r, s$), fertility and education decisions ($n_{r,s,t}, q_{r,s,t}\ \forall r, s$) determine the size and structure of the native population before migration ($N_{r,s,t+1}\ \forall r, s$) at time $t+1$. We have the following:

$$\left\{\begin{array}{l} N_{r,h,t+1} = L_{r,h,t}n_{r,h,t}p_{r,h,t} + L_{r,l,t}n_{r,l,t}p_{r,l,t} \\ N_{r,l,t+1} = L_{r,h,t}n_{r,h,t}\left[1 - p_{r,h,t}\right] + L_{r,l,t}n_{r,l,t}\left[1 - p_{r,l,t}\right] \end{array}\right. \quad \forall r, t. \tag{15}$$

### 3.2.2 Migration and population dynamics

Given their taste characteristics (captured by $\xi$), individuals choose the location that maximizes her/his utility, defined in Eq. (7). Under the Type I Extreme Value distribution for $\xi$, McFadden (1974) shows that the solution to a discrete choice problem (that is, in our context, a decision to migrate from region $r$ to $r^*$) is governed by a logit expression. The emigration rate is given by the following:

$$\frac{M_{r^*r,s,t}}{N_{r^*,s,t}} = \frac{\exp\left(\frac{\ln v_{r,s,t}+\ln(1-x_{r^*r,s,t})}{\mu}\right)}{\sum_k \exp\left(\frac{\ln v_{k,s,t}+\ln(1-x_{r^*k,s,t})}{\mu}\right)} = \frac{(v_{r,s,t})^{1/\mu}(1-x_{r^*r,s,t})^{1/\mu}}{\sum_k (v_{k,s,t})^{1/\mu}(1-x_{r^*k,s,t})^{1/\mu}}.$$

Skill-specific emigration rates are endogenous and restricted between 0 and 1. Staying rates $(M_{r^*r^*,s,t}/N_{r^*,s,t})$ are governed by the same logit model. It follows that the emigrant-to-stayer ratio $(m_{r^*r,s,t})$ is governed by the following expression:

$$m_{r^*r,s,t} \equiv \frac{M_{r^*r,s,t}}{M_{r^*r^*,s,t}} = \left(\frac{v_{r,s,t}}{v_{r^*,s,t}}\right)^{1/\mu}(1-x_{r^*r,s,t})^{1/\mu}. \quad (16)$$

Equation (16) is a gravity-like migration equation, which states that the ratio of emigrants from region $r^*$ to location $r$ to stayers in region $r^*$ (i.e., individuals born in $r^*$ who remain in $r^*$) is an increasing function of the utility achievable in the destination location $r$ and a decreasing function of the utility attainable in $r^*$. The proportion of migrants from $r^*$ to $r$ also decreases with the bilateral migration cost $x_{r^*r,s,t}$. Heterogeneity in migration tastes implies that emigrants select all destinations for which $x_{r^*r,s,t} < 1$ (if $x_{r^*r,s,t} = 1$, the corridor is empty).

Individuals born in region $n$ (resp. $a$) have the choice between staying in their region of origin $n$ (resp. $a$), moving to the other region $a$ (resp. $n$), or emigrating to a foreign country $f$. Contrary to Hansen and Prescott (2002) or Lucas (2009), labor is not perfectly mobile across sectors/regions; internal migration costs ($x_{an,s,t}$ and $x_{na,s,t}$) capture all private costs that migrants must incur to move between regions. In line with Young (2013), internal mobility is driven by self-selection, i.e., skill-specific disparities in utility across regions as well as heterogeneity in individual unobserved characteristics ($\xi$). Overall, if $v_{n,s,t} > v_{a,s,t}$, net migration is in favor of urban areas, but migration is limited by the existence of migration costs, whose sizes govern the sectoral misallocations of workers (Rodrik, 2013). Similarly, international migration costs ($x_{af,s,t}$ and $x_{nf,s,t}$) capture private costs and the legal/visa costs imposed by the destination countries. They are also assumed to be exogenous.

Using Eq. (16), we can characterize the equilibrium structure of the resident population at time $t$:

$$\begin{cases} L_{n,s,t} = \frac{N_{n,s,t}}{1+m_{na,s,t}+m_{nf,s,t}} + \frac{m_{an,s,t}N_{a,s,t}}{1+m_{an,s,t}+m_{af,s,t}} + I_{n,s,t} \\ L_{a,s,t} = \frac{N_{a,s,t}}{1+m_{an,s,t}+m_{af,s,t}} + \frac{m_{na,s,t}N_{n,s,t}}{1+m_{na,s,t}+m_{nf,s,t}} + I_{a,s,t} \end{cases} \quad \forall s, \quad (17)$$

where $I_{r,s,t}$ stands for the inflow of immigrants (which only applies to migration from developing to OECD member states). For simplicity, we assume that the distribution of immigrants by OECD destination is time-invariant and calibrated on the year 2010. Eq. (16) also determines the outflow of international migrants by education level ($O_{s,t}$):

$$\begin{aligned} O_{s,t} &= M_{nf,s,t} + M_{af,s,t} \\ &= \frac{m_{nf,s,t}N_{n,s,t}}{1+m_{na,s,t}+m_{nf,s,t}} + \frac{m_{af,s,t}N_{a,s,t}}{1+m_{an,s,t}+m_{af,s,t}} \quad \forall s, \end{aligned} \tag{18}$$

where $N_{r,s,t}$ is a predetermined variable given by (15).

## 3.3 Intertemporal equilibrium

An intertemporal equilibrium for the world economy can be defined as following:

**Definition 1** *For a set* $\{\gamma, \theta, \lambda, \phi, \mu\}$ *of common parameters, a set* $\{\sigma_r, \epsilon_r, \kappa_r\}$ *of sector-specific elasticities, a set* $\{\overline{A}_{r,t}, \overline{R}^{\varpi}_{r,t}, \overline{x}_{r^*r,s,t}, \psi_r, \pi_r\}$ *of country- and region-specific exogenous characteristics, and a set* $\{N_{r,s,0}\}$ *of predetermined variables, an intertemporal equilibrium is a reduced set of endogenous variables* $\{A_{r,t}, \varpi_{r,h,t}, w_{r,s,t}, n_{r,s,t}, q_{r,s,t}, v_{r,s,t}, E_{r,t}, m_{r^*r,s,t}, N_{r,s,t+1}, L_{r,s,t}\}$*, which simultaneously satisfies technological constraints (4), (6) and (12), profit maximization conditions (2), utility maximization conditions (13), (14) and (16) in all countries and regions of the world, and such that the equilibrium structure and dynamics of population satisfy (15) and (17).*

The equilibrium level of the other variables described above (in particular, $\ell_{r,s,t}$, $R^{\ell}_{r,t}$, $R^{\varpi}_{r,t}$, $R^{w}_{r,t}$, $\Gamma_t$ as well as urbanization rates and international migration outflows) can be computed as a by-product of the reduced set of endogenous variables. Note that equilibrium wage rates are obtained by substituting the labor force variables into the wage equation (2), thereby assuming full employment. By the Walras law, the market for goods is automatically balanced.

# 4 Data and parameterization

In this section, we describe our parameterization strategy for 145 developing countries and for the entire set of 34 OECD countries.[11] Our parameterization strategy consists in calibrating a few common elasticities and a large number of region-specific parameters in order to (perfectly) match socio-demographic and economic data for the years 1980 and 2010 (including internal and international migrations) and to be in line with official socio-demographic projections for the year 2040.[12] We use all the degrees of freedom of the data to identify the parameters needed. Consequently, our model is exactly identified and cannot produce a test of its assumptions. However, it is worth noticing that we use relatively consensual specifications for the production and migration technologies and that we test the robustness of our results in the Appendix. We start describing how we estimate the geographic distribution of skills. Then, the parameterization of the technological

[11] With the exceptions of Macao, North-Korea, Somalia and Taiwan, all countries that are not covered by our sample have less than 100,000 inhabitants.

[12] Our set of region-specific parameters includes TFP and skill-bias levels, education costs, internal and international migration costs.

and preference parameters is outlined. More details about the calibration can be found in Section A.1 in the Appendix. We finally explain the general hypotheses used to initialize our baseline projections for the 21st century.

*Estimating the geography of skills.* To construct labor force data by education level and by sector ($L_{r,s,t}$), we follow the four steps described below.

In the *first step*, we extract population data by age group from the United Nations Population Division and combine it with the database on educational attainment described in Barro and Lee (2013). For the years 1980 and 2010, we proxy the working age population with the number of residents aged 25 to 60. To proxy the number of high-skilled workers in each country, we multiply the working age population by Barro and Lee's estimates of the proportion of individuals aged 25 and over with tertiary education completed (denoted by $H_t$). The rest of the working age population is treated as a homogeneous group of less educated workers. Barro and Lee's data are available for 143 countries. For the other countries, we make use of estimated data from Artuç et al. (2015). Note that Barro and Lee (2013) also document the average years of schooling of the working age population ($\text{Yo}S_t$), a variable that we use in the third step of our estimation strategy. We are able to characterize the total number of workers ($\Sigma_{r,s} L_{r,s,t}$) and the total number of college-educated and less educated workers ($\Sigma_r L_{r,h,t}$ and $\Sigma_r L_{r,l,t}$) by country. The same strategy has been applied to all decades between 1970 and 2010 to compute the between-country index of inequality depicted in Figure 1.

In the *second step*, we split the total population data by region/sector. When it is possible, we use the share of employment in agriculture, which is available from the World Development Indicators. This variable is available for 134 countries in 2010 and for 61 in 1980. However, the same database also provides information on the share of people living in rural areas, which is highly correlated with the share of employment in agriculture (correlation of 0.71 in 2010 and 0.75 in 1980). When the share of employment in agriculture is not available, we predict it using estimates from year-specific regressions as a function of the share of people living in rural areas. This determines the total number of workers ($\Sigma_s L_{r,s,t}$) in both sectors.

The major problem is that, to the best of our knowledge, there is no database documenting the share of college graduates by region or by sector ($H_{r,t}$). We estimate these shares and compare them with nationally representative data from the Gallup World Polls. More details on the Gallup World Polls are provided in Section A.1 in the Appendix. To compute these shares, we collect or construct data on the years of schooling by sector ($\text{YoS}_{r,t}$)) and use them to predict the sector-specific shares of college graduates as a function of $\text{YoS}_{r,t}$. Hence, our *third step* consists of gathering data on $\text{YoS}_{r,t}$ and imputing the missing values. Gollin et al. (2014) and Ulubaşoğlu and Cardak (2007) provide incomplete data on the countrywide average years of schooling and on the average years of schooling in agriculture and nonagricultural for different years.[13] We have data for 20 countries around the year 1980 and for 65 countries around the year 2010. We match these data to the closest year that marks the beginning of the 1980 and 2010 decades.

[13] In Gollin et al. (2014) and Vollrath (2009), the nonagriculture/agriculture ratio of years of schooling varies between 2.0 or 1.5 in poor countries and is close to 1.0 in rich countries.

For the missing countries, we take advantage of the high correlation between the gap in years of schooling, $\text{YoS}_{n,t}/\text{YoS}_{a,t}$, and the average years of schooling in the country, $\text{YoS}_t$. We predict the schooling gap by using estimates from year-specific regressions of this gap on $\text{YoS}_t$.[14]

Finally, in the *fourth step*, we take advantage of the high correlation between the average years of schooling and the proportion of college graduates in the labor force at the national level. We estimate the relationship between these variables, $H_t = f(\text{YoS}_t)$, using Barro and Lee's data, and then use the estimated coefficients to predict the share of college graduates in the urban sector, $H_{r,t} = f(\text{YoS}_{r,t})$.[15] We then fit the average share of college graduates from Barro and Lee by adjusting the share of college graduates in the rural sector.

To validate our estimation strategy, we compute the correlation between the sector-specific estimated shares of college graduates and the shares obtained from household surveys. Using the Gallup World Poll data (available for approximately 145 countries), we can estimate the skill-ratio $R^{\ell}_{r,t}$ in the number of respondents by country and region (corrected by sample weights); on average, the correlation between the Gallup sample and our estimates is equal to 0.70 in the urban region and to 0.73 in the rural region. The same imputation strategy can be used to identify the sector-specific shares of college graduates in total employment for all decades between 1970 and 2010. We use it to compute the within-country index of inequality depicted in Figure 1. Additional stylized facts are provided in Section A.1 in the Appendix.

*Technology parameters.* The output in each sector depends on the size and skill structure of employment. Below, we explain how fertility rates are calibrated for each skill group and for each region/sector. Combining labor force data ($L_{r,s,t}$) with fertility rates ($n_{r,s,t}$) allows us to quantify the employment levels ($\ell_{r,s,t}$) and the total employment in efficiency unit using Eq. (11).

To calibrate the set of technological parameters $\left\{\sigma_r, \epsilon_r, \kappa_r, \overline{R}^{\varpi}_r, \overline{A}_{r,t}\right\}$, we proceed in two steps. First, we calibrate the parameters affecting the private returns to higher education. For each sector, we combine our estimates for $\ell_{r,s,t}$ with cross-country data on the income gap between college graduates and the less educated. This enables us to parameterize the elasticities of substitution between workers ($\sigma_r$), the relative productivity of college graduates ($R^{\varpi}_r$), the magnitude of the skill-biased externalities ($\kappa_r$), and the scale factors of the skill-bias technology ($\overline{R}^{\varpi}_r$). In the second step, we focus on the social returns to education. We use output data by sector and identify the level of total factor productivity that matches the GDP data by sector. We then investigate the relationship between TFP and the skill ratio, which enables us to estimate the size of the aggregate TFP externalities ($\epsilon_r$) and the TFP scale factors ($\overline{A}_{r,t}$). Figure A2 in the Appendix summarizes our main findings.

In the *first step*, we calibrate the elasticity of substitution between college graduates and less educated workers, relying on existing studies. For the nonagricultural sector, there is a large number of influential papers that propose specific estimates for industrialized countries (i.e., countries where the employment share

[14] Simple OLS regressions give $\log \frac{\text{YoS}_n}{\text{YoS}_a} = 1.944 - 0.744 \log \text{YoS}$ ($R^2$=0.809) in 2010, and $\log \frac{\text{YoS}_n}{\text{YoS}_a} = 1.464 - 0.550 \log \text{YoS}$ ($R^2$=0.905) in 1980.

[15] Simple OLS regressions give $\log H = -4.804 + 0.279 \log \text{YoS}$ ($R^2 = 0.496$) in 2010, and $\log H = -5.133 + 0.306 \log \text{YoS}$ ($R^2 = 0.575$) in 1980.

of agriculture is small). Johnson (1970) and Murphy et al. (1998) obtain values for $\sigma_n$ of approximately 1.3. Ciccone and Peri (2005) and Krusell et al. (2000) find values of approximately 1.6, and Ottaviano and Peri (2012) suggest setting $\sigma_n$ close to 2.0. Angrist (1995) recommends a value above 2 to explain the trends in the college premium in the Palestinian labor market. For the agricultural sector, it is usually assumed that the elasticity of substitution is much larger. For example, Vollrath (2009) or Lucas (2009) consider that labor productivity is determined by the average level of human capital of workers (thus assuming perfect substitution between skill groups). In line with the existing literature, we assume $\sigma_n = 2$ and $\sigma_a = \infty$.

Once the elasticities are chosen, we use sector-specific data on returns to schooling to calibrate the relative productivity of college-educated workers. In the agricultural sector, we rule out the possibility of a skill-biased technical change in agriculture ($\kappa_a = 0$), and assume a linear technology with a constant $R_a^{\varpi}$ for all countries and all periods (we use $\varpi_a = 0.57$). For the nonagricultural sector, we use data on the skill premium and calibrate $R_n^{\varpi}$ as a residual of Eq. (3). Regressing $R_n^{\varpi}$ on $R_n^{\ell}$ yields an estimate of 0.38. Given the bidirectional causation relationship between the skill bias and education decisions, we consider this estimate as an upper bound for the skill-bias externality. In our baseline projections, we assume that half the correlation is due to the skill-bias externality (i.e., $\kappa_n = 0.19$). We calibrate the scale factor $\overline{R}_n^{\varpi}$ as a residual from Eq. (6).

In the *second step*, we use data on the national gross domestic product (GDP) and on the agriculture share in value added. We obtain data on output by sector in the year 2010 and identify the TFP levels ($A_{r,t}$) by dividing the sector-specific output by the quantity of labor in efficiency unit using Eq. (1). There is a clear positive relationship between TFP and the share of college-educated workers in both sectors. Regressing the log of $A_{r,t}$ on the log of $R_{r,t}^{\ell}$ gives a coefficient of 0.57 in the nonagricultural sector and 0.66 in agriculture. Given the reverse causation relationship between productivity and the education decision, we consider these estimates as upper bounds for the aggregate TFP externality. In our baseline scenario, we assume that half the correlation between TFP and the share of college-educated workers is due to the schooling externality (i.e., $\epsilon_n = 0.28$ and $\epsilon_a = 0.33$). We calibrate the scale factor $\overline{A}_n$ as a residual from Eq. (4).

