Geography of skills and global inequality☆

Michał Burzynski a, Christoph Deuster b,c, Frédéric Docquier a,b,d,∗

a LISER, Luxembourg Institute of Socio-Economic Research, Luxembourg
b IRES, Université Catholique de Louvain, Belgium
c Universidade Nova de Lisboa, Portugal
d National Fund for Scientific Research (FNRS), Belgium

ARTICLE INFO

JEL codes:
E24
J24
O15

Keywords:
Human capital
Migration
Urbanization
Growth
Inequality

ABSTRACT

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 influence 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 have a long-term impact on demographic growth and global inequality.

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 regions/sectors (urban and rural or equivalently, non-agriculture 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

https://doi.org/10.1016/j.jdeveco.2019.02.003
Received 23 November 2017; Received in revised form 16 February 2019; Accepted 20 February 2019
Available online XXX

© 2019 Elsevier B.V. All rights reserved.

Please cite this article as: Burzynski, M., et al., Geography of skills and global inequality, Journal of Development Economics, https://doi.org/10.1016/j.jdeveco.2019.02.003
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 play a major role. 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 a tightening of international migration policies. We also assess the robustness of our results to the technological and preference assumptions of the model.

Fig. 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, Zim-
Fig. 1d shows that the ratio of human capital between agriculture and 358 regions (i.e., rural and urban regions of the 179 countries) educated workers in the year 2010 for a sample of 179 countries oraging) and economic variables (such as trade, unemployment, or conges-
tions). Distinguishing between urban and rural regions allows us to model the differential in the access to education across regions (as in Lucas, 2009). Hence, in poor countries, the share of college graduates is remarkably low in the rural areas, 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. 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 growth 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 (i.e., aggregate and skill-biased technical changes induced by human capital accumulation). 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 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. 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 Hsieh and Klenow, 2009; Bryan et al., 2014). 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. 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 Fig. 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 a possible tightening of international migration policies.

Within the context of the convergence literature, 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 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 have a long-term impact on demographic growth 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 (Glaeser et al., 2004; Acemoglu et al., 2014; 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 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

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 (Castelló-Climent, 2008; Bobba and Coviello, 2007; 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 (see Mokyr, 2005; Mokyr and Voth, 2009). 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-Cabrallina, 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 to agglomerate in countries/regions with high rewards to skill (Grogger and Hanson, 2011; Belot and Hatton, 2012; Docquier and Rappoport, 2012; 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 (Hsieh and Klenow, 2009; Bryan et al., 2014). 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 formalizes the interactions between the geography of skills and the distribution of income. It endogenizes the accumulation

---

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 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.
of skills and its implications for economic development. 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 Boppard (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 Galor and Weil, 2000; Galor, 2011; de la Croix and Døpke, 2003, 2004). 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 accumulation, 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. 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.\(^8\) 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:

\[
Y_{r,t} = A_{r,t} \left( \sum_s \sigma_{r,s,t} \ell_{r,s,t}^{\sigma_{r,s,t}} \right)^{1/\sum_s \sigma_{r,s,t}} \forall r, t,
\]

where \( A_{r,t} \) denotes the productivity scale in sector \( r \) at time \( t \), \( \sigma_{r,s,t} \) is a sector-specific variable governing the relative productivity of workers of type \( s \) (such that \( \sigma_{r,h,t} + \sigma_{r,l,t} = 1 \)) 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:

\[
w_{r,s,t} = A_{r,t} \left( \sum_s \sigma_{r,s,t} \ell_{r,s,t}^{\sigma_{r,s,t}} \right)^{1/\sum_s \sigma_{r,s,t}} \forall r, s, t.
\]

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

\[
R_{r,t} \equiv \frac{w_{r,h,t}}{w_{r,l,t}} = \frac{R_{r,t}^m \left( R_{r,t}^e \right)^{\gamma_t}}{\forall r, t},
\]

where \( R_{r,t}^m \equiv \frac{\ell_{r,h,t}}{\ell_{r,l,t}} \) is the skill ratio in the labor force of region \( r \) at time \( t \) and \( R_{r,t}^e \equiv \frac{\ell_{r,h,t}^e}{\ell_{r,l,t}^e} \) measures the skill bias in relative productivity. Although human capital is used in agriculture, the literature has emphasized that the marginal product of human capital is greater in the nonagricultural sector (see Lucas, 2009; Vollrath, 2009; Gollin et al., 2014).

