Adam Alexander Oppenheimer
June 5, 2020

***Productivity Differences* by Daron Acemoglu and Fabrizio Zilibotti: Replication and Extension**

1 Introduction

In this paper I replicate *Productivity Differences* by Daron Acemoglu and Fabrizio Zilibotti then extend their model. In the replication, I re-derive the theoretical results of the model. The extension introduces worker-technology mismatch into their model. Additionally, I include a critical review of their theoretical model.

Productivity Differences by Daron Acemoglu and Fabrizio Zilibotti addresses an issue of key concern in economics: the source of between-country per capita income inequality. The prevailing view at the time this paper was published was that differences in technological knowledge could explain much of this inequality. This view was bolstered by large gaps in measured total factor productivity (TFP) between high-income and low-income countries. Acemoglu and Zilibotti argue that this view is difficult to support, as technological knowledge and ideas spread quickly between countries in practice, and new technologies can easily be imported by low-income countries.

In order to reconcile the seeming contradiction between accessible technology but low TFP in low-income countries, they propose a new multi-country model where all countries have access to the same technology. They distinguish between two sets of countries: *the North* and *the South*. The North is a single, advanced economy. The South is a set of small, less developed economies and is considered in the model as an aggregate. The North has a greater proportion of high-skill workers relative to low-skill workers than the South. Investment in technology is made with respect to the needs of the North, leading the marginal payoff to investment between low- and high-skill industries to become equal for the North. For the South, however, this level of investment is too low in low-skill industries while it is too high in high-skill industries. This misalignment with respect to the needs of the South (what they refer to as *technology-skill mismatch*) leads to notable between-country differences in TFP and output per worker. It also leads to counterintuitive cross-country, cross-industry TFP behavior, where relative to the North, the South is less productive in low-skill industries but is more productive in high-skill industries. This behavior is driven by the equal level of technology but relative scarcity of high-skill workers in the South compared to the North, which drives up high-skill labor prices and production value, and consequently high-skill TFP. They finally verify that these results (including the cross-country, cross-industry TFP behavior) are consistent with the data and give preliminary estimates of the amount of inequality that this can explain using regression analysis.

In this paper, I re-derive their theoretical results. I then propose an extension to their model and derive some initial results. The extension allows for worker-technology mismatch, which is missing from their model. The initial derivations seem to indicate that including worker-technology mismatch does not have any direct implications on the model's results.

The paper is organized as follows: Section 2 outlines the fundamentals of their economic model and critically evaluates its assumptions. Section 3 discusses the theoretical results of their model. Derivations can be found in Appendix A.1. Section 4 includes a general description of the data used in their empirical analysis. Due to time constraints, I was unable to complete the empirical replication. Section 5 includes the initial results for a theoretical extension to their model. Section 6 concludes.

2 Model Setup and Discussion

This section provides the initial setup of the economic model and then discusses its assumptions. All equations in this section apply to both the North and the South. Note that derivations for all results from this section can be found in Appendix A.1.1. The model consists of the North and the South. The North is a large, advanced economy; the South is a set of small, less developed economies that are assumed to be identical and are considered in aggregate. The North has H^n high-skill and L^n low-skill workers. The South has H^s high-skill and L^s low-skill workers. We assume $\frac{H^n}{L^n} > \frac{H^s}{L^s}$.

The model itself is made of consumers, producers who purchase labor and rent capital (referred to as *machines*), and monopolists who build and rent out machines (while Acemoglu and Zilibotti refer to these as *monopolists*, for clarity we will refer to them as *patent owners*).

2.1 Output Aggregate

We consider a Cobb-Douglas output aggregate, with consumption C (price of consumption is normalized to 1 in each period), investment in machines I, expenditure on R&D X, and where sector-level output is given by $y(i)$:

$$C + I + X \leq Y \equiv \exp\left[\int_0^1 \ln(y(i)) \cdot di\right] \tag{1}$$

2.2 Consumers

Each country has a representative consumer with CRRA utility that maximizes over consumption $C(\tau)$ at time τ and discounts at rate ρ:

$$\int_t^\infty \frac{C(\tau)^{1-\sigma} - 1}{1-\sigma} \exp(-\rho(\tau - t)) \cdot d\tau$$

2.3 Producers

Each sector i can produce output through a combination of low- and high-skill labor and machines. We denote the quantity of low- and high-skill labor utilized by sector i by $l(i)$ and $h(i)$, respectively. We denote the quantity of machines of variety v for skill type $z \in \{L, H\}$ utilized by sector i by $k_z(i, v)$.[1] The measure of machines v that is accessible to skill type z is given by N_z. Finally, there are exogenous productivity levels i, $(1 - i)$, and Z. i leads to higher productivity for higher-indexed sectors, while $(1 - i)$ leads to higher productivity for lower-indexed sectors. $Z \geq 1$ gives the relative productivity of high-skilled workers. Together, this gives us the following output in sector i:

$$\begin{aligned} y(i) = &\left[\int_0^{N_L} k_L(i,v)^{1-\beta} \cdot dv\right] \cdot [(1-i) \cdot l(i)]^\beta \\ &+ \left[\int_0^{N_H} k_H(i,v)^{1-\beta} \cdot dv\right] \cdot [i \cdot Z \cdot h(i)]^\beta \end{aligned} \tag{2}$$

Firms are price takers and maximize the following profit function, taking prices $p(i)$, wages w_L and w_H, and rental prices of machines $\chi_L(v)$ and $\chi_H(v)$ as given:

$$\begin{aligned} \max_{\substack{l(i),\, h(i), \\ \{k_L(i,v)\}_{v\in[0,N_L]} \\ \{k_H(i,v)\}_{v\in[0,N_H]}}} \quad & p(i)y(i) - w_L l(i) - w_H h(i) - \int_0^{N_L} \chi_L(v) k_L(i,v) \cdot dv \\ & - \int_0^{N_H} \chi_H(v) k_H(i,v) \cdot dv \end{aligned}$$

[1] While Acemoglu and Zilibotti denote both low- and high-skill technologies as v, each skill type can only use its own associated technology. Consequently, it may be clearer to differentiate technologies as v_L and v_H. However, in the spirit of replication, and given this clarification, I use their notation throughout.

2.4 Patent Owners

Patent owners produce machines of the patent variety, v, that they own. For simplicity, it is assumed machines depreciate instantly. The marginal cost of producing each machine is assumed to be a constant fraction θ for all varieties. A patent owner who owns a patent of variety v maximizes profits for each skill type z by setting prices as follows, using the optimal $k_z(i,v)$ from the producer's problem:

$$\max_{\chi_z(v)} \pi_z(v) = (\chi_z(v) - \theta) \int_0^1 k_z(i,v) \cdot di \tag{3}$$

Patent owners in the North are able to invest in R&D. Denote by $V_z^n(v,t)$ the value of producing a new machine, v, for skill type z at time t, where n denotes the North. Acemoglu and Zilibotti argue that by symmetry, we have $V_z^n(v,t) = V_z^n(t) \quad \forall v$. Finally, $r(\tau)$ denotes the interest rate and $\pi_z^n(\tau)$ gives flow profits at time τ. Together, this gives us the following expression:

$$V_z^n(t) = \int_t^\infty \exp\left[-\int_t^\tau r(\omega) \cdot d\omega\right] \pi_z^n(\tau) \cdot d\tau$$

The total investment made for technologies usable by skill type z is denoted by X_z, where it costs μ to develop a new variety of machine. This leads to the law of motion of N_z being given as:

$$\dot{N}_z = \frac{X_z}{\mu}$$

2.5 Discussion

Acemoglu and Zilibotti set out to find a model where countries can have the same level of technology but different levels of TFP and output per worker. While they technically succeeded, they made many assumptions on the way.

There are some assumptions that apply to the paper in general. As Acemoglu and Zilibotti mention, they assume all technology is designed in the North and targets the North's needs.[2] Technology is transferred to the South via knowledge diffusion but not via trade.[3] It is adopted by countries in the South at cost $\epsilon > 0$ (this ensures Bertrand competition does not arise, as that would cause negative profits, meaning the South also has local monopolists and firms solve the same profit maximization problem).[4] Technological innovation does not happen in the South because each country is individually too small for firms to profit from patents, as patent owners cannot sell their machines to other countries. They mention that while this fits the data, the reason why most innovation happens in developed countries is for reasons such as property rights and other distortions, but that this simplification captures the overall effect.[5] As they mention, this is to focus on how biased technological development impacts between-country TFP and output per worker. The main mechanism of this bias is between-country disparities in the proportion of low- and high-skill labor. They claim that this may be a primary factor in the generation of between-country inequalities and may be more important than capital-labor ratios or size of plants, as discussed in previous literature.[6] These assumptions should be taken at face value: they are meant to outline a model that can address the question of the importance of relative availability of labor types between countries. To help answer this question, these assumptions seem reasonable.

