Innovation, Growth, and Dynamic Gains from Trade

Wen-Tai Hsu Raymond G. Riezman Ping Wang*†

April 19, 2020

Abstract

How large are the welfare gains from trade? Would such gains be significantly amplified in the long run when productivity is endogenously enhanced? We address these questions by focusing on the dynamic effect of trade, in particular, how trade affects the incentives for technological advancement. We construct an innovation-based endogenous growth model of North-South trade. There are two types of innovation: one by the North to upgrade the general purpose technology (GPT) and another by both countries to advance entrepreneurial knowledge for developing differentiated products. We find sizable welfare gains from trade, about 5.3% when compared to autarky. The gains in our dynamic model are much higher than the static estimates where the effets of GPT-driven innovation are eliminated. The share of dynamic gains from trade is about 78% of the total gains in our benchmark economy – much higher than comparable figures identified in previous studies. Comparative statics indicate that GPT innovation efficacy, entrepreneurial talent distribution and trade elasticity are crucial for dynamic gains from trade.

*Hsu: School of Economics, Singapore Management University. E-mail: wentaihsu@smu.edu.sg. Riezman: Aarhus University, University of California, Santa Barbara, University of Iowa. Email: raymond-riezman@uiowa.edu. Wang: Department of Economics, Washington University in St. Louis; NBER. E-mail: pingwang@wustl.edu.

†We are grateful for valuable comments and suggestions from Costas Arkolakis, Eric Bond, Been-Lon Chen, Jonathan Eaton, Chang-Tai Hsieh, Sam Kortum, Erzo G. J. Luttmer, Paul Segerstrom and Michael Waugh, as well as conference participants at the Midwest Trade Meetings, the Society for Advanced Economic Theory Conference, and the Taiwan Economic Research Conference. Hsu would like to acknowledge the financial support provided by the Sing Lun Fellowship of Singapore Management University. We also acknowledge grant support from Taiwan's Ministry of Science and Technology (MOST 104=2811-H-001-007) to facilitate the coauthor work at Academia Sinica.

> The essence of technological modernity is non-stationarity: ... as far as future technological progress and economic growth are concerned, not even the sky is the limit. (Mokyr 2005; p. 32, p. 62)

1 Introduction

How large are the welfare gains from trade? This has been one of the central questions in economics dating back to Adam Smith. The influential study by Arkolakis, Costinot, and Rodriguez-Clare (2012; henceforth ACR) and related literature have shown that the gains from trade in a static framework are typically small. In the empirical development literature, however, trade liberalization has been perceived to lead to higher productivity and faster growth.[1,2] This suggests that gains from trade would be amplified in a dynamic framework in which productivity is endogenously improved.

To properly address this important issue, we shall determine: (i) *what the main driving forces are for productivity enhancement and growth* and (ii) *how trade may play a role in this dynamic process*. Regarding the first issue, we generalize the endogenous growth framework *a là* Aghion and Howitt (1992; henceforth AH) in which the sustained growth of the economy is driven by R&D with creative destruction. In light of the historical evolution of technology (cf. Mokyr 1992 and his handbook article 2005), we consider two types of innovation. One is a general purpose technology (GPT) that evolves in a similar way to that in AH. We think of GPT as a technology that is pervasive, potentially beneficial to all products and having potential for continuous technical advancement (cf. Helpman and Trajtenberg 1998). Historical examples of GPT include internal combustion engines, electricity, digital data processing, macromolecules composition and semiconductors. Concerning the role of trade, we focus on innovation that occurs while developing differentiated traded products. Similar to Bernard, Eaton, Jensen, and Kortum (2003;

[1] In most cross-country studies, there is a positive relationship between trade openness and economic growth (e.g., Frenkle and Romer 1999 find that the effect of trade on income growth is large though the standard errors are also large in IV estimation leading to only marginal statistical significance). In the decade after liberalization, economic growth has been significantly higher than in the previous decade. For example, using Sachs and Warner liberalization index in panel regressions, Greenaway, Morgan, and Wright (2002) find a growth effect of major trade reforms amounted to just under 1% in year 1 after the reform, 1.5% in year 2, and 2% in subsequent years.