Figure A2 in the Appendix shows that these assumptions are consistent with the macro and microdata. Nevertheless, alternative technological scenarios are considered in the robustness checks (see Section A.4 in the Appendix).

*Preference parameters.* The literature indicates some common values of several preference parameters. We assign the following values to the parameters that are time-invariant and equal for all countries: $\theta = 0.25$, $\lambda = 0.5$ and $\phi = 0.14$.[16] From Eq. (14) and Eq. (16), the scale parameter of the distribution of migration tastes ($\mu$) is the inverse of the elasticity of bilateral migration to the wage rate. Bertoli and Fernández-Huertas Moraga (2013) find a value between 0.6 and 0.7 for this elasticity. Hence, we use $\mu = 1.4$.

Parameters $\pi_r$ and $\psi_{r,t}$ are country- and sector-specific. They govern the fertility and education decisions. We calibrate them to match the population dynamics

[16] Given the expression in Eq. (10), this assumption reflects setting the bound of the maximal number of children equal to 7 (i.e., 14 children per couple). See Docquier et al. (2017) for a brief review of studies using similar parameter values.

between the years 1980 and 2010, i.e., the transition from the resident population in 1980 and the native population in 2010. We begin by estimating the size of the *before-migration* population in 2010 by skill group ($\sum_r N_{r,s,2010}$). The average (national) fertility rate ($\overline{n}_{1980}$) is thus obtained by dividing the total native population of adults in 2010 ($\sum_{r,s} N_{r,s,2010}$) by the total resident population of adults in 1980 ($\sum_{r,s} L_{r,s,1980}$). We also observe the skill structure of the native population in 2010 ($N_{r,s,2010}$), which helps identifying education decisions in 1980 ($\overline{q}_{1980}$). We use the Gallup World Polls and extract the average number of children per household by region and by skill level for 2010 to proxy the fertility differentials. We calibrate $\pi_r$ and $\psi_{r,t}$ to match $n_{r,s,1980}$ and $\overline{q}_{1980}$. From 2010 onwards, the number of children and education decisions are endogenous.

We then estimate the skill and regional distribution of workers in 1980 and 2010 and calibrate internal migration costs as a residual from Eq. (16). For this, we assume there is only migration from rural to urban regions (i.e., $x_{an,s,t} < 1$ and $x_{na,s,t} = 1$). Similarly, we compute the average utility achievable in OECD destination countries and calibrate the international migration costs ($x_{af,s,t}$ and $x_{nf,s,t}$) as a unique solution from Eq. (16) to match the DIOC migration data. Again, more details are provided in Section A.1.

*Baseline trajectory for the 21st century.* Our parameter set is such that the model matches the geographic disparities in income, population and human capital in the year 2010, and their evolution between 1980 and 2010. Our baseline also includes technological externalities, assuming that half the correlation between TFP (and skill bias) and the share of college-educated workers is due to the schooling externality. Alternative technological and preference scenarios are considered in Section A.4 in the Appendix.

The philosophy of our baseline projection exercise is to predict the future trends in income, population and human capital if all parameters, with the exception of the TFP scale factor (assumed to grow at a constant rate of 1.5% per year in all countries) and the parameters governing access to education, remain constant. More precisely, we constrain our baseline trajectory to be compatible with official socio-demographic projections for the year 2040 for each country. The rationale for matching medium-term projections is that the size and skill structure of the national population in 2040 are determined by fertility and education decisions in the contemporaneous period (i.e., the years 2010 to 2040). Hence, the reliability of medium-term projections is high, and their consistency with the economic environment is presumably good. Nevertheless, we let the micro-founded model predict the sectoral allocation of labor and international migration rates in 2040 as well as the evolution of socio-demographic variables beyond 2040. The comparison between our simulations and official projections is discussed in Section 5.2.

To match the size and skill structure of the national population in 2040, we allow for country-specific proportional adjustments in $\psi_{r,t}$ ($r = a, n$) (i.e., the same relative change in both sectors, keeping $\psi_{a,t}/\psi_{n,t}$ constant) that minimizes the sum of squared differences in total population and in its skill structure between the baseline simulations and the UN projections for the year 2040. Remember $\psi_{r,t}$ determines the access to education in the region. Comparing the new levels of $\psi_{r,2010}$ with those obtained in 1980 (i.e., $\psi_{r,1980}$), we identify a conditional convergence process in the access to education. We see it as a likely consequence of the Millennium Development policy. More precisely, we estimate two quadratic, region-specific convergence equations, considering the US as the benchmark frontier:

$$\ln\left(\psi_{r,t+1}/\psi_{r,t}\right) = \alpha_r + \beta_r \ln\left(\psi_{r,t}^{USA}/\psi_{r,t}\right) + \gamma_r \left(\ln\left(\psi_{r,t}^{USA}/\psi_{r,t}\right)\right)^2. \quad (19)$$

We obtain $\gamma_a = 0.032$, $\gamma_n = 0.046$, $\beta_a = -0.195$ and $\beta_n = -0.223$, in which all parameters are highly significant. This quadratic convergence process implies that middle-income countries converge more rapidly than low-income countries do. For subsequent years, our baseline scenario assumes a continuation of this quadratic convergence process, in line with the new Sustainable Development Agenda. Alternative (i.e., more and less optimistic) convergence scenarios will also be considered in Section 5.3.

# 5 Results

Our model is used to investigate the interactions between the current/future distributions of skills and global inequality worldwide. First, in line with the development accounting methodology, we only use the (parameterized) technological block of the model and disregard the endogeneity of human capital accumulation. Section 5.1 describes a set of counterfactual experiments that allow identifying the causal impact of skills accumulation on inequality. More precisely, we quantify the fraction of contemporaneous development inequality that is explained by differences in the national proportion of highly educated workers, by their allocation across sectors, and by international migration. Second, our attention is turned to the determinants of the geography of skills. In Section 5.2, we provide integrated projections of worldwide population, urbanization, human capital and income per capita for the 21st century. Then, we assess the sensitivity of our projections to future educational policies (Section 5.3) and to future mobility frictions (Section 5.4).[17] Section 5.5 describes the underlying income inequality prospects and discusses their sensitivity.

## 5.1 How much does the current geography of skills matter for global inequality?

In line with the development accounting methodology (Jones, 2014), we consider the US as the base-case economy and proceed with three static counterfactual experiments to quantify the economic implications of skill accumulation in the year 2010. The advantage of our two-sector model is that we can separately quantify the development implications of skill accumulation, of the sectoral allocation of labor, and of international labor mobility. For each country, we first simulate the counterfactual level of national income per worker ($y_{\text{CF}}$) obtained after transposing the US shares of college-educated workers in each sector. We then compare it with the observed level ($y_{\text{obs}}$). The second counterfactual consists of keeping the country-specific share of college-educated residents constant but allocating high-skilled and low-skilled workers across sectors based on their allocation in the US

[17]Section A.4 in the Appendix shows that our socio-demographic projections are highly robust to the size of technological externalities as well as to the way preferences for agricultural and nonagricultural goods are modeled.

economy. In the third counterfactual, we keep the country-specific share of college-educated natives constant but simulate a no-migration scenario (US international emigration rates are almost nil). The results are depicted in Figure 2.

Figures 2a and 2b give the counterfactual levels of income per capita and the smoothed growth factor ($y_{\text{CF}}/y_{\text{obs}}$) obtained under three technological scenarios after transposing the US shares of college-educated workers. Under these scenarios, all countries have the same national fraction of college graduates as the US has and the same regional shares by educational level. In Figure 2a, the bold line shows the observed income levels; countries are ranked by ascending order with respect to the observed level of income per worker. Most studies in development accounting disregard technological externalities (see Jones, 2016) or consider that externalities are small (Caselli and Ciccone, 2013). In contrast, our baseline scenario (solid line) assumes that externality sizes are equal to 50% of the correlations between human capital and technological characteristics (i.e., $\kappa_n = 0.19$, $\kappa_a = 0$, $\epsilon_n = 0.28$ and $\epsilon_a = 0.33$). The variants (dashed line) assume no externality, or externalities equal to 100% of the correlations (i.e., $\kappa_n = 0.38$, $\kappa_a = 0$, $\epsilon_n = 0.56$ and $\epsilon_a = 0.66$). Figure 2b gives the smoothed growth factor induced by the counterfactual under the same externality variants.

We show that the geography of skills matters for development, regardless of the size of technological externalities. In the absence of any externality, transposing the US educational structure increases income per worker by a factor of 2.5 for countries in the lowest quartile of the income distribution (i.e., from \$5,000 to \$12,500). The growth factor decreases with economic development, as the distance to the technology frontier gets smaller. This is in line with Jones (2014), who finds a growth factor of 2 for poor countries with the same elasticity of substitution. As in Jones, the effect is mainly driven by the fact that high-skilled workers are more productive and by the labor market complementarity with less educated workers. In addition, our model accounts for the sector allocation of labor. Transposing the US skill shares and the US sectoral allocation of workers not only increases the level of education but also increases the size of the urban (more productive) sector. This is equivalent to raising the average TFP level in a one-sector model and explains our greater success rate. In our baseline scenario with conservative externalities, transposing the US skill shares increases income per worker by a factor of 5 in the poorest countries (i.e., $y$ increases from \$5,000 to \$25,000) after transposing the US educational structure. In the full-externality scenario, human capital almost becomes the single determining factor for economic development. Unsurprisingly, the size of technological externalities has a strong influence on the global inequality effect of the geography of skills.[18]

Figures 2c and 2d illustrate the role of the sector allocation of skills under the same externality scenarios. We simulate the effect of transposing the US skill-specific urban shares (keeping the country-wide share of college graduates at the observed levels). The baseline scenario is shown as the solid line, while

[18] In unreported simulations, we used the baseline externality scenario (50% of correlations) and included one externality at a time. The results are highly sensitive to the aggregate TFP externality (almost equivalent to the baseline with both externalities). However, the skill-biased externality affects within-country wage disparities but plays a negligible role in explaining income per capita differentials (almost equivalent to the no-externality scenario). Directed technical changes slightly exacerbate income disparities across countries (poorest countries are better off in the absence of skill-biased technical changes, unlike richest countries).

the zero- and the full-externality scenarios are shown as dashed lines. Under the baseline, transposing the US urban shares for each category of worker increases income per worker by a factor of 1.7 in the lowest quartile of the distribution (i.e., about one third of the total effect identified above). Transposing the US shares in employment means increasing the urban share from 20% to 95% in the poorest countries. This shock drastically increases the mean levels of productivity and income. Poor countries are unable to realize these gains because individuals have no incentives to move due to liquidity constraints, imperfect information, or congestion effects (Bryan et al., 2014; Hsieh and Klenow, 2009). In line with Rodrik (2013), our results suggest that internal mobility frictions are responsible for a large misallocation of workers in poor countries and shows the relevance of a two-sector approach.

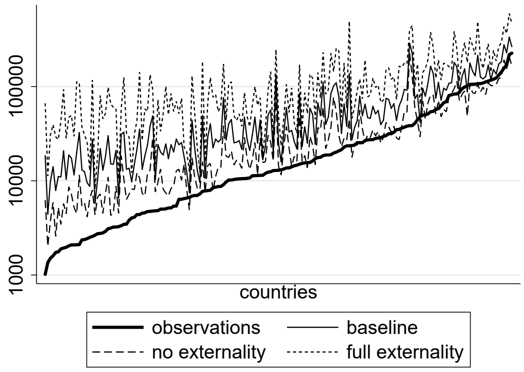
Under the same externality scenarios, Figures 2e and 2f illustrate the role of international migration. We simulate the effect of returning all expatriates to their home country (no-migration scenario). The baseline scenario is shown as the solid line, while the zero- and the full-externality scenarios are shown as dashed lines. With the exception of Small Island Developing States (corresponding to the peaks on Figure 2e), the effect of international migration on global inequality is small. On average, returning all international migrants to origin countries in the bottom quartile of the distribution increases income per workers by a factor of 1.2 in the baseline case (and by a factor of 1.5 with full externalities). This is because average emigration rates to the OECD are small in developing countries (approximately 5% for college graduates and less than 1% for the low-skilled). Contrary to the previous experiments, the global inequality response to international migration is rather limited.[19]

## 5.2 The changing geography of skills: baseline prospects

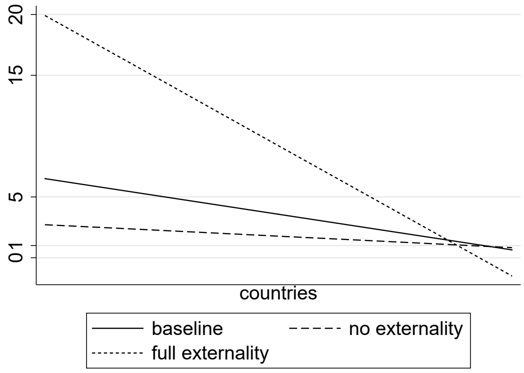
Disparities in the level and in the sector allocation of skills explain a significant fraction of economic inequality across countries. We now turn our attention to the factors governing the long-term trend in the geography of skills. This section compares our baseline socio-demographic prospects for the 21st century with the widely used projections of the United Nations (the UN medium variant).

The UN projections assume a long-term convergence in fertility, mortality and education attainment, and constant immigration flows. Given the high correlation between socio-demographic and economic variables, the UN medium variant implicitly assumes income convergence between countries. In the medium term, the UN projections also predict higher demographic growth in developing countries. These facts are incompatible with the hypothesis of constant migration flows. In contrast, our micro-founded model provides consistent projections of fertility, education, migration and income inequality. As explained above, our baseline projections rely on a minimum of assumptions. Note that we assume a quadratic, region-specific convergence process in access to education (i.e., in $\psi_{r,t}$).

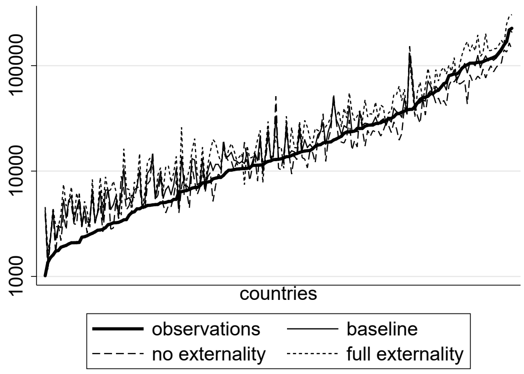
[19] Table A1 and Table A2 in the appendix give a more detailed description of the effect of the different static counterfactual experiments for the US and for the $15^{th}$ (Cambodia), $25^{th}$ (Ghana), $50^{th}$ (Tunisia), $75^{th}$ (Mexico) and $85^{th}$ (Greece) percentiles of the income distribution. The presentation is organized as in Jones (2014). Table A1 focuses on the average level of income per worker, while Table A2 distinguishes between the two production sectors.



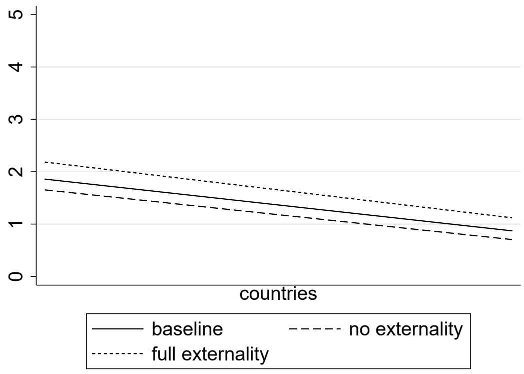

(a) Distribution of GDP per capita: Transposing US skill shares



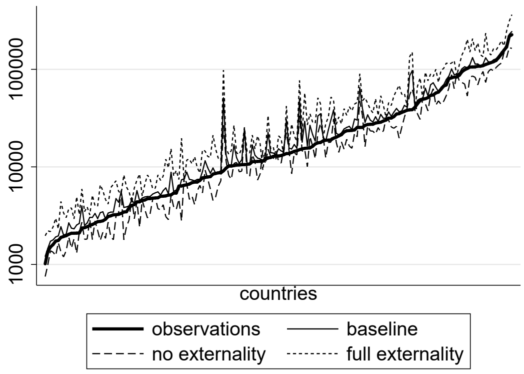

(b) Smoothed growth factor: Transposing US skill shares



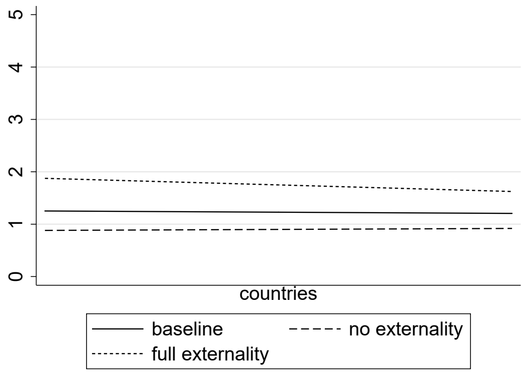

(c) Distribution of GDP per capita: Transposing US urban shares




(d) Smoothed growth factor: Transposing US urban shares




(e) Distribution of GDP per capita: No international migration




(f) Smoothed growth factor: No international migration

Figure 2: Geography of skills and income per worker: static counterfactuals

Notes: On the horizontal axis, countries are ranked by ascending order with respect to the observed level of GDP per capita and the respective scenario. Results are depicted for the baseline, zero- and full-externality scenarios.

