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 1, 2017

Abstract

Research aimed at understanding cross-country income differences finds that inputs of human and physical capital play a limited role in explaining those differences. However, most of this work assumes workers with different education levels are perfect substitutes. Does moving away from this assumption affect our conclusions about the causes of long run development? To answer this question we construct measures of skill-specific productivity and barriers to innovation for a large sample of countries over the period 1910-2010. We use a model of endogenous directed technological change together with a new data set on output and labor force composition across countries. We find that rich countries use labor of all skill categories more efficiently, however, in the absence or barriers to entry, poor countries would actually be more efficient at using low-skill labor. Our estimates imply that after 1950 the world technology frontier expanded much faster for college-educated workers than for those with lower skill sets. This technology diffused to many countries, allowing even poorer countries to experience relatively robust growth of high-skill-specific productivity. Their GDP growth failed to reflect that because of their labor composition; they have very few workers in the higher skilled category. Finally, we investigate the relative importance of factor endowments versus barriers to technology in explaining the current disparities of standards of living and find it to depend crucially on the value of the elasticity of substitution between skill-types. Under a lower value of 1.6, our model yields barrier estimates that are lower and relatively less important in explaining cross-country income differences: in this scenario physical and human capital account almost 70% of variance in 2010 GDP per worker in our sample. Using elasticity of 2.6, we find barriers that are higher and explain most of the variation in output. We provide some evidence that the higher value of elasticity is preferred.

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 education attainment, which goes back to 1910 for many countries.[1] We explore quantitatively how the key mechanism of this theory – the link from skill composition of the labor force to 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 146 countries over the period 1950-2010. Additionally, our calculations let us back out the levels of barriers to adoption of new technologies, a potentially important determinant of long run relative level of development in the model. With these measures in hand we explore how efficiency of workers with different skill levels differs across countries and how it has evolved over the last century. We also characterize the distribution and evolution of barriers to adoption of technology. 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, that is the relative efficiency of workers with different skill sets always remains constant. However, there is is recent evidence suggesting that relaxing the factor-neutrality assumption may be a

[1]We use the new data from Tamura, Dwyer, Devereux and Baier (2016). That data covers 168 countries, and provides information on schooling, physical capital and output per worker in census years. We have complete information on these variables for 69 countries in 1910, 66 countries in 1870, 58 countries in 1820, and 18 countries in 1800. We concentrate on the century 1910-2010 because we have additional information on relative wages of high skilled, medium skilled and low skilled workers for this period.

fruitful avenue for understanding cross-country income differences. Caselli (2005) and Caselli and Coleman (2006) provide pioneering empirical analyses of the importance of factor-bias for aggregate productivity. Jerzmanowski (2007), using non-parametric techniques, argues that factor-neutral productivity is not a good approximation of the world production possibilities frontier. Moreover, the large 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 last 30 or so years. Subsequent contributions by Acemolgu (1998, 2002) provide a rich theoretical analysis of the mechanisms through which the direction of the technology bias is endogenously determined by relative supplies of factors, specifically the skill composition of the labor force. Unfortunately, unlike in the case of factor-neutral TFP, calculating skill-specific productivity levels poses a greater 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 important 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, and international technology diffusion. We calibrate the parameter values of the model and use its equilibrium conditions to back out skill-specific productivity levels and measures of barriers to innovation, requiring only data on output, factor inputs, and shares of the labor force with primary, secondary, and college education. These variables come from a new data set constructed by Tamura, Dwyer, Devereux, Baier (2016), which covers 168 countries over the period 1950-2010 with information for some going back to 1910.[2]

We use our productivity and barriers measurements to (1) study the historical patterns of directed technological change and (2) evaluate the contribution of non-neutral technology and barriers to cross-country income differences. One of the key parameters of our empirical strategy is the elasticity of substitution between labor of different skill types. We consider two values of this parameter. The value of 1.6 is closer to earlier estimates found in the literature on skill-bias (Katz and Murphy 1992) but under this scenario our model does not exhibit the so-called strong skill-bias, a condition necessary for the directed technological change to be a successful explanation of the recent rise in US skill premium. Since we want to explore the implications of the directed technology paradigm, including strong skill-bias, for the world income distribution, we also compute our results under a higher value of elasticity of 2.6. This value is more in line with some more recent findings (Acemoglu and Autor 2015), as well as

[2]The fact that we have three categories of skill may itself be an important improvement since evidence suggest binary division of the labor force into skilled and unskilled groups may be too restrictive (See Acemoglu and Autor, 2015)

our own estimates, found using EU KLEMS data and the equilibrium expression for the skill premium from our model (Jerzmanowski and Tamura 2017). The elasticity parameter turns out to be important for some of the results and we argue the lower elasticity scenario often leads to less plausible findings.

Regardless of the value of elasticity, our estimates of the world technology frontier imply that in the early part of the 20th century technological progress favored high-school-educated workers over those with more education, while after 1950 – with a brief interruption in the 1970's – college-specific frontier productivity growth outstrips the other categories as technological progress turns decisively college-biased. Comparing skill-specific productivity across countries, we find that rich countries use labor of all three skill categories more efficiently. However, low productivity levels in poor countries arise from high barriers to entry. A counterfactual calculation that removes barriers, reveals that in their absence poor countries would actually achieve higher productivity in the lower skilled sectors compared to their richer counterparts. Our results also shed light on the puzzle of low (or negative) TFP growth in many countries often found using the factor-neutral approach.[3] This 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, our calculations show that poor countries also experience relatively robust growth of college-specific productivity. In some instance it is indeed offset by increases in barriers but by and large barriers have been falling, even in poor countries. Their GDP growth fails to reflect that 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 frontier has 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 barriers to technology in explaining the current disparities of standards of living, what turns out to matter crucially is the magnitude of barriers in poor countries. And this is where the

[3]When researchers look at growth rates instead of levels, they usually find that it is TFP growth that mostly accounts for long run growth rates of output per worker (Easterly and Levine 2001). Of course, this means that developing countries, many of which have had mediocre output growth, must have had very slow TFP growth. In fact, significant fraction of countries are usually found to have negative TFP growth even over very long periods of time. In our sample 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. Tamura, Dwyer, Devereux and Baier (2016) however 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.

value of the elasticity of substitution between skill-types makes a critical difference. Under a lower value of 1.6, our model yields barrier estimates that are lower and relatively less important in explaining cross-country income differences: under this scenario endowments of physical and human capital account almost 70% of variance in 2010 GDP per worker in our sample. In contrast, when we use a higher elasticity of 2.6, we find barriers that are larger and explain almost 80% of the variation in output.[4] To the extent that reduction of barriers to entry and increasing educational attainment involve distinct approaches, our results imply that policies most effective at raising income levels in poor countries may be quite different depending on how difficult it is to substitute unskilled labor for its more skilled counterpart. Given that, under the constraints imposed by our theoretical approach, the preferred value of elasticity seems to be higher, we see our findings as providing more support for the barrier reduction approach.

2 Related Literature

Our work is part of the large literature which tries to understand the causes 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 suggest strongly that factors are not as important as the largely unexplained total factor productivity (Hall and Jones, 1999, and Klenow and Rodriguez-Clare, 1998). For example, Hsieh and Klenow (2010) using the standard development accounting approach 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. This has led to a large literature trying to understand TFP differences. The prevailing approach within this literature has been to treat productivity as factor-neutral, that is affecting the efficiency of all factors equally. 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 con-

[4] While this is similar to the finding under traditional factor-neutral approach, which attributes most of the cross-country income variation to productivity, we emphasize that even in the high elasticity scenario our results suggest a slightly more important role of human capital in case of some countries.

ditions 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. Caselli computes two counterfactuals. First, he endows all countries with the US productivity levels for both factors. He finds that if the elasticity of substitution between the two types of capital is sufficiently below 1, he is able to explain a very large fraction of cross-country income variation using factor endowments alone without the need for enormous technological differences across countries. His second counterfactual allows countries to choose any pair of physical/human capital productivities, in effect treating all the computed pair as a "menu" of technological choices. Here the the importance of factors is somewhat smaller as countries offset their unfavorable factor endowments by appropriate choice from the technology menu. 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, precisely as we do in this paper. Using binary skill classification, they find that rich countries use skilled labor more efficiently than than their low-income counterparts. However, the opposite true for unskilled labor, which they find to be relatively –and in some specifications even absolutely – more productive in poor countries. We relate our findings to Caselli and Coleman's by computing the measures of cross-country skill bias. Two recent papers that explore the consequences of imperfect substitutability between workers with different skill levels for understanding international income differences are Caselli and Ciccone (2013) and Jones (2014). The former paper demonstrates 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. We compute labor force composition counterfactuals using our approach and compare them to their results. Jones develops a generalized method of aggregating human capital, which allows for the marginal product of worker types to be non-constant and depend on supplies of workers of other categories, and demonstrates that under certain assumptions, his approach implies a much larger role of capital inputs in explaining output variation across countries. Like us, Jones acknowledges that the available estimates of elasticity of substitution between skill types might not be right in cross-country setting and he too computes his decompositions under different values of elasticity. We find that, although our approaches are different, our results align quite closely when we use comparable elasticity values.

In their computation of skill-specific productivity levels Caselli and Coleman (2006), Caselli and Ciccone (2013), and Jones (2014) rely on international wage data, while Caselli (2005) uses capital's income share data.[5] All of these have the drawback of sparse coverage and often questionable quality. The advantage of our approach to gauging the role of factors 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), construct a model with the same key elements as ours: directed technological change, capital accumulation, and technology diffusion subject to barriers. Like us, they are interested in quantitative implication of directed technology model for understanding international income differences, especially the role of barriers to technology diffusion, factor endowment and the shape of the world technology frontier.[6] Among their key findings is that barriers to technology adoption 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. Their approach, especially its empirical part, differs from ours in several important 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 barriers to technology adoption over a longer time period. Second, our data distinguishes between three levels of skills: primary, high school 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 productivity of workers with primary and secondary education in the US are not the highest in the world. There are also differences in our theoretical models. In our specification of the technology diffusion process we follow Howitt (2000) and Klenow and Rodriguez (2005) and choose to neutralize scale effects. Also, in our formulation machines are produced using physical capital not labor. However, the most important differences are in how we take the model to the data and what element of the model we associate with barriers to innovation. Like us Gancia et al. use the US skill premium to calibrate their model. In particular, they use the skill premium equation for 1970 and 2000 to back out the slope (i.e. the elasticity of substitution between skill types) and the intercept of that equation. Similar equations arise in our model, however, we cannot simultaneously back out the elasticity and the intercept, because we allow the latter to vary across decades. This is natural since the intercept represents relative productivity of R&D in different sectors, something that surely changed during our sample period.[7] Interestingly enough, their method leads them to a value

[5]Strictly speaking, some of these papers use data on returns to education data.

[6]They also consider trade openness and IPR which we abstract from in this paper.

[7]Assuming it to be constant may be better justified over a shorter period of time, as in Gancia et al, who

of elasticity of about 2.3, close to our preferred value of 2.6. Additionally, working with a similar structural equations to ours, Gancia et al. estimate the level of barriers assuming uniform values within the three groups: OECD, sub-Saharan and non-OECD economies. Our approach allows us to compute exactly the country-specific relative levels of barriers. Compared to Gancia et al. we also find that barriers are very important for understanding international income differences under our preferred elasticity of substitution value. However, their removal would not lead to income gains as big as those reported in their paper. Some of this may be related to the differences in the approach outlined above, however the final crucial difference seems to be more important. Gancia et al. identify barriers with the parameter that controls the speed of technology diffusion. As a consequence, they equate barrier removal with the speed of diffusions becoming infinite. One important consequence of this approach is that in the absence of barriers all countries share the same technology. In contrast, we model barriers as a cost of entry into innovative activity. The removal of this cost boost technology levels but local conditions such as market size and skill supplies continue to matter. As a result even in the absence of barriers, not all countries end up with identical technologies.[8]

We are also directly related to the mostly theoretical literature which attempts to explain the low TFP in poor countries. The inefficiency view, emphasized by Parente and Prescott (1994,1999) and Olson (1982, 1996) among others, seeks to explain low technology with barriers to entry, which prevent countries from adopting otherwise readily available best-practice (frontier) technologies or cause them to use such technologies below full potential. Such barriers include corruption, licensing, excessive regulation, labor laws, or other arrangements that limit competition, outright prohibit entry of new firms or otherwise distort allocation of resources away from most productive uses. Our approach enables us to compute the measures of barriers to entry and thus allows us to shed light on the relative importance of this hypothesis.

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 In their model technology is factor-neutral and so their goal is not to back out TFP levels but to see if the model can be successfully calibrated to match the world income distributions. They conclude that it can, provided strong technology spillovers and sufficient barriers to innovation (taxes on R&D) are included. In their approach, as in ours, factors feedback onto TFP by creating incentives to innovate. This is also true in Cordoba and Ripoll (2008) who conduct a development accounting in an endogenous growth model. However unlike us, they assume factor-neutral productivity differences. They report

only look at 1970 and 2000.

[8]Under the lower elasticity scenario, we find barriers to be to far less important than in Gancia et al.

that under certain parametrizations of the technology diffusion process, their model implies factors can account for between 64% and 74% of cross-country variation in incomes. However, the unattractive feature of their result is that in order for factors to play such a large role, implied technology differences are very small: the average country has 90% of US technology. In our approach, we find international technology differences that are large and varying across sectors.

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 barriers to entry on the one side and the levels of skill-specific productivity and relative wage of workers with different skill types on the other.

3.1 The Model

Our model incorporates physical capital, technology diffusion and barriers to technology adoption 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 to improve technology even though they enjoy the benefit of being able to tap into the world pool of knowledge. This cost of innovation reflects the resources necessary to implement a given technology and adapt it to the local conditions as well as the barriers to such activity, such as compliance with regulatory requirements, licensing fees, bribes, etc. 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 two skill types: high skilled (H) and low skilled (L).[9] Population and type shares, s_i, are constant

[9] In the empirical implementation of the model we will use three skill categories: those exposed to higher education, H, those exposed to secondary school (but not higher education), S, and those with no education or exposed to at most primary school, P. For clarity of exposition, we focus on a version with just two skill categories: H (skilled) and L (unskilled) when we derive the key conditions of the model. Key results are restated for the three-skill type model at the end of this section.

$$H = s_H N,$$
$$L = s_L N$$

and

$$s_H + s_L = 1.$$

The households own physical capital and patents rights on innovation and maximize the present discounted value of an infinite steam 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 the tax rate on capital income 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[10]

$$Y = \{Y_H^{\frac{\varepsilon-1}{\varepsilon}} + Y_L^{\frac{\varepsilon-1}{\varepsilon}}\}^{\frac{\varepsilon}{\varepsilon-1}} \tag{2}$$

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

$$\left[P_L^{1-\varepsilon} + P_H^{1-\varepsilon}\right]^{\frac{1}{1-\varepsilon}} = 1. \tag{3}$$

Intermediate Goods & Machines

The intermediate goods are produced using labor of a single skill type and machines designed for workers with that specific skill type. At a point in time, there is a continuum A_i of machine

[10] In the empirical implementation, where we use three skill categories mentioned earlier, the production function becomes:

$$Y = \{Y_H^{\frac{\varepsilon-1}{\varepsilon}} + Y_S^{\frac{\varepsilon-1}{\varepsilon}} + Y_P^{\frac{\varepsilon-1}{\varepsilon}}\}^{\frac{\varepsilon}{\varepsilon-1}} \tag{1}$$

varieties available for each skill type $i = H, L$. These evolve over time as innovators invent new machine varieties and are distinct in the sense that machines designed to be used by one skill type cannot be used by another skill type. Since the the sectors are symmetric, we will conserve space by presenting equations for the L-type sector only. Corresponding equations for the H sector can be easily obtained by replacing L with H.

