DIRECTED TECHNOLOGICAL CHANGE: A QUANTITATIVE ANALYSIS

Michal Jerzmanowski & Robert Tamura

February 15, 2013

Abstract

We explore the role of skill non-neutral productivity differences and barriers to technology adoption in explaining cross-country income differences. Using a model of directed technological change modified to include physical capital, technology diffusion and barriers to entry, and a new data set on output and labor force composition we construct measures of productivity by worker skill type for a large cross-section of countries going back as far as the 19th century in some cases. Additionally, the equilibrium conditions of the model allow us to back out the measure of barriers to adoption of technology. We use these 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. We find that allowing for non-neutrality of productivity increases the importance of human capital differences in explaining income variation across countries. Additionally, we find that barriers to adoption of technology are important for understanding productivity differences but this importance has been declining in recent decades. Importantly, in the presence of technology diffusion, barriers are only important if the elasticity of substitution between skill types is well above 2. We argue that other predictions of the model, especially about the cross-country wage distributions, are also consistent with data only for large values of elasticity of substitution. We provide some auxiliary empirical evidence using EU KLEMS data which suggest that the elasticity of substitution between skill types may indeed be considerably greater than 2. Finally, we find that historical patterns of skill bias in technological change which we measure are largely consistent with priors.
1 Introduction

This paper takes the theory of directed technological progress (Acemoglu, 2002, 200x), seriously. We use the theory, combined with a new data set, to address one of the most important challenges of macroeconomics: understanding cross-country income differences. Specifically, we are interested in the relative importance of technology differences versus factor endowments, such as physical and human capital, in explaining income variation around the world. Much of the large literature on this topic employs the Cobb-Douglas production function approach and thus imposed factor-neutral productivity differences. Recently, there is growing evidence that in order to fully understand productivity differences, one needs to go beyond the factor-neutral assumption (Caselli and Coleman, 2006; Jerzmanowski 2007). One way to do this is to allow technological progress to differentially affects the productivity of workers different skill sets. In fact, the literature on the evolution of U.S. wage inequality has long argued that this skill-bias has characterized technological progress in the last 30 or so years. In the economic growth literature, Acemoglu (1998, 2002, 2003, 2007) provide a rich theoretical analysis of cross-country income differences resulting from directed technological progress. Caselli (2005) and Caselli and Coleman (2006) provides pioneering empirical analyses of the importance of factor-bias for cross-country productivity differences. However, these papers relay on international wage data calculate skill specific productivity levels. Such data is often not available for many countries and the time series is usually not very long. Moreover, the quality of cross-country wage data is often low. In this paper, we propose to employ the theory of directed technological progress for quantitative analysis of cross-country productivity differences. We use the model’s equilibrium conditions to back out skill-specific productivity levels requiring only data on output, factor inputs and skill composition of the labor force. We make use of a unique new data set on output, physical capital and education levels of the labor force constructed by Tamura, Dwyer, Devereux, Baier (2011). This data set contains observations on real output per worker, real physical capital per worker, and schooling per worker for 168 countries. For 50 countries we have time series from 1820-2007, while most of our sample covers the years 1950-2007.

We set up a model of directed technological change (Acemoglu; 1998, 2002) modified to include physical capital, technology diffusion and barriers to entry. In our model there are three skill groups which correspond to the education categories in our data: workers exposed to no more the primary school, those exposed with some secondary school (but no more) and those exposed to higher education. We derive the balanced growth path equilib-
rium conditions, calibrate the parameter values and use our data to back out skill-specific productivity levels and measure soy barriers to innovation. We use these 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. We find that allowing for non-neutrality of productivity increases the importance of human capital differences in explaining income variation across countries. Additionally, we find that barriers to adoption of technology are important for understanding productivity differences but this importance has been declining in recent decades. Importantly, in the presence of technology diffusion, barriers are only important if the elasticity of substitution between skill types is well above 2. We argue that other predictions of the model, especially about the cross-country wage distributions, are also consistent with data only for large values of elasticity of substitution. We provide some auxiliary empirical evidence using EU KLEMS data which suggest that the elasticity of substitution between skill types may indeed be considerably greater than 2. Finally, we find that historical patterns of skill bias in technological change which we measure are largely consistent with priors.

2 Model

2.1 Households

There is a continuum of infinitely lived households with CRRA preferences and a discount rate of $\rho$.

2.2 Final Good

In the empirical implementation of the model we will consider the following general production technology:

$$Y_t = \{\gamma_H Y_H^{\varepsilon - 1} + \gamma_S Y_S^{\varepsilon - 1} + \gamma_p Y_p^{\varepsilon - 1}\}^{\frac{1}{\varepsilon - 1}}$$

(1)

where $Y_i, i = H, S, p$ are intermediate outputs produced by skill workers of type $i$. We assume that there are three skill types of workers: those exposed to higher education, $H$, those exposed to secondary school, $S$, and those with no education or exposed to at most primary school.\(^1\)

\(^1\)In Tamura, Dwyer, Devereux, and Baier [2011] the final category is broken out separately into no education, and those exposed to at most primary school. However by the end of the time frame, 2007, 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.
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. This allows for a direct comparison with Acemoglu’s (200X) original model which we are modifying. Key results are restated for the three-skill type model at the end of this section.

Let final output be produced using two types of intermediate goods

\[ Y_t = \left\{ \gamma_H Y_H^{\frac{\epsilon - 1}{\epsilon}} + \gamma_L Y_L^{\frac{\epsilon - 1}{\epsilon}} \right\}^{\frac{1}{\epsilon - 1}} \]  

(2)

These intermediate outputs are produced by competitive firms, and sold to competitive final output producers for prices \( p_i, i = H, Sp \). The typical final goods producer has the following optimization problem:

\[ \max_{\{Y_H, Y_L\}} \{ Y_t - \sum_{i \in \{H, L\}} P_i Y_i \} \]  

(3)

The first order conditions for profit maximization are:

\[ Z \gamma_L Y_L^{\frac{1}{\epsilon}} = P_L \]
\[ Z \gamma_H Y_H^{\frac{1}{\epsilon}} = P_H \]
\[ Z = \left\{ \gamma_L Y_L^{\frac{\epsilon - 1}{\epsilon}} + \gamma_H Y_H^{\frac{\epsilon - 1}{\epsilon}} \right\}^{-\frac{1}{\epsilon - 1}} \]

The final good is the numereire so that

\[ \left[ \gamma_L P_L^{1-\epsilon} + \gamma_H P_H^{1-\epsilon} \right]^{\frac{1}{1-\epsilon}} = 1 \]  

(4)

2.3 Intermediate Goods

The intermediate outputs are produced combining machines and labor of skill type \( i \) in a standard variety of machine inputs manner:

\[ Y_L = \frac{1}{1 - \beta} \int_0^{A_L} \chi_i L^{1-\beta} dL^\beta \]  

(5)
\[ Y_H = \frac{1}{1 - \beta} \int_0^{A_H} \chi_i H^{1-\beta} dH^\beta \]  

(6)

where \( A_i \) is the level of directed technology for skill factor \( i = p, S, H, K_i \), where \( K_i \) is the total capital allocated to skill set \( i \), \( N_i \) is the supply of skill type \( i \) workers (later we use \( H = s_H N \), where \( N \) is total labor force).
The typical intermediate goods firm using skill type \(i\) has the following maximization problem:

\[
\max_{\{\chi_i(o), L_i\}} \left\{ \frac{P_i}{1 - \beta} \int_0^{A_i} \chi_i(o)^{1-\beta} \, do L_i^\beta - \int_0^{A_i} p_i(o)\chi_i(o)\, do - w_i L_i \right\}
\]  

(7)

Relative to Acemoglu (1998) we make the following modification: machines are produced one for one using capital rented at the rate \(R\). That is to say one unit of physical capital can produce one machine. For a representative firm hiring skill type \(i\), the inverse derived demand for a typical machine \((o)\) is given by:

\[
P_i \chi_i(o)^{-\beta} L_i^\beta = p_i(o)
\]

(8)

We also assume (following Agion and Howitt, XXXX) that each machine producing monopolist faces an potential imitator with cost \(\mu > 1\) times higher the original innovator’s own marginal cost, i.e the imitator uses \(\mu\) units of capital to produce one machine. This implies that the profit maximizing monopolist will set the price equal to a \(\mu\) markup over her own marginal cost\(^2\)

\[
p_i(o) = \mu R
\]

(9)

The equilibrium supply of machines of type \((o)\) to skill \(i = L, H\), and the equilibrium quantities of machines are:

\[
\chi_i(o) = \left( \frac{P_i}{\mu R} \right)^{1/\beta} N_i
\]

(10)

which means the (derived) production functions become

\[
Y_i = \frac{1}{1 - \beta} \left( \frac{P_i}{\mu R} \right)^{\frac{1-\beta}{\beta}} A_i N_i
\]

(11)

\(^2\text{This is try as long as } \mu < \frac{1}{1 - \beta}, \text{ which we assume to be true.}\)
And the profit per line of machines is given by

\[ \pi_i(o) = \left( \frac{\mu - 1}{\mu} \right) P_i^{1/\beta} N_i (\mu R) \frac{\sigma - 1}{\sigma} \]  

(12)

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

\[ \frac{P_H}{P_L} = \left( \frac{\gamma_H}{\gamma_L} \right)^{\frac{\sigma}{\beta}} \left( \frac{A_H}{A_L} \right)^{-\frac{1}{\sigma}} \]  

(13)

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

2.4 Wages and Technology

Returning to the intermediate goods producing firm in skill i, their F.O.C. for hiring workers of skill i is:

\[ \frac{\beta P_i}{1 - \beta} \int_0^{A_i} \chi_i(o)^{1-\beta} doL_i^{\beta-1} = w_i \]  

(14)

Substituting for the equilibrium quantities of machines, and available workers of skill i produces:

\[ w_i = \frac{\beta}{1 - \beta} A_i \beta P_i^{1/\beta} (\mu R)^{-\frac{1-\beta}{\sigma}} \]  

(15)

With that we get, as in Acemoglu, the relative wages of skill types:

\[ \frac{w_H}{w_L} = \left( \frac{\gamma_H}{\gamma_L} \right)^{\frac{\sigma}{\beta}} \left( \frac{A_H}{A_L} \right)^{\frac{\sigma - 1}{\sigma}} \left( \frac{H}{L} \right)^{-\frac{1}{\sigma}} \]  

(16)

where \( \sigma \) is the elasticity of substitution between \( N_H \) and \( N_p \) and analogously between \( N_S \) and \( N_p \).

