Directed Technological Change Cross Country Income Differences: A Quantitative Analysis

Jerzmanowski, Michal and Tamura, Robert

Clemson University, Clemson University

1 August 2017

Directed Technological Change & Cross-Country Income Differences: A Quantitative Analysis

Michal Jerzmanowski & Robert Tamura

August 12, 2019

Abstract

Understanding international income differences requires measuring supplies of multiple production factors and their productivity. Recent work suggests that heterogeneous workers should be treated as imperfect substitutes. Using a model of endogenous directed technological change and a new data set on labor force composition we construct productivity for workers in three skill categories for 63 countries from 1910 to 2010 (up to 83 additional countries for 1950-2010). Rich countries use all skill categories more efficiently. Poor countries have a large technology adoption wedge, which prevents them from using low-skill labor more efficiently. Reducing the technology adoption wedges would have a much larger impact on living standards of living than skill upgrading of the workforce.

1 Introduction

Explaining the enormous disparities in incomes across countries is one of the most important goals of macroeconomics. In this paper, we apply the model of directed technological progress (Acemoglu, 1998) to think about these disparities by combining the theory with a new data set on output, capital, and educational attainment, which goes back to 1910 for many countries. We explore quantitatively how the critical mechanism of this theory – the link from skill composition of the labor force to the accumulation of knowledge and productivity – contributes to our understanding of the sources and evolution of cross-country variation in standards of living. Specifically, we use the model to help us compute the skill-specific productivity levels for 63 countries from 1910 to 2010, 128 countries over the shorter period of 1950 - 2010, and 146 countries in the most recent decades. We explore the levels and growth rates of these productivity measures and document the gaps between developed and developing countries and how they have changed over the last century. Additionally, our theoretical model features a technology adoption wedge, which prevents countries from adopting frontier-level technologies. This is similar to the TFP measure in the factor-neutral approach but is more tightly linked to the theoretical model. We calculate the value of this wedge and compare it across countries. Finally, we conduct counterfactual experiments in our model to shed light on the fundamental issue of the relative roles of productivity differences versus human capital endowments in shaping long-run development.

Since the seminal work of Hall and Jones (1999) and Klenow and Rodriguez-Clare (1998) we have accumulated substantial evidence suggesting that a very large fraction of differences in the level of development across countries can be attributed to differences in total factor productivity (TFP) and that endowments of physical and human capital play a much smaller role. Much research effort has been devoted to understanding where such large differences in productivity originate but, despite important progress, recent surveys conclude that there is still no consensus explanation for the causes of cross-country variation in TFP (Hsieh and Klenow 2010, Jones 2016).

We contribute to this literature by departing from the standard approach of treating productivity as factor-neutral and by using a new data set on labor force skill composition. It turns out that much of the research attempting to understand TFP differences employs the Cobb-Douglas production function approach and thus assumes that labor inputs of different skills are perfect substitutes. This implies that productivity is factor-neutral, i.e. the *relative* efficiency of workers with two different skill sets always remains constant. However, theoretical models starting with Atkinson and Stiglitz (1969), through more recent contributions in Basu and Weil (1997) and Acemoglu (1998, 2002), provide a rich theoretical framework for thinking about the levels of productivity and the direction of technological progress as being endogenously determined by relative supplies of factors, specifically the skill composition of the labor force. Subsequent empirical studies provide evidence suggesting that relaxing the factor-neutrality assumption may be a fruitful avenue for understanding cross-country income differences (Caselli 2005, Caselli and Coleman 2006, Jerzmanowski 2007). Moreover, the extensive literature on the evolution of U.S. wage inequality has convincingly demonstrated that different types of labor are not perfect substitutes and that skill-bias has characterized technological progress in the U.S. for the last several decades.

Unfortunately, unlike in the case of factor-neutral TFP, calculating skill-specific productivity levels poses a more significant challenge as it generally requires data on wages or returns to education (Caselli and Coleman 2006, Caselli and Ciccone 2013). However, such data often have sparse coverage and questionable quality. An essential part of our contribution lies in combining the directed technical change theory (DTC) with a new data set to quantify non-skill-neutral productivity levels without the need to use cross-country wage data. We extend the basic DTC model to include physical capital, barriers to innovation (technology adoption wedge), and international technology diffusion. We calibrate the parameter values of the model and use its equilibrium conditions to back out the skill-specific productivity levels and measures of the technology adoption wedge, requiring only data on output, factor inputs, and shares of the labor force with primary, secondary, and college education. Of note is the fact that, unlike most other papers in this literature, we are able to use three skill categories, which may by itself be a substantial improvement since evidence suggests that binary division of the labor force into skilled and unskilled groups may be too restrictive (See Acemoglu and Autor, 2015). All of our variables come from a new data set constructed by Tamura, Dwyer, Devereux, Baier (2019), which covers 168 countries, 147 of them over the period 1950-2010 and 63 over the period 1910-2010.[1]

We use our productivity and technology wedge measurements to: (1) study the historical patterns of directed technological change, (2) compare levels and rates of change of skill-specific productivity measures across countries, and (3) evaluate the contribution of non-neutral technology and adoption wedges to cross-country income differences. We find that rich countries use labor of all three skill categories more efficiently. However, low productivity levels in poor countries arise from high technology wedges. A counterfactual calculation that removes adoption wedges reveals that in their absence poor countries would, in fact, achieve higher productivity in the lower skilled sectors compared to more developed economies. The magnitude of this effect is significant for the most impoverished countries; some of them would see 10-15 fold increase in their output just from boosting the productivity of their least skilled workers.

Our results shed light on the puzzle of low (or negative) factor-neutral TFP growth. Most studies that look at growth rates of TFP, instead of levels, report that they also vary greatly across countries and often a significant fraction of countries is found to have negative TFP growth even over very long periods of time (Easterly and Levine 2001).[2] For example, in our data, 12.5% of countries had negative average annual factor-neutral TFP growth during the period 1950-2010, and 25% had growth that did not exceed a half percentage point. This

[1] For the empirical analysis we typically drop approximately 20 countries which are primarily resource-based economies.

[2] There are exceptions however, Tamura, Dwyer, Devereux, and Baier (2019) find that using intergenerational human capital accumulation reverses this result. When human capital accumulates across generations, the long run variation in log living standards and cross-sectional variation in growth rates are mostly captured by variations in log input or growth rate variations in inputs.

finding, if we think of TFP as a measure of technological knowledge, is hard to reconcile with the notion of diffusion of ideas and may even imply technological regress. It is sometimes explained by arguing that the catch-up potential – created by technological advances at the frontier – is offset by deteriorating institutional quality in developing countries. However, the adoption wedges we computed have by and large been falling, even in less developed economies. In fact, our calculations show that poor countries experienced relatively robust growth of college-specific productivity. Their GDP per worker growth remained low mostly because of their labor composition; they have very few workers in the higher skilled category, where frontier growth has been fast, and many more in the lower skill groups, where frontiers have stagnated and so, even in the presence of technology diffusion, potential for growth has been limited.

When we take up the question of the relative importance of factor endowments versus the technology wedge in explaining the current disparities of standards of living, we find that the wedge accounts for most of the variation in output per worker. Our counterfactual calculations show that reducing the technology wedge leads to substantial income gains in poor countries, whereas endowing those countries with the skill distribution of their more developed counterparts has surprisingly little effect on their standard of living. The reason for this is that despite growth in college-specific productivity in poor countries, their productivity in this sector is still very far behind that of the developed economies. Increasing their share of skilled workers would therefore not boost their overall output very much. On the other hand, removing adoption wedges would yield large productivity gains in the lower-skill sectors where most of their labor force is found.

2 Related Literature

Our work is part of the large literature which studies the causes of long-run economic development. Within this literature, much emphasis has been placed on understanding total factor productivity differences among countries (Hsieh and Klenow 2010, Jones 2016). The reason for this is that despite early evidence in favor of an important role of factors of production, such as physical and human capital (Mankiw et al., 1992), subsequent research strongly suggests that factors are not as important as the largely unexplained total factor productivity (Hall and Jones, 1999, and Klenow and Rodriguez-Clare, 1997). For example, Hsieh and Klenow (2010) use the standard development accounting approach to conclude that as much as 70% of cross-country income differences are due to TFP. It is common to interpret TFP as representing technology or knowledge more broadly making such large and persistent differences between economies surprising given the natural propensity of ideas to diffuse.

The prevailing approach in this literature has been to treat productivity as factor-neutral. However, several recent studies suggest that accounting for factor non-neutral productivity could be an important step in understanding such large TFP differences. For example, Caselli (2005) uses a CES production function with physical and human capital and backs out the productivity of each factor using profit maximization conditions and data on capital's income share. He finds that rich countries use human capital more efficiently but are less efficient at using physical capital than their poorer counterparts. Jerzmanowski (2007) follows an even more flexible approach by constructing a non-parametric estimate of the world technological frontier and finds that the role of factors can increase to as much as 50%. Caselli and Coleman (2006) study cross-country productivity differences allowing imperfect substitution between skilled and unskilled labor, as we do in this paper. They find that rich countries use skilled labor more efficiently than their low-income counterparts while the opposite is true for unskilled labor. Caselli and Ciccone (2013) demonstrate how development accounting with the assumption of perfect substitutability between skill types can be interpreted as the upper bound on potential income gains from changing labor force composition.[3]

In their computation of skill-specific productivity levels, Caselli and Coleman (2006), as well as Caselli and Ciccone (2013), rely on international data on wages and returns to education, while Caselli (2005) uses capital's income share data. All of these have the drawback of sparse coverage and often questionable quality. The advantage of our approach to gauging the extent of skill-bias under imperfect substitutability between skills is that by using theory-based equilibrium conditions our calculations bypass the requirement for cross-country wage or income share data. In doing so, we are most closely related to Gancia et al. (2013). The authors, building on earlier work by Gancia and Zilibotti (2009) and Acemoglu and Zilibotti (2001), construct a model with the same key elements as ours: directed technological change, capital accumulation, and technology diffusion subject to barriers. Among their key findings is that the technology wedges are large and have not fallen much among non-OECD countries since 1970 and that removing those barriers would significantly increase income levels in this group of countries. Some of our findings are similar to theirs while others are quite different. We discuss the differences between their approach and ours later in the paper.

We are also directly related to the mostly theoretical literature which attempts to ex-

[3]In another related paper, Malmberg (2016) shows that with imperfect substitutability of worker skill types, one way to explain large productivity differences between rich and poor countries is to assume that quality differences in skilled workers, with rich countries possessing higher quality skilled workers than poor countries. Malmberg does admit the possibility that this quality difference could be from differences in skill augmenting technology, which would be consistent with our approach.

plain the low TFP in poor countries. Acemoglu and Zilibotti (2001) provide a theoretical model (and some empirical evidence) supporting the view that poor countries cannot make use of best-practice technologies because they are incompatible with their low-skilled labor force. An alternative view, emphasized by Parente and Prescott (1994,1999) and Olson (1982, 1996) among others, argues the importance of the technology adoption wedge, a distortion which prevents countries from adopting otherwise readily available best-practice technologies or causes them to use such technologies below full potential. Such distortions could include corruption, licensing, excessive regulation, labor laws, or other arrangements that limit competition, or outright prohibit entry or otherwise distort the allocation of resources away from most productive uses. Our approach includes both a technology wedge and technologies that are skill-specific.[4]

3 Theoretical Approach

In this section, we set up our theoretical model, characterize its balanced growth path equilibrium, and discuss the relationship between relative skill supplies and the technology wedge on the one side and the levels of skill-specific productivity and relative wage of workers with different skill types on the other. We provide only the key equations of the model in this section leaving more detailed derivations to Appendix A.

3.1 The Model

Our model incorporates physical capital, technology diffusion, and a technology adoption wedge into the directed technology framework of Acemoglu (2002). The economy is closed to capital flows, and there is no trade of final or intermediate goods. Technology is allowed to diffuse across borders, but this process is not automatic; local innovators must spend resources on improving technology even though they enjoy the benefit of being able to tap into the world pool of knowledge. This is a direct cost of innovation, and it reflects the resources necessary to implement a given technology and adapt it to the local conditions. There is an additional cost that innovators must bear, which captures various obstacles to entrepreneurship, such as compliance with regulatory requirements, licensing fees, bribes, etc. We refer to this friction as the "technology adoption wedge" because it limits the entry of new innovating firms and

[4]Finally, we note that we are not the first to use calibrated theory to understand cross-country income differences. Klenow and Rodriguez-Clare (2005) calibrate an endogenous growth model with technology spillovers and Cordoba and Ripoll (2008) conduct a development accounting using an endogenous growth model. Both of those papers work with factor-neutral technology, an assumption we want to relax.

slows down the rate of technology adoption. The economy evolves in continuous time, but we drop time from equations where this causes no confusion.

Households

There is a continuum N of infinitely lived representative households with CRRA preferences, a discount rate of ρ and one of three skill types: high skilled (H), medium skilled (M), and low skilled (L). Population and type shares, s_i, are constant

$$\begin{aligned} H &= s_H N, \\ M &= s_M N, \\ L &= s_L N \end{aligned}$$

and

$$s_H + s_M + s_L = 1.$$

The households own physical capital and patents rights on innovation and maximize the present discounted value of an infinite stream of utility. The optimal consumption path obeys the familiar Euler equation

$$\frac{\dot{C}}{C} = \frac{1}{\theta}\left[r - \rho\right] = \frac{1}{\theta}\Big[(1-\tau)R - \delta - \rho\Big]$$

where ρ is the discount rate, θ is the CRRA coefficient, and τ is a capital wedge, and the interest rate r is equal to the rental rate minus the rate of depreciation $r = (1-\tau)R - \delta$.

Final Good

Final output is produced using intermediate goods which are skill-specific according to the following production function

$$Y = \{Y_H^{\frac{\varepsilon-1}{\varepsilon}} + Y_M^{\frac{\varepsilon-1}{\varepsilon}} + Y_L^{\frac{\varepsilon-1}{\varepsilon}}\}^{\frac{\varepsilon}{\varepsilon-1}} \tag{1}$$

Competitive firms (characterized below) produce the intermediate goods Y_H, Y_M and Y_L and sell them to competitive final output producers at prices P_i, $i = H, M, L$. We take the final good to be the *numereire* so that

$$\left[P_H^{1-\varepsilon} + P_M^{1-\varepsilon} + P_L^{1-\varepsilon}\right]^{\frac{1}{1-\varepsilon}} = 1. \tag{2}$$

Intermediate Goods, Machines & Capital

The sectors in our model are symmetric, so we can conserve space by presenting equations for the L-type sector only. Corresponding equations for the H and M sectors can be easily obtained by replacing L with H or M. The intermediate goods are produced using labor of a single skill type and machines designed for workers with that specific skill type in the standard "variety of machine inputs" manner. At a point in time, there is a continuum A_i of machine varieties available for each skill type $i = H, M, L$ and the number of varieties expand over time through an innovators process described below. Machine types are distinct in the sense that a skill type cannot use machines designed to be used by another skill type.

Specifically, the intermediate goods production functions are given by

$$\begin{aligned} Y_H &= \frac{1}{1-\beta} \int_0^{A_H} \chi_{jH}^{1-\beta} dj \, H^{\beta} \\ Y_M &= \frac{1}{1-\beta} \int_0^{A_M} \chi_{jM}^{1-\beta} dj \, M^{\beta} \\ Y_L &= \frac{1}{1-\beta} \int_0^{A_L} \chi_{jL}^{1-\beta} dj \, L^{\beta} \end{aligned} \tag{3}$$

where χ_{ji} is quantity of machines of variety j rented by the i-type intermediate goods producer.

The representative L-type intermediate goods firm maximizes profits by selecting optimal employment of unskilled workers L and the quantity of corresponding machines to buy. Labor is supplied in a competitive market, but machines are purchased from monopolists who own the blueprints and manufacture the machines. To produce machines, the monopolist machine suppliers use capital, which is rented in a competitive market at a rental rate R. One unit of physical capital can produce one machine of any variety, and machines depreciate at a rate of 100%. Following Aghion and Howitt (2008), each machine producing monopolist faces a potential imitator with cost $\upsilon > 1$ times higher than the original innovator's own marginal cost, i.e. the imitator uses υ units of capital to produce one machine. This implies that the profit-maximizing monopolist will set the price equal to a υ markup over her own marginal cost to just price the imitators out of the market (assuming $\upsilon^{-1} > 1 - \beta$, which is satisfied by the values we use). Profit maximization choices of the intermediate goods producers imply that relative prices of intermediate goods of two different types are given by

$$\begin{aligned}\frac{P_H}{P_L} &= \left(\frac{A_H H}{A_L L}\right)^{-\frac{\beta}{\sigma}}, \\ \frac{P_M}{P_L} &= \left(\frac{A_M M}{A_L L}\right)^{-\frac{\beta}{\sigma}},\end{aligned} \tag{4}$$

and relative wages of workers of two different types are

$$\begin{aligned}\frac{w_H}{w_L} &= \left(\frac{A_H}{A_L}\right)^{\frac{\sigma-1}{\sigma}} \left(\frac{H}{L}\right)^{-\frac{1}{\sigma}}, \\ \frac{w_M}{w_L} &= \left(\frac{A_M}{A_L}\right)^{\frac{\sigma-1}{\sigma}} \left(\frac{M}{L}\right)^{-\frac{1}{\sigma}},\end{aligned} \tag{5}$$

where $\sigma = 1 + (\varepsilon - 1)\beta$, is the elasticity of substitution between worker types (see Appendix A for details). Profit per variety of machines is equated and given by

$$\begin{aligned}\pi_H &= \left(\frac{\upsilon - 1}{\upsilon}\right) P_H^{1/\beta} H \left(\upsilon R\right)^{\frac{\beta-1}{\beta}}, \\ \pi_M &= \left(\frac{\upsilon - 1}{\upsilon}\right) P_M^{1/\beta} M \left(\upsilon R\right)^{\frac{\beta-1}{\beta}} \\ \pi_L &= \left(\frac{\upsilon - 1}{\upsilon}\right) P_L^{1/\beta} L \left(\upsilon R\right)^{\frac{\beta-1}{\beta}}\end{aligned} \tag{6}$$

Market clearing in the capital market pins down the relative share of physical capital used in production of machines of different varieties. Denoting by K_L the amount of physical capital devoted to production of L-type machines, we have

$$\begin{aligned}\frac{K_H}{K_L} &= \left(\frac{A_H}{A_L}\right)^{\frac{\sigma-1}{\sigma}} \left(\frac{H}{L}\right)^{\frac{\sigma-1}{\sigma}}, \\ \frac{K_M}{K_L} &= \left(\frac{A_M}{A_L}\right)^{\frac{\sigma-1}{\sigma}} \left(\frac{M}{L}\right)^{\frac{\sigma-1}{\sigma}}\end{aligned} \tag{7}$$

Using the above conditions, we can show that the reduced-form production function of intermediate goods production is given by

$$
\begin{aligned}
Y_H &= \frac{K_H^{1-\beta}(A_H H)^\beta}{1-\beta}, \\
Y_M &= \frac{K_M^{1-\beta}(A_M M)^\beta}{1-\beta}, \\
Y_L &= \frac{K_L^{1-\beta}(A_L L)^\beta}{1-\beta},
\end{aligned} \tag{8}
$$

Finally, we show that the marginal product of capital is equalized across intermediate good industries and that the equilibrium rental rate on capital is

$$R = \left(\frac{1-\beta}{\upsilon}\right)\frac{Y}{K}, \tag{9}$$

which together with the fact that the interest rate in this economy is given by $r = (1-\tau)R - \delta$ imply

$$r = (1-\tau)\left(\frac{1-\beta}{\upsilon}\right)\frac{Y}{K} - \delta, \tag{10}$$

where δ is the rate of depreciation of capital.