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 the following:

\[
A_{r,t} = \left( \sum_s \sigma_{r,s,t} \ell_{r,s,t}^{\gamma_t} \right)^{1/\sum_s \sigma_{r,s,t}} \forall r, t.
\]

where \( y_t^\gamma \) is a time trend in productivity that is common to all countries \((\gamma > 1)\), \( A_{r,t} \) 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 \( \gamma_t \in (0, 1) \) 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:

\[
\Gamma_t = \frac{A_{r,t} - A_{n,t}}{A_{r,t}} \frac{\tilde{A}_{n,t} \left( R_{n,t}^e \right)^{\gamma_t}}{\tilde{A}_{r,t} \left( R_{r,t}^e \right)^{\gamma_t}} \forall r, t.
\]

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, \( \tilde{A}_{n,t} \neq \tilde{A}_{r,t} \), (ii) the differences in the elasticity of TFP to human capital, \( \gamma_t \neq \gamma_a \), or (iii) the disparities in human capital across sectors, \( R_{n,t} \neq R_{r,t} \). 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
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 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:

\[
U_{\lambda r,s,t} = \ln v_{r,s,t} + \ln (1 - x_{r,t,s}) + \xi_{r,t,s} \quad \forall r,s,t,
\]

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

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

where \( \mu > 0 \) is a common scale parameter governing the responsiveness of migration decisions to changes in \( v_{r,s,t} \) and \( x_{r,t,s} \) at the rate \( \theta \approx 0.577 \) (or Euler’s constant. Although \( \xi_{r,t,s} \) 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 \( v_{r,s,t} \) is a function of consumption \( c_{r,s,t} \), fertility \( n_{r,s,t} \) and the probability that each child becomes highly skilled \( p_{r,s,t} \):

\[
\ln v_{r,s,t} = \ln c_{r,s,t} + \theta \ln (n_{r,s,t} p_{r,s,t}) \quad \forall r,s,t,
\]

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

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

\[
p_{r,s,t} = (\pi_t + q_{r,s,t})^\delta \quad \forall r,s,t,
\]

where \( \pi_t \) is an exogenous parameter that is region-specific and \( \delta \) 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 \( E_r \). The budget constraint is written as follows:

\[
c_{r,s,t} = w_{r,s,t}(1 - \delta n_{r,s,t}) - n_{r,s,t} p_{r,s,t} E_r;
\]

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

\[
\ell_{r,s,t} = E_r(1 - \delta n_{r,s,t}).
\]

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.
3.2.2. 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:

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

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

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

Eq. (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 $x_{r \rightarrow s,t}^*$. Heterogeneity in migration tastes implies that emigrants select all destinations for which $x_{r \rightarrow s,t}^* < 1$ (if $x_{r \rightarrow s,t} = 1$, 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 ($x_{a,n,t}$ and $x_{a,a,t}$) 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 $v_{a,n,t} > v_{a,a,t}$, 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 ($x_{f,a,t}$ and $x_{f,n,t}$) 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 $t$:

$$
\begin{align*}
I_{n,s,t} &= \frac{N_{n,s,t}}{1 + m_{a,n,s,t} + m_{a,s,n,t}} + \frac{m_{a,n,s,t} N_{a,s,t}}{1 + m_{a,n,s,t} + m_{a,s,n,t}} + I_{n,s,t} \quad \forall s, \\
I_{a,s,t} &= \frac{N_{a,s,t}}{1 + m_{a,n,s,t} + m_{a,s,n,t}} + \frac{m_{a,n,s,t} N_{a,s,t}}{1 + m_{a,n,s,t} + m_{a,s,n,t}} + I_{a,s,t} \quad \forall s,
\end{align*}
$$

where $I_{n,s,t}$ 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 ($O_{s,t}$):

$$
O_{s,t} = M_{a,f,s,t} + M_{f,a,s,t} = \frac{m_{f,a,s,t} N_{a,f,s,t}}{1 + m_{a,n,s,t} + m_{a,s,n,t}} + \frac{m_{f,a,s,t} N_{a,f,s,t}}{1 + m_{a,n,s,t} + m_{a,s,n,t}} \forall s,
$$