2.5 Capital Allocation

Capital is used to make the three types machines, thus we have

\[ K_i = \int_0^{A_i} \chi_i(o) do = A_i \left( \frac{P_i}{\mu R} \right)^{1/\beta} N_i \]  

(17)

so that
\[
\frac{K_H}{K_L} = \left(\frac{\gamma_H}{\gamma_p}\right)^{\varepsilon/\sigma} \left(\frac{A_H}{A_p}\right)^{\frac{\sigma-1}{\sigma}} \left(\frac{H}{L}\right)^{\frac{\sigma-1}{\sigma}}
\] (18)

In addition we have an adding restriction that

\[K_L + K_H = K\] (19)

Note that the marginal product of capital is equalized across sectors. To see this note that for intermediate goods producers we have that machines take one unit of capital to produce. Since all machines within a skill industry are symmetric it must be the case that:

\[K_L = A_L\chi_L = A_L k_L\] (20)
\[K_H = A_H\chi_H = A_H k_H\] (21)

Substituting for \(k_i\) produces:

\[Y_L = \frac{1}{1-\beta} K_L^{1-\beta} (A_L L)^{\beta}\] (22)
\[Y_H = \frac{1}{1-\beta} K_H^{1-\beta} (A_H H)^{\beta}\] (23)

so

\[MPK_L = P_L K_L^{1-\beta} (A_L L)^{\beta} = P_L (1 - \beta) \frac{Y_L}{K_L}\] (24)
\[MPK_H = P_H K_H^{1-\beta} (A_H N_H)^{\beta} = P_H (1 - \beta) \frac{Y_H}{K_H}\] (25)

and so using the expression for \(P_i/P_p\) and setting marginal product of capital equal yields the above expression (33)-(34) for capital ratios.

Also note that when intermediate producers rent machines, they pay a rental rate of \(R\mu\) per unit of capital where \(\mu\) is the mark-up over the rental rate. This means that

\[R\mu = MPK_H = P_H K_H^{1-\beta} (A_H N_H)^{\beta} = P_H (1 - \beta) \frac{Y_H}{K_H}\]

and we have
\[ R = \frac{MPK}{\mu} = P_H \left( \frac{1 - \beta}{\mu} \right) \frac{Y_H}{K_H} \]

Recall that since the final good is the numereire, we have
\[ [\gamma_L^{\epsilon} P_L^{1-\epsilon} + \gamma_H^{\epsilon} P_H^{1-\epsilon}] \frac{1}{\epsilon} = 1 \]  
(26)

Also, since the final good sector is competitive, it must be that
\[ P_L Y_L + P_H Y_H = Y \]

We have
\[ R = P_H \left( \frac{1 - \beta}{\mu} \right) \frac{Y_H}{K_H} = \left( \frac{1 - \beta}{\mu} \right) \frac{Y - P_L Y_L}{K_H} \]

and we can use the MPK equalization across sectors
\[ R = P_p \left( \frac{1 - \beta}{\mu} \right) \frac{Y_L}{K_L} \]
to substitute for \( P_L Y_L \) to obtain
\[ R = \left( \frac{1 - \beta}{\mu} \right) \frac{Y}{K} \]  
(27)

To understand this expression, note that the equilibrium MPK is given by \( (1 - \beta)Y/K \)
and the rental rate is equal to \( MPK/\mu \) where \( \mu \) is the markup in the machine market.

Finally, the interest rate in this economy is given by \( r = (1 - \tau)R - \delta \), where \( \delta \) is the rate of depreciation of capital and \( \tau \) is the tax on capital income.
\[ r = (1 - \tau) \left( \frac{1 - \beta}{\mu} \right) \frac{Y}{K} - \delta \]

2.6 Innovation

Discovery of new blueprints for sector \( i \) in country \( k \) is governed by the following process
\[ \dot{A}_{Hk} = \bar{\eta}_i \left( \frac{A_W^i}{A_{Hk}} \right)^{\phi} \frac{Z_{Hk}}{H_k^\lambda} \]  
(28)
\[ \dot{A}_{Lk} = \bar{\eta}_i \left( \frac{A_W^i}{A_{Lk}} \right)^{\phi} \frac{Z_{Lk}}{H_k^\lambda} \]

where represents \( A_W^i \) is the world frontier technology for sector \( i \), \( \bar{\eta}_i \) is the productivity
of research effort, and \( Z_{ik} \) is the R&D expenditure on innovation or technology adoption in sector \( i \). In Acemoglu’s original model \( \phi = 0 \). 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 skilled worker raised to the power \( \lambda \). We do this to reduce the scale effects in levels, namely a situation where countries with larger populations have higher levels of productivity. This can be motivated by a duplication of R&D effort argument. See Klenow and Rodriguez-Clare (2004, handbook). A value of \( \lambda = 1 \) completely eliminates scale effects, a value of \( \lambda = 0 \) is equivalent to Acemoglu’s original formulation and features strong scale effects. Intermediate values mitigate scale effects. Finally, we assume that research productivity depends on the share of skilled workers

\[
\tilde{\eta}_i = \eta_i \left( \frac{H}{H+L} \right)^\psi
\]

where the \( \eta_i \) is sector specific but common across countries while the term \( (H/(H+L))^\psi \) reflects the idea that higher concentration of skilled workers facilitates spillovers in research; this term is country-specific.

The cost of innovation/adoption is given by \( \zeta_k \) and is specific to the country and not the sector, it represents the barriers to entry and other institutions that increase the costs of introducing new technologies (Parente and Prescott 1994, 1999). Note that in Acemoglu’s original formulation \( \zeta_k = 1 \). We again relax this assumption to allow the model to account for cross country differences in productivity.

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

\[
\tilde{\eta}_H \left( \frac{A_W^H}{A_{Hk}} \right)^\phi \frac{V_{Hk}}{H_k^\lambda} = \zeta_k
\]

(29)

\[
\tilde{\eta}_L \left( \frac{A_W^L}{A_{Lk}} \right)^\phi \frac{V_{Lk}}{H_k^\lambda} = \zeta_k
\]

where \( V_{ik} \) is the value of a blueprint for a machine in sector \( i = H, L \). Defining \( \mu_i = \frac{A_i^W}{A_i^H} \) and dropping the country indicator, this equation implies that

\[
\frac{V_H}{V_L} = \left( \frac{\eta_H}{\eta_L} \right)^{-1} \left( \frac{\mu_H}{\mu_L} \right)^\phi
\]

(30)
2.7 BGP Growth Rate and Interest Rate

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

\[
g = \frac{1}{\theta} [r - \rho] = \frac{1}{\theta} [(1 - \tau)R - \delta - \rho]
\]

where \( \rho \) is the discount rate, \( \theta \) is the CRRA coefficient, and \( \tau \) is the tax rate on capital income. The interest rate \( r \) will be equal to the rental rate minus the rate of depreciation \( r = (1 - \tau)R - \delta \). In the model the rental rate would adjust to ensure that the above Euler equation holds. In our empirical analysis we will appeal to the literature that finds no significant differences in the marginal product of capital across countries and use the equation to back out \( \tau \).

Also, using the no-arbitrage conditions for \( i = H, S \) or \( p \).

\[
r V_i = \pi_i + \dot{V}_i
\]

and the fact that along the BGP the value of a patent must be stationary (\( \dot{V}_i = 0 \)) we get the following relationship between the value of a patent, profits and the interest rate

\[
V_i = \frac{\pi_i}{r}
\]

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

\[
\frac{V_H}{V_L} = \frac{\pi_H}{\pi_L} = \left( \frac{\mu - 1}{\mu} \right) P_H^{1/\beta} H (\mu R)^{\beta - 1} / \sigma = \left( \frac{P_H}{P_L} \right)^{1/\beta} \frac{H}{L}
\]

and using the expression for relative prices we get

\[
\frac{V_H}{V_L} = \left( \frac{\gamma_H}{\gamma_L} \right)^{\frac{\sigma}{\sigma - 1}} \left( \frac{A_H}{A_L} \right)^{-\frac{1}{\sigma}} \left( \frac{H}{L} \right)^{\frac{\sigma - 1}{\sigma - 1}}
\]

(31)

Finally, combining equations (30), (31), and the expression for relative prices \( P_H/P_L \) derived above yields

\[
A_H / A_L = \left( \frac{\eta_H}{\eta_L} \right)^{\sigma} \left( \frac{\gamma_H}{\gamma_L} \right)^{\epsilon} \left( \frac{H}{L} \right)^{\sigma - 1}.
\]

\( ^{3}\) Notice that in our baseline specification (with \( \phi = 0 \)) this collapses to the expression familiar from Acemoglu
Thus the relative levels of productivity are increasing in the relative supply of skilled workers $H/L$ as long as $\sigma > 1$.

### 3 Discussion of the Model

#### Skill Bias

Recall from Acemoglu (1998) that we call technological change skill biased ($H$-biased in our notation) whenever an increase in the level of technology increase the relative marginal product of skilled workers. From equation (16) it is clear that an increase in $A_H/A_L$ is a skill-biased technological change as long as $\sigma > 1$. From equation (32) it is also apparent that an increase in $H/L$ induces skill-biased technological change whenever $\sigma > 1$; that the model always displays the weak equilibrium skill-bias.\(^4\) Notice in equation (16) 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.