Innovation

Blueprints for new machines varieties are specific to the economy. They are invented by domestic entrepreneurs who hold perpetual monopoly rights over a given variety they have invented within the country. Technology diffusion means that these innovators benefit from the world pool of knowledge (i.e. the machine varieties invented elsewhere); however, their innovations are always specific to the local economy and cannot be traded internationally, neither as blueprints nor as physical machines. We can think of these innovations consisting in large part of adaptation of existing technology to local conditions. The following process governs discovery of new blueprints for each sector

$$
\begin{aligned}
\dot{A}_H &= \eta_H \left(\frac{A_H^W}{A_H}\right)^\varphi \frac{Z_H}{N}, \\
\dot{A}_M &= \eta_M \left(\frac{A_M^W}{A_M}\right)^\varphi \frac{Z_M}{N}, \\
\dot{A}_L &= \eta_L \left(\frac{A_L^W}{A_L}\right)^\varphi \frac{Z_L}{N},
\end{aligned} \tag{11}
$$

where A_i^W represents the world frontier technology for sector $i = H, M, L$, η_i is the productivity of research spending, and Z_i is the R&D expenditure on innovation or technology adoption in sector $i = H, M, L$. In Acemoglu's original directed technology formulation $\varphi = 0$. Since our focus is understanding productivity differences across countries, we relax these assumptions to allow for diffusion of technology ($\varphi > 0$). Another difference from Acemoglu's model, is that we assume research outcome is proportional to R&D expenditure *per worker*. We do this to eliminate the level scale effect, that is a situation where countries with larger populations have higher levels of productivity and income. One way to motivate such scaling of the R&D expenditure is to appeal to the notion of duplication of innovative effort, as for example in Klenow and Rodriguez-Clare (2005).

In order to innovate, the entrepreneurs must incur an entry cost X, which is the same in all sectors and represents expenses on overcoming obstacles to the introduction of new technologies, such as compliance with regulatory requirements, licensing fees, resistance from labor unions, bribes, etc. (Parente and Prescott 1994, 1999). This is what we referred to above as the "technology adoption wedge" because, as we will show below, it affects the rate of technology adoption and thus the equilibrium distance to the world frontier. Note that in Acemoglu's original formulation $X = 1$ but we need to relax this assumption to allow the model to account for cross-country differences in productivity.

Free entry into research implies that marginal benefit of extra innovation/adoption effort Z is equal to the cost,

$$\begin{aligned} \eta_H \left(\frac{A_H^W}{A_H}\right)^{\varphi} \frac{V_H}{N} &= X, \\ \eta_M \left(\frac{A_M^W}{A_M}\right)^{\varphi} \frac{V_M}{N} &= X, \\ \eta_L \left(\frac{A_L^W}{A_L}\right)^{\varphi} \frac{V_L}{N} &= X, \end{aligned} \tag{12}$$

which in turn pins down the relative value of blueprints as

$$\begin{aligned} \frac{V_H}{V_L} &= \left(\frac{\eta_H}{\eta_L}\right)^{-1} \left(\frac{A_H/A_H^W}{A_L/A_L^W}\right)^{\varphi}, \\ \frac{V_M}{V_L} &= \left(\frac{\eta_M}{\eta_L}\right)^{-1} \left(\frac{A_M/A_M^W}{A_L/A_L^W}\right)^{\varphi}, \end{aligned} \tag{13}$$

where V_i, $i = H, M, L$ is the value of a blueprint for a machine in sector i.

Finally, the value of a blueprint must satisfy the no-arbitrage condition

$$r\,V_L = \pi_L + \dot{V}_L \tag{14}$$

BGP Growth Rate & Interest Rate

We assume that there is an exogenously given constant growth rate of the frontier technology g and it is the same for technologies specific all skill types. It follows that along the balanced growth path the economy grows at this rate.

$$g = \frac{1}{\theta}\left[r^{*} - \rho\right]$$

where where ρ is the discount rate and θ is the CRRA coefficient. The BGP interest rate r^{*} therefore given by $r^{*} = \theta g + \rho$ and is common across countries.[5] Using the no-arbitrage conditions from (14) and the fact that along the BGP the value of a blueprint must be stationary ($\dot{V}_i = 0$), we get the following relationship between the value of a blueprint, profits and the interest rate

$$\begin{aligned} V_H &= \frac{\pi_H}{r^{*}}, \\ V_M &= \frac{\pi_M}{r^{*}}, \\ V_L &= \frac{\pi_L}{r^{*}}, \end{aligned} \tag{15}$$

It follows, using the expressions for profits (6) and for relative prices (4), that

$$\begin{aligned} \frac{V_H}{V_L} &= \left(\frac{A_H}{A_L}\right)^{-\frac{1}{\sigma}} \left(\frac{H}{L}\right)^{\frac{\sigma-1}{\sigma}}, \\ \frac{V_M}{V_L} &= \left(\frac{A_M}{A_L}\right)^{-\frac{1}{\sigma}} \left(\frac{M}{L}\right)^{\frac{\sigma-1}{\sigma}}, \end{aligned} \tag{16}$$

Finally, combining equations (13) and (16), yields

[5] The free capital tax wedge parameter τ (equation (10)) ensures that the interest rate parity is consistent with cross-country variation capital-output ratios.

$$
\begin{aligned}
\frac{A_H}{A_L} &= \left(\frac{\eta_H}{\eta_L}\right)^{\frac{\sigma}{1+\sigma\varphi}} \left(\frac{H}{L}\right)^{\frac{\sigma-1}{1+\sigma\varphi}} \left(\frac{A_H^W}{A_L^W}\right)^{\frac{\varphi\sigma}{1+\sigma\varphi}}, \\
\frac{A_M}{A_L} &= \left(\frac{\eta_M}{\eta_L}\right)^{\frac{\sigma}{1+\sigma\varphi}} \left(\frac{M}{L}\right)^{\frac{\sigma-1}{1+\sigma\varphi}} \left(\frac{A_M^W}{A_L^W}\right)^{\frac{\varphi\sigma}{1+\sigma\varphi}}
\end{aligned} \tag{17}
$$

Thus the relative levels of productivity are increasing in the relative supply of skilled workers H/L or semi-skilled workers, M/L, as long as $\sigma > 1$. Recall that we refer to technological change as skill biased (H-biased in our notation) whenever an increase in the level of technology raises the relative marginal product of skilled workers. From equation (17) it is clear that an increase in A_H/A_L is a skill-biased (an increase in A_M/A_L is a semi-skill biased) technological change as long as $\sigma > 1$. In Acemoglu's terminology, weak equilibrium skill-bias occurs whenever an increase in H/L or (M/L) induces skill-biased (semi-skill biased) technological change. Equation (17) thus implies that we have weak equilibrium skill-bias whenever $\sigma > 1$.

Substituting the expression for relative productivity levels (17) into the relative wage formula (5) we obtain

$$
\begin{aligned}
\frac{w_H}{w_L} &= \left(\frac{\eta_H}{\eta_L}\right)^{\frac{\sigma-1}{1+\sigma\varphi}} \left(\frac{H}{L}\right)^{\frac{\sigma-2-\varphi}{1+\sigma\varphi}} \left(\frac{A_H^W}{A_L^W}\right)^{\frac{\varphi(\sigma-1)}{1+\sigma\varphi}}, \\
\frac{w_M}{w_L} &= \left(\frac{\eta_M}{\eta_L}\right)^{\frac{\sigma-1}{1+\sigma\varphi}} \left(\frac{M}{L}\right)^{\frac{\sigma-2-\varphi}{1+\sigma\varphi}} \left(\frac{A_M^W}{A_L^W}\right)^{\frac{\varphi(\sigma-1)}{1+\sigma\varphi}}
\end{aligned} \tag{18}
$$

Notice in equation (18) that an increase in H/L (M/L), besides its effect on A_H/A_L (A_M/A_L), also works to reduce the relative skilled wage through the standard supply effects. When the increase in relative productivity is strong enough to offset this supply effect and lead to an increase in the relative wage of skilled workers following a rise in their relative supply, we refer to it as strong equilibrium skill bias. Clearly the *strong skill bias* is present as long as

$$\sigma > 2 + \varphi$$

which reduces to $\sigma > 2$, a result familiar from Acemoglu (2009), when $\varphi = 0$. Notice that the presence of international technology diffusion ($\varphi > 0$) implies a higher value of σ is required for strong bias to exist. This follows because the presence of technology diffusion means it is easier to free ride on the technological progress of the world technology frontier. Thus for an "upward sloping relative demand"for an increasing relative supply of factor requires an even higher bar in terms of elasticity of substitution between factors.

Using equations (12) and (15) we can show that on the BGP productivity relative to the frontier is given by

$$\mu_H \equiv \frac{A_H}{A_H^W} = \left[\frac{\eta_H \left(\frac{\upsilon-1}{\upsilon}\right) (H/N) P_H^{*1/\beta} \left(\upsilon R^*\right)^{\frac{\beta-1}{\beta}}}{r^* X} \right]^{1/\varphi}, \tag{19}$$

$$\mu_M \equiv \frac{A_M}{A_M^W} = \left[\frac{\eta_M \left(\frac{\upsilon-1}{\upsilon}\right) (M/N) P_M^{*1/\beta} \left(\upsilon R^*\right)^{\frac{\beta-1}{\beta}}}{r^* X} \right]^{1/\varphi},$$

$$\mu_L \equiv \frac{A_L}{A_L^W} = \left[\frac{\eta_L \left(\frac{\upsilon-1}{\upsilon}\right) (L/N) P_L^{*1/\beta} \left(\upsilon R^*\right)^{\frac{\beta-1}{\beta}}}{r^* X} \right]^{1/\varphi},$$

These equations imply that, all else equal, countries with a greater technology wedge (X) will find themselves further away from the frontier. Additionally, the productivity level will be further from the frontier when the price of machines is higher (υR), the supply of i-type workers is greater (i/N), and the price of i-type intermediate good is higher (P_i), $i = H, M, L$. The latter two terms are affected in opposite directions by a change in the skill supply ratio H/L or M/L. When this ratio increases, L/N falls but, as we demonstrate below, P_L increases.

Finally, substituting the above expressions into the aggregate final goods production function we get that (see Appendix A for derivation)

$$\frac{Y}{N} = \frac{A_L^{\beta} (\frac{K}{N})^{1-\beta} \left\{ \left[\left(\frac{K_H}{K_L}\right)^{1-\beta} \left(\frac{A_H s_H}{A_L}\right)^{\beta} \right]^{\frac{\varepsilon-1}{\varepsilon}} + \left[\left(\frac{K_M}{K_L}\right)^{1-\beta} \left(\frac{A_M s_M}{A_L}\right)^{\beta} \right]^{\frac{\varepsilon-1}{\varepsilon}} + \left[s_L^{\beta} \right]^{\frac{\varepsilon-1}{\varepsilon}} \right\}^{\frac{\varepsilon}{(\varepsilon-1)}}}{(1-\beta)(\frac{K_H}{K_L} + \frac{K_M}{K_L} + 1)^{1-\beta}} \tag{20}$$

We have data on output Y, labor force N, physical capital K, and skill-type shares s_i. However, these are not sufficient to solve for A_L in the above equation, which is our objective. This is where the theoretical machinery we have deployed makes a contribution to our empirical strategy. At the heart of this contribution are equations (17) and (18). They allow us to use historical data on U.S. skill premia to back out the value of $\frac{\eta_H}{\eta_L}$ and $\frac{\eta_M}{\eta_L}$ and use them to compute $\frac{A_H}{A_L}$ and $\frac{A_M}{A_L}$ and in turn find the allocation of capital across sectors (7), all of which are unobserved but necessary inputs for goal: the computation of skill-specific productivity levels. The next section gives the details of our approach.

4 Empirical Approach & Data

Our ultimate goal is to take the model developed above to the data and back out the skill-specific technology levels (A_i's) and the measures of the technology wedge (X). To achieve this goal we use data on output (Y), capital (K) and labor supply by three skill categories (at most primary school, some secondary, and some college) from Tamura et al. (2019).[6] From our point of view the most important aspect of this dataset is that it allows us to construct supplies of workers at three distinct education levels: at most primary, at least some secondary (but no college), and at least some college. The data is at a decadal frequency and covers 58 countries for the years 1820-2010 and 149 countries for the period 1950-2010, which is where we focus most of our analysis. However, we also need our model since the above data and the production function itself are not sufficient to identify the productivity levels (unlike in the case in the factor neutral approach). The model delivers the values of relative productivity (eq. 17) and capital allocation across sectors (eq. 7), both of which are unobserved but necessary for computing absolute values of productivity. In fact, it turns out that we need even more information than the model and the Tamura et al. (2019) data are able to supply. In order to back out the A_i's we also need to know the (relative) efficiency of the innovation production function (η's in 11), which we do not observe. To recover these parameters, we combine the wage equation of the model with historical data on the skill premia in the US from Goldin and Katz (2008). Our approach is, therefore, to choose values for the model's parameters and feed the data into its BGP conditions to produce country-specific productivity levels and the technology adoption wedge, the (A_L, A_M, A_H, X). One final difficulty that arises is that in order to perform this calculation we need to know the value of productivity at the world technology frontier. Since these are unobserved, we compute them by an iterative procedure. Essentially, we start with a (naive) guess of the frontier productivity levels and then we compute individual countries' productivity levels. Then we update the frontier to be a maximum productivity in the sample and we go back to computing country individual levels. We continue in this manner until the frontier values do not change. The details of this procedure are laid out in the next subsection.

[6] Those authors construct their series of estimates of real output, physical capital stock and labor force composition by education level based on multiple sources, which include Benavot and Riddle (1988), Bolt and van Zanden (2013), Mitchell (2003a,b,c), Picketty and Zucman (2014), Sabillon (2005) as well as *World* and *Human Development 2010* reports. See their Data Appendix (2018) for more detail.

4.1 Computing Productivity Levels & the World Technology Frontier

In order to back out the skill-specific technology levels (A_i's), we assume that the economies are on their balanced growth paths each time we observe them (i.e. every decade).

Our first step is to pick values for the following parameters: $\sigma, \varphi, \upsilon, \eta_L, \eta_M, \eta_H, \beta$. Unfortunately, for some of them there is very little guidance in the existing literature forcing us to make some judgment calls. β is labor's income share. We choose a value of 2/3, in agreement with Gollin (2002). We use a value of 1.4 for the markup based on the work of Ramey and Nekarda (2013) and Jones and Williams (2000). The technology diffusion parameter φ is set equal to 0.5. We choose this value to match the speed of convergence to the steady state of about 2.5% per year, see Barro and Sala-i-Martin (2003) (Appendix F discusses the model's transition dynamics). Lastly, we follow Gourinchas and Jeanne (2004) and choose the intertemporal elasticity of substitution to be 1, the time discount rate to be 0.04, the depreciation rate to be 6% and finally the frontier productivity growth to be 1.2%. Together these choices produce the equilibrium BGP interest rate (in the absence of capital distortions) of 5.2%.

The parameter σ, which measures the elasticity of substitution between labor of different skill types, is set to 2.6. This choice warrants more discussion as it may seem like a high value, given that many estimates in the literature, from the seminal work of Katz and Murphy (1992) to the more recent paper by Ciccone and Perri (2005), put it at around 1.6. However, the most recent work concludes that higher values of σ are also plausible. For example, using Katz and Murphy's regression on updated data Acemoglu and Autor (2011) find values around 2.9. Additionally, it is important to note that most of the papers in this literature are based on equation (5), whereas in the presence of directed technology and international knowledge spillovers, equation (18) provides the appropriate interpretation of the coefficient on the relative skill supplies, which in this case is not the inverse of the elasticity of substitution σ (and additionally depends on the diffusion parameter φ.) In a companion paper Jerzmanowski and Tamura (2017), we estimate σ using data from the EU KLEMS Growth and Productivity Accounts panel data set (O'Mahony and Timmer, 2009), and we interpret the estimated coefficient in a way that is consistent with the presence of technology diffusion and directed technological change. Generally, we estimate the elasticity to be considerably above 2 and, in some specifications, we are not able to reject the hypothesis that it is above 2.5, the value required for strong skill bias given our choice of $\varphi = .50$. Since we want to explore the implications of the directed technology paradigm, including strong skill-bias, for the world income distribution, we choose this higher value for σ. Appendix C briefly explains our approach and summarizes our estimates of σ using the EU KLEMS data.[7] Finally, recall that we calculate the relative research efficiencies η_H/η_L and η_M/η_L using equation (18) and the U.S. data historical series on relative wages and supplies of workers with different educational attainment. The values of these relative productivities are discussed in Appendix D.

Table 1 summarizes our parameter value choices.

Table 1: Parameter values used in the calculation of skill-specific productivity levels

Parameter	Value	Parameter's Role	Value Source
g	0.012	BGP growth rate of frontier technology	Gourinchas and Jeanne (2001)
ρ	0.02	time discount rate	Gourinchas and Jeanne (2001)
θ	1	intertemporal elasticity of substitution	Gourinchas and Jeanne (2001)
σ	2.6	elasticity of sub. between skill types	Acemoglu and Autor (2011); see text
β	2/3	labor's income share	Gollin (2002)
ϕ	0.5	strength of technology diffusion	see text
υ	1.4	mark-up of innovators	Jones and Williams (2000)
η's	varying	efficiency of R&D	Goldin and Katz (2008) with eq. (18)

We then proceed as follows.[8]

1. In the first iteration we assume $\frac{A_H^W}{A_L^W} = \frac{A_M^W}{A_L^W} = 1$.

2. Using the Goldin and Katz (2008) data on U.S. historical wage premia (with an update for 2010 using Acemoglu and Autor (2011)), the data from Tamura et al. (2019) for the associated relative skill supplies, we invert the two equations in (18) to solve for η's, where H is identified as those with at least some college, M is identified with more than primary schooling (some secondary) but less than college, and L which includes at most primary and those with no schooling at all. Technology adoption is driven by the productivity of investment resources, the η's, the technology frontier, A_W, and the country's skill distribution, (L, M, H). So the solution finds the η's that produce the evolution of technology in the US consistent with the evolution of wages.