Intermediate goods producers combine machines and labor in the standard "variety of machine inputs" manner

$$Y_L = \frac{1}{1-\beta} \int_0^{A_L} \chi_{jL}^{1-\beta} dj \, L^\beta \tag{4}$$

$$\tag{5}$$

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

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

$$\max_{\{\chi_{jL}, L\}} \{ \frac{P_L}{1-\beta} \int_0^{A_L} \chi_{jL}^{1-\beta} dj L^\beta - \int_0^{A_L} p_{jL} \chi_{jL} dj - w_L L \}, \tag{6}$$

where p_{jL} is the price of variety j, L-type machine.

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

$$P_L \chi_{jL}^{-\beta} L^\beta = p_{jL} \tag{7}$$

Blueprints for machines varieties are specific to the economy. They are invented by local entrepreneurs who hold perpetual monopoly rights over a given variety they have invented within the country. As we will see later, technology diffusion means that these innovators benefit form 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.[11]

Machines are supplied to the intermediate goods producers by the monopolists who own the blueprints and rent capital to manufacture the machines. Capital is rented in a competitive market at the capital rental rate R. One unit of physical capital can produce one machine of any variety and machines depreciate at a rate of 100%.[12] We also assume (following Aghion and Howitt, 2008) that each machine producing monopolist faces a potential imitator with

[11]For example, we can think of a innovations consisting in large part of how to adapt a given variety to local conditions. For example electrical systems around the world differ in their reliability, fluctuations in amperage, etc.

[12]In the specification of Acemoglu (1998), machines are produced using the final good.

cost $\upsilon > 1$ times higher 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[13]

$$p_{jL} = \upsilon R \tag{8}$$

The equilibrium supply of machines of type j to skill L, and the equilibrium quantities of machines are:

$$\chi_{jL} = \left(\frac{P_L}{\upsilon R}\right)^{1/\beta} L \tag{9}$$

which means the (derived) production functions of intermediate goods become

$$Y_L = \frac{1}{1-\beta}\left(\frac{P_L}{\upsilon R}\right)^{\frac{1-\beta}{\beta}} A_L L \tag{10}$$

And the profit per line of machines is given by

$$\pi_{jL} = \left(\frac{\upsilon - 1}{\upsilon}\right) P_L^{1/\beta} L \left(\upsilon R\right)^{\frac{\beta-1}{\beta}} \tag{11}$$

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

$$\frac{P_H}{P_L} = \left(\frac{A_H H}{A_L L}\right)^{-\frac{\beta}{\sigma}} \tag{12}$$

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

Wages & Technology

Intermediate goods producers hire labor according to the following first order condition:

$$\frac{\beta P_L}{1-\beta}\int_0^{A_L} \chi_{jL}^{1-\beta} dj L^{\beta-1} = w_L, \tag{13}$$

[13]This is true as long as $\upsilon < 1/(1-\beta)$, which we assume to be true.

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

$$w_L = \frac{\beta}{1-\beta} A_L \beta P_L^{\frac{1}{\beta}} (\upsilon R)^{-\frac{1-\beta}{\beta}} \tag{14}$$

Thus the relative wages of workers with different skill levels are given by

$$\frac{w_H}{w_L} = \left(\frac{A_H}{A_L}\right)^{\frac{\sigma-1}{\sigma}} \left(\frac{H}{L}\right)^{-\frac{1}{\sigma}} \tag{15}$$

where σ is the elasticity of substitution between H and L.

Capital Allocation & Rental Rate

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

$$K_L = \int_0^{A_L} \chi_{jL} dj = A_L \left(\frac{P_L}{\upsilon R}\right)^{1/\beta} L \tag{16}$$

and it follows that

$$\frac{K_H}{K_L} = \left(\frac{A_H}{A_L}\right)^{\frac{\sigma-1}{\sigma}} \left(\frac{H}{L}\right)^{\frac{\sigma-1}{\sigma}} \tag{17}$$

In addition, total capital stock by K is given by

$$K = K_L + K_H. \tag{18}$$

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

$$K_L = A_L \chi_L \tag{19}$$

$$\tag{20}$$

Substituting this into the intermediate goods production functions

$$Y_L = \frac{1}{1-\beta} K_L^{1-\beta} (A_L L)^{\beta} \tag{21}$$

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

$$P_L MPK_L = P_L K_L^{-\beta}(A_L L)^{\beta} = P_L(1-\beta)\frac{Y_L}{K_L} \tag{22}$$

It is easy to show, using the expressions for P_H/P_L and K_H/K_L 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 mark-up over cost of producing machines (the rental rate). This implies that

$$\upsilon R = P_L MPK_L = P_L K_L^{-\beta}(A_L L)^{\beta} = P_L(1-\beta)\frac{Y_L}{K_L}$$

and we have

$$R = \frac{P_L MPK_L}{\upsilon} = P_L\left(\frac{1-\beta}{\upsilon}\right)\frac{Y_L}{K_L} == \left(\frac{1-\beta}{\upsilon}\right)\frac{Y}{K} \tag{23}$$

where the last equality follows from the zero-profit condition in the final goods sector. To understand this expression, note that the equilibrium MPK is given by $(1-\beta)Y/K$ and the rental rate is equal to MPK/υ where υ is the markup in the machine market. Finally, the interest rate in this economy is given by $r = (1-\tau)R - \delta$, where δ is the rate of depreciation of capital and τ is the tax on capital income

$$r = (1-\tau)\left(\frac{1-\beta}{\upsilon}\right)\frac{Y}{K} - \delta. \tag{24}$$

Innovation

Discovery of new blueprints for sector i is governed by the following process

$$\dot{A}_L = \eta_L\left(\frac{A_L^W}{A_L}\right)^{\phi}\frac{Z_L}{N} \tag{25}$$

where represents A_L^W is the world frontier technology for sector L, η_L is the productivity of research effort, and Z_L is the R&D expenditure on innovation or technology adoption in sector L. In Acemoglu's original directed technology formulation $\phi = 0$. Since our focus is understanding productivity differences across countries, we relax this assumptions to allow for diffusion of technology ($\phi > 0$). 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 ζ, which is the same in both sectors and represents the costs of implementation/adaptaion of new technology as well as expenses on overcoming barriers to introducing new technologies, such as compliance with regulatory requirements, licensing fees, bribes, etc. (Parente and Prescott 1994, 1999).[14] Note that in Acemoglu's original formulation $\zeta = 1$. We again relax this assumption to allow the model to account for cross country differences in productivity.[15]

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

$$\eta_L \left(\frac{A_L^W}{A_L} \right)^{\phi} \frac{V_L}{N} = \zeta \tag{26}$$

where V_L is the value of a blueprint for a machine in sector L. Defining $\mu_L = \frac{A_L}{A_L^W}$ and dropping the country indicator, this equation implies that

$$\frac{V_H}{V_L} = \left(\frac{\eta_H}{\eta_L} \right)^{-1} \left(\frac{\mu_H}{\mu_L} \right)^{\phi} \tag{27}$$

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

$$r\, V_L = \pi_L + \dot{V}_L \tag{28}$$

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).[16]

[14] To avoid clutter we omit the country subscripts but it's important to keep in mind that ζ is country-specific. In the empirical part below, we will compute the values of this parameter for a large sample of countries.

[15] As explained in the introduction, this is one of the key differences between our approach and that of Gancia et al. (2013). In their interpretation, barriers are measured by the diffusion parameter ϕ, with removing barriers equivalent to setting $\phi \to \infty$. This has important consequences for the findings about the importance of barriers and we come back to it below.

[16] Since innovation for the two skill types and capital accumulation technologies are linear, the transitional

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

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

$$r^{*} = \theta g + \rho, \tag{29}$$

and, using equation equations (23) and (24), the BGP rental rate is

$$R^{*} = \frac{\theta g + \rho + \delta}{1 - \tau} \tag{30}$$

Using the no-arbitrage conditions from (28) and the fact that along the BGP the value of a patent must be stationary ($\dot{V}_L = 0$) we get the following relationship between the value of a patent, profits and the interest rate

$$V_L = \frac{\pi_L}{r} \tag{31}$$

where profits are given by $\pi_i = \left(\frac{\upsilon-1}{\upsilon}\right) P_i^{1/\beta} N_i \left(\upsilon R\right)^{\frac{\beta-1}{\beta}}$. It follows that

$$\frac{V_H}{V_L} = \frac{\pi_H}{\pi_L} = \frac{\left(\frac{\upsilon-1}{\upsilon}\right) P_H^{1/\beta} H \left(\upsilon R\right)^{\frac{\beta-1}{\beta}}}{\left(\frac{\upsilon-1}{\upsilon}\right) P_L^{1/\beta} L \left(\upsilon R\right)^{\frac{\beta-1}{\beta}}} = \left(\frac{P_H}{P_L}\right)^{1/\beta} \frac{H}{L},$$

which can be further simplified using the expression for relative prices to obtain

$$\frac{V_H}{V_L} = \left(\frac{A_H}{A_L}\right)^{-\frac{1}{\sigma}} \left(\frac{H}{L}\right)^{\frac{\sigma-1}{\sigma}} \tag{32}$$

Finally, combining equations (27) and (32), yields[17]

$$\frac{A_H}{A_L} = \left(\frac{\eta_H}{\eta_L}\right)^{\frac{\sigma}{1+\phi\sigma}} \left(\frac{H}{L}\right)^{\frac{\sigma-1}{1+\phi\sigma}} \left(\frac{A_H^W}{A_L^W}\right)^{\frac{\phi\sigma}{1+\phi\sigma}} \tag{33}$$

Thus the relative levels of productivity are increasing in the relative supply of skilled workers H/L as long as $\sigma > 1$. Recall that we refer to technological change as skill biased

dynamics may involve 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 this section.

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

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

(H-biased in our notation) whenever an increase in the level of technology raises the relative marginal product of skilled workers. From equation (15) it is clear that an increase in A_H/A_L is a skill-biased technological change as long as $\sigma > 1$. In Acemoglu's terminology, weak equilibrium skill-bias occurs whenever an increase in H/L induces skill-biased technological change. Equation (33) thus implies that we have weak equilibrium skill-bias whenever $\sigma > 1$.[18]

Substituting the expression for relative productivity levels (33) into the relative wage formula (15) we obtain

$$\frac{w_H}{w_L} = \left(\frac{\eta_H}{\eta_L}\right)^{\frac{\sigma-1}{1+\phi\sigma}} \left(\frac{H}{L}\right)^{\frac{\sigma-2-\phi}{1+\phi\sigma}} \left(\frac{A_H^W}{A_L^W}\right)^{\frac{\phi(\sigma-1)}{1+\phi\sigma}}. \tag{34}$$

Notice in equation (15) that an increase in H/L, besides its effect on A_H/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 refer to it as strong equilibrium skill bias. Clearly the *strong skill bias* is present as long as

$$\sigma > 2 + \phi$$

which reduces to $\sigma > 2$, a result familiar from Acemoglu (2009), when $\phi = 0$. Notice that the presence of international technology diffusion ($\phi > 0$) 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.

From equations (26) and (31) [DO YOU MEAN (27) and (33)?] we can see that on the BGP productivity relative to the frontier is given by

$$\mu_L = \left[\frac{\eta_L \left(\frac{\upsilon-1}{\upsilon}\right)(L/N)P_L^{*1/\beta}\left(\upsilon R\right)^{\frac{\beta-1}{\beta}}}{r^*\zeta}\right]^{1/\phi}, \tag{35}$$

This equation implies that, all else equal, countries with greater barriers to entry (ζ) will find themselves further away from the frontier. Additionally, the productivity level will be closer to the frontier when the price of machines of L-type is higher (νR), the supply of L-type workers is greater (L/N), and the price of L-type intermediate good is higher (P_L). The latter two terms are affected in opposite directions by a change in the skill supply ratio H/L. When this

[18] When $\sigma < 1$, we always have weak equilibrium unskill-bias.

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

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

where

$$s_H = \frac{H}{H+L} = \frac{H}{N}$$

.

The appendix shows that this reduces to

$$\frac{Y}{N} = \left(A_L^W\right)^{\frac{\phi}{1+\phi}}\left(\frac{1-\beta}{\upsilon}\right)^{\frac{\beta-1}{(1+\phi)\beta}}\left(\frac{K}{Y}\right)^{\frac{1-\beta}{\beta}}\left(\frac{\zeta}{Y/N}\right)^{-\frac{1}{1+\phi}}\Omega\left(\frac{H}{L},\frac{A_H^W}{A_L^W}\right)^{\frac{\phi}{(1+\phi)\beta}} \tag{36}$$

Thus on the BGP output per worker depends on the world technology frontier (A_H^W and A_L^W) , capital accumulation (K/Y), domestic relative supply of skills (H and L), and barriers to entry relative to GDP per worker (ζ/y).[19]

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 barriers to entry (ζ). To achieve this goal we use data on output (Y), capital (K) and labor supply by three skill categories (primary school only, high school and college) from Tamura et al. (2016). Those authors construct their series of estimates of real output, physical capitals 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. Most important for us is that these data allow us to construct supplies of workers at three distinct education levels: at most primary,

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

at least some secondary (but no college), and at least some college.[20] The data is at decadal frequency and covers 58 countries for the years 1820-2010 and 157 countries for the period 1950-2010, which is where we focus most of our analysis.

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). We then proceed as follows. First, we pick values for the following parameters: $\sigma, \phi, \upsilon, , \lambda' s$, and β (see next section for details of parameter value choices) and next we iterate on the following procedure.

1. In the first iteration we assume $\frac{A_C^W}{A_P^W} = \frac{A_S^W}{A_P^W} = 1$.

2. We solve for η's using the the Katz and Goldin (2008) data on U.S. historical wage premiums and associated relative skill supplies and the following two equations

$$\frac{w_C}{w_S} = \left(\frac{\eta_C}{\eta_S}\right)^{\frac{\sigma-1}{1+\phi\varepsilon}} \left(\frac{s_C LF}{s_S LF}\right)^{\frac{\sigma-2-\phi}{1+\phi\varepsilon}} \left(\frac{A_C^W}{A_S^W}\right)^{\frac{\phi(\sigma-1)}{1+\phi\varepsilon}}, \tag{37}$$

$$\frac{w_C}{w_P} = \left(\frac{\eta_C}{\eta_P}\right)^{\frac{\sigma-1}{1+\phi\varepsilon}} \left(\frac{s_C LF}{s_P LF}\right)^{\frac{\sigma-2-\phi}{1+\phi\varepsilon}} \left(\frac{A_C^W}{A_P^W}\right)^{\frac{\phi(\sigma-1)}{1+\phi\varepsilon}}. \tag{38}$$

   where C stands for "college", S for "high school", and P for "primary" (which includes primary and those with no schooling at all); LF is labor force and s_i is the share of the education group i in labor force. (See the discussion below for details)

3. Next we solve for the relative productivity levels using versions of

$$\frac{A_i}{A_P} = \left(\frac{\eta_i}{\eta_P}\right)^{\frac{\sigma}{1+\phi\sigma}} \left(\frac{s_i}{s_P}\right)^{\frac{\sigma-1}{1+\phi\sigma}} \left(\frac{A_i^W}{A_P^W}\right)^{\frac{\phi\sigma}{1+\phi\sigma}}$$

   for $i = C, S$.