This implies that regions at an intermediate level of development converge towards the US frontier more rapidly than do the poor ones. We keep all other parameters constant, including the medium level of technological externalities.

Prospective results are described in Figure 3. The simulated (dashed lines) and official (continuous lines) trajectories of population, share of college graduates, and share of the urban population are depicted in Figures 3a, 3c and 3e, respectively. Separate curves are provided for OECD countries, for developing countries, and for the entire world.[20] The cross-country correlations between our simulations (Y-axis) and official projections (X-axis) for population, share of college graduates, and share of the urban population for the year 2100 are described in Figures 3b, 3d and 3f, respectively. Bubbles are proportional to country size (OECD countries in light gray and developing countries in dark gray). The 45-degree line allows visualizing whether our long-term simulations are greater or smaller than official projections.

Figures 3a and 3b show that our baseline trajectory is very much in line with official socio-demographic projections. Although we only initialize our simulations to be compatible with the 2040 national population levels, our long-term level of the adult population is almost equal to official projections. Furthermore, the cross-country correlation between simulated and UN population sizes in the year 2100 equals 0.98.[21]

Nevertheless, we obtain significant differences when focusing on the evolution of education and urbanization. As far as education is concerned, we are less optimistic than the United Nations. Figure 3c shows that the long-term, worldwide share of college graduates is smaller than that reflected in official projections. This share increases from 8.8% in 2010 to 17.3% in 2100 in our model, against 21.4% in the UN medium scenario. Similar differences are obtained for OECD and developing countries. As shown on Figure 3d, the cross-country correlation between simulated and UN shares of college graduates in the year 2100 is large (0.91).[22] However, most countries are below the 45 degree line, and for a large number of small OECD countries, compared with the UN projections, the simulated shares of college graduates is multiplied by a factor between 0.7 and 0.9. According to our baseline prospects for the 21st century, the share of college graduates increases from 20.5% to 48% in OECD countries, and from 5.1% to 12.5% in the developing world. Assuming a continuation of the ongoing convergence in access to education, the ratio of skill shares between OECD and developing countries increases from 3.3 to 3.8.

Similarly, Figure 3e shows that our predicted share of the population living in urban areas is smaller than the UN projections. The worldwide urban share increases slightly from 53.0% in 2010 to 58.3% in 2100. These trends are the outcomes of two opposing forces: the rural/urban fertility differential and the net internal mobility towards cities (driven by the rising educational attainment). The former is important and imprecisely modeled in official projections. In Figure 3f,

[20]The definition of the developing countries follows the official definition of the United Nations. The remaining 29 countries (not reported) are neither classified as an OECD nor as a developing country.

[21]The regression line of Figure 3b is given by: $\textit{baseline} = -0.33 + 1.02 \cdot \textit{official}$ ($R^2 = 0.97$).

[22]The regression line of Figure 3d is given by the following: $\textit{baseline} = -0.03 + 0.92 \cdot \textit{official}$ ($R^2 = 0.82$).

the cross-country correlation between simulated and UN urban shares in the year 2100 equals 0.83.[23] Again, most countries are below the 45-degree line, and for a large number of developing countries, our simulated urban share is multiplied by a factor between 0.5 and 0.8, compared with the UN one. Comparing OECD member states with developing countries, our baseline prospects predict fairly stable disparities in urbanization.

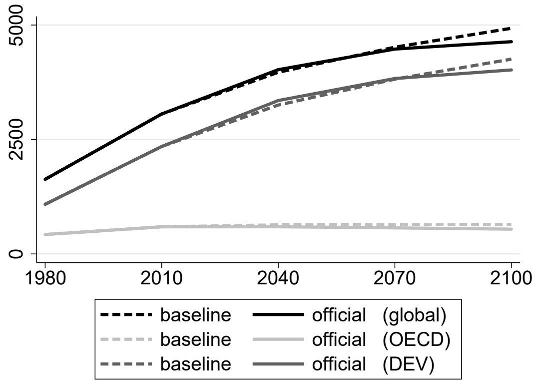
These comparisons give suggestive evidence that our stylized model does a good job in generating realistic and consistent, although less optimistic, projections of population, human capital, and urbanization for the coming decades. Despite a convergence in access to education, our baseline scenario neither predicts a fall in human capital inequality nor a strong convergence in the sector allocation of skills. Importantly, as it is micro-founded, the model also enables us to identify the key factors that will govern the future of the world population and global inequality. In particular, we can assess whether the evolution of population and global inequality is sensitive to future educational policies (i.e., convergence in the access to education) and geographic mobility costs. In Section A.4 in the Appendix, we show that our socio-demographic prospects are highly robust to technological externalities and to the structure of preferences.

## 5.3 Sensitivity to education policies

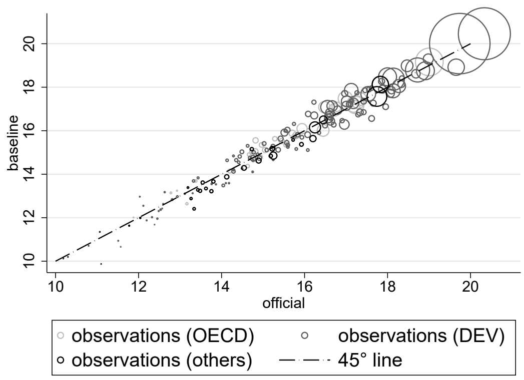
We first assess whether our socio-demographic prospects are sensitive to policies affecting future access to education. In line with the recent *Sustainable Development Agenda*, the baseline scenario assumes a continuation of the quadratic convergence process in education costs observed between 1980 and 2010; this implies that middle-income countries catch up more rapidly than low-income countries do. Figure 4 compares the baseline trajectories of population, education and urbanization with those obtained with a smaller magnitude of the quadratic convergence (we divide the convergence speed by two compared to the baseline) or when there is an unconditional, linear convergence process.

Under the linear convergence scenario, the poorest countries are the most prone to converge. We investigate this possibility by estimating a linear convergence equation for education cost (instead of a second-order polynomial in the baseline): $\ln\left(\psi_{r,t+1}/\psi_{r,t}\right) = \alpha_r + \beta_r \ln\left(\psi_{r,t}^{USA}/\psi_{r,t}\right)$. We obtain the following estimates: $\beta_a = 0.056$ for rural regions, and $\beta_n = 0.074$ for urban regions. Compared to the baseline, this scenario predicts faster human capital accumulation and urbanization in the poorest countries of the world, as shown on Figures 4b, 4d and 4f. Looking at worldwide aggregates, in the long run, this implies a significantly smaller population size, a greater share of college graduates and a greater urban share of the population. Nevertheless, Figures 4a, 4c and 4e show that these aggregate changes are relatively small due to the small demographic size of the low-income countries. With the exception of the poorest countries, our projections are almost identical when using a well-fitted linear or a quadratic convergence model. In other words, when extrapolating current trends in education costs, socio-demographic prospects are fairly robust to the specification of the estimated convergence process.

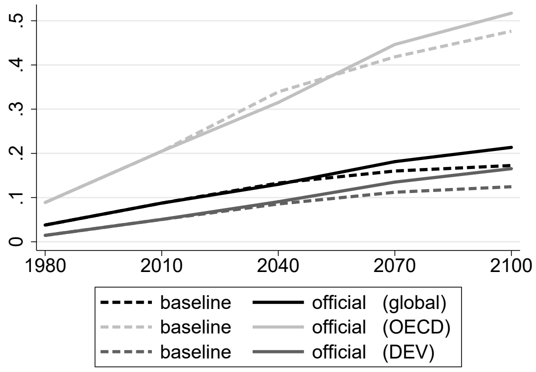
[23] The regression line of Figure 3f is given by: $\textit{baseline} = -0.15 + 1.11 \cdot \textit{official}$ ($R^2 = 0.69$).



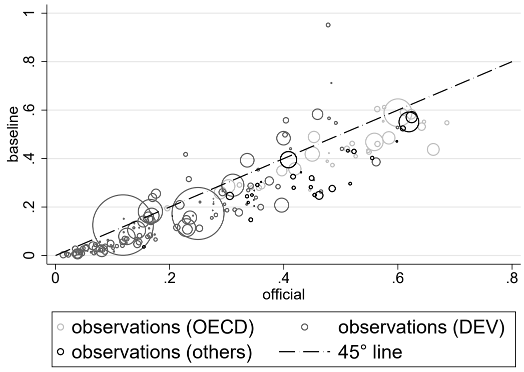

(a) Population (in million of people)



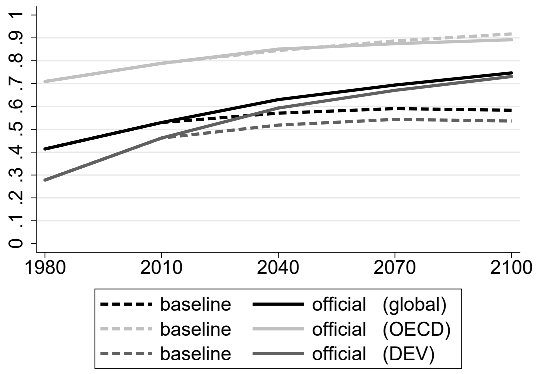

(b) Population in 2100 by countries



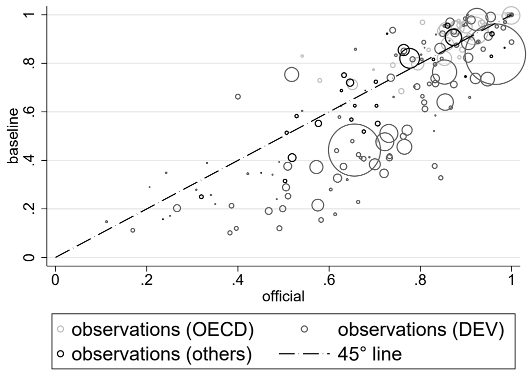

(c) Worldwide share of college-educated workers




(d) Share of college-educated in 2100 by countries




(e) Urban share




(f) Urban share in 2100 by countries

Figure 3: Comparison of the baseline trajectory with official projections by the UN

Notes: The left panel reports the projected population size, the share of college educated workers, and the share of urban population for the baseline and for offical projections (UN medium variant). Results are depicted for the worlwide averages, for countries in the OECD and for developing countries (DEV). The right panel compares the simulated levels in 2100 with official projections. Bubbles are proportional to country size.

However, if we assume a slow-down of convergence (i.e., if we divide by two the convergence speed), it drastically affects the geography of skills and long-term population growth. In the developing world, the proportion of college graduates and the share of the urban population stagnate after 2040. The long-term level of the population is 20% to 25% greater than in the baseline. These changes are noticeable in all developing countries, including the largest ones (see Figures 4b, 4d and 4f). Hence, Figures 4a, 4c and 4e show that the changes in the size and skill structure of the world population are important. In line with the *Sustainable Development Agenda*, our results suggest that policies targeting access to all levels of education and education quality are vital to reduce the demographic pressures and to stimulate human capital accumulation.[24]

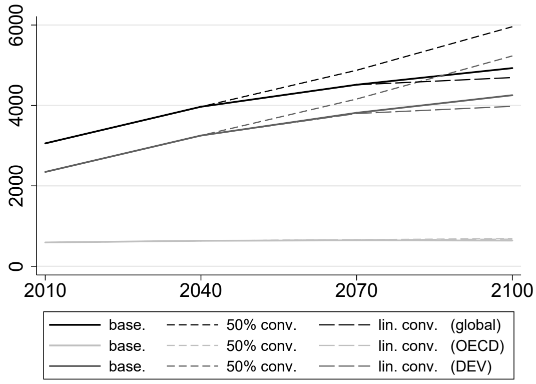
## 5.4 Sensitivity to mobility constraints

We now investigate whether our socio-demographic prospects are sensitive to future mobility frictions. The baseline scenario assumes constant international and internal migration costs in the future. It predicts that the international migration pressures drastically intensify in the OECD countries (see Table A1 in Appendix). We consider here an extreme no-international migration scenario for the future ($x_{rf,s,t} = 1$ after 2010). In the same vein, our static experiments suggest that internal mobility frictions drastically affect the (mis-)allocation of workers between sectors. We consider a no-internal migration scenario with maximal frictions ($x_{an,s,t} = 1$ after 2010). Figure 5 compares the baseline trajectories of population, education, and urbanization with those obtained without international or internal mobility.

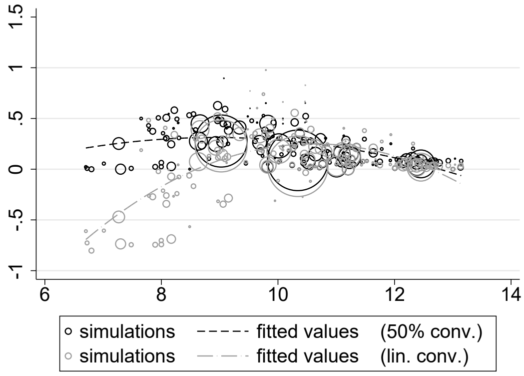
In line with the static development accounting exercise, we find that international migration has limited (and often negligible) effects on aggregated socio-demographic prospects (Figure 5a, 5c and 5e). In the no-migration scenario, Figure 5d shows that the share of college-educated workers increases in developing countries and that the effect is particularly strong in the poorest countries in which emigrants are highly positively selected. However, in general, the trend is mostly governed by small countries (and small developing islands in particular), exhibiting large emigration rates. The effect is small in large countries.[25] Comparing OECD member states with developing countries, the ratio of skill shares in the year 2100 reaches 3.4 (instead of 3.8 in the baseline), but this change is mostly due to the decrease in human capital in OECD countries. Figures 5e and 5f show that the urbanization responses are small, except in OECD countries. This is because immigrants to OECD countries usually reside in urban regions. As far as population is concerned, the no-migration scenario predicts a substantial decrease in the size of the population in Western economies, which is completely balanced out by an increase in developing countries.

[24]Changing demographic shares have drastic implications in terms of immigration and emigration (see Table A2 in the Appendix). In the half-convergence scenario, the number of international migrants increases by 22% compared to the baseline, due to the larger population in developing countries.

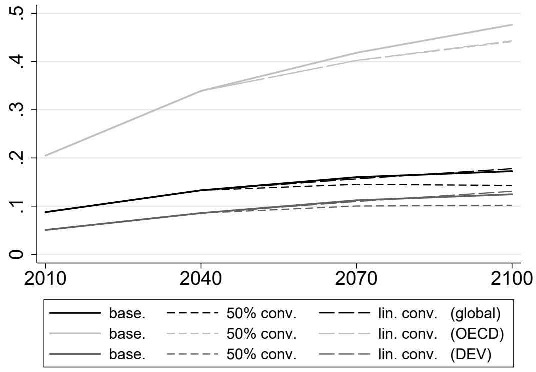
[25]Migration barriers can also affect an individual's incentives to acquire higher education. However, Docquier and Machado (2016) and Delogu et al. (2018) numerically demonstrate that the latter brain gain mechanism has little impact on the world distribution of skills.



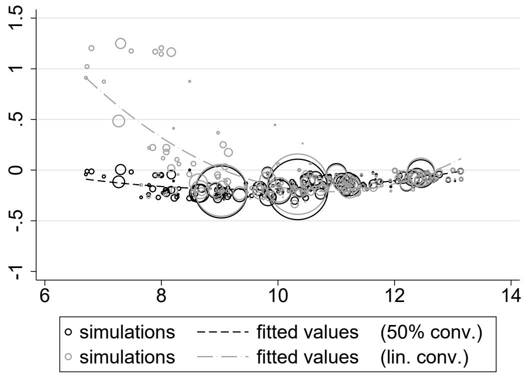

(a) Population (in million of people)



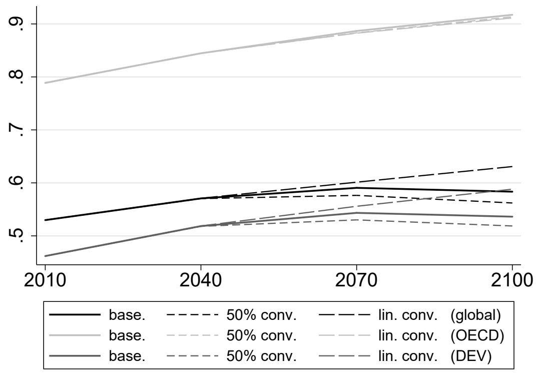

(b) Relative deviations from the baseline in 2100



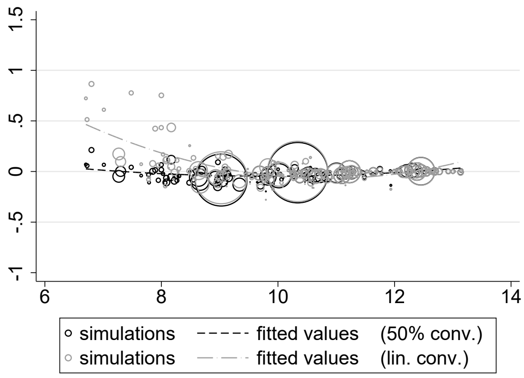

(c) Share of college educated workers




(d) Relative deviations from the baseline in 2100




(e) Share of urban population




(f) Relative deviations from the baseline in 2100

Figure 4: Sensitivity to educational policies

Notes: This figure reports the projected population size, the share of college educated workers, and the share of urban population for the baseline and the respective counterfactual scenario. The "lin. conv." scenario assumes a monotonic convergence in $\psi_{r,t}$. The "50% conv." scenario assumes a slower conditional convergence process.