[7] We have also produced a set of results with $\sigma = 1.6$. They are less plausible than under our preferred value. The appendix discusses those results briefly. Under this lower value, some of the results – especially the measures of the technology adoption wedge – seem highly implausible. Specifically, out of 146 countries, the U.S. ranks only the 26th in terms of the wedge, behind countries such as Hong Kong and Singapore (plausible), and Sudan and Mozambique (implausible). The magnitudes of the wedge differences also seem questionable, for example, Singapore's wedge at only about 32% of those in the U.S. seems too low. These results are available upon request.

[8] We generously thank an anonymous referee for providing a more compact, transparent and intuitive description of our solution algorithm.

3. Next we solve for the relative productivity levels using equations in (17). Thus while all countries have the same frontier technology access, and same η's, they have different factor endowments.

4. We next choose X to get the correct value of technology for the A's consistent with the country's output per worker. The X's determine the proportion of the frontier technology is used for each skill type in the country. See equation (66) in Appendix E for the expression we use in this step.

5. We assume the frontier in year t to be the maximum of observed productivity up to year t, i.e. $A_{it}^W = \max(A_{ih}|h \leq t)$. We replace the frontier numbers set in step 1 to the new values.

We iterate on steps (3)-(5) until the frontier values are no longer changing. That is we find the fixed point of the following problem

$$A_{itn+1}^W = \max\left\{\mathcal{A}_{ih}\left(\frac{A_{Htn}^W}{A_{Ltn}^W}, \frac{A_{Mtn}^W}{A_{Ltn}^W}, \mathcal{D}\right)|h \leq t\right\}$$

where $\mathcal{D}$ stands for our data and $\mathcal{A}_{ih}$ is the vector of sector i productivity levels in year h for all countries in our sample. The values of $\mathcal{A}_{ih}$ are those computed in steps (3)-(4), given the relative frontier values $\frac{A_{Htn}^W}{A_{Ltn}^W}$, and $\frac{A_{Mtn}^W}{A_{Ltn}^W}$ obtained in step (5) of the previous iteration of the algorithm (they equal 1 if this is the first iteration).[9]

Finally, we calculate the level of the technology adoption wedge relative to the US using the equations (4) and (55)

$$\frac{X_k}{X_{US}} = \left(\frac{P_{L,k}}{P_{L,US}}\right)^{1/\beta}\left(\frac{R_k^*}{R_{US}^*}\right)^{\frac{\beta-1}{\beta}}\left(\frac{A_{L,US}}{A_{L,k}}\right)^{\varphi}\left(\frac{s_{L,k}}{s_{L,US}}\right) \tag{21}$$

which allows us, using the A's computed previously as well as expressions (9) and the equations for intermediate good prices (see Appendix B), to compute the (relative) level of the wedge for each country.

Before proceeding to the findings, note that our ability to compute the productivity levels and wedges in a tractable manner rests in an important way on the assumption that countries

[9]Computing the frontier as the maximum of productivity levels in only the current year, not current and past years, produces very similar results. The only change is that the frontier productivity for the lowest skilled workers decreases in most recent decades. The individual country productivity levels are invariant because any change of the frontier values is offset by the values of η calibrated form the U.S. wage premia via eq. (18).

are on their respective balanced growth paths. When we carry out our calculations, we find that the growth rates of productivity differ across countries, sectors, and decades. We interpret them as movement from one BGP to another, which we feel is a reasonable (recall that there are 10 years between observations in our sample) but inevitably a somewhat limiting assumption.

5 Results

We begin by presenting the skill-specific productivity levels and examining how they differ across countries, and what they imply about the evolution of the factor bias of technology during our sample period.[10] Next, we present and discuss the measures of the technology wedge implied by our model. Finally, we evaluate the role of adoption wedge and skill-bias in explaining cross-country income differences. We confine our analysis to the years 1910-2010 since this is the period for which we have the most reliable US skill premium data, a crucial ingredient into our calibration exercise.

5.1 World Technology Frontier and Cross-Country Skill Bias

How has the skill-bias of the technology frontier evolved over the last century? Figure 1 shows the change of the relative skill bias of the world technology frontier by plotting the (log) of A_i/A_j where i and j each denote one of the three skill groups. During the first several 50 years of the 1900s, excluding the first two decades, when it briefly appears low-skill-biased, technological change is middle-skill-biased. Things change after 1950, at which point the productivity of college-educated workers begins to outpace the other two categories and high-skill-bias emerges (with a brief respite in 1970s when new technology briefly ceases to favor college-educated workers over secondary-educated workers).

Figure 2 plots the college/secondary productivity ratio for a selected group of countries. As is clear from these plots, the endogenous and directed nature of technological progress in our model is essential for our findings. Despite the diffusion of frontier technology across economies, the individual countries' paths of relative productivity levels do not closely mirror the frontier but also reflect changing skill composition and movements of the technology adoption wedge.

[10] Most papers in this area use a specification with only two skill types and call technological progress *skill-biased* (*unskill-biased*) if it increases (decreases) A_H/A_L. With three skill groups, we will say that technological progress *favors college-educated workers relative to secondary-school-educated workers* when A_H/A_M goes up, without any restriction on the concurrent change in A_H/A_L. We will say that technological progress is *college-biased*, if it favors college educated groups relative to both remaining skill groups, that is both A_H/A_M and A_H/A_L increase. Recall that , since we have assumed $\sigma > 1$, an increase in A_H/A_M will favors H over M.

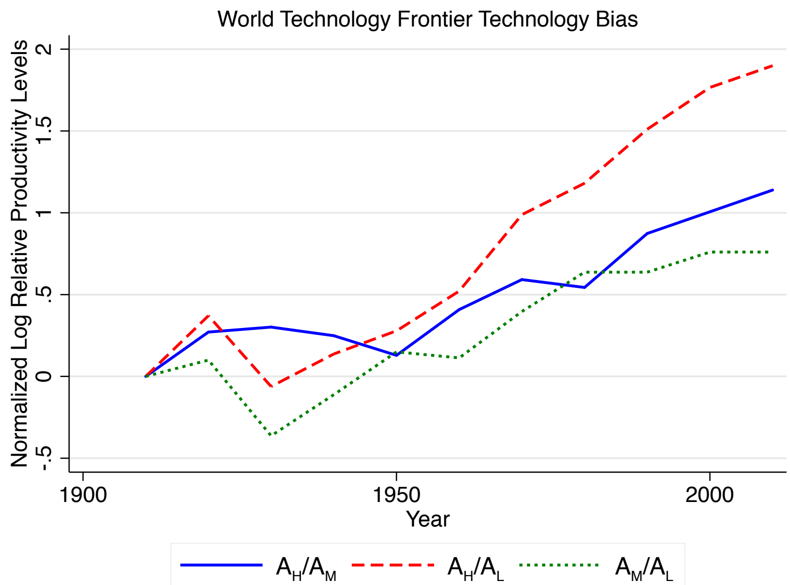
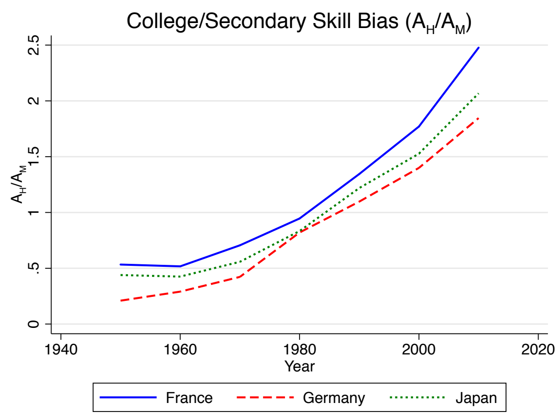


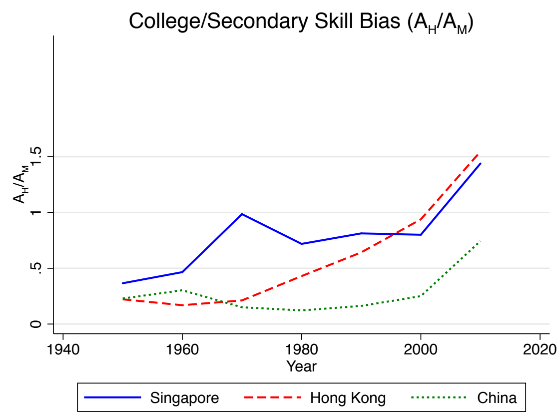
Figure 1: The skill bias of world technology frontier from 1910 to 2010

To investigate cross-country patterns in skill bias more systematically, we turn to regression analysis. Caselli and Coleman (2006) translate statements about the bias in the direction of technological change over time into statements of the bias across countries by replacing time with income per worker. Specifically, they refer to their finding that A_H/A_L increases with output per worker as *relative skill bias* and call *absolute skill bias* the instance when A_H is increasing and A_L decreasing with income level. They find strong evidence of the former and slightly weaker evidence of the latter in their data set. To investigate whether our results also indicate the presence of a skill bias across countries, in the sense of Caselli and Coleman, we regress the (log) productivity levels on (log) of output per worker. We do this using OLS for the most recent year in our sample (2010), as well as for our entire sample using fixed effects and OLS. Table 2 reports the results, which broadly imply that richer countries operate more productive technology at all skill levels. We show the fit of the above regressions in Appendix G. As is clear from these results, we do not find any support for *absolute skill bias* in the form of lower low-skill labor productivity in high-income countries, as reported by Caselli and Coleman (2006). There is, however evidence of the weaker version, i.e. the *relative skill bias*, in the sense that poorer countries use lower-skilled types of labor *relatively* more efficiently (i.e. A_M/A_H and A_L/A_H are decreasing in income).

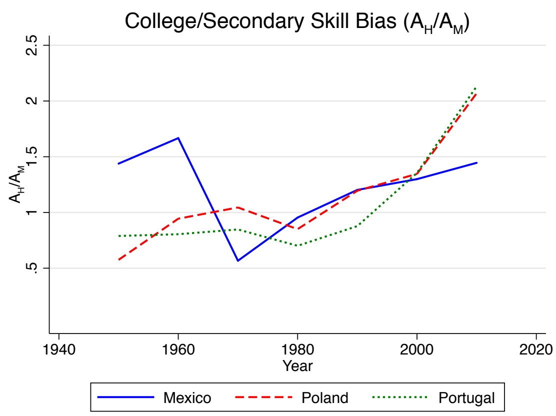
The caveat here is that the productivity levels used in the above analysis are the actual levels computed for each country, that is they are determined both by the country's skill



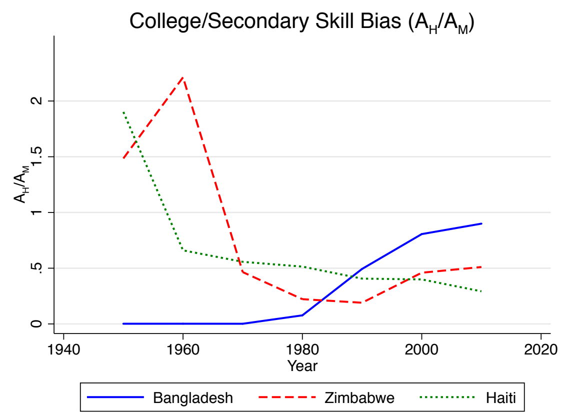
(a) A_H/A_M; France, Germany, and Japan.



(b) A_H/A_M; Singapore, Hong Kong, and China.



(c) A_H/A_M; Mexico, Poland, and Portugal.



(d) A_H/A_M; Bangladesh, Zimbabwe, and Haiti.

Figure 2: Evolution of the skill bias (college/secondary) in selected countries

endowments, via the directed technology channel, and by the level of the technology wedge. So our findings could arise simply due to richer countries having systematically lower wedges – a fact we indeed confirm in Section 5.3. If not for the distortion due to the adoption wedge, would rich countries still operate better technologies for low-skilled workers than the poor economies, where such labor type is much more abundant? Later we conduct a counterfactual exercise of removing the technology wedge and recomputing the productivity levels. We return to this question at that point.

	OLS (2010)	OLS (All Yrs.)	FE (All Yrs.)
College (A_H)	1.441***	1.915***	0.999***
	(0.037)	(0.030)	(0.067)
Secondary (A_M)	0.941***	1.471***	0.705***
	(0.025)	(0.021)	(0.055)
Primary (A_L)	0.353***	0.684***	0.721***
	(0.041)	(0.013)	(0.032)
N	146	1252	1252

Table 2: Skill Bias of Technology Across Countries. Regressions of the log productivity level for each skill type ($\log A_i$, $i = H, M, L.$) on a constant and the log of GDP per worker. OLS (2010) uses only 2010 data, OLS (All Yrs.) and FE (All Yrs.) use the entire 1950-2010 sample, with the addition of country fixed effects in the latter regression

5.2 Cross-Country Relative Levels and Growth of Productivity

A well-established finding in the growth literature is that both total factor productivity levels and growth rates, computed using the factor-neutral Cobb-Douglas production function, differ vastly across countries (Easterly and Levine 2001, Jones 2016). In our methodology, there are three distinct levels of productivity instead of the one factor neutral TFP. In this section, we summarize our cross-sectional findings about these skill-specific productivity levels and their growth rates and compare them to results obtained with the traditional factor-neutral TFP measures.[11]

Levels

Figure 3 shows the distributions of (log) productivity levels relative to the US in 2010. The black dash-dot line, which shows the factor-neutral TFP, illustrates the common findings referenced above: most countries have factor-neutral productivity levels lower than that of the US, and the dispersion is fairly large. The median TFP is only 35% of the US level, and the coefficient of variation is 77%. The US has one of the highest factor-neutral TFP levels, but nine countries – including Singapore, Hong Kong, and Taiwan, but also Sweden and Israel – have somewhat higher productivity. In our approach, each skill group has its own productivity level, and they do not all resemble the factor-neutral TFP distribution. First, we find that the U.S. has the highest productivity of college-educated workers (A_H) in the world in every decade since 1910. Comparing across countries, college-specific productivity has a greater dispersion (coefficient of variation of 104%) and a lower median (15% of the US level) than TFP. The productivity of primary-educated workers' is much less widely dispersed (coeff. of var. = 70%) and – even more importantly – has a median value of 144%, indicating that more than half of the countries have at least a 44% advantage in this sector over the US. Secondary school productivity levels fall in between, with the median country equal to 69% of US productivity and variation somewhat higher than that of TFP. About one-third of the countries are more productive in this sector than the US, including most East Asian economies and Japan, some European economies such as Italy and Ireland, as well as many former communist countries, among them the Czech Republic, Poland, and Latvia.

[11] The factor-neutral results obtained using a Cobb-Douglas production function

$$Y_i = K_i^{\alpha}(A_i h_i L_i)^{1-\alpha},$$

with human capital is given by

$$log(h_{it}) = .10E_{it}.$$

where E_{it} stand for average years of schooling.

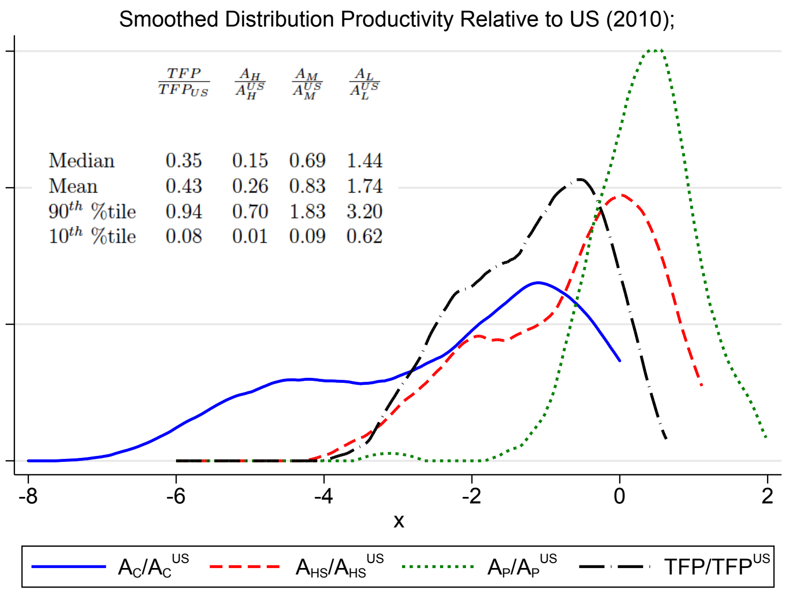


Figure 3: Smoothed distribution of the (log) productivity levels relative to the US (2010). Numbers reported in the table are levels of relative A's instead of logs.

Growth Rates

Table 3 summarizes the average annual growth rates of our skill-specific productivity measures and the factor-neutral TFP during the period 1950-2010. Consistent with many studies in this area, we find that factor-neutral TFP growth has a high variation across counties. While the median annual rate of increase was about 1%, in the top 10% of our sample TFP grew at least 2.7% per year while in the bottom 10% it has, in fact, declined. Of the countries we study, about 12.5% had negative average annual factor-neutral TFP growth during the period 1950-2010, and 25% had growth that did not exceed a half percentage point. If we think of TFP as a measure of technological knowledge, this set of facts is hard to reconcile with the notion of diffusion of ideas and may even imply technological regress. On the other hand, if TFP includes institutional quality, or the ability to prevent rent-seeking, then declining TFP is consistent with a group of countries suffering from social decay of market favoring institutions. Our skill-specific productivity growth rates display somewhat different patterns. First, the rate of change in college productivity was the highest, reflecting the fact that the world frontier was biased towards this skill group during the sample period. The median country's growth rate was 6.4%, and only the 10% slowest growers recorded annual increases of less than 4.1%. There was also more variation in the growth rates of college productivity than one would conclude from looking at TFP: countries in the top decile enjoyed annual college productivity growth about 6% faster than those in the bottom decile. The distribution of medium-skill productivity growth rates was similar but shifted to the left relative to college. However, the top decile had nearly an identical growth advantage over the bottom decile, 5.7% as the top-bottom decile gap for college, 6%. Finally, low-skill workers' productivity grew much slower on average, with roughly half of the countries registering declines in this measure. This is likely because many countries experienced a reduction – to nearly zero – of the share of workers with only primary level of education.

	TFP	A_H	A_M	A_L
Median	1.1	6.4	4.3	-0.1
Mean	1.2	6.9	4.6	0.0
90^{th} %tile	2.7	10.1	7.6	1.8
10^{th} %tile	-0.1	4.1	1.9	-1.9
Frontier Growth	1.0	3.2	1.5	0.5

Table 3: Summary statistics of annual growth rate of factor-neutral TFP and skill-specific productivity measurers during the period 1950-2010 (in percentage points per year) The bottom row provides the average growth rate of the world technology frontier (in case of TFP, it is defined as the U.S. TFP.)