Substituting the expression for relative productivity levels (32) into the relative wage formula (16) we obtain

$$\frac{A_H}{A_L} = \left( \frac{\eta_H}{\eta_L} \right)^{\frac{\sigma-1}{1+\phi\sigma}} \left( \frac{\gamma_H}{\gamma_L} \right)^{\frac{\epsilon(1+\phi)}{1+\phi\sigma}} \left( \frac{H}{L} \right)^{\frac{\sigma-2-\phi}{1+\phi\sigma}} \left( \frac{A_W}{A_L} \right)^{\frac{\phi(\sigma-1)}{1+\phi\sigma}} \quad (32)$$

Clearly the strong skill bias is present as long as

$$\sigma > 2 + \phi$$

which reduces to the result familiar from Acemoglu: $\sigma > 2$ when $\phi = 0$. Notice that the presence of international technology diffusion ($\phi > 0$) implies a higher value of $\sigma$ is required to strong bias to exist.

Notice also that technically the reason for the skill bias is slightly different than in Acemolglu. Here, as in Acemoglu, the profits form innovating for both types of workers

\(^4\)When $\sigma < 1$, we always have weak equilibrium unskill-bias.
increase with the supply of each type (the market-size effect), however the cost of innovation also increase with the size of the skilled labor force $H$. And since an increase in $H$ decreases the price of H-type intermediates, this means it decreases the return, per unit of R&D expenditure, to innovation for skilled workers. However it reduces that return for innovation in L-type technology even more... need to add something about $\eta$ so return may not actually go down.

**Distance to the Technology Frontier and Output per Worker on BGP**

Note that using equations (4) and (13) we can show that along the BGP price of L-type intermediate goods will be given by (check the math here)

$$P_L^* = \left[ \gamma_L^\varepsilon + \frac{\gamma_H^\varepsilon}{\gamma_H^\varepsilon} \left( \frac{H}{L} \right)^{-\frac{\beta(\phi+1)}{\sigma+1}} \left( \frac{\gamma_H}{\gamma_L} \right)^{\frac{\beta\phi}{(1+\sigma)\sigma}} \left( \frac{\eta_H}{\eta_L} \right)^{\frac{-\beta}{1+\sigma\phi}} \left( \frac{A_H}{A_L^{\frac{1}{\sigma}}} \right)^{\frac{-\beta}{1+\sigma\phi}} \right]^{\frac{1}{1-\varepsilon}}$$

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

$$P_H^* = \left[ \gamma_H^\varepsilon - \left( \frac{\gamma_L}{\gamma_H^\varepsilon} \right) \cdot \left( P_L^* \right)^{1-\varepsilon} \right]^\frac{1}{1-\varepsilon}$$

which is decreasing in $H/L$.

Using (29) this equation implies that the BGP distance to the frontier is given by

$$\mu_H = \left[ \eta_H \left( \frac{\mu-1}{\mu} \right) H^{1-\lambda} P_H^{1/\beta} (\mu R)^{\frac{\beta-1}{\beta}} \right]^{\frac{1}{1/\phi}}$$
\[\mu_L = \left[ \frac{\bar{\eta}_L \left( \frac{\mu - 1}{\mu} \right) (L/H) P_L^{1/\beta} (\mu R)^{\frac{1}{\phi}}}{r^* \zeta} \right]^{1/\phi}, \]  
\tag{37}

where the BGP interest rate is given by

\[r^* = \theta g + \rho. \]  
\tag{38}

and the BGP rental rate is

\[R^* = \frac{\theta g + \rho + \delta}{1 - \tau} \]  
\tag{39}

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

where

\[s_H = \frac{H}{H + L} = \frac{H}{N} \]

The appendix shows that this reduces to

\[Y = A_L^W \left( \frac{\mu}{1 - \beta} \right)^{\frac{1}{\phi}} \zeta^{\frac{1}{2}} R(\tau)^{\frac{(\beta - 1)(1 + \phi)}{\phi \beta}} \Omega \left( \frac{H}{L}, \frac{A_H^W}{A_L^W} \right)^{\frac{1}{2}} \]

Thus on the BGP output per worker depends on the world technology frontier, domestic relative supply of labor and distortions: barriers to entry \(\zeta\) and taxes on investment \(\tau\). Notice that it is decreasing in both distortions. Whether it is also 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 \(\mu_H\) and \(\mu_L\) as the composition of the labor force \(H/L\) changes.

**Technological Possibilities Frontier**

Historically, in most economies and time periods we have observed steady increases in education attainments (Barro and Lee, 200x), and thus a rise in the relative supply of skilled workers.\(^5\) An interesting question that arises in our model is what happens with

\(^5\)With three skill groups, as is the case in our empirical application the dynamics of relative labor supplies can be more complex. We ignore this issue here.
the two productivity levels $A_H$ and $A_L$ as the composition of the labor force changes in such away. Of course, in our model productivity levels are growing at a constant rate along the balanced growth path so the proper focus is the distances between a country’s productivity level and the world frontier ($\mu_H$ and $\mu_L$), which holding constant the level of barriers to innovation and investment distortions and treating the world productivity frontier as exogenous from the point of view of a single country, productivity levels are proportional to the quittances from the frontier

$$A_H(H/L) = A_H^W \mu_H(\zeta, \tau, H/L, A_H^W/A_L^W)$$

$$A_L(H/L) = A_L^W \mu_L(\zeta, \tau, H/L, A_H^W/A_L^W)$$

Let us focus on a version of the model with $\lambda = 1$, which is what we assume – following Klenow and Rodriguez-Claire (2003) – when we take the model to the data, and compare the balanced growth productivity levels of economies with different labor force composition.\(^6\)

Taking logs of equations (36) and (37) and differentiating with respect to $\ln(H/L)$ (holding barriers to entry and investment distortions constant) we get the following expressions

$$\frac{d \ln(\mu_H)}{d \ln(H/L)} = \frac{1}{\phi} \left( \frac{\psi}{1 + H/L} + \frac{1}{\beta} \frac{d \ln(P_H)}{d \ln(H/L)} \right) \tag{40}$$

$$\frac{d \ln(\mu_L)}{d \ln(H/L)} = \frac{1}{\phi} \left( \left[ \frac{1}{1 + H/L} \right] + \frac{1}{\beta} \frac{d \ln(P_L)}{d \ln(H/L)} \right) \tag{41}$$

To illustrate the discussion we will use Figure 1, whose two columns correspond to the two cases: no externality from skilled workers on research productivity ($\psi = 0$, right column) and positive externality ($\psi = 1$, left column). Consider the left column, its first two panels 1(a) and 1(c) graph separately the two parts in the bracket of equations (40) and (41), respectively. In each case, the second part, which is the effect of $H/L$ on the price of the skill-specific intermediate good, is plotted as an inverse so that the gaps between the solid and the dashed line indicate the sign of the expression: whenever the dashed red line is above the solid blue line, productivity is decreasing. The last panel, 1(e), shows the resulting “production possibilities frontier“, i.e. the possible configuration of skill specific productivity levels for different labor force composition values. The right columns shows the corresponding graphs for the case of $\psi = 0$.\(^7\)

\(^6\)For the sake of brevity we will abuse terminology slightly and use the term “productivity levels” instead of “distances from the world productivity frontier”.

\(^7\)See Appendix for details.
Let us consider first the case of no externalities. Recall from the previous analysis that the price of the unskilled-intensive good always rises with $H/L$ while the price of the skilled-intensive good always falls (because greater relative supply of skills increases the relative supply of skilled-intermediates), i.e. $\frac{d\ln(P_L)}{d(H/L)} > 0$ and $\frac{d\ln(P_H)}{d(H/L)} < 0$. Figure 1(b) shows that if there is no externality from the skilled workers on the productivity of research ($\psi = 0$), the productivity of skilled labor is strictly decreasing in its relative supply since $\frac{d\ln(\mu H)}{d(H/L)} < 0$. The reason for this is that our method for eliminating the scale effect—normalizing research productivity by $H$—also neutralizes the market size effect. Thus as the relative supply of skilled workers rises, the price of the good they produce falls but there is no offsetting effect from the growing size of the market for skill-specific machines, which makes $H$-specific innovation less attractive. The level of unskilled-specific productivity $A_L$ is also strictly decreasing with $H/L$ in this case. This happens because while the price of unskilled-specific intermediate good $P_L$ rises with $H/L$, the profit per unit of research expenditure falls faster due to the scale-effect eliminating normalization. Since both productivity levels decrease as the proportion of skilled workers rises, the possible configuration of skill specific productivity levels are given by the non-negatively sloped curve in Figure 1(f): as $H/L$ increases the economy moves down and to the left along this curve.

Now consider, what happens when higher proportion of skilled workers makes research more productive ($\psi = 1$; left column of Figure 1). In this case, the productivity of skilled workers is always increasing in $H/L$. This is because the external effect on research productivity offsets the negative effect on $P_H$. The effect of skilled workers supply on the productivity of unskilled workers can also be positive for low levels of $H/L$. This is due to the combined effect of the rising price $P_L$ and externality effect, which make innovation, including for unskilled workers, more attractive. The strength of both of these effect however diminishes with the proportion or skilled in population and for high enough levels researching unskilled-technologies is again less attractive and the BGP productivity level falls. The corresponding, “production possibilities frontier” is depicted in Figure 1(e). In this case, starting from a low level, increasing the share of skilled workers increases the balanced growth path productivity in both sectors. For economies with higher shares of skilled work force however there emerges a trade-off: further increases in the proportion of skilled workers means higher skilled productivity but lower unskilled productivity, as depicted by the downward sloping portion of the curve.

Finally, note that the preceding analysis held the level of barriers to innovation $\zeta$ and

---

8These hod for $\sigma > 1$, which we assume throughout.