4. We then compute A_P, A_S, and A_C. To do this note that output per worker can be expressed as

[20] In Tamura, Dwyer, Devereux, and Baier (2016) the lowest education category is broken out separately into no education, and those exposed to at most primary school. However by the end of the time frame, 2010, no rich country has any population without schooling. In order not to deal with this issue, we combined the bottom two skill categories into one.

$$\frac{Y}{N} = A_P^{\beta}\left(\frac{K}{N}\right)^{1-\beta}\Omega\left(s_C, s_S, s_P, \frac{A_C^W}{A_P^W}, \frac{A_S^W}{A_P^W}\right)$$

so that the productivity of P-type workers is given by

$$A_p = \left(\frac{(Y/N)/(K/N)^{1-\beta}}{\Omega\left(s_C, s_S, s_P, \frac{A_C^W}{A_P^W}, \frac{A_S^W}{A_P^W}\right)}\right)^{1/\beta}$$

.

We then use equations equivalent to (33) to compute A_C and A_S.

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

6. We and solve for the frontier by iteration on steps (3)-(6) to find the fixed point of the following problem

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

where $\mathcal{D}$ stands for our data, $\mathcal{A}_{ih}$ is the vector of sector i productivity levels for all countries in our sample in year t computed using the BGP conditions of the model and our data as outlined below, and A_{itn}^W is the value of the frontier productivity level for sector i in year t found in the n-th iteration of our algorithm.

7. We normalize the level of barriers to entry in the US to be equal to one and using the equation (35) we get

$$\frac{\zeta_k}{\zeta_{US}} = \left(\frac{P_{L,k}^*}{P_{L,US}^*}\right)^{1/\beta}\left(\frac{R_k^*}{R_{US}^*}\right)^{\frac{\beta-1}{\beta}}\left(\frac{A_{L,US}}{A_{L,k}}\right)^{\phi}\left(\frac{s_{L,k}}{s_{L,US}}\right) \tag{39}$$

which allows us, using the A's computed previously as well as expressions (23) and the equations for intermediate good prices (see Appendix), to compute the (relative) level of barriers for each country.

4.2 Choice of Parameter Values

We have to pick values for the following parameters in our model: η's, σ, ϕ, υ and β Unfortunately for many there is very little guidance in the existing literature. If this is the case, we make some judgment calls and experiment with several possible values. β is the labor's income share and we choose a value of 2/3, in agreement with Gollin (2002). We use a values of 1.4 for the markup based on 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%, see Barro and Sala-i-Martin (2003).[21] Lastly, we follow Gourinchas and Jeanne (2004) and choose the inter-temporal 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 world interest rate (in the absence of capital distortions) of 5.2%.

The choice of σ turns out to be very important for some of our results. Based on estimates of Katz and Murphy (1992) and more recently Ciccone and Perri (2005) is about 1.4 or 1.6, but could be above 2.

In a seminal contribution, Katz and Murphy (1992) estimate σ using college/high school wage premium for the years 1963-87 and get 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 instrumental variables strategy (since H/L responds to shock to wages, OLS may be inconsistent) and data across US states. They find σ close to 1.5. Most recently however, Acemoglu and Autor (2011) argue that higher values of σ are also plausible. For example, using Katz and Murphy's regression on updated data they find $\sigma = 2.9$. An additional concern, shared by Jonea (2014), is whether the existing US-basedl estimates are appropriate for a cross-country setting. In a companion paper Jermanowski and Tamura (2017), we estimate estimate σ using a 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 High School degree), medium-skill (High School degree) and high-skills (College). 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 2.6, the value required for strong skill bias given our calibration of the other parameters. Since we want to explore the implications of the directed technology paradigm, including strong skill-bias, for the world income distribution, we compute and compare the results under two different values of elasticity: 1.6 and 2.6.[22] Finally, we calculate the relative research efficiencies η_C/η_P and η_{HS}/η_P using the data from

[21] See the Appendix for the discussion of the model's dynamics.

[22] Our Appendix briefly explains our approach and summarizes our estimates of σ using the EU KLEMS data.

Goldin and Katz (2008), who report relative wages and supplies of workers with different educational attainment for the US economy since 1910 based information from the censuses.[23].

Table 1: College and High School Premia from Goldin and Katz

Year	w_C/w_P	w_C/w_{HS}	w_{HS}/w_P
1915	2.74	1.89	1.45
1940	2.33	1.65	1.41
1950	1.69	1.37	1.24
1960	1.87	1.49	1.26
1970	2.00	1.59	1.26
1980	1.86	1.48	1.26
1990	2.26	1.73	1.31
2000	2.67	1.83	1.45
2005	2.62	1.81	1.44

Using the above data, the parameter values chosen above and our relative wage equation

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

we recover the relative η's.[24]

5 Results

We begin by presenting the skill-specific productivity levels, which we have computed and discussing how they differ across countries and what they imply about the evolution of the factor bias of technology during our sample period.[25] Next we present and discuss the measures

[23]We update the data to 2010 using Acemoglu and Autor (2011)

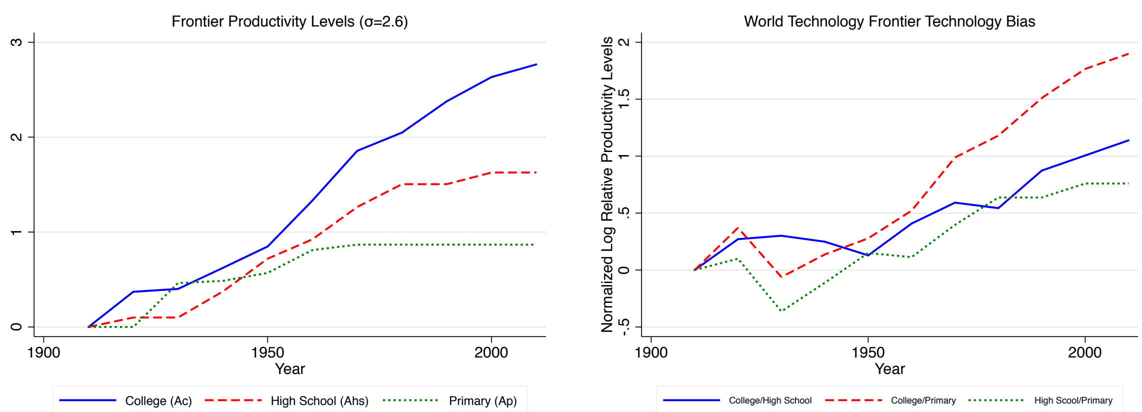
[24]This is a different approach than the one taken by Gancia et al. (2013). They arrive at an equation equivalent to the above formula (with the exception that they treat the U.S. as the frontier country, which by definition does not benefit from diffusion so $\phi = 0$). They use observations on the skill premium and 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 the 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 (either 2.6 or 1.6) 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 are allowed to change over time.

[25]Recall that, in the two-skill version of the model, technological progress is said to be skilled-labor augmenting if it increases A_H and skill-biased if it increases A_H/A_L. In the empirical analysis we have three skill groups, so we obviously have more possible combinations of effects. To avoid confusion we will adopt the following terminology. We will say that technological progress, for example, favors college educated workers relative to high school educated worker when A_C/A_{HS} goes up, without any restriction on the concurrent change in A_C/A_P. We will only say that technological progress is college-biased, if it favors college educated

of barriers to technology adoption implied by our model. Finally, we evaluate the role of barriers and technology-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

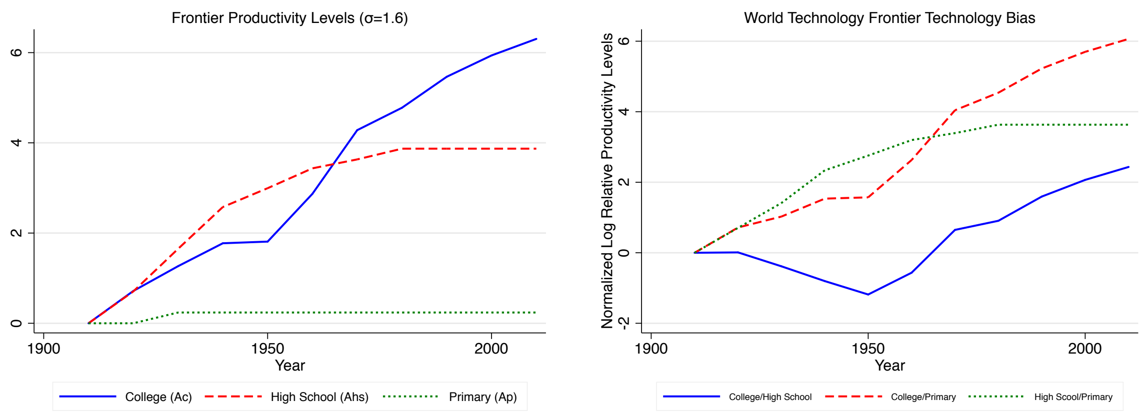
How has the skill-bias of the technology frontier evolved over the last century? Figures 1 and 2 plot the time-series of the world frontier productivity levels for the case of $\sigma = 2.6$ and $\sigma = 1.6$, respectively. Panels 1(a) and 2(a) show the evolution of the (normalized) frontier productivity levels for each of the three skill groups. Panels 1(b) and 2(b) translate these numbers into measures of relative skill bias, that is they report (log) of A_i/A_j where i and j each denote one of the three skill groups. The first lesson from examining these graphs is that the value of elasticity of substitution matters for conclusions about the level of world technology frontier and its evolution. Under the higher elasticity case ($\sigma = 2.6$; Figure 1), the frontier productivity for all skill levels rises until the mid-century, at which point the primary-specific productivity begins to stagnate. With lower elasticity ($\sigma = 1.6$; Figure 2), the primary productivity hardly changes at all during the entire sample period. In contrast, while the productivity of college-educated workers experiences the greatest increase under both elasticity scenarios, the magnitude of the increase is much higher in the low elasticity case (average 6% per year increase vs. average 2.6% per year).



(a) (b)

Figure 1: The (log) of productivity levels (A's) and the skill bias for $\sigma = 2.6$ from 1910 to 2010.

groups relative to both remaining skill groups, that is both A_C/A_{HS} and A_C/A_P increase.



(a) (b)

Figure 2: The (log) of productivity levels (A's) and the skill bias for $\sigma = 1.6$ from 1910 to 2010.

Recall that, in the two-skill version of the model, technological progress is said to be skilled-labor augmenting if it increases A_H and skill-biased if it increases A_H/A_L.[26] Caselli and Coleman (2006) who translate the above statements about direction of technology bias over time into statements of the bias across countries by replacing time with income per capita. 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.[27] To investigate whether there is 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. Tables 3 and 2 report the results, which are similar for both values of σ. Generally, these results imply that richer countries operate more productive technology at all skill levels. Figures 3 and 4 show the fit of the above regressions for the college and primary groups in 2010 data. 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).[28] Although there does seem to be a relative bias, in sense that poorer countries use lower-skilled types of labor relatively more

[26]Strictly speaking, the latter statement is true as long as $\sigma > 1$, which is both empirically relevant and assumed though out our empirical analysis.

[27]As described earlier, they find strong evidence of the former and slightly weaker evidence of the latter in their data set.

[28]The closest we come to it, is in the case of primary-specific productivity in 2010 under the low elasticity assumption. There there does not appear to be any significant relationship between income and productivity level.

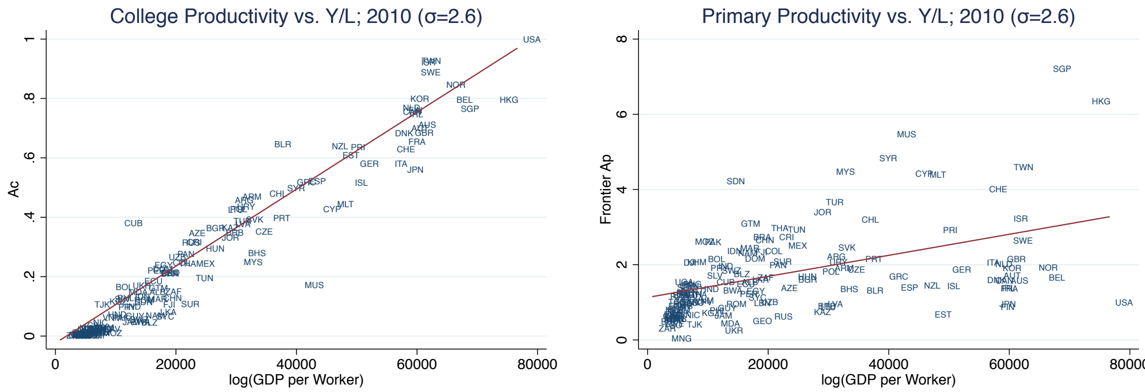
efficiently (i.e. A_{HS}/A_C and A_P/A_C are decreasing in income). The caveat here is that the productivity levels used in the above analysis are the actual levels computed for each country, that is they are determined both by the country's skill endowments, via the directed technology channel, and by the level of barriers to entry. So our findings could simply arise due to the richer countries having systematically lower barriers – a fact we indeed confirm in Section 5.4. If not for the distortion due to barriers, would rich countries still operate better technologies for low skilled workers than the poor countries, where such workers are much more abundant? Later we conduct a counterfactual exercise of removing the barriers to entry and recomputing the productivity levels. We return to this question at that point.

Table 2: $\sigma = 2.6$;

	OLS (2010)	OLS (All Yrs.)	FE (All Yrs.)
College (A_C)	1.434***	1.794***	1.048***
	(0.038)	(0.034)	(0.104)
High School (A_{HS})	0.938***	1.289***	0.938***
	(0.026)	(0.019)	(0.068)
Primary (A_P)	0.353***	0.570***	0.937***
	(0.043)	(0.019)	(0.048)
N	146	1769	1769

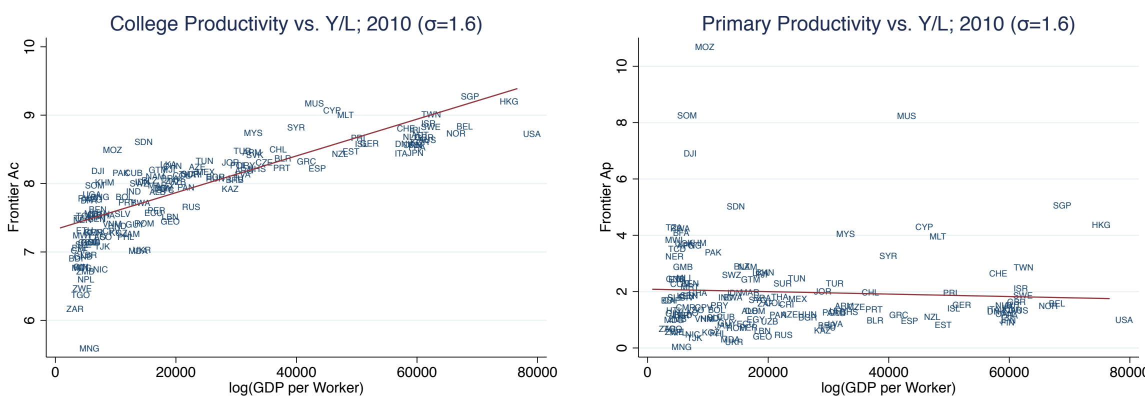
Table 3: Skill Bias of Technology Across Countries ($\sigma = 1.6$)

	OLS (2010)	OLS (All Yrs.)	FE (All Yrs.)
College (A_C)	0.510***	0.753***	0.954***
	(0.026)	(0.017)	(0.064)
High School (A_{HS})	0.272***	0.511***	0.901***
	(0.040)	(0.017)	(0.054)
Primary (A_P)	−0.009	0.166***	0.901***
	(0.053)	(0.024)	(0.053)
N	146	1769	1769



(a) College Productivity (b) Primary productivity

Figure 3: Skill-specific productivity levels versus 2010 GDP per capita; $\sigma = 2.6$.