[2] Productivity gains from trade liberalization have been identified in varous empirical studies, such as in Brazil (Ferreira and Rossi 2003), Chile (Pavcnik 2002), Colombia (Fernandes 2007), India (Topalova and Khandelwal 2011), Indonesia (Amiti and Konings 2007), Korea (Kim 2000), and even in developed countries such as the U.S. (such productivity gain amounts to 1.2% in Corrado, Hulten and Sichel 2009). More recently, Halpern, Koren, and Szeidl (2015) find large productivity gains from tariff reduction in Hungary where imported inputs contribute to about a quarter of its productivity growth. Bloom, Draca, and van Reenen (2016) also identify large productivity gains from Chinese firms since China's accession to the WTO.

henceforth BEJK), each entrepreneur in our economy enters a market of a specific product to compete globally in a Bertrand fashion. All differentiated products require the use of GPT to produce. Hence, when GPT occurs, the ideas (or blueprints) used to produce the set of differentiated traded products needs to be redrawn. Better GPT allows more ideas to be drawn. In our North-South trade framework, GPT innovation occurs in the North but is used by both the North and the South to develop products.[3]

Next we address how international trade affects R&D incentives and growth. Our model uses the BEJK framework as a starting point, supplemented with GPT innovation. Here, the rents earned by differentiated-product firms are shared with the North GPT firm according to Nash bargaining. The static gains from lowering trade barriers over time will thus affect the incentive for GPT innovation. Thus, through trade there are interactions between the forces affecting GPT innovation and the general equilibrium effects in the BEJK trade environment. We identify three channels through which trade liberalization affects welfare: (i) a typical static channel comparable to ACR, (ii) an income-gains channel for the North due to payments the GPT firm receives from the South, and (iii) a growth channel via GPT innovation.

We will show that the link between GPT innovation and trade costs can be captured by a multiplier. Trade liberalization induces this multiplier to increase, hence encouraging GPT innovation and increasing economic growth. The intuition is that aggregate world revenue and hence aggregate labor demand (in North labor units) increase in face of trade liberalization. This occurs because a more integrated economy implies a stronger demand for the South's labor (relative to the North's) which reduces the wage gap (via a factor-price-equalization effect of trade). When R&D effort increases as a result, economic growth becomes higher. That is, trade amplifies a spiral effect of growth. This mechanism of trade and growth, with globalization characterized by falling tariffs and improving transport technology, contributes to the narrowing of income and wage gaps between the North and the South and is consistent with recent research (see, e.g., Caliendo et al. 2017).

The central focus of our analysis is on the total gains from trade and the contribution of dynamics to those gains. In our model, the long-run welfare change in response to a change in trade cost can be written as a product of a growth-rate (GR) effect, an income-gains (IG) effect, and the ACR statistic. That is, in addition to the usual static ACR gains, there is a GR effect measuring how growth amplifies the gains from trade and an IG effect via payments received by the North GPT firms. We investigate two scenarios: (i) the first

[3]Innovation on GPT is generally more difficult than other innovation. Empirical studies have shown that the spatial distribution of innovation (let alone the GPT innovation) is highly uneven, with the North countries having a dominant share (e.g., see Egger and Loumeau 2018).

studies moving from autarky to the benchmark trade cost, and (ii) the second looks at the move from the benchmark trade cost to frictionless or free trade.

In scenario (i), the total gains from trade are 5.34% and the GR share of total gains is 79.6%. In scenario (ii), the total gains from trade is 3.18% and the GR counts for 49.7% of those gains. However, this simple decomposition ignores the general equilibrium effects. To incorporate the general equilibrium effects, we compare the benchmark model with a purely static model. We find that in scenario (i), the static gains are 1.15%, and the total gains are 4.7 times larger than the static gains, implying a dynamic share of 78.1% which is very close to 79.6%. In scenario (ii), the static gains are 1.62, and the total gains are 1.97 times larger, yielding a dynamic share of 48.7% which again is very close to 49.7%. This shows that the general equilibrium effects of shutting down the dynamic mechanism are quite small and we can use the simple decomposition to understand the intuition underlying these channels.