The socio-demographic effects of internal mobility are greater. Preventing the movement of people from rural to urban areas has larger implications for human capital accumulation in large countries (access to education is better in cities), for the continuation of the urbanization process, and for population growth. Without internal mobility, the long-term level of the population increases by 16% in the developing world, the share of college graduates peaks at 10%, and the urban share declines by half compared to the baseline. This confirms that internal mobility frictions are important to reduce the demographic pressures and to boost human capital accumulation worldwide. Without internal mobility, the long-term ratio of skill shares between OECD member states and developing countries reaches 4.3 (instead of 3.8 in the baseline), and the sector allocation of skills drastically deteriorates in the developing world.

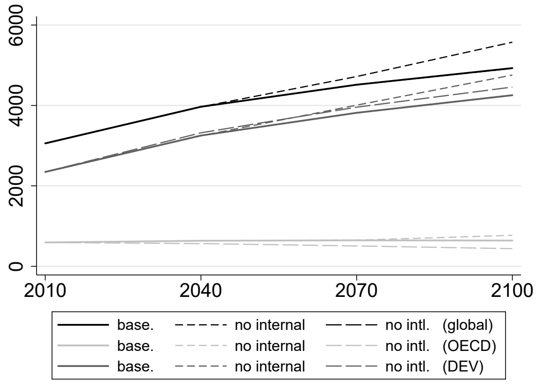
## 5.5 Geography of skills and geography of income

This last section connects the results of the static development accounting experiments with our socio-demographic prospects. Our static analysis shows that global inequality is influenced by the geography of skills. The prospective part shows that a continuation of ongoing trends should neither lead to a drastic fall in human capital inequality nor to strong improvement in the sector allocation of skills.[26] Nevertheless, the geography of skills can be affected by public policies affecting education and internal labor mobility. We now examine how these policies impact the world distribution of income. Our baseline prospects involve a variation of the Theil index of income inequality from 0.81 in 1980 to 1.14 in 2100 (see Figure A3a in Appendix A.3). Figure 6 illustrates this result and analyzes its sensitivity to education and mobility policies. The left panel depicts the trajectory of the average level of income per capita in the OECD member states, in developing countries and in the world. The right panel depicts the sensitivity of the Theil index of income inequality.

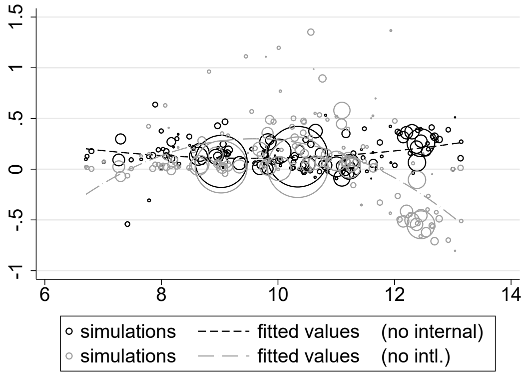
Figures 6a and 6b show the sensitivity of the world distribution of income to education policies. Compared to the baseline, the Theil index is unsurprisingly smaller when we assume linear convergence in the access to education and greater when we divide the coefficients of the quadratic convergence equation by two. However, as illustrated in Figure 6a, the trajectory of income per capita in all regions is not greatly affected by the convergence assumption. Variations in the Theil index are rather mechanical and linked to the construction of the index: the variations are mostly explained by the changing demographic shares of the developed and developing world (as illustrated in Figure 4a).

Figures 6c and 6d show the sensitivity of the world distribution of income to future mobility frictions. Preventing people from migrating internationally markedly reduces the world GDP (as it prevents individuals to move from low-productivity to high-productivity countries) and reduces global income inequality. However, Figure 6c shows that it has a negligible effect on income per capita in the developing world. In other words, development prospects are robust to future international

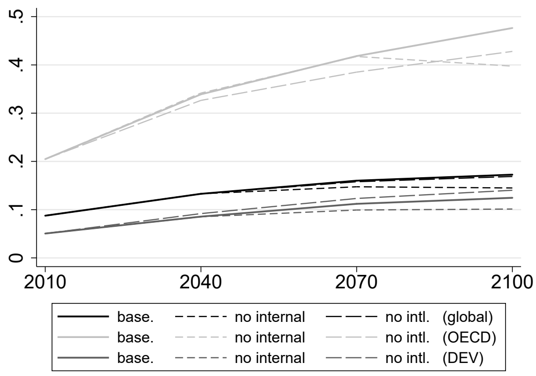
[26] The results reported in Appendix A.3 indicate that the Theil index of human capital inequality remains almost stable over the 21st century. It ranges from 0.63 in 1980 to 0.56 in 2100.



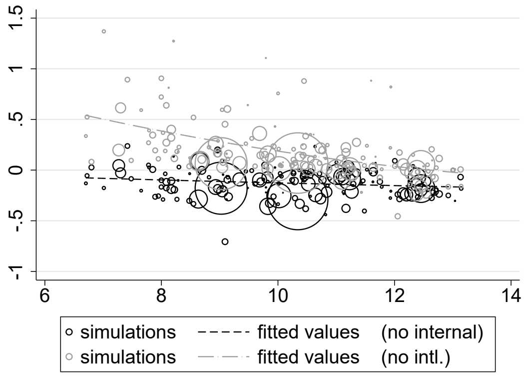

(a) Population (in million of people)



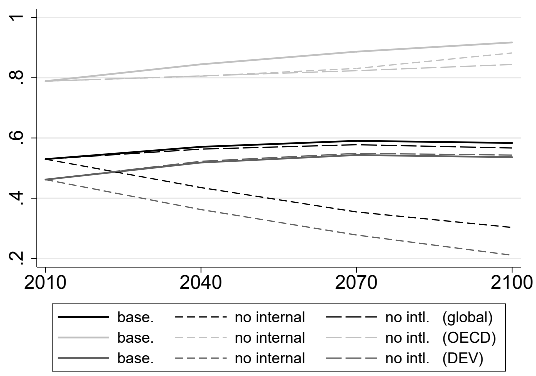

(b) Relative deviations from the baseline in 2100



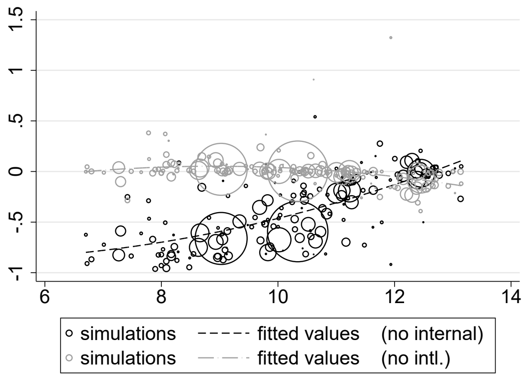

(c) Share of college educated workers




(d) Relative deviations from the baseline in 2100




(e) Share of urban population

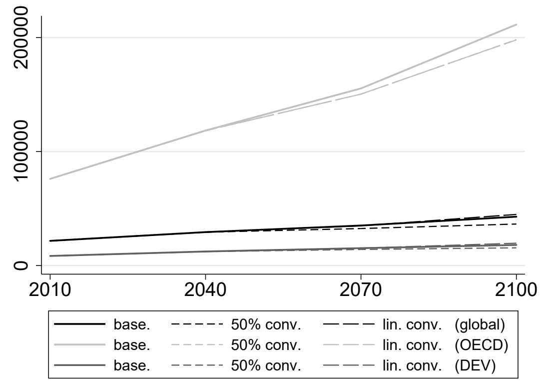



(f) Relative deviations from the baseline in 2100

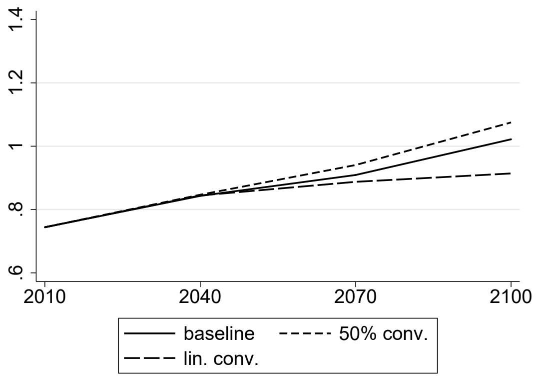
Figure 5: Sensitivity to mobility constraints

Notes: This figure reports the projected population size, the share of college educated workers, and the share of urban population for the baseline and the respective counterfactual scenario. The scenario "no intl." refers to the scenario with prohibitively high international migration costs ($x_{rf,s,t} = 1$) after 2010. The scenario "no internal" refers to the scenario with prohibitively high internal migration costs ($x_{an,s,t} = 1$) after 2010.

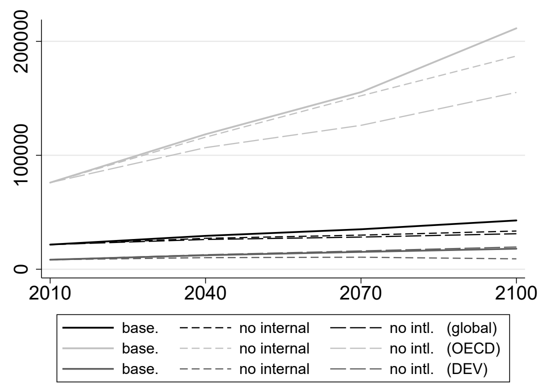
migration barriers.[27] Again, the effect on global inequality is rather mechanical and linked to the construction of the Theil index: cutting migration decreases the demographic share of industrialized countries and increases the share of developing countries. In contrast, the level of income per capita in developing countries is more sensitive to internal migration policies. Preventing rural-to-urban migration reduces income and drastically increases the Theil index of income inequality. In line with our static numerical experiments, internal mobility frictions can induce a large misallocation of workers in poor countries (Rodrik, 2013). Policies targeting sustainable urban development are vital to reduce the demographic pressure and global inequality.



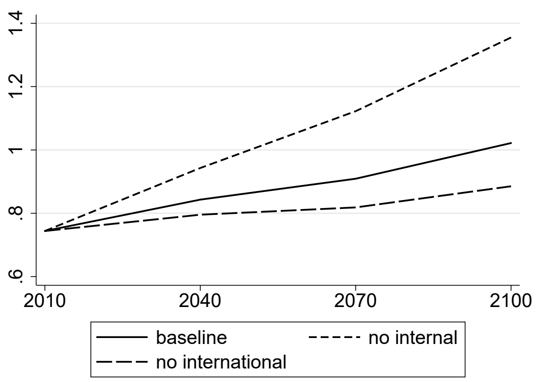

(a) Income per capita - access to education




(b) Theil index - access to education




(c) Income per capita - mobility variant




(d) Theil index - mobility variant

Figure 6: Implications for global income inequality

Section A.4 in the Appendix demonstrates that these conclusions are highly robust to the modeling assumptions. If we change the size of technological externalities or if we consider that agricultural and nonagricultural goods are imperfect substitutes, as in Boppart (2014), we obtain similar trajectories for the Theil index of income inequality. The size of technological externalities affects the levels of income per capita in developing and developed countries but has negligible effect

[27]We are aware that the real contribution of international migration to development might be underestimated here, as the model disregards diaspora externalities (Docquier and Rapoport, 2012) and the link between education decisions and migration prospects.

on inequality. The structure of preferences has little effect on the levels of income per capita and on its distribution.

# 6 Conclusion

This paper analyzes the root drivers of the geographic distribution of skills and its effect on current and future development disparities. We use a multi-country, two-sector, two-class, dynamic model of the world economy that endogenizes population growth, human capital formation and income in all countries and regions. We consider various sizes for technological externalities, alternative structures of preferences, as well as scenarios of access to education, internal and international mobility. Overall, we argue that the geography of skills explains a non-negligible fraction of development disparities between countries and regions. An important part of this effect is due to disparities in the (national) average level of schooling. Nevertheless, when considering the bottom quartile of the income distribution, one third of the total effect is due to disparities, which result from internal mobility frictions, in the sector allocations of workers. Compared to results from the standard, one-sector development accounting model, taking into account within-country disparities in human capital reinforces the role of the geographic allocation of skills. However, although migrants are positively selected in terms of their education level, international migration has little effect on the world distribution of skills and income.

Assuming a continuation of the ongoing convergence process in the access to schooling, we provide unified projections of socio-demographic and economic variables for the 21st century. Our baseline prospects show fairly stable disparities in the world's distribution of skills and slow urbanization in developing countries. This implies that the future geography of skills *per se* is unlikely to bring down global income inequality if access to education does not converge faster than it has over the last 30 years. On the contrary, increasing inequality occurs if the speed of convergence in education cost decreases or if internal mobility frictions increase. In line with the Sustainable Development Agenda, our analysis clearly suggests that policies targeting access to all levels of education, education quality and sustainable urban development are vital to reduce the demographic pressure and global inequality. These conclusions are highly robust to the technological and preference assumptions and to future international migration policies.

# References


Acemoglu, D. (2002). Technical change, inequality, and the labor market. *Journal of Economic Literature*, 40(1):7–72.

Acemoglu, D. and Angrist, J. (2000). How large are human-capital externalities? evidence from compulsory schooling laws. *NBER Macroeconomics Annual*, 15:9–59.

Acemoglu, D., Gallego, F. A., and Robinson, J. A. (2014). Institutions, human capital, and development. *Annual Review of Economics*, 6(1):875–912.

Angrist, J. D. (1995). The economic returns to schooling in the west bank and gaza strip. *American Economic Review*, 85(5):1065–1087.

Artuç, E., Docquier, F., Özden, Ç., and Parsons, C. (2015). A global assessment of human capital mobility: the role of non-oecd destinations. *World Development*, 65:6–26.

Autor, D. H., Levy, F., and Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics*, 118(4):1279–1333.

Barro, R. J. and Lee, J. W. (2013). A new data set of educational attainment in the world, 1950–2010. *Journal of Development Economics*, 104:184–198.

Belot, M. V. and Hatton, T. J. (2012). Immigrant selection in the oecd. *The Scandinavian Journal of Economics*, 114(4):1105–1128.

Benhabib, J. and Spiegel, M. M. (1994). The role of human capital in economic development: Evidence from aggregate cross-country data. *Journal of Monetary Economics*, 34(2):143–173.

Bertoli, S. and Fernández-Huertas Moraga, J. (2013). Multilateral resistance to migration. *Journal of Development Economics*, 102:79–100.

Bertoli, S. and Ruyssen, I. (2018). Networks and migrants' intended destination. *Journal of Economic Geography*, 18(4):705–728.

Biavaschi, C., Burzyński, M., Elsner, B., and Machado, J. (2016). The gain from the drain: Skill-biased migration and global welfare.

Bobba, M. and Coviello, D. (2007). Weak instruments and weak identification, in estimating the effects of education, on democracy. *Economics Letters*, 96(3):301–306.

Boppart, T. (2014). Structural change and the kaldor facts in a growth model with relative price effects and non-gorman preferences. *Econometrica*, 82(6):2167–2196.

Bourguignon, F. and Morrisson, C. (2002). Inequality among world citizens: 1820-1992. *American Economic Review*, 92(4):727–744.

Bryan, G., Chowdhury, S., and Mobarak, A. M. (2014). Underinvestment in a profitable technology: The case of seasonal migration in bangladesh. *Econometrica*, 82(5):1671–1748.

Card, D. (2009). Immigration and inequality. *American Economic Review*, 99(2):1–21.

Caselli, F. and Ciccone, A. (2013). The contribution of schooling in development accounting: Results from a nonparametric upper bound. *Journal of Development Economics*, 104:199–211.

Caselli, F. and Coleman, W. J. (2006). The world technology frontier. *American Economic Review*, 96(3):499–522.

Castelló-Climent, A. (2008). On the distribution of education and democracy. *Journal of Development Economics*, 87(2):179–190.

Castelló-Climent, A. and Hidalgo-Cabrillana, A. (2012). The role of educational quality and quantity in the process of economic development. *Economics of Education Review*, 31(4):391–409.

