Abstract

In this paper, we characterize the recent evolution of the geographic distribution of educated elites, and study its implications for development inequality. Assuming continuation of recent educational policies, we then produce integrated projections of income, population, urbanization and human capital for the 21st century. For 179 countries, we parameterize a two-sector, two-class model that endogenizes knowledge accumulation, population growth, labor mobility, and income disparities across countries and across regions/sectors (agriculture vs nonagriculture). We show that the geography of educated elites matters for development, whatever the size of technological externalities. Disparities in access to education and internal mobility frictions explain a large fraction of income disparities between countries, while the effect of international migration is small. As a proof of concept, the model predictions are in line with official socio-demographic projections. Population, human capital, urbanization and income inequality prospects are highly robust to technological externalities and to future immigration policies. However, they are highly sensitive to future education policies and to internal mobility costs, suggesting 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 in the long-run.

JEL codes: E24, J24, O15.

1 Introduction

It is commonly accepted that human capital acts as a proximate cause of development. In particular, the recent literature has shown that educated elites are instrumental to facilitating innovation and technology diffusion when knowledge becomes economically useful. This was the case during the industrial revolution (Mokyr, 2005; Scquicciarini and Voigtlander, 2015) and it is still relevant in the
modern world (Castelló-Climent and Mukhopadhyay, 2013). In this paper, we measure educated elites as the number of adults with completed tertiary/college education, as percentage of the working-age population. In many countries, college graduates represent a minority elite. Indeed, the worldwide average proportion of college graduates increased from only 3.5% in 1980 to 8.0% in 2010, and this proportion is currently smaller than 1% in fifteen developing countries such as Niger, Malawi, Zambia, Zimbabwe, Tanzania (Barro and Lee, 2013). The size of the educated elite is clearly endogenous as higher-education investments are costly, returns to schooling are endogenous, and educated elites are highly mobile across nations and regions. Figure 1(a) shows the evolution of inequality in educated elites from 1970 to 2010 in ten-year intervals. We use the Theil index of inequality and distinguish between its between-country component (capturing differences in the country average proportion of college graduates) and the within-country component (capturing differences between rural and urban regions). Inequality in educated elites is almost totally explained by the between-country component; this means that cross-country disparities are much greater than domestic disparities between regions. Since 1970, educated elites have grown faster in poor countries. Hence, the Theil index decreased by 0.27, reflecting unconditional convergence in the share of college graduates (around 0.7% per year). This is partly due to the fact that the average level of schooling is becoming relatively stable in the richest countries. However, large differences persist between the tails of the distribution, and between regions. This is illustrated in Figure 1(b), which depicts the kernel density for the shares of educated elites by country and by region in the year 2010.

What explains the recent disparities in educated elites? How are they affected by internal and international migrations? How does inequality in educated elites affect development disparities between countries? What are the prospects for the 21st century? These are the questions addressed in this paper.

![Theil indices of inequality in educated elites 1970-2010](a) and Kernel density of the share of educated elites in 2010 (b)

Figure 1: World distribution of educated elites

To address them, we build a model that endogenizes the size of educated

---

1This is in line with Meisenzahl and Mokyr (2012), who 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, they highlight the role of the top 3-5 percent of the labor force in terms of skills, including artisans, entrepreneurs and employees.
elites and income disparities within and across countries, and confront it with the data. 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 with completed college education and the less educated) and by their offspring. Distinguishing between urban and rural regions allows us modeling the differential in the access to education across regions, in line with Lucas (2009), as well as misallocations of workers (in line with Rodrik (2013). Adults decide how much to consume, what fraction of their children to provide with higher education, and where to live. Internal and international migration decisions depend on geographic disparities in income and on moving costs. The model is stylized and omits several features of the real world. It does not account for all demographic variables (such as mortality or aging) and economic variables (such as trade, unemployment, or redistribution). However, it accounts for the long-run interactions between human capital accumulation, migration and growth. We believe such a quantitative theory framework is an appropriate tool to identify the key factors governing the future disparities in the size of the educated elites across countries and regions. Our quantitative strategy consists of parametrizing the model for 179 countries so as to match the evolution of population, human capital, urbanization, productivity and income between the years 1980 and 2010, and then simulating the trajectory of these variables over the 21st century. Our goal is to investigate how the geography of educated elites affects current inequality, and identify the key forces that govern the future geography of educated elites, population growth and global inequality.

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 (2008), Card (2009) and Ottaviano and Peri (2012) who find high substitutability between workers with no schooling and high school degree, but small substitutability between those and workers with college education. In the context, increasing the size of the educated elite not only affects the average skills and cognitive abilities of workers, 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., a level of 2), human capital explains around 50% of the ratio of income per worker between the richest and poorest countries; this is much greater than what was found in earlier studies assuming perfect substitution between all categories of workers.

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 the US cities (Moretti, 2004) or the US states (Acemoglu and Angrist, 2004; Iranzo and Peri, 2009), some instrumental-variable approaches give substantial externalities (Moretti, 2004) while others do not (Acemoglu and Angrist, 2004). 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 identified 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 (Castello-Climente, 2008; Bobba and Coviello, 2007; Murtin and Wacziarg, 2014). Comparative development studies suggest that focusing on educated elites is more appropriate to account for such externalities. Squicciarini and Voigtländer (2015) show that upper-tail human capital was instrumental to explaining the process of technology diffusion during the French Industrial Revolution; on the contrary, mass education (proxied by the average level of literacy) is positively associated with development at the onset of the Industrial Revolution, but does not explain growth. Confirming Mokyr’s findings for the British Revolution, they conclude that the effect of educated elites 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 and poor countries struggle to adopt them. The contemporaneous effect of educated elites in poor countries is confirmed 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 size of the educated elite (proportion of individuals with tertiary education completed) has the same effect on growth as a 13% decrease in illiteracy rates (equivalently, a one standard deviation in educated elites has the same effect as a three standard deviation in literacy). Aggregate and skill-biased externalities cannot be ignored when dealing with long-run growth and inequality. However, given the uncertainty about their size, our analyses and projections cover several plausible scenarios.

As far as the source of human capital disparities is concerned, we treat the geography of educated elites as 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 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 Rapoport, 2012). Positive selection 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 get stuck in the wrong sectors, and that these sectoral and/or regional misallocations are likely to be important in poor countries. Our model will be used to approximate the effect of international migration on global inequality, and the fraction of income disparities explained by internal mobility frictions. We will also shed light on the implications of labor mobility for future
development.

Our main findings can be summarized as following. First, we use the model to quantify the fraction of contemporaneous inequality that is explained by differences in the size of educated elites. We show that the geography of educated elites matters for development, whatever the size of technological externalities. In the absence of technological externality, transposing the US educational structure to all countries reduces the Theil index by 33%, and reduces the income ratio between the US and countries in the lowest quartile of the income distribution by about 60%. This success rate is very much in line with Jones (2014); we obtain slightly greater success rates because in our two-sector model, a rise in human capital induces labor flows from the rural to the urban (more productive) sector. 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, disparities in educated elites explains 50% of the Theil index and more than 80% of the income ratio. And in a maximalist scenario where externality sizes are proxied by the correlation, human capital almost becomes the single determining factor for global inequality. Coming back to the baseline scenario, we show that keeping the size of the educated elite constant but transposing the US skill-specific urban shares reduces the income ratio by 30% (i.e., about one third of the total effect of human capital). This suggests that internal mobility frictions can generate misallocations of workers in poor countries, and shows the relevance of a two-sector approach. On the contrary, with the exceptions of small islands, the effect of international migration on global inequality is very small.