Moving to their economic model: initially, it is very general. They use conventional models for both the output aggregate and consumers. Their contributions begin with the introduction of producers. From their description, the setup of the producer's problem allows for a good from sector i to be produced using a combination of low- and high-skill labor, each of which utilizes a variety of machines.[7] In their words, "[t]he

[2] Acemoglu and Zilibotti (2001), p. 563-564
[3] Acemoglu and Zilibotti (2001), p. 568
[4] Acemoglu and Zilibotti (2001), p. 574
[5] Acemoglu and Zilibotti (2001), p. 573
[6] Acemoglu and Zilibotti (2001), p. 566
[7] Note: they show that in equilibrium, each sector uses either low- or high-skill labor, but not both; however, this does not

key assumption is that some machines can only be used by unskilled workers, while some other machines can only be used by skilled workers."[8] I believe this assumption warrants more discussion than Acemoglu and Zilibotti provide. For clarity, the producer's output is reproduced below:

$$y(i) = \left[\int_0^{N_L} k_L(i,v)^{1-\beta} \cdot dv\right] \cdot [(1-i) \cdot l(i)]^\beta + \left[\int_0^{N_H} k_H(i,v)^{1-\beta} \cdot dv\right] \cdot [i \cdot Z \cdot h(i)]^\beta$$

We first attempt to understand what the authors meant by the statement above. The most reasonable interpretation seems to be N_L and N_H are measures of technologies that are accessible exclusively to low and high skill workers, respectively. This seems reasonable: in general, low and high skill workers likely utilize different technologies. However, we must also consider the context of this model: at the start of their paper, Acemoglu and Zilibotti discuss an example from the 1960s where an American technological company began a venture with a Japanese and an Indian firm. The Japanese firm ended up succeeding, but the Indian firm failed because the Indian workers lacked the requisite skills to properly utilize the existing technology. This fits in Acemoglu and Zilibotti's framework only if we assume both low and high skill workers work in the same industry, but cannot use the same technology. However, the point of their example appears to be about mismatch between workers and technology, not about mismatch between workers and particular industries.

Is there a way to integrate worker-technology mismatch into the model? One possible alternative to their model would be to allow all skill types to access all varieties of machines, but at a penalty if the variety is not yet optimized for their skill type. Further, there could also be a persistent productivity gap, Z. This is a slight adjustment to the model described by Acemoglu and Zilibotti but is not subject to the previous criticism. This modified output can be written as follows, where $\gamma_z \in [0,1]$ gives the penalty to skill type z of using a technology that has not yet been optimized for its use:

$$y(i) = \left[\int_0^{N_L} k_L(i,v)^{1-\beta} \cdot dv + \mathbb{1}\{N_H > N_L\}\gamma_L \int_{N_L}^{N_H} k_L(i,v)^{1-\beta} \cdot dv\right] \cdot [(1-i) \cdot l(i)]^\beta + \left[\int_0^{N_H} k_H(i,v)^{1-\beta} \cdot dv + \mathbb{1}\{N_L > N_H\}\gamma_H \int_{N_H}^{N_L} k_H(i,v)^{1-\beta} \cdot dv\right] \cdot [i \cdot Z \cdot h(i)]^\beta$$

We consider one more assumption made in the producer's problem. Acemoglu and Zilibotti assume that firms can purchase any variety of machine available to them. This assumption is understandable. However, there is no penalty to using too many varieties, and in equilibrium firms use all varieties available to them. While in regard to the paper's questions of interest this simplification may be harmless, in general it may be preferable to introduce a "complexity penalization," which increases the firm's costs for utilizing too many varieties of machines (this could implemented similarly to the Lasso).

Moving finally to the patent owner's problem, in general the authors' assumptions seem reasonable. It is a relatively simple profit maximization problem. They clarify that they assume machines depreciate instantaneously for simplicity, but reference a previous paper that solves out the problem with slow depreciation. They assume constant marginal cost to produce machines for all varieties of machines. For the question at hand, this is likely a reasonable assumption, as the question does not investigate the effects of variable costs for machinery. One somewhat difficult to understand assumption is that prices for the same variety of machine can differ by skill type. It is possible that this is justified by the assumption that patent owners are selling expensive machines and work with each customer individually, allowing them to engage in price discrimination by labor type. While in general this seems unlikely and it may be more reasonable to assume that costs cannot vary by skill type, the problem ultimately solves such that prices equalize across skill types.

Considering the paper as a whole, their model setup and assumptions seem reasonable. The consumer's and patent owner's problems are relatively standard. However, the producer's problem does not allow for worker-technology mismatch. Including small adjustments that allow low- and high-skill labor to access all

have any bearing on the current discussion.

[8] Acemoglu and Zilibotti (2001), p. 568

available technology would potentially make the model more believable. However, it seems unlikely this would qualitatively change the results (we show this in Section 5). Therefore, it seems that the model is well structured to answer the questions at hand: is it possible to generate between-country inequalities in TFP and output per worker even when access to technology is equal? And if so, how much inequality can this explain?

3 Model Results

3.1 Country-Level Results

This section essentially re-iterates the model results described in Acemoglu and Zilibotti (2001). These results apply to both the North and the South. Note that derivations for all results from this section can be found in Appendix A.1.1. From the firm's problem, we derive the following:

$$\begin{aligned} k_L(i,v) &= [(1-\beta)p(i)((1-i)\cdot l(i))^\beta/\chi_L(v)]^{1/\beta} \\ k_H(i,v) &= [(1-\beta)p(i)(i\cdot Z\cdot h(i))^\beta/\chi_H(v)]^{1/\beta} \end{aligned} \tag{4}$$

As discussed by Acemoglu and Zilibotti, we can see that as prices and employment increase, and as costs decrease, number of machines demanded increases. In particular, they note that the more abundant type of labor will have higher machine demand, causing what they call a "market size effect."

As Acemoglu and Zilibotti explain, (4) defines isoelastic demands with elasticity β, meaning the patent owner's profit maximizing price is given by $\chi_z(v) = \theta/(1-\beta) = \chi$ (note that I re-derive this result). Without loss of generality, they normalize $\theta = \delta^{\beta/(1-\beta)}(1-\beta)^2$. Plugging in gives $\chi = \delta^{\beta/(1-\beta)}(1-\beta)$. They include δ to allow for cross-country differences in the price of capital. They set $\delta = 1$ for the North, where $\delta > 1$ for the South (they argue their results still hold even with $\delta = 1$ in the South, but include this parameter for their empirical estimation).

Plugging machine prices into machine demand from (4), and consequently plugging this into output from (2), we get:

$$y(i) = \delta^{-1}p(i)^{(1-\beta)/\beta}[N_L\cdot(1-i)\cdot l(i) + N_H\cdot i\cdot Z\cdot h(i)]$$

From this result, Acemoglu and Zilibotti explain that increases in N_L and N_H improve productivity for their respective labor type in all sectors. Further, they assert that they will measure skill-bias through the ratio N_H/N_L.

Acemoglu and Zilibotti then consider the equilibrium within a particular country, taking N_L and N_H as given. They initially assert there exists some J such that industries $i \le J$ use only low-skill workers while industries $i \ge J$ use only high-skill workers. While they reference the working version of their paper for the proof, the proof they reference is missing an essential component. A complete proof is included in Appendix A.1.1. With J, we can simplify output to the following:

$$y(i) = \begin{cases} \delta^{-1}p(i)^{(1-\beta)/\beta}N_L\cdot(1-i)\cdot l(i) & \text{if} \quad 0\le i\le J \\ \delta^{-1}p(i)^{(1-\beta)/\beta}N_H\cdot i\cdot Z\cdot h(i) & \text{if} \quad J\le i\le 1 \end{cases} \tag{5}$$

In Appendix A.1.1, we show that $\delta^{-1}p(i)^{1/\beta}N_L\cdot(1-i)$ is constant if $0\le i\le J$ and $\delta^{-1}p(i)^{1/\beta}N_H\cdot i\cdot Z$ is constant if $J\le i\le 1$. From this, we can define $p(i) = P_L\cdot(1-i)^{-\beta}$ for $0\le i\le J$ and $p(i) = P_H\cdot i^{-\beta}$ for $J\le i\le 1$. Further, if we consider $p(i)y(i)$, then all variability in $p(i)y(i)$ must come from $l(i)$ or $h(i)$. However, because (1) is Cobb-Douglas, $p(i)y(i)$ must be constant $\forall i$, meaning $l(i)$ and $h(i)$ are also constant $\forall i$. Considering this, as well as market clearing $\left(\int_0^J l(i)di = L \text{ and } \int_J^1 h(i)di = H\right)$, we have:

$$p(i) = P_L\cdot(1-i)^{-\beta} \quad \text{and} \quad l(i) = L/J \quad \text{if} \quad 0\le i\le J \tag{6}$$

$$p(i) = P_H\cdot i^{-\beta} \quad \text{and} \quad h(i) = H/(1-J) \quad \text{if} \quad J\le i\le 1 \tag{7}$$

Note that P_L is given by $p(0)$ and P_H is given by $p(1)$. Further, prices for low-skill labor increase in technology index, whereas prices for high-skill labor decrease in technology index. To find J, we note that both (6) and

(7) should hold:

$$P_L \cdot (1-J)^{-\beta} = P_H \cdot J^{-\beta}$$
$$\Leftrightarrow \frac{P_H}{P_L} = \left(\frac{J}{1-J}\right)^{\beta} \tag{8}$$