Castelló-Climent, A. and Mukhopadhyay, A. (2013). Mass education or a minority well educated elite in the process of growth: The case of india. *Journal of Development Economics*, 105:303–320.

Ciccone, A. and Papaioannou, E. (2009). Human capital, the structure of production, and growth. *The Review of Economics and Statistics*, 91(1):66–82.

Ciccone, A. and Peri, G. (2005). Long-run substitutability between more and less educated workers: Evidence from us states, 1950–1990. *The Review of Economics and Statistics*, 87(4):652–663.

De La Croix, D. and Doepke, M. (2003). Inequality and growth: why differential fertility matters. *American Economic Review*, 93(4):1091–1113.

De la Croix, D. and Doepke, M. (2004). Public versus private education when differential fertility matters. *Journal of Development Economics*, 73(2):607–629.

Delogu, M., Docquier, F., and Machado, J. (2018). Globalizing labor and the world economy: the role of human capital. *Journal of Economic Growth*, 23(2):223–258.

Docquier, F., Lowell, B. L., and Marfouk, A. (2009). A gendered assessment of highly skilled emigration. *Population and Development Review*, 35(2):297–321.

Docquier, F. and Machado, J. (2016). Global competition for attracting talents and the world economy. *World Economy*, 39(4):530–542.

Docquier, F., Müller, T., and Naval, J. (2017). Informality and long-run growth. *The Scandinavian Journal of Economics*, 119(4):1040–1085.

Docquier, F. and Rapoport, H. (2012). Globalization, brain drain, and development. *Journal of Economic Literature*, 50(3):681–730.

Foellmi, R. and Zweimüller, J. (2008). Structural change, engel's consumption cycles and kaldor's facts of economic growth. *Journal of Monetary Economics*, 55(7):1317–1328.

Galor, O. (2011). *Unified Growth Theory*. Princeton University Press.

Galor, O. and Weil, D. N. (2000). Population, technology, and growth: From malthusian stagnation to the demographic transition and beyond. *American Economic Review*, 90(4):806–828.

Galor, O. and Zeira, J. (1993). Income distribution and macroeconomics. *The Review of Economic Studies*, 60(1):35–52.

Glaeser, E. L., La Porta, R., Lopez-de Silanes, F., and Shleifer, A. (2004). Do institutions cause growth? *Journal of Economic Growth*, 9(3):271–303.

Goldin, C. and Katz, L. F. (2007). The race between education and technology, cambridge. *MA: Harvard University Press*.

Gollin, D., Lagakos, D., and Waugh, M. E. (2014). The agricultural productivity gap. *The Quarterly Journal of Economics*, 129(2):939–993.

Grogger, J. and Hanson, G. H. (2011). Income maximization and the selection and sorting of international migrants. *Journal of Development Economics*, 95(1):42–57.

Hansen, G. D. and Prescott, E. C. (2002). Malthus to solow. *American Economic Review*, 92(4):1205–1217.

Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and manufacturing tfp in china and india. *The Quarterly Journal of Economics*, 124(4):1403–1448.

Iranzo, S. and Peri, G. (2009). Schooling externalities, technology, and productivity: Theory and evidence from us states. *The Review of Economics and Statistics*, 91(2):420–431.

Johnson, G. E. (1970). The demand for labor by educational category. *Southern Economic Journal*, 37(2):190–204.

Jones, B. F. (2014). The human capital stock: a generalized approach. *American Economic Review*, 104(11):3752–77.

Jones, C. I. (2016). The facts of economic growth. In *Handbook of Macroeconomics*, volume 2, pages 3–69. Elsevier.

Kennan, J. (2013). Open borders. *Review of Economic Dynamics*, 16(2):L1–L13.

Kerr, S. P., Kerr, W., Özden, Ç., and Parsons, C. (2016). Global talent flows. *Journal of Economic Perspectives*, 30(4):83–106.

Klein, P. and Ventura, G. (2009). Productivity differences and the dynamic effects of labor movements. *Journal of Monetary Economics*, 56(8):1059–1073.

Krusell, P., Ohanian, L. E., Ríos-Rull, J.-V., and Violante, G. L. (2000). Capital-skill complementarity and inequality: A macroeconomic analysis. *Econometrica*, 68(5):1029–1053.

Lucas, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, 22(1):3–42.

Lucas, R. E. (2009). Trade and the diffusion of the industrial revolution. *American Economic Journal: Macroeconomics*, 1(1):1–25.

McFadden, D. L. (1974). Conditional logit analysis of qualitative choice behavior. *Frontiers in Econometrics, P. Zarembka (ed.)*, Chapter 2:105–142.

Meisenzahl, R. R. and Mokyr, J. (2011). The rate and direction of invention in the british industrial revolution: Incentives and institutions. In *The Rate and Direction of Inventive Activity Revisited*, pages 443–479. University of Chicago Press.

Mokyr, J. (2005). The intellectual origins of modern economic growth. *The Journal of Economic History*, 65(2):285–351.

Mokyr, J. and Voth, H.-J. (2009). Understanding growth in early modern europe. In *S. Broadberry and K. O'Rourke (Eds.)*. The Cambridge Economic History of Europe. Cambridge: Cambridge University Press.

Mookherjee, D. and Ray, D. (2003). Persistent inequality. *The Review of Economic Studies*, 70(2):369–393.

Moretti, E. (2004). Workers' education, spillovers, and productivity: evidence from plant-level production functions. *American Economic Review*, 94(3):656–690.

Mountford, A. and Rapoport, H. (2016). Migration policy, african population growth and global inequality. *World Economy*, 39(4):543–556.

Murphy, K. M., Riddell, W. C., and Romer, P. M. (1998). Wages, skills, and technology in the united states and canada. In *E. Helpman (ed.), General Purpose Technology and Economic Growth*. MIT Press, Cambridge, MA.

Murtin, F. and Wacziarg, R. (2014). The democratic transition. *Journal of Economic Growth*, 19(2):141–181.

Ngai, L. R. and Pissarides, C. A. (2007). Structural change in a multisector model of growth. *American Economic Review*, 97(1):429–443.

Ortega, F. and Peri, G. (2014). Openness and income: The roles of trade and migration. *Journal of International Economics*, 92(2):231–251.

Ottaviano, G. I. and Peri, G. (2012). Rethinking the effect of immigration on wages. *Journal of the European Economic Association*, 10(1):152–197.

Pritchett, L. (1997). Divergence, big time. *Journal of Economic Perspectives*, 11(3):3–17.

Ramsey, F. P. (1928). A mathematical theory of saving. *Economic Journal*, 38(152):543–559.

Restuccia, D. and Vandenbroucke, G. (2013). The evolution of education: A macroeconomic analysis. *International Economic Review*, 54(3):915–936.

Rodrik, D. (2013). Unconditional convergence in manufacturing. *The Quarterly Journal of Economics*, 128(1):165–204.

Sala-i Martin, X. (2006). The world distribution of income: falling poverty and convergence, period. *The Quarterly Journal of Economics*, 121(2):351–397.

Samir, K., Barakat, B., Goujon, A., Skirbekk, V., Sanderson, W., and Lutz, W. (2010). Projection of populations by level of educational attainment, age, and sex for 120 countries for 2005-2050. *Demographic Research*, 22(15):383–472.

Solow, R. M. (1956). A contribution to the theory of economic growth. *The Quarterly Journal of Economics*, 70(1):65–94.

Squicciarini, M. P. and Voigtländer, N. (2015). Human capital and industrialization: Evidence from the age of enlightenment. *The Quarterly Journal of Economics*, 130(4):1825–1883.

Ulubaşoğlu, M. A. and Cardak, B. A. (2007). International comparisons of rural–urban educational attainment: Data and determinants. *European Economic Review*, 51(7):1828–1857.

United Nations (2014). Probabilistic population projections based on the world population prospects: The 2012 revision. Population Division: New York.

Vollrath, D. (2009). How important are dual economy effects for aggregate productivity? *Journal of Development Economics*, 88(2):325–334.

World Bank (2007). World development report 2008: Agriculture for development. The World Bank: Washington.

Young, A. (2013). Inequality, the urban-rural gap, and migration. *The Quarterly Journal of Economics*, 128(4):1727–1785.

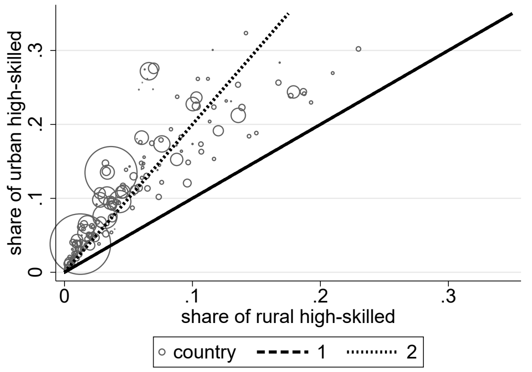
# A Appendix

## A.1 Calibration details

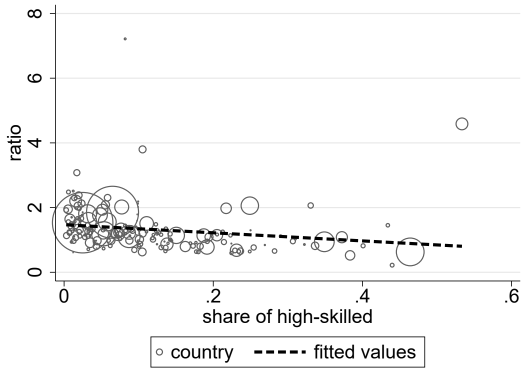
*Data from the Gallup World Polls.* – The data sources used to parameterize our model are described in Section 4. To identify the structure of income, fertility and migration intentions by region and by skill level, we use individual data from the Gallup World Poll (GWP) surveys. GWP covers about 150 countries between the years 2007 and 2016. For the majority of countries, the data are collected through face-to-face interviews. In some cases, interviews were conducted through phone calls. On average, the sample includes about 1,000 randomly selected respondents per year and per country. Data weighting is used to ensure a nationally representative sample for each country and is intended to be used for calculations within a country. To construct post-stratification weights, Gallup uses population statistics by gender, age, education or socioeconomic status, and region. The sampling frame is such that GWP data are representative of the entire population aged 15 and over (including populations from rural areas). However, in line with our model and with the macro databases used in the calibration, we only consider individuals aged 25 to 64. We aggregate the available 10 waves and assume they correspond to the year 2010 of our model. Hence, our income, fertility and migration intention proxies are drawn from about 10,000 responses per country.

*Estimated geography of skills.* – Figure A1 characterizes the geography of skills in the year 2010, and describes the worldwide evolution of urbanization and human capital between 1970 and 2010. Figure A1a shows that the urban share of college graduates is larger than the rural share in all countries. This is particularly true in poor countries. In line with Gollin et al. (2014), Figure A1b shows that the gap between regions decreases with the economy-wide proportion of college graduates. Figure A1c shows that the college-educated minority is predominantly and increasingly employed in the nonagricultural sector. As far as less educated workers are concerned (i.e., the large majority of people in the world), the fraction of them employed in the nonagricultural sector increased from 37.8% in 1970 to 50.5% in 2010. Figure A1d is the mirror image of Figure A1c: it depicts the evolution of the share of the college graduates in the labor force of each sector. On average, the world average proportion of college graduates increased from 2.4% to 8.8% between 1970 and 2010. In relative terms, the rise is greater in agriculture (from 1.1% to 3.9%) than in nonagriculture (from 4.6% to 13.1%). In absolute terms, the magnitude of the change is reversed; the small share of college-educated professionals and technicians in agriculture limits the capacity for innovation in poor countries (as argued in World Bank, 2007).

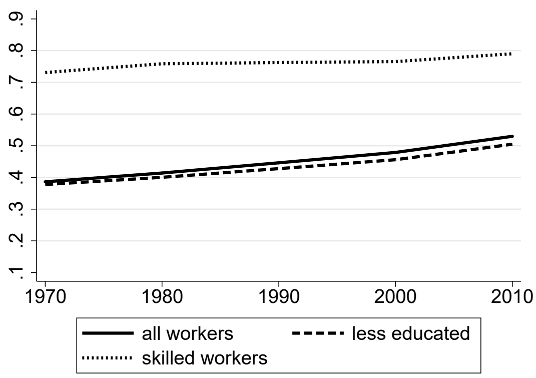
*Technology parameters.* – Figure A2 provides stylized facts on technological differences across countries, and summarizes the main findings of our calibration strategy. In line with the existing literature, we assume $\sigma_n = 2$ and $\sigma_a = \infty$. Once the elasticities are chosen, we use sector-specific data on returns to schooling to calibrate the relative productivity of college-educated workers. In the agricultural sector, we use the Gallup World Polls and compute the average household income per adult member as a function of the education level of the household head. As a proxy for the wage ratio in rural regions ($R^w_{a,t}$), we divide the average income of households with a college-educated household head by the average income of



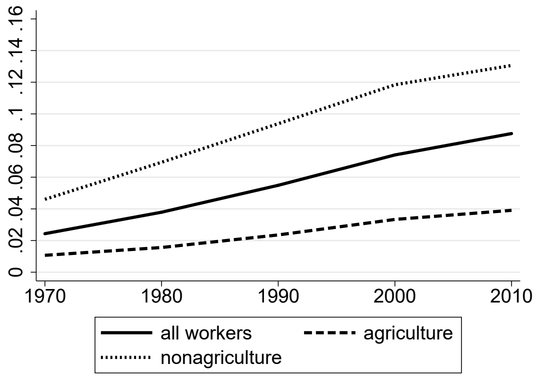

(a) Share of college graduates in agriculture ($H_{a,t}$) and nonagriculture ($H_{n,t}$) in 2010




(b) Regional ratio of skills ($H_{n,t}/H_{a,t}$) and national share of college graduates ($H_t$) in 2010




(c) World population share in nonagriculture by skill group




(d) World share of college graduates in population by sector

Figure A1: Additional stylized facts on the geography of skills

Note: In Figure A1a and A1b, bubble size is proportional to the population of the country.

households with a less educated household head. Combining (3) and (6), the elasticity of $R_a^w$ to $R_a^\ell$ is equal to $\kappa_a - 1/\sigma_a$. Assuming $\sigma_a = \infty$, this elasticity boils down to $\kappa_a$. Figure A2a shows that the correlation between $R_a^\varpi$ and $R_a^\ell$ is virtually nil. We thus rule out the possibility of skill-biased technical change in agriculture ($\kappa_a = 0$), and assume a linear technology with a constant $R_a^\varpi$ for all countries and all periods. The value of $R_a^\varpi$ is given by the population-weighted average of $R_a^w$, leading to $\varpi_a = 0.57$. We use this value for all countries and assume it is time-invariant.

As for the nonagricultural sector, we use data on the wage ratio from Biavaschi et al. (2016) for 143 countries.[28] We calibrate $R_n^\varpi$ using (3). Regressing $R_n^\varpi$ on $R_n^\ell$ yields a correlation of 0.38. Given the bidirectional causation relationship between the skill bias and education decisions, we consider this estimate as an upper bound for the skill-bias externality. In our baseline projections, we assume that half

[28]For the missing countries we predict the wage ratio using the estimated relationship between the log wage ratio on the log skill ratio.

the correlation is due to the skill-bias externality (i.e., $\kappa_n = 0.19$). Alternative scenarios are also considered in the simulation section. We calibrate $\overline{R}_n^{\varpi}$ as a residual from (6). Again, from (3) and (6), the elasticity of the $R_n^w$ to $R_n^{\ell}$ is equal to $\kappa_n - 1/\sigma_n$, which is equal to -0.37. Figure A2b shows that this elasticity is in line with the Gallup data on income per adult member.

In the *second step*, we use data on national Gross Domestic Product (GDP) for all countries from the Economic Research Service of the United States Department of Agriculture (USDA).[29] Data on the agriculture share in the value added are taken from the Food and Agriculture Organization of the UN (FAOSTAT).[30] We construct data on output by sector in the year 2010, and identify the TFP levels ($A_{r,t}$) by dividing the sector-specific output by the quantity of labor in efficiency unit using (1). There is a clear positive relationship between TFP and the share of college-educated workers in both sectors. Indeed, regressing the log of $A_{r,t}$ on the log of $R_{r,t}^{\ell}$ gives a coefficient of 0.57 in the nonagricultural sector, and 0.66 in agriculture, as shown in Figures A2c and A2d. Given the reverse causation relationship between productivity and education decision, we consider these estimates as upper bounds for the aggregate TFP externality. In our baseline scenario, we assume that half the correlation between TFP and the share of college-educated workers is due to the schooling externality (i.e., $\epsilon_n = 0.28$ and $\epsilon_a = 0.33$). Alternative scenarios are also considered in the simulation section. We calibrate $\overline{A}_n$ as a residual from (4).