Second, we use the model to predict the evolution of population, human capital, urbanization and income over the 21st century. Accounting for interdependencies between these variables has rarely been done in projection exercises. In the baseline scenario, the model predictions are in line with official socio-demographic projections. This is a proof of concept that our stylized model does a very good job in generating reasonable projections of population, human capital, urbanization and income inequality. Furthermore, its micro-founded structure enables to identify the key factors that will govern the future of global inequality. In particular, we show that population, urbanization, human capital and income inequality prospects are highly robust to the size of technological externalities; changing the size of these externalities affects the long-run level of income per worker, but has small effect on its distribution. Socio-demographic and economic prospects are also highly robust to future international migration scenarios. Given demographic imbalances, the migration pressure to OECD destination will intensify; immigration

---

2For 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 IIASA projections 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 enrolment 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, or Sala-i-Martin, 2006).
policy responses (as drastic as totally cutting future migration flows) have negligible effects on future population and global inequality. However, our projections are highly sensitive to future education policies, and to future internal mobility frictions. Our baseline assumes a continuation of the convergence process in access to education observed during the last decades (as a possible consequence of the Millennium Development Goals). Attenuating or eliminating this convergence in education costs induces dramatic effects on population growth, urbanization and the world distribution of income. 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, education quality and sustainable urban development are vital to reduce the demographic pressure and global inequality.

The rest of this paper is organized as following. Section 2 describes our two-sector, two class model of human capital accumulation and income inequality. In Section 3, we parametrize this model to match historical data over the period 1980-2010 and the socio-demographic prospects for 2040. Section 4 discusses our simulation results, distinguishing between the contemporaneous implications of knowledge inequality, the projections for the 21st century, and a sensitivity analysis. Finally, Section 5 concludes.

2 Model

Our model depicts 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 the educated elite ($h$) and the less educated ($l$). We assume two-period lived agents (adults and children). 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. Goods produced in the two sectors are perfect substitutes from the point of view of consumers, and their price is normalized to unity. Considering goods as heterogeneous in a small open economy context with exogenous relative prices would lead to similar results. 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 the quantity and quality of their children. The latter decisions are governed by a warm-glow motive; adults directly value the quality and quantity of children, but they do not anticipate the future income and utility of their children (as in de la Croix and Doepke, 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, internal and international labor mobility. This section describes our assumptions and defines the intertemporal equilibrium.

2.1 Technology

We assume that output 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.\(^3\) 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 the labor market participation rate. Output levels at time \(t\) are given by:

\[
Y_{r,t} = A_{r,t} \left( \sum_s \varpi_{r,s,t} \ell_{r,s,t}^{\frac{\sigma}{\sigma-1}} \right)^{\frac{\sigma-1}{\sigma}} \forall r,t,
\]

where \(A_{r,t}\) denotes the scale productivity 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 worker 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:

\[
w_{r,s,t} = A_{r,t} \left( \sum_s \varpi_{r,s,t} \ell_{r,s,t}^{\frac{\sigma}{\sigma-1}} \right)^{\frac{\sigma-1}{\sigma}} \forall r,s,t,
\]

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

\[
R_{w,r,t}^w \equiv \frac{w_{r,h,t}}{w_{r,l,t}} = R_{w,r,t}^\varpi \left( R_{w,r,t}^{\ell \sigma} \right)^{\frac{\varpi-1}{\varpi}} \forall r,t,
\]

where \(R_{w,r,t}^{\ell \sigma} \equiv \frac{\ell_{r,h,t}}{\ell_{r,l,t}}\) is the skill ratio in the labor force of region \(r\) at time \(t\), and \(R_{r,t}^\varpi \equiv \frac{\varpi_{r,h,t}}{\varpi_{r,l,t}}\) 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 Lucas-type, aggregate externality (see Lucas, 1988) and assume that the scale of total productivity factor (TFP) in each sector is a concave function of the skill-ratio in the resident labor force. This externality captures the fact that educated elites facilitate innovation and the adoption of advanced technologies. We have:

\[
A_{r,t} = \gamma^t \overline{A}_{r,t} \left( R_{r,t}^{\ell \sigma} \right)^{\epsilon_r} \forall r,t,
\]

where \(\gamma^t\) is a time trend in productivity which 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, geography, soil fertility, etc.), and \(\epsilon_r \in (0,1)\) is a pair of elasticities of TFP to the skill-ratio in the sector. The productivity gap between the two sectors is thus given by

\[
\Gamma_t \equiv \frac{A_{n,t}}{A_{a,t}} = \frac{\overline{A}_{n,t} \left( R_{n,t}^{\ell \sigma} \right)^{\epsilon_n}}{\overline{A}_{a,t} \left( R_{a,t}^{\ell \sigma} \right)^{\epsilon_a}}.
\]

\(^3\)This elasticity plays a key role in development accounting and is shown to vary across sectors (Jones, 2014; Caselli and Ciccone, 2014; Lucas, 2009).
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, these findings can be driven by the correlation between the productivity gap with exogenous characteristics affecting development, $A_{n,t} \neq A_{a,t}$, by the effect of development on disparities in educated elites across sectors, $R_{\ell,n,t} \neq R_{\ell,a,t}$, or by differences in the elasticity of TFP to educated elites, $\epsilon_n \neq \epsilon_a$.

Second, we assume a skill-biased technical change. As the technology improves, the relative productivity of educated elites 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 declining relative industry demand for routine manual and cognitive tasks, and increased relative demand for nonroutine cognitive tasks. The observed relative demand shift favors college versus non-college labor. We write:

$$R_{\omega,r,t} = \overline{R}_{r} \left( R_{\ell,r,t} \right)^{\kappa_r} \forall r, t,$$

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

### 2.2 Preferences

Individual decisions to emigrate result from the comparison of discrete alternatives, staying in the region of birth, emigrating to the other region or 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 $s$, born in region $r^*$, moving to region/country $r$ is given by:

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

where $\ln v_{r,s,t} \in \mathbb{R}$ 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), $x_{r^*,r,s,t} \in (0, 1)$ 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, also known as the double exponential 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 change in $v_{r,s,t}$ and $x_{r^*,r,s,t}$, and $\vartheta \approx 0.577$ is the Euler’s constant. Although $x_{r^*,r,s,t}$ is individual-specific, we omit individual subscripts for notational convenience.

In line with de la Croix and Doepke (2003, 2004), the inner utility function $\ln v_{r,s,t}$ is a function of consumption ($c_{r,s,t}$), fertility ($n_{r,s,t}$) and the probability that each child becomes highly skilled ($p_{r,s,t}$):

$$\ln v_{r,s,t} = \ln c_{r,s,t} + \theta \ln (n_{r,s,t} p_{r,s,t}) \forall r, s,$$
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 ($q_{r,s,t}$):

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

where $\pi_r$ is an exogenous parameter that is region-specific as well as country-specific and $\lambda$ governs the elasticity of knowledge acquisition to 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 writes as:

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

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

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

In the following sub-sections, we solve this problem backward. We first determine the optimal fertility rate and investment in education in a given location $r$, which characterizes the optimal level of deterministic utility that can be reached in any location. We then characterize the choice of the optimal location.

2.2.1 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:

$$\frac{\phi w_{r,s,t} + q_{r,s,t}E_{r,t}}{w_{r,s,t}(1 - \phi n_{r,s,t}) - n_{r,s,t}q_{r,s,t}E_{r,t}} = \frac{\theta}{n_{r,s,t}},$$

$$\frac{n_{r,s,t}E_{r,t}}{w_{r,s,t}(1 - \phi n_{r,s,t}) - n_{r,s,t}q_{r,s,t}E_{r,t}} = \frac{\theta \lambda}{\pi_r + q_{r,s,t}}.$$  \hspace{1cm} (12)

Solving this system gives

$$\left\{ \begin{array}{l}
q_{r,s,t} = \frac{\lambda \phi w_{r,s,t} - \pi_r E_{r,t}}{(1 - \lambda)E_{r,t} - \pi_r w_{r,s,t}} \\
n_{r,s,t} = \frac{\theta(1 - \lambda)E_{r,t}}{1 + \theta} \cdot \frac{w_{r,s,t}}{\phi w_{r,s,t} - \pi_r E_{r,t}}
\end{array} \right. \quad \forall r, s.$$  \hspace{1cm} (13)

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.) which is the same for high-skilled and low-skilled individuals:

$$E_{r,t} = \psi_{r,t}w_{r,h,t} \quad \forall r, s.$$  \hspace{1cm} (12)