Combining (5) and (8) (this result is derived in Appendix A.1.1), we have

$$J = \left(1 + \left(\frac{N_H}{N_L} \cdot \frac{Z \cdot H}{L}\right)^{1/2}\right)^{-1} \tag{9}$$

We can see that as technology weights towards high-skilled labor, or as the ratio of high- to low-skill laborers increases, J decreases. This means that more industries use high-skill labor. Next, we solve P_L and P_H. Derivations are in Appendix A.1.1. We get:

$$\begin{aligned} P_L &= \exp(-\beta) \cdot \left(1 + \left(\frac{N_H}{N_L} \cdot \frac{Z \cdot H}{L}\right)^{1/2}\right)^{\beta} \\ P_H &= \exp(-\beta) \left(1 + \left(\frac{N_H}{N_L} \cdot \frac{Z \cdot H}{L}\right)^{-1/2}\right)^{\beta} \end{aligned} \tag{10}$$

Similar to J, prices move according to the relative availability of technology and labor: as high-skill technology and labor increase proportionally to low-skill technology and labor, products produced by low-skilled labor become more expensive while products produced by high-skilled labor become less expensive. Moving to relative wages, we derive the following:

$$\frac{w_H}{w_L} = Z \left(\frac{N_H}{N_L}\right)^{1/2} \left(\frac{Z \cdot H}{L}\right)^{-1/2} \tag{11}$$

We see that relative wages are proportional to relative to technology but inversely related to proportion of labor. Finally, we consider aggregate output, $Y = \int_0^1 p(i)y(i) \cdot di$. We derive:

$$Y = \exp(-1) \cdot \delta^{-1} \left((N_L \cdot L)^{1/2} + (N_H \cdot Z \cdot H)^{1/2}\right)^2 \tag{12}$$

3.2 Cross-Country Results

This section describes results that differ by country. First, Acemoglu and Zilibotti consider R&D. The model assumes that patent owners cannot sell their technology. The North has sufficient demand to incentivize R&D. Because the South is a set of small countries, each individual country has insufficient demand to spur investment in R&D. As described in Section 2.4, this causes technologies to be researched in the North and coopted by the South. There is a marginal cost of adopting these technologies in the South, ϵ, that ensures only a single firm enters each market in the South. Otherwise, Bertrand competition would cause all firms to have negative profits. As described in Section 2.4, the present value of a patent for machine v used by skill type z at time t is given by:

$$V_z^n(v,t) = \int_t^{\infty} \exp\left[-\int_t^{\tau} r(\omega) \cdot d\omega\right] \pi_z^n(\tau) \cdot d\tau \tag{13}$$

where n denotes the North, and $r(\tau)$ denotes the interest rate and $\pi_z^n(\tau)$ gives flow profits at time τ. To reiterate, Acemoglu and Zilibotti argue that by symmetry, we have $V_z^n(v,t) = V_z^n(t) \quad \forall v$. We next write simplified expressions for $\pi_z^n(\tau)$, noting that machines are produced in the North accounting for the North's needs:

$$\pi_L^n(\tau) = (\chi^n - \theta)\int_0^{J^n} k_L^n(i)\cdot di = \beta(1-\beta)(P_L^n(\tau))^{1/\beta}L^n$$
$$\pi_H^n(\tau) = (\chi^n - \theta)\int_{J^n}^{1} k_H^n(i)\cdot di = \beta(1-\beta)(P_H^n(\tau))^{1/\beta}Z\cdot H^n \tag{14}$$

Because the market for patents has free entry, Acemoglu and Zilibotti argue that the present value of a patent cannot exceed the cost to develop a new patent in equilibrium, as this will cause new firms to enter, which will subsequently lower the value of any particular patent. Further, they assert that along the BGP, N_L and N_H must grow at the same rate. Because research occurs only in the North, this means that the ratio of research expenditure relative to existing technology must be equal for both skill types. Their line of arguments then asserts this implies $V_L^n = V_H^n = \mu$, and further, that this implies $\pi_L^n = \pi_H^n$. Using this condition with (14) gets us the following:

$$(P_L^n)^{1/\beta}L^n = (P_H^n)^{1/\beta}Z\cdot H^n$$
$$\Leftrightarrow \frac{L^n}{Z\cdot H^n} = \left(\frac{P_H^n}{P_L^n}\right)^{1/\beta} \tag{15}$$

Combining with (8) and (9), we get the following (this is derived in Appendix A.1.2):

$$\frac{N_H}{N_L} = \frac{1-J^n}{J^n} = \frac{Z\cdot H^n}{L^n} \tag{16}$$

From this, we can see that along the BGP the ratio of skilled to unskilled technology is increasing in the ratio of skilled to unskilled workers in the North. Acemoglu and Zilibotti describe this as a market size effect, where technologies that are used by more workers are more profitable to research.

With these results, Acemoglu and Zilibotti can now prove there exists a unique, stable BGP. It is characterized by (6), (7), (8), (9), and (16). Along the BGP, GDP, consumption, N_L, and N_H grow at the following rate:

$$g = (1/\sigma)\cdot[\exp(-1)\cdot\beta\cdot(1-\beta)\cdot\mu^{-1}\cdot(L^n + Z\cdot H^n) - \rho] \tag{17}$$

They show the BGP is unique and any initial values of N_L and N_H converge to the BGP. They additionally assert that net output and consumption in the North, but not in the South, are maximized along the the BGP. This is because technological research is directed toward the needs of the North but not the South.

3.3 Productivity Differences

Once Acemoglu and Zilibotti characterize the equilibrium, they begin to compare the productivity of the North and the South. They define the following:

$$p^c(i)\cdot y_L^c(i) = \alpha_L^c(i)\cdot K_L^c(i)^{1-\beta}\cdot l^c(i)^\beta$$
$$p^c(i)\cdot y_H^c(i) = \alpha_H^c(i)\cdot K_H^c(i)^{1-\beta}\cdot [Z\cdot h^c(i)]^\beta \tag{18}$$

where c gives country, $K_z^c(i) \equiv \int_0^{N_z} k_z^c(i,v)\cdot dv$ gives capital input for industry i, and $\alpha_L^c(i) \equiv p^c(i)\cdot[(1-i)\cdot N_L]^\beta$ and $\alpha_H^c(i) \equiv p^c(i)\cdot[i\cdot N_H]^\beta$ give TFP. Using (6) and (7), we get:

$$\alpha_L^c(i) = \alpha_L^c = P_L^c\cdot N_L^\beta$$
$$\alpha_H^c(i) = \alpha_H^c = P_H^c\cdot N_H^\beta \tag{19}$$

Using (15) and (16), we note that in the North, $\alpha_L^n = \alpha_H^n = \alpha^n$. We also have from (10) that $P_H^s > P_H^n$ and $P_L^s < P_L^n$, which leads into Acemoglu and Zilibotti's main result:

$$\frac{H^s}{L^s} < \frac{H^n}{L^n} \Longrightarrow \alpha_H^s(i \geq J^n) > \alpha^n > \alpha_L^s(i \leq J^s)$$

Therefore, we expect that countries with skill scarcity will have higher productivity in high-skill industries and lower productivity in low-skill industry than countries without skill scarcity. They argue this result is driven by the scarcity itself: skill scarcity increases prices for skill intensive industries, increasing TFP.

Finally, Acemoglu and Zilibotti describe two more measures of productivity. They begin by defining physical productivity as:

$$\frac{a_z^c(i)}{p^c(i)} = \begin{cases} ((1-i)\cdot N_L)^\beta & \text{if} \quad 0 \leq i \leq J^c \\ (i \cdot N_H)^\beta & \text{if} \quad J^c \leq i \leq 1 \end{cases}$$

This measure indicates that physical productivities are equal in the North and South for industries where they use the same technologies - that is, for $i \leq J^n$ and $i \geq J^s$. However, for $i \in [J^n, J^s]$, they show that physical productivity is higher in the North. Using this measure, they define a measure of aggregate TFP, using the following representation of output:

$$\begin{aligned} Y^c &= \exp\left(\int_0^{J^c} \ln(y_L^c(i)) \cdot di + \int_{J^c}^1 \ln(y_H^c(i)) \cdot di\right) \\ &= A(J^c, N_L, N_H) \cdot (K_L^{1-\beta} l^\beta)^J (K_H^{1-\beta}(Z \cdot h)^\beta)^{1-J} \end{aligned} \tag{20}$$

where $A(J^c, N_L, N_H) = \exp\left(\int_0^{J^c} \ln(\alpha_L(i)/p(i)) \cdot di + \int_{J^c}^1 \ln(\alpha_H(i)/p(i)) \cdot di\right)$ gives aggregate TFP. They also define two normalized outputs, output per worker (y^c) and output per efficiency unit of labor ($y^{eff,c}$):

$$\begin{aligned} y^c(H^c, L^c, N_L, N_L|\delta) &\equiv \frac{Y^c}{L^c + H^c} \\ &= \exp(-1) \cdot \delta^{-1} \frac{\left((N_L \cdot L)^{1/2} + (N_H \cdot Z \cdot H)^{1/2}\right)^2}{L^c + H^c} \\ y^{eff,c}(H^c, L^c, N_L, N_L|\delta) &\equiv \frac{Y^c}{L^c + Z \cdot H^c} \\ &= \exp(-1) \cdot \delta^{-1} \frac{\left((N_L \cdot L)^{1/2} + (N_H \cdot Z \cdot H)^{1/2}\right)^2}{L^c + Z \cdot H^c} \end{aligned}$$

The remainder of the paper discusses extensions. In the following section, I describe the data for the empirical results in Acemoglu and Zilibotti's paper. Section 5 derives initial results for a theoretical extension to their model that considers worker-technology mismatch.