Let us make two remarks on the calibration of the technology. First, Figure A2e and A2f show the distribution of $A_r$ and $\overline{A}_r$ in the agricultural and nonagricultural sector and for the year 2010. These distributions are relatively similar, meaning that a large fraction of TFP differences is explained by exogenous determinants. Remember that we assume a TFP externality equal to half of the correlation between TFP and the skill ratio. Second, the methodology used to calibrate the TFP parameters can be also used for the year 1980. Comparing the calibrated scale factors ($\overline{A}_n$) in 1980 and 2010, we obtain a high correlation of 0.78 and no sign of convergence or divergence (i.e., log changes in $\overline{A}_n$ are not significantly correlated with their initial level). It follows that we can reasonably consider these scale factors as time-invariant in our numerical experiments.

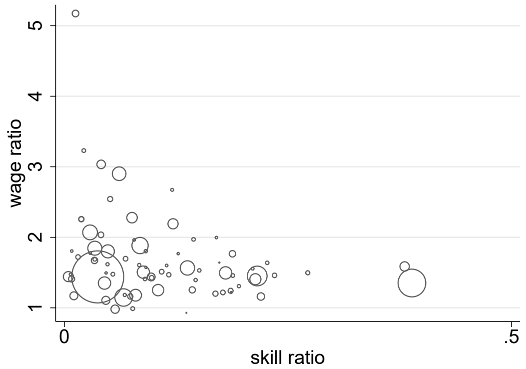
*Preference parameters.* – We assign the following values to the parameters that are time-invariant and equal for all countries: $\theta = 0.25$, $\lambda = 0.5$ and $\phi = 0.14$.[31] From (14) and (16), the scale parameter of the distribution of migration tastes ($\mu$) is the inverse of the elasticity of bilateral migration to the wage rate. Bertoli and Fernández-Huertas Moraga (2013) find a value between 0.6 and 0.7 for this elasticity. Hence, we use $\mu = 1.4$.

Let us now explain how we calibrate the values of $\pi_r$ and $\psi_{r,t}$. These two parameters are country- and sector-specific, and affect the fertility and education

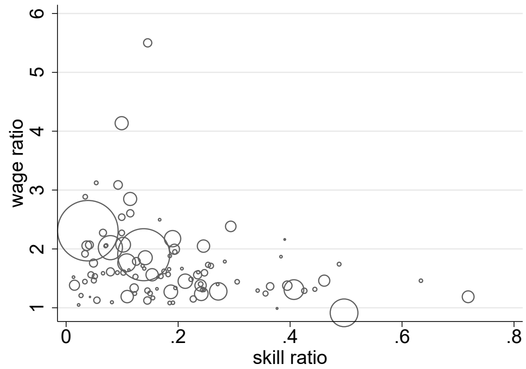
[29]For a few missing observations we impute values by making use of the Maddison data base and data from the World Bank.

[30]For a few missing observations we impute values by making use of data from the World Bank. Since data is volatile for several countries, the average of five data points around the data point of interest is used.

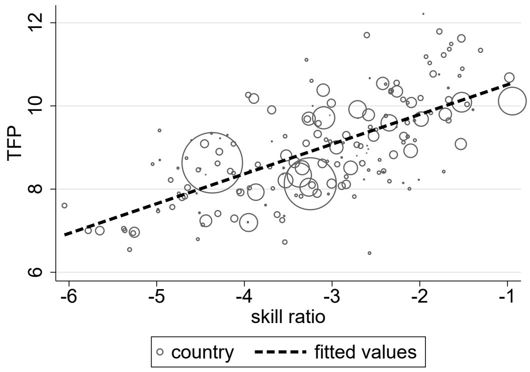
[31]Given the expression in (10), this assumptions reflects setting the bound of the maximal number of children equal to 7 (i.e., 14 children per couple). See Docquier et al. (2017) for a brief review of studies using similar parameter values.



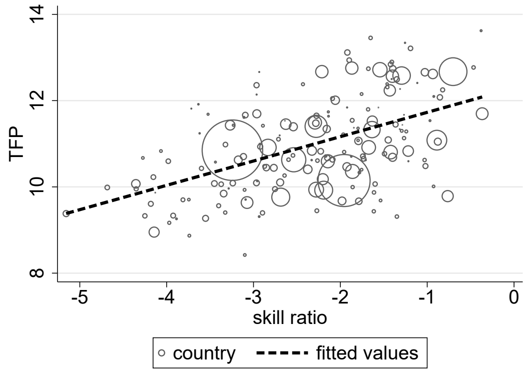

(a) Correlation between skill ratio ($R_a^\ell$) and wage ratio ($R_a^w$) in agriculture



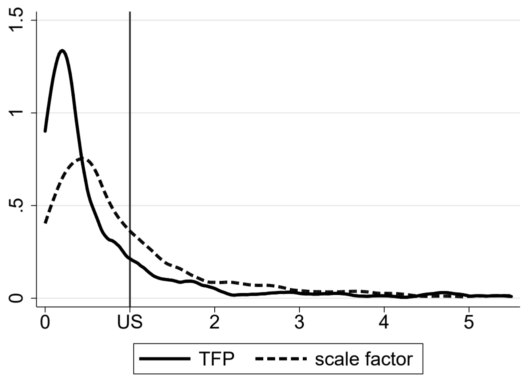

(b) Correlation between skill ratio ($R_n^\ell$) and wage ratio ($R_n^w$) in nonagriculture



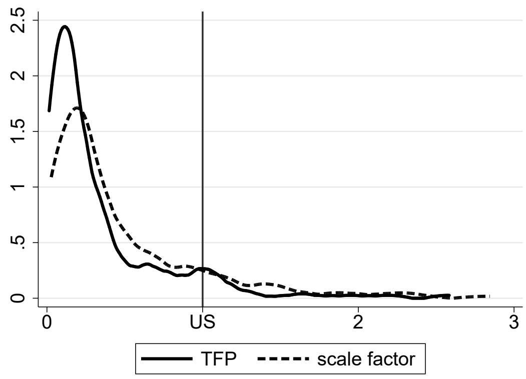

(c) Correlation between skill ratio ($log(R_a^\ell)$) and TFP ($log(A_a)$)




(d) Correlation between skill ratio ($log(R_n^\ell)$) and TFP ($log(A_n)$)




(e) Kernel density of TFP ($A_a$) and its scale factor ($\overline{A}_a$) in agriculture




(f) Kernel density of TFP ($A_n$) and its scale factor ($\overline{A}_n$) in nonagriculture

Figure A2: Calibration of the technological parameters in 2010

Notes: In Figures A2a-A2d, bubble size is proportional to the population of the country. Figures A2e and A2f assume that the elasticity of TFP or skill bias to the skill ratio is equal to 50% of the correlation between these variables.

decisions. We calibrate them to match the population dynamics between the years 1980 and 2010, i.e., the transition from the resident population in 1980 and the native population in 2010. We begin by estimating the size of the *before-migration* population in 2010 by skill group ($\sum_r N_{r,s,2010}$). We do this by adding the number of international migrants by region and skill level to the respective number of high-skilled and low-skilled workers by region of our basic data set, the after-migration population ($L_{r,s,2010}$). For simplicity, we focus on international migration to OECD countries only. From the Database on Immigrants in OECD and non-OECD countries (DIOC), we extract the number of emigrants by education level to OECD countries for all countries in our sample and for the year 2010. The DIOC does not identify the region of origin of migrants (urban vs. rural). However, for the majority of countries in our sample, skill- and region-specific information on the desire to emigrate can be extracted from the Gallup World Polls. Assuming the structure of migration aspirations is reflected in actual emigration stocks, we split the number of emigrants to OECD countries by region of origin and by education level.[32] The average fertility rate ($\overline{n}_{1980}$) is thus obtained by dividing the total native population of adults in 2010 ($\sum_{r,s} N_{r,s,2010}$) by the total resident population of adults in 1980 ($\sum_{r,s} L_{r,s,1980}$).[33] Moreover, our calibration requires data on the skill- and region-specific fertility for each country. By construction, we have $\overline{n}_t \equiv \sum_{r,s} L_{r,s,t} n_{r,s,t} / \sum_{r,s} L_{r,s,t}$. We use the Gallup World Polls and extract the Gallup-based average number of children per household by region and skill level for 2010.[34] We compute the fertility of the college educated workers by fitting the sector-specific low/high-skilled fertility differentials from the Gallup database. In this way, we obtain the fertility rates for each country for the year 1980. From 2010 onwards, the number of children is endogenous.

The last moment to fit in the procedure is the number of internal migrants between the years 1980 and 2010. Two factors may determine the difference in the evolution of skills in both sectors. First, this evolution may be brought about by the differences in educational prospects (given the already computed fertility differential). Second, it might be caused by the selectivity of rural-to-urban migrants. We decided to pin down the first of the two factors. This draws on the different probabilities to become high-skilled in urban and rural areas. These probabilities are calibrated by assuming a log-normal distribution of years of schooling in both sectors. The location parameters simply match the mean years of schooling in rural/urban areas, while the dispersion parameter is identical across sectors and is set to fit the country-specific share of high-skilled individuals (defined as the percentage of population with more than 17 years of schooling). Finally, the quested ratio of probabilities is the quotient of two respective probabilities of obtaining more than 17 years of schooling, derived from region-specific distributions. We set the ratio of the probabilities so that net internal migration is computed as a residual in the model. We arbitrary impose that the process of urbanization is the dominant one (which is the case in almost all countries). The matched number

[32] Bertoli and Ruyssen (2018) show that aspirations to emigrate are correlated with emigration flows within five years.

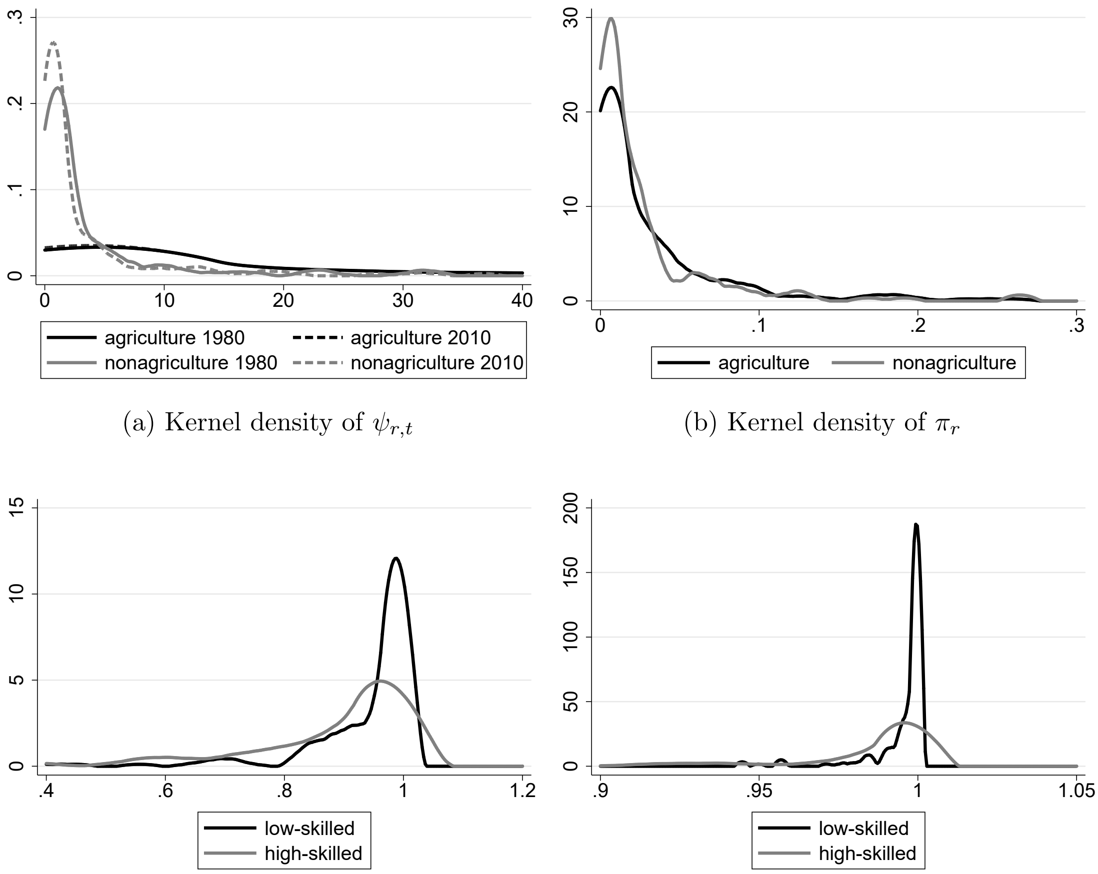
[33] There is no mortality in the model. The average fertility rate at time $t$, $\overline{n}_t$, should be seen as a net population growth rate. Note that the average fertility rate is not affected by internal migration, so that we need to only account for international migration at this stage.

[34] We only include countries with at least ten respondents. When data are missing, crude birth rates from the World Health Organization are used.

represents the net migration from rural to urban region. The net internal migration is then the difference between the "before-migration" population ($N_{r,s,2010}$) in 2010 and the sum of the resident population and the international migrants ($\sum_{r,s}(L_{r,s,2010} + M_{rf,s,2010})$) in 2010. In this way, the model perfectly matches the skill and regional distribution of workers in 1980 and 2010.

From Eq. (13), the fertility rate in the model depends on the product of $\pi_r\psi_{r,t}$. Once fertility rates are matched we are able to identify the product $\pi_r\psi_{r,t}$. We then calibrate $\pi_r$ and $\psi_{r,t}$ in order to match the educational structure of the native population in 2010, imposing the given value to the ratio of probabilities of becoming high-skilled across regions. Figures A3a and A3b show the distributions of $\pi_r$, $\psi_{r,t}$ for the two regions. Figure A3a depicts the distributions for two periods (1980 and 2010). The distribution of $\pi_r$ is stable over time. As far as $\psi_{r,t}$ is concerned, the mean levels decreased between 1980 and 2010, reflecting expansive education policies that can be related to the Millennium Development Goals. As for internal migration costs, we assume there is only migration from rural to urban regions (i.e., $x_{an,s,t} < 1$ and $x_{na,s,t} = 1$). We obtain internal migration costs for rural-urban migration from Eq. (16). Figure A3c shows that moving costs are usually smaller for highly educated workers than for the less educated.

In order to determine the international migration costs ($x_{af,s,t}$ and $x_{nf,s,t}$), we begin by retrieving the utilities achievable abroad. We set these utilities equal to the skill-specific weighted average utilities of the OECD countries. The weights consist in the respective population sizes of the OECD countries. We then obtain the international migration costs from Eq. (16). In line with Figure A3c, Figure A3d shows that international migration costs are smaller for college-educated workers.




(a) Kernel density of $\psi_{r,t}$

(b) Kernel density of $\pi_r$

(c) Kernel density of $x_{an,s}$

(d) Kernel density of $x_{af,s}$

Figure A3: Calibration of the preference parameters in 1980 and 2010

## A.2 Geography of skills and current income inequality (static experiments)

In Section 5.1, we consider the US as the base-case economy and proceed with three static counterfactual experiments to quantify the role of skills accumulation in the year 2010. Figure 2 in the main text describes the changes in income per capita induced by an increase in the average skill level, or by a better geographical allocation of national skills. Results are presented in a different manner in the development accounting literature. For example, Jones (2014) uses the concept of success rate ($SR$), defined as the share of the income ratio explained by the counterfactual. In other words, $SR$ equals one minus the counterfactual-to-observed ratio of income with the US (i.e., \$100,000 per year). Equivalently, the success rate measures the national income loss due to the lower level of human capital and/or to the sectoral allocation of workers when compared to the US:

$$SR = 1 - \frac{y_{US}/y_{CF}}{y_{US}/y_{obs}} = \frac{y_{CF} - y_{obs}}{y_{CF}}.$$

In development accounting studies, the success rate is usually provided for selected countries located at various percentiles of the income distribution. Table A1 describes our static simulation results likewise. This table also reports the Theil index for each of the counterfactual experiments. For the baseline, the Theil index takes a value of 0.744. In case the US shares are transposed, the value falls to 0.354. If the US urban shares are transposed, the index takes a value of 0.607. With the repatriation of emigrant workers the Theil index is very similar to the baseline with a value of 0.735. In addition, Table A2 decomposes the aggregate results by sector.