Note that the dynamic shares, especially those in scenario (i), are substantially larger than the figures found in related recent literature on dynamic gains from trade discussed below. First, by the envelope theorem (or Shepherd's Lemma), the ACR statistic reflects essentially a direct effect of trade cost on prices of the imported goods. Under lower trade costs, consumed quantities of imported goods tend to be larger, and this amplifies the direct price effect. This explains why the ACR effect is much larger in scenario (ii) than (i). Second, with a large trade cost, the wage gap between the North and South is quite large, and reducing trade costs from autarky to the benchmark level reduces this wage gap drastically. This results in a sharp increase in the relative demand for the South's labor, and, with more equal wages, a larger global market. The increase in GPT R&D is therefore larger in scenario (i), yielding a larger GR effect.

The two main takeaways from this paper are (i) Total gains from trade associated with GPT-driven innovations are substantial and the dynamic share of these gains are much larger than the share reported in the recent literature and (ii) Trade liberalization substantially reduces the North-South welfare gap, which is almost halved when moving from autarky to a frictionless world. We will show that these results are quite robust. Taken together these results reflect the views of Mokyr (2005) in that trade is an amplifier of the spiral effect that leads to non-stationarity and endogenous growth: aggregate global sales go up in face of trade liberalization and hence the value of innovation rises, thereby inducing more R&D and faster growth, which in turns promotes further growth of aggregate global sales.

Related Literature

Our study is closely related to a series of recent studies on dynamic welfare gains from trade. Sampson (2016) develops an endogenous growth model in which the productivity distribution of firms constantly moves to the right over time as entrants learn from the incumbent firms that survived selection previously.[4] Perla, Tonetti and Waugh (2015) quantify a model of trade and growth in which all firms (entrants and incumbents) learn from each other (not just the entrants learning from the incumbents as in Sampson 2016). Impullitti and Licandro (2017) construct a two-country model with process innovation. Hsieh, Klenow and Nath (2019) construct a North-North trade model allowing global competition in innovation and growth exclusively from global creative destruction. Our model mechanism differs from all of the above studies. Whereas the total gains from trade compared with autarky in our benchmark economy, 5.34%, is near the low end (higher than 3.6% found by Sampson (2016) but lower than the other three papers,) the dynamic share, 78%, is higher than all of the above except Perla et al. (2015), who found a dynamic share of 85%.

Also related are the following four studies on dynamic gains from trade which focuses on different counter-factual exercises from ours, and hence the numbers are not directly comparable. Alessandria, Choi and Ruhl (2014) evaluate welfare gains from tariff reductions in a model of exporting firms' life cycle which features a trade-off between new firm creation and export capacity expansion. Bloom, Romer, Terry, and Van Reenen (2016) examine gains from trade in the presence of frictions that impede factor mobility. These frictions lead to more innovation and increased growth in the medium run, resulting in larger welfare gains from trade. Buera and Oberfield (2019) quantifies a multi-country model of trade with global diffusion of ideas in a BEJK-type model. Similar to our model, country-specific distributions of the frontier knowledge are described by Fréchet and endogenously evolve over time. Ravikumar, Santacreu, and Sposi (2018) investigate a trade model with capital accumulation and fit the model to a sample of 44 countries.[5]

Using a static framework, Atkeson and Burstein (2010) show that the welfare gains from trade are typically small and similar to models with homogenous firms because the positive effects on aggregate productivity by firm entry/exit are offset by product innovation.[6] Also related is the study by Grossman and Helpman (2018) that probes the

[4]The growth mechanism in our model is different from Sampson's, and the productivity distribution in our model also moves to the right over time as the scaling factors in the Fréchet distribution grow endogenously over time.

[5]See also the earlier work by Baldwin (1992) who showed that the gains in output in the steady state from a trade liberalization in a model with capital accumulation can be several times larger than the static gains.