Plugging (12) into the first-order conditions gives:

$$\left\{ \begin{array}{l}
q_{r,h,t} = \frac{\lambda \phi}{(1 - \lambda)\psi_r} - \frac{\pi_r}{1 - \lambda} \\
q_{r,l,t} = \frac{\lambda \phi}{(1 - \lambda)\psi_r R_{r,t}} - \frac{\pi_r}{1 - \lambda}
\end{array} \right. \quad \text{and} \quad \left\{ \begin{array}{l}
n_{r,h,t} = \frac{\theta(1 - \lambda)}{1 + \theta} \cdot \frac{1}{\phi - \pi_r \psi_r} \\
n_{r,l,t} = \frac{\lambda \phi}{1 + \theta} \cdot \frac{1}{\phi - \pi_r \psi_r R_{r,t}}
\end{array} \right.$$  \hspace{1cm} (13)
note that $R_{r,t}^{w} > 1$ 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 nonagriculture region. Under the plausible condition $\psi_{a}/\psi_{n} > 1$, college-educated workers living in urban areas have fewer and more educated children ($q_{n,h,t} > q_{n,l,t}$ and $n_{n,h,t} < n_{n,l,t}$). Finally, when $\psi_{n}R_{r,t}^{w}/(\psi_{a}R_{r,t}^{w}) > 1$, 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 (13) into (8):

$$\begin{align*}
\ln v_{r,h,t} &= \chi + \ln (w_{r,h,t}) + \theta \lambda \ln \left(\frac{1}{\psi_{r}}\right) - \theta (1 - \lambda) \ln (\phi - \pi_{r,\psi_{r}}) \\
\ln v_{r,l,t} &= \chi + \ln (w_{r,l,t}) + \theta \lambda \ln \left(\frac{1}{\psi_{r}}\right) - \theta (1 - \lambda) \ln (\phi - \pi_{r,\psi_{r}}R_{r,t}^{w}) \\
&\quad + \ln \left(\frac{\phi(1 + \theta(1 - 1/R_{r,t}^{w})) - \pi_{r,\psi_{r}}R_{r,t}^{w}(1 + \theta(1 - 1/R_{r,t}^{w}))}{\phi - \pi_{r,\psi_{r}}R_{r,t}^{w}}\right)
\end{align*}
$$

(14)

where $\chi = \theta \ln \left(\frac{1}{\theta + \theta + (1 - \lambda)^{1 - \lambda}^{1 - \lambda}}\right) - \ln (1 + \theta)$ 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 or ”before-migration” population ($N_{r,s,t+1}$ $\forall r, s$) at time $t + 1$. We have:

$$\begin{align*}
N_{r,h,t+1} &= L_{r,h,t}n_{r,h,t}p_{r,h,t} + L_{r,l,t}n_{r,l,t}p_{r,l,t} \\
N_{r,l,t+1} &= L_{r,h,t}n_{r,h,t} [1 - p_{r,h,t}] + L_{r,l,t}n_{r,l,t} [1 - p_{r,l,t}]
\end{align*}
$$

(15)

### 2.2.2 Migration and population dynamics

Given their taste characteristics (captured by $\xi$), each individual chooses the location that maximizes her/his utility (7). Under the Type I Extreme Value distribution for $\xi$, McFadden (1984) shows that the emigration rate from region $r^{*}$ to a particular destination $r$ is governed by a logit expression. The emigration rate is given by

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

Skill-specific emigration rates are endogenous and comprised 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:

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

(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 in the region of origin \( r^* \). 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 \( x_{a,n,s,t} \) and \( x_{n,a,s,t} \) 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 are 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 (16), we can characterize the equilibrium structure of the resident population at time \( t \):

\[
\begin{align*}
L_{n,s,t} &= \frac{N_{a,s,t}}{1 + m_{a,n,s,t} + m_{nf,s,t}} + \frac{m_{an,s,t} N_{a,s,t}}{1 + m_{an,s,t} + m_{af,s,t}} \\
L_{a,s,t} &= \frac{N_{n,s,t}}{1 + m_{a,n,s,t} + m_{nf,s,t}} + \frac{m_{an,s,t} N_{n,s,t}}{1 + m_{an,s,t} + m_{af,s,t}}
\end{align*}
\forall s, \tag{17}
\]

as well as the outflow of international migrants by education level \( (O_{s,t}) \),

\[
O_{s,t} = M_{nf,s,t} + M_{af,s,t}
= \frac{m_{nf,s,t} N_{n,s,t}}{1 + m_{a,n,s,t} + m_{nf,s,t}} + \frac{m_{af,s,t} N_{a,s,t}}{1 + m_{a,n,s,t} + m_{af,s,t}} \forall s.
\tag{18}
\]

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

2.3 Intertemporal equilibrium

An intertemporal equilibrium for a developing economy can be defined as following:

**Definition 1** For a set \( \{ \gamma, \theta, \lambda, \phi, \rho, \mu \} \) of common parameters, a set \( \{ \sigma_r, \epsilon_r, \kappa_r \} \) of sector-specific elasticities, a set \( \{ A_r,t, \bar{R}_r,t, \bar{R}_r^w, \bar{R}_r^e, \bar{\gamma}_r, \bar{\pi}_r \} \) of country- and region-specific exogenous characteristics, and a set \( \{ N_{r,s,0} \} \) of predetermined variables, an intertemporal equilibrium is a reduced set \( \{ A_r,t, \bar{R}_r,t, w_{r,s,t}, \ell_{r,s,t}, T_{r,s,t}, N_{r,s,t+1}, L_{r,s,t} \} \) of endogenous variables satisfying technological constraints (4), (6) and (12), profit maximization conditions (2), utility maximization conditions (13), (14) and (16), 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, \( \ell_{r,s,t}, \bar{R}_r,t, \bar{R}_r^w, \bar{R}_r^e, \bar{\gamma}_t \) 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.
3 Parameterization

In this section, we describe our parametrization strategy for 145 developing countries and the set of 34 OECD countries as a single entity. We use socio-demographic and economic data for 1980 and 2010, as well as socio-demographic prospects for the year 2040. For each country, our baseline trajectory matches the recent trends in human capital accumulation, income disparities, and population movements (including internal and international migration). We start describing how the geographic distribution of educated elites and less educated workers is estimated in Section 3.1. We then calibrate the technological and preference parameters in Sections 3.2 and 3.3, respectively. Finally, Section 3.4 explains the general hypotheses governing our baseline projections for the 21st century.

3.1 Estimating the geography of educated elites

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 them with the widely used 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 quantify the size of educated elites of 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 homogenous group of less educated workers. Barro and Lee’s data are available for 143 countries. In addition, we make use of estimated data from Artuc et al. (2015). Note that Barro and Lee (2013) also document the average years of schooling of the working age population ($YoS_t$), a variable that we use in the third step of our estimation strategy. Without imputation, we are able to characterize the total number of workers ($\Sigma_{r,t}L_{r,s,t}$) and the total number of college-educated and less educated workers ($\Sigma_{r,t}L_{r,h,t}$ and $\Sigma_{r,t}L_{r,l,t}$) by country. The same strategy has been applied to all decades between 1970 and 2010 to compute the between-country index of inequality depicted on 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, 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. The latter variable is available in all countries and 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 ($\Sigma_{s,t}L_{r,s,t}$) 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 ($H_{r,t}$). To

---

4With the exceptions of Macao, North-Korea, Somalia and Taiwan, all countries that are not covered by our sample have less than 100,000 inhabitants.
impute these shares, we use data on years of schooling by sector \((YoS_{r,t})\), and predict the sector-specific shares of college graduates as a function of \(YoS_{r,t}\). Our third step consists of collecting data on \(YoS_{r,t}\) and imputing the missing values. Gollin et al. (2014) and Ulubasoglu and Cardak (2007) provide incomplete data on the average years of schooling and the average years of schooling in agriculture and nonagriculture for different years.\(^5\) We have data on 20 countries around 1980 and 65 countries around 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, \(YoS_{a,t}/YoS_{n,t}\), and the average years of schooling in the country, \(YoS_t\). We predict the schooling gap using estimates from year-specific regressions of this gap on \(YoS_t\).\(^6\)

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, \(H_t = f(YoS_t)\), using Barro and Lee’s data,\(^7\) and then use the estimated coefficients to predict the share of college graduates in the urban sector, \(H_{r,t} = f(YoS_{r,t})\). 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 calibration strategy, we compute the correlation between the sector-specific imputed shares of college graduates and the shares obtained from household survey. Using the Gallup data (available for about 145 countries), we can estimate the skill-ratio \((R'_{r,t})\) 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 educated elites in total employment for all decades between 1970 and 2010. We use it to compute the within-country index of inequality depicted on Figure 1.