The slow – and sometimes even negative – growth of the factor-neutral TFP has puzzled many researchers because it implied lack of technological progress (or even a regress), which are hard to reconcile with the natural tendency of knowledge to diffuse. Our findings indicate that slow or negative TFP growth is not as puzzling. In the sectors where the world technology frontier has increased the most, that is high-skill (college) and – to a lesser degree – medium-skill (secondary), even poorer countries recorded sizable productivity improvements. However, since these countries have relatively low proportions of their labor force in those sectors, the impact on output was small, leading to small or negative numbers when factor-neutral TFP growth was computed. This is similar to Rodrik's (2012) finding that manufacturing productivity levels – unlike GDP per capita levels – have generally been converging between poor and rich countries.

5.3 Technology Adoption Wedges

One of the outcomes of our computation is a measure of the technology adoption wedge for each country and period. This measure is akin to the standard TFP measure in the factor-neutral development accounting; it is the unobserved cross-country residual that is needed for the model to fit relative income levels observed in the data. Two features distinguish our wedge from the traditional TFP measures. First, it is more directly linked to a theoretical model and therefore contains more precise interpretation. Specifically, in the theoretical context of our model this wedge is most directly interpreted as a barrier cost of entry for new innovators but in reality is probably best understood as the compounded effect of the multiple frictions that inhibit incentives and abilities to innovate by new and existing firms. Second, and more crucial contribution, is that our wedge measures are based on factors-specific productivity production function and could, therefore, have different levels and change over time compared to TFP. In this section, we briefly discuss the measures of the technology wedge we have calculated.

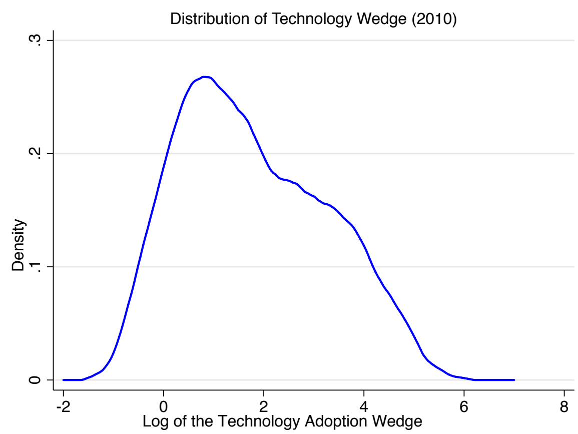


Figure 4: Smoothed distribution of the 2010 (log) technology adoption wedge relative to its value in the U.S.

Figure 4 shows the smoothed distributions of the log of the technology wedge in 2010. There is a lot of variation and a large number of economies have a wedge significantly higher than in the U.S. Only ten developed economies such as Singapore, Hong Kong, Taiwan, Ireland, and Norway had a wedge lower than the U.S. with Singapore's wedge equal to 55% of the US value and most others in this low-barrier group between 80% and 96% of the U.S. level. In general, these measures of the technology adoption wedge have a strong inverse relationship with income levels.

Over the last 60 years, the growth of global trade and capital flows, together with significant efforts to deregulate economies around the world, suggest that we should see the technology wedge fall over time for many, if not the majority, of world's countries. Our calculations imply this was indeed the case; during the period 1950-2010, the wedge declined steadily for most countries with the median rate of change of about -1% per year.

5.4 Productivity and Incomes without Adoption Wedges

As we discussed in the introduction, there is a large number of studies documenting the dominant role of the unobserved residual – TFP – in explaining the income gap between rich and poor countries. Is this also the case under the skill-specific productivity assumption of our model? To address this question, we compute the world income distribution under counterfactual productivity levels, which arise from reducing the technology wedge as a share of GDP per worker (X/y) to the U.S. level for all countries in the sample.[12] Throughout the calculations, we keep the level of frontier technologies constant and equal to the one recovered in our baseline calibration. We do this mainly because we do not want to take a stance on how the world technology frontier evolves, which would be required to construct the counterfactual values of the frontier. We will denote these counterfactuals by $\tilde{A}_i$, $i = H, M, L$.

This calculation is similar in spirit to the exercise of awarding every country in the sample the U.S. level of TFP, an approach typically followed in factor-neutral development accounting (Caselli 2005). However, since our model has more structure, reducing the adoption wedge to the U.S. level, does not necessarily imply productivity levels will be equalized because they depend on both the wedge and the factor supplies. In both counterfactuals, we keep the capital-output ratio unchanged, which corresponds to keeping all investment distortion (τ in equation 10) unchanged. With the counterfactual productivity levels in hand, we also revisit the question of cross-country skill bias, that is we ask whether after we have removed the burden of the technology wedge, do poor countries exhibit absolute advantage in lower-skill technologies. We then discuss what the removal of wedges does for international productivity and income differences.

Productivity Differences in Absence the Technology Adoption Wedge

The endogenous and directed aspect of technological progress in our model suggests that countries should enjoy higher productivity in sectors where they possess abundant labor endowments. Since poor countries have much higher proportions of their labor force in the lower skilled categories, we might expect their productivity in those sectors to be higher than those of rich countries, as reported – in the context of a different model – by Caselli and Coleman (2006). Recall however that in Section 5.1, where we compared skill-specific productivity levels across countries, we concluded that rich countries enjoy higher productivity levels for all three

[12] Since we are only able to compute relative wedges we choose the U.S. – because it is generally recognized as a relatively low barrier, high productivity economy – and construct the counterfactuals by setting the technology adoption wedge so that, after we calculate the new level of output per worker, the wedge-to-GDP per worker ratio is the same as in the U.S. We leave the wedge unchanged for countries with lower wedge-to-GDP per worker ratio than in the U.S., which happens in a handful of cases.

skill-types. We have just seen in Section 5.3, that poor countries have, on average, considerably higher technology adoption wedges. How does removing those wedges affect productivity differences between rich and poor countries? To answer this question we again look at the relationship between productivity levels and GDP per worker, except this time we use our *counterfactual* productivity levels computed by forcing the technology adoption wedge (as a share of GDP per worker) to be the same across countries and equal to that of the U.S. Table 4 reports the regressions of the log *counterfactual* productivity levels on GDP per worker. College productivity ($\tilde{A}_H$) is again higher in rich countries. The medium skilled worker (secondary school) productivity ($\tilde{A}_M$) is also higher in rich countries under the OLS estimation, but the coefficients are significantly lower implying that there is a relative bias across countries in the sense that poorer countries have higher secondary-specific productivity relative to the college-specific one (i.e. $\tilde{A}_M/\tilde{A}_H$ is decreasing with income). This is similar to what we found in section 5 using actual (not counterfactual) productivity numbers. However, for primary-specific productivity ($\tilde{A}_L$) – and in the case of fixed-effects estimates, also (secondary school) productivity ($\tilde{A}_M$) – we now find evidence of absolute bias in the sense that poorer countries would (in the absence of adoption wedges) enjoy a higher absolute primary-specific productivity level than their more developed counterparts. We conclude that, according to our data and the structure imposed by our model, in the absence of the technology wedge differences, poor countries would actually be able to operate their low-skill labor at higher efficiency than more developed economies. Since many of these economies have a large share of their labor force in this category, one would expect that reduction of the technology adoption wedges would produce a significant increase in their standards of living and this is the calculation we turn to next.

Counterfactual Income Levels

What is the impact of the technology wedge on the world income distribution? Table 5 reports the summary statistics for the data (first column), our wedge-equalization counterfactual where adoption wedges are set to the U.S. level (column two), and the factor-neutral counterfactual, where all countries receive the U.S. TFP value (column three).

Note first that the 2010 data clearly show the high degree of income disparities; the mean output per worker in the sample is \$21,890, and the median is \$14,592, while the coefficient of variation is 0.92 and the 90/10 percentile ratio is almost 24. Removing wedges shifts a significant portion of the distribution to the right; median income almost triples while the 90/10 gap falls to 2.2. In fact, the counterfactual distribution from our model is quite similar to the one obtained using the traditional factor-neutral model, where the counterfactuals are computed

	OLS (2010)	OLS (All Yrs.)	FE (All Yrs.)
College ($\tilde{A}_H$)	0.649^{***}	1.028^{***}	0.141^{**}
	(0.038)	(0.029)	(0.063)
Secondary ($\tilde{A}_M$)	0.150^{***}	0.584^{***}	-0.155^{***}
	(0.023)	(0.021)	(0.050)
Primary ($\tilde{A}_L$)	-0.438^{***}	-0.204^{***}	-0.142^{***}
	(0.035)	(0.010)	(0.024)
N	146	1252	1252

Table 4: Skill Bias of Technology Across Countries (In Absence of Adoption Wedge). Regressions of the log *counterfactual* productivity level for each skill type ($\log \tilde{A}_i$, $i = H, M, L.$) on a constant and the log of GDP per worker. OLS (2010) uses only 2010 data, OLS (All Yrs.) and FE (All Yrs.) use the entire 1950-2010 sample, with the addition of country fixed effects in the latter regression. Counterfactual productivity levels are computed using the same approach as actual levels but additionally forcing the technology adoption wedge (as a share of GDP per worker) to be the same across countries and equal to that of the U.S.

by increasing every country TFP to the level of the U.S. However, the underlying mechanisms are quite different. Obviously, in the factor-neutral counterfactual, low-productivity countries see their output rise because of an increase in the catch-all TFP, which uniformly increases the productivity of all labor-skill types. In our directed technology model, the transformation is different. Figures 5-7 illustrate the change in productivity levels that result from wedge reduction for each of the skill-types by plotting the actual level versus the counterfactual one. Note first that after reducing the wedges, countries with low college-specific productivity (mostly poor countries with low shares of skilled labor) experience relatively small productivity gains in that sector (Figure 5). This is because of the strong skill-bias present in our model; these countries have low shares of college labor and thus a small market for machines compatible with high skills. As a consequence, even when the technology wedge is low, the incentives to innovate for this sector are weak. However, countries with higher shares of highly skilled labor can register quite large productivity gains. Specifically, notice that there are possibilities of leap-frogging; countries such as Spain or Russia whose actual college-specific productivity is lower than, say, Germany's, end up with higher efficiency in that sector once wedges are reduced. This is a consequence of their higher shares of college-educated workforce.[13] We can also see that the results for the high-skilled sector contrast with those for the two lower skill sectors, especially the primary one (Figures 6 and 7). In these sectors, less-developed

[13]The obvious caveat here is that we are not controlling for the quality of education.

economies experience substantial increases in productivity once wedges are reduced. This, of course, is just the opposite of what happened in the college sector; they have relatively large shares of labor in the lower-skill sectors and that, combined with the reduction in the technology wedge, creates strong incentives to innovate and adopt frontier technology.

	Data	DTC	Neutral
Median	14,592	42,481	43,418
Mean	21,890	44,479	44,791
Coeff. Of Variation	0.92	0.30	0.33
90/10	23.8	2.2	2.5

Table 5: Summary statistics for the data (first column), our wedge-equalization counterfactual where adoption wedges are set to the U.S. level (column two), and the factor-neutral counterfactual, where all countries receive the U.S. TFP value (column three).

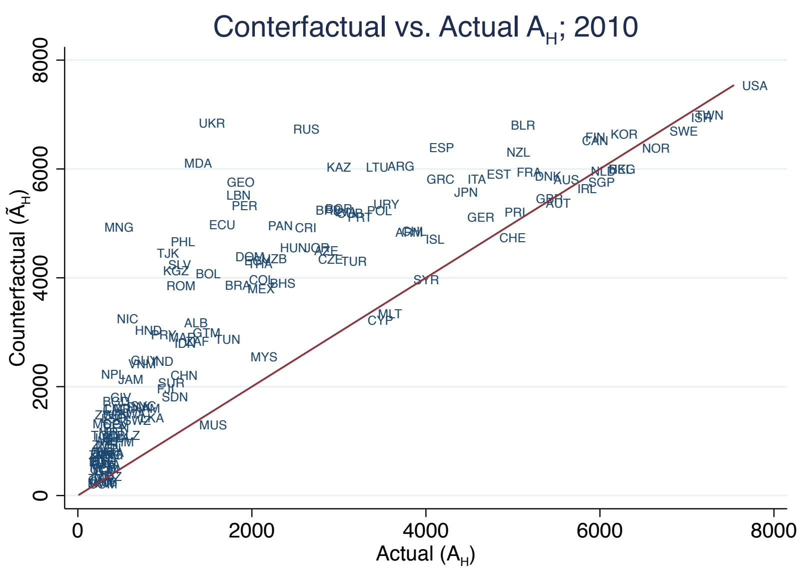


Figure 5: Counterfactual versus actual 2010 college-specific productivity level in 2010 (A_H). Counterfactual productivity levels are computed using the same approach as actual levels but additionally forcing the technology adoption wedge (as a share of GDP per worker) to be the same across countries and equal to that of the U.S.

The above calculations imply that reducing the technology wedge has a powerful effect on the world distribution of standards of living, with an especially substantial shift in the lower part of the distribution as a consequence of big income gains in poor countries. Because of the endogenous and directed innovation process, removal of technology wedges causes (most)

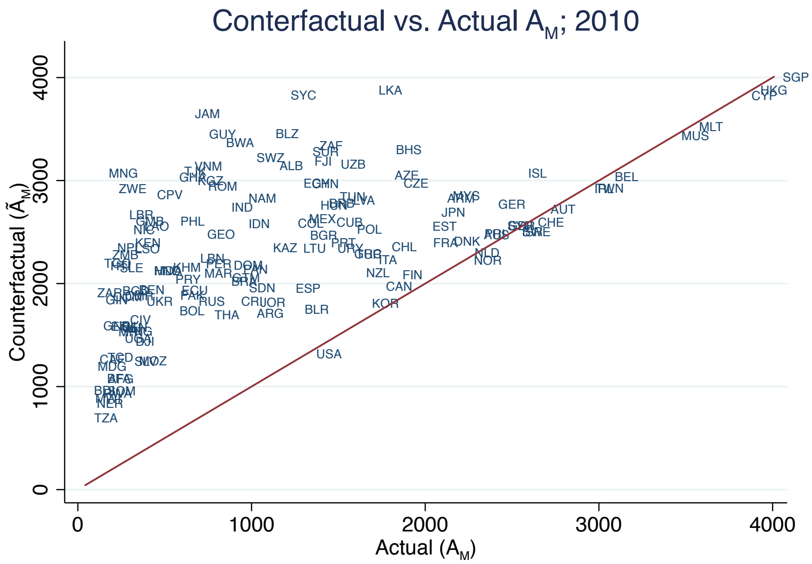


Figure 6: Counterfactual versus actual 2010 secondary-specific productivity level in 2010 (A_M). Counterfactual productivity levels are computed using the same approach as actual levels but additionally forcing the technology adoption wedge (as a share of GDP per worker) to be the same across countries and equal to that of the U.S.

poorer countries to adopt better technologies for the lower-skill sector, which often end up surpassing the low-skilled technologies currently being used by more developed economies. The magnitude of this effect can be illustrated by calculating output lost by poor countries because their unskilled workers' productivity is artificially depressed by technology wedges.[14] Suppose we calculate a slightly different counterfactual output from the one above; instead of giving each country the reduced-wedge productivity in all sectors, we only do it in the lowest skill sector. That is, we compute the new counterfactual GDP by setting each country's primary-specific productivity at the reduced-wedge counterfactual value computed above ($\tilde{A}_L$) but we keep productivity values for the other two sectors at their actual levels. We can then define output loss due to depressed unskilled productivity as the ratio of this counterfactual to the actual output. Figure 8 plots this measure against GDP per worker relative to the U.S. The magnitude of this effect is very large for the most impoverished economies; some of them would see 10-15 fold increase in their output just from boosting the productivity of their least skilled workers.

We note that these findings are consistent with the theoretical results in Acemoglu and Zilibotti (2001).[15] In those models, the consequence of different skill distributions in rich

[14]The authors wish to thank Daron Acemoglu for suggesting this calculation.

[15]Earlier models such as Atkinson and Stiglitz (1969) and Stewart (1977) considered appropriateness of

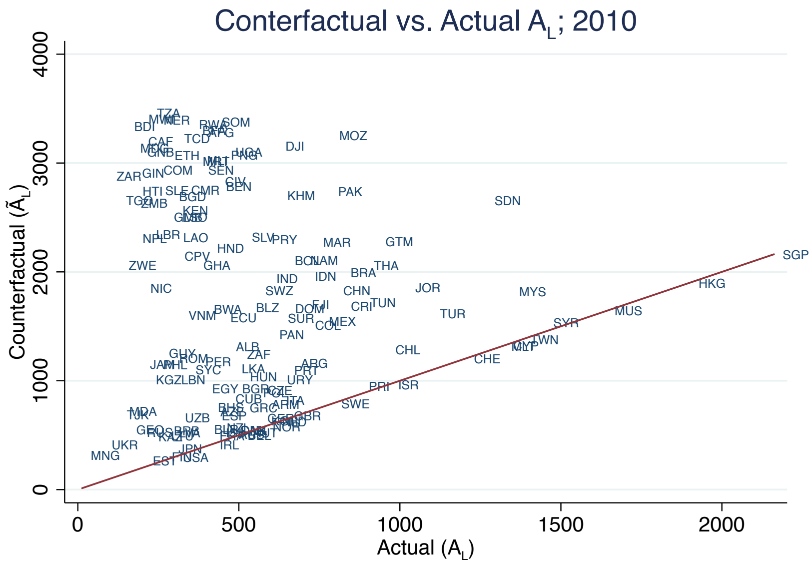


Figure 7: Counterfactual versus actual 2010 primary-specific productivity level in 2010 (A_L). Counterfactual productivity levels are computed using the same approach as actual levels but additionally forcing the technology adoption wedge (as a share of GDP per worker) to be the same across countries and equal to that of the U.S.

and poor countries is that the rich-invented technologies that are an inappropriate fit for the workers in poor countries. This results in low measured TFP even when they have equal access to technology. The high-skill productivity level does not increase in those countries by nearly as much. Therefore, the growth from wedge removal is generated not by adopting the same set of technologies that are currently used in developed countries but instead technologies more suited to the local factor endowments.[16] Our model with skill-specific sectors does not readily map into one with a manufacturing/services distinction, but it is worth emphasizing this last result in the context of recent work, which suggests that developing countries miss out on growth opportunities by de-industrializing prematurely and adopting the sectoral composition of their more developed counterparts (Rodrik 2015).[17]

technology centered around capital-labor ratios or plant size. They specifically did not consider differential labor skill.

[16] Of course, these technologies were presumably used in the currently-rich countries at some point in the past.