(a) College Productivity (b) Primary productivity

Figure 4: Skill-specific productivity levels versus 2010 GDP per capita; $\sigma = 1.6$.

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 greatly 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 the traditional factor-neutral TFP measures.[29]

Levels

Figure 5 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 productivity levels lower than that of the US and the dispersion is fairly large. Table 4, with the corresponding summary statistics, reports that median TFP is only 35% of the US level and the coefficient of variation is 77%.[30] These patterns are not uniformly shared by the distributions of relative skill-specific productivity we have computed. When using the higher elasticity of substitution, $\sigma = 2.6$, we find that the U.S. has the highest productivity of college-educated workers (A_C) in the world in every decade since 1910. Comparing across countries, college-specific productivity has an even greater dispersion (coefficient of variation of 104%) and lower median (15% of the US level) than TFP, the productivity of primary-educated workers' is much less widely dispersed (coeff. of var. = 70%) but – 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

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

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

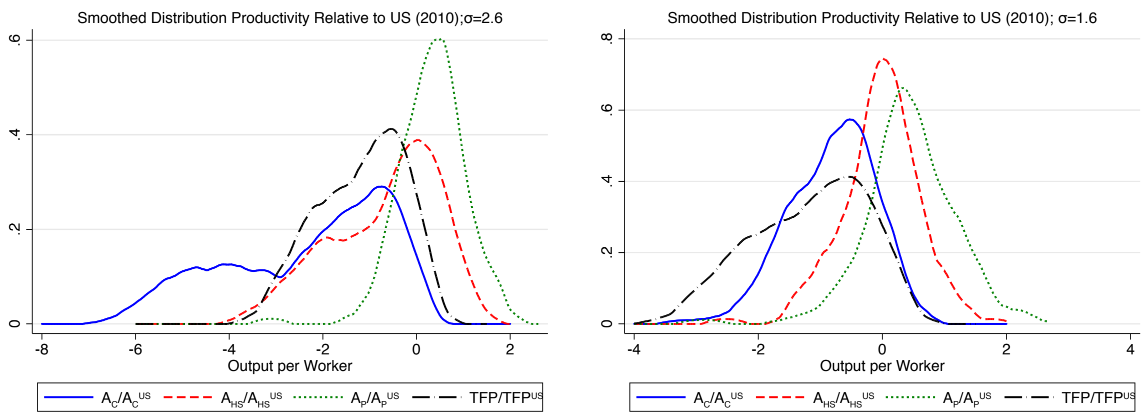
with human capital is given by

$$log(h_i) = \phi_P P_i + \phi_S S_i + \phi_T T_i.$$

where P_{it}, S_{it} and T_{it} stand for years of primary schooling, years of secondary schooling and years of tertiary schooling, respectively and we assume $\phi_P = \phi_S = \phi_T = 0.10$.

[30]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.

former communist countries, among them the Czech Republic, Poland, and Latvia. In the case of lower elasticity, the distribution of college-specific productivity is quite different from that under higher elasticity. First, the gap between the median and the US is much smaller, with half of the countries at 49% of the US level. Moreover, the US is only a technological leader in this category in 1920 and 1950. The variation is also much smaller. The primary-specific and secondary-specific productivity distributions are similar to the higher elasticity case.



(a) $\sigma = 2.6$

(b) $\sigma = 1.6$

Figure 5: Smoothed distribution of the (log) productivity levels relative to the US (2010).

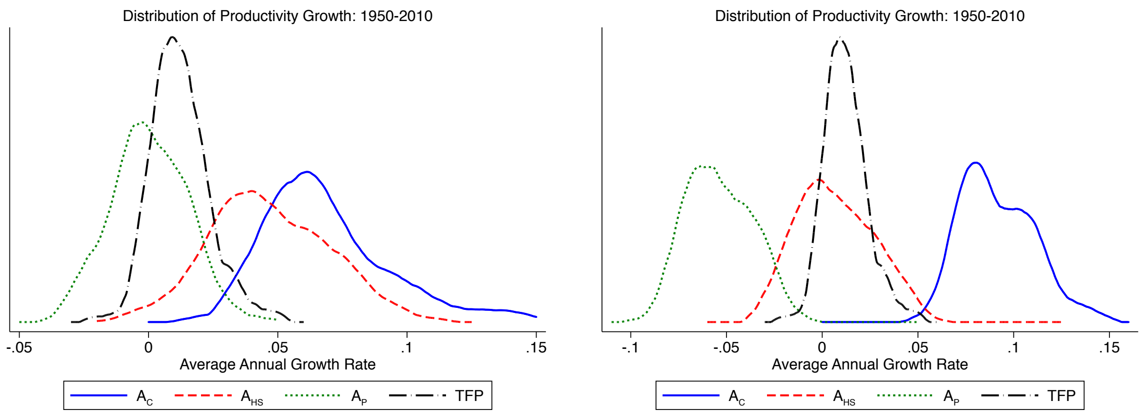
		$\sigma = 2.6$			$\sigma = 1.6$		
	$\frac{TFP}{TFP_{US}}$	$\frac{A_C}{A_C^{US}}$	$\frac{A_{HS}}{A_{HS}^{US}}$	$\frac{A_P}{A_P^{US}}$	$\frac{A_C}{A_C^{US}}$	$\frac{A_{HS}}{A_{HS}^{US}}$	$\frac{A_P}{A_P^{US}}$
Median	0.35	0.15	0.69	1.44	0.49	0.99	1.53
Mean	0.43	0.26	0.83	1.74	0.54	1.16	2.00
Coeff. Of Variation	0.77	1.04	0.84	0.70	0.62	0.68	0.78
90^{th} %tile	0.94	0.70	1.83	3.20	1.00	2.02	3.84
10^{th} %tile	0.08	0.01	0.09	0.62	0.18	0.45	0.69

Table 4: Moments of the 2010 distribution of relative productivity levels, data and various counterfactuals when barriers to innovation are removed. Capital is held constant when barriers are removed.

Growth Rates

Figure 6 and Table 5 summarize the average annual growth rates of our skill-specific productivity measurers 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 in fact declined. In our sample, 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. 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. With skill-specific productivity, our computed growth rates of TFP display somewhat different patterns. Growth of college productivity was the highest, reflecting the fact that the world frontier was biased towards this skill group in the sample period. Under the assumption of $\sigma = 2.6$, 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 secondary productivity growth rates was similar, but shifted to the left relative to college. Finally, primary productivity grew much slower on average, with roughly half of the countries registering declines in this measure. This is because for many countries, they had a rapid reduction – to nearly zero – of the share of workers with only primary level of education. With lowered elasticity, these distributions look different. College productivity grew much faster on average (reflecting the faster growth of college productivity frontier under this assumption, as already discussed). However, the secondary-specific and primary-specific productivity levels grew much slower under this scenario, with the median country experiencing an implausible 5.6% average annual rate of decline in the latter case. The implausibility of the low elasticity scenario is seen even more clearly in Figure 6(b), which shows that under this scenario nearly all of the countries experienced a decline in primary-specific productivity.

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 finding indicate that, especially under the higher elasticity assumption, the slow or negative TFP growth is not as puzzling. In sectors where world technology was in fact growing, that is college and to a lesser degree secondary-specific, even poorer countries recored sizable productivity improvements. However, since these countries have relatively low proportions of their labor



(a) $\sigma = 2.6$ (b) $\sigma = 1.6$

Figure 6: Distribution of the average annual growth rates of skill-specific productivity; 1950-2010.

force in those sectors, the impact on output was small, leading to small or negative numbers when factor-neutral TFP growth was computed.

		$\sigma = 2.6$			$\sigma = 1.6$		
	TFP	A_C	A_{HS}	A_P	A_C	A_{HS}	A_P
Median	1.05	6.40	4.35	-0.08	8.82	0.48	-5.56
Mean	1.22	6.86	4.66	0.03	9.16	0.70	-5.37
Coeff. Of Variation	1.10	2.41	2.24	1.44	1.88	1.94	1.69
90^{th} %tile	2.74	10.15	7.60	1.78	11.50	3.56	-3.12
10^{th} %tile	-0.05	4.12	1.90	-1.92	6.97	-1.89	-7.44
Frontier Growth	0.98	3.20	1.50	0.50	7.49	1.46	0.00

Table 5: Summary statistics of annual growth rate of factor-neutral TFP and skill-specific productivity measurers during the period 1950-2010. The bottom row provides the average growth rate of the world technology frontier (in case of TFP, it is defined as US TFP.

5.3 Barriers

In this section we discuss the measures of barriers to entry we have calculated. Recall that we cannot identify the absolute level of barriers but instead compute their value relative to that of the U.S. using equation (39). We will focus our analysis on the level of barriers relative to GDP per worker or ζ/y in the notation of our model.. This is a natural choice if we want to compare the burden of barriers across countries. Below we often refer to our object of analysis as "barriers to entry" for short but the reader should keep in mind that they actually represent a ratio of the cost of entry to GDP per worker, relative to that ratio in the U.S.

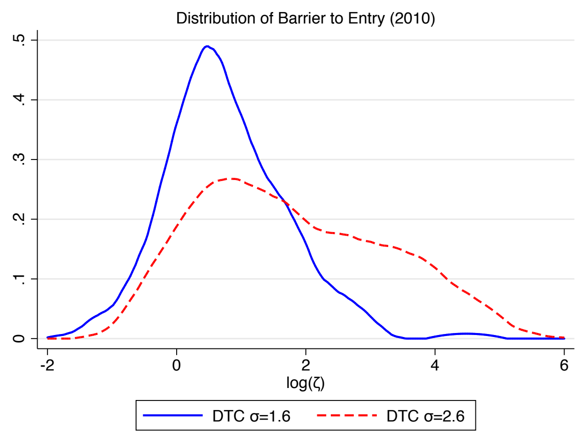


Figure 7: Smoothed distribution of the (log) barriers to entry in 2010.

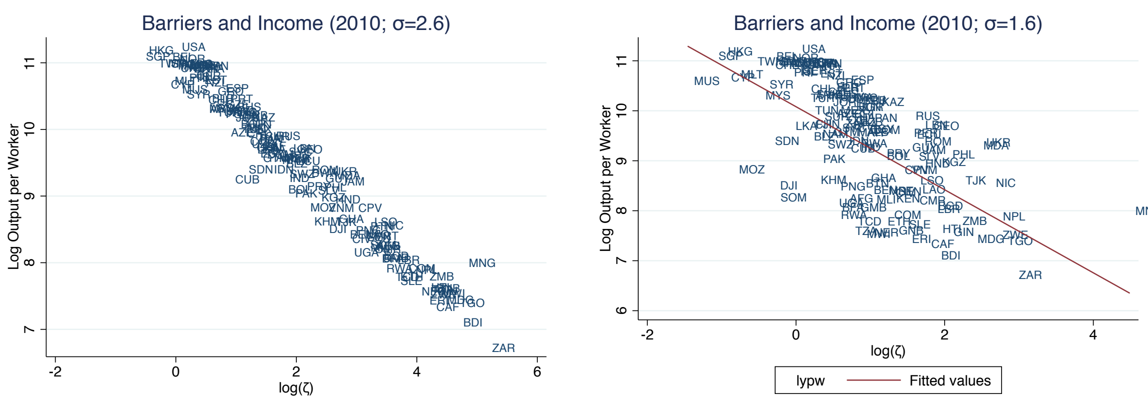
Figure 7 shows the smoothed distributions of the log of barriers to entry in 2010 for the two values of σ we consider. For both values of elasticity there is a lot of variation and large number of economies with barriers significantly higher than in the U.S.[31] However, the distribution under low elasticity is both tighter and, more importantly, has a lot more mass to the left of zero (i.e. the U.S. value). Specifically, out 146 countries, the U.S. ranks only the 26^{th} in terms of barriers under $\sigma = 1.6$.[32] Moreover, while some of the countries found to have lower costs of entry 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 entry barriers at only about 32% of those in the U.S. seem unrealistically low. The distribution for the case of $\sigma = 2.6$ seems more plausible. Not only is the U.S. ranked 11^{th}, but the countries we find to have lower barriers include only developed economies such as Singapore, Hong Kong, Taiwan, Ireland,

[31] Since barriers are relative to the U.S., a negative value implies barriers less than in the U.S.

[32] In the Appendix we list all countries and their 2010 barrier levels.[ADD THIS APPENDIX]

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-barrier group between 80% and 96%.

Are our measures of barriers systematically related to the countries' level of development? Figures 8(a) and 8(b) plot the relationship between barriers and log GDP per worker. In both cases the relationship is strongly negative, implying rich countries have systematically lower barriers. However, the relationship is less tight under $\sigma = 1.6$. In fact, a closer inspection reveals that richer countries (log GDP per worker greater than $22,000) the relationship is much closer than for the group of lower income countries (Figures 9(a) and 9(b)).

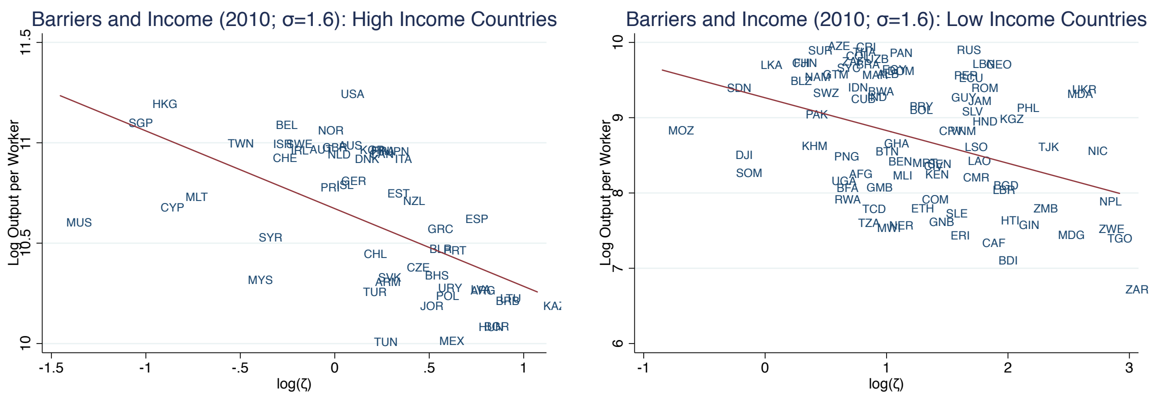


(a) $\sigma = 2.6$ (b) $\sigma = 1.6$

Figure 8: Relationship between log GDP per worker and the (log) level of barriers to technology adoption in 2010.

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 barriers to entry fall over time for many, if not the majority, of world's countries. How did our measured barriers evolve over the sample period? Figure 10 plots the smooth distributions of average annual change in barriers during the period 1950-2010 under the two elasticity assumptions. We report the moments of the two distributions in Table 6. In both cases, we find that barrier declined steadily for most countries with median rate of change of about -1% and -1.7% per year, for the high and low elasticity cases respectively.[33] In both cases a minority of countries

[33] These are relative to the US. It is possible, although perhaps not plausible that the U.S. increased absolute barriers to technology adoption at a faster rate than the measured decline mentioned above. We consider this unlikely as through most of this period tariffs on manufactured goods declined greatly after the end of World War II. It is the case that there has been an increase in regulations, environmental, occupational safety, etc., which could be a rise in barriers. However if the air and water are cleaner, and work places are increasingly safe and were properly valued in addition to GDP, (see Murphy and Topel (2006) and Jones and Klenow (2016)) ,



(a) $\sigma = 1.6$; High income countries.