[6]Since ACR and Atkeson and Burstein, the literature on the welfare gains from trade is fast-growing and already large. This includes the discussion on the role of pro-competitive effects such as Arkolakis et al. (2018), Edmond et al. (2015), Feenstra et al. (2016) and Holmes et al. (2014), the role of estimation of trade elasticity such as Simonovska and Waugh (2014a, b), the role of productivity distribution such as Melitz

effect of trade and growth on income inequality.

2 The Model

There are two countries, $i = 1, 2$, each with population N_i, which is constant over time. Without loss of generality, we shall label country 1 and 2 as the North and South, respectively. Time is continuous, with calendar time indexed by t and innovation time by ν.

2.1 Preferences

The life time utility of each individual in country i is:

$$U_i = \int_0^\infty Q_{it} e^{-\rho t} dt,$$

where ρ is the time preference rate, and instantaneous utility Q_{it} is given by the standard CES aggregator:

$$Q_{it} = \left(\int_0^1 (q_{it}(\omega))^{\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}},$$

where σ is the elasticity of substitution among differentiated goods which are indexed by ω.

2.2 Innovation

GPT R&D and the subsequent innovation only occur in the North. Each good ω requires a blueprint of the production process and production workers to produce. Each production process requires use of GPT, which is embodied in a "chip" (or an operation system, or engine, etc.). Here, a new generation of GPT has two effects. First, the production process of every good has to be revamped to accommodate this new chip, and hence new ideas about how to produce have to be drawn again. Second, because chips become better over time and allow more flexibility in how to produce the final good ω, the number of ideas that entrepreneurs can draw from grows by a constant factor $\gamma > 1$. This is similar

and Redding (2015), and the role of intermediate goods and sectoral linkages such as Caliendo and Parro (2015). Also, Fajgelbaum and Khandelwal (2016) discuss the differential effects of trade liberalization on consumers with different income. di Giovanni, Levchenko, and Zhang (2014) and Hsieh and Ossa (2016) investigate the global welfare impact of China's trade integration and productivity growth.

to the quality ladders in AH and Anant, Dinopoulos, and Segerstrom (1990). The GPT monopoly can make each chip good-specific at zero cost.

In the North, there is a unit continuum of GPT innovating firms, each of which engages in hiring R&D labor (researchers.) Innovation occurs according to a Poisson arrival process. This market is perfectly competitive, and innovating firms take the wages for R&D labor as given. As these firms are ex ante identical and the GPT arrival rate for an individual firm is the same as the GPT arrival rate for the whole continuum. Once the GPT innovation occurs, without loss of generality, one firm is randomly selected as the new GPT monopoly.[7] We will formally describe the GPT innovation process in Section 2.4.

Let $\tilde{t}_{i0}$ denote the number of ideas per unit of the M_{i0} entrepreneurial labor at time zero. When ν-th generation of GPT is invented, at time ν_+ each unit of the $M_{i\nu}$ entrepreneurial labor for each good ω has $\gamma^\nu \tilde{t}_{i0}$ ideas, and the productivity of each idea is drawn from a Fréchet distribution, $F_i^{\text{draw}}(z) = e^{-z^{-\theta}}$. Firms engage in Bertrand competition as in BEJK, and the best idea prevails. Then, at ν_{++} the GPT monopoly makes the chip ω-specific to each ω and engages in Nash bargaining with each entrepreneur who owns the best production process to produce ω. For each ω, the bargaining is over the expected profits of the ω-firm for the entire duration of the GPT (until it is replaced by the next GPT innovation), and the bargaining power of the GPT firm is $\beta \in (0, 1)$. For simplicity, marginal cost of producing each additional chip is assumed to be zero. In expected terms, this bargaining outcome can also be interpreted as the GPT monopoly selling the right at every instant to use ω-chip for each ω at the price of β times the expected profit of the ω-firm. One can think of this as a royalty payment.