4 Data

Historical UN industrial data are retrieved from United Nations Statistics Division (n.d.). Codebooks to interpret the data are retrieved from Economic Statistics and Classification Section, United Nations Statistics Division (2007). Historical data is identical to that used by Acemoglu and Zilibotti (2001). All variables are generated (to the best of my ability) as they describe. Modern UN industrial data are retrieved from United Nations Industrial Development Organization (2020). While historical data is available at the level of 3-digit ISIC codes, the free version of the modern data is available only at the level of 2-digit ISIC codes. Due to time constraints, I was unable to complete this empirical replication.

5 Extension

I propose an extension to Acemoglu and Zilibotti's model that includes worker-technology mismatch. Section 2.5 includes a description of the extension and a justification for its relevance. Worker-technology mismatch is included through a slight adjustment of the model. Rather than having each skill type have independent technologies, technologies are consistent across skill types. N_L gives technologies that have been optimized for low skill workers and N_H gives technologies that have been optimized for high skill workers. Any technology beyond the maximum of N_L and N_H has not yet been developed. However, each skill type can access technologies above its own maximum so long as those technologies have been developed for the other skill type. This allows for worker-technology mismatch, where a penalty is applied for a worker using a technology that has not been optimized for their use. Output in industry i can be described as follows:

$$\begin{aligned} y(i) = &\left[\int_0^{N_L} k_L(i,v)^{1-\beta} \cdot dv + \mathbb{1}\{N_H > N_L\}\gamma_L \int_{N_L}^{N_H} k_L(i,v)^{1-\beta} \cdot dv\right] \cdot [(1-i)\cdot l(i)]^\beta \\ &+ \left[\int_0^{N_H} k_H(i,v)^{1-\beta} \cdot dv + \mathbb{1}\{N_L > N_H\}\gamma_H \int_{N_H}^{N_L} k_H(i,v)^{1-\beta} \cdot dv\right] \cdot [i \cdot Z \cdot h(i)]^\beta \end{aligned} \tag{21}$$

where $\gamma_z \in [0,1]$. If $\gamma_L = \gamma_H = 0$ or $N_L = N_H$, this reduces to Acemoglu and Zilibotti's model. However, we consider a case when $\gamma_L, \gamma_H > 0$ and $N_L \neq N_H$. Assume without loss of generality that $N_H > N_L$ (this is somewhat motivated by (16), which shows that if $H^n > L^n$ then we expect $N_H > N_L$). While at first glance it may appear that we should assume $\gamma_H > \gamma_L$, this may be unjustified - there is no reason that high-skill workers should necessarily be more productive at low-skill jobs than low-skill workers (or vice-versa, if you believe we should assume $\gamma_L > \gamma_H$).

From the firm's problem, we derive the following:

$$\begin{aligned} k_L(i,v) &= [(1-\beta)(1+\gamma_L)p(i)((1-i)\cdot l(i))^\beta / \chi_L(v)]^{1/\beta} \\ k_H(i,v) &= [(1-\beta)p(i)(i \cdot Z \cdot h(i))^\beta / \chi_H(v)]^{1/\beta} \end{aligned} \tag{22}$$

This is relatively similar to Acemoglu and Zilibotti's result, although it includes a scalar. We see that the scalar increases the amount of machines purchased for use by low skill workers.

We next derive machine prices. We find that the results are identical to Acemoglu and Zilibotti's. Plugging machine prices into machine demand from (22), and consequently plugging this into output from (21), we get:

$$y(i) = \delta^{-1} p(i)^{(1-\beta)/\beta} [N_L \cdot (1-\gamma_L) \cdot (1+\gamma_L)^{1-\beta} \cdot (1-i) \cdot l(i) + N_H \cdot (1+\gamma_L)^{2-\beta} \cdot i \cdot Z \cdot h(i)]$$

We see that introducing worker-technology mismatch has competing effects on output. First, we see that it somewhat dampens the effects of increasing N_L. This is because when N_L increases, it does not actually give access to new technologies, it is only making the technologies more productive. At the same time, it enhances the output of N_L by giving low skill workers access to high skill technologies. Second, it enhances the effects of increasing N_H. This is because when N_H increases, it grants access to new technologies to low skill workers in addition to high skill workers. What is particularly noteworthy is that even if $\beta = 1$, we see a relationship between γ_L and the marginal effects of N_L and N_H. This is because it actively affects the amount of technology that low skill workers can access.

We next assume that a unique J exists (the proof is likely identical to Acemoglu and Zilibotti's model). This gets us the following:

$$y(i) = \begin{cases} \delta^{-1} p(i)^{(1-\beta)/\beta} N_L \cdot (1-\gamma_L) \cdot (1+\gamma_L)^{1-\beta} \cdot (1-i) \cdot l(i) & \text{if} \quad 0 \le i \le J \\ \delta^{-1} p(i)^{(1-\beta)/\beta} N_H \cdot (1+\gamma_L)^{2-\beta} \cdot i \cdot Z \cdot h(i) & \text{if} \quad J \le i \le 1 \end{cases} \tag{23}$$

From here, results will be identical to Acemoglu and Zilibotti's model. This is because we can derive P_L and P_H exactly as before. While P_L and P_H will now include the γ_L terms, any further analysis should be unchanged. In essence, we can think of γ_L as being absorbed by P_L and P_H. Therefore, equilibrium results should be unchanged.

6 Conclusion

This paper replicates the main theoretical results of *Productivity Differences* by Daron Acemoglu and Fabrizio Zilibotti. It then proposes an extension to the model that includes worker-technology mismatch. Initial derivations indicate that this does not alter model results.

Productivity Differences proposes a dynamic economic model that includes developed and undeveloped countries, where all countries have access to the same technologies but output differences can still arise. While this model is parsimonious and demonstrates the results that the authors intended, its representation of technology appears to require improvements. The paper is motivated by an example of low skill workers using high skill technology and consequently having low productivity, but this scenario does not fit into the authors' proposed model. Surprisingly, introducing a representation of worker-technology mismatch that can fit this scenario does not appear to change the main results of the model.

That including worker-technology mismatch into Acemoglu and Zilibotti's model does not change the model's fundamental results raises many questions. For instance, is this because of the particular form that worker-technology mismatch takes in this case? Does the particular form of Acemoglu and Zilibotti's model have any effect, or would many similar economic models be unchanged when considering this type of mismatch? Future research should consider these questions.

References

Acemoglu, Daron and Fabrizio Zilibotti, "Productivity Differences," *The Quarterly Journal of Economics*, 2001, *116* (2), 563–606.

Economic Statistics and Classification Section, United Nations Statistics Division, "General Industrial Statistics Dataset," 2007.

United Nations Industrial Development Organization, "INDSTAT 2 2020 - ISIC Revision 3," 2020.

United Nations Statistics Division, "United Nations General Industrial Statistics (Vol 1), 1967-1993."

A Appendix

A.1 Model Results

A.1.1 Country-Level Results

These results apply to both the North and the South. We start by considering the Producer's problem. Firms solve the following problem:

$$\max_{\substack{l(i),\,h(i),\\ \{k_L(i,v)\}_{v\in[0,N_L]}\\ \{k_H(i,v)\}_{v\in[0,N_H]}}} p(i)y(i) - w_L l(i) - w_H h(i) - \int_0^{N_L} \chi_L(v) k_L(i,v)\cdot dv \\ - \int_0^{N_H} \chi_H(v) k_H(i,v)\cdot dv$$

where

$$\begin{aligned} y(i) = &\left[\int_0^{N_L} k_L(i,v)^{1-\beta}\cdot dv\right]\cdot[(1-i)\cdot l(i)]^\beta \\ &+ \left[\int_0^{N_H} k_H(i,v)^{1-\beta}\cdot dv\right]\cdot[i\cdot Z\cdot h(i)]^\beta \end{aligned} \tag{2}$$

Taking FOC with respect to $k_z(i,v)$, we get:

$$\begin{aligned} [k_L(i,v)]: \quad & 0 = p(i)(1-\beta)k_L(i,v)^{-\beta}[(1-i)\cdot l(i)]^\beta - \chi_L(v) \\ \Leftrightarrow & k_L(i,v) = [(1-\beta)p(i)((1-i)\cdot l(i))^\beta/\chi_L(v)]^{1/\beta} \\ [k_H(i,v)]: \quad & 0 = p(i)(1-\beta)k_H(i,v)^{-\beta}[i\cdot Z\cdot h(i)]^\beta - \chi_H(v) \\ \Leftrightarrow & k_H(i,v) = [(1-\beta)p(i)(i\cdot Z\cdot h(i))^\beta/\chi_H(v)]^{1/\beta} \end{aligned} \tag{4}$$