(b) $\sigma = 1.6$; Low income countries.

Figure 9: Relationship between log GDP per worker and the (log) level of barriers to technology adoption in 2010 for the case of low elasticity($\sigma = 1.6$); high vs. low income countires.

experienced an increase in barriers. Of course, these 60-year averages mask many interesting patterns including sharp changes in barriers, reversals, etc. Figures 11(a) and 11(b) show examples of time paths of barriers for two countries, China and Zimbabwe, the first of which was one of the leaders in barrier reduction during this period while the other experienced one of the largest increases. Notable, of course, is the uninterrupted decline in Chinese barriers which began around 1980 and the rapid increase in costs of entry in Zimbabwe, roughly coinciding with the Mugabe regime.

	$\sigma = 2.6$	$\sigma = 1.6$
Median	-1.00	-1.70
Mean	-1.30	-1.73
Std. Dev.	1.69	1.93
90^{th} %tile	0.68	0.59
10^{th} %tile	-3.48	-4.16

Table 6: Moments of the average annual change of barriers ($\Delta \log(\zeta)$); 1950-2010.

To get a better sense of how the entire distribution of barriers evolved, we plot – in Figures 12(a) and 12(b) – these distributions for selected dates in our sample. Over time, for both values of σ, we observe a similar evolution of barriers in the sense that between 1950 and 2010 the distributions shift to the left indicating declining barriers in most countries. However, for the case of $\sigma = 1.6$ this tendency is much more strongly manifested by the movement

then even with these increasing regulations barriers to technology adoption could have fallen.

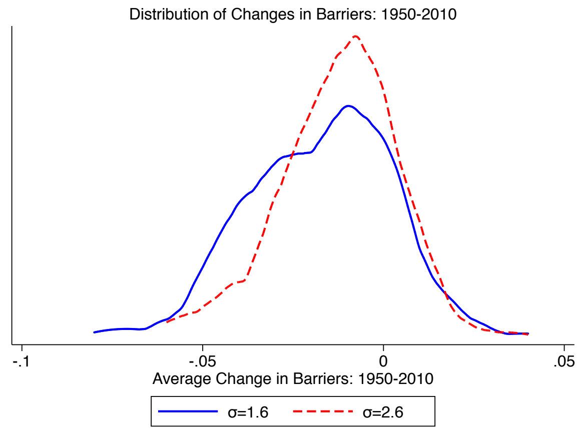
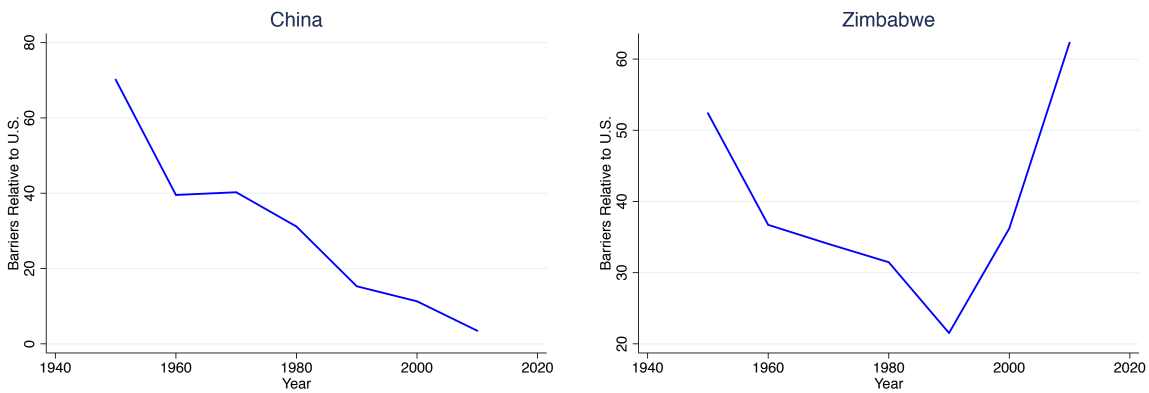


Figure 10: Smoothed distribution of the change in (log) barriers to technology adoption 1950-2010.



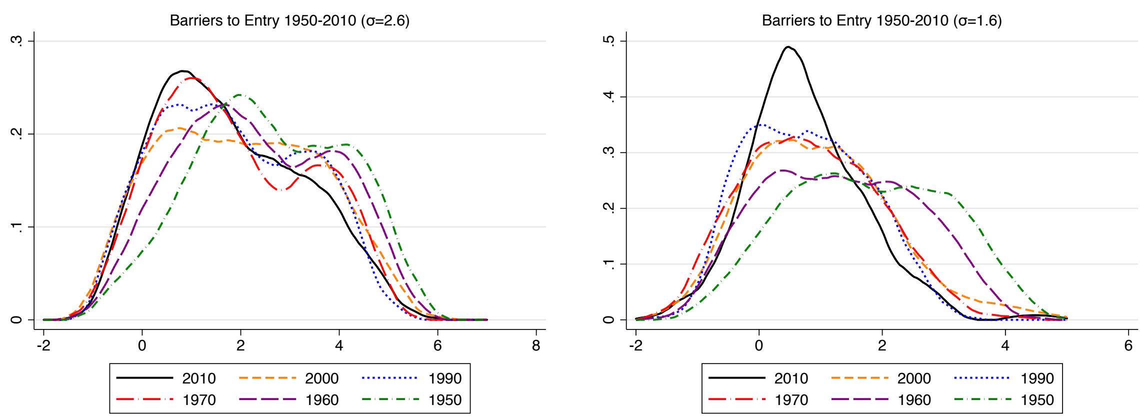
(a) Barriers to entry in China.

(b) Barriers to entry in Zimbabwe.

Figure 11: Evolution of barriers to entry; China and Zimbabwe.

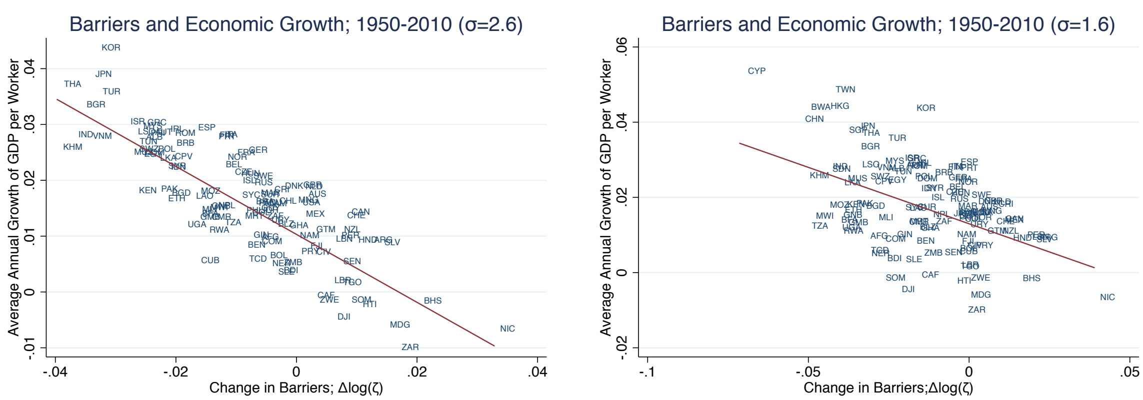
of the right tail of the distribution, especially after 1960. This implies very large reductions in barriers in many countries that started out the sample period with a relatively high level of them. This happens also under for $\sigma = 2.6$ but is less pronounced. Also of note is the fact that until 2010 the distribution of barriers (under the high elasticity case) seems to be bi-modal. This is reminiscent of the well-know twin-peaks finding in the literature on the evolution of the world income distribution (Quah 1996, Feyrer 2008). Finally, recall that using a similar theoretical approach, but a different interpretation of barriers and empirical approach to measuring them, Gancia et al. (2013) find that barriers increased significantly in non-OECD countries between 1970 and 2000. While our findings suggest a different general trend; barriers have been falling over time. For the same time period, 1970 to 2000, (under high elasticity, which is close to the value used by Gancia et al.), we also find a slight increase in barriers.

Finally, in Figure 13 we plot the change in barriers versus growth of GDP per capita during the same period. For both values of the elasticity parameter, the correlation is negative suggesting that, consistent with our priors and the model, reduction of barriers contributes to growth. However, as was the case with levels, the relationship is much closer under the higher value of elasticity, especially after we remove very fast growing countries. To illustrate this, in Figure 14 we plot the relationship for only those economies that grew at less than 3% per year. Clearly, under $\sigma = 1.6$, growth and barrier reduction are much more weakly related.



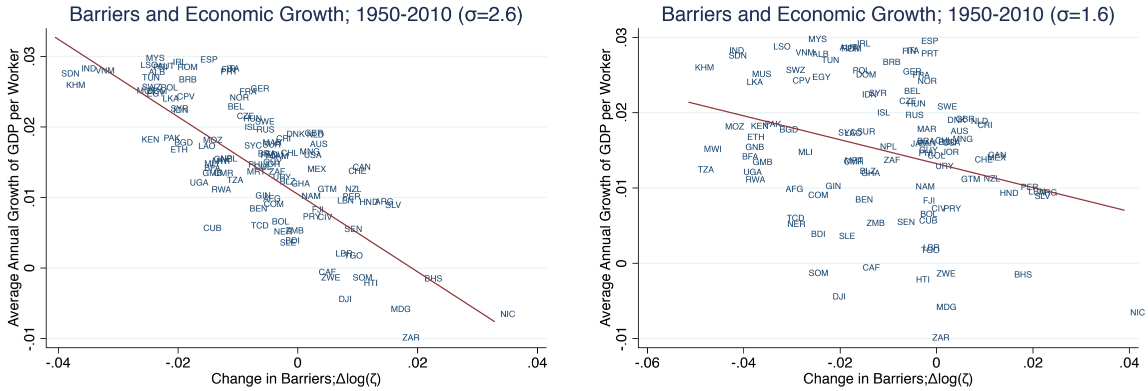
(a) $\sigma = 2.6$ (b) $\sigma = 1.6$

Figure 12: Time series of log relative barriers per worker.



(a) $\sigma = 2.6$ (b) $\sigma = 1.6$

Figure 13: Average annual growth of GDP per worker vs. change in barriers to technology; 1950-2010.



(a) $\sigma = 2.6$

(b) $\sigma = 1.6$

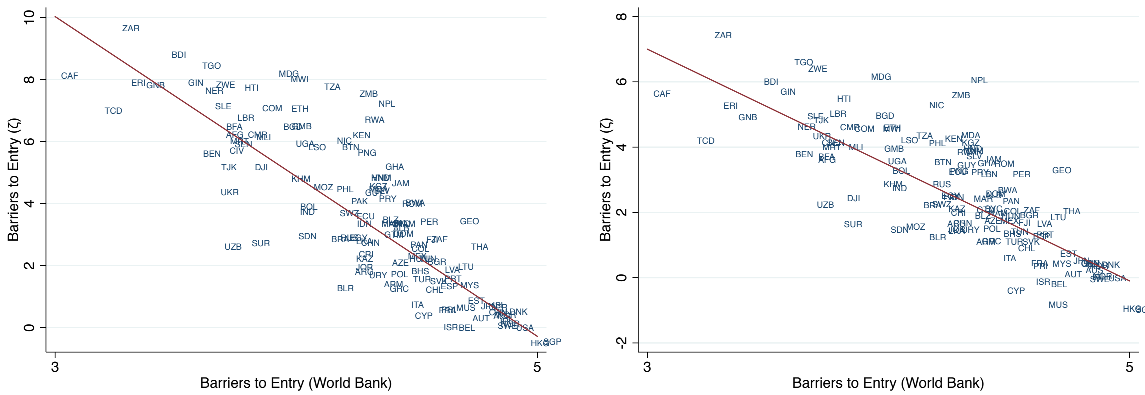
Figure 14: Average annual growth of GDP per worker vs. change in barriers to technology, excluding very fast growing countries; 1950-2010.

Before proceeding we consider two ways to check the plausibility of our barrier measures. First we correlate their 2010 values with the World Bank *Doing Business: Distance to Frontier*, 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 15 with higher score on the Doing Business scale indicating more entrepreneur-friendly environment. Our measures of barriers, under both elasticity specifications, 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 high barrier's are usually also deemed unfriendly to business by the World Bank.

The second test of our barrier measures uses tariff data over the 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 barriers.[34] This is an imperfect test since barriers encompass much more that 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

[34]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.

problem. We regress our measures of barriers to entry on the average tariff rate relative to the U.S. rate, since our barriers are also measured in this manner. The results are shown in Table 7 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 barriers. However, this result is robust to inclusion of country fixed effects only for the higher elasticity case. Moreover, in the latter case the relationship is stronger, as reflected by the higher partial correlation coefficients. We conclude that the measures of barriers we have identified correspond well to other measures available in the literature, both more recent and in more historical data, and that this relationship is stronger for the higher elasticity case of $\sigma = 2.6$.



(a) $\sigma = 2.6$ (b) $\sigma = 1.6$

Figure 15: World Bank Doing Business indicator vs. our measures of barriers (2010).

Table 7: Relative Tariffs vs. Barriers; 1880-1990.

	$\sigma = 2.6$		$\sigma = 1.6$	
	OLS	FE	OLS	FE
Relative Tariffs	1.404***	0.807**	0.839*	0.262
	(0.461)	(0.394)	(0.435)	(0.351)
Trend	−0.036***	−0.029***	−0.032***	−0.025***
	(0.008)	(0.006)	(0.007)	(0.005)
Constant	70.450***	57.925***	62.458***	50.349***
	(14.539)	(10.967)	(13.723)	(9.765)
R^2	0.123	0.125	0.119	0.186
N	144	144	144	144

5.4 Counterfactuals

What do our results – the skill-specific productivity levels and measures of barriers that we have computed – imply for our understanding of cross-country income differences? An influential strand of the economic growth literature emphasizes the importance of barriers to entrepreneurship and entry as key determinants of productivity and therefore standards of living in the long run (Olson 1982, Parente and Prescott 1994, 1999). Under this theory, poor countries' inadequate factor endowments (low physical or human capital levels) or lack of technologies suitable for those factor endowments are not the main forces holding these countries from achieving economic prosperity. If only barriers that protect monopoly incumbents from the competition and innovation of new entrants were removed, rapid economic growth would follow. In our model output levels differ across countries because of endowments of skill-types and physical capital, the available world frontier technologies for each skill type, and the incentives to adopt the frontier technology, themselves determined by skill endowments (via the directed technology channel) and barriers. Therefore it is natural to ask what our approach and its empirical implementation imply about the importance of barriers for long run development. We address this by computing the world income distribution under two counterfactual experiments. The first one is constructed after endowing countries with counterfactual productivity levels, which arise from reducing barriers to technology adoption, ζ, for all countries. Since we do not actually compute the absolute level of barriers, but only their relative levels, we must use a reference point. We choose the U.S., since it is generally recognized as a relatively low-barrier economy with robust innovation and high productivity levels. To be specific, we construct the counterfactuals by setting barriers so that, after we recompute the new output per worker, the barriers as a proportion of per worker output are the same as the U.S. barriers per worker output. We leave the barriers unchanged for countries with lower barriers-to-GDP per worker ratio than in the U.S., which happens in a handful of cases for $\sigma = 2.6$ and and many more for $\sigma = 1.6$.[35] 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 barriers to the U.S. level, does not necessarily imply productivity levels will be equalized because they depend both on barriers and on factor supplies.[36] We construct the second counterfactual by keeping barriers unchanged (as a % of

[35]Throughout the calculations, we keep the level of frontier technologies constant. We do this mainly because we do not want to take a stance on how the world technology frontier evolves over time, which would be required to construct the counterfactual values of the frontier.