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 of $\{\gamma, \theta, \lambda, \phi, \mu\}$ of common parameters, a set $\{\sigma, \epsilon, \lambda_Y\}$ of sector-specific elasticities, a set $\{\Gamma_{s,t}, R_{s,t}^p, \Gamma_{s,t}, \psi_{s,t}\}$ of country- and region-specific exogenous characteristics, and a set $\{N_{r,s,t}\}$ of predetermined variables, an intertemporal equilibrium is a reduced set of endogenous variables $\{\Lambda_{s,t}, \Lambda_{r,s,t}, I_{r,s,t}, N_{r,s,t}\}$, 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, $\epsilon_{s,t}, \Gamma_{s,t}, R_{s,t}^p$, $R_{s,t}^b$, $\Gamma_1$ 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 Eq. (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. Our methodology consists in calibrating common elasticities and 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. We start first describing how we estimate the geographic distribution of skills. Secondly, the parameterization of the technological and preference parameters is outlined. Thirdly, we illustrate the relevance of our modelling approach by showing that region-specific parameters exhibit desired correlations with traditional explanatory variables, and by assessing the ability of our model to match historical migration data. More details about the calibration can be found in Section A.1 in Appendix. We finally explain the general hypotheses used to initialize our baseline projections for the 21st century. The robustness of our results to specification choices is investigated in the Appendix.

**Estimating the geography of skills.** To construct labor force data by education level and by sector ($I_{r,s,t}$), 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 (YoS), a variable that we use in the third step of our estimation strategy. We are able to characterize the total number of workers ($\Sigma r, I_{r,s,t}$) and the total number of college-educated and less educated workers ($\Sigma r, I_{r,s,t}$) and $\Sigma I_{r,s,t}$) by

---

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.
country. The same strategy has been applied to all decades between 1970 and 2010 to compute the between-country index of inequality depicted in Fig. 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 (\(\sum L_{r,s,t}\)) 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 (\(H_{r,s,t}\)). 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 Appendix. To compute these shares, we collect or construct data on the years of schooling by sector (YoS, ) and use them to predict the sector-specific shares of college graduates as a function of YoS, . Hence, our third step consists of gathering data on YoS, and imputing the missing values. Gollin et al. (2014) and Ulubasoglu 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. 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. For the missing countries, we take advantage of the high correlation between the gap in years of schooling, YoS, /YoS, and the average years of schooling in the country, YoS. We predict the schooling gap by using estimates from year-specific regressions of this gap on YoS, .

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, \(H_{t,s} = f(YoS, )\), using Barro and Lee’s data, and then use the estimated coefficients to predict the share of college graduates in the urban sector, \(H_{t,s} = f(YoS, )\). 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 \(R^n_{s}\), 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 Fig. 1. Additional stylized facts are provided in Section A.1 in 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.

---

12 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.

13 Simple OLS regressions give \(\log YoS = 1.944 - 0.744 \log YoS (R^2 = 0.809)\) in 2010, and \(\log YoS = 1.464 - 0.550 \log YoS (R^2 = 0.905)\) in 1980.

14 Simple OLS regressions give \(\log YoS = 1.944 - 0.744 \log YoS (R^2 = 0.809)\) in 2010, and \(\log YoS = 1.464 - 0.550 \log YoS (R^2 = 0.905)\) in 1980.
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 \( \phi = 0.14. \) From Eqs. (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 \( \mu = 1.4. \)

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 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 (\( n_r(1980) \)). The average (national) fertility rate (\( \bar{r}_n \)) is thus obtained by dividing the total native population of adults in 2010 (\( \sum_r n_r(x_{1980,2010}) \)) by the total resident population of adults in 1980 (\( \sum_r n_r(x_{1980,1980}) \)). We also observe the skill structure of the native population in 2010 (\( n_r(x_{2010}) \)), which helps identifying education decisions in 1980 (\( \bar{r}_n \)). 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 \( n_r(1980) \) and \( \bar{r}_n \). From 2010 onwards, the number of children and education decisions are endogenous.

We estimate the skill and regional distribution of workers in 1980 and 2010 and calibrate internal migration costs as a residual from Eq. (16). In each country, we thus exactly match the variation in the urban population-share between the years 1980 and 2010. Similarly, we compute the average utility achievable in OECD destination countries and calibrate the international migration costs (\( x_{r,t} = x_{r,t}^{(I)} \) and \( x_{r,t}^{(F)} \)) as a unique solution from Eq. (16). We then calibrate the DIIO data on international migration stocks by education level for the year 2010. Again, more details are provided in Section A.1.