Next, we consider the patent owner's problem. Patent owners solve the following problem:

$$\max_{\chi_z(v)} \pi_z(v) = (\chi_z(v)-\theta)\int_0^1 k_z(i,v)\cdot di \tag{3}$$

Taking FOC with respect to $\chi_z(v)$, we get:

$$\begin{aligned} [\chi_L]: \quad 0 &= \int_0^1 k_L(i,v)\cdot di + (\chi_L(v)-\theta)\int_0^1 \frac{\partial k_L(i,v)}{\partial \chi_L(v)}\cdot di \\ &= \left(\frac{1-\beta}{\chi_L(v)}\right)^{1/\beta}\int_0^1 p(i)^{1/\beta}(1-i)\cdot l(i)\cdot di + (\chi_L(v)-\theta)(1-\beta)^{1/\beta}\left(-\frac{1}{\beta}\right)\chi_L(v)^{\frac{-1-\beta}{\beta}}\int_0^1 p(i)^{1/\beta}(1-i)\cdot l(i)\cdot di \\ &= 1 + (\chi_L(v)-\theta)\left(-\frac{1}{\beta}\right)\chi_L(v)^{-1} \\ \Leftrightarrow 1-\beta &= \frac{\theta}{\chi_L(v)} \\ \Leftrightarrow \chi_L(v) &= \frac{\theta}{1-\beta} \end{aligned}$$

$$
\begin{aligned}
[\chi_H]: \quad 0 &= \int_0^1 k_H(i,v)\cdot di + (\chi_H(v)-\theta)\int_0^1 \frac{\partial k_H(i,v)}{\partial \chi_H(v)}\cdot di \\
&= \left(\frac{1-\beta}{\chi_H(v)}\right)^{1/\beta}\int_0^1 p(i)^{1/\beta} i\cdot Z\cdot h(i)\cdot di + (\chi_H(v)-\theta)(1-\beta)^{1/\beta}\left(-\frac{1}{\beta}\right)\chi_H(v)^{\frac{-1-\beta}{\beta}}\int_0^1 p(i)^{1/\beta} i\cdot Z\cdot h(i)\cdot di \\
&= 1 + (\chi_H(v)-\theta)\left(-\frac{1}{\beta}\right)\chi_H(v)^{-1} \\
\Leftrightarrow 1-\beta &= \frac{\theta}{\chi_H(v)} \\
\Leftrightarrow \chi_H(v) &= \frac{\theta}{1-\beta}
\end{aligned}
$$

We therefore have that $\chi_L(v)=\chi_H(v)=\chi$. As explained in Section 3.1, we set $\theta=\delta^{\beta/(1-\beta)}(1-\beta)^2$, giving us $\chi=\delta^{\beta/(1-\beta)}(1-\beta)$.

Now plugging these prices into (4), we have:

$$
\begin{aligned}
k_L(i,v) &= p(i)^{1/\beta}(1-i)\cdot l(i)\cdot\delta^{1/(\beta-1)} \\
k_H(i,v) &= p(i)^{1/\beta} i\cdot Z\cdot h(i)\cdot\delta^{1/(\beta-1)}
\end{aligned}
$$

Plugging these into (2), we have:

$$
\begin{aligned}
y(i) = &\left[\int_0^{N_L}\left(p(i)^{1/\beta}(1-i)\cdot l(i)\cdot\delta^{1/(\beta-1)}\right)^{1-\beta}\cdot dv\right]\cdot[(1-i)\cdot l(i)]^\beta \\
&+\left[\int_0^{N_H}\left(p(i)^{1/\beta} i\cdot Z\cdot h(i)\cdot\delta^{1/(\beta-1)}\right)^{1-\beta}\cdot dv\right]\cdot[i\cdot Z\cdot h(i)]^\beta \\
= &\ \delta^{-1}p(i)^{(1-\beta)/\beta}[N_L\cdot(1-i)\cdot l(i)+N_H\cdot i\cdot Z\cdot h(i)]
\end{aligned}
$$

We now show that given these results, in a particular country where N_L and N_H are taken as given, there exists a unique $J\in(0,1)$ where all industries lower than J use only low-skill workers, while all industries above J use only high-skill workers. We start by considering firm profits:

$$
\pi(i)=p(i)y(i)-w_L l(i)-w_H h(i)-\int_0^{N_L}\chi_L(v)k_L(i,v)\cdot dv-\int_0^{N_H}\chi_H(v)k_H(i,v)\cdot dv
$$

Plugging in the optimal values for $y(i)$, $k_L(i,v)$, $k_H(i,v)$, $\chi_L(v)$, and $\chi_H(v)$, this gives us:

$$
\begin{aligned}
\pi(i) = &\ \delta^{-1}p(i)^{1/\beta}[N_L\cdot(1-i)\cdot l(i)+N_H\cdot i\cdot Z\cdot h(i)]-w_L l(i)-w_H h(i) \\
&-N_L\cdot\delta^{-1}p(i)^{1/\beta}(1-i)\cdot l(i)\cdot(1-\beta)-N_H\cdot\delta^{-1}p(i)^{1/\beta} i\cdot Z\cdot h(i)\cdot(1-\beta)
\end{aligned}
$$

Denoting profits from workers of type z in firm i as $\pi_z(i)$, and taking per-worker profits, we have

$$
\begin{aligned}
\frac{\pi_L(i)}{l(i)} &= \delta^{-1}p(i)^{1/\beta}N_L\cdot(1-i)-w_L-N_L\cdot\delta^{-1}p(i)^{1/\beta}(1-i)\cdot(1-\beta) \\
&= \beta\delta^{-1}p(i)^{1/\beta}N_L\cdot(1-i)-w_L \\
\frac{\pi_H(i)}{h(i)} &= \delta^{-1}p(i)^{1/\beta}N_H\cdot i\cdot Z-w_H-N_H\cdot\delta^{-1}p(i)^{1/\beta} i\cdot Z\cdot(1-\beta) \\
&= \beta\delta^{-1}p(i)^{1/\beta}N_H\cdot i\cdot Z-w_H
\end{aligned}
$$

Competition implies that in equilibrium, $\pi_L(i),\pi_H(i)\le 0$. As we are using Cobb-Douglas technology in (1), we have that all goods $i\in(0,1)$ are produced. Therefore, we have either $\pi_L(i)=0$, $\pi_H(i)=0$, or both $\forall i$. Finally, we cannot have either type of worker be unemployed, otherwise $w_z=0$ and firms would profit from hiring this labor, meaning this is not an equilibrium outcome. Therefore, each type of labor is used and all goods are produced. Finally, we have that $\frac{\pi_H(i)}{h(i)} - \frac{\pi_L(i)}{l(i)}$ is strictly increasing in i (this is not proven by Acemoglu and Zilibotti, but I prove it below). Taken together, there must be some $J \in (0,1)$ such that $\frac{\pi_H(i)}{h(i)} = \frac{\pi_L(i)}{l(i)}$ for $i = J$, $\frac{\pi_H(i)}{h(i)} < \frac{\pi_L(i)}{l(i)}$ for $i < J$, and $\frac{\pi_H(i)}{h(i)} > \frac{\pi_L(i)}{l(i)}$ for $i > J$.

Now, we prove that $\frac{\pi_H(i)}{h(i)} - \frac{\pi_L(i)}{l(i)}$ is strictly increasing in i:

$$\frac{\pi_H(i)}{h(i)} - \frac{\pi_L(i)}{l(i)} = \beta\delta^{-1}p(i)^{1/\beta}(N_H \cdot i \cdot Z - N_L \cdot (1-i)) - w_H + w_L$$

Taking the derivative with respect to i, we have

$$[i]: \quad \delta^{-1}p'(i) \cdot p(i)^{(1-\beta)/\beta}(N_H \cdot i \cdot Z - N_L \cdot (1-i)) + \beta\delta^{-1}p(i)^{1/\beta}(N_H \cdot Z + N_L)$$

At this point, Acemoglu and Zilibotti assert that the difference in profits per worker is strictly increasing in i. In general, this statement is not true, as we do not know the functional form of $p(i)$; however, we assume $p(i)$ is of the following form (note that this is the form Acemoglu and Zilibotti prove p_i takes, assuming there exists a unique J):

$$p(i) = \begin{cases} P_L \cdot (1-i)^{-\beta} & \text{if} \quad i \leq J \\ P_H \cdot i^{-\beta} & \text{if} \quad i \geq J \end{cases}$$
$$\Rightarrow p'(i) = \begin{cases} \beta \cdot P_L \cdot (1-i)^{-\beta-1} & \text{if} \quad i \leq J \\ -\beta \cdot P_H \cdot i^{-\beta-1} & \text{if} \quad i \geq J \end{cases}$$