[17] Gancia et al. (2013), whose model is similar to ours, estimate the structural equations to obtain measures of the technology wedge, instead of calibrating them as we have done. In their preferred specification, the removal of the wedge results in an increase in GDP per worker relative to the U.S. from 19% to 61% in an average non-OECD country and from 68% to 91% for an average OECD country. These are comparable but somewhat higher than our findings of increases from 21% to 53% for non-OECD countries and 71% to 82% for OECD economies. The fact that these increases are smaller is at least in part due to the fact that in our model removal of the wedge does not propel countries all the way to the technology frontier (productivity

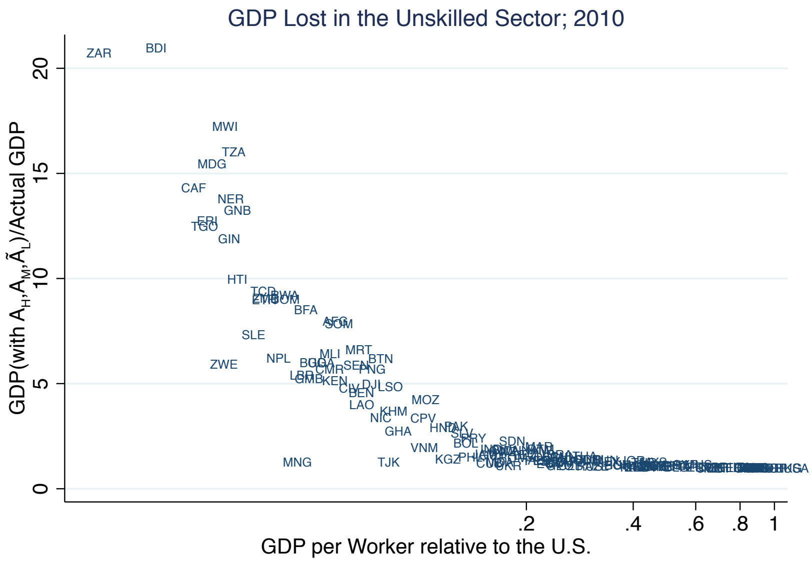


Figure 8: Output loss ratio vs. GDP per worker (2010). The output loss ratio is the ratio of counterfactual GDP (calculated with primary technology at the reduced-wedge counterfactual value ($\tilde{A}_L$) but productivity values for the other two sectors kept at their actual levels) to actual GDP.

Adoption Wedge vs. Human Capital

Finally, we consider how the income gains from reduction of the technology wedge compare to what would happen if factor endowments were equalized across countries? To do this, we compute a second set of counterfactuals, which is analogous to the one above except that we now keep wedges (as a percentage of GDP per worker) unchanged and instead endow every country with the skill distribution of the U.S.[18] These counterfactuals are similar to Caselli and Ciccone's (2013) calculations. We find that the income gains from changing the skill composition of the labor force are considerably lower than those we saw when we removed the wedges: For example, the median GDP per worker rises to \$22,357 compared to \$42,481 achieved by reducing technology adoption wedges. Thus even when endowed with highly skilled workers, the adoption wedge prevents most poor countries from operating close to the technology frontier. Appendix H presents a more detailed comparison of our results to those in Caselli and Ciccone (2013).

levels continue to depend on local skill endowments.) See Appendix D for more detailed comparison of the two approaches.

[18] As before we allow physical capital to respond so that the K/Y ratio remains unchanged.

6 Conclusions

We have explored quantitatively how the theory of directed technological progress contributes to our understanding of the sources and evolution of cross-country variation in standards of living. Specifically, we set up a model of endogenous directed technology with human capital, physical capital, and barriers to innovation/technology diffusion (technology adoption wedges). Using a new dataset, we calibrated the equilibrium conditions of the model and computed the skill-specific productivity levels and measures of the technology adoption wedge for 128 economies over the period 1950-2010. We used these skill-specific productivity and wedge measurements to study the historical patterns of directed technological change across countries and over time. We then evaluated the role of non-neutral technology and adoption wedges in explaining the current cross-country of GDP per worker.

Our estimates of the world technology frontier imply that in the early part of the 20th-century technological progress favored secondary-school-educated workers over those with more education. However, after 1950 – with a brief interruption in the 1970s – college-specific frontier productivity growth outstripped the other categories. Comparing skill-specific productivity across countries, we find that rich countries use labor of all three skill categories more efficiently. However, the gaps are not the same for each skill category. They are largest in the productivity of high-skilled workers but considerably smaller in the lower skilled sectors. Overall, greater technology adoption wedges are the reason behind low aggregate productivity in poor countries. Our counterfactual calculation reveals that in their absence developing economies would achieve considerable income gains, most of which would come from higher productivity in the lower-skilled sectors. This happens as a consequence of the interaction between the directed nature of technological change of our model and the skill endowments of poor countries. The magnitude of this effect is very large for the most impoverished economies; some of them would see 10-15 fold increase in their output just from boosting the productivity of their least skilled workers.

Furthermore, our calculations show that poor countries enjoyed relatively robust growth of high-skill-specific productivity. Their GDP growth failed to reflect that because of their labor composition. Developing countries have very few workers in the higher skilled category, where the world technology frontier growth has been fast, and many more in the lower skill groups where the frontier has stagnated. Thus even in the presence of technology diffusion, the potential for growth has been limited.

When we take up the question of the relative importance of factor endowments versus the technology wedge in explaining the current disparities of standards of living, we find that it is the wedge that explains most of the variation in output. The reason for this finding is that despite impressive growth in college-specific productivity in many poor countries, their productivity in this sector is still very far behind that of the developed economies. Increasing their share of skilled workers would therefore not boost their overall output very much. On the other hand, removing adoption wedges would yield substantial productivity gains in the lower-skill sectors where most of their labor force is found. Our theory suggests interpreting the adoption wedges as barriers to entry of new firms or obstacles to innovation and transformation at existing businesses, of regulatory, bureaucratic or institutional nature. However, we should be cautious as the wedges we computed remain unobserved residuals, and more work is needed to identify their precise nature and to reach more tangible policy prescriptions.

References

Acemoglu, Daron. "Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality " *Quarterly Journal of Economics* 113, 1998: 1055-1090.

Acemoglu, Daron. "Directed Technical Change " *Review of Economic Studies* 69, 2002: 781-809.

Acemoglu, Daron. "Patterns of Skill Premia " *Review of Economic Studies* 70, 2003: 199-230.

Acemoglu, Daron. "Equilibrium Bias of Technology " *Econometrica* 75, 2007: 1371-1410.

Acemoglu, Daron. *Introduction to Modern Economic Growth.* Princeton University Press, New York, 2009.

Acemoglu, Daron, and Autor, David. "Skills, Tasks and Technologies: Implications for Employment and Earnings" in *Handbook of Labor Economics 4B* Orley Ashenfelter and David Card editors, North Holland, San Diego, 2011.

Acemoglu, Daron, Zilibotti, F. "Productivity Differences," *Quarterly Journal of Economics* 116, 2001: 563-606.

Aghion, Philippe, and Howitt, Peter. *The Economics of Growth*, MIT Press: Cambridge, MA, 2008.

Atkinson, A.B., Stiglitz, J. "A New View of Technological Change", *Economic Journal* 79, 1969: 573-578.

Autor, David, Murnane, Richard J., and Levy, Frank. "The Skill Content of Recent Technological Change: An Empirical Exploration", *Quarterly Journal of Economics* 118, 2003: 1279-1334.

Bairoch, Paul. "The Paradoxes of Economic History: Economic Laws and History", *European Economic Review* 33, 1989: 225-249.

Barro, Robert J., and Sala-i-Martin, Xavier. *Economic Growth.* MIT Press, Cambridge, 2003.

Barro, Robert J.. "Convergence and Modernisation", *Economics Journal* Volume125, Issue 585, June 2015, Pages 911-942

Benavot, A., Riddle, P. "The Expansion of Primary Education, 1870-1940: Trends and Issues " *Sociology of Education* 61, 1988: 191-210.

Bloom, Nick, Jones, Charles, Reenen, John Van, Webb, Michael, "Are Ideas Getting Harder to Find?", working paper, 2018

Bolt, Jolt, and van Zanden, Jan Luiten. "The First Update of the Maddison Project: Reestimating Growth Before 1820", `http://www.ggdc.net/maddison/maddison-project/home.htm`, 2013 version.

Casseli, Francesco. "Technological Revolutions", *American Economic Review* 89, 1999: 78-102.

Casseli, Francesco. "Accounting for Cross-Country Income Differences "in *Handbook of Economic Growth*, Philippe Aghion and Steven Durlauf, eds. North Holland, 2005.

Caselli, Francesco and Ciccone, Antonio. "The contribution of schooling in development accounting: results from a nonparametric upper bound", *Journal of Development Economics*, 104, 2013: 199-211.

Caselli, Francesco and Coleman, Wilbur John "The world technology frontier" *American Economic Review*, 96, 2006: 499-522.

Ciccone, Antonio, and Peri, Giovanni. "The Long-Run Substitutability Between More and Less Educated Workers: Evidence from U.S. States, 1950-1990", *The Review of Economics and Statistics* 87, 2005: 652-663.

Cordoba, Juan-Carlos, Ripoll, Marla. "Endogenous TFP and Cross-Country Income Differences", *Journal of Monetary Economics* 55, 2008: 1158-1170.

Easterly, William, and Levine, Ross. "It's Not Factor Accumulation: Stylized Facts & Growth Models", *World Bank Economic Review* 15, 2001: 177-219.

Feyrer, James. "Convergence by Parts", *B.E. Journal of Macroeonomics: Contributions* 2008: article 19.

Gancia, Gino A, Müller, A., Zilibotti, Fabrizio. "Structural Development Accounting" in *Advances in Economics and Econometrics: Tenth World Congress*, eds. Daron Acemoglu, Manuel Arellano; Econometric Society, 2013.

Gancia, Gino A., and Zilibotti, Fabrizio. "Technological Change and the Wealth of Nations", *Annual Review of Economics* 1, 2009: 93-120.

Goldin, Claudia, and Katz, Lawrence F. *The Race Between Education and Technology* Cambridge, MA: Harvard University Press, 2008.

Gollin, Douglas. "Getting Income Shares Right " *Journal of Political Economy* 110, 2002: 458-474.

Goos, Maarten, and Manning, Alan. "Lousy and Lovely Jobs: The Rising Polarization of Work in Britain", *Review of Economics & Statistics* 89, 2007: 118-133.

Gourinchas, Pierre-Olivier, and Jeanne, Olivier. "The Elusive Gains from International Financial Integration", *IMF* working paper, 2004.

Hall, Robert E. "The Relation Between Price and Marginal Cost in U.S. Industry "*Journal of Political Economy* 96, 1988: 921-947.

Hall, Robert, and Jones, Charles I. "Why Do Some Countries Produce So Much More Output per Worker Than Others?" *Quarterly Journal of Economics* 114, 1999: 83-116.

Howitt, Peter. "Endogenous Growth and Cross Country Income Differences", *American Economic Review* 90, 2000: 829-846.

Hsieh, C.T., Klenow, Peter J. "Development Accounting", *American Economic Journal: Macroeconomics* 2, 2010: 207-223.

Hsieh, C.T., Klenow, Peter J., 2009. "Misallocation and manufacturing TFP in China and India", *Quarterly Journal of Economics* 124 (4), 1403?1448.

Jerzmanowski, Michal. "TFP Differences: Appropriate Technology vs. Efficiency", *European Economic Review* , Volume 51, Issue 8, November 2007, p. 2080-2110.

Jerzmanowski, Michal, and Tamura, Robert. "Substitutability Between Workers With Different Skill Levels: Evidence from Cross-Country Manufacturing Data", Clemson University working paper, 2017.

Jones, Charles I., "R&D-Based Models of Economic Growth,"Journal of Political Economy, August 1995, 103(4), 759?784.

Jones, Charles I., "The Facts of Economic Growth", *Handbook of Macroeconomics*, 2016, Vol. 2A, pp. 3-69.

Jones, Charles I., and Klenow, Peter J. "Beyond GDP? Welfare Across Countries and Time", *American Economic Review* 106, 2016: 2426-2457.

Jones, Charles I., and Williams, John C. "Too Much of a Good Thing? The Economics of Investment in R&D", *Journal of Economic Growth* 5, 2000: 65-85.

Katz, Lawrence F., and Murphy, Kevin M. "Changes in Relative Wages, 1963-1987: Supply and Demand Factors "*Quarterly Journal of Economics* 107, 1992: 35-78.

Klenow, Peter J., and Rodríguez-Clare, Andrés. "The Neoclassical Revival in Growth Economics: Has It Gone Too Far?" *NBER Macroeconomics Annual 1997*, 73-114.

Klenow, Peter J., and Rodríguez-Clare, Andrés. "Externalities and Growth", *Handbook of Economic Growth* 1A P. Aghion and S. Durlauf editors, Elsevier, New York, 2005.

Malmberg, Hannes. "Human Capital and Development Accounting Revisited ", *Working Paper* 2016.

Mankiw, N. Gregory, Romer, David, and Weil, David N. "A Contribution to the Empirics of Economic Growth", *Quraterly Journal of Economics* 107, 1992: 407-437.

Mitchell, B.R. *International Historical Statistics: Africa, Asia & Oceania: 1750-2000.* Fourth Edition, New York: Palgrave Macmillan, 2003a.

Mitchell, B.R. *International Historical Statistics: The Americas: 1750-2000.* Fourth Edition. New York: Palgrave Macmillan, 2003b.

Mitchell, B.R. *International Historical Statistics: Europe: 1750-2000.* Fourth Edition. New York: Palgrave Macmillan, 2003c.

Murphy, Kevin M., and Topel, Robert H. "Value of Health and Longevity", *Journal of Political Economy* 114, 2006: 871-904.

O'Mahony, Mary and Marcel P. Timmer, "Output, Input and Productivity Measures at the Industry Level: the EU KLEMS Database ", *Economic Journal*, 119(538) 2009: F374-F403.

Olson, Mancur Jr., "The Rise and Decline of Nations: Economic Growth, Stagflation, and Social Rigidities."New Haven: Yale University Press. 1982.

Olson, M., 1996. "Big Bills Left on the Sidewalk: Why Some Nations Are Rich, and Others Poor."in Kähkönen, S., Olson, M. (Eds.), 1996. *A Not-so-Dismal Science: A Broader View of Economies and Societies*. Oxford University Press, Oxford.

O'Rourke, Kevin H. "Tariffs and Growth and in the Late 19th Century ", *Economic Journal* 110, 2000: 456-483.

O'Rourke, Kevin H. and Jeffrey G. Williamson. *Globalization and History. The Evolution of a Nineteenth-Century Atlantic Economy*; MIT Press 1999.

Parente, Stephen L., and Prescott, Edward C. "Barriers to Technology Adoption and Development", *Journal of Political Economy* 102, 1994: 298-321.

Parente, Stephen L., and Prescott, Edward C. "Monopoly Rights: A Barrier to Riches", *American Economic Review* 89, 1999: 1216-1233.

Piketty, T., and Zucman, G. "Capital is Back: Wealth-Income Ratios in Rich Countries 1700-2010" *Quarterly Journal of Economic* 129, 2014: 1255-1310.

Quah, Danny T. “Twin Peaks: Growth and Convergence in Models of Distribution Dynamics”,*Economic Journal* 106, 1996: 1045-1055.

Ramey, Valerie, A. and Nekarda, Christopher J. “The Cyclical Behavior of the Price-Cost Markup”, University of California at San Diego working paper, 2013.

Rodrik, Dani. “Unconditional Convergence in Manufacturing ”. *Quarterly Journal of Economics* , 2012:

Rodrik, Dani. “Premature Deindustrialization ”. *Journal of Economic Growth* 21, 2015: 1-33.

Sabillon, C. *World Economic Historical Statistics* New York: Algora Publishing 2005.

Stewart, F. *Technology and Underdevelopment* London: The Macmillan Press, 1977.

Tamura, Robert. “From Decay to Growth: A Demographic Transition to Economic Growth”, *Journal of Economic Dynamics and Control* 20, 1996: 1237-1262.

Tamura, Robert, Dwyer, Gerald P., Devereux, John, and Baier, Scott L. “Economic Growth in the Long Run ”*Journal of Development Economics* 137, 2019: 1-35.

Tamura, Robert, Dwyer, Gerald P., Devereux, John, and Baier, Scott L. “Data Appendix for Economic Growth in the Long Run ”Clemson University working paper, 2018.

United Nations Development Programme. *Human Development Report 2010* New York, Palgrave Macmillan, 2010.

World Bank. *Doing Business: Distance to Frontier*. `http://www.doingbusiness.org/data/distance-to-frontier`, accessed Feb 2017.

World Bank. *World Development Report*. World Bank, New York, various years.

Appendices

A Detailed Model Derivations

In this section of the appendix, we provide more detailed derivations of some of the equations on the production side of the model.

Final Good

Final output is produced using intermediate goods which are skill-specific according to the following production function

$$Y = \{Y_H^{\frac{\varepsilon-1}{\varepsilon}} + Y_M^{\frac{\varepsilon-1}{\varepsilon}} + Y_L^{\frac{\varepsilon-1}{\varepsilon}}\}^{\frac{\varepsilon}{\varepsilon-1}} \tag{22}$$

Competitive firms (characterized below) produce the intermediate goods Y_H, Y_M and Y_L and sell them to competitive final output producers at prices P_i, $i = H, M, L$. We take the final good to be the *numereire* so that

$$\left[P_H^{1-\varepsilon} + P_M^{1-\varepsilon} + P_L^{1-\varepsilon}\right]^{\frac{1}{1-\varepsilon}} = 1. \tag{23}$$

Intermediate Goods & Machines

The intermediate goods production functions are given by

$$\begin{aligned} Y_H &= \frac{1}{1-\beta}\int_0^{A_H} \chi_{jH}^{1-\beta} dj\, H^\beta \\ Y_M &= \frac{1}{1-\beta}\int_0^{A_M} \chi_{jM}^{1-\beta} dj\, M^\beta \\ Y_L &= \frac{1}{1-\beta}\int_0^{A_L} \chi_{jL}^{1-\beta} dj\, L^\beta \end{aligned} \tag{24}$$

where χ_{ji} is quantity of machines of variety j rented by the i-type intermediate goods producer.

The representative i-type intermediate goods firm solves the following maximization problem:

$$\max_{\{\chi_{ji}, N_i\}} \{\frac{P_i}{1-\beta}\int_0^{A_i} \chi_{ji}^{1-\beta} dj i^\beta - \int_0^{A_i} p_{ji}\chi_{ji} dj - w_i N_i\}, \tag{25}$$

where p_{ji} is the price of variety j, i-type machine.