Figure 2 characterizes the geography of educated elites in the year 2010, and describes the worldwide evolution of urbanization and human capital between 1970 and 2010. Figure 2(a) 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 2(b) shows that the gap between regions decreases with the economy-wide proportion of college graduates. Figure 2(c) shows that the college-educated minority is predominantly and increasingly employed in the non-agricultural 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 non-agricultural sector increased from 37.8% in 1970 to 50.5% in 2010. Figure 2(d) is the mirror image of Figure 2(c): it depicts the evolution of the share of the educated elite 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

\(^5\)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.

\(^6\)Simple OLS regressions give \(\log \frac{YoS_{n}}{YoS_{a}} = 1.944 - 0.744 \log YoS \ (R^2=0.809)\) in 2010, and \(\log \frac{YoS_{n}}{YoS_{a}} = 1.464 - 0.550 \log YoS \ (R^2=0.905)\) in 1980.

\(^7\)Simple OLS regressions give \(\log H = -4.804 + 0.279 \log YoS \ (R^2 = 0.496)\) in 2010, and \(\log H = -5.133 + 0.306 \log YoS \ (R^2 = 0.575)\) in 1980.
in nonagriculture (from 4.6% to 13.1%). In absolute terms, the magnitude of the change is reversed.

![Graph](image1)

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

![Graph](image2)

(b) Ratio of educated elites ($H_{a,t}/H_{n,t}$) and national share of college graduates ($H_t$) in 2010

![Graph](image3)

(c) World population share in nonagriculture by skill group

![Graph](image4)

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

Figure 2: Geography of educated elites

Notes: On Figure 2(a) and 2(b), bubble size is proportional to the population of the country.

### 3.2 Technology parameters

Output in each sector depends on the size and skill structure of employment. In the next section, 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 quantifying the employment levels ($\ell_{r,s,t}$) and the total employment in efficiency unit.

To calibrate the technological parameters a set $\{\sigma_r, \epsilon_r, \kappa_r, R^C_r, A_{r,t}\}$, we proceed in two steps. First, we calibrate the parameters affecting the private returns to higher education. For each sector, we combine cross-country data on income disparities across skill groups, with our estimates for $\ell_{r,s,t}$. This enables us parametrizing the elasticities of substitution between workers ($\sigma_r$), the relative
productivity of college graduates ($R^\omega_a$), the magnitude of the skill-biased externalities ($\kappa_a$), and the scale factors of the skill-bias technology ($\overline{R}^\omega_a$). In the second step, we focus on the social return to education. We use output data by sector and identify the level of total factor productivity. We then investigate the relationship between TFP and the skill ratio, which enables us defining an upper-bound for the aggregate TFP externalities ($\epsilon_r$) and the TFP scale factors ($\overline{A}_{r,t}$). Figure 3 summarizes our main findings.

In the first step, we calibrate the elasticity of substitution between educated elites and less educated workers relying on existing studies. As for the non agricultural sector, there is a large number of influential papers that propose specific estimated values for industrialized countries (i.e. countries where the employment share of agriculture is small). Johnson (1970) and Murphy et al. (1998) estimate values for $\sigma_n$ around 1.3. Ciccone and Peri (2005) and Krusell et al. (2000) estimate values around 1.6, and Ottaviano and Peri (2012) estimate a value close to 2.0. Angrist (1995) recommends a value above 2 to explain the trends in the college premium on the Palestinian labor market. As for the agriculture sector, it is usually assumed that the elasticity of substitution is 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 elite workers. In the agriculture 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 households with a less educated household head. Combining (3) and (6), the elasticity of $R^w_a$ to $R^\ell_a$ is equal to $\kappa_a - 1/\sigma_a$. Assuming $\sigma_a = \infty$, this elasticity boils down to $\kappa_a$. Figure 3(a) shows that the correlation between $R^w_a$ and $R^\ell_a$ is virtually nil. We thus rule out the possibility of skill-biased technical changes in agriculture ($\kappa_a = 0$), and assume a linear technology with a constant $R^\omega_a$ for all countries and all periods. The value of of $R^\omega_a$ is given by the population-weighted average of $R^w_a$, leading to $\omega_a = 0.57$. We use this value for all countries and assume it is time-invariant. As for the non-agricultural sector, we use data on the wage ratio from Biavaschi et al. (2016) for 143 countries. We calibrate $R^\omega_n$ using (3). Regressing $R^\omega_n$ on $R^\ell_n$ gives 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 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 $\overline{R}^\omega_n$ as a residual from (6). Again, from (3) and (6), the elasticity of the $R^\omega_n$ to $R^\ell_n$ is equal to $\kappa_n - 1/\sigma_n$, which is equal to -0.37. Figure 3(b) 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
Notes: On Figure 3(a)-3(d), bubble size is proportional to the population of the country.

of Agriculture (USDA).\textsuperscript{9} Data on the agriculture share in the value added are

\textsuperscript{9}For a few missing observations we impute missing values by making use of the Maddison
data base and data from the World Bank.
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 elite 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 on Figures 3(c) and 3(d). 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 educated elites 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 3(e) and 3(f) 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 parametrize 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 projection exercise.

### 3.3 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: \(\theta = 0.3\), \(\lambda = 0.4\) and \(\phi = 0.14\).\(^{11}\) 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 an elasticity 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 \(\psi_r\). These two parameters are country- and sector-specific, and affect 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 \((\sum_r N_{r,s,2010})\). 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

---

\(^{10}\)For a few missing observations we impute missing 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.

\(^{11}\)Given the expression in (10), this assumptions reflects setting the bound of the maximal number of children equal to seven. See Docquier et al. (2016) for a brief review of studies using similar parameter values.
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 database 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.\textsuperscript{12} The average fertility rate ($\bar{n}_{1980}$) 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 ($\sum_{r,s} L_{r,s,1980}$).\textsuperscript{13} Moreover, our calibration requires data on the skill- and region-specific fertility for each country. By construction, we have $\bar{n}_t \equiv \sum_{r,s} L_{r,s,t} n_{r,s,t} / \sum_{r,s} L_{r,s,t}$. We use the Gallup World Polls and extract the Gallup-based average number of children per household by region and skill level for 2010.\textsuperscript{14} We match the fertility rates of the low-skilled workers in both sectors, and compute the fertility of the educated elites. We then extrapolate these fertility measures to obtain all the respective values for each country between 1980 and 2010. The last moment to fit in the procedure is related with determining the number of internal migrants between years 1980 and 2010.

Since the evolution of educated elites in both sectors may be brought about by the differences in educational prospects (given the already computed fertility differential), or the selectivity of rural-urban migrants, we decided to pin down the first of the two above mentioned factors. We set the ratio of probabilities of becoming high-skilled in urban and rural areas, so that net internal migration is computed as a residual in the model. We arbitrarily impose that the process of urbanization is the dominant one (which is the case in almost all countries), so the matched number represents the net migration from rural to urban region. In this way, the model perfectly fits 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 $\pi_r \psi_r$. Once fertility rates are matched and the product $\pi_r \psi_r$ is identified, we calibrate $\pi_r$ and $\psi_r$ to match the educational structure of the native population in 2010, imposing a given value to the ratio of probabilities of becoming high-skilled across regions.\textsuperscript{15} Figures 4(a) and 4(b) show the distributions of $\pi_r, \psi_r$ for the two regions. Figure 4(a) depicts the distributions for two periods (1980 and 2010).

\begin{itemize}
\item\textsuperscript{12}Bertoli and Ruyssen (2016) show that aspirations to emigrate are correlated with emigration flows within five years.
\item\textsuperscript{13}There is no mortality in the model. The average fertility rate at time $t$, $\bar{n}_t$, 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.
\item\textsuperscript{14}We only include countries with at least ten respondents. For missing data countries, data from the World Health Organization on the crude birth rates is used.
\item\textsuperscript{15}The country-specific ratio of probabilities of becoming highly-skilled is computed using the overall shares of highly-educated, years of schooling, and the computed ratio of years of schooling in rural and urban sectors (using the data from Gollin et al. 2015). Assuming that the distribution of years of schooling comes from the log-normal family, we compute their standard deviations in a way that we fit perfectly the share of highly-skilled workers in each country. Then, using the estimated measures of dispersion and region-specific means, we retrieve the probabilities of becoming skilled in rural and urban sectors. A quotient of the two is the required ratio, which allows to separately identify $\psi_n$ and $\psi_u$.
\end{itemize}
distribution of $\pi_r$ is stable over time. As far as $\psi_r$ 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 assumed 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 4(c) shows that moving costs are usually smaller for highly educated workers than for the less educated.