Table A1: Geography of skills and income per worker in 2010

| | $15^{th}$ (Cambodia) | $25^{th}$ (Ghana) | $50^{th}$ (Tunisia) | $75^{th}$ (Mexico) | $85^{th}$ (Greece) | Theil Index |
|---|---|---|---|---|---|---|
| I. Observed levels and ratios of income per worker | | | | | | |
| Income pw | 2,018 | 3,651 | 9,032 | 20,761 | 55,262 | 0.744 |
| US/ctry ratio | 50.7 | 28.0 | 11.3 | 4.9 | 1.9 | - |
| II. Counterfactual: Transposing the US skill shares in each sector | | | | | | |
| Income pw | 16,010 | 13,709 | 19,968 | 34,968 | 63,328 | 0.354 |
| US/ctry ratio | 6.4 | 7.5 | 5.2 | 2.9 | 1.6 | - |
| Success | 0.873 | 0.732 | 0.544 | 0.402 | 0.121 | 0.525 |
| III. Counterfactual II with exogenous TFP ($A_r$) and exogenous skill bias ($R_r^\omega$) | | | | | | |
| Income pw | 5,300 | 5,932 | 9,186 | 21,936 | 42,087 | 0.488 |
| US/ctry ratio | 16.9 | 15.1 | 9.7 | 4.1 | 2.1 | - |
| Success | 0.668 | 0.463 | 0.141 | 0.174 | -0.146 | 0,345 |
| IV. Counterfactual II with full TFP externality ($A_r$) and full skill bias externality ($R_r^\omega$) | | | | | | |
| Income pw | 58,737 | 37,430 | 50,151 | 61,843 | 103,421 | 0.267 |
| US/ctry ratio | 2.1 | 3.3 | 2.5 | 2.0 | 1.2 | - |
| Success | 0.958 | 0.881 | 0.781 | 0.592 | 0.350 | 0.642 |
| V. Counterfactual: Transposing the US urbanization share | | | | | | |
| Income pw | 4,681 | 4,028 | 10,007 | 20,785 | 54,482 | 0.607 |
| US/ctry ratio | 22.0 | 25.6 | 10.3 | 5.0 | 1.9 | - |
| Success | 0.566 | 0.087 | 0.091 | -0.006 | -0.021 | 0.184 |
| VI. Counterfactual V with exogenous TFP ($A_r$) and exogenous skill bias ($R_r^\omega$) | | | | | | |
| Income pw | 3,469 | 3,490 | 6,377 | 16,259 | 37,709 | 0.594 |
| US/ctry ratio | 25.2 | 25.1 | 13.7 | 5.4 | 2.3 | - |
| Success | 0.503 | 0.106 | -0.211 | -0.092 | -0.253 | 0.202 |
| VII. Counterfactual V with full TFP externality ($A_r$) and full skill bias externality ($R_r^\omega$) | | | | | | |
| Income pw | 6,351 | 4,710 | 15,740 | 26,678 | 79,050 | 0.626 |
| US/ctry ratio | 19.1 | 25.8 | 7.7 | 4.6 | 1.5 | - |
| Success | 0.623 | 0.080 | 0.319 | 0.076 | 0.170 | 0.159 |
| VIII. Counterfactual: Repatriation of emigrant workers | | | | | | |
| Income pw | 2,481 | 4,164 | 9,789 | 22,911 | 57,745 | 0.735 |
| US/ctry ratio | 41.8 | 24.9 | 10.6 | 4.5 | 1.8 | - |
| Success | 0.176 | 0.112 | 0.065 | 0.082 | 0.031 | 0.012 |
| IX. Counterfactual VIII with exogenous TFP ($A_r$) and exogenous skill bias ($R_r^\omega$) | | | | | | |
| Income pw | 1,290 | 2,988 | 5,815 | 17,750 | 38,584 | 0.772 |
| US/ctry ratio | 68.1 | 29.4 | 15.1 | 5.0 | 2.3 | - |
| Success | -0.343 | -0.049 | -0.334 | -0.004 | -0.230 | -0.038 |
| X. Counterfactual VIII with full TFP externality ($A_r$) and full skill bias externality ($R_r^\omega$) | | | | | | |
| Income pw | 4,775 | 5,855 | 16,495 | 29,606 | 86,428 | 0.707 |
| US/ctry ratio | 25.6 | 20.9 | 7.4 | 4.1 | 1.4 | - |
| Success | 0.495 | 0.254 | 0.345 | 0.162 | 0.236 | 0.049 |

Table A2: Productivity by sector - Development accounting

| | $15^{th}$ (Cambodia) | $25^{th}$ (Ghana) | $50^{th}$ (Tunisia) | $75^{th}$ (Mexico) | $85^{th}$ (Greece) | $99^{th}$ (US) |
|---|---|---|---|---|---|---|
| I. Observed levels and ratios of income per worker | | | | | | |
| Income pw (n) | 7,169 | 5,020 | 12,904 | 25,726 | 68,259 | 125,133 |
| Income pw (a) | 807 | 1,987 | 1,707 | 4,310 | 15,026 | 10,214 |
| US/ctry ratio (n) | 17.5 | 24.9 | 9.7 | 4.9 | 1.8 | 1.0 |
| US/ctry ratio (a) | 12.7 | 5.1 | 6.0 | 2.4 | 0.7 | 1.0 |
| College grads in n | 0.036 | 0.030 | 0.105 | 0.148 | 0.297 | 0.326 |
| II. Counterfactual: Transposing US skill shares in each sector | | | | | | |
| Success (n) | 0.624 | 0.671 | 0.458 | 0.378 | 0.060 | - |
| Success (a) | 0.743 | 0.726 | 0.548 | 0.468 | 0.385 | - |
| III. Counterfactual II with exogenous TFP ($A_r$) and exogenous skill bias ($R_r^{\omega}$) | | | | | | |
| Success (n) | 0.032 | 0.371 | -0.011 | 0.151 | -0.206 | - |
| Success (a) | -0.136 | 0.139 | -0.103 | 0.060 | 0.091 | - |
| IV. Counterfactual II with full TFP externality ($A_r$) and full skill bias externality ($R_r^{\omega}$) | | | | | | |
| Success (n) | 0.873 | 0.845 | 0.737 | 0.569 | 0.293 | - |
| Success (a) | 0.942 | 0.913 | 0.815 | 0.699 | 0.585 | - |
| V. Counterfactual: Transposing the US urbanization share | | | | | | |
| Success (n) | -0.286 | -0.132 | -0.085 | -0.051 | -0.096 | - |
| Success (a) | 0.121 | 0.143 | 0.175 | 0.178 | 0.305 | - |
| VI. Counterfactual V with exogenous TFP ($A_r$) and exogenous skill bias ($R_r^{\omega}$) | | | | | | |
| Success (n) | -0.458 | -0.074 | -0.436 | -0.129 | -0.325 | - |
| Success (a) | -0.267 | -0.010 | -0.207 | 0.021 | 0.074 | - |
| VII. Counterfactual V with full TFP externality ($A_r$) and full skill bias externality ($R_r^{\omega}$) | | | | | | |
| Success (n) | -0.133 | -0.193 | 0.181 | 0.022 | 0.094 | - |
| Success (a) | 0.390 | 0.273 | 0.436 | 0.337 | 0.478 | - |
| VIII. Counterfactual: Repatriation of emigrant workers | | | | | | |
| Success (n) | 0.090 | 0.116 | 0.019 | -0.027 | -0.006 | - |
| Success (a) | 0.136 | 0.151 | 0.110 | 0.035 | 0.103 | - |
| IX. Counterfactual VIII with exogenous TFP ($A_r$) and exogenous skill bias ($R_r^{\omega}$) | | | | | | |
| Success (n) | -0.461 | 0.009 | -0.394 | -0.118 | -0.283 | - |
| Success (a) | -0.267 | -0.010 | -0.214 | -0.041 | 0.046 | - |
| X. Counterfactual VIII with full TFP externality ($A_r$) and full skill bias externality ($R_r^{\omega}$) | | | | | | |
| Success (n) | 0.434 | 0.211 | 0.309 | 0.056 | 0.211 | - |
| Success (a) | 0.411 | 0.286 | 0.348 | 0.105 | 0.157 | |

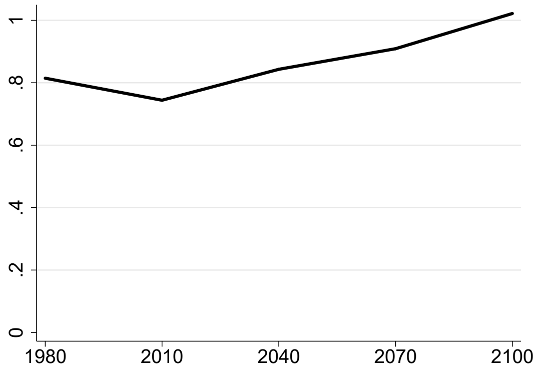
Notes: Tables A1 and A2 give the level of income per worker of Cambodia, Ghana, Tunisia, Mexico, and Greece for the baseline and the respective counterfactual scenario. Part I reports the observed level of income per worker and the US-to-country ratio. Part II reports the income levels and ratios obtained if the US shares were observed in each sector. Part V reports the income levels and ratios obtained if the US urbanization share was transposed. Part VIII reports the income levels and ratios obtained if emigration rates were nil. The remaining parts are variants of the respective scenario with different assumptions on the technological externalities. For each simulation, the success rate is the share of the wage ratio explained by the counterfactual, i.e., one minus the counterfactual-to-observed ratio of income differential with the US (in col. 2-6), and one minus the counterfactual-to-observed ratio of Theil index (in col. 7).

## A.3 Baseline prospects: geopolitical implications

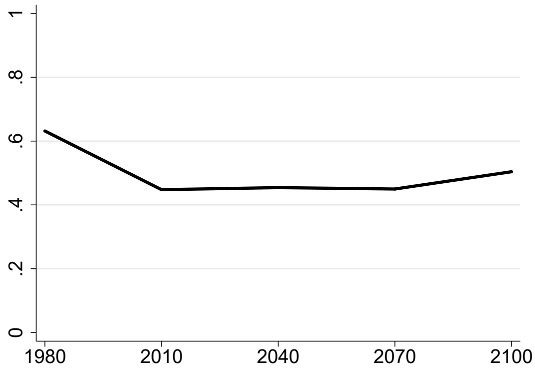
We examine the main geopolitical implications of the baseline projections described in Section 5.2. The model does not predict convergence in income per worker and in the share of college graduates across countries. The Theil index of human capital inequality remains almost stable over the 21st century. It ranges from 0.63 in 1980 to 0.5 in 2100 as illustrated in Figure A4b. Similarly, income per capita does not converge. On the contrary, the Theil index of income inequality varies from 0.81 in 1980 to 1.02 in 2100 as depicted in Figure A4a.

Figure A4c depicts the evolution of the region/continent shares in the worldwide working-age population. The share of sub-Saharan Africa increases from 7.2% in 1980 to 34.0% in 2100. The share of OECD countries decreases from 25.8% to 13.0% over the same period of time. In addition, the OECD share in the college-educated population shrinks markedly, as illustrated in Figure A4d. This is caused by the progress in higher education in the other regions, in particular in Asia, and by the rise of the demographic share of the developing world. Figure A4e shows that the speed of urbanization is faster in Africa than in the other regions. Finally, Figure A4f depicts the evolution of income shares. The OECD income share decreases by more than 13 percentage points (from 77.4% in 1980 to 64.1% in 2100) whereas the Asian share increases from 9.1% to 17.0% over the same period.

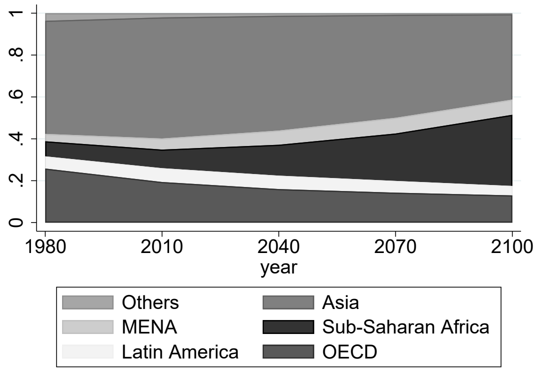
Table A3 describes the international migration implications of our baseline projections. Assuming constant migration policies, we predict slight decreases in future emigration rates from the OECD member states. On the contrary, emigration rates from Latin America, from the Middle East and North Africa, from sub-Saharan Africa and from Asia increase. This is due to the rising share of college-educated workers (the most mobile individuals) in the population. Given its rising share in the world population, sub-Saharan Africa is responsible for drastic changes in worldwide migration pressures. As a result, the proportion of foreigners increases in European countries. In particular, the average immigration rate to the EU15 is expected to rise from 13.6% in 2010 to 21.2% in 2100. This is explained by four factors: (i) Europe is the main destination for African emigrants; (ii) the demographic ratio between Africa and Europe increases sharply; (iii) college-educated workers are more mobile than the less educated and the rise in African human capital has limited effects on income disparities between Africa and Europe; (iv) urbanization increases and international migration costs are lower for urban citizens than for villagers. Reinforcement of immigration restrictions are likely to be observed in European countries to curb the migration pressure; their implications are investigated in Section 5.4. Note that the share of immigrants increases less drastically in the US (from 16.0% to 19.2%), Australia (from 24.9% to 25.9%) and Canada (from 18.7% to 25.0%).



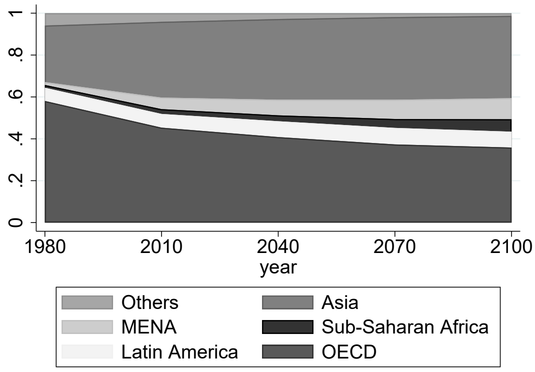

(a) Theil index of income inequality



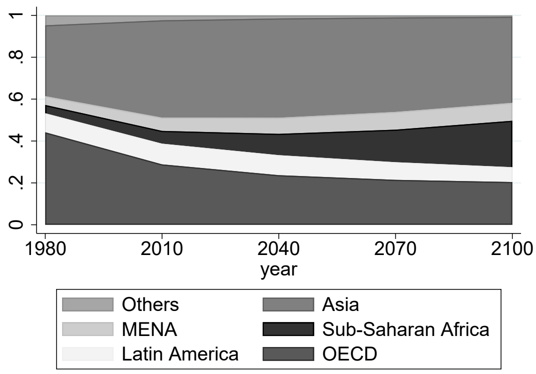

(b) Theil index of inequality in skills



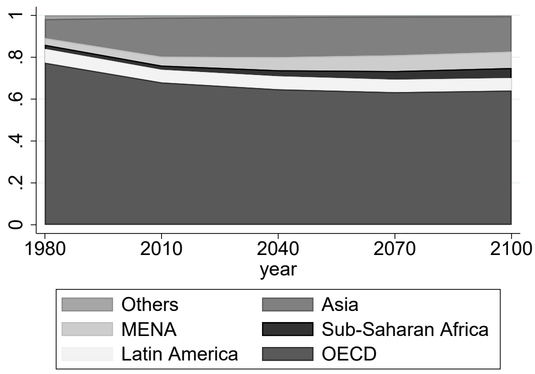

(c) Share of working age population




(d) Share of college-educated workers




(e) Share of urban population




(f) Share of GDP

Figure A4: Global inequality and regional shares (1980-2100)

Notes: This figure reports the Theil index of income inequality, the Theil index of inequality in the share of skilled workers, the regional shares of global labor force, high-skilled workers, urban workers and GDP. In Figures A4c-A4f countries are exclusively and completely assigned to one of six groups: OECD, Latin America, sub-Saharan Africa, Middle East and North Africa (MENA), Asia and Others

Table A3: Projections of immigration and emigration rates

| | Baseline scenario | | | | Half | Linear | No internal |
|---|---|---|---|---|---|---|---|
| | 2010 | 2040 | 2070 | 2100 | 2100 | 2100 | 2100 |
| Emigration rates (as percent of native population) | | | | | | | |
| OECD | 4.3% | 4.4% | 4.2% | 3.7% | 3.9% | 4.0% | 3.7% |
| LAC | 3.9% | 4.5% | 4.9% | 5.2% | 5.4% | 5.1% | 6.1% |
| SSA | 1.4% | 1.5% | 1.7% | 1.7% | 1.7% | 2.2% | 7.2% |
| MENA | 2.9% | 3.6% | 3.8% | 3.9% | 3.8% | 4.0% | 4.9% |
| Asia | 1.1% | 1.4% | 1.8% | 2.1% | 1.9% | 2.0% | 1.6% |
| Others | 13.9% | 15.1% | 16.0% | 16.3% | 16.6% | 16.8% | 17.1% |
| Immigration rates (as percent of resident population) | | | | | | | |
| EU | 12.1% | 16.1% | 18.7% | 20.1% | 21.8% | 20.5% | 29.4% |
| EU 15 | 13.6% | 17.7% | 20.1% | 21.2% | 22.9% | 21.6% | 30.4% |
| *GER* | *15.0%* | *19.0%* | *21.2%* | *22.0%* | *24.0%* | *22.6%* | *32.4%* |
| *FRA* | *12.2%* | *15.9%* | *18.4%* | *19.7%* | *21.3%* | *20.1%* | *28.9%* |
| *GBR* | *14.6%* | *20.0%* | *23.2%* | *24.4%* | *25.4%* | *24.4%* | *30.0%* |
| *ITA* | *10.9%* | *14.5%* | *16.9%* | *18.3%* | *20.4%* | *19.0%* | *29.5%* |
| *ESP* | *14.0%* | *17.3%* | *19.1%* | *19.8%* | *21.7%* | *20.4%* | *29.2%* |
| USA | 16.0% | 18.6% | 19.5% | 19.2% | 21.0% | 19.8% | 27.3% |
| CAN | 18.7% | 23.0% | 24.9% | 25.0% | 25.8% | 24.9% | 29.2% |
| AUS | 24.9% | 27.0% | 26.9% | 25.9% | 27.3% | 26.2% | 32.9% |

Notes: The upper part of the table gives the share of emigrants in the total native population for the OECD, Latin America and the Caribbean (LAC), sub-Saharan Africa (SSA), Middle East and North Africa (MENA), Asia, and Others. The bottom part of the table gives the share of immigrants in the working-age population for the European Union (EU), the 15 countries of the European Union (EU 15), Germany (GER), France (FRA), Great Britain (GBR), Italy (ITA), Spain (ESP), the United States (USA), Canada (CAN), and Australia (AUS). The first to fourth columns give the respective values for the baseline scenario for the years 2010-2100. Column "Half" gives the respective values for the counterfactual scenario where the coefficients of the (baseline) quadratic convergence equation are divided by two for the year 2100. Column "Linear" gives the respective values for the counterfactual scenario with the linear convergence in education costs for the year 2100. Column "No internal" gives the respective values for the counterfactual scenario with no internal mobility for the year 2100.