9As the appendix shows, $\lim_{H/L \to \infty} \frac{1}{\beta} \frac{d\ln(P_L)}{d(H/L)} = \frac{1+\phi}{1+\phi\sigma}$, which – under our assumption of $\sigma > 1$ – is
investment distortions $\tau$ constant. An increase in both of these results in a decrease in innovation as described before. In the current context, such increases would result in a shift of the frontier depicted in Figure 2.

There are two important conclusions we draw from the above analysis. First, it is possible for a trade-off between between $A_L$ and $A_H$ to be observed in the data. As long as the externality is large enough, countries with higher $H/L$ will have higher skilled labor productivity but low unskilled labor productivity. This result mirrors the finding of Caselli and Coleman’s (2006) finding of a trade-off between productivity of human and physical capital across countries discussed above. Second and related point is that because we find Caselli and Coleman’s results convincing and we take the case where increases in the proportion of skills workers result in a smaller gasp to the world skill-productivity frontier (i.e. $\frac{d\ln(\mu H)}{d\ln(H/L)} > 0$) to be more plausible, we will impose externalities in innovation in our empirical computations.

\[10\text{For } \psi \in (0, 1) \text{ there are intermediate cases; we ignore them here.}\]
Figure 1: Production possibilities frontier: the relationship between distance to the high-skilled frontier ($\mu_H$) and the low-skilled frontier ($\mu_L$) for different level of relative skill supply $H/L$. The left column shows the case of $\psi = 1$; the right column shows the case of $\psi = 0$. 

(a) Elements of $\frac{d\ln(\mu_H)}{d\ln(H/L)}$ for $\psi = 1$. 
(b) Elements of $\frac{d\ln(\mu_H)}{d\ln(H/L)}$ for $\psi = 0$. 
(c) Elements of $\frac{d\ln(\mu_L)}{d\ln(H/L)}$ for $\psi = 1$. 
(d) Elements of $\frac{d\ln(\mu_L)}{d\ln(H/L)}$ for $\psi = 0$. 
(e) Production possibilities frontier: $\mu_H$ vs. $\mu_L$ for $\psi = 1$. 
(f) Production possibilities frontier: $\mu_H$ vs. $\mu_L$ for $\psi = 0$. 

The left column shows the case of $\psi = 1$; the right column shows the case of $\psi = 0$. 

17
Figure 2: The effect of an increase in \( \zeta \) or \( \tau \) on the production possibilities frontier (relationship between distance to the high-skilled frontier (\( \mu_H \)) and the low-skilled frontier (\( \mu_L \)) for different level of relative skill supply \( H/L \)). The left panel shows the case of \( \psi = 1 \); the right panel shows the case of \( \psi = 0 \).
Cross Country Comparisons

Using the above results it is easy to show that

\[ Y_H = \frac{1}{1 - \beta} \left( \frac{P_H}{\mu R_H} \right)^{(1-\beta)/\beta} \mu_H A_H^W H \]

\[ Y_L = \frac{1}{1 - \beta} \left( \frac{P_L}{\mu R_L} \right)^{(1-\beta)/\beta} \mu_L A_L^W L \]

recalling that wages are equal to the marginal product of labor \( w_i = P_i \frac{dy_i}{di} \) for \( i = H, L \) it follows that the US workers’ wage relative to that in a country \( k \) is given by

\[
\frac{w_{H,US}}{w_{H,k}} = \left( \frac{P_{H,k}}{P_{H,US}} \right)^{1+\phi \over \phi} \left( \frac{\zeta_k}{\zeta_{US}} \right)^{1/\phi} \left( \frac{R_{k}}{R_{US}} \right)^{(1-\beta)(1+\phi) \over \beta \phi} \left( \frac{H_{US}}{H_{k}} \right)^{1-\lambda \over \phi} \left( \frac{s_{H,US}}{s_{H,k}} \right) \psi \] (42)

\[
\frac{w_{L,US}}{w_{L,k}} = \left( \frac{P_{L,k}}{P_{L,US}} \right)^{1+\phi \over \phi} \left( \frac{\zeta_k}{\zeta_{US}} \right)^{1/\phi} \left( \frac{R_{k}}{R_{US}} \right)^{(1-\beta)(1+\phi) \over \beta \phi} \left( \frac{H_{US}}{H_{k}} \right)^{1-\lambda \over \phi} \left( \frac{H_{US}/L_{US}}{H_{k}/L_{k}} \right)^{-1} \left( \frac{s_{H,US}}{s_{H,k}} \right) \psi \] (43)

Suppose that in the long run all countries converge to the same \( H/L \) ratio and in addition (to simplify the formulas) ignore depreciation. In that case these equations collapse to

\[
\frac{w_{H,US}}{w_{H,k}} = \left( \frac{\zeta_k}{\zeta_{US}} \right)^{1/\phi} \left( \frac{1 - \tau_{US}}{1 - \tau_k} \right)^{(1-\beta)1+\phi \over \beta \phi} \left( \frac{H_{US}}{H_{k}} \right)^{1-\lambda \over \phi} \]

\[
\frac{w_{L,US}}{w_{L,k}} = \left( \frac{\zeta_k}{\zeta_{US}} \right)^{1/\phi} \left( \frac{1 - \tau_{US}}{1 - \tau_k} \right)^{(1-\beta)1+\phi \over \beta \phi} \left( \frac{H_{US}}{H_{k}} \right)^{1-\lambda \over \phi} \]

That is the relative wages depend on barriers to entry for innovators, relative capital tax and the relative size of the skilled labor force. The latter is the scale effect. Recall that when \( \lambda = 1 \) there is no scale effect and the last term disappears.
4 Empirical Approach

In this section we take the model developed above to the data. Our goal is to back out the skill-specific technology levels ($A_i$’s) and use them to (a) study the historical patterns of directed technological change and (b) evaluate the contribution of non-neutral technology to cross-country income differences. To achieve this goal we use data on output ($Y$), capital ($K$) and labor supply by three skill categories (high school dropouts, high school and college) from Baier et al. (2011).

4.1 Raw data from Tamura, Dwyer, Devereux and Baier 2011

In this section we present the graphs for output per worker, physical capital per worker, schooling exposure per worker for nine geographic regions in the world. These are the inputs used to produce measures of directed technological levels as well as earnings by schooling category.

[output graph goes here]
4.2 Algorithm for Computing Productivity Levels and the World Technology Frontier

In order to back out the skill-specific technology levels \((A_i's)\) we assume that the economies are on the balanced growth path and proceed as follows:

1. First, we pick values for the following parameters: \(\sigma, \phi, \mu, \lambda, \gamma's,\) and \(\beta\). See next section for details of parameter value choices.

2. In the first iteration we assume \(\frac{A_W}{A_P} = \frac{A_S}{A_P} = 1\).

3. We solve for \(\eta's\) using the the Katz and Goldin data and the following equations

\[
\frac{w_H}{w_S} = \left( \frac{\eta_H}{\eta_S} \right)^{\frac{\sigma-1}{1+\phi}} \left( \frac{\gamma_H}{\gamma_S} \right)^{\frac{\sigma(\sigma+\phi)}{\sigma(1+\phi)}} \left( \frac{s_H LF}{s_S LF} \right)^{\frac{\sigma-2-\phi}{1+\phi}} \left( \frac{A_H}{A_W} \right)^{\frac{\phi(\sigma-1)}{1+\phi}}, \tag{44}
\]

\[
\frac{w_H}{w_P} = \left( \frac{\eta_H}{\eta_P} \right)^{\frac{\sigma-1}{1+\phi}} \left( \frac{\gamma_H}{\gamma_P} \right)^{\frac{\sigma(\sigma+\phi)}{\sigma(1+\phi)}} \left( \frac{s_H LF}{s_P LF} \right)^{\frac{\sigma-2-\phi}{1+\phi}} \left( \frac{A_H}{A_W} \right)^{\frac{\phi(\sigma-1)}{1+\phi}}. \tag{45}
\]
Figure 5: Share Exposed to Less than Secondary Schooling: by Region

where $H$ 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)

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

$$\frac{A_i}{A_p} = \left( \frac{\eta_i}{\eta_p} \right)^{\frac{\sigma}{1+\rho}} \left( \frac{\gamma_i}{\gamma_p} \right)^{\frac{\epsilon - 1}{\epsilon + \rho}} \left( \frac{s_i}{s_p} \right)^{\frac{\sigma - 1}{1+\rho}} \left( \frac{A_i^W}{A_p^W} \right)^{\frac{\rho \sigma}{1+\rho}}$$

for $i = C, S$.

5. We then compute $A_p, A_s$, and $A_H$ as follows:

$$Y = \frac{1}{1-\beta} (A_{LN})^\beta K_P^{1-\beta} \times \left\{ \frac{\gamma_H}{(K_H/K_P)^{1-\beta}} \left( \frac{A_{CS}}{A_p} \right)^{\frac{\epsilon - 1}{\epsilon + \rho}} + \frac{\gamma_S}{(K_S/K_P)^{1-\beta}} \left( \frac{A_{SS}}{A_p} \right)^{\frac{\epsilon - 1}{\epsilon + \rho}} + \frac{\gamma_P}{s_P} \left[ s_P^\beta \right]^{\frac{\epsilon - 1}{\epsilon + \rho}} \right\}$$
where

\[ s_i = \frac{N_i}{N} \]

furthermore, since and the fact that

\[ K = K_P \left( 1 + \frac{K_C}{K_P} + \frac{K_S}{K_P} \right) \]

\[ Y = \frac{1}{1-\beta}(A_LN)^\beta K^{1-\beta} \times \]

\[ \left\{ \frac{\gamma_H \left[ \left( \frac{K_H}{K_P} \right)^{1-\beta} \left( \frac{A_{iC}}{A_P} \right) \right]^{\frac{\epsilon-1}{\epsilon}} + \gamma_S \left[ \left( \frac{K_S}{K_P} \right)^{1-\beta} \left( \frac{A_{iS}}{A_P} \right) \right]^{\frac{\epsilon-1}{\epsilon}} + \gamma_P [s_P]^{\frac{\epsilon-1}{\epsilon}} \left( 1+\frac{K_C}{K_P} + \frac{K_S}{K_P} \right)^{1-\beta} }{\left( 1+\frac{K_C}{K_P} + \frac{K_S}{K_P} \right)^{1-\beta}} \right\}^{\frac{\epsilon}{\epsilon-1}} \]

and since from (18) and (32) we know that \( \frac{A_i}{A_P} = f \left( s_C, s_S, s_P, \frac{A_{iC}}{A_P}, \frac{A_{iS}}{A_P} \right) \), and \( \frac{K_i}{K_P} = g \left( s_C, s_S, s_P, \frac{A_{iC}}{A_P}, \frac{A_{iS}}{A_P} \right) \) for \( i = C, S \).

\[ \frac{Y}{N} = A_P^\beta \left( \frac{K}{N} \right)^{1-\beta} \Omega \left( s_C, s_S, s_P, \frac{A_{iC}}{A_P}, \frac{A_{iS}}{A_P} \right) \]

\[ A_p = \left( \frac{(Y/N)/(K/N)^{1-\beta}}{\Omega \left( s_C, s_S, s_P, \frac{A_{iC}}{A_P}, \frac{A_{iS}}{A_P} \right)} \right)^{1/\beta} \]

We then use equations equivalent to (32) to compute \( A_C \) and \( A_S \).