Validation. As shown in Definition 1, our model includes several region-specific parameters (namely TFP and skill bias levels, education cost and quality, internal and international migration costs) that are calibrated to match some moments. We use all the degrees of freedom of the data to identify the parameters needed, and the calibrated parameters exhibit a lot of variability (see Figs. A2 and A3 in Appendix A.1). This variability captures the determinants of demographic growth and global inequality, which are not directly or indirectly related to the geography of skills. Consequently, our model is exactly identified.

In order to establish the relevance of our parameterization strategy, we first demonstrate that our identified parameters exhibit realistic correlations with traditional explanatory variables from the econometric literature. Regression results are provided in Appendix A.1. As far as technological parameters are concerned, we find that the TFP scale factor \( \varphi \) exhibits realistic correlations with the cost of starting a business and with two proxies for infrastructure. In addition, the scale factor of the skill bias function \( \psi_{r,t} \) is positively associated with the share of researchers in the labor force. As far as educational parameters are concerned, we find that the cost of education \( \psi_{r,t} \) decreases with government spending in education, increases with the pupil-to-teacher ratio, and is greater in the rural sector. The quality of education \( \pi_r \) exhibits opposite correlations.

Regarding calibrated migration costs, Delogu et al. (2018) use a similar model and show that \( x_{r,t} \) decreases with educational attainment, and exhibits expected correlations with control variables used in the standard gravity equations (distance, common language, colonial links, visa restrictions). Instead of replicating these results, we assess the ability of our model to match historical mobility trends. In Appendix, we use our migration technology to predict migration stocks in the year 1980 (as well as the variations in the stock of migrants between 1980 and 2010), and compare our predictions with observations. We show that our predicted levels and variations almost coincide with the observed ones (see Fig. A4). This “backcasting” exercise shows that our model does a good job in explaining the long term evolution of international migration stocks.

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 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 2010–2040). 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 (i.e., in the years 2070 and 2100). 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 \( \psi_{r,t} \) (i.e., the same relative change in both sectors, keeping \( \psi_{r,t} / \psi_{n,t} \) 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 \( \psi_{r,t} \) with those obtained in 1980 (i.e., \( \psi_{r,t} \)), 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:

\[
\ln \left( \psi_{r,t+1} / \psi_{r,t} \right) = a_t + \beta_1 \ln \left( \psi_{r,t+1} / \psi_{r,t} \right) + \gamma_t \ln \left( \psi_{r,t+1} / \psi_{r,t} \right)^2. \tag{19}
\]

We obtain \( \gamma_t = 0.032, \gamma_n = 0.046, \) and \( \beta_n = 0.195 \) 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

---

15 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.

16 For this, we assume there is only migration from rural to urban regions (i.e., \( x_{r,t} < 1 \) and \( x_{r,t} = 1 \)).
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 \( y_{CF} \) obtained after transposing the US shares of college-educated workers in each sector. We then compare it with the observed level \( y_{obs} \). 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 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 Fig. 2.

Fig. 2a and b give the counterfactual levels of income per capita and the smoothed growth factor \( \frac{y_{CF}}{y_{obs}} \) 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 Fig. 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 externalities sizes are equal to 50% of the correlations between human capital and technological characteristics (i.e., \( \kappa_{n} = 0.19, \kappa_{s} = 0, \kappa_{l} = 0.28 \) and \( \kappa_{e} = 0.33 \)). The variants (dashed line) assume no externality, or externalities equal to 100% of the correlations (i.e., \( \kappa_{n} = 0.38, \kappa_{s} = 0, \kappa_{l} = 0.56 \) and \( \kappa_{e} = 0.66 \)). Fig. 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

Fig. 2c and d 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 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 (Hsich and Klenow, 2009; Bryan et al., 2014). 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, Fig. 2e and f 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 Fig. 2e), the effect of international migration on global inequality is small. On average, emigration rates are low in developing countries (approximately 5% for college graduates and less than 1% for the low-skilled). Despite positive selection in emigration, 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). Contrary to the previous experiments, the global inequality response to international migration is rather limited.19

---

17 Section A.4 in 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 non-agricultural goods are modeled.

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 (poorer countries are better off in the absence of skill-biased technical changes, unlike richest countries).