![Kernel density of $\psi_r$](a)

![Kernel density of $\pi_r$](b)

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

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

Figure 4: Calibration of the preference parameters in 1980 and 2010

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 equation (16). In line with Figure 4(c), Figure 4(d) shows that international migration costs are smaller for the educated elite.

### 3.4 Baseline scenario

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 of the correlation between TFP (and skill bias) and the share of educated elites is due to the schooling externality; alternative technological scenarios are considered in section 4.4. The philosophy of our baseline projection exercise is to predict the future trends in income, population and human capital if all parameters remain constant, with the exception of the parameters governing access to education. More precisely, we constrain our baseline trajectory to be compatible with medium-term official demographic projections, as reflected by the UN projections of the national adult population and proportion of college graduates for the year 2040. Hence, we allow for country-specific proportional adjustments in $\psi_{r,t}$ ($r = a, n$) (i.e. the same relative change in both sectors) that minimizes the sum of squared differences in population and human capital between the baseline simulations and the UN projections for the year 2040. Remember $\psi_{r,t}$ determines the cost of 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 clear unconditional convergence process in the access to education. We see it as a likely consequence of the Millennium Development policy. We estimate two standard, region-specific convergence equations considering the US as the benchmark frontier: $\ln (\psi_{r,t+1}/\psi_{r,t}) = \alpha_r + \beta_r \ln (\psi_{USA}^{t+1}/\psi_{r,t})$. We obtain $\beta_a = 0.085$ and $\beta_n = 0.126$, and both are highly significant. For subsequent years, our baseline scenario assumes a continuation of this convergence process, in line with the new Sustainable Development Agenda. Alternative (less optimistic) educational scenarios will also be considered in Section 4.5.

4 Results

In this section, we investigate how disparities in the size of educated elites affect current and future development and inequality. In line with the development accounting literature, Section 4.1 uses a set of counterfactual experiments to quantify the fraction of contemporaneous development inequality that is explained by differences in the proportion of elite workers, internal migration, and internal mobility frictions. Then, Section 4.2 compares our baseline projections of worldwide population, urbanization, human capital and income per capita, with official projections for the 21st century. Section 4.3 examines the main geopolitical implications of our baseline projections. The sensitivity of our projections to the size of technological externalities, to future educational policies, and to future mobility frictions is then assessed in Sections 4.4, 4.5 and 4.6, respectively.

4.1 Educated elites and development in 2010

In this section, we use the parametrized model and proceed to a set static counterfactual experiments to identify the role of educated elites and of their geographic mobility. Results are depicted in Figure 5. In line with the development accounting literature, we first simulate the counterfactual (CF) levels of national income per worker ($y$) obtained after transposing the US shares of educated elite workers in each sector, and we compare it with the observed levels (obs). Building on Jones (2014), we compute the success rate (SR) as the share of the income ratio explained by the counterfactual (i.e. one minus the counterfactual-to-observed ratio of income differential with the US). Equivalently, the success rate measures
the national income loss due to the lower level of human capital and the sectoral (mis)allocation of workers when compared to the US:

$$SR = 1 - \frac{(y_{US}/y)_{CF}}{(y_{US}/y)_{obs}} = \frac{y_{CF} - y_{obs}}{y_{CF}}$$

Figures (5a) and (5b) give the counterfactual levels of income per capita and the success rates obtained under three technological scenarios after transposing the US shares of educated elite workers. On Figure 5(a), the bold line shows the observed income levels; countries are ranked by ascending order with respect to the observed level of income per worker. The baseline scenario (solid blue line) assumes that externality sizes are equal to 50% of the correlations between human capital and technological characteristics (i.e., $\kappa_n = 0.19$, $\epsilon_n = 0.28$, and $\epsilon_a = 0.33$). The variants (red lines) assume no externality, or externalities equal to 100% of the correlations (i.e., $\kappa_n = 0.38$, $\epsilon_n = 0.56$ and $\epsilon_a = 0.66$). We show that the geography of educated elites matters for development, whatever the size of technological externalities.

In the absence of any externality, transposing the US educational structure reduces the income ratio between the US and countries in the lowest quartile of the income distribution by about 60% per worker. The effect decreases with development, as the distance to the frontier gets smaller. Our success rate is in line with Jones (2014), who finds a success rate around 50% with the same elasticity of substitution. As in Jones, the effect is mainly driven by the fact that educated elites are more productive, and by the complementarity between them and less educated workers. In addition, our model accounts for the role of urbanization. Transposing the US skill shares and the US sectoral allocation of workers increases the relative size of the urban (more productive) sector; this explains our greater success rate. In our baseline scenario with conservative externalities, human capital disparities generates a success rate of about 80% in the poorest countries. In the full-externality scenario, human capital almost becomes the single determining factor for global inequality.

Figures (5c) and (5d) use the baseline externality scenario (50% of correlations) and include each externality at a time. It clearly shows that the results are highly sensitive to the aggregate TFP externality (almost equivalent to the baseline with both externalities). On the contrary, the skill biased technical change externality plays a negligible role in explaining the growth potential (almost equivalent to the no-externality scenario), though it slightly exacerbates disparities income across countries (poorest countries are better off in the absence of skill biased technical changes, unlike richest countries).

Then, Figures (5e) and (5f) illustrate the role of the geographic mobility of educated elites. Using the baseline externality levels, we simulate the effect of transposing the US skill-specific urban shares (keeping the country-wide size of the educated elite at the observed levels) and of returning all expatriates to their home country (no migration scenario). Although small islands exhibit large high-skilled emigration rates, the effect of international migration on global inequality is very small. On average, international migration explains an income loss of 15%.

\[1^{16}\text{Most studies in development accounting disregard technological externalities (see Jones, 2015), or consider that externalities are small (Caselli and Ciccone, 2014).}\]
Figure 5: Income per worker and success rate in static counterfactuals

Notes: Countries are ranked by ascending order with respect to the observed level of GDP per capita and the respective scenario.
to 18% with externalities. On the contrary, transposing the US urban shares for each category of worker reduces the income ratio between the US and countries in the lowest quartile of the income distribution by 30% to 40% (i.e., more than a third of the total effect of transposing the US skill shares in each sector). In line with Rodrik (2013), this suggests that internal mobility frictions might be responsible for large misallocation of workers in poor countries, and shows the relevance of a two-sector approach. 17

4.2 Baseline projections

We now turn our attention to the prospective analysis. This section compares our baseline simulations with official projections (medium variant of the UN). Results are described in Figure 6. The simulated and official trajectories of worldwide population, share of college graduates, and share of urban population are depicted in Figures 6(a), 6(c) and 6(e), respectively. The cross-country correlations between long-term simulations and official projections are described in Figures 6(b), 6(d) and 6(f), respectively.

On the one hand, the UN projections assume long-term, unconditional convergence in fertility, mortality and education attainment. On the other hand, our stylized model assumes unconditional convergence in access to education (i.e., in $\psi_{r,t}$), a constant growth rate of the scale TFP factor, and keeps all other parameters constant; it also assumes conservative (albeit positive) technological externalities. Figure 6 shows that our baseline trajectory is very much in line with official socio-demographic projections. The long-run level of the adult population slightly exceeds official projections; however, the cross-country correlation between simulated and UN population sizes in the year 2100 equals 0.875.