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

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

\[ A_{i(t+1)}^W = \max(A_{ih}(A_{i(t)}^W, D)| h \leq t) \]

where \( D \) stands for our data, \( A_{ih} \) 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.

8. We normalize the level of barriers to entry in the US to be equal to one and using the equation (36) we get

$$\frac{\zeta_k}{\zeta_{US}} = \zeta_k = \left( \frac{P^*_H,k}{P^*_H,US} \right)^{1/\beta} \left( \frac{R^*_k}{R^*_US} \right)^{\frac{\beta-1}{\beta}} \left( \frac{A_{H,US}}{A_{H,k}} \right) \phi \left( \frac{H_k}{H_{US}} \right)^{1-\lambda} \left( \frac{s_{H,k}}{s_{H,US}} \right)^{\psi}$$

which allows us, using the $A$'s computed previously as well as expressions (34) and (27), to compute the (relative) level of barriers for each country.

### 4.3 Choice of Parameter Values

We have to pick values for the following parameters in our model: $\eta$'s, $\sigma$, $\phi$, $\mu$, $\psi$, $\lambda$, $\gamma$'s and $\beta$. 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. $\beta$ is the capital’s income share and we choose a value of $1/3$, in agreement with Gollin (200x). We follow Kelnow and Rodriguez-Clare (200x), who calibrate a endogenous growth models with technology diffusion, by choosing $\phi = 1$ and $\lambda = 1$. As discussed above, a value of $\psi$ close to 1 ensures that a country’s technology frontier exhibits a trade-off between productivity levels in the different sectors. Since Caselli and Coleman argue countries with high productivity of skilled labor exhibit low productivity of unskilled labor, we choose $\psi = 1$. We use a values of 1.4 for the markup based on work of Ramey and X (2012) and Jones and Williams (200x).

The choice of $\gamma$’s turns out not to matter much for our main exercise (except perhaps for the levels of $A$’s): they only affect the level of the relative productivity paths, but not its shape. We’ll start with 1/3 for each.

The choice of $\sigma$ is important for our results so we are especially careful. Based on estimates of Katz and Murphy (1992) and more recently Perri and ?? is about 1.4 or 1.6, but could be above 2. More recent estimates suggest even higher values. We undertake some estimation of our own, described in the next section. In the end we pro dive results for $\sigma$ of 1.6, 2.6 and 3.1.

We calculate the relative research efficiencies $\eta_C/\eta_P$ and $\eta_{HS}/\eta_P$ using the data from Katz and Goldin (200x), who report relative wages and supplies of workers with different educational attainment for the US economy since 1910. Their data is based on censuses
and.....

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

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

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

\[
\frac{w_H}{w_L} = \left( \frac{\eta_H}{\eta_L} \right)^{\frac{\sigma-1}{\sigma+\phi}} \left( \frac{\gamma_H}{\gamma_L} \right)^{\frac{\sigma-2-\phi}{\sigma+2+\phi}} \left( \frac{A_H}{A_L} \right)^{\frac{2(\sigma-1)}{1+\gamma_1}}
\]

we recover the relative \(\eta\)’s.

4.3.1 Estimating \(\sigma\) using EU KLEMS

In this section we used 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 exercise has been performed before, mainly using the US data. Katz and Murphy (1992) start with an equation equivalent to our

\[
\frac{w_H}{w_L} = \left( \frac{\gamma_H}{\gamma_L} \right)^{\frac{\sigma-1}{\sigma+\phi}} \left( \frac{A_H}{A_L} \right)^{\frac{\sigma-1}{\sigma+2+\phi}} \left( \frac{H}{L} \right)^{\frac{\gamma_1}{\sigma+\phi}}
\]

Taking logs and assuming that \(A_H/A_L\) (skill-bias of technology) is growing at a smooth exponential rate \(\gamma_1\)

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

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 and the assumption is. Katz and Murphy estimate the above using data on college/high school wage premium for the hers 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 $\sigma$ close to 1.5. Most recently however, Autor and Acemoglu (2011) argue that higher values of $\sigma$ are also plausible. For example, using Katz and Murphy’s regression on updated sample they find $\sigma = 2.9$.

Here we estimate $\sigma$ using a EU KLEMS panel dataset. Before proceeding however, notice that according to our model – once the endogenous direction of technological change is taken into account – the relative wages are given by

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

(47)

Taking logs and assuming that $A_H^W/A_L^W$ (skill-bias of world technology frontier) is growing at a smooth exponential rate $\gamma_1$

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

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 $\phi = 1$ we arrive at

$$
\log \left( \frac{w_H}{w_L} \right) = \alpha + \frac{\sigma-1}{1+\sigma} \gamma_1 t + \frac{\sigma-3}{1+\sigma} \log \left( \frac{H}{L} \right)
$$

This the coefficient on the relative skill supplies in the above regression corresponds to $\frac{\sigma-3}{1+\sigma}$ not $-1/\sigma$; these two do not coincide unless $\sigma = 1$.

Tables 1-4 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).. Tables 5-8 present results for college and primary; they are similar. The implied $\sigma$ refers to
\[ \sigma_1 = \frac{\beta + 3}{1 - \beta}, \]  
(48)

\[ \sigma_2 = -\frac{1}{\beta}, \]  
(49)

where \( \beta \) is the coefficient on \( H/L \) in the regression. The standard errors on the implied \( \sigma \)'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 2.3 – 3.1. The estimates using the \( \sigma_1 \) interpretation of \( \beta \) are also much closer together; using the \( \sigma_2 \) interpretation the estimates are all over the place and very implausible.

Conclusion: taken to together with the Autor and Acemoglu results this suggest we use a value of \( \sigma \) close to 3. We choose two values: 2.6 and 3.1 and show that calibrated wages are much more reasonable for the case of \( \sigma = 3.1 \) so we make this our preferred value.

<table>
<thead>
<tr>
<th>Table 2: College and High School (OLS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N_{C/N_{HS}} )</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Constant</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Trend Squared</td>
</tr>
<tr>
<td>Country Trend</td>
</tr>
<tr>
<td>Country Trend Sq</td>
</tr>
<tr>
<td>Implied ( \sigma ) (1)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Implied ( \sigma ) (2)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>( R^2 )</td>
</tr>
<tr>
<td>( N )</td>
</tr>
</tbody>
</table>
Table 3: College and High School (OLS)

<table>
<thead>
<tr>
<th>$N_C/N_{HS}$</th>
<th>0.030</th>
<th>0.029</th>
<th>−0.086</th>
<th>−0.088</th>
<th>−0.224**</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(0.045)</td>
<td>(0.044)</td>
<td>(0.058)</td>
<td>(0.055)</td>
<td>(0.101)</td>
</tr>
<tr>
<td>Constant</td>
<td>3.710</td>
<td>1707.357*</td>
<td>2.719</td>
<td>690.185</td>
<td>766.366</td>
</tr>
<tr>
<td></td>
<td>(5.115)</td>
<td>(806.414)</td>
<td>(5.911)</td>
<td>(510.232)</td>
<td>(592.130)</td>
</tr>
<tr>
<td>Implied $\sigma$</td>
<td>3.122</td>
<td>3.119</td>
<td>2.682</td>
<td>2.676</td>
<td>2.269</td>
</tr>
<tr>
<td></td>
<td>(0.189)</td>
<td>(0.189)</td>
<td>(0.196)</td>
<td>(0.187)</td>
<td>(0.269)</td>
</tr>
<tr>
<td>Trend Squared</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Country Trend</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Country Trend Sq</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

| Implied $\sigma$ (1) | 2.68*** | 2.67*** | 2.30*** | 2.26*** | 2.23*** |
|                       | (.191)  | (.183)  | (.191)  | (.262)  | (.186)  |
| Implied $\sigma$ (2) | 11.42   | 11.21   | 4.68*** | 4.48**  | 4.24*** |
|                       | (7.368) | (6.85)  | (1.55)  | (1.94)  | (1.28)  |

| $R^2$ | 0.031 | 0.069 | 0.829 | 0.835 | 0.949   |
|       | (0.056)| (0.054)| (0.071)| (0.099)| (0.071)|
| $p(\sigma < 1.6)$ | 0.000 | 0.000 | 0.000 | 0.000 | 0.003   |
| $p(\sigma < 2.7)$ | 0.013 | 0.013 | 0.537 | 0.550 | 0.912   |
| $p(\sigma < 3.0)$ | 0.258 | 0.264 | 0.923 | 0.934 | 0.979   |
| N    | 320   | 320   | 320   | 320   | 320     |

Table 4: College and High School (Fixed Effects)