For a representative firm hiring workers of skill type i, the inverse derived demand for a typical machine j is given by

$$P_i \chi_{ji}^{-\beta} i^{\beta} = p_{ji} \tag{26}$$

Machine producing monopolist will set the price equal to a υ markup over her own marginal cost.

$$p_{ji} = \upsilon R \tag{27}$$

The equilibrium supply of machines of type j to skill i is::

$$\chi_{ji} = \left(\frac{P_i}{\upsilon R}\right)^{1/\beta} i \tag{28}$$

which means the (derived) production functions of intermediate good of type i become

$$Y_i = \frac{1}{1-\beta}\left(\frac{P_i}{\upsilon R}\right)^{\frac{1-\beta}{\beta}} A_i i, \tag{29}$$

and the profits per variety of machines are equated and given by

$$\begin{aligned} \pi_H &= \left(\frac{\upsilon-1}{\upsilon}\right) P_H^{1/\beta} H \left(\upsilon R\right)^{\frac{\beta-1}{\beta}}, \\ \pi_M &= \left(\frac{\upsilon-1}{\upsilon}\right) P_M^{1/\beta} M \left(\upsilon R\right)^{\frac{\beta-1}{\beta}}, \\ \pi_L &= \left(\frac{\upsilon-1}{\upsilon}\right) P_L^{1/\beta} L \left(\upsilon R\right)^{\frac{\beta-1}{\beta}}. \end{aligned} \tag{30}$$

Finally, it also follows that the relative prices of the two intermediate goods are given by:

$$\begin{aligned} \frac{P_H}{P_L} &= \left(\frac{A_H H}{A_L L}\right)^{-\frac{\beta}{\sigma}}, \\ \frac{P_M}{P_L} &= \left(\frac{A_M M}{A_L L}\right)^{-\frac{\beta}{\sigma}}, \end{aligned} \tag{31}$$

where $\sigma = 1 + (\varepsilon - 1)\beta$.

Wages

Producers of type-i Intermediate good hire labor according to the following first order condition:

$$\frac{\beta P_i}{1-\beta} \int_0^{A_i} \chi_{ji}^{1-\beta} dj i^{\beta-1} = w_i, \tag{32}$$

which, after substituting for the equilibrium quantities of machines and available workers of type L, produces

$$w_i = \frac{\beta}{1-\beta} A_i \beta P_i^{\frac{1}{\beta}} (\upsilon R)^{-\frac{1-\beta}{\beta}} \tag{33}$$

These conditions imply that relative prices of intermediate goods of two different types are given by

$$\begin{aligned} \frac{P_H}{P_L} &= \left(\frac{A_H H}{A_L L}\right)^{-\frac{\beta}{\sigma}}, \\ \frac{P_M}{P_L} &= \left(\frac{A_M M}{A_L L}\right)^{-\frac{\beta}{\sigma}}, \end{aligned} \tag{34}$$

and relative wages of workers of two different types are

$$\begin{aligned} \frac{w_H}{w_L} &= \left(\frac{A_H}{A_L}\right)^{\frac{\sigma-1}{\sigma}} \left(\frac{H}{L}\right)^{-\frac{1}{\sigma}}, \\ \frac{w_M}{w_L} &= \left(\frac{A_M}{A_L}\right)^{\frac{\sigma-1}{\sigma}} \left(\frac{M}{L}\right)^{-\frac{1}{\sigma}}, \end{aligned} \tag{35}$$

where $\sigma = 1 + (\varepsilon - 1)\beta$, is the elasticity of substitution between worker types.

Capital Allocation & Rental Rate

Capital is used to manufacture machines. Denoting by K_i the amount of physical capital devoted to production of i-type machines, we have

$$K_i = \int_0^{A_i} \chi_{ji} dj = A_i \left(\frac{P_i}{\upsilon R}\right)^{1/\beta} i \tag{36}$$

and it follows that

$$\frac{K_H}{K_L} = \left(\frac{A_H}{A_L}\right)^{\frac{\sigma-1}{\sigma}} \left(\frac{H}{L}\right)^{\frac{\sigma-1}{\sigma}}, \tag{37}$$
$$\frac{K_M}{K_L} = \left(\frac{A_M}{A_L}\right)^{\frac{\sigma-1}{\sigma}} \left(\frac{M}{L}\right)^{\frac{\sigma-1}{\sigma}}$$

In addition, total capital stock by K is given by

$$K = K_H + K_M + K_L. \tag{38}$$

Since machines take one unit of capital to produce and all machines within a skill industry are symmetric it must be the case that:

$$K_i = A_i \chi_i \tag{39}$$
$$\tag{40}$$

Using the above conditions, we can show that the reduced-form production function of intermediate goods production is given by

$$Y_H = \frac{K_H^{1-\beta}(A_H H)^\beta}{1-\beta}, \tag{41}$$
$$Y_M = \frac{K_M^{1-\beta}(A_M M)^\beta}{1-\beta},$$
$$Y_L = \frac{K_L^{1-\beta}(A_L L)^\beta}{1-\beta},$$

Differentiating the above with respect to capital, and multiplying by the sector price, P_i, we obtain expressions for the value marginal product of capital in the i sector:

$$P_i MPK_i = P_i K_i^{-\beta}(A_i i)^\beta = P_i(1-\beta)\frac{Y_i}{K_i} \tag{42}$$

It is easy to show, using the expressions for P_H/P_L, P_H/P_M,K_H/K_L, and K_H/K_M derived above, that this implies the value marginal product of capital is equal across sectors.

Further, note that when intermediate producers buy machines, they pay υR per unit of capital where υ is the markup over the cost of producing machines (the rental rate). This implies that

$$\upsilon R = P_i MPK_i = P_i K_i^{-\beta}(A_i i)^{\beta} = P_i(1-\beta)\frac{Y_i}{K_i}$$

and we have

$$R = \frac{P_i MPK_i}{\upsilon} = P_i\left(\frac{1-\beta}{\upsilon}\right)\frac{Y_i}{K_i} = \left(\frac{1-\beta}{\upsilon}\right)\frac{Y}{K} \tag{43}$$

This last result, together with the fact that $r = (1-\tau)R - \delta$, where δ is the rate of depreciation of capital and τ is the tax on capital income

$$r = (1-\tau)\left(\frac{1-\beta}{\upsilon}\right)\frac{Y}{K} - \delta. \tag{44}$$

Innovation

The following process governs the discovery of new blueprints for sector i

$$\begin{aligned}
\dot{A}_H &= \eta_H\left(\frac{A_H^W}{A_H}\right)^{\varphi}\frac{Z_H}{N}, \qquad (45)\\
\dot{A}_M &= \eta_M\left(\frac{A_M^W}{A_M}\right)^{\varphi}\frac{Z_M}{N},\\
\dot{A}_L &= \eta_L\left(\frac{A_L^W}{A_L}\right)^{\varphi}\frac{Z_L}{N},
\end{aligned}$$

where represents A_i^W is the world frontier technology for sector $i = H, M, L$, η_i is the productivity of research effort, and Z_i is the R&D expenditure on innovation or technology adoption in sector $i = H, M, L$.

In order to innovate, the entrepreneurs must incur an entry cost X. Free entry into research implies that marginal benefit of extra innovation/adoption effort Z is equal to the cost,

$$\begin{aligned}
\eta_H\left(\frac{A_H^W}{A_H}\right)^{\varphi}\frac{V_H}{N} &= X, \qquad (46)\\
\eta_M\left(\frac{A_M^W}{A_M}\right)^{\varphi}\frac{V_M}{N} &= X,\\
\eta_L\left(\frac{A_L^W}{A_L}\right)^{\varphi}\frac{V_L}{N} &= X,
\end{aligned}$$

which in turn pins down the relative value of blueprints as

$$\begin{aligned}\frac{V_H}{V_L} &= \left(\frac{\eta_H}{\eta_L}\right)^{-1}\left(\frac{A_H/A_H^W}{A_L/A_L^W}\right)^{\varphi}, \\ \frac{V_M}{V_L} &= \left(\frac{\eta_M}{\eta_L}\right)^{-1}\left(\frac{A_M/A_M^W}{A_L/A_L^W}\right)^{\varphi},\end{aligned} \tag{47}$$

where V_i, $i = H, M, L$ is the value of a blueprint for a machine in sector i, which must satisfy the the no-arbitrage condition

$$r\,V_L = \pi_L + \dot{V}_L \tag{48}$$

BGP Growth Rate & Interest Rate

Along the balanced growth path the economy grows at a constant growth rate g, equal to the growth rate of the technology frontier (assumed to be the same for all types of skills).

$$g = \frac{1}{\theta}\left[r^* - \rho\right],$$

where ρ is the discount rate and θ is the CRRA coefficient. The BGP interest rate r^* therefore given by

$$r^* = \theta g + \rho, \tag{49}$$

and, using equation equations (9) and (10), the BGP rental rate is

$$R^* = \frac{\theta g + \rho + \delta}{1 - \tau}. \tag{50}$$

Using the no-arbitrage conditions from (48) and the fact that along the BGP the value of a patent must be stationary ($\dot{V}_i = 0$ for $i = H, M, L.$) we get the following relationship between the value of a patent, profits and the interest rate

$$\begin{aligned}V_H &= \frac{\pi_H}{r^*}, \\ V_M &= \frac{\pi_M}{r^*}, \\ V_L &= \frac{\pi_L}{r^*},\end{aligned} \tag{51}$$

where profits are given by $\pi_i = \left(\frac{\upsilon-1}{\upsilon}\right) P_i^{1/\beta} i \left(\upsilon R^*\right)^{\frac{\beta-1}{\beta}}$. This leads to

$$\begin{aligned}\frac{V_H}{V_L} &= \left(\frac{A_H}{A_L}\right)^{-\frac{1}{\sigma}}\left(\frac{H}{L}\right)^{\frac{\sigma-1}{\sigma}}, \\ \frac{V_M}{V_L} &= \left(\frac{A_M}{A_L}\right)^{-\frac{1}{\sigma}}\left(\frac{M}{L}\right)^{\frac{\sigma-1}{\sigma}},\end{aligned} \tag{52}$$

Finally, combining equations (47) and (52), yields[19]

$$\begin{aligned}\frac{A_H}{A_L} &= \left(\frac{\eta_H}{\eta_L}\right)^{\frac{\sigma}{1+\sigma\varphi}}\left(\frac{H}{L}\right)^{\frac{\sigma-1}{1+\sigma\varphi}}\left(\frac{A_H^W}{A_L^W}\right)^{\frac{\varphi\sigma}{1+\sigma\varphi}}, \\ \frac{A_M}{A_L} &= \left(\frac{\eta_M}{\eta_L}\right)^{\frac{\sigma}{1+\sigma\varphi}}\left(\frac{M}{L}\right)^{\frac{\sigma-1}{1+\sigma\varphi}}\left(\frac{A_M^W}{A_L^W}\right)^{\frac{\varphi\sigma}{1+\sigma\varphi}}\end{aligned} \tag{53}$$

Substituting the expression for relative productivity levels (53) into the relative wage formula (35) we obtain

$$\begin{aligned}\frac{w_H}{w_L} &= \left(\frac{\eta_H}{\eta_L}\right)^{\frac{\sigma-1}{1+\sigma\varphi}}\left(\frac{H}{L}\right)^{\frac{\sigma-2-\varphi}{1+\sigma\varphi}}\left(\frac{A_H^W}{A_L^W}\right)^{\frac{\varphi(\sigma-1)}{1+\sigma\varphi}}, \\ \frac{w_M}{w_L} &= \left(\frac{\eta_M}{\eta_L}\right)^{\frac{\sigma-1}{1+\sigma\varphi}}\left(\frac{M}{L}\right)^{\frac{\sigma-2-\varphi}{1+\sigma\varphi}}\left(\frac{A_M^W}{A_L^W}\right)^{\frac{\varphi(\sigma-1)}{1+\sigma\varphi}}\end{aligned} \tag{54}$$

Using equations (46) and (51) we can show that on the BGP productivity relative to the frontier is given by

$$\begin{aligned}\mu_H \equiv \frac{A_H}{A_H^W} &= \left[\frac{\eta_H\left(\frac{\upsilon-1}{\upsilon}\right)(H/N)P_H^{1/\beta}(\upsilon R)^{\frac{\beta-1}{\beta}}}{r^*X}\right]^{1/\varphi}, \\ \mu_M \equiv \frac{A_M}{A_M^W} &= \left[\frac{\eta_M\left(\frac{\upsilon-1}{\upsilon}\right)(M/N)P_M^{1/\beta}(\upsilon R)^{\frac{\beta-1}{\beta}}}{r^*X}\right]^{1/\varphi}, \\ \mu_L \equiv \frac{A_L}{A_L^W} &= \left[\frac{\eta_L\left(\frac{\upsilon-1}{\upsilon}\right)(L/N)P_L^{1/\beta}(\upsilon R)^{\frac{\beta-1}{\beta}}}{r^*X}\right]^{1/\varphi},\end{aligned} \tag{55}$$

[19] Notice that in our baseline specification (with $\varphi = 0$) this collapses to the expression familiar from Acemoglu

$$\frac{A_H}{A_L} = \left(\frac{\eta_H}{\eta_L}\right)^{\sigma}\left(\frac{H}{L}\right)^{\sigma-1}.$$

Finally, substituting the above expressions into the aggregate final goods production function, we get that (see Appendix E for derivation)

$$\frac{Y}{N} = \left(A_L^W\right)^{\frac{\varphi}{1+\varphi}} \left(\frac{1-\beta}{\upsilon}\right)^{\frac{\beta-1}{(1+\varphi)\beta}} \left(\frac{K}{Y}\right)^{\frac{1-\beta}{\beta}} \left(\frac{X}{Y/N}\right)^{-\frac{1}{1+\varphi}} \Omega\left(\frac{H}{L}, \frac{M}{L}, \frac{A_H^W}{A_L^W}, \frac{A_M^W}{A_L^W}\right)^{\frac{\varphi}{(1+\varphi)\beta}} \tag{56}$$

Thus on the BGP output per worker depends on the world technology frontier (A_H^W, A_M^W, and A_L^W) , capital accumulation (K/Y), domestic relative supply of skills (H, M, and L), and the technology wedge relative to GDP per worker (X/y).[20]

B Intermediate Good Prices

This section derives expressions for equilibrium intermediate good prices. Using equations (23) and (4) we can show that along the BGP price of intermediate goods will be given by:

$$\begin{aligned} P_H &= \left\{1 + \left[\left(\frac{H}{M}\right)^{1+\varphi} \frac{\eta_H}{\eta_M}\left(\frac{\mathrm{A}_H^W}{\mathrm{A}_M^W}\right)^{\varphi}\right]^{\frac{\beta(1-\varepsilon)}{1+\sigma\varphi}} + \left[\left(\frac{H}{L}\right)^{1+\varphi} \frac{\eta_H}{\eta_L}\left(\frac{\mathrm{A}_H^W}{\mathrm{A}_L^W}\right)^{\beta}\right]^{\frac{\beta(1-\varepsilon)}{1+\sigma\varphi}}\right\}^{\frac{1}{\varepsilon-1}} \\ P_M &= \left\{\left[\left(\frac{M}{H}\right)^{1+\varphi} \frac{\eta_M}{\eta_H}\left(\frac{\mathrm{A}_M^W}{\mathrm{A}_H^W}\right)^{\varphi}\right]^{\frac{\beta(1-\varepsilon)}{1+\sigma\varphi}} + 1 + \left[\left(\frac{M}{L}\right)^{1+\varphi} \frac{\eta_M}{\eta_L}\left(\frac{\mathrm{A}_M^W}{\mathrm{A}_L^W}\right)^{\varphi}\right]^{\frac{\beta(1-\varepsilon)}{1+\sigma\varphi}}\right\}^{\frac{1}{\varepsilon-1}} \\ P_L &= \left\{\left[\left(\frac{L}{H}\right)^{1+\varphi} \frac{\eta_L}{\eta_H}\left(\frac{\mathrm{A}_L^W}{\mathrm{A}_H^W}\right)^{\varphi}\right]^{\frac{\beta(1-\varepsilon)}{1+\sigma\varphi}} + \left[\left(\frac{L}{M}\right)^{1+\varphi} \frac{\eta_L}{\eta_M}\left(\frac{\mathrm{A}_L^W}{\mathrm{A}_M^W}\right)^{\varphi}\right]^{\frac{\beta(1-\varepsilon)}{1+\sigma\varphi}} + 1\right\}^{\frac{1}{\varepsilon-1}} \end{aligned} \tag{57}$$

C Estimating σ using EU KLEMS

Jerzmanowski and Tamura (2017) use data on skill composition and compensation across 18 OECD countries during the period 1970-2000 to estimate the elasticity of substitution between skill types. This section briefly describes the approach and highlights the main results related to the present paper.

[20] Output is clearly decreasing in the level of barrier. Whether it is increasing in H/L and M/L depends on the world technology frontier A_H^W/A_L^W and A_H^W/A_M^W and what happens to the gaps to this frontier μ_H , μ_M, and μ_L as the composition of the labor force changes.

Previous work can be best described by starting with the equation for skill premium

$$\frac{w_H}{w_L} = \left(\frac{A_H}{A_L}\right)^{\frac{\sigma-1}{\sigma}} \left(\frac{H}{L}\right)^{-\frac{1}{\sigma}} \tag{58}$$

Taking logs and assuming that A_H/A_L (skill-bias of technology) is growing at a smooth exponential rate γ_1

$$\log\left(\frac{A_H}{A_L}\right) = \gamma_0 + \gamma_1 t$$

produces the following expression

$$\log\left(\frac{w_H}{w_L}\right) = \alpha + \frac{\sigma-1}{\sigma}\gamma_1 t - \frac{1}{\sigma}\log\left(\frac{H}{L}\right)$$

where t is time. Katz and Murphy (1992) estimated the above using data on college/secondary wage premium for the years 1963-87 and found an estimate of $\sigma = 1.4$. However, they observe that including a square and higher order polynomials of t (i.e. allowing for A_H/A_L to grow at variable rate) affects the estimate and they conclude that values as high as 2.6 are consistent with the data. More recently Ciccone and Peri (2005) use an instrumental variables strategy (since H/L responds to wage shocks, OLS may be inconsistent) and data across US states. They find σ close to 1.5. Most recently however, Autor and Acemoglu (2011) argue that higher values of σ are also plausible. For example, using Katz and Murphy's regression on updated sample they find $\sigma = 2.9$.