As far as education is concerned, the worldwide share of college graduates in the labor force is virtually identical to the UN estimates. This share increases from 8.8% in 2010 to 21.8% in 2100 (from 20.5 to 42.7% in the OECD countries, and from 5.9 to 18.0% in the non-OECD countries). The cross-country correlation between simulated and UN shares of college graduates in the year 2100 is equal to 0.881.

Finally, the share of the population living in urban areas is also in line with the UN projections. The worldwide urban share increases from 53.0% in 2010 to 73.0% in 2100 (from 79.0 to 90.1% in the OECD countries, and from 46.8 to 70.0% in the non-OECD countries). These trends are the outcomes of two opposing forces, i.e. rural/urban fertility differential and the net internal mobility towards cities (driven by the rising educational attainment). The latter strongly dominates. The cross-country correlation between simulated and UN urban shares in the year 2100 is equal to 0.787.

Overall, these comparisons give suggestive evidence that, albeit small in scale, our stylized model does a very good job in generating reasonable projections of population, human capital, urbanization and income inequality for the coming

---

17Table 2 and Table 3 in the 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 2 focuses on the average level of income per worker, while Table 3 distinguishes between the two production sectors.
Figure 6: Comparison of the baseline trajectory with official projections

decades. Importantly, as it is micro-founded, the model enables to identify the key factors that will govern the future of the world population and global inequality. In particular, our goal is to assess whether the evolution of population and global inequality is robust to technological externalities, educational policies and geographic mobility costs.
4.3 World economy implications

Before moving into the presentation of the robustness analysis, this section analyzes the geopolitical implications of our baseline projections. First, the model predicts convergence in income per worker and in the size of educated elites across countries. As a consequence of the converging access to education, the Theil index of human capital inequality decreases threefold over the 21st century, from 0.63 to 0.21 as illustrated in Figure 7(b). This implies that income per capita also converges, but much less rapidly due to the large disparities in the TFP scale factor depicted in Figure 3(f). The Theil index of income inequality decreases from 0.81 in 1980 to 0.64 in 2100 as depicted in Figure 7(a).

Figure 7(c) 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 12.8% in 2100. The share of OECD countries decreases from 26.0% to 15.7% over the same period of time. In addition, the OECD share in the college-educated population shrinks markedly, as illustrated in Figure 7(d). 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 7(e) shows that the speed of urbanization is faster in Africa than in the other regions. Finally, Figure 7(f) depicts the evolution of income shares. The OECD income share decreases by more than 20 percentage points (from 77.4% in 1980 to 55.2% in 2100) whereas the Asian share increases from 9.1% to 24.0% over the same period.

Table 1 describes the international migration implications of our baseline projections. Assuming constant migration policies, we predict little variations in future emigration rates from the OECD member states, from Latin America and from the Middle East and North Africa. On the contrary, emigration rates from Sub-Saharan Africa and Asia substantially increase (from 1.4% to 2.5% in sub-Saharan Africa between 2010 and 2100). 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 20.9% 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) educated elites 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 4.6. Note that the share of immigrants increases slightly in the US (from 16.0% to 18.9%), while migration pressures are greater in Canada and, to a lesser extent, in Australia.

4.4 Technological variants

Let us now get back to the global projections and assess their robustness to modeling assumptions. The static counterfactual experiments conducted in Section 4.1 show that the effect of educated elites on global inequality quantitatively depends
(a) Theil index of income inequality

(b) Theil index of inequality in educated elites

(c) Share of working population

(d) Share of college-educated workers

(e) Share of urban population

(f) Share of GDP

Figure 7: Shares by regions (1980-2100)

Notes: This figure reports the Theil index of income inequality, the Theil index of inequality in educated elites, the regional shares of global labor force, high-skilled workers, urban workers and GDP. In Figure 7(c)-(f) 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 1: Projections of immigration and emigration rates

<table>
<thead>
<tr>
<th></th>
<th>Baseline scenario</th>
<th>No ext.</th>
<th>No conv.</th>
<th>No urb.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2010</td>
<td>2040</td>
<td>2070</td>
<td>2100</td>
</tr>
<tr>
<td>Emigration rates (as percent of native population)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OECD</td>
<td>0.043</td>
<td>0.046</td>
<td>0.043</td>
<td>0.040</td>
</tr>
<tr>
<td>LAC</td>
<td>0.039</td>
<td>0.041</td>
<td>0.040</td>
<td>0.039</td>
</tr>
<tr>
<td>SSA</td>
<td>0.014</td>
<td>0.016</td>
<td>0.020</td>
<td>0.025</td>
</tr>
<tr>
<td>MENA</td>
<td>0.029</td>
<td>0.032</td>
<td>0.033</td>
<td>0.033</td>
</tr>
<tr>
<td>Asia</td>
<td>0.011</td>
<td>0.013</td>
<td>0.015</td>
<td>0.017</td>
</tr>
<tr>
<td>Others</td>
<td>0.139</td>
<td>0.143</td>
<td>0.143</td>
<td>0.140</td>
</tr>
<tr>
<td>Immigration rates (as percent of resident population)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EU</td>
<td>0.121</td>
<td>0.153</td>
<td>0.174</td>
<td>0.190</td>
</tr>
<tr>
<td>EU 15</td>
<td>0.136</td>
<td>0.170</td>
<td>0.193</td>
<td>0.209</td>
</tr>
<tr>
<td>GER</td>
<td>0.150</td>
<td>0.184</td>
<td>0.206</td>
<td>0.224</td>
</tr>
<tr>
<td>FRA</td>
<td>0.122</td>
<td>0.155</td>
<td>0.177</td>
<td>0.194</td>
</tr>
<tr>
<td>GBR</td>
<td>0.146</td>
<td>0.210</td>
<td>0.258</td>
<td>0.297</td>
</tr>
<tr>
<td>ITA</td>
<td>0.109</td>
<td>0.131</td>
<td>0.144</td>
<td>0.152</td>
</tr>
<tr>
<td>ESP</td>
<td>0.140</td>
<td>0.159</td>
<td>0.167</td>
<td>0.171</td>
</tr>
<tr>
<td>USA</td>
<td>0.160</td>
<td>0.183</td>
<td>0.190</td>
<td>0.189</td>
</tr>
<tr>
<td>CAN</td>
<td>0.187</td>
<td>0.241</td>
<td>0.270</td>
<td>0.288</td>
</tr>
<tr>
<td>AUS</td>
<td>0.249</td>
<td>0.285</td>
<td>0.300</td>
<td>0.304</td>
</tr>
</tbody>
</table>

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). Columns one to four give the respective values for the baseline scenario for the years 2010-2100. Column five gives the respective values for the counterfactual scenario with no technological externalities (No ext.) for the year 2100. Column six gives the respective values for the counterfactual scenario with no convergence in education costs (No conv.) for the year 2100. Column seven gives the respective values for the counterfactual scenario with no internal mobility (No urb.) for the year 2100.

on the level of technological externalities. In the section, we assess whether these externalities also affect our socio-demographic and income projections. Figure 8 compares the baseline trajectories of population, education, urbanization, income per capita and inequality with those obtained without or with full externalities.

Given the steady increase in the proportion of college graduates, long-run projections of income per capita are sensitive to the technological scenario (Figure 8(d)). However, the evolution of socio-demographic variables is highly robust to the technological environment (Figures 8(a) to 8(c)). The fifth column of Table 1 also shows the migration projections are robust to the technological variants, although long-term migration pressures are greater in the absence of externality. More importantly, externalities have a rather limited impact on the Theil index. Compared to the baseline, a moderate short-term decrease in inequality
Figure 8: Sensitivity to technological externalities

Notes: This figure reports the worldwide projected population size, the share of college educated workers, the share of urban population, GDP per capita, Theil index of income inequality, and Theil index of inequality in educated elites for the baseline and the respective counterfactual scenario. The scenario "no externalities" refers to the scenario with no technological externalities. The scenario "full externalities" refers to the scenario with full technological externalities.

is obtained in the no-externality scenario; this is due to the fact that the skill
biased externality increases wage inequality and the cost of higher education for the poor (remember the cost of higher education is proportional to the high-skilled wage). On the contrary, a long-run decrease in inequality is obtained in the full-externality scenario; this is driven by the stronger change in the TFP in poor countries. Overall, technological externalities have a negligible effect on future demographic pressure, and a moderate impact on development inequality.