<table>
<thead>
<tr>
<th>$N_C/N_{HS}$</th>
<th>−0.088</th>
<th>−0.089</th>
<th>−0.213***</th>
<th>−0.225**</th>
<th>−0.235***</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(0.056)</td>
<td>(0.054)</td>
<td>(0.071)</td>
<td>(0.099)</td>
<td>(0.071)</td>
</tr>
<tr>
<td>Constant</td>
<td>2.666</td>
<td>687.815</td>
<td>−4.705</td>
<td>767.240</td>
<td>627.046***</td>
</tr>
<tr>
<td></td>
<td>(5.732)</td>
<td>(500.970)</td>
<td>(3.326)</td>
<td>(577.981)</td>
<td>(107.634)</td>
</tr>
<tr>
<td>Trend Squared</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Country Trend</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Country Trend Sq</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

| Implied $\sigma$ (1) | 2.68*** | 2.67*** | 2.30*** | 2.26*** | 2.23*** |
|                       | (.191)  | (.183)  | (.191)  | (.262)  | (.186)  |
| Implied $\sigma$ (2) | 11.42   | 11.21   | 4.68*** | 4.48**  | 4.24*** |
|                       | (7.368) | (6.85)  | (1.55)  | (1.94)  | (1.28)  |

| $R^2$ | 0.190 | 0.218 | 0.730 | 0.758 | 0.876   |
|       | (0.513)| (0.511)| (1.037)| (1.038)| (1.038)|
| N    | 320   | 320   | 320   | 320   | 320     |

28
### Table 5: College and High School (GMM)

<table>
<thead>
<tr>
<th></th>
<th>$N_C/N_{HS}$</th>
<th>Trend Squared</th>
<th>Country Trend</th>
<th>Country Trend Sq</th>
<th>Implied $\sigma$ (1)</th>
<th>Implied $\sigma$ (2)</th>
<th>$R^2$</th>
<th>$N$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.022**</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>3.09***</td>
<td>−44.76**</td>
<td></td>
<td>269</td>
</tr>
<tr>
<td></td>
<td>0.022*</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>3.09***</td>
<td>−44.59*</td>
<td></td>
<td>269</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.011)</td>
<td>(0.017)</td>
<td>(0.017)</td>
<td>(0.031)</td>
<td>(0.011)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.001</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>3.00**</td>
<td>−789.6</td>
<td></td>
<td>269</td>
</tr>
<tr>
<td></td>
<td>0.001</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>3.00***</td>
<td>−874.6</td>
<td></td>
<td>269</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.017)</td>
<td>(0.017)</td>
<td>(0.031)</td>
<td>(0.031)</td>
<td>(0.017)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>−0.081**</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>2.69**</td>
<td>12.29**</td>
<td></td>
<td>269</td>
</tr>
<tr>
<td></td>
<td>(0.031)</td>
<td>(0.031)</td>
<td>(0.031)</td>
<td>(0.031)</td>
<td>(0.031)</td>
<td>(0.031)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 6: College and High School (System GMM)

<table>
<thead>
<tr>
<th></th>
<th>$N_C/N_{HS}$</th>
<th>Trend Squared</th>
<th>Country Trend</th>
<th>Country Trend Sq</th>
<th>Implied $\sigma$ (1)</th>
<th>Implied $\sigma$ (2)</th>
<th>$R^2$</th>
<th>$N$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.003</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>3.01***</td>
<td>−316</td>
<td></td>
<td>285</td>
</tr>
<tr>
<td></td>
<td>0.003</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>3.01***</td>
<td>−316</td>
<td></td>
<td>285</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.011)</td>
<td>(0.011)</td>
<td>(0.009)</td>
<td>(0.009)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.022*</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>3.09***</td>
<td>−45.6**</td>
<td></td>
<td>285</td>
</tr>
<tr>
<td></td>
<td>0.022*</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>3.09***</td>
<td>−45.5**</td>
<td></td>
<td>285</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.011)</td>
<td>(0.0439)</td>
<td>(0.0444)</td>
<td>(0.0439)</td>
<td>(0.0444)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.011)</td>
<td>(0.0439)</td>
<td>(0.0444)</td>
<td>(0.0439)</td>
<td>(0.0444)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

N 269 269 269 269 269 269

5 Results (Preliminary and Incomplete)

For now we focus on the period 1910-2007, since we have no way of backing out the relative $\eta$'s prior to 1910.

5.1 Productivity Levels & the Evolution of Skill Bias

This section presents and discusses the skill-specific productivity levels implied by our model and the data. In particular the patterns of skill (college vs high school, high school vs primary etc.) Figures 6-8 show the evolution of two world frontier productivity levels for each of the three skill groups under alternative assumptions about the elasticity of substitution. Figure ?? plots the (log) of productivity levels ($A$'s) for $\sigma = 3.1$ all the way back to 1820. Figures 9 - translate these numbers into measures of relative skill bias, that is they report (log) of $A_i/A_j$ where $i$ and $j$ are one of the three skill groups. The interesting feature of these graphs is that the early first part of the 20th century appears to have been characterized by technological change biased towards less skilled workers: the productivity of college workers relative to high school workers did not begin to rise until the post WWII period. Additionally, for $\sigma = 3.1$ and $\sigma = 2.7$ the relative productivity of college to primary workers was fairly stable between 1920 and 1960 9it did increase in the $\sigma = 1.6$ calibration.
Figure 6: The (log) of productivity levels ($A$'s); $\sigma = 3.1$
Figure 7: The (log) of productivity levels (A’s); $\sigma = 2.7$
Figure 8: The (log) of productivity levels ($A$'s); $\sigma = 1.6$
Figure 9: Skill bias for $\sigma = 3.1$. 
Figure 10: Skill bias for $\sigma = 2.7$. 

Skill Bias

Log of Relative Frontier Productivity Levels

College/HS

High School/P

College/Primary

Year

1920 1940 1960 1980 2000

−0.5 0.0 0.5 1.0 1.5

Figure 10: Skill bias for $\sigma = 2.7$. 

35
Figure 11: Skill bias for $\sigma = 1.6$. 
5.2 Barriers

In this section we analyzed the measures of barriers to entry $\zeta$ backed out using our empirical approach. As above, we consider three values of the elasticity of substitution between skill types. Additionally, we compute barriers using a factor-neutral version of our model. This model has the same basic set up as the main model with the exception that aggregate output is produced using a Cobb-Douglas production function

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

where $L$ is the labor force and $h$ denotes per worker human capital stock calculated using the Mincerian approach as

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

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 $\phi_P = \phi_S = \phi_T = 0.10$. The factor neutral productivity level $A$ is determined endogenously via a process of R&D subject to diffusion and barriers to entry in analogously to our main model. The details are in the appendix.

Figure 12 shows the smoothed distributions of the log go barriers to entry in 2005 under the three different scenarios for the value of $\sigma$ as well as under factor neutral productivity differences. Since we normalized US barrier to 1, a negative value implies barriers less than in the US. Clearly the low (1.6) value of $\sigma$ and the neutral technology both deliver implausibly high proportion of countries with barriers lower than those in the US. This can also be seen in figures 13(a) - 13(d): with low $\sigma$ or factor neutral technology the relationship between the level of barriers and productivity becomes positive.
Figure 12: Smoothed distribution of the (log) of barriers to entry in 2005.
Figure 13: Output per worker vs. the (log) of barriers to entry in 2005
Figure 14: Barriers to entry in 1960-2005.
Average Barriers Over Time; OECD

Figure 15: Barriers to entry in 1960-2005; OECD countries
5.3 Counterfactuals

In this section we experiment with removing barriers technology adoption to see how much
of the variation in GDP per worker is explained by barriers and how much is due to skill
and capital endowments.

5.3.1 Benchmarks

We now consider a thought experiments of removing all the barriers to innovation, but
keep capital unchanged.\(^\text{11}\) We compare the resulting counterfactual distribution of GDP
per worker with the actual one from the data to gauge the importance of barriers to tech-
nology adoption and factors of production in explaining the dispersion of incomes across
countries. Initially, only allow levels of productivity to respond to the policy change but
not factor inputs. In the next section we also allow physical capital to adjust. For each
experiment, we follow the development accounting literature which looks at the variance in
the counterfactual (log) incomes as a percentage of the variance found in the data, ie. the
variance ratio

\[
VR = \frac{Var(\ln(y^*))}{Var(\ln(y))}
\]

where \(y^*\) is the counterfactual distribution of GDP per worker and \(y\) is the observed
one.

For example, in a standard Cobb-Douglas-based development accounting, researchers
have constructed the counterfactual by endowing all countries with the highest (factor-
neutral) TFP level observed in the world. The resulting variance ratio is the percentage of of
variation in incomes that would remain after productivity differences have been eliminated
and is interpreted as the contribution of everything except productivity – usually taken to
mean physical and human capital – to income differences.\(^\text{12}\) For example, following this
method Hall and Jones (1999) and Caselli (2005) reports that physical and human capital
together explain about 33% of income variation. Allowing for capital vs. labor -specific
productivity (Caselli, 2005) and skilled vs. unskilled -specific productivity (Caselli and
Coleman, 2006) show that the role of factor can increase to as much as 50%. A similar,
result is found by Jerzmanowski (2007) who uses a non-parametric estimate of the world

\(^{11}\)Note: we remove barriers by setting them equal to 1 (the US value) however several countries have
\(\zeta < 1\); we allow them to keep their true, lower \(\zeta\).