19 Tables A3 and A4 in Appendix give a more detailed description of the effect of the different static counterfactual experiments for the US and for the 15th (Cambodia), 25th (Ghana), 50th (Tunisia), 75th (Mexico) and 85th (Greece) percentiles of the income distribution. The presentation is organized as in Jones (2014). Table A3 focuses on the average level of income per worker, while Table A4 distinguishes between the two production sectors.
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 medium variant in United Nations (2014)).

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 $\psi_{r,t}$). This implies that regions at an intermediate level of development converge towards the US fron-
tier 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 Fig. 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 Fig. 3a, c and e, respectively. Separate curves are provided for OECD countries, for developing countries, and for the entire world.\textsuperscript{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 Fig. 3b, d and f, 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.

Fig. 3a and b 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

\textsuperscript{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.
equal to official projections. Furthermore, the cross-country correlation between simulated and UN population sizes in the year 2100 equals 0.98.\textsuperscript{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 (2014). Fig. 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 Fig. 3d, the cross-country correlation between simulated and UN shares of college graduates in the year 2100 is large (0.91).\textsuperscript{22} However, most countries are below the 45° 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, Fig. 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 Fig. 3f, the cross-country correlation between simulated and UN urban shares in the year 2100 equals 0.83.\textsuperscript{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 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. Fig. 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): \( \ln(\psi_{r,t+1}/\psi_{r,t}) = \alpha_r + \beta_r \ln(\psi_{r,t+1}/\psi_{r,t}) \). We obtain the following estimates: \( \beta_r = 0.056 \) for rural regions, and \( \beta_u = 0.074 \) 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 Fig. 4b, d and f. 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, Fig. 4a, c and e 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.

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%–25% greater than in the baseline. These changes are noticeable in all developing countries, including the largest ones (see Fig. 4b, d and f). Hence, Fig. 4a, c and e 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 have a significant impact on demographic growth and global inequality in the long term.\textsuperscript{24}

5.4. Sensitivity to a tightening of 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 A3 in Appendix), which could induce a tightening of immigration policies. We consider here an extreme no-international migration scenario for the future (\( x_{nt,tt} = 1 \) after 2010).\textsuperscript{25} 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 (\( x_{nt,tt} = 1 \) after 2010). Fig. 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 reinforcing international migration restrictions has limited (and often negligible) effects on aggregated socio-demographic prospects (Fig. 5a, c and e). In the no-migration scenario, Fig. 5d shows that the share of college-educated workers increases in developing countries and that

\textsuperscript{21} The regression line of Fig. 3b is given by the following: baseline = 0.33 + 1.02-official (\( R^2 = 0.97 \)).

\textsuperscript{22} The regression line of Fig. 3d is given by the following: baseline = 0.03 + 0.92-official (\( R^2 = 0.82 \)).

\textsuperscript{23} The regression line of Fig. 3f is given by the following: baseline = 0.15 + 1.11-official (\( R^2 = 0.69 \)).

\textsuperscript{24} Changing demographic shares have drastic implications in terms of immigration and emigration (see Table A4 in 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.

\textsuperscript{25} On the contrary, Delogu et al. (2018) shed light on the dynamic implication of a complete relaxation of immigration restrictions.
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. Emigration rates are much smaller in larger countries and the effects of cutting (selective) migration are negligible. 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. Fig. 5e and f 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.

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 demographic growth 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. Nevertheless, the geography of skills can be affected by public policies affecting education and internal labor mobility. We now examine how these policies influence the world distribution of income. Our baseline prospects involve a variation of the Theil index of income inequality from 0.81 in

---

27 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.
Fig. 6. Implications for global income 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

1980 to 1.14 in 2100 (see Fig. A3a in Appendix A.3). Fig. 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.

Fig. 6a and b 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 Fig. 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 Fig. 4a).

Fig. 6c and d 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, Fig. 6c shows that it has a negligible effect on income per capita in the developing world. In other words, development prospects are robust to a tightening of international migration barriers. 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 demographic growth and global inequality.

Section A.4 in Appendix demonstrates that these conclusions are highly robust to the modelling assumptions. If we change the size of technological externalities or if we consider that agricultural and non-agricultural 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 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

28 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.
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 have a long-term impact on demographic growth and global inequality. These conclusions are highly robust to the technological and preference assumptions and to future international migration policies.

Appendix A

The Appendix provides material that supplements the paper. I can be found online at http://doi.org/10.1016/j.jdeveco.2019.02.003.

References

Young, A., 2013. Inequality, the urban-rural gap, and migration. Q. J. Econ. 128 (4), 1727-1785.