Then the derivative of the difference in profits per worker for $i \leq J$ becomes:

$$\begin{aligned}
[i]: \quad & \delta^{-1}\beta \cdot P_L \cdot (1-i)^{-\beta-1} \cdot P_L^{(1-\beta)/\beta} \cdot (1-i)^{\beta-1}(N_H \cdot i \cdot Z - N_L \cdot (1-i)) + \beta\delta^{-1}P_L^{1/\beta} \cdot (1-i)^{-1}(N_H \cdot Z + N_L) \\
&= \delta^{-1}\beta \cdot P_L^{1/\beta} \cdot (1-i)^{-2}(N_H \cdot i \cdot Z - N_L \cdot (1-i)) + \beta\delta^{-1}P_L^{1/\beta} \cdot (1-i)^{-1}(N_H \cdot Z + N_L) \\
&= \beta\delta^{-1}P_L^{1/\beta} \cdot (1-i)^{-1}((1-i)^{-1}(N_H \cdot i \cdot Z - N_L \cdot (1-i)) + (N_H \cdot Z + N_L)) \\
&= \beta\delta^{-1}P_L^{1/\beta} \cdot (1-i)^{-1}N_H Z\left(1 + \frac{i}{1-i}\right) \\
&= \frac{\beta\delta^{-1}P_L^{1/\beta}N_H Z}{(1-i)^2}
\end{aligned}$$

Further, the derivative of the difference in profits per worker for $i \geq J$ becomes:

$$\begin{aligned}
[i]: \quad & -\delta^{-1}\beta \cdot P_H \cdot i^{-\beta-1} \cdot P_H^{(1-\beta)/\beta} \cdot i^{\beta-1}(N_H \cdot i \cdot Z - N_L \cdot (1-i)) + \beta\delta^{-1}P_H^{1/\beta} \cdot i^{-1}(N_H \cdot Z + N_L) \\
&= -\delta^{-1}\beta \cdot P_H^{1/\beta} \cdot i^{-2}(N_H \cdot i \cdot Z - N_L \cdot (1-i)) + \beta\delta^{-1}P_H^{1/\beta} \cdot i^{-1}(N_H \cdot Z + N_L) \\
&= \beta\delta^{-1}P_H^{1/\beta} \cdot i^{-1}((N_H \cdot Z + N_L) - i^{-1}(N_H \cdot i \cdot Z - N_L \cdot (1-i))) \\
&= \beta\delta^{-1}P_H^{1/\beta} \cdot i^{-1}N_L\left(1 + \frac{1-i}{i}\right) \\
&= \frac{\beta\delta^{-1}P_H^{1/\beta}N_L}{i^2}
\end{aligned}$$

Therefore, for $i \in (0,1)$, we have that the derivative is positive, meaning that the difference in profits per worker is strictly increasing in i. Finally, we do not have to worry about the discontinuity, as J is defined such that the profits from the two types of workers are equal. Therefore, while the discontinuity applies to the derivative, profits themselves, and consequently the different in profits per worker, do not have jumps.

We now show that the marginal output of workers is equalized across sectors. First, we write output taking into consideration J:

$$y(i) = \begin{cases} \delta^{-1} p(i)^{(1-\beta)/\beta} N_L \cdot (1-i) \cdot l(i) & \text{if} \quad 0 \leq i \leq J \\ \delta^{-1} p(i)^{(1-\beta)/\beta} N_H \cdot i \cdot Z \cdot h(i) & \text{if} \quad J \leq i \leq 1 \end{cases} \tag{5}$$

We now consider profits in a low-skill market:

$$\begin{aligned} \pi_L(i) &= p(i)y(i) - w_L l(i) - \int_0^{N_L} \chi_L(v) k_L(i,v) \cdot dv \\ &= \delta^{-1} p(i)^{1/\beta} N_L \cdot (1-i) \cdot l(i) - w_L l(i) - N_L \cdot \delta^{-1} p(i)^{1/\beta} (1-i) \cdot l(i) \cdot (1-\beta) \\ &= \beta \delta^{-1} p(i)^{1/\beta} N_L \cdot (1-i) \cdot l(i) - w_L l(i) \end{aligned}$$

Taking FOC with respect to $l(i)$, we have

$$\begin{aligned} [l(i)]: \quad 0 &= \beta \delta^{-1} p(i)^{1/\beta} N_L \cdot (1-i) - w_L \\ \Leftrightarrow w_L &= \beta \delta^{-1} p(i)^{1/\beta} N_L \cdot (1-i) \end{aligned}$$

We next consider profits in a high-skill market:

$$\begin{aligned} \pi_H(i) &= p(i)y(i) - w_H h(i) - \int_0^{N_H} \chi_H(v) k_H(i,v) \cdot dv \\ &= \delta^{-1} p(i)^{1/\beta} N_H \cdot i \cdot Z \cdot h(i) - w_H h(i) - N_H \cdot \delta^{-1} p(i)^{1/\beta} i \cdot Z \cdot h(i) \cdot (1-\beta) \\ &= \beta \delta^{-1} p(i)^{1/\beta} N_H \cdot i \cdot Z \cdot h(i) - w_H h(i) \end{aligned}$$

Taking FOC with respect to $h(i)$, we have

$$\begin{aligned} [h(i)]: \quad 0 &= \beta \delta^{-1} p(i)^{1/\beta} N_H \cdot i \cdot Z - w_H \\ \Leftrightarrow w_H &= \beta \delta^{-1} p(i)^{1/\beta} N_H \cdot i \cdot Z \end{aligned}$$

We therefore have the following: $\delta^{-1} p(i)^{1/\beta} N_L \cdot (1-i)$ is constant if $0 \leq i \leq J$; and $\delta^{-1} p(i)^{1/\beta} N_H \cdot i \cdot Z$ is constant if $J \leq i \leq 1$. For $0 \leq i \leq J$, define the constant $P_L^{1/\beta} = p(i)^{1/\beta} \cdot (1-i) \Leftrightarrow p(i) = P_L \cdot (1-i)^{-\beta}$. Similarly, for $J \leq i \leq 1$, define the constant $P_H^{1/\beta} = p(i)^{1/\beta} \cdot i \Leftrightarrow p(i) = P_H \cdot i^{-\beta}$.

Next, we solve for J. We have the following:

$$\begin{aligned} P_L \cdot (1-J)^{-\beta} &= P_H \cdot J^{-\beta} \\ \Leftrightarrow \frac{P_H}{P_L} &= \left(\frac{J}{1-J} \right)^{\beta} \end{aligned} \tag{8}$$

Noting that $p(i)y(i)$ is constant, we have $p(i)/p(i') = y(i')/y(i) \ \ \forall i, i'$. Without loss of generality, consider

$P_H/P_L = y(0)/y(1)$. Plugging in (5) and recalling that $l(i) = L/J$, $h(i) = H/(1-J)$, we have:

$$\frac{\delta^{-1}p(0)^{(1-\beta)/\beta}N_L \cdot (1-0)\cdot l(i)}{\delta^{-1}p(1)^{(1-\beta)/\beta}N_H \cdot 1 \cdot Z \cdot h(i)} = \left(\frac{J}{1-J}\right)^{\beta}$$
$$\Leftrightarrow \left(\frac{P_L}{P_H}\right)^{(1-\beta)/\beta} \frac{N_L \cdot L/J}{N_H \cdot Z \cdot H/(1-J)} = \left(\frac{J}{1-J}\right)^{\beta}$$
$$\Leftrightarrow \left(\frac{1-J}{J}\right)^{1-\beta} \frac{N_L \cdot L}{N_H \cdot Z \cdot H} = \left(\frac{J}{1-J}\right)^{1+\beta}$$
$$\Leftrightarrow \left(\frac{J}{1-J}\right) = \left(\frac{N_L \cdot L}{N_H \cdot Z \cdot H}\right)^{1/2}$$
$$\Leftrightarrow J\left(1 + \left(\frac{N_L \cdot L}{N_H \cdot Z \cdot H}\right)^{1/2}\right) = \left(\frac{N_L \cdot L}{N_H \cdot Z \cdot H}\right)^{1/2}$$
$$\Leftrightarrow J = \left(1 + \left(\frac{N_H}{N_L} \cdot \frac{Z \cdot H}{L}\right)^{1/2}\right)^{-1}$$

By construction, the numeraire rule gives us:

$$\exp\left[\int_0^1 \ln(p(i)) \cdot di\right] = 1$$
$$\Leftrightarrow \exp\left[\int_0^J \ln\left(P_L \cdot (1-i)^{-\beta}\right) \cdot di + \int_J^1 \ln\left(P_H \cdot i^{-\beta} \cdot di\right)\right] = 1$$
$$\Leftrightarrow \exp\left[J \cdot \ln(P_L) - \beta \int_0^J \ln(1-i) \cdot di + (1-J) \cdot \ln(P_H) - \beta \int_J^1 \ln(i) \cdot di\right] = 1$$

We solve the two remaining integrals:

$$\begin{aligned}\int_0^J \ln(1-i) \cdot di &= (i-1) \cdot \ln(1-i) - i\Big|_{i=0}^{i=J} \\ &= (J-1) \cdot \ln(1-J) - J \\ \int_J^1 \ln(i) \cdot di &= i \cdot (\ln(i) - 1)\Big|_{i=J}^{i=1} \\ &= -1 - J \cdot (\ln(J) - 1)\end{aligned}$$