\(^{12}\)There is an issue of what to do with covariance terms as well as whether to account for productivity-
induced capital accumulation. We ignore the covariance terms following Caselli (2005). The issue
productivity-induced capital accumulation will be discussed below.
technological frontier. Here we

Recall that

\[ y_i = A_W^L \left( \frac{\mu}{1 - \beta} \right)^{\frac{\beta - 1}{\beta}} \zeta_i^{\frac{1}{1 - \beta}} R(\tau_i) \frac{(\beta - 1)(1 + \phi)}{\phi \beta} \Omega \left( \frac{H_i A_H^W}{L_i A_L^W} \right)^{\frac{1}{\beta}} \]

If the removal of barriers produces a distribution that is not much less dispersed than the observed distribution, our measure of success is close to zero and we conclude that the model of factor-non-neutral productivity differences and barriers to innovation cannot account for much of the cross-country dispersion of income per worker observed in the data. Conversely, if the dispersion of the counterfactual incomes is considerably smaller than that in the data, our measure of success will be closer to one and we will conclude that the model is capable of accounting for a large fraction of dispersion in incomes.

We then compare this measure of success with two benchmarks, both of which assume factor-neutral technology differences. The first one is the standard approach where TFP is calculated assuming factor neutral technology (Cobb-Douglas production function). That is TFP, or \( A \), is backed out from

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

where \( L \) is the labor force and \( h \) denotes per worker human capital stock calculated as

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

\( P_i, S_i \) and \( T_i \) stand for years of primary schooling, years of secondary schooling and years of tertiary schooling, respectively and we assume \( \phi_P = \phi_S = \phi_T = 0.10 \). The counterfactual \( Y^* \) is then computed by endowing every country in the sample with the highest (usually the US) level of \( A \) and comparing the resulting variance in GDP per worker to that in the data. As in our model this can be done with or without letting the physical capital adjust and we use the appropriate version when comparing our model’s implications to this benchmark.\(^{13}\)

This is the canonical approach to measuring the importance of productivity differences in explaining cross-country income dispersion. However, since our model departs from the canonical approach by explicitly modeling endogenous technological change, we introduce another benchmark: a factor neutral version of our model. In this version, detail of which

\(^{13}\)Allowing capital to adjusts is equivalent to transforming the production function and keeping \( K/Y \) ratio constant. See Caselli (2004) for details.
can be found in the Appendix, final output is produced according to

\[ Y = \frac{1}{1-\beta} \int_0^A \chi_i^{1-\beta} di (hL)^\beta, \]

where \( L \) is the labor force and \( h \) denotes per worker human capital stock and is calculated as above. Innovation follows a one-sector version of the same process we use in our model

\[ \dot{A}_i = \dot{\eta}_i \left( \frac{A_i^{WR}}{A_i} \right)^\phi Z_i \left( \frac{h_i L_i}{h_i L_i} \right)^\lambda \]

where represents \( A^{WR} \), \( \dot{\eta}_i \) is the productivity of research effort, and \( Z_i \) is the R&D expenditure on innovation or technology adoption in country \( i \). Again, we assume that research productivity depends on the share of college educated workers \( s_C \)

\[ \dot{\eta}_i = \eta_i (s_{C,i})^\psi \]

Assuming that the cost of entry into innovation is \( \zeta_i \), and \( \lambda = 1 \) (as we do in our the calibration of our model), the BGP distance to the world frontier will be

\[ \mu \equiv \frac{A_i}{A^{WR}} = \left[ \frac{\dot{\eta}(s_C) \left( \frac{\mu-1}{\mu} \right) (\mu R)^{\frac{\phi-1}{\phi}}} {r^* \zeta} \right]^{1/\phi}, \] (50)

We use the reduced for of the production function

\[ Y = \frac{1}{1-\beta} K^{1-\beta} (A_i hL)^\beta, \]

to compute the values of total factor productivity \( A_i \). Again we normalize barriers in the US to be one and compute

\[ \frac{\zeta_i}{\zeta_{US}} = \zeta_k = \left( \frac{R^*_i}{R^*_US} \right)^{\frac{\phi-1}{\phi}} \left( \frac{A_{US}}{A_i} \right)^\phi \left( \frac{s_{C,i}}{s_{C,US}} \right)^\psi \]

To compute the counterfactual we set barriers equal to 1, and use equation (48) to back out the level of \( A \). We then use the reduced from output equation (or its \( K/Y \) version) to get the counterfactual \( Y \)'s.
5.4 Results

Figure 17 shows the smoothed distributions of GDP per worker in 2005. The solid line is that data, the dashed line is the counterfactual form the directed technology model when we remove barriers to innovation (but keep physical and human capital stocks unchanged) and the dotted line represents scenario where TFP is calculated assuming factor neutral technology and the maximum is assigned to all countries. The data clearly shows the extent of dispersion on income levels. The mean output per worker in the sample is $19,830 and the median is $10,600 while the standard deviation is $20,200 and then 90/10 percentile ratio is 27. The distribution displays the familiar bi-model distribution emphasized by Jones, Quah and others (kurtosis is 2.77 and skewness = 1.06) Both counterfactual scenarios shift much of the distribution to the right however the effect is larger for the factor neutral technology: the mean increase to $34,500 and median to $32,230 compared with only $29,430 and $24,420, respectively, in the DTC model. Similarly removing barriers to technology in the DTC model leaves the distribution less skewed than the data, with kurtosis of 2.24, skewness of 0.6 and the 90/10 percentile range equal to 9.3, but more than the factor neutral model where kurtosis falls to 2.0 and skewness to 0.44 while the 90/10 percentile range drops to 5.25.
The following figures show how successful the following three scenarios are in accounting for observed variation of log output per worker in 2005: (a) barriers to innovation are removed in the directed technology model (DTC; solid line), (b) in the factor neutral TFP model the maximum TFP is given to every country (dashed line), and (c) we give the frontier technology from the DTC model to each country. Actual number plotted is one minus the variation in the counterfactual log GDP per worker relative to the actual variation in the data.

Figure 18 shows this for the group of OECD countries (regions 1 and 2) from 1820 to 2005. There are several things to note about this graphs. First, in the early period of the sample, roughly between 1820 and 1850, barriers to innovation explained only about 40% of the observed differences in worker productivity. In the next 40 years the importance of barriers rises. On the one hand this is a period of rapid emergence of new technologies (electricity, etc.) on the other hand it first wave of globalization with all the trade and technology diffusion. The result suggest that despite this increased interactions among countries, barriers to innovation and adoption of technology became more important during this period... From 1900 until 1950 the importance of barriers didn’t change much; they accounted, according to the model, for between 70 and 80% of the OECD income differences.
Figure 18: The fraction of variation in log GDP per worker accounted for when (a) barriers to innovation are removed in the directed technology model (DTC; solid line), (b) in the factor neutral TFP model the maximum TFP is given to every country (dashed line), and (c) we give the frontier technology from the DTC model to each country. OECD countries only.

This fraction drops rapidly in the post WWII period to reach a about 50 - 60% by the end of the 20th century.
Figure 19: Fraction of observed variation of log output per worker in 2005 accounted for when barriers to innovation are removed in the directed technology model under different values of $\sigma$. 
6 Emigration to the US: Robustness

In this section we present some evidence on the robustness of our results. We use emigration numbers from countries to the US contained in *Historical Statistics of the United States: Millennial Edition* to compute annualized rates of emigration to the US for multiple countries. We use the computations above to construct an average earnings for country i in year t, as well as the expected gain in earnings for a typical worker in country i in year t. We then compare these with the rates of emigration in the following time period.\(^\text{14}\) Expected earnings are given by:

\[
\text{Expected Earnings} = \bar{y}_{it} = s_H(i,t)w_H(i,t) + s_S(i,t)w_S(i,t) + s_p(i,t)w_p(i,t) \quad (51)
\]

where the shares are from the TDDB data, and the earnings are computed from above. The expected wage gain is given by:

\[
\text{Expected Earnings Gain} = \Delta \bar{y}_{it} = s_H(i,t)\Delta w_H(i,t) + s_S(i,t)\Delta w_S(i,t) + s_p\Delta w_p(i,t) \quad (52)
\]

\[
\Delta w_H(i,t) = w_H(US,t) - w_H(i,t) \quad (53)
\]

\[
\Delta w_S(i,t) = w_S(US,t) - w_S(i,t) \quad (54)
\]

\[
\Delta w_p(i,t) = w_p(US,t) - w_p(i,t) \quad (55)
\]

Finally we computed the annualized emigration rate from country i between year t and t+k-1 as:

\[
\text{Annualized Emigration Rate} = \text{emrate} = \frac{\sum_{j=t}^{t+k} \frac{n_j}{\text{pop}_{it}k}} \quad (56)
\]

We report the results of regressions on the following specification:

\[
\ln(\text{emrate}_{it}) = \alpha + \beta_1\ln(\bar{y}_{it}) + \beta_2\ln(\Delta \bar{y}_{it}) + \beta_3\ln(\text{pop}_{it}) \quad (57)
\]

Tables 3 and 4 report the results for both \(\sigma = 1.6\) and \(\sigma = 2.03\).

Perhaps not surprisingly, there is not much difference between the two results. However it is clear that \(\ln(\Delta \bar{y})\) is positively related to future rates of emigration to the US. This is a simple confirmation that our estimates produce useful information about the potential

\(^{14}\)Typically we observe countries every 10 years, so we are generally using 10 years of emigration data to produce an estimate of the annualized rate of emigration to the US.
Table 7: Emigration Rates, $\sigma = 1.6$

<table>
<thead>
<tr>
<th>Variable</th>
<th>$\ln(\text{Annualized emigration rates})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\ln(\text{pop})$</td>
<td>-0.7854*** -0.8140*** -0.4106*** 1.0893***</td>
</tr>
<tr>
<td></td>
<td>(0.0653) (0.0630) (0.1256) (0.2083)</td>
</tr>
<tr>
<td>$\ln(\overline{\Delta y})$</td>
<td>0.3461*** 0.3173 0.4619** 0.4588***</td>
</tr>
<tr>
<td></td>
<td>(0.1033) (0.1959) (0.1821) (0.1706)</td>
</tr>
<tr>
<td>$\ln(\overline{y})$</td>
<td>1.1143*** 1.1073*** 0.0967 -0.8499***</td>
</tr>
<tr>
<td></td>
<td>(0.1075) (0.2167) (0.2473) (0.2694)</td>
</tr>
<tr>
<td>$N$</td>
<td>455 455 455 455</td>
</tr>
<tr>
<td>$\overline{R}$</td>
<td>.3881 .4937 .3657 .4668</td>
</tr>
<tr>
<td>decade dummies</td>
<td>no yes yes yes</td>
</tr>
<tr>
<td>fixed effects</td>
<td>no no no yes</td>
</tr>
<tr>
<td>random effects</td>
<td>no no yes</td>
</tr>
</tbody>
</table>

earnings gains of departing from one’s home country to come to the United States. Obviously more will be done to check on robustness of these results, and to see if there is a way to discriminate between the $\sigma = 1.6$ and $\sigma = 2.03$.