We estimate σ using international data since our model, which accounts for technology diffusion across countries, implies the value of the diffusion parameter, φ, affects the relationship between the coefficient estimates from equations like the one above and the elasticity of substitution. Specifically, we use the EU KLEMS Growth and Productivity Accounts panel data set (O'Mahony and Timmer, 2009). This is a detailed database of industry-level measures of output, inputs, and productivity for 25 European countries, Japan and the US for the period from 1970 to 2005. We use the information on hours worked and wages in manufacturing broken down into three skill groups: low-skill (less than secondary school degree), medium-skill (secondary school degree) and high-skills (at least some college). To derive our specification recall that according to our model – once the endogenous direction of technological change is taken into account – the relative wages are given by

$$\frac{w_H}{w_L} = \left(\frac{\eta_H}{\eta_L}\right)^{\frac{\sigma-1}{1+\sigma\varphi}} \left(\frac{H}{L}\right)^{\frac{\sigma-2-\varphi}{1+\sigma\varphi}} \left(\frac{A_H^W}{A_L^W}\right)^{\frac{\varphi(\sigma-1)}{1+\sigma\varphi}}. \tag{59}$$

Taking logs and assuming that A_H^W/A_L^W (skill-bias of world technology frontier) is growing at a smooth exponential rate γ_1

$$\log\left(\frac{A_H^W}{A_L^W}\right) = \gamma_0 + \gamma_1\, t$$

we get the following expression

$$\log\left(\frac{w_H}{w_L}\right) = \alpha + \frac{\varphi(\sigma-1)}{1+\sigma\varphi}\gamma_1\, t + \frac{\sigma-2-\varphi}{1+\sigma\varphi}\log\left(\frac{H}{L}\right)$$

and imposing our preferred value of $\varphi = 0.5$ we arrive at

$$\log\left(\frac{w_H}{w_L}\right) = \alpha + \frac{\sigma/2-1}{1+\sigma/2}\gamma_1\, t + \frac{\sigma-2.5}{1+\sigma/2}\log\left(\frac{H}{L}\right) \tag{60}$$

Notice that this is the same regression that most papers in the literature on the elasticity of substitution between skills estimate. However, the structural interpretation of the estimated coefficient on the relative skill supplies (H/L)is different. In the presence of directed technological change and technology diffusion, this coefficient is not the inverse of the elasticity of substitution σ (and the additionally depends on the diffusion parameter φ.) We use the estimate of this coefficient ($\frac{\sigma-2.5}{1+\sigma/2}$) to back out the estimate of σ.

Tables 6 and 7 below show the results of estimating the above equation using our data using college-educated workers as H and the sum of the remaining two skill groups as L. We use OLS, fixed effects, GMM, and system GMM (where we instrument H/L with lagged values). The standard errors on the implied σ's are calculated using the delta method. (Using year effects in place of trend does not change the results). The bottom line is that many of our point estimates are well above 2 and the 95% confidence intervals contain values close or even above the 2.5 required for strong skill bias in our quantitative exercise.

We have computed our results under a lower elasticity of 1.6 and found that some of the magnitudes – especially the measures of the technology adoption wedge – seem highly implausible. Specifically, out 146 countries, the U.S. ranks only the 26th in terms of the wedge Moreover, while some of the countries found to have lower adoption wedges than the U.S., such as Hong Kong and Singapore, are plausible, others are decidedly less so (for example, Sudan and Mozambique). The magnitudes on this side of the distribution also seem questionable; for example, Singapore's wedge at only about 32% of those in the U.S. seem unrealistically low. The distribution of wedges for the case of $\sigma = 2.6$ seems more plausible. Not only is the U.S. ranked 11th, but the countries we find to have lower wedges include only developed economies such as Singapore, Hong Kong, Taiwan, Ireland, and Norway. Equally important, the magnitudes seem more realistic with the lowest level (Singapore) equal to 55% of the US value and most others in this low-wedge group between 80% and 96%. These results are available upon request.

OLS					
$\ln(H/(M+L))$	-0.101	-0.100	-0.274^{***}	-0.277^{***}	-0.262^{***}
	(0.064)	(0.063)	(0.084)	(0.085)	(0.091)
Trend Squared	*No*	*Yes*	*No*	*Yes*	*Yes*
Country Trend	*No*	*No*	*Yes*	*Yes*	*Yes*
Country Trend Sq	*No*	*No*	*No*	*No*	*Yes*
σ	2.28	2.29	1.96	1.95	1.98
95% CI	[1.42, 2.54]	[1.41, 2.55]	[1.32, 2.25]	[1.28, 2.25]	[0.95, 2.30]
R^2	0.07	0.07	0.73	0.75	0.89
N	645	645	645	645	645
Fixed Effects					
$\ln(H/(M+L))$	-0.271^{***}	-0.273^{***}	-0.265^{***}	-0.263^{***}	-0.158
	(0.081)	(0.083)	(0.081)	(0.089)	(0.108)
Trend Squared	*No*	*Yes*	*No*	*Yes*	*Yes*
Country Trend	*No*	*No*	*Yes*	*Yes*	*Yes*
Country Trend Sq	*No*	*No*	*No*	*No*	*Yes*
σ	1.96	1.96	1.97	1.98	2.17
95% CI	[1.42, 2.25]	[1.41, 2.25]	[1.32, 2.26]	[1.28, 2.29]	[0.95, 2.59]
R^2	0.18	0.23	0.66	0.67	0.78
N	645	645	645	645	645

Table 6: Estimates of the elasticity of substitution between skill types based using equation (62), using college-educated workers as H and the sum of the other two skill categories as L. Standard errors for σ competed using the delta method.

GMM					
$\ln(H/(M+L))$	-0.255^{***}	-0.270^{***}	-0.266^{***}	-0.266^{***}	-0.173
	(0.077)	(0.075)	(0.084)	(0.091)	(0.107)
Trend Squared	*No*	*Yes*	*No*	*Yes*	*Yes*
Country Trend	*No*	*No*	*Yes*	*Yes*	*Yes*
Country Trend Sq	*No*	*No*	*No*	*No*	*Yes*
σ	1.99	1.96	1.97	1.97	2.14
95% CI	[1.42, 2.26]	[1.41, 2.23]	[1.32, 2.27]	[1.28, 2.29]	[0.95, 2.55]
N	609	609	609	609	609
System GMM					
$\ln(H/(M+L))$	-0.112	-0.110	-0.285^{***}	-0.284^{***}	-0.096
	(0.075)	(0.074)	(0.090)	(0.089)	(0.182)
Trend Squared	*No*	*Yes*	*No*	*Yes*	*Yes*
Country Trend	*No*	*No*	*Yes*	*Yes*	*Yes*
Country Trend Sq	*No*	*No*	*No*	*No*	*Yes*
σ	2.26	2.26	1.94	1.94	2.29
95% CI	[1.42, 2.57]	[1.41, 2.57]	[1.32, 2.25]	[1.28, 2.25]	[0.95, 3.04]
N	645	645	645	645	645

Table 7: Estimates of the elasticity of substitution between skill types based using equation (62), using college-educated workers as H and the sum of the other two skill categories as L. Standard errors for σ competed using the delta method.

D Relative Research Productivity

In our quantitative computations we back out the value of relative research productivity ($\frac{\eta_H}{\eta_L}$) by imposing a value for σ and solving equation (18) for each decade in our sample. Here we discuss the resulting values of these relative productivities. We also note that our quantitative approach is different from the one taken by Gancia et al. (2013). They arrive at an equation equivalent our equation (18) – with the exception that they treat the U.S. as the frontier country, which by definition does not benefit from diffusion so $\varphi = 0$). They use observations on the skill premium and relative supply of skills in 1970 and 2000 to solve for the intercept ($\frac{\eta_H}{\eta_L}$) and the slope of this line defined by taking logs of the above equation. This gives them a value of about 2.3 for σ and a common value for the intercept. We instead back out ($\frac{\eta_H}{\eta_L}$) by imposing a value for σ and solving the above equation for each decade in our sample. The key difference is of course that in our approach, the relative productivity of different research directions is allowed to change over time. The values of these relative productivities are plotted in Figure 9. For both high skill (college) and medium skill (secondary-educated) workers, the value of η has declined steadily throughout our sample. This means that a hypothetical country at the frontier with equal amount of innovation expenditure on each sector, would experience smaller growth rates in productivity of the higher-skilled sectors than in the lowest skilled sector (and since η_H/η_M also declines, the same applies to relative growth of high (college-specific) productivity relative to that of medium (secondary-specific) productivity.) This might at first seem counterintuitive but is, in fact, consistent with findings such as Jones (1995) and Bloom et al. (2018), which suggests that large historical increases R&D efforts over time have not produced an acceleration of growth rates. We explored an alternative calibration where we forced the η's to be constant over time. This approach produces a path of skill premia in the US that is vastly at odds with their historical patterns.[21]

[21] Gancia, et al. (2013) differs from our approach in other vital ways. First, Gancia et al. only examines the last several decades and use data from 1970 to 2000. Our data goes back as far as 1910 for some countries, which allows us to investigate the evolution of skill-bias technology and the level of the technology wedge over a longer time period. Second, our data distinguishes between three levels of skills: primary, secondary, and college, whereas Gancia et al. work with binary skilled/unskilled distinction. Third, Gancia et al. assume the US is the technology leader and thus defines the world productivity frontier. We do not impose this restriction, and indeed we find that in recent years the U.S. does not have the highest productivity of workers with primary and secondary education in the world. Finally, In their interpretation, adoption wedges are measured by the diffusion parameter φ, and removing them is equivalent to setting $\varphi \to \infty$, which forces convergence to the frontier), something that does not happen in our model.

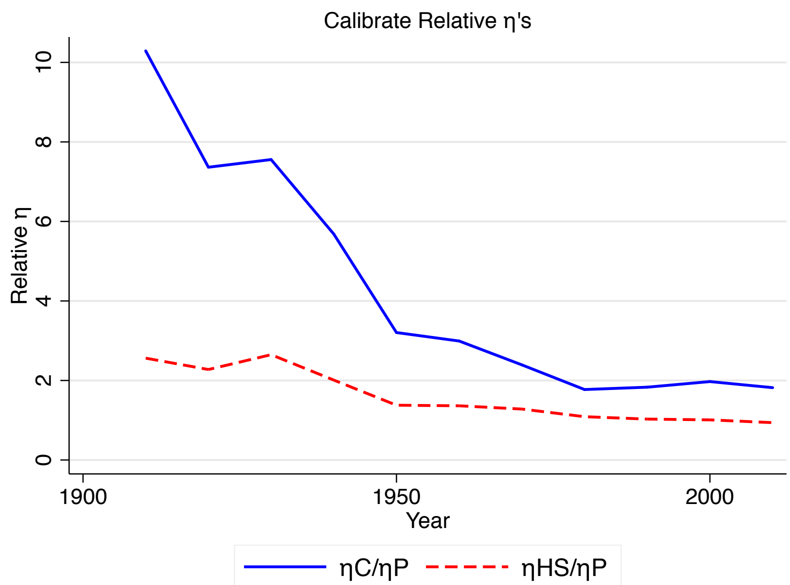


Figure 9: Calibrated relative η's.

E BGP Output per Worker: Two Useful Expressions

This section derives the expressions for the balanced growth path GDP per worker used in the main text. First, note that starting with the production function

$$Y = \frac{1}{1-\beta}\left\{ \left[K_H^{1-\beta}(A_H N s_H)^\beta\right]^{\frac{\varepsilon-1}{\varepsilon}} + \left[K_M^{1-\beta}(A_M N s_M)^\beta\right]^{\frac{\varepsilon-1}{\varepsilon}} + \left[K_L^{1-\beta}(A_L N s_L)^\beta\right]^{\frac{\varepsilon-1}{\varepsilon}} \right\}^{\frac{\varepsilon}{\varepsilon-1}},$$

and factoring out A_L, the GDP of an economy can be written as

$$Y = \frac{(A_L N)^\beta K_L^{1-\beta}\left\{ \left[\left(\frac{K_H}{K_L}\right)^{1-\beta}\left(\frac{A_H s_H}{A_L}\right)^\beta\right]^{\frac{\varepsilon-1}{\varepsilon}} + \left[\left(\frac{K_M}{K_L}\right)^{1-\beta}\left(\frac{A_M s_M}{A_L}\right)^\beta\right]^{\frac{\varepsilon-1}{\varepsilon}} + \left[s_L^\beta\right]^{\frac{\varepsilon-1}{\varepsilon}} \right\}^{\frac{\varepsilon}{\varepsilon-1}}}{1-\beta}$$

where

$$\begin{aligned} s_H &= \frac{H}{H+M+L} = \frac{H}{N}, \\ s_M &= \frac{M}{H+M+L} = \frac{M}{N}, \\ s_L &= \frac{L}{H+M+L} = \frac{L}{N} \end{aligned} \tag{61}$$

Furthermore, using the fact that $K = K_L(\frac{K_H}{K_L} + \frac{K_M}{K_L} + 1)$, we get

$$Y = \frac{(A_L N)^\beta K^{1-\beta}\left\{ \left[\left(\frac{K_H}{K_L}\right)^{1-\beta}\left(\frac{A_H s_H}{A_L}\right)^\beta\right]^{\frac{\varepsilon-1}{\varepsilon}} + \left[\left(\frac{K_M}{K_L}\right)^{1-\beta}\left(\frac{A_M s_M}{A_L}\right)^\beta\right]^{\frac{\varepsilon-1}{\varepsilon}} + \left[s_L^\beta\right]^{\frac{\varepsilon-1}{\varepsilon}} \right\}^{\frac{\varepsilon}{(\varepsilon-1)}}}{(1-\beta)(\frac{K_H}{K_L} + \frac{K_M}{K_L} + 1)^{1-\beta}} \tag{62}$$

With the aid of equations (7), (17) we can see that, except A_L, all of the terms on the right had side are either data (N, K, and s_i's) or a function of data, parameters, and $(\frac{A_H^W}{A_L^W}, \frac{A_M^W}{A_L^W})$. Thus given our data, parameter values, and values for $(\frac{A_H^W}{A_L^W}, \frac{A_M^W}{A_L^W})$, we can solve the above equation for A_L, which is one of the steps in our empirical algorithm.

Another useful expression for GDP per worker can be obtained in a similar way by factoring out the world frontier productivity level (A_L^W) instead of A_L

$$Y = \frac{(A_L^W N)^\beta K^{1-\beta} \left\{ \left[\left(\frac{K_H}{K_L}\right)^{1-\beta} \left(\frac{A_H^W \mu_H s_H}{A_L^W}\right)^\beta \right]^{\frac{\varepsilon-1}{\varepsilon}} + \left[\left(\frac{K_M}{K_L}\right)^{1-\beta} \left(\frac{A_M^W \mu_M s_M}{A_L^W}\right)^\beta \right]^{\frac{\varepsilon-1}{\varepsilon}} + \left[(\mu_L s_L)^\beta \right]^{\frac{\varepsilon-1}{\varepsilon}} \right\}^{\frac{\varepsilon}{(\varepsilon-1)}}}{(1-\beta)(\frac{K_H}{K_L} + \frac{K_M}{K_L} + 1)^{1-\beta}} \tag{63}$$

Again, equations (7), (17), (55), and (57) tell us that the productivity and capital ratios depend on s_i $i = H, M, L$ and $(\frac{A_H^W}{A_L^W}, \frac{A_M^W}{A_L^W})$. The fact that (from 55)

$$\mu_i \equiv \frac{A_i}{A_i^W} \quad = \quad \left[\frac{\eta_i \left(\frac{\upsilon-1}{\upsilon}\right) (i/N) P_i^{1/\beta} (\upsilon R)^{\frac{\beta-1}{\beta}}}{rX} \right]^{1/\varphi}$$

and the expressions for P_i in 57 (also recalling that R depends on investment distortion τ), we can write

$$\mu_i = X^{-\frac{1}{\varphi}} R(\tau)^{\frac{\beta-1}{\varphi\beta}} \times (\text{terms that depend on } \frac{H}{L}, \frac{M}{L}, \frac{A_H^W}{A_L^W}, \frac{A_M^W}{A_L^W} \text{ and } \eta\text{'s}),$$

which in turn, allows us to simplify the formula for GDP per worker to

$$\frac{Y}{N} = (A_L^W)^\beta \left(\frac{K}{N}\right)^{1-\beta} X^{-\frac{\beta}{\varphi}} R(\tau)^{\frac{\beta-1}{\varphi}} \Omega \left(\frac{H}{L}, \frac{M}{L}, \frac{A_H^W}{A_L^W}, \frac{A_M^W}{A_L^W}, \text{ and } \eta\text{'s}\right) \tag{64}$$

or

$$\frac{Y}{N} = A_L^W \left(\frac{K}{Y}\right)^{\frac{1-\beta}{\beta}} X^{-\frac{1}{\varphi}} R(\tau)^{\frac{\beta-1}{\varphi\beta}} \Omega \left(\frac{H}{L}, \frac{M}{L}, \frac{A_H^W}{A_L^W}, \frac{A_M^W}{A_L^W}, \text{ and } \eta\text{'s}\right)^{\frac{1}{\beta}}$$

Note: in what follows, we omit the dependence of Ω on η's. It is not essential for the exposition at hand but should be kept in mind because otherwise, equation (67) below suggest that, once we know the value of the frontier A_L^W, we could calculate the *absolute* level of X for each country and not just the *relative* value of X we have analyzed in the paper. This is not possible since it would require the computation of the Ω and thus a knowledge of absolute values of η's (we only know the relative values from fitting Goldin and Katz's skill premia).

Using the equation for the rental rate R

$$\frac{Y}{N} = A_L^W \left(\frac{K}{Y}\right)^{\frac{1-\beta}{\beta}} X^{-\frac{1}{\varphi}} \left(\frac{1-\beta}{\mu} \frac{Y}{K}\right)^{\frac{\beta-1}{\varphi\beta}} \Omega \left(\frac{H}{L}, \frac{M}{L}, \frac{A_H^W}{A_L^W}, \frac{A_M^W}{A_L^W}\right)^{\frac{1}{\beta}}$$

$$\frac{Y}{N} = A_L^W \left(\frac{1-\beta}{\upsilon}\right)^{\frac{\beta-1}{\varphi\beta}} \left(\frac{K}{Y}\right)^{\frac{\varphi-\varphi\beta+1-\beta}{\varphi\beta}} X^{-\frac{1}{\varphi}} \Omega\left(\frac{H}{L}, \frac{M}{L}, \frac{A_H^W}{A_L^W}, \frac{A_M^W}{A_L^W}\right)^{\frac{1}{\beta}}$$

$$\frac{Y}{N} = A_L^W \left(\frac{1-\beta}{\upsilon}\right)^{\frac{\beta-1}{\varphi\beta}} \left(\frac{K}{Y}\right)^{\frac{\varphi-\varphi\beta+1-\beta}{\varphi\beta}} \left(\frac{X}{Y/N}\right)^{-\frac{1}{\varphi}} \left(\frac{Y}{N}\right)^{-\frac{1}{\varphi}} \Omega\left(\frac{H}{L}, \frac{M}{L}, \frac{A_H^W}{A_L^W}, \frac{A_M^W}{A_L^W}\right)^{\frac{1}{\beta}}$$

$$\frac{Y}{N} = \left(A_L^W\right)^{\frac{\varphi}{1+\varphi}} \left(\frac{1-\beta}{\upsilon}\right)^{\frac{\beta-1}{(1+\varphi)\beta}} \left(\frac{K}{Y}\right)^{\frac{1-\beta}{\beta}} \left(\frac{X}{Y/N}\right)^{-\frac{1}{1+\varphi}} \Omega\left(\frac{H}{L}, \frac{M}{L}, \frac{A_H^W}{A_L^W}, \frac{A_M^W}{A_L^W}\right)^{\frac{\varphi}{(1+\varphi)\beta}}$$

$$y = \text{constant} \times \left(\frac{K}{Y}\right)^{\frac{1-\beta}{\beta}} \left(\frac{X}{y}\right)^{-\frac{1}{1+\varphi}} \Omega\left(\frac{H}{L}, \frac{M}{L}, \frac{A_H^W}{A_L^W}, \frac{A_M^W}{A_L^W}\right)^{\frac{\varphi}{(1+\varphi)\beta}}, \tag{65}$$

where the last three terms are country-specific and correspond to the contribution of physical capital, technology wedge/income ratio, and human capital endowment. The constant term is the same for all countries in a given time period as it depends only on parameters of the models and the world technology frontier (A_L^W).