[36]This is an important difference between our approach and that of Gancia et al. as already mentioned in section 3, footnote 15. In their approach eliminating barriers amounts to making ϕ go to infinity and make

GDP per worker) but instead adjusting each country's skill distribution to match that of the U.S. and recomputing the implied levels of skill-specific technology. In both cases we keep the capital-output ratio unchanged, which corresponds to keeping all investment distortion (τ in equation 24) unchanged. With the counterfactual productivity levels in hand, we re-visit the question of cross-country skill bias, that is we ask whether now that we have removed barriers to entry, do poor countries have an absolute advantage in lower-skill technologies.[37] We then discuss what the removal of barriers does for international productivity and income differences. Finally, we conduct a development accounting exercises to see how its results differ from the factor-neutral counterpart. We continue to present results under both low and high elasticity scenarios, however, we note that the results so far favors the higher value. The estimates of productivity levels and growth rates as well as barriers to entry and their correlations with alternative measures of barriers, economic growth and income levels were all more plausible for $\sigma = 2.6$, making it our preferred value.

Productivity Differences in Absence of Barriers

The endogenous and directed aspect of technological progress in our model suggest 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 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 skill-types. We have just seen in Section 5.6, that poor countries have, on average, considerably higher barriers. How does removing those barriers affect productivity differences between rich and poor countries?

The 2010 relationships for college- and primary-specific productivity are plotted in Figures 16(a) and 16(b).

Tables 8 and 9 report regressions of counterfactual productivity levels on GDP per worker for the two elasticity values. Under the higher elasticity (Table 8), college productivity (A_C) is again higher is rich countries. The secondary-specific productivity (A_{HS}) 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

every economy's convergence to the frontier instantaneous and full.

[37]Recall that in Section 5.1, we reported that actual technologies used (with barriers) were superior for all skill levels in richer countries.

(a) College Productivity (b) Primary productivity

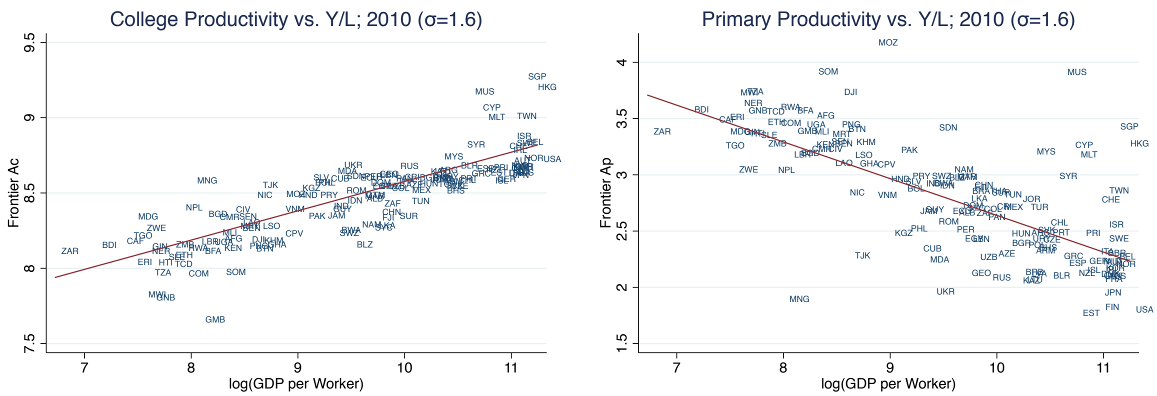
Figure 16: College and Primary specific technologies (given skill endowments) versus 2010 GDP per capita. Computed by setting barriers to the U.S. level.

secondary-specific productivity relative to the college-specific one (i.e. A_{HS}/A_C is decreasing with income). However, for primary-specific productivity (A_P) we now find an indication of absolute bias in the sense that poorer countries would (in the absence of barriers) enjoy a higher absolute primary-specific productivity level than their richer counterparts.

Table 8: Skill Bias of Technology Across Countries ($\sigma = 2.6$)

	OLS (2010)	OLS (All Yrs.)	FE (All Yrs.)
College (A_C)	0.645***	0.935***	0.057
	(0.039)	(0.034)	(0.097)
High School (A_{HS})	0.148***	0.430***	−0.053
	(0.022)	(0.018)	(0.058)
Primary (A_P)	−0.437***	−0.289***	−0.054
	(0.034)	(0.014)	(0.034)
N	146	1764	1764

In the lower elasticity case, we find similar results for the college-specific productivity, namely that it increases with income. There is an even stronger indication of absolute bias in both secondary- and primary-specific productivity levels, which – at least in the most recent data (column one of Table 9) – are decreasing with income. When using all years to estimate the relationship, it is no longer negative for secondary-specific productivity level but remains so for the primary one. We conclude that according to our data and the model imposed structure, that, in the absence of barriers to technology adoption, poor countries would actually be able



(a) College Productivity

(b) Primary productivity

Figure 17: College and Primary specific technologies (given skill endowments) versus 2010 GDP per capita. Computed by setting barriers to the U.S. level.

Table 9: Skill Bias of Technology Across Countries ($\sigma = 1.6$)

	OLS (2010)	OLS (All Yrs.)	FE (All Yrs.)
College (A_C)	0.194***	0.361***	0.137***
	(0.012)	(0.014)	(0.050)
High School (A_{HS})	−0.044**	0.119***	0.085**
	(0.018)	(0.009)	(0.035)
Primary (A_P)	−0.325***	−0.225***	0.084***
	(0.028)	(0.012)	(0.029)
N	146	1764	1764

to operate their low skill labor at greater 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 barriers would produce a big increase in their standards of living.

Counterfactual Income Levels

What is the impact of barriers to technology adoption on the world income distribution? Figure 18 shows the smoothed distributions of GDP per worker in 2010. The solid line is that data, the other two lines correspond to counterfactuals from the directed technology model with elasticity of substitution of 2.6 (dotted line) and 1.6 (dashed line). The right panel contains the smoothed distributions of GDP per worker in 2010, the directed technology model with $\sigma = 2.6$ and neutral TFP (where all countries receive the US TFP value).

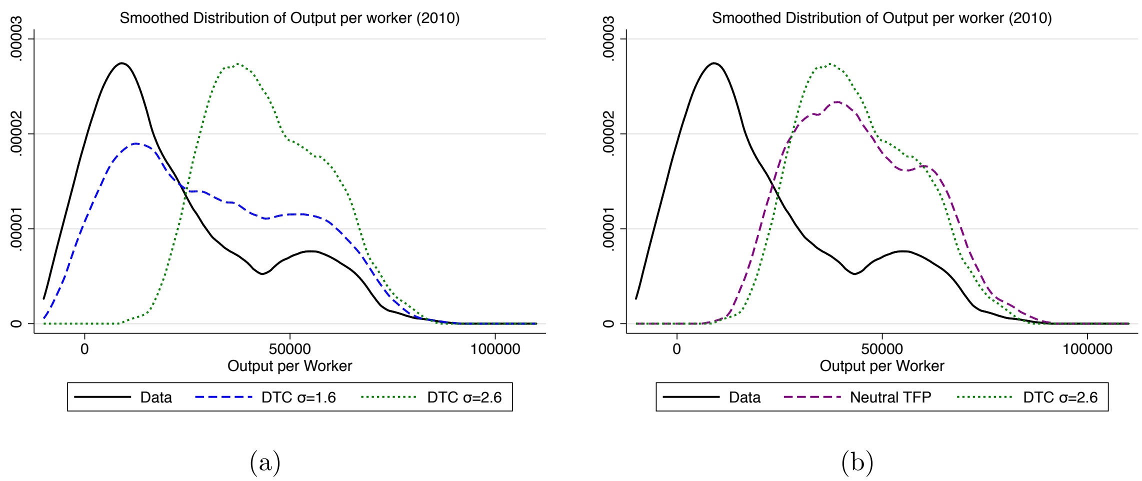


Figure 18: Smoothed distributions of GDP per worker in 2010.
Panel (a): Data, $\omega = 2.6$, $\omega = 1.6$, Panel (b): Data, $\omega = 2.6$, Neutral TFP.

Table 10 reports the summary statistics for the data (first column) and our three counterfactual distributions; the second and third columns for the two different values of σ, and the last column for the factor-neutral model. Table 11 summarizes the size of output gains due to reduction of barriers in 2010, defined as the difference between the counterfactual and actual GDP per worker, as a share of the latter.

	Data	$\sigma = 2.6$	$\sigma = 1.6$	Neutral
Median	14,592	42,481	28,420	43,418
Mean	21,890	44,479	30,709	44,791
Coeff. Of Variation	0.92	0.30	0.67	0.33
90/10	23.6	2.2	10.2	2.5

Table 10: Moments of the 2010 world distribution of output per worker, data and various counterfactuals when barriers to innovation are removed.

Note first that the 2010 data clearly show the high degree of income disparities and the familiar bi-modal nature of the distribution emphasized by Quah (1996), Tamura (1996) and others. 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 barriers to innovation shifts a significant portion of the distribution to the right but the effect is much stronger when elasticity of substitution is assumed to be 2.6. This follows from the finding discussed above; with low elasticity of substitution, we find that barriers in many low-income countries are actually lower than in the U.S. (Figure 8(b)). With $\sigma = 1.6$, median

	$\sigma = 2.6$	$\sigma = 1.6$	Neutral
Mean	4.23	1.10	4.04
Max	29.96	18.88	28.00
90th percentile	13.20	2.51	11.90
75th percentile	6.06	1.38	6.10
Median	1.83	0.57	1.87
Min	0.00	0.00	0.00

Table 11: Moments of the 2010 gain in GDP per capita when barriers to innovation are removed.

income per worker just about doubles and the 90/10 ratio falls to 10.2. In contrast, with $\sigma = 2.6$, median income almost triples while the 90/10 gap falls to 2.2. The two resulting distributions of GDP per worker are displayed in the left-hand panel of Figure 18(a). The right-hand panel compares the higher elasticity case with the traditional factor-neutral model, where the counterfactuals are computed by increasing every country TFP to the maximum of it own level or that of the U.S. As can be seen from the plot and Table 10, the resulting counterfactual world income distributions are quite similar. In both cases, removing barriers to entry in our directed technological progress model and moving low-TFP countries up to the U.S. productivity level in the factor-neutral one, we see large rightward shift of the world income distribution. However, the underlying mechanisms are quite different. Obviously, in the factor-neutral setting, low-productivity countries see their output rise because of a uniform increase in the catch-all TFP, which increases the productivity of all labor-skill types equally. In our directed technology model, the transformation is different. Figures 19-21 illustrate the change in productivity levels that result from barrier reduction for each of the skill-types by plotting the actual level versus the counterfactual one (left-hand panels refer to the high elasticity of substitution case while the right-hand panels are for the lower elasticity). Note first that after removing barriers in the high elasticity environment, countries with low college-specific productivity (mostly poor countries with low shares of skilled labor) experience relatively small productivity gains in that sector; Figure 19, panel (a). This is because of the strong bias present in this case; these countries have low shares of college labor and thus a small market for machines compatible with high skills. As a consequence, even when barriers to entry are 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 barriers are reduced. This is a consequence of their higher shares of college educated workforce.[38] Continuing to focus on the high elasticity case, we see that the results for the high skilled sector contrast with those for the two lower skill sectors, especially the primary one. Here poor countries experience very large increases in productivity once barriers are removed. This is of course just the opposite of what happened in the college sector; they have relatively large shares of labor in these sectors and this, combined with the reduction in barriers to technology adoption, create strong incentives to innovate. Productivity gains from removal of barriers are much more modest under the lower elasticity scenario (left-hand panels). This is because skill endowments matter less for innovation and because barriers are identified to be much lower, which many countries already below the U.S. level.

Thus in our preferred scenario of $\sigma = 2.6$, we see that reducing barriers to entry has a powerful effect on the world distribution of standards of living, with an especially large shift in the lower part of the distribution as a consequence of large income gains in poor countries. Because of endogenous and directed innovation process, removal of barriers causes (most) poorer countries to adopt better technologies for the lower skill sector, where many of them actually end up surpassing the low-skilled productivity in richer economies. The high-skill productivity level does not increase in those countries by nearly as much. Therefore potential for growth from barrier removal occurs not by adopting the same set of technologies that are used in the rich countries but instead technologies more suited to the local factor endowments, (which presumably were used in the currently-rich countries at some point in the past). 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 suggesting that developing countries miss out on growth opportunities by de-industrializing prematurely and adopting the output mix of their more developed counterparts (Rodrik 2015).

Recall that Gancia et al. (2013) follow a similar approach to ours except that they interpret barriers as the speed of technology diffusion and estimate the structural equation of their model, instead of calibrating them as we have done, to compute barriers. They also are interested in the counterfactual of removing barriers to technology adoption. How do our results compare to theirs? In their preferred specification, they report that removal of barriers to entry 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 significantly higher than in our preferred specification ($\sigma = 2.6$). The equivalent changes in our counterfactual calculation are an increase from 21% to 53% for non-OECD countries and 71% to 82% for OECD economies. The fact that these increases are

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

smaller is at least in part due to the different way we model barriers. Gancia et al. (2013) equate barriers with the diffusion parameter ϕ and consider removal of barriers to be equivalent to this parameter going to infinity, which implies instantaneous convergence to the technology frontier and equalization of productivity levels across countries. In our setting, removal of barriers corresponds to making everyone's barrier as a share of GDP per worker the same as in the U.S. and does not imply convergence to the frontier.[39]

How do the income gains from reduction of barriers compare to what would happen if factor endowments were equalized across countries? To answer this question, we present the results of a counterfactual exercise analogous to the one above except that we now keep barriers (as a percentage of GDP per worker) constant and endow every country with the skill distribution of the U.S.[40] These are the counterfactuals that Caselli and Ciccone (2013) and Jones (2014) look at and we compare our findings to theirs. Summaries of the resulting income distributions and the sizes of income gains are reported in Tables 12 and 13. Notice that the results are reversed relative to those in Table 10: income gains under the high elasticity case ($\sigma = 2.6$) are considerably lower than those under low elasticity. This is intuitive since, as we have found above, barriers are on average lower and less important in determining the skill-specific productivity levels under the $\sigma = 1.6$ assumption. Additionally, in this case the low substitutability of skill-types makes the skill composition a more important determinant of relative output levels. Caselli and Ciccone (2013) use a non-parametric formulation to show that the output gain for the kind of counterfactual we have just considered, when assuming perfect substitution among skill-types (the standard approach) is – under certain assumptions – an upper bound on the role of human capital.[41] We do not assume perfect substitutability but our approach is also different from that of Caselli and Ciccone. Table 13 summarizes the percentage gains in GDP per worker from the skill composition change. Both the mean and the median gain was 55% under the high elasticity case while the mean was 325% and the median 144% with low elasticity. Caselli and Ciccone reported a mean of 61% and median of 45%, which is not very different from our preferred specification of $\sigma = 2.6$. However,

[39] The difference between their results and ours is even greater under the low elasticity of substitution scenario. Here our counterfactual barrier reduction leads to increases from 21% to only 33% for non-OECD countries and from 71% to 78% for the OECD group. As noted earlier, their value of elasticity of substitution between skill types is close to our preferred value.