4.5 Educational variants

We now assess whether our socio-demographic and income projections 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 convergence process in education costs observed between 1980 and 2010. Figure 9 compares the baseline trajectories of population, education, urbanization, income per capita and inequality with those obtained without such a convergence, or when the convergence speed is halved.

The results clearly show that the evolution of the size and structure of population, as well as of the world distribution of income are highly sensitive to future educational policies. In the absence of convergence in access to education, population growth is almost twice as large as in the baseline (and the world population increases by about 100%), the proportion of college graduates stagnates after 2040, and urbanization grows much less rapidly. As far as migration pressures are concerned, the sixth column of Table 1 shows that destination countries exhibit higher immigration rates (due to the larger population in developing countries). Income growth is also affected (by the year 2100, income per capita is 30% smaller than in the baseline), and the Theil index of income inequality increases by about 10 percentage points. The latter effect is mainly driven by the increasing share of developing countries in the world population, and has drastic implications in terms of immigration and emigration. Hence, the number of international migrants increases by 30% compared to the the baseline. Interestingly, strong effects on population and income are also obtained when the convergence speed is halved. 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 pressure and global inequality.

4.6 Mobility variants

Finally, this section investigates whether our socio-demographic and income projections are sensitive to future migration costs. The baseline scenario assumes constant international and internal migration costs in the future. It shows that the international migration pressure drastically intensifies in the OECD countries. 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 full frictions ($x_{an,s,t} = 1$ after 2010). Figure 10 compares the baseline trajectories of population, education, urbanization, income per capita and inequality with those obtained without international or internal mobility.
Notes: This figure reports the worldwide projected population size, the share of college educated workers, the share of urban population, GDP per capita, Theil index of income inequality, and Theil index of inequality in educated elites for the baseline and the respective counterfactual scenario. The "half convergence" scenario assumes that the speed of unconditional convergence in $\psi_{r,t}$ is divided by two after 2010. The "no convergence" scenario assumes constant $\psi_{r,t}$ after the same year.
Notes: This figure reports the worldwide projected population size, the share of college educated workers, the share of urban population, GDP per capita, Theil index of income inequality, and Theil index of inequality in educated elites for the baseline and the respective counterfactual scenario. The scenario "no internal" refers to the scenario with prohibitively high internal migration costs ($x_{rf,s,t} = 1$ after 2010). The scenario "no international" refers to the scenario with prohibitively high international migration costs ($x_{an,s,t} = 1$) after 2010.
In line with the static development accounting exercise, we find that international migration has a negligible impact on socio-demographic and economic prospects. The real contribution of migration might be underestimated here as the model disregards diaspora externalities (Docquier and Rapoport, 2012) and the link between education decisions and migration prospects. However, Docquier and Machado (2016) and Delogu et al. (2015) show that the latter brain gain mechanism has little impact on the world distribution of income.

On the contrary, internal mobility plays a key role. Preventing people to move from rural to urban areas has drastic implications for human capital accumulation (access to education is better in cities), for population growth (fertility is greater in rural regions), for income (productivity is greater in cities), and for income inequality. The effect on the world distribution of income is even larger than that of the absence of convergence in education. In line with Figure 5, this confirms that internal mobility frictions might be responsible for large misallocation of workers in poor countries (Rodrik, 2013), and that policies targeting sustainable urban development are vital to reduce the demographic pressure and global inequality.

5 Conclusion

Educated elites shape the prospects of development. In this paper we investigate and quantify this statement using a multi-country macroeconomic model with endogenous fertility and human capital formation, externalities to productivity and skill-bias, as well as individual decisions concerning internal and international migrations. We parameterize this model to match the characteristics of global economy in 1980 and 2010. Overall, we find that the allocation and evolution of educated elites induces drastic effects on global inequality.

The static development accounting analysis suggests that a hypothetical equalization of the (relative) size of educated elites across countries and regions to the US levels allows the poorest countries to make up even 80% of current development gap (as measured by income per worker). The extent to which this factor impacts inequality among economies decreases with countries’ development, and is almost nil in the right tail of the distribution of wealth. Interestingly, in the poorest regions, human capital is instrumental to explaining the development gap even after neutralizing its positive externalities on productivity. The contrary may be stated about the richest parts of the world, where the human capital had already been accumulated, and externalities matter most for further growth. Our results show that the human capital externality to TFP is responsible for a large level effect in GDP per capita, while skill biased technical changes slightly deepen the divergence between the richest and the poorest. Finally, we confirm that the less industrialized states suffer from misallocation of labour across sectors (which may explain up to 30% of the income gap), whereas international migration is of less importance apart from several countries hit by the brain drain.

Our prospective analysis reveals that the size of the educated elite will remain an important factor in enhancing growth opportunities and in reducing income gaps between countries. A convergence in access to education affects fertility rates, individual development opportunities and migration options. We show that demographic and economic prospects are highly sensitive to future trends in ed-
ucational policies. In contrast, the magnitudes of TFP externalities affect future income inequalities only slightly (even though, it significantly shifts the levels of observable average output per worker). The last set of counterfactuals, in which we abolish future internal and international migration flows, reveals that the intra-country mobility is reducing inequality in income and is a significant factor in socio-demographic convergence. International mobility of people has much more limited impact on future demographic pressure and global inequality.