F Transitional Dynamics and the Calibration of φ

We calibrate the value of φ to match the dynamic behavior of the model. Specifically, we choose a value of this parameter, which governs the strength of technology diffusion, to match the rate of convergence to the BGP. This section briefly describes the dynamics of our model. For this calibration, we with a version of the model with two skill types denoted H and L. We feel this is sufficient since in the long run, most countries are likely to look like the present-day developed economies with lowest skill sector share practically equal to zero.

Even with the assumption of constant supplies of skilled and unskilled labor (H and L) the dynamics of the model can be complicated. Because innovation for the two skill types and capital accumulation technologies are linear, the transitional dynamics may involve initial periods when only some of these activities take place. Eventually, the rates of return to all three activities are equalized, and the economy converges to the BGP characterized in the paper. Characterizing the entire transitional dynamics of the model is beyond the scope of our analysis. Here we briefly discuss the dynamics of the system once all investment activities yield the same rate of return (and thus all are undertaken). We show how to linearize the model around the BGP and discuss the implied speed of convergence which we use to choose the value for the diffusion parameter φ.

To characterize the dynamics of the model, we start by re-writing the free entry condition (where the equations are symmetric for the two skill types, we conserve space by presenting only one version)

$$V_H = \eta_H^{-1} X H \mu_H^{\varphi} \tag{66}$$

we can differentiate the free entry condition to yield

$$\frac{\dot{V}_H}{V_H} = \varphi \frac{\dot{\mu}_H}{\mu_H} \tag{67}$$

Also since $\mu_H = A_H / A_H^W$, and the frontier is assumed to grow at the rate g, it follows from the expressions for the growth rate of productivity that

$$\frac{\dot{A}_H}{A_H} = \eta_H \mu_H^{-(1+\varphi)} \frac{\tilde{Z}_H}{H} \tag{68}$$

and

$$\frac{\dot{A}_L}{A_L} = \eta_L \mu_L^{-\varphi} \frac{A_H^W \tilde{Z}_L}{H}, \tag{69}$$

where $\tilde{X} \equiv X / A_H^W$, which yields the dynamic equations for the gaps to the frontier

$$\frac{\dot{\mu}_H}{\mu_H} = \eta_H \mu_H^{-(1+\varphi)} \frac{\tilde{Z}_H}{H} - g \tag{70}$$

$$\frac{\dot{\mu}_L}{\mu_L} = \eta_L \mu_L^{-(1+\varphi)} \frac{A_H^W / A_L^W \tilde{Z}_L}{H} - g \tag{71}$$

Additionally, recall that the no-arbitrage conditions are

$$\frac{\dot{V}_H}{V_H} = r - \frac{\pi_H}{V_H}$$

$$\frac{\dot{V}_L}{V_L} = r - \frac{\pi_L}{V_L}$$

Combining these conditions with the expression for profit rates derived earlier and equations (69), (72), and (73) we get

$$\tilde{Z}_H = \eta_H^{-1}\mu_H^{1+\varphi}H\left(g + \left(r - \eta_H\left(\frac{\mu-1}{\mu}\right)\frac{P_H^{1/\beta}\left(\mu R\right)^{\frac{\beta-1}{\beta}}}{X\mu_H^{\varphi}}\right)/\varphi\right)$$

$$\tilde{Z}_L = \frac{A_L^W}{A_H^W}\eta_L^{-1}\mu_L^{1+\varphi}H\left(g + \left(r - \eta_H\left(\frac{\mu-1}{\mu}\right)\frac{P_L^{1/\beta}(H/L)^{-1}\left(\mu R\right)^{\frac{\beta-1}{\beta}}}{X\mu_L^{\varphi}}\right)/\varphi\right)$$

Finally, using the budget constraint and the capital accumulation equation, we can derive the dynamics of K

$$I = Y - X(Z_H + Z_L) - C$$

$$\frac{\dot{\tilde{K}}}{\tilde{K}} = I/K - \delta - g = \frac{\mu R}{1-\beta} - X\left(\frac{\tilde{Z}_H}{\tilde{K}} + \frac{\tilde{Z}_L}{\tilde{K}}\right) - \frac{\tilde{C}}{\tilde{K}} - \delta - g$$

The Euler equation completes the dynamical system

$$\frac{\dot{\tilde{C}}}{\tilde{C}} = \frac{1}{\theta}(r - \rho - \theta g)$$

The four difference equations in $\tilde{C}, \tilde{K}, \mu_H, \mu_L$ define the dynamics of the system. We linearize them around the BGP. It is tempting to view this system as one with one control ($\tilde{C}$) and three state variables. However, recall that we have assumed that innovation and capital accumulation are all taking place (i.e. free entry conditions are binding), which, for a given value of initial physical capital, forces the values of μ_H, μ_L. This system only has one negative root, and this root determines the speed of convergence to the BGP.

Setting all the other parameters equal to their calibrated values we let the speed of technology diffusion φ vary and calculate the speed of convergence to the BGP by solving the linearized system described above. Figure 10 plots the results. Our target value for the rate of convergence is 2.5% (Barro, 2015). Values of φ larger than 0.5 produce speeds of convergence well over that target, while those much below it result in too slow convergence. We therefore choose to set $\varphi = 0.5$.

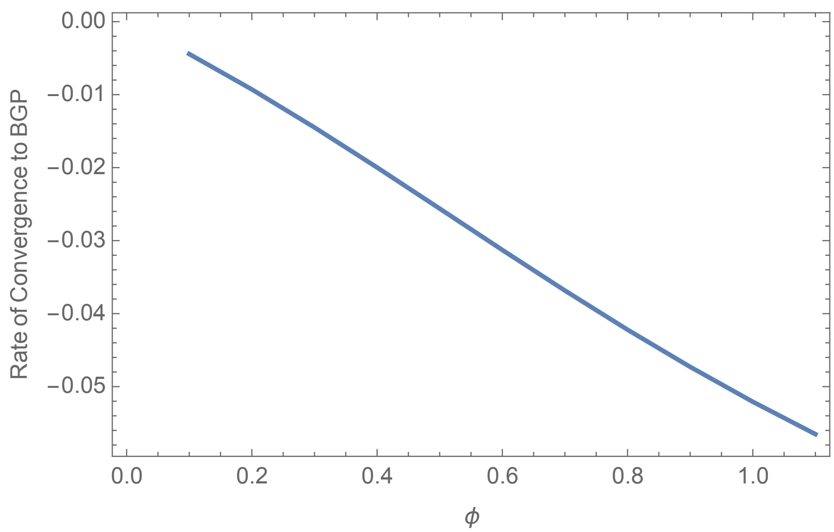


Figure 10: Speed of convergence to the BGP for different values of φ.

G Fit of the Cross-Country Skill-Bias Regressions

Table 2 in the main text reports the results of the Caselli-Coleman cross-country skill bias regressions.In Section 5.4 we re-ran the Caselli-Coleman regressions using counterfactual productivity levels (Table 4). The 2010 cross-sectional relationships are plotted in Figures 11(a) - 11(f) below.

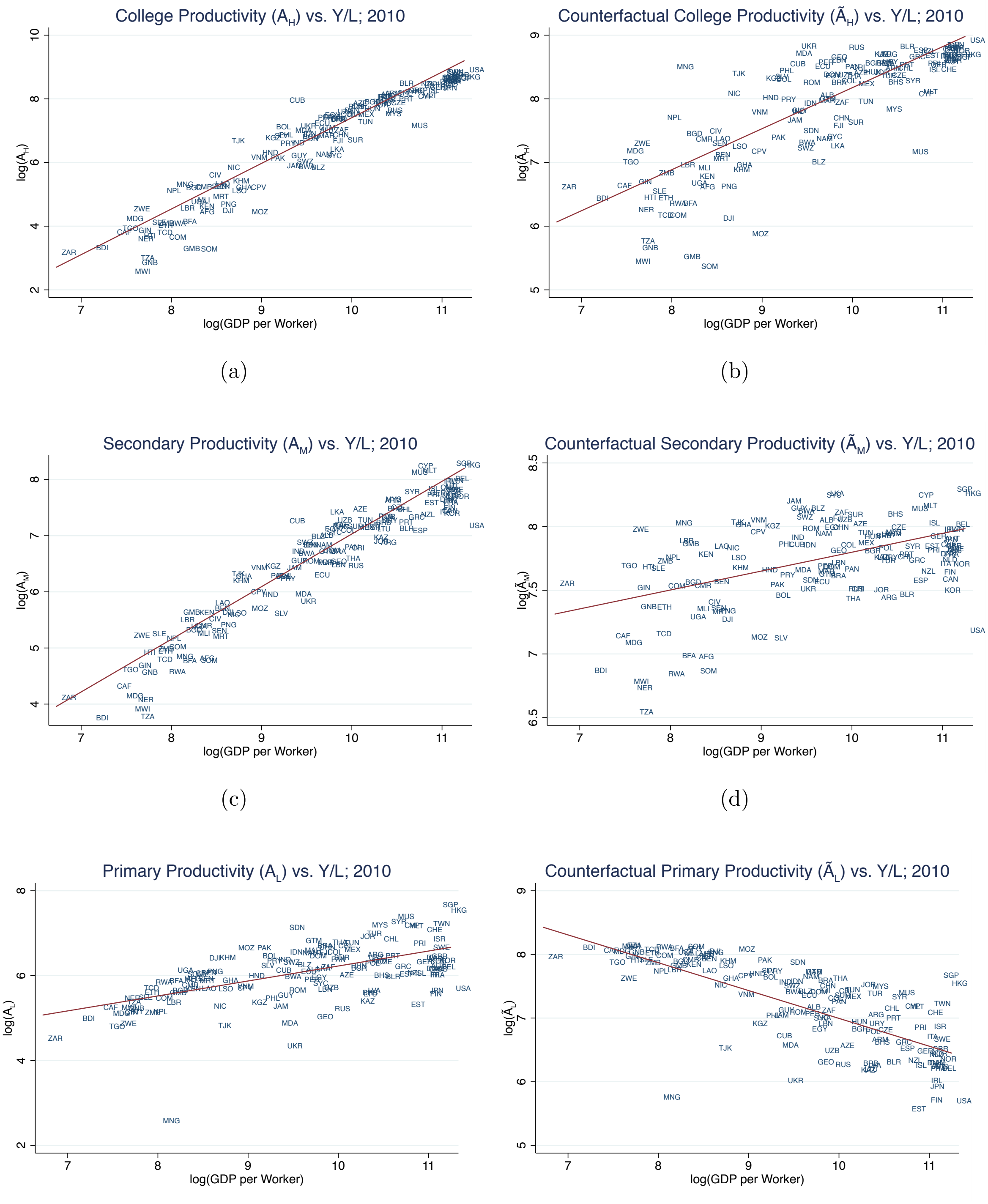


Figure 11: 2010 actual (left column) and counterfactual (right column) productivity levels vs. GDP per worker. Counterfactual productivity levels are computed using the same approach as actual levels but additionally forcing the technology adoption wedge (as a share of GDP per worker) to be the same across countries and equal to that of the U.S.

Wedges & Other Measures of Costs of Entry

As we have discussed in the main text, our measure of technology adoption wedge is a residual that makes the model fit cross-country output per worker data. However, a narrower interpretation, suggested by the theory but not implied by our empirical work, is that of barriers to entry or innovation by existing firms. In this section we consider two ways to check the plausibility of this interpretation. First we correlate our 2010 adoption wedge measures with the World Bank's *Doing Business: Distance to Frontier* indicators, which are designed to measure the cost of opening and operating a new business venture through assessment of entry costs and fees for new firms, licensing requirements, number of legal procedures required and how much time they take, etc. The results are plotted in Figure 12 with a higher score on the Doing Business scale indicating more entrepreneur-friendly environment. Our measures of wedges are the quite strongly correlated with World Bank's assessment, and the direction of the relationship is as expected, that is countries we identified as having a high adoption wedge are usually also deemed unfriendly to business by the World Bank.

The second test of our wedge measures uses tariff data over the period between 1870 and 1990. As is well known, during and after the Great Depression the world witnessed a surge of protectionist policies, especially tariff rate increases, that caused a massive collapse in world trade (O'Rourke and Williamson, 1999). We use the data on tariffs for 12 countries in the period 1875-1987 from Bairoch (1989) and O'Rourke (2000) to test whether the countries that increased their tariffs most are also identified by our methodology as those that have experienced the largest increases in adoption wedges.[22] This is an imperfect test since wedges, even in the narrower interpretation as barriers to entry, encompass much more than just tariffs on imports, but to the extent that protectionist policies were correlated with other policies that restricted entry and competition, this may not be a problem. We regress our wedges on the average tariff rate relative to the U.S. rate since our wedges are also measured in this manner. The results are shown in Table 8 for the two values of elasticity. Using OLS, we find that in both cases there is a positive relationship between relative tariff level and our measure of adoption wedge. This result is robust to the inclusion of country fixed effects. We conclude that the measures of adoption wedge we have identified correspond well to other indicators associated with barriers to entry/innovation available in the literature, both more recent and in more historical data.

[22]The earliest date for which we have tariff data for all 12 countries is 1875. Since our measures of barriers are at decadal frequencies, we interpolate the 1880 tariff rates. We also treat the 1987 tariff values as proxies for 1990 values.

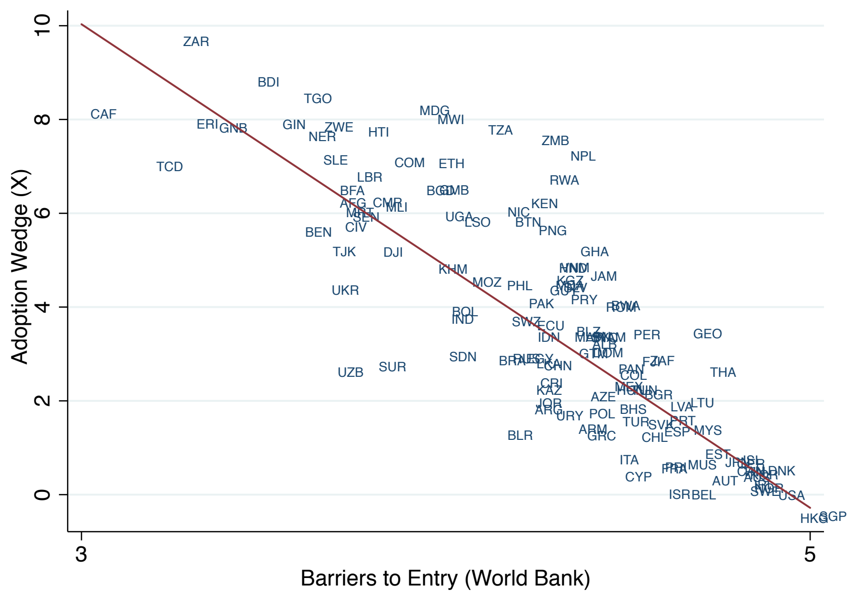


Figure 12: World Bank Doing Business indicator vs. our measures of wedges (2010).

Table 8: Relative Tariffs vs. Wedges; 1880-1990.

	OLS	FE
Relative Tariffs	1.404***	0.807**
	(0.461)	(0.394)
Trend	−0.036***	−0.029***
	(0.008)	(0.006)
Constant	70.450***	57.925***
	(14.539)	(10.967)
R^2	0.123	0.125
N	144	144

H Comparison to Caselli and Ciccone (2013)

Caselli and Ciccone (2013) use a non-parametric formulation to show that the output gains from equating labor skill composition while maintaining the assumption of perfect substitutability among skill-types (the Cobb-Douglas standard approach) is – under certain assumptions – an upper bound on the role of human capital. The is a similar counterfactual to the one we presented at the end of section 5.4, except, of course, we do not assume perfect substitutability.[23] How do our findings compare with theiir lower bound calculations? Table 10 summarizes the percentage gains in GDP per worker from the skill composition change. Both the mean and the median gain was 55%. Caselli and Ciccone reported a mean of 61% and a median of 45%, which is not very different. Removing wedges has a much bigger effect, with a mean gain of 325% and median of 183%. Looking at individual countries, we find that for some our results are fairly close while for others they are not; Table 11 shows a selection of cases.

	Data	Human Capital	DTC
Median	14,541	22,367	42,481
Mean	21,697	29,626	44,479
Coeff. Of Variation	0.92	0.83	0.30
90/10	23.8	16.5	2.2

Table 9: Moments of the 2010 world distribution of output per worker; Column one: data, Column two: counterfactuals computed using the directed technology model with every country's skill composition changed to equal that of the U.S. (without any change in wedges), Column three: counterfactuals computed using the directed technology model and forcing the technology adoption wedge (as a share of GDP per worker) to be the same across countries and equal to that of the U.S.

[23] Our approach is also different from that of Caselli and Ciccone in other ways. The obvious advantage of their method is that it doesn't require the strong parametric assumptions we have made. The downside is that it provides only the upper bound on the gains (under the assumption that skill-types are perfect substitutes, an assumption we depart from based on much evidence from the macro-labor literature as well as our own estimates in Jerzmanowski and Tamura (2017)). Additionally, their method requires data on wages by skill-type and hence limits their analysis to only those countries and years where such data is available (and relies on their quality). Note finally that their approach is robust to only certain kinds of endogenous technological change, which of course is at the heart of our approach.

	Human Capital	DTC (Tech. Wedge)
Mean	0.55	4.23
Max	0.96	29.96
90th percentile	0.90	2.51
75th percentile	0.81	1.38
Median	0.55	1.83
Min	0.09	0.00

Table 10: Moments of the 2010 gain in GDP per worker, defined as $(\tilde{y} - y)/y$, $\tilde{y}$ stands for counterfactual GDP per worker and y represents its actual observed value. Column one: counterfactuals computed using the directed technology model with every country's skill composition changed to equal that of the U.S. (without any change in wedges), Column two: counterfactuals computed using the directed technology model and forcing the technology adoption wedge (as a share of GDP per worker) to be the same across countries and equal to that of the U.S.

Country	Year	CC Bound	DTC
Brazil	2000	0.90	0.56
South Africa	2000	0.71	0.48
Japan	1995	0.26	0.15
Taiwan	1995	0.33	0.31
Vietman	1995	0.41	0.45
India	1995	1.07	0.42

Table 11: The percentage of GDP per worker gained when skill distribution is changed to equal that of the U.S. Caselli and Ciccone (2013) (labeled CC bound) vs this paper (labeled DTC)