Plugging in, we now have:

$$\Leftrightarrow \exp\left[J \cdot \ln(P_L) + (1-J) \cdot \ln(P_H) - \beta((J-1) \cdot \ln(1-J) - J - 1 - J \cdot (\ln(J) - 1))\right] = 1$$
$$\Leftrightarrow \exp\left[J \cdot \ln(P_L) + (1-J) \cdot \ln(P_H) - \beta((J-1) \cdot \ln(1-J) - J \cdot \ln(J) - 1)\right] = 1$$
$$\Leftrightarrow P_L^J \cdot P_H^{1-J} \cdot (1-J)^{\beta \cdot (1-J)} \cdot J^{\beta \cdot J} \cdot \exp(\beta) = 1$$

Recall from (8) that we have the following:

$$\frac{P_H}{P_L} = \left(\frac{J}{1-J}\right)^{\beta}$$
$$\Leftrightarrow P_H = \left(\frac{J}{1-J}\right)^{\beta} P_L$$
$$\Leftrightarrow P_L = \left(\frac{J}{1-J}\right)^{-\beta} P_H$$

Plugging in P_H, we have:

$$\Leftrightarrow P_L^J \cdot \left(\left(\frac{J}{1-J} \right)^{\beta} P_L \right)^{1-J} \cdot (1-J)^{\beta \cdot (1-J)} \cdot J^{\beta \cdot J} \cdot \exp(\beta) = 1$$
$$\Leftrightarrow P_L \cdot J^{\beta} \cdot \exp(\beta) = 1$$
$$\Leftrightarrow P_L = \exp(-\beta) \cdot \left(1 + \left(\frac{N_H}{N_L} \cdot \frac{Z \cdot H}{L} \right)^{1/2} \right)^{\beta} \tag{10}$$

Now plugging in P_L, we have:

$$\Leftrightarrow \left(\left(\frac{J}{1-J} \right)^{-\beta} P_H \right)^{J} \cdot P_H^{1-J} \cdot (1-J)^{\beta \cdot (1-J)} \cdot J^{\beta \cdot J} \cdot \exp(\beta) = 1$$
$$\Leftrightarrow P_H \cdot (1-J)^{\beta} \cdot \exp(\beta) = 1$$
$$\Leftrightarrow P_H = \exp(-\beta) \left(1 - \left(1 + \left(\frac{N_H}{N_L} \cdot \frac{Z \cdot H}{L} \right)^{1/2} \right)^{-1} \right)^{-\beta}$$
$$\Leftrightarrow P_H = \exp(-\beta) \left(\frac{1 + \left(\frac{N_H}{N_L} \cdot \frac{Z \cdot H}{L} \right)^{1/2} - 1}{1 + \left(\frac{N_H}{N_L} \cdot \frac{Z \cdot H}{L} \right)^{1/2}} \right)^{-\beta}$$
$$\Leftrightarrow P_H = \exp(-\beta) \left(1 + \left(\frac{N_H}{N_L} \cdot \frac{Z \cdot H}{L} \right)^{-1/2} \right)^{\beta} \tag{10}$$

Now, recall that we derived:

$$w_L = \beta \delta^{-1} p(i)^{1/\beta} N_L \cdot (1-i)$$
$$w_H = \beta \delta^{-1} p(i)^{1/\beta} N_H \cdot i \cdot Z$$

Substituting in prices, we have:

$$\begin{aligned} w_L &= \beta \delta^{-1} (P_L \cdot (1-i)^{-\beta})^{1/\beta} N_L \cdot (1-i) \\ &= \beta \delta^{-1} P_L^{1/\beta} N_L \\ w_H &= \beta \delta^{-1} (P_H \cdot i^{-\beta})^{1/\beta} N_H \cdot i \cdot Z \\ &= \beta \delta^{-1} P_H^{1/\beta} N_H \cdot Z \end{aligned}$$

Taking relative wages gives us:

$$\frac{w_H}{w_L} = \frac{N_H \cdot Z}{N_L} \left(\frac{P_H}{P_L} \right)^{1/\beta}$$

Plugging in from (10):

$$
\begin{aligned}
\frac{w_H}{w_L} &= \frac{N_H \cdot Z}{N_L} \left(\frac{1 + \left(\frac{N_H}{N_L} \cdot \frac{Z \cdot H}{L}\right)^{-1/2}}{1 + \left(\frac{N_H}{N_L} \cdot \frac{Z \cdot H}{L}\right)^{1/2}} \right) \\
&= \frac{N_H \cdot Z}{N_L} \left(\frac{1 + \left(\frac{N_H}{N_L} \cdot \frac{Z \cdot H}{L}\right)^{1/2}}{\left(\frac{N_H}{N_L} \cdot \frac{Z \cdot H}{L}\right)^{1/2} \left(1 + \left(\frac{N_H}{N_L} \cdot \frac{Z \cdot H}{L}\right)^{1/2}\right)} \right) \\
&= \frac{N_H \cdot Z}{N_L} \left(\frac{N_H}{N_L} \cdot \frac{Z \cdot H}{L} \right)^{-1/2} \\
&= Z \left(\frac{N_H}{N_L} \right)^{1/2} \left(\frac{Z \cdot H}{L} \right)^{-1/2}
\end{aligned}
$$

Finally, we consider the definition $Y = \int_0^1 p(i)y(i) \cdot di$. Recalling that $p(i)y(i)$ is constant $\forall i$, and looking at $p(0) = P_L$, we have the following:

$$
\begin{aligned}
Y &= \int_0^1 p(i)y(i) \cdot di \\
&= P_L y(0) \\
&= P_L \cdot \delta^{-1} P_L^{(1-\beta)/\beta} N_L \cdot L/J \\
&= \exp(-1) \cdot \left(1 + \left(\frac{N_H}{N_L} \cdot \frac{Z \cdot H}{L} \right)^{1/2} \right) \cdot \delta^{-1} N_L \cdot L/J \\
&= \exp(-1) \cdot \delta^{-1} \left(1 + \left(\frac{N_H}{N_L} \cdot \frac{Z \cdot H}{L} \right)^{1/2} \right)^2 N_L \cdot L \\
&= \exp(-1) \cdot \delta^{-1} \left((N_L \cdot L)^{1/2} + (N_H \cdot Z \cdot H)^{1/2} \right)^2
\end{aligned}
$$

A.1.2 Cross-Country Results

We start by considering expressions for $\pi_z^n(\tau)$:

$$
\begin{aligned}
\pi_L^n(\tau) &= (\chi^n - \theta) \int_0^{J^n} k_L^n(i) \cdot di \\
\pi_H^n(\tau) &= (\chi^n - \theta) \int_{J^n}^1 k_H^n(i) \cdot di
\end{aligned}
$$

Recall the following:

$$
\begin{aligned}
\chi &= \delta^{\beta/(1-\beta)}(1-\beta) \\
\theta &= \delta^{\beta/(1-\beta)}(1-\beta)^2 \\
k_L(i,v) &= p(i)^{1/\beta}(1-i) \cdot l(i) \cdot \delta^{1/(\beta-1)} \\
k_H(i,v) &= p(i)^{1/\beta} i \cdot Z \cdot h(i) \cdot \delta^{1/(\beta-1)} \\
p(i) &= P_L \cdot (1-i)^{-\beta} \quad \text{if} \quad 0 \le i \le J \\
p(i) &= P_H \cdot i^{-\beta} \quad \text{if} \quad J \le i \le 1 \\
l(i) &= L/J \\
h(i) &= H/(1-J)
\end{aligned}
$$

Further, recall that δ is normalized to 1 in the North. This gets us $\chi^n = 1 - \beta$ and $\theta^n = (1-\beta)^2$. Now simplify:

$$\begin{aligned}
k_L^n(i,v) &= \delta^{1/(\beta-1)}(P_L^n)^{1/\beta} L^n / J^n \\
&= (P_L^n)^{1/\beta} L^n / J^n \\
k_H^n(i,v) &= \delta^{1/(\beta-1)}(P_H^n)^{1/\beta} Z \cdot H^n / (1 - J^n) \\
&= (P_H^n)^{1/\beta} Z \cdot H^n / (1 - J^n)
\end{aligned}$$

Now simplify profits:

$$\begin{aligned}
\pi_L^n(\tau) &= (\chi^n - \theta) \int_0^{J^n} k_L^n(i) \cdot di \\
&= (1 - \beta - (1-\beta)^2)(P_L^n(\tau))^{1/\beta} L^n \\
&= (1 - \beta - 1 + 2\beta - \beta^2)(P_L^n(\tau))^{1/\beta} L^n \\
&= \beta(1-\beta)(P_L^n(\tau))^{1/\beta} L^n \\
\pi_H^n(\tau) &= (\chi^n - \theta) \int_{J^n}^{1} k_H^n(i) \cdot di \\
&= \beta(1-\beta)(P_H^n(\tau))^{1/\beta} Z \cdot H^n
\end{aligned}$$