References


35


## 6 Appendix

Table 2: Educated elites and income per worker - Development accounting

<table>
<thead>
<tr>
<th></th>
<th>15&lt;sup&gt;th&lt;/sup&gt;</th>
<th>25&lt;sup&gt;th&lt;/sup&gt;</th>
<th>50&lt;sup&gt;th&lt;/sup&gt;</th>
<th>75&lt;sup&gt;th&lt;/sup&gt;</th>
<th>85&lt;sup&gt;th&lt;/sup&gt;</th>
<th>Theil Index</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Cambodia)</td>
<td>(Ghana)</td>
<td>(Tunisia)</td>
<td>(Mexico)</td>
<td>(Greece)</td>
<td></td>
</tr>
<tr>
<td>I. Observed levels and ratios of income per worker</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income pw</td>
<td>2,457</td>
<td>5,044</td>
<td>10,350</td>
<td>25,638</td>
<td>63,027</td>
<td>0.724</td>
</tr>
<tr>
<td>US/ctry ratio</td>
<td>49.4</td>
<td>24.1</td>
<td>11.7</td>
<td>4.7</td>
<td>1.9</td>
<td>-</td>
</tr>
<tr>
<td>Success</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>II. Counterfactual: Transposing the US skill shares in each sector</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income pw</td>
<td>19,116</td>
<td>17,842</td>
<td>22,722</td>
<td>42,308</td>
<td>71,752</td>
<td>0.342</td>
</tr>
<tr>
<td>US/ctry ratio</td>
<td>6.4</td>
<td>6.8</td>
<td>5.3</td>
<td>2.9</td>
<td>1.7</td>
<td>-</td>
</tr>
<tr>
<td>Success</td>
<td>0.871</td>
<td>0.717</td>
<td>0.544</td>
<td>0.394</td>
<td>0.122</td>
<td>0.528</td>
</tr>
<tr>
<td>III. Counterfactual II with exogenous TFP ($A_r$) and exogenous skill bias ($\mathcal{R}_w$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income pw</td>
<td>6,263</td>
<td>7,595</td>
<td>10,335</td>
<td>26,208</td>
<td>47,541</td>
<td>0.479</td>
</tr>
<tr>
<td>US/ctry ratio</td>
<td>16.8</td>
<td>13.8</td>
<td>10.2</td>
<td>4.0</td>
<td>2.2</td>
<td>-</td>
</tr>
<tr>
<td>Success</td>
<td>0.660</td>
<td>0.425</td>
<td>0.133</td>
<td>0.153</td>
<td>-0.148</td>
<td>0.338</td>
</tr>
<tr>
<td>IV. Counterfactual II with full TFP externality ($A_r$) and full skill bias externality ($\mathcal{R}_w$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income pw</td>
<td>71,042</td>
<td>49,173</td>
<td>57,837</td>
<td>75,694</td>
<td>117,671</td>
<td>0.253</td>
</tr>
<tr>
<td>US/ctry ratio</td>
<td>2.1</td>
<td>3.0</td>
<td>2.5</td>
<td>1.9</td>
<td>1.2</td>
<td>-</td>
</tr>
<tr>
<td>Success</td>
<td>0.958</td>
<td>0.876</td>
<td>0.784</td>
<td>0.591</td>
<td>0.352</td>
<td>0.650</td>
</tr>
<tr>
<td>V. Counterfactual II with exogenous skill bias ($\mathcal{R}_w$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income pw</td>
<td>22,355</td>
<td>19,860</td>
<td>25,463</td>
<td>45,341</td>
<td>75,736</td>
<td>0.327</td>
</tr>
<tr>
<td>US/ctry ratio</td>
<td>5.5</td>
<td>6.2</td>
<td>4.9</td>
<td>2.7</td>
<td>1.6</td>
<td>-</td>
</tr>
<tr>
<td>Success</td>
<td>0.888</td>
<td>0.741</td>
<td>0.585</td>
<td>0.423</td>
<td>0.150</td>
<td>0.548</td>
</tr>
<tr>
<td>VI. Counterfactual II with exogenous TFP ($A_r$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income pw</td>
<td>5,324</td>
<td>6,768</td>
<td>9,165</td>
<td>24,278</td>
<td>44,964</td>
<td>0.503</td>
</tr>
<tr>
<td>US/ctry ratio</td>
<td>19.3</td>
<td>15.2</td>
<td>11.2</td>
<td>4.2</td>
<td>2.3</td>
<td>-</td>
</tr>
<tr>
<td>Success</td>
<td>0.609</td>
<td>0.369</td>
<td>0.043</td>
<td>0.105</td>
<td>-0.188</td>
<td>0.305</td>
</tr>
<tr>
<td>VII. Counterfactual: Transposing the US urbanization share</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income pw</td>
<td>5,572</td>
<td>5,770</td>
<td>11,392</td>
<td>25,453</td>
<td>61,690</td>
<td>0.586</td>
</tr>
<tr>
<td>US/ctry ratio</td>
<td>21.8</td>
<td>21.0</td>
<td>10.7</td>
<td>4.8</td>
<td>2.0</td>
<td>-</td>
</tr>
<tr>
<td>Success</td>
<td>0.559</td>
<td>0.126</td>
<td>0.091</td>
<td>-0.007</td>
<td>-0.022</td>
<td>0.190</td>
</tr>
<tr>
<td>VIII. Counterfactual: Repatriation of emigrant workers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income pw</td>
<td>2,949</td>
<td>5,873</td>
<td>11,148</td>
<td>28,056</td>
<td>65,397</td>
<td>0.715</td>
</tr>
<tr>
<td>US/ctry ratio</td>
<td>41.4</td>
<td>20.8</td>
<td>11.0</td>
<td>4.4</td>
<td>1.9</td>
<td>-</td>
</tr>
<tr>
<td>Success</td>
<td>0.162</td>
<td>0.136</td>
<td>0.066</td>
<td>0.081</td>
<td>0.031</td>
<td>0.012</td>
</tr>
<tr>
<td>Part</td>
<td>Scenario Description</td>
<td>15th</td>
<td>25th</td>
<td>50th</td>
<td>75th</td>
<td>85th</td>
</tr>
<tr>
<td>------</td>
<td>---------------------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>I.</td>
<td>Observed levels</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>of income per worker</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Income pw (n)</td>
<td>8,112</td>
<td>6,529</td>
<td>14,916</td>
<td>31,476</td>
<td>77,508</td>
</tr>
<tr>
<td></td>
<td>Income pw (a)</td>
<td>1,044</td>
<td>3,102</td>
<td>2,303</td>
<td>6,170</td>
<td>17,531</td>
</tr>
<tr>
<td></td>
<td>US/ctry ratio (n)</td>
<td>18.1</td>
<td>22.5</td>
<td>9.8</td>
<td>4.7</td>
<td>1.9</td>
</tr>
<tr>
<td></td>
<td>US/ctry ratio (a)</td>
<td>11.5</td>
<td>3.9</td>
<td>5.2</td>
<td>1.9</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>College grads in n</td>
<td>0.524</td>
<td>0.726</td>
<td>0.823</td>
<td>0.880</td>
<td>0.911</td>
</tr>
<tr>
<td>II.</td>
<td>Counterfactual:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Transposing US skill shares in each sector</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Success (n)</td>
<td>0.612</td>
<td>0.630</td>
<td>0.442</td>
<td>0.349</td>
<td>0.059</td>
</tr>
<tr>
<td></td>
<td>Success (a)</td>
<td>0.738</td>
<td>0.703</td>
<td>0.535</td>
<td>0.447</td>
<td>0.385</td>
</tr>
<tr>
<td>III.</td>
<td>Counterfactual II</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>with exogenous TFP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>((A_r))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>and exogenous skill bias ((R^c_{\omega r}))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Success (n)</td>
<td>-0.015</td>
<td>0.280</td>
<td>-0.061</td>
<td>0.089</td>
<td>-0.208</td>
</tr>
<tr>
<td></td>
<td>Success (a)</td>
<td>-0.180</td>
<td>0.112</td>
<td>-0.146</td>
<td>0.032</td>
<td>0.086</td>
</tr>
<tr>
<td>IV.</td>
<td>Counterfactual II</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>with full TFP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>externality ((A_r))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>and full skill bias externality ((R^c_{\omega r}))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Success (n)</td>
<td>0.871</td>
<td>0.827</td>
<td>0.733</td>
<td>0.552</td>
<td>0.297</td>
</tr>
<tr>
<td></td>
<td>Success (a)</td>
<td>0.942</td>
<td>0.901</td>
<td>0.812</td>
<td>0.684</td>
<td>0.587</td>
</tr>
<tr>
<td>V.</td>
<td>Counterfactual II</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>with exogenous skill bias ((R^c_{\omega r}))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Success (n)</td>
<td>0.661</td>
<td>0.663</td>
<td>0.489</td>
<td>0.373</td>
<td>0.094</td>
</tr>
<tr>
<td></td>
<td>Success (a)</td>
<td>0.738</td>
<td>0.703</td>
<td>0.535</td>
<td>0.447</td>
<td>0.385</td>
</tr>
<tr>
<td>VI.</td>
<td>Counterfactual II</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>with exogenous TFP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>((A_r))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Success (n)</td>
<td>-0.165</td>
<td>0.208</td>
<td>-0.162</td>
<td>0.049</td>
<td>-0.254</td>
</tr>
<tr>
<td></td>
<td>Success (a)</td>
<td>-0.180</td>
<td>0.112</td>
<td>-0.146</td>
<td>0.032</td>
<td>0.086</td>
</tr>
<tr>
<td>VII.</td>
<td>Counterfactual:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Transposing the US</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>urbanization share</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Success (n)</td>
<td>-0.252</td>
<td>-0.084</td>
<td>-0.078</td>
<td>-0.043</td>
<td>-0.095</td>
</tr>
<tr>
<td></td>
<td>Success (a)</td>
<td>0.121</td>
<td>0.141</td>
<td>0.172</td>
<td>0.170</td>
<td>0.305</td>
</tr>
<tr>
<td>VIII.</td>
<td>Counterfactual:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Repatriation of</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>emigrant workers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Success (n)</td>
<td>0.085</td>
<td>0.087</td>
<td>0.017</td>
<td>-0.023</td>
<td>-0.006</td>
</tr>
<tr>
<td></td>
<td>Success (a)</td>
<td>0.136</td>
<td>0.149</td>
<td>0.108</td>
<td>0.033</td>
<td>0.103</td>
</tr>
</tbody>
</table>

Notes: These tables gives the level of income per worker of Cambodia, Ghana, Tunisia, Mexico and Greece for the baseline and the respective counterfactual scenario. In each table the last column reports the effect on the Theil index. 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. Parts III-VI are variants of Part II, with different assumptions on the technological externalities. Part VII 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. For each simulation, the success rate is the share of the wage ratio explained by 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).