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

[41] 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.

under the lower elasticity, human capital matters more than their bound indicates. Looking at individual countries, we find that for some our results are fairly close while for others they are not; Table 14 shows a selection of cases.

	Data	$\sigma = 2.6$	$\sigma = 1.6$
Median	14,541	22,367	40,075
Mean	21,697	29,626	46,481
Coeff. Of Variation	0.92	0.83	0.61
90/10	23.8	16.5	5.0

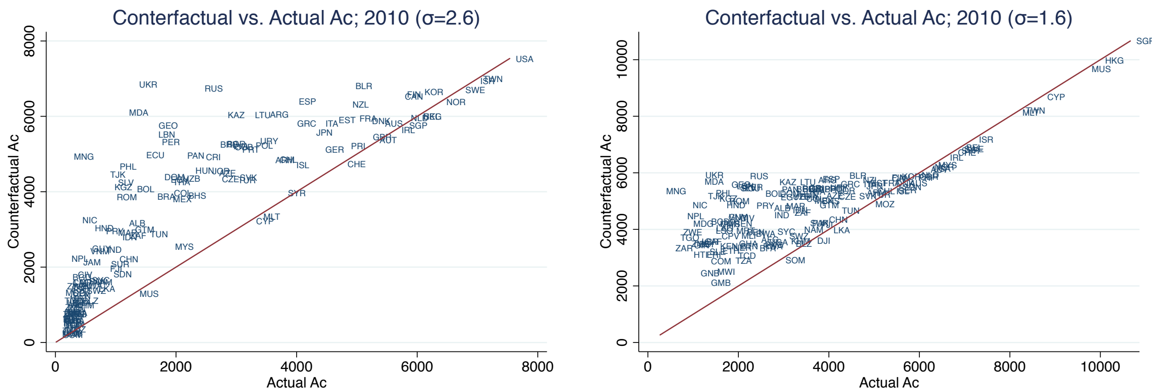
Table 12: Moments of the 2010 world distribution of output per worker, data and various counterfactuals when skill distribution is equating to that of the U.S.

	$\sigma = 2.6$	$\sigma = 1.6$
Mean	0.55	3.25
Max	0.96	18.32
90th percentile	0.90	9.05
75th percentile	0.81	4.27
Median	0.55	1.44
Min	0.09	0.11

Table 13: Moments of the 2010 gain in GDP per capita from equating skill distribution to that of the U.S.

Country	Year	CC Bound	$\sigma = 2.6$	$\sigma = 1.6$
Brazil	2000	0.90	0.56	1.57
South Africa	2000	0.71	0.48	1.24
Japan	1995	0.26	0.15	0.25
Taiwan	1995	0.33	0.31	0.70
Vietman	1995	0.41	0.45	2.30
India	1995	1.07	0.42	1.50

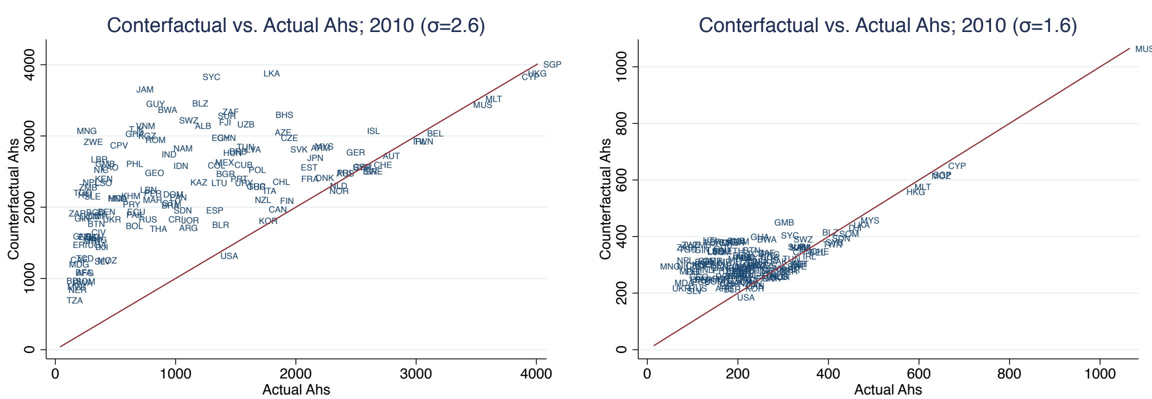
Table 14: The percentage of GDP per worker gained when skill distribution is changed to equal that of the U.S. Caselli and Ciccone (2013) (labeled CC bound) vs this paper.



(a) $\sigma = 2.6$.

(b) $\sigma = 1.6$.

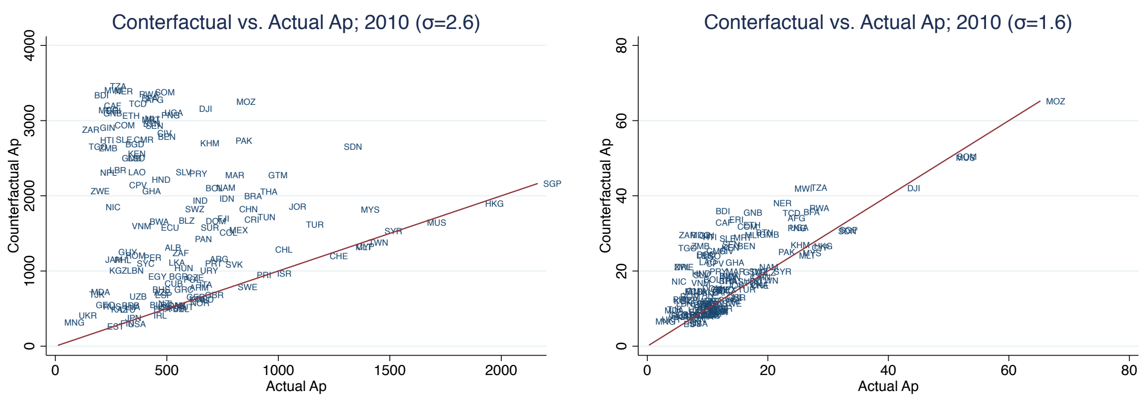
Figure 19: Counterfactual versus actual 2010 college-specific productivity level in 2010 (A_C). Counterfactual calculated by reducing the level of barriers to equal that in the U.S.



(a) $\sigma = 2.6$.

(b) $\sigma = 1.6$.

Figure 20: Counterfactual versus actual 2010 college-specific productivity level in 2010 (A_C). Counterfactual calculated by reducing the level of barriers to equal that in the U.S.



(a) $\sigma = 2.6$.

(b) $\sigma = 1.6$.

Figure 21: Counterfactual versus actual 2010 college-specific productivity level in 2010 (A_C). Counterfactual calculated by reducing the level of barriers to equal that in the U.S.

Development Accounting

To summarize our findings from the counterfactual exercise presented above, we conduct a development accounting exercise and compare its results to those found using a standard factor-neutral approach as well as those in Jones (2014). Development accounting assigns a measure to the importance of a certain characteristic – for example, physical capital –in explaining the variation in incomes (or the gap between rich and poor countries) by constructing a counterfactual distribution of income and comparing its variance with that of the actual distribution. In the counterfactual, countries keep their individual values of the characteristic in question (e.g. physical and human capital)but are otherwise made identical (e.g. by assuming they all have the TFP of the U.S.) The ratio of the variance of counterfactual incomes to that of actual ones (or the portion of the gap between rich and poor accounted for) is then interpreted as the percentage of variation in the cross-country income distribution accounted for by the characteristic that was left unchanged (physical capital). This is exactly the kind of counterfactuals we have computed earlier in this section. To see this note that in our model GDP per worker on the balanced growth path can be expressed, using equation (36), as

$$
\begin{aligned}
y &\propto \left(A_L^W\right)^{\frac{\phi}{1+\phi}} \left(\frac{K}{Y}\right)^{\frac{1-\beta}{\beta}} \left(\frac{\zeta}{y}\right)^{-\frac{1}{1+\phi}} \Omega\left(\frac{H}{L}, \frac{A_H^W}{A_L^W}\right)^{\frac{\phi}{(1+\phi)\beta}} \qquad (40)\\
&= \left(A_L^W\right)^{\frac{\phi}{1+\phi}} \kappa_i z_i \omega_i, \qquad (41)
\end{aligned}
$$

where the last three terms are country-specific and correspond to the contribution of physical capital (κ_i), barriers (z_i), and human capital endowment (ω_i).[42] The counterfactuals we have computed were $y_i^* \propto \kappa_i z_{US}\, \omega_i$ (when we reduced barriers) and $y_i^* \propto \kappa_i z_i \omega_{US}$ (when we changed labor force compositions). We can sue these to calculate the variance ratio as[43]

$$
\text{VR} = \frac{Var(\ln(y^*)) + \frac{1}{2} \times Cov(\ln y^*, \ln(\frac{y}{y^*}))}{Var(\ln(y))},
$$

which, for example, in the case of the barrier reduction exercise becomes

$$
\text{VR} = \frac{Var(\ln(\kappa_i\omega_i)) + \frac{1}{2} \times Cov(\ln(\kappa_i\omega_i), \ln(z_i))}{Var(\ln(y))},
$$

Figure 22 below shows the results of this variance ratio for all the countries in our sample

[42]See Appendix for details of the derivation.

[43]See Appendix for details

during the period 1950-2010, with left- and right-hand panels corresponding to $\sigma = 2.6$ and $\sigma = 1.6$, respectively. In each case, we plot for comparison the results of a standard development accounting, i.e. the variance ratio computed using the factor-neutral model. First note that in our sample, when we use the standard factor-neutral decomposition, the importance of factors in explaining the 2010 variance of GDP per capita is just short of 25%, which is in line with the results found in the literature. How does departing from factor-neutrality affect our conclusion about the role of factors? When we assume $\sigma = 2.6$ in panel (a), we see that factors account for even less of the output variation more under the endogenous factor-biased technological progress model. This of course is just a summary of what we found in the previous subsection when we looked at the counterfactual distributions of GDP per worker in detail. In our model, when barriers are reduced, poor countries channel significantly more resources into innovation for the lower-skill sectors and enjoy very large productivity gains there, often exceeding the levels observed in more developed economies. This allows poorer countries to partially offset their disadvantage due to lower skill endowments. Balancing this effect of course is a symmetric one on the high-skilled end of the distribution: rich countries channeling resources to college-specific technology adoption. However, since those economies start with relatively low barriers, the counterfactual income distribution is much more compact, suggesting a small role for factors in explaining income disparities.[44] In contrast, when we use $\sigma = 1.6$ (in panel (b)), we see that factors account for even more of the income variation than under factor neutrality. Again, this is not surprising given the discussion in the previous subsection; in the absence of strong technology bias, as we have seen in the previous sub-section, reduction in barriers does not result in large productivity gains. Second, because a large number of countries were found to have barriers lower than the U.S., and thus did not have them reduced in order to compute the counterfactual.[45] Thus, according to our preferred specification ($\sigma = 2.6$), barriers to technology adoption account for the bulk of cross-country differences in GDP per worker, with only about a quarter of the variation due to a combination of factors (direct and through technology adoption) and frontier technology availability. However, while it appears that our results under higher elasticity and the traditional factor neutral approach both imply a similarly large role of productivity differences versus inputs in explaining average income variation, the averaging masks some interesting differences. These are best seen using the rich-poor gap approach to development accounting decomposition. Following Jones' (2016)

[44]Note that in our model the remaining variation (after we give everyone the same level of barriers) is due to countries' individual skill and physical capital endowments, as well as the level of world frontier technologies.

[45]One could use another reference country, say, one with the lowest level of barriers to conduct the counterfactual analysis. However, as noted earlier, the number and the identity of countries found to have barriers lower than the US under $\sigma = 1.6$ makes us doubt that specification.

methodology, we decompose the fraction of the difference in GDP per worker between a given country and the US into components "due to" input and barriers as follows. Using equation (40), we have

$$\frac{y_i}{y_{US}} = \frac{\kappa_i}{\kappa_{US}} \times \frac{z_i}{z_{US}} \times \frac{\omega_i}{\omega_{US}} \tag{42}$$

We can then compute the fraction of this gap "due to barriers" ($\mathfrak{z}$) as the proportion of the gap that would be removed if we equated barriers for all countries, that is

$$\mathfrak{z}_i \equiv \frac{\frac{z_i}{z_{US}}}{\frac{\kappa_i}{\kappa_{US}} + \frac{z_i}{z_{US}} + \frac{\omega_i}{\omega_{US}}}$$

with analogous definitions for physical capital and human capital composition.[46] Table 15 below shows the results under the two values of elasticity and the contribution of factor-neutral TFP (taken from Jones, 2016) for a selection of countries. Notice, as we have already seen in the variance decomposition, under the higher elasticity scenario barriers explain most of the average income gap, while under low elasticity human capital plays the key role. The factor-neutral results are similar to our high elasticity case on average, barriers contribute 63% in our approach and TFP contributes 69% in the factor-neutral calculation. However, there are important differences. Notice that for the developed economies, even under the high elasticity case, our approach implies much smaller contribution of barriers than TFP in the traditional approach. For example, in the case of Singapore, our calculation never assigns more than 22% of the gap to barriers whereas under perfect substitutability between skill types TFP explains almost half of the gap. The results are more similar for less developed countries but here too we find cases were barriers are considerably less important than traditional TFP , regardless of the assumed value of elasticity (for example China). Finally, note that Ben Jones (2014) finds role of inputs (physical and human capital together) 79-88% with $\sigma = 1.6$, which is similar to our result of 74%.

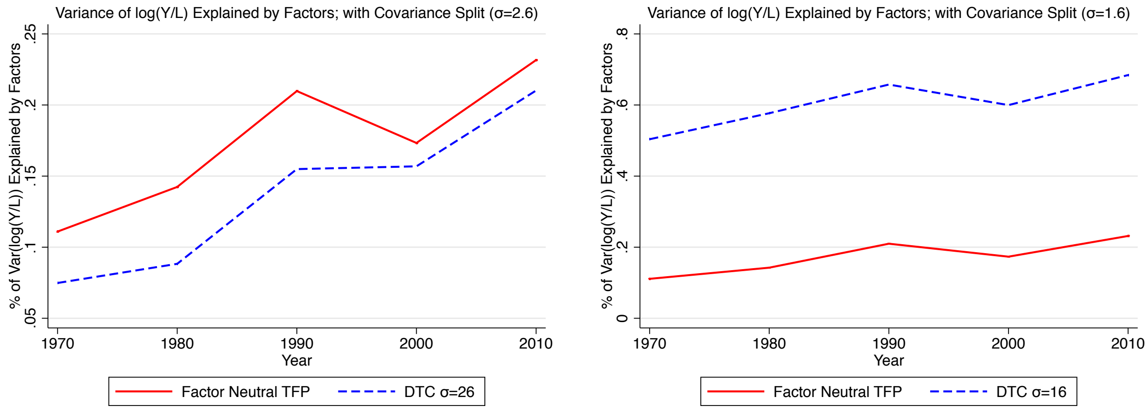
6 Conclusions

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

[46]Technically, this is a bit fuzzy, since the last term includes contribution of the shape of the frontier, which itself is determined by human capital differences in the past.