## A.4 Sensitivity to technological externalities and to the preference structure

We assess the sensitivity of our socio-demographic projections to modelling assumptions. Firstly, we assess the extent to which technological externalities influence our socio-demographic and income projections. The static counterfactual experiments conducted in Section 5.1 show that the effect of human capital on global inequality quantitatively depends on the size of technological externalities.

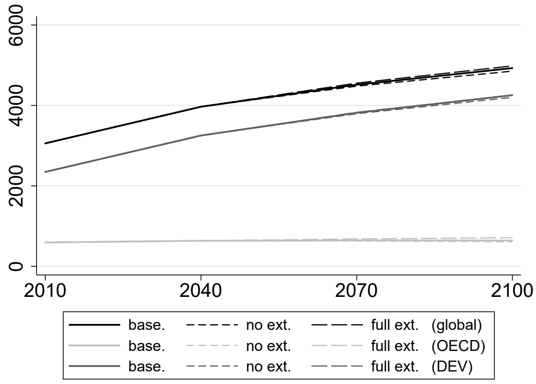
Figure A5 compares the baseline trajectories of population, education and urbanization with those obtained without or with full externalities. The evolution of socio-demographic variables is highly robust to the technological environment. The only exception is the share of college graduates in OECD countries, which depends on the intensity directed technical change. With full externalities, the skill premium and the cost of education increase. This makes access to education more difficult for poor households. At the world level, technological externalities have a negligible effects on future demographic pressures, urbanization and human capital accumulation.

Secondly, we challenge the assumption of homogenous consumption goods produced across sectors and of the homothetic preference structure. It is well documented in the macroeconomic literature on structural change that relative prices (Ngai and Pissarides, 2007) and income effects (Foellmi and Zweimüller, 2008) can influence consumption choices and welfare. In our framework, urbanization and human capital accumulation affect the quantity of goods produced in the agricultural and nonagricultural sectors, with potential implications on relative prices. In particular, if the relative price of agricultural goods increases, this may attenuate the process of urbanization and increasing access to education.

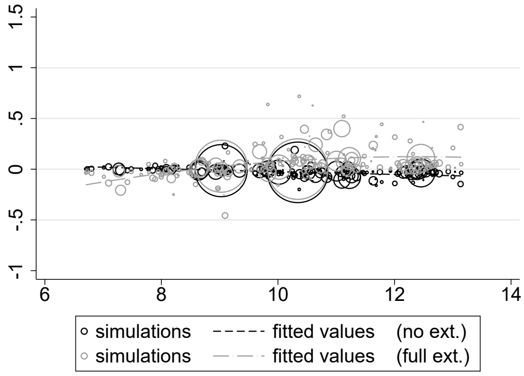
To investigate this mechanism, we extend our model and rely on the preference structure described in Boppart (2014). In each region, we assume that the utility of total consumption is a nonlinear transformation of the quantity of agricultural ($c^a_{r,s,t}$) and nonagricultural ($c^n_{r,s,t}$) goods produced in the country: $c_{r,s,t} = (c^a_{r,s,t})^\alpha + c^n_{r,s,t}$. We thus disregard trade, which would attenuate the average relative price variations. More precisely, consumption of agriculture goods is subject to diminishing marginal utility, as long as $\alpha \in (0, 1)$. Knowing that each good is characterized by a separate price level (with $p^a$ serving as a numeraire in each country), and the total consumption expenditure is labeled by $c_{r,s,t}$ (as in Eq. (10)), one can solve for the share of each good in consumption basket, e.g. for the agricultural good: $c^a_{r,s,t}/c_{r,s,t} = \alpha^{\frac{1}{1-\alpha}}(p^n_{r,t})^{\frac{1}{1-\alpha}} c^{-1}_{r,s,t}$. The latter resembles Eq. (20) in Boppart (2014), which is then structurally estimated to retrieve the value of $\alpha$. According to his regressions, $\alpha \approx 0.67$, which we take as the reference value of the non-homotheticity parameter in individuals' utility. We also consider a scenario with a smaller substitutability between goods (i.e., $\alpha = 0.50$).

Figure A6 compares the baseline trajectories of population, education and urbanization with those obtained with imperfectly substitutable goods and non homothetic preferences (using $\alpha = 0.67$ and $\alpha = 0.50$). Again, the evolution of socio-demographic variables is highly robust to preferences. At the world level, accounting for imperfect substitution as in Boppart (2014) has negligible effects on future demographic pressures, urbanization and human capital accumulation.

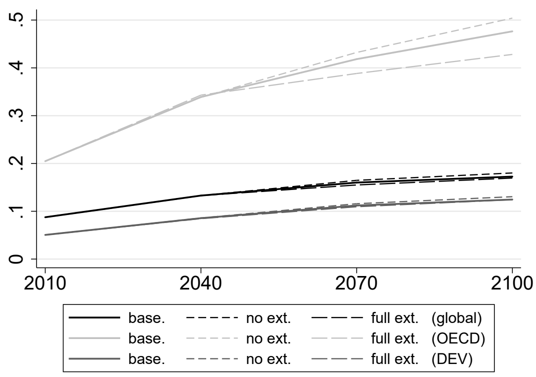
On Figure A7, we illustrate the effect of our modelling assumption on income



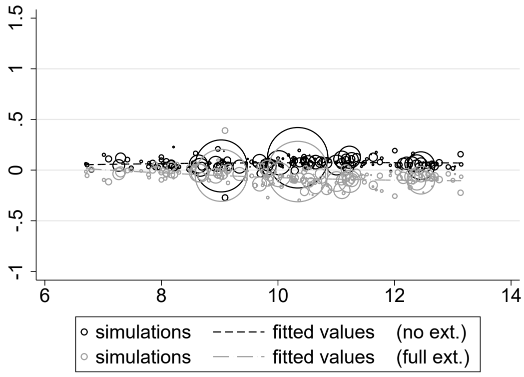

(a) Population (in million of people)



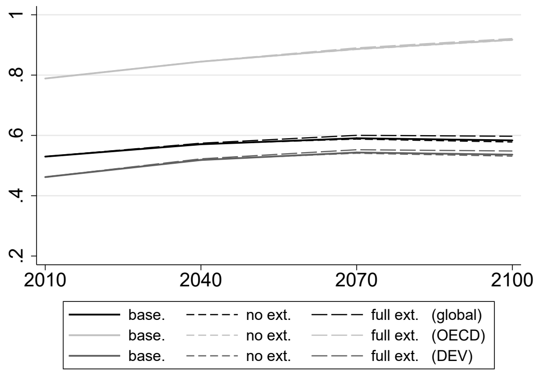

(b) Relative deviations from the baseline in 2100



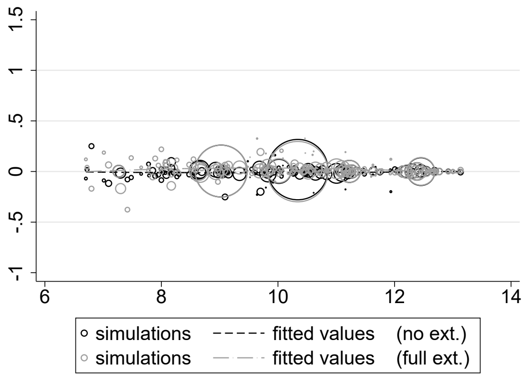

(c) Share of college educated workers




(d) Relative deviations from the baseline in 2100



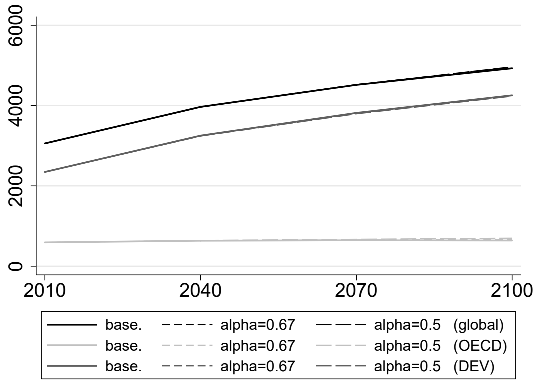

(e) Share of urban population



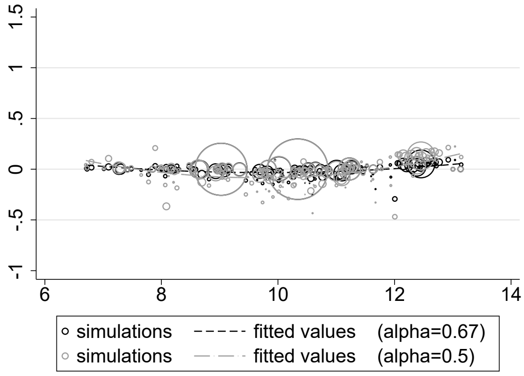

(f) Relative deviations from the baseline in 2100

Figure A5: Sensitivity to technological scenarios

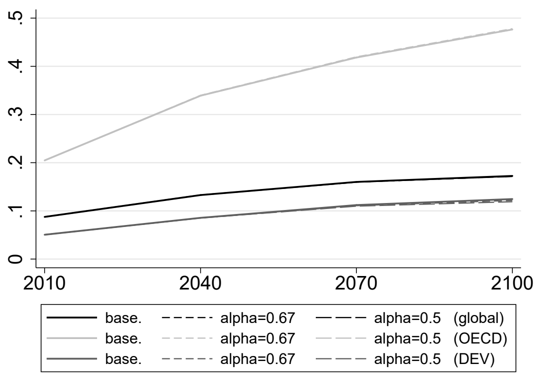
Notes: This figure reports the projected population size, the share of college educated workers, and the share of urban population for the baseline and the respective counterfactual scenario. The scenario "no ext." refers to the scenario with no technological externalities. The scenario "full ext." refers to the scenario with full technological externalities.



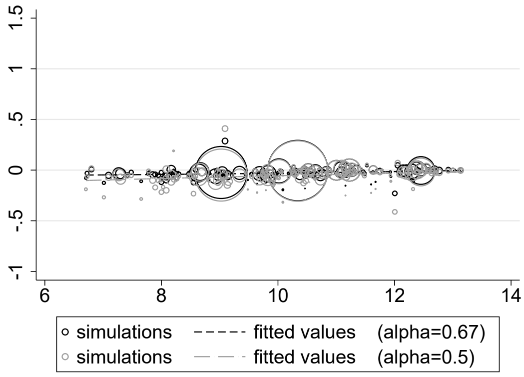

(a) Population (in million of people)



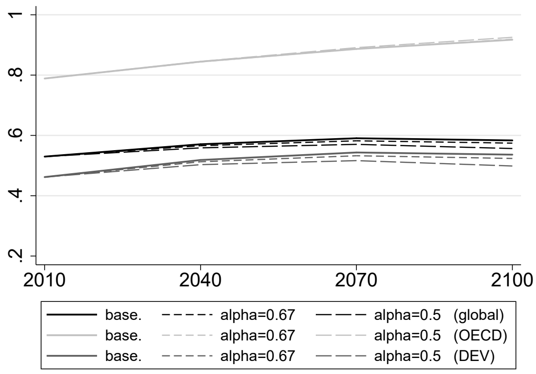

(b) Relative deviations from the baseline in 2100



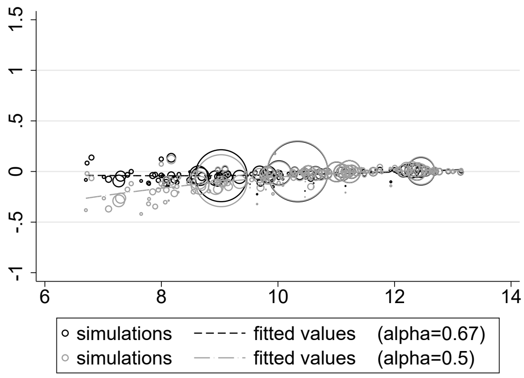

(c) Share of college educated workers




(d) Relative deviations from the baseline in 2100




(e) Share of urban population

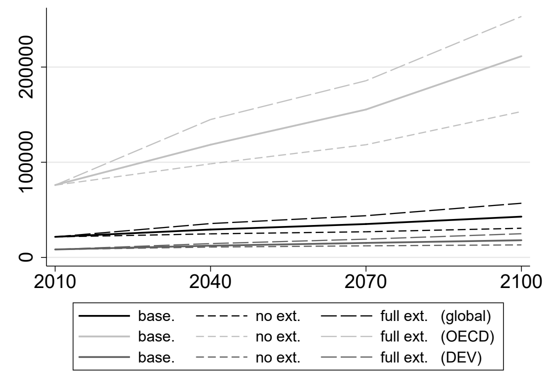



(f) Relative deviations from the baseline in 2100

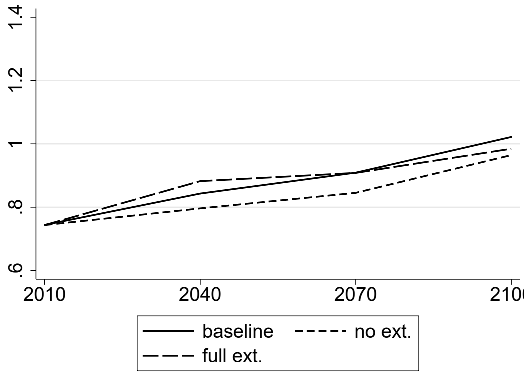
Figure A6: Sensitivity to preference structures

Notes: This figure reports the projected population size, the share of college educated workers, and the share of urban population for the baseline and the respective counterfactual scenario. The scenario "alpha=0.67" refers to the scenario with imperfectly substitutable goods and non homothetic preferences and a value of 0.67 for $\alpha$. The scenario "alpha=0.5" refers to the scenario with imperfectly substitutable goods and non homothetic preferences and a value of 0.5 for $\alpha$.

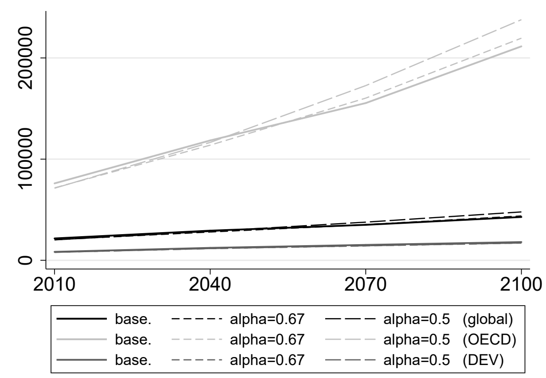
inequality. Figures A7a and A7b assess the sensitivity of the income distribution to the size of technological externalities. We find that the size of technological externalities affects the trajectory of income per capita in developing and developed countries, but has negligible effect on income inequality. Figures A7c and A7d assess the sensitivity of global inequality to the structure of preferences. Assuming imperfectly substitutable goods and non-homothetic preferences has little effect on the levels of income per capita and on the Theil index.



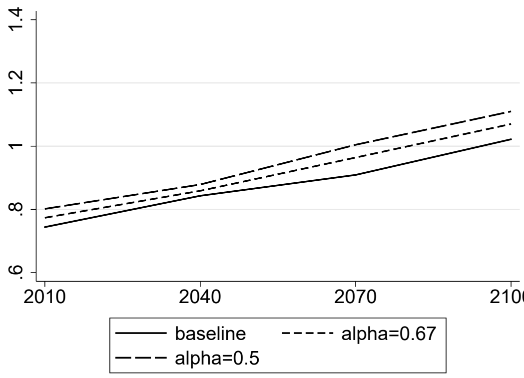

(a) Income per capita - externalities




(b) Theil index - externalities




(c) Income per capita - preferences




(d) Theil index - preferences

Figure A7: Income inequality prospects under alternative modeling assumptions