7 Conclusion

To be added.
Table 8: Emigration Rates, $\sigma = 2.03$

<table>
<thead>
<tr>
<th>Variable</th>
<th>ln(Annualized emigration rates)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln(pop)</td>
<td>$-0.7905^{<em><strong>}$ $-0.8109^{</strong></em>}$ $-0.3998^{<em><strong>}$ $1.1018^{</strong></em>}$</td>
</tr>
<tr>
<td></td>
<td>(0.0656) (0.0633) (0.1262) (0.2082)</td>
</tr>
<tr>
<td>ln($\Delta y$)</td>
<td>$0.3841^{<em><strong>}$ $0.4715^{</strong>}$ $0.6209^{</em><strong>}$ $0.6337^{</strong>*}$</td>
</tr>
<tr>
<td></td>
<td>(0.1045) (0.2087) (0.1982) (0.1865)</td>
</tr>
<tr>
<td>ln($\bar{y}$)</td>
<td>$1.0790^{<em><strong>}$ $1.1974^{</strong></em>}$ $0.1757$ $-0.7630^{***}$</td>
</tr>
<tr>
<td></td>
<td>(0.1070) (0.2244) (0.2512) (0.2703)</td>
</tr>
<tr>
<td>$N$</td>
<td>455 455 455 455</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.3833 .4895 .3467 .4670</td>
</tr>
<tr>
<td>decade dummies</td>
<td>no yes yes yes</td>
</tr>
<tr>
<td>fixed effects</td>
<td>no no no yes</td>
</tr>
<tr>
<td>random effects</td>
<td>no no yes no</td>
</tr>
</tbody>
</table>

References


Appendices

A  BGP Output per Worker

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

where

\[ s_H = \frac{H}{H + L} = \frac{H}{N} \]

Furthermore, since and the fact that

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

\[
Y = \frac{1}{1 - \beta} (A_L^W N)^\beta K_L^{1-\beta} \left\{ \gamma_H \left( \frac{A_H^W}{A_L^W} \right)^{\frac{(\epsilon-1)}{\epsilon}} \left[ \left( \frac{K_H}{K_L} \right)^{1-\beta} \left( \frac{\mu_H s_H}{s_H} \right)^{\beta} \right]^{\frac{\epsilon-1}{\epsilon}} \right. \\
+ \left. \gamma_L \left[ \left( \frac{\mu_L s_L}{s_L} \right)^{\beta} \right]^{\frac{\epsilon-1}{\epsilon}} \right\} \]

\[
(1 + \frac{K_H}{K_L})^{1-\beta}
\]
and since from (18) and (32) we know that 
\[ \mu_i = \zeta \frac{1}{\phi} R(\tau) \frac{\beta - 1}{\phi} f_i \left( \frac{H}{L}, \frac{A_H^{W}}{A_L^{W}} \right) \]
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

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

or

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

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

\[
\frac{Y}{N} = A_L^{W} \left( \frac{\mu}{1 - \beta} \right)^{\frac{\beta - 1}{\phi}} \zeta^{\frac{1}{\phi}} R(\tau) \frac{(\beta - 1)(1 + \phi)}{\phi \sigma} \Omega \left( \frac{H}{L}, \frac{A_H^{W}}{A_L^{W}} \right) \]

### B Details of the Technology Frontier; PRELIMINARY AND INCOMPLETE

In this section we show that in order for the trade-off between \( A_H \) and \( A_L \) to exist (as in Figure ?? and unlike in Figure ??) we need a relatively high value of \( \psi \).

Here we derive conditions on \( \lambda \) and \( \psi \) that ensure \( \frac{d \ln(\mu_H)}{d \ln(H/L)} > 0 \) and \( \frac{d \ln(\mu_L)}{d \ln(H/L)} < 0 \), i.e. we have a frontier with a trade-off between \( A_H \) and \( A_L \) like in Caselli and Coleman.

Recall from equations (13) and (32) that

\[
\frac{d \ln(P_H/P_L)}{d \ln(H/L)} = \frac{1}{\phi} \left( \psi + 1 - \lambda \right) \left( \frac{1 - \beta}{1 + H/L} + \frac{1}{\beta} \frac{d \ln(P_H)}{d \ln(H/L)} \right)
\]

\[
\frac{d \ln(\mu_L)}{d \ln(H/L)} = \frac{1}{\phi} \left( \psi + 1 - \lambda \right) \left( \frac{1 - \beta}{1 + H/L} - 1 + \frac{1}{\beta} \frac{d \ln(P_L)}{d \ln(H/L)} \right)
\]

In this section we show that in order for the trade-off between \( A_H \) and \( A_L \) to exist (as in Figure ?? and unlike in Figure ??) we need a relatively high value of \( \psi \).

Here we derive conditions on \( \lambda \) and \( \psi \) that ensure \( \frac{d \ln(\mu_H)}{d \ln(H/L)} > 0 \) and \( \frac{d \ln(\mu_L)}{d \ln(H/L)} < 0 \), i.e. we have a frontier with a trade-off between \( A_H \) and \( A_L \) like in Caselli and Coleman.

Recall from equations (13) and (32) that

\[
\frac{d \ln(P_H/P_L)}{d \ln(H/L)} = \frac{\beta(1 + \phi)}{1 + \phi \sigma}
\]

which must mean that

\[
\frac{d \ln(P_H)}{d \ln(H/L)} = \frac{d \ln(P_L)}{d \ln(H/L)} - \frac{\beta(1 + \phi)}{1 + \phi \sigma}
\]

We can also show that
\[
\frac{d \ln(P_L)}{d \ln(H/L)} = \frac{\beta(1+\phi)}{\gamma_L(1+\phi \sigma)} \left( \frac{P_H}{P_L} \right)^{\epsilon-1} \\
\frac{d \ln(P_H)}{d \ln(H/L)} = -\frac{\beta(1+\phi)}{\gamma_H(1+\phi \sigma)} \left( \frac{P_H}{P_L} \right)^{1-\epsilon}
\]

C  Factor-Neutral Model

We then compare this measure of success with two benchmarks, both of which assume factor-neutral technology differences. The first one is the standard approach where TFP is calculated assuming factor neutral technology (Cobb-Douglas production function). That is TFP, or \(A\), is backed out from

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

where \(L\) is the labor force and \(h\) denotes per worker human capital stock calculated as

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

\(P_i, S_i\) and \(T_i\) stand for years of primary schooling, years of secondary schooling and years of tertiary schooling, respectively and we assume \(\phi_P = \phi_S = \phi_T = 0.10\). The counterfactual \(Y^*\) is then computed by endowing every country in the sample with the highest (usually the US) level of \(A\) and comparing the resulting variance in GDP per worker to that in the data. As in our model this can be done with or without letting the physical capital adjust and we use the appropriate version when comparing our model’s implications to this benchmark.\(^{15}\)

This is the canonical approach to measuring the importance of productivity differences in explaining cross-country income dispersion. However, since our model departs from the canonical approach by explicitly modeling endogenous technological change, we introduce another benchmark: a factor neutral version of our model. In this version, detail of which can be found in the Appendix, final output is produced according to

\[
Y = \frac{1}{1-\beta} \int_0^A \chi_i^{1-\beta} d(hL)^\beta,
\]

\(^{15}\)Allowing capital to adjust is equivalent to transforming the production function and keeping \(K/Y\) ratio constant. See Caselli (2004) for details.
where $L$ is the labor force and $h$ denotes per worker human capital stock and is calculated as above. Innovation follows a one-sector version of the same process we use in our model

$$\dot{A}_i = \bar{\eta}_i \left( \frac{A^W}{A_i} \right)^\phi \frac{Z_i}{(h_i L_i)^\lambda}$$

where represents $A^W$, $\bar{\eta}_i$ is the productivity of research effort, and $Z_i$ is the R&D expenditure on innovation or technology adoption in country $i$. Again, we assume that research productivity depends on the share of college educated workers $s_C$

$$\tilde{\eta}_i = \eta_i \left( s_{C,i} \right)^\psi$$

Assuming that the cost of entry into innovation is $\zeta_i$, and $\lambda = 1$ (as we do in our the calibration of our model), the BGP distance to the world frontier will be

$$\mu \equiv \frac{A_i}{A^W} = \left[ \frac{\bar{\eta}(s_C) \left( \frac{\mu - 1}{\mu} \right) (\mu R)^{\frac{\beta-1}{\beta}}} {r^* \zeta} \right]^{1/\phi}, \quad (58)$$

We use the reduced form of the production function

$$Y = \frac{1}{1 - \beta} K^{1-\beta} (A_i h L)\beta,$$

to compute the values of total factor productivity $A_i$. Again we normalize barriers in the US to be one and compute

$$\frac{\zeta_i}{\zeta_{US}} = \zeta_k = \left( \frac{R^*_i}{R^*_{US}} \right)^{\frac{\beta - 1}{\beta}} \left( \frac{A_{US}}{A_i} \right)^\phi \left( \frac{s_{C,US}}{s_{C,i}} \right)^\psi$$

To compute the counterfactual we set barriers equal to 1, and use equation (48) to back out the level of $A$. We then use the reduced from output equation (or its $K/Y$ version) to get the counterfactual $Y$’s.