	$\sigma = 2.6$			$\sigma = 1.6$			Factor-neutral
	Barriers	*HC*	*K*	*Barriers*	*HC*	*K*	TFP
Hong Kong	0.22	0.46	0.33	0.15	0.56	0.29	0.49
Singapore	0.20	0.45	0.35	0.13	0.57	0.31	0.46
France	0.36	0.36	0.28	0.33	0.39	0.28	0.56
Germany	0.34	0.37	0.29	0.28	0.43	0.29	0.57
United Kingdom	0.32	0.38	0.30	0.28	0.43	0.29	0.46
Japan	0.39	0.36	0.24	0.35	0.41	0.24	0.64
South Korea	0.34	0.34	0.32	0.32	0.36	0.32	0.65
Argentina	0.42	0.30	0.29	0.38	0.33	0.29	0.67
Mexico	0.44	0.32	0.24	0.30	0.47	0.23	0.60
Botswana	0.62	0.25	0.14	0.25	0.61	0.14	0.74
South Africa	0.48	0.30	0.22	0.28	0.50	0.22	0.65
Brazil	0.47	0.30	0.23	0.31	0.46	0.23	0.75
Thailand	0.47	0.31	0.22	0.32	0.46	0.22	0.79
China	0.45	0.32	0.23	0.20	0.59	0.21	0.83
Indonesia	0.48	0.28	0.24	0.26	0.51	0.23	0.78
India	0.53	0.27	0.21	0.27	0.52	0.21	0.67
Kenya	0.74	0.16	0.11	0.23	0.65	0.12	0.87
Malawi	0.85	0.09	0.06	0.08	0.87	0.05	0.94
Entire Sample	0.63	0.22	0.15	0.26	0.58	0.16	0.69

Table 15: Development Accounting. Shares of the gap between each country's GDP per worker and that of the U.S. ($\frac{y_i}{y_{US}}$) accounted for by barriers ($\frac{z_i}{z_{US}}$), human ($\frac{\omega_i}{\omega_{US}}$), and physical capital ($\frac{\kappa_i}{\kappa_{US}}$) – under the directed technology approach (computed using equation (42)), or TFP – under the conventional factor neutral approach (from Jones, 2016).



(a) Factors; $\sigma = 2.6$

(b) Barriers; $\sigma = 1.6$

Figure 22: Development accounting. Variance ratio VR for the directed technology model (dashed lines) and the standard factor-neutral model (solid lines).

capital, physical capital, barriers to innovation and international technology diffusion. Using a unique new dataset we calibrated equilibrium conditions of the model and simultaneously computed the skill-specific productivity levels and measures of barriers to innovation for 168 countries over the period 1950-2010. We used these skill specific technologies and barriers measurements to study the historical patterns of directed technological change. We then evaluates the contribution of non-neutral technology and barriers to cross-country income differences.

Our estimates of the world technology frontier imply that in the early part of the 20th century technological progress favored high-school-educated workers over those with more education, while after 1950 – with a brief interruption in the 1970's – college-specific frontier productivity growth outstrips the other categories. Comparing skill-specific productivity across countries, we find that rich countries use labor of all three skill categories more efficiently. However, low productivity levels in poor countries were driven by high barriers to technological adoption. A counterfactual calculation that removes barriers, reveals that in their absence poor countries would actually achieve higher productivity in the lower skilled sectors that their richer counterparts. Our results also shed light on the puzzle of low (or negative) TFP growth found in many countries using the factor-neutral approach. This finding, if we think of TFP as a measure of technology knowledge, is hard to reconcile with the notion of technology diffusion 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, our calculations show that poor countries also experience relatively robust growth college-specific productivity. In some cases it is indeed offset by increases in barriers, but typically barriers have been falling, even in poor countries. Their GDP growth fails to reflect that 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 the frontier has stagnated. Thus 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 barriers to technology in explaining the current disparities of standards of living, the value of the elasticity of substitution between skill-types is crucial. Under a lower value of 1.6, our model yields barrier estimates that are lower and relatively less important in explaining cross-country income differences: under this scenario endowments of physical and human capital account almost 70% of variance in 2010 GDP per worker. Using elasticity of 2.6, we find barriers that are higher and explain almost all of the variation in output. To the extent that reduction of barriers to technological adoption and increasing educational attainment involve distinct approaches, our results imply that policies most effective at raising income levels in poor countries may be quite different depending on how difficult it is to substitute unskilled labor for its more skilled counterpart. Guided by our version of directed endogenous technological progress, the preferred value of elasticity is one which implies strong bias technological progress. Thus providing more support for the barrier reduction approach. However, more research on the elasticity of substitution between skill-types, especially in a cross-country setting, is fundamental.

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.

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

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.

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

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 Macroeconomics: 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, Benjamin F. "The human capital stock: a generalized approach". *American Economic Review*. 104, 2014: 3752-3777.

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.

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 19^{th} 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. "Premature Deindustrialization ". *Journal of Economic Growth* 21, 2015: 1-33.

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

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 "Clemson University working paper, 2016.

Tamura, Robert, Dwyer, Gerald P., Devereux, John, and Baier, Scott L. "Data Appendix for Economic Growth in the Long Run "Clemson University working paper, 2016.

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 Intermediate Good Prices

This section derives expressions for equilibrium intermediate good prices. Note that using equations (3) and (12) we can show that along the BGP price of L-type intermediate goods will be given by (check the math here)

$$P_L^* = \left[\gamma_L^\epsilon + \gamma_H^\epsilon \left[\left(\frac{H}{L}\right)^{-\frac{\beta(\phi+1)}{\sigma\phi+1}} \left(\frac{\gamma_H}{\gamma_L}\right)^{\frac{\beta\epsilon\phi}{(1+\sigma\phi)\sigma}} \left(\frac{\eta_H}{\eta_L}\right)^{-\frac{\beta}{1+\sigma\phi}} \left(\frac{\mathrm{A}_H^W}{\mathrm{A}_L^W}\right)^{-\frac{\beta\phi}{1+\sigma\phi}}\right]^{1-\epsilon}\right]^{\frac{1}{\epsilon-1}}$$

which is increasing in H/L. While, the BGP price of H-type intermediate goods will be given by

$$P_H^* = \left[\gamma_H^{-\epsilon} - \left(\frac{\gamma_L}{\gamma_H}\right)^\epsilon P_L^{*\,1-\epsilon}\right]^{\frac{1}{1-\epsilon}} \tag{43}$$

$$= \left[\gamma_H^{-\epsilon} - \left(\frac{\gamma_L}{\gamma_H}\right)^\epsilon \left[\left(\gamma_H^\epsilon \left(\left(\frac{\eta_L}{\eta_L}\right)^{-\frac{\beta}{\sigma\phi+1}} \left(\frac{H}{L}\right)^{-\frac{\beta(\phi+1)}{\sigma\phi+1}} \left(\frac{\mathrm{A}_H^W}{\mathrm{A}_L^W}\right)^{-\frac{\beta\phi}{\sigma\phi+1}} \left(\frac{\gamma_H}{\gamma \mathrm{L}}\right)^{\frac{\beta\epsilon\phi}{\sigma\phi+1}}\right)^{1-\epsilon} + \gamma_L^\epsilon\right)^{\frac{1}{\epsilon-1}}\right]^{1-\epsilon}\right]^{\frac{1}{1-\epsilon}} \tag{44}$$

which is decreasing in H/L.

B Estimating σ using EU KLEMS

In Jerzmanowski and Tamura (2017) we 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 main results related to the present paper.

Our approach can be best described by starting with the equation for skill premium

$$\frac{w_H}{w_L} = \left(\frac{A_H}{A_L}\right)^{\frac{\sigma-1}{\sigma}} \left(\frac{H}{L}\right)^{-\frac{1}{\sigma}} \tag{45}$$

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

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

they get 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) estimateed the above using data on college/high school 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 instrumental variables strategy (since H/L responds to shock to wages, 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 High School degree), medium-skill (High School degree) and high-skills (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

$$\frac{w_H}{w_L} = \left(\frac{\eta_H}{\eta_L}\right)^{\frac{\sigma-1}{1+\phi\sigma}} \left(\frac{H}{L}\right)^{\frac{\sigma-2-\phi}{1+\phi\sigma}} \left(\frac{A_H^W}{A_L^W}\right)^{\frac{\phi(\sigma-1)}{1+\phi\sigma}}. \tag{46}$$

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

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

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

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

We use the estimate of the coefficient on the relative skill supplies in the above regression, $\frac{\sigma-2.5}{1+\sigma/2}$, to back out the estimate of σ.

Table 16 below shows the results of estimating the above equation using our data for college and high school groups. We sue 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 most of our estimates fall in the range $2.30 - 2.57$.

Table 16: College and High School

σ (OLS)	2.568	2.566	2.314	2.310
s.e.	0.103	0.103	0.119	0.114
R^2	0.031	0.069	0.829	0.835
p($\sigma < 1.6$)	0.000	0.000	0.000	0.000
p($\sigma < 2.6$)	0.619	0.491	0.966	0.973
σ (FE)	2.311	2.308	2.066	2.044
s.e.	0.117	0.112	0.130	0.179
R^2	0.190	0.218	0.730	0.758
p($\sigma < 1.6$)	0.000	0.000	0.040	0.120
p($\sigma < 2.6$)	0.970	0.975	0.998	0.989
σ (GMM)	2.290	2.294	2.058	2.038
s.e.	0.120	0.117	0.137	0.189
R^2				
p($\sigma < 1.6$)	0.000	0.000	0.053	0.138
p($\sigma < 2.6$)	0.976	0.977	0.997	0.987
σ (Sys. GMM)	2.571	2.571	2.299	2.298
s.e.	0.106	0.106	0.111	0.111
R^2				
p($\sigma < 1.6$)	0.000	0.000	0.000	0.000
p($\sigma < 2.6$)	0.470	0.470	0.979	0.979
Time Trend	*Yes*	*Yes*	*Yes*	*Yes*
Trend Squared	*No*	*Yes*	*No*	*Yes*
Country Trend	*No*	*No*	*Yes*	*Yes*
N	320	320	320	320

C BGP Output per Worker

This section derives the expression for balanced growth path GDP per worker used in the main text. We start with the equation of the text

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

where

$$s_H = \frac{H}{H+L} = \frac{H}{N}$$

furthermore, since and the fact that

$$K = K_L(1+\frac{K_H}{K_L})$$

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

and since from (17) and (33) we know that $\mu_i = \zeta^{-\frac{1}{\phi}}R(\tau)^{\frac{\beta-1}{\phi\beta}}f_i\left(\frac{H}{L},\frac{A_H^W}{A_L^W}\right)$ and $s_i = h_i(H/L)$ for $i = H, L$, and $\frac{K_H}{K_L} = g\left(\frac{H}{L},\frac{A_H^W}{A_L^W}\right)$ we can write

$$\frac{Y}{N} = (A_L^W)^\beta\left(\frac{K}{N}\right)^{1-\beta}\zeta^{-\frac{\beta}{\phi}}R(\tau)^{\frac{\beta-1}{\phi}}\Omega\left(\frac{H}{L},\frac{A_H^W}{A_L^W}\right)$$

or

$$\frac{Y}{N} = A_L^W\left(\frac{K}{Y}\right)^{\frac{1-\beta}{\beta}}\zeta^{-\frac{1}{\phi}}R(\tau)^{\frac{\beta-1}{\phi\beta}}\Omega\left(\frac{H}{L},\frac{A_H^W}{A_L^W}\right)^{\frac{1}{\beta}}$$

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

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

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

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

$$y \propto \left(A_L^W\right)^{\frac{\phi}{1+\phi}} \left(\frac{K}{Y}\right)^{\frac{1-\beta}{\beta}} \left(\frac{\zeta}{y}\right)^{-\frac{1}{1+\phi}} \Omega\left(\frac{H}{L}, \frac{A_H^W}{A_L^W}\right)^{\frac{\phi}{(1+\phi)\beta}}$$

D Variance Decomposition

This section describes the variance decomposition used in the main text. Let y denote the observed level of GDP per worker and y^* be the counterfactual. Then we can write

$$y = y^* \times \frac{y}{y^*}$$

or taking logs and variance of both sides we have

$$Var(\ln y_i) = Var(\ln y^*) + Var\left(\ln(\frac{y}{y^*})\right) + 2 \times Cov(\ln y^*, \ln(\frac{y}{y^*}))$$

The *variance ratio* can then be defined as

$$\text{VR} = \frac{Var(\ln(y^*)) + \frac{1}{2} \times Cov(\ln y^*, \ln(\frac{y}{y^*}))}{Var(\ln(y))},$$

where we have split the covariance term following Klenow and Rodriguez-Clare (1998) (see Caselli (2005) and Baier et al. (200x) for alternatives to such even split of the covariance).

For example, in the standard Cobb-Douglas-based development accounting, researchers construct the counterfactual by endowing all countries with the highest (factor-neutral) TFP level observed in the world (usually the U.S. level). The resulting variance ratio is the percentage of variation in incomes that would remain after all productivity differences have been eliminated and is interpreted as the contribution of everything except TFP – usually taken to mean physical and human capital – to income differences. Specifically, using $y = TFP^{1-\alpha} \times k^{\alpha}h^{1-\alpha}$,, where k and h denote physical and human capital per worker, and setting the counterfactual as $y^* = TFP_{US} \times (\frac{k}{y})^{\alpha/(1-\alpha)}h$, one obtains VR $= \frac{Var(\ln(k^{\alpha}h^{1-\alpha}+1/2Cov(\ln(k^{\alpha}h^{1-\alpha},\ln(TFP)))}{Var(\log y)}$. In an influential paper, Klenow and Rodriguez-Clare (1998) use this approach and report the variance ratio of about 1/3, which implies that physical and human capital together explain only about one third of variation in output per worker (see Jones (2016) for a more updated discussion). In our setting, the variance ratio will be given by the variance of (log) counterfactual GDP per worker (y^*) plus one half of its covariance with the (log) gap between actual and counterfactual output levels (y/y^*), divided by the variance of (log) actual GDP per worker. For example, in the barrier reduction counterfactual we have $y_i^* \propto \kappa_i z_{US} \omega_i$, and so

$$Var(\ln y_i) = Var(\ln(\kappa_i \omega_i)) + Var\left(\ln(z_i)\right) + 2 \times Cov(\ln(\kappa_i \omega_i), \ln(z_i))$$

and

$$\text{VR} = \frac{Var(\ln(\kappa_i \omega_i)) + \frac{1}{2} \times Cov(\ln(\kappa_i \omega_i), \ln(z_i))}{Var(\ln(y))},$$

E Transitional Dynamics and the Calibration of ϕ

This section describes the dynamics of the model, which we use to calibrate ϕ, the parameter that governs the strength of technology diffusion. 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)

$$V_H = \eta_H^{-1} \zeta H \mu_H^{\phi} \tag{47}$$

we can differentiate the free entry condition to yield

$$\frac{\dot{V}_H}{V_H} = \phi \, \frac{\dot{\mu}_H}{\mu_H} \tag{48}$$

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

$$\frac{\dot{A}_H}{A_H} = \eta_H \mu_H^{-(1+\phi)} \frac{\tilde{Z}_H}{H^\lambda} \tag{49}$$

and

$$\frac{\dot{A}_L}{A_L} = \eta_L \mu_L^{-\phi} \frac{A_H^W \tilde{Z}_L}{H^\lambda} \tag{50}$$

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

$$\frac{\dot{\mu}_H}{\mu_H} = \eta_H \mu_H^{-(1+\phi)} \frac{\tilde{Z}_H}{H^\lambda} - g \tag{51}$$

$$\frac{\dot{\mu}_L}{\mu_L} = \eta_L \mu_L^{-(1+\phi)} \frac{A_H^W/A_L^W \tilde{Z}_L}{H^\lambda} - g \tag{52}$$

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 (48), (51), and (52) we get

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

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

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

$$I = Y - \zeta(Z_H + Z_L) - C$$

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

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 the system only has one negative root and this root determines the speed of convergence to the BGP.