We therefore have the following two expressions:

$$\begin{aligned}
\pi_L^n(\tau) &= (\chi^n - \theta) \int_0^{J^n} k_L^n(i) \cdot di = \beta(1-\beta)(P_L^n(\tau))^{1/\beta} L^n \\
\pi_H^n(\tau) &= (\chi^n - \theta) \int_{J^n}^{1} k_H^n(i) \cdot di = \beta(1-\beta)(P_H^n(\tau))^{1/\beta} Z \cdot H^n
\end{aligned} \tag{14}$$

We now derive the BGP relationship between technology levels, labor ratios, and J in the North. We have the following:

$$\frac{P_H}{P_L} = \left(\frac{J}{1-J}\right)^{\beta} \tag{8}$$

$$\frac{L^n}{Z \cdot H^n} = \left(\frac{P_H^n}{P_L^n}\right)^{1/\beta} \tag{15}$$

Combining gives us:

$$\frac{Z \cdot H^n}{L^n} = \frac{1 - J^n}{J^n}$$

Finally, we have the following from our derivation of J:

$$\left(\frac{J}{1-J}\right) = \left(\frac{N_L \cdot L}{N_H \cdot Z \cdot H}\right)^{1/2}$$

Substituting in gives us:

$$\begin{aligned}
\left(\frac{J^n}{1-J^n}\right) &= \left(\frac{N_L \cdot J^n}{N_H \cdot (1-J^n)}\right)^{1/2} \\
\Leftrightarrow \frac{N_H}{N_L} &= \frac{1-J^n}{J^n}
\end{aligned}$$

Taken together, we have the following:

$$\frac{N_H}{N_L} = \frac{1-J^n}{J^n} = \frac{Z \cdot H^n}{L^n} \tag{16}$$

A.2 Extension

These results apply to both the North and the South. We start by considering the Producer's problem. Firms solve the same problem as in Acemoglu and Zilibotti's model, except for output. We also assume that $N_H > N_L$. Firms solve the following problem:

$$\max_{\substack{l(i),\, h(i), \\ \{k_L(i,v)\}_{v\in[0,N_L]} \\ \{k_H(i,v)\}_{v\in[0,N_H]}}} \quad p(i)y(i) - w_L l(i) - w_H h(i) - \int_0^{N_L} \chi_L(v) k_L(i,v) \cdot dv$$
$$- \int_0^{N_H} \chi_H(v) k_H(i,v) \cdot dv$$

where output is given as follows:

$$\begin{aligned} y(i) = &\left[\int_0^{N_L} k_L(i,v)^{1-\beta} \cdot dv + \mathbb{1}\{N_H > N_L\}\gamma_L \int_{N_L}^{N_H} k_L(i,v)^{1-\beta} \cdot dv\right] \cdot [(1-i)\cdot l(i)]^\beta \\ &+ \left[\int_0^{N_H} k_H(i,v)^{1-\beta} \cdot dv + \mathbb{1}\{N_L > N_H\}\gamma_H \int_{N_H}^{N_L} k_H(i,v)^{1-\beta} \cdot dv\right] \cdot [i \cdot Z \cdot h(i)]^\beta \end{aligned} \tag{21}$$

Taking FOC with respect to $k_z(i,v)$, we get:

$$\begin{aligned} [k_L(i,v)]: \quad & 0 = p(i)(1-\beta)k_L(i,v)^{-\beta}[(1-i)\cdot l(i)]^\beta(1+\gamma_L) - \chi_L(v) \\ \Leftrightarrow\ & k_L(i,v) = [(1-\beta)(1+\gamma_L)p(i)((1-i)\cdot l(i))^\beta / \chi_L(v)]^{1/\beta} \\ [k_H(i,v)]: \quad & 0 = p(i)(1-\beta)k_H(i,v)^{-\beta}[i \cdot Z \cdot h(i)]^\beta - \chi_H(v) \\ \Leftrightarrow\ & k_H(i,v) = [(1-\beta)p(i)(i \cdot Z \cdot h(i))^\beta / \chi_H(v)]^{1/\beta} \end{aligned} \tag{22}$$

Next, we consider the patent owner's problem. Patent owners solve the following problem:

$$\max_{\chi_z(v)} \pi_z(v) = (\chi_z(v) - \theta) \int_0^1 k_z(i,v) \cdot di$$

Taking FOC with respect to $\chi_z(v)$, we get:

$$\begin{aligned} [\chi_L]: \quad 0 &= \int_0^1 k_L(i,v) \cdot di + (\chi_L(v) - \theta) \int_0^1 \frac{\partial k_L(i,v)}{\partial \chi_L(v)} \cdot di \\ &= \left(\frac{(1-\beta)(1+\gamma_L)}{\chi_L(v)}\right)^{1/\beta} \int_0^1 p(i)^{1/\beta}(1-i)\cdot l(i) \cdot di \\ &\qquad + (\chi_L(v)-\theta)((1-\beta)(1+\gamma_L))^{1/\beta}\left(-\frac{1}{\beta}\right)\chi_L(v)^{\frac{-1-\beta}{\beta}} \int_0^1 p(i)^{1/\beta}(1-i)\cdot l(i)\cdot di \\ &= 1 + (\chi_L(v)-\theta)\left(-\frac{1}{\beta}\right)\chi_L(v)^{-1} \\ \Leftrightarrow 1-\beta &= \frac{\theta}{\chi_L(v)} \\ \Leftrightarrow \chi_L(v) &= \frac{\theta}{1-\beta} \end{aligned}$$

$$
\begin{aligned}
[\chi_H]: \quad 0 &= \int_0^1 k_H(i,v)\cdot di + (\chi_H(v)-\theta)\int_0^1 \frac{\partial k_H(i,v)}{\partial \chi_H(v)}\cdot di \\
&= \left(\frac{1-\beta}{\chi_H(v)}\right)^{1/\beta}\int_0^1 p(i)^{1/\beta} i\cdot Z\cdot h(i)\cdot di + (\chi_H(v)-\theta)(1-\beta)^{1/\beta}\left(-\frac{1}{\beta}\right)\chi_H(v)^{\frac{-1-\beta}{\beta}}\int_0^1 p(i)^{1/\beta} i\cdot Z\cdot h(i)\cdot di \\
&= 1 + (\chi_H(v)-\theta)\left(-\frac{1}{\beta}\right)\chi_H(v)^{-1} \\
\Leftrightarrow 1-\beta &= \frac{\theta}{\chi_H(v)} \\
\Leftrightarrow \chi_H(v) &= \frac{\theta}{1-\beta}
\end{aligned}
$$

We therefore have that $\chi_L(v) = \chi_H(v) = \chi$, which is unchanged from Acemoglu and Zilibotti's model. As explained in Section 3.1, we set $\theta = \delta^{\beta/(1-\beta)}(1-\beta)^2$, giving us $\chi = \delta^{\beta/(1-\beta)}(1-\beta)$.

Now plugging these prices into (22), we have:

$$
\begin{aligned}
k_L(i,v) &= ((1+\gamma_L)p(i))^{1/\beta}(1-i)\cdot l(i)\cdot \delta^{1/(\beta-1)} \\
k_H(i,v) &= p(i)^{1/\beta} i\cdot Z\cdot h(i)\cdot \delta^{1/(\beta-1)}
\end{aligned}
$$

Plugging these into (21), we have:

$$
\begin{aligned}
y(i) &= \left[\int_0^{N_L}\left(((1+\gamma_L)p(i))^{1/\beta}(1-i)\cdot l(i)\cdot \delta^{1/(\beta-1)}\right)^{1-\beta}\cdot dv\right]\cdot[(1-i)\cdot l(i)]^\beta \\
&\qquad + \gamma\left[\int_{N_L}^{N_H}\left(((1+\gamma_L)p(i))^{1/\beta}(1-i)\cdot l(i)\cdot \delta^{1/(\beta-1)}\right)^{1-\beta}\cdot dv\right]\cdot[(1-i)\cdot l(i)]^\beta \\
&\qquad + \left[\int_0^{N_H}\left(p(i)^{1/\beta} i\cdot Z\cdot h(i)\cdot \delta^{1/(\beta-1)}\right)^{1-\beta}\cdot dv\right]\cdot[i\cdot Z\cdot h(i)]^\beta \\
&= \delta^{-1}p(i)^{(1-\beta)/\beta}[N_L\cdot(1+\gamma_L)^{1-\beta}\cdot(1-i)\cdot l(i) + (N_H-N_L)\cdot\gamma_L\cdot(1+\gamma_L)^{1-\beta}\cdot(1-i)\cdot l(i) \\
&\qquad + N_H\cdot(1+\gamma_L)^{1-\beta}\cdot i\cdot Z\cdot h(i)] \\
&= \delta^{-1}p(i)^{(1-\beta)/\beta}[N_L\cdot(1-\gamma_L)\cdot(1+\gamma_L)^{1-\beta}\cdot(1-i)\cdot l(i) + N_H\cdot(1+\gamma_L)^{2-\beta}\cdot i\cdot Z\cdot h(i)]
\end{aligned}
$$