Entrepreneurs in both countries need to purchase the ω-chip from GPT innovators. The evolution of the total number of idea draws in each country i is $\widetilde{T}_{i\nu} = M_{i\nu}\gamma^v \tilde{t}_{i0}$. Let Z_1 and Z_2 denote top two productivities (clearly random variables). To utilize the limiting joint distribution of the top two productivities in BEJK, assume that $\tilde{t}_{i0} = t_{i0}K$, where K is an arbitrarily large number. Defining $T_{i\nu} \equiv \widetilde{T}_{i\nu}/K$, we have:

$$T_{i\nu} = M_{i\nu}\gamma^v t_{i0}.$$

While details are relegated to Appendix A, the limiting joint distribution of the scaled

[7]The model can be straightforwardly modified to the Aghion-Howitt setting with R fixed and φ function being constant returns to scale so that there are numerous R&D firms to compete for the lone monopolist in the future.

order statistics $K^{-1/\theta}Z_1$ and $K^{-1/\theta}Z_2$ is given by,[8]

$$\lim_{K\to\infty} \Pr\left[K^{-1/\theta}Z_1 \le z_1, K^{-1/\theta}Z_2 \le z_2\right] = \left[1 + T_{i\nu}\left(z_2^{-\theta} - z_1^{-\theta}\right)\right] e^{-T_{i\nu} z_2^{-\theta}}.$$

Thus, with proper rescaling, the c.d.f. of the maximum productivity of these draws and that of the top two productivities are, respectively,

$$F_{i\nu}(z) = e^{-T_{i\nu} z^{-\theta}}, \quad z \ge 0, \tag{1}$$

$$F_{i\nu}(z_1, z_2) = [1 + T_{i\nu}(z_2^{-\theta} - z_1^{-\theta})] e^{-T_{i\nu} z_2^{-\theta}}. \tag{2}$$

The BEJK trade model starts with (1) and (2) with given $T_{i\nu}$. Here, $T_{i\nu}$ is endogenous and evolves with $M_{i\nu}$ and GPT innovations.[9] The cross-country difference in $T_{i\nu}$ is also affected by t_{i0}. We assume that the North has a larger initial knowledge stock so that $t_{10} > t_{20}$. We now turn to the production and trade environment in BEJK.

2.3 Production and Trade

We briefly explain how the cost and price distributions are determined in BEJK and then collect several useful results from BEJK. For each good ω, the unit cost of supplying to consumers in location n by the k-th most efficient producers located in i is:

$$C_{kni}(\omega) = \left(\frac{w_i}{Z_{ki}(\omega)}\right) \tau_{ni},$$

where $\tau_{ni} = 1$ if $n = i$, $\tau_{ni} = \tau$ if $n \ne i$, and the $Z_{1i}(\omega)$ and $Z_{2i}(\omega)$ are random variables whose joint distribution is given by (2). The producer serving market n has unit cost $C_{1n}(\omega) = \min_i\{C_{1ni}(\omega)\}$, and the second lowest cost to supply to n is $C_{2n}(\omega) = \min\{C_{2i^*n}(\omega), \min_{i \ne i^*}\{C_{1in}(\omega)\}\}$, where i^* is the country in which the lowest cost supplier to n is located. Bertrand competition implies that the producer with $C_{1n}(\omega)$ charges $C_{2n}(\omega)$, resulting in a markup of $C_{2n}(\omega)/C_{1n}(\omega)$. Under the CES specification of the utility function, this markup cannot exceed the monopoly markup $\sigma/(\sigma - 1)$ if $\sigma > 1$. Bertrand limit pricing yields:

$$P_n(\omega) = \begin{cases} \min\{C_{2n}(\omega), \frac{\sigma}{\sigma-1}C_{1n}(\omega)\} & \text{if } \sigma > 1 \\ C_{2n}(\omega) & \text{if } \sigma \le 1 \end{cases}.$$

[8]We thank Erzo Luttmer for his comments which help clarify the setup here.

[9]Note that whereas GPT innovation enhances productivities of all goods, it is not a typical sense of process innovation, which is often considered as a firm-level effort.

Using (1) and (2), the joint distribution function of C_{1n} and C_{2n} is $H_n(c_1, c_2) = 1 - e^{-\Phi_n c_1^\theta} - \Phi_n c_1^\theta e^{-\Phi_n c_2^\theta}$, where $\Phi_n = \sum_{i=1}^{2} T_i(w_i\tau_{ni})^{-\theta}$, which distills the parameters of productivity distributions, wages, and the trade cost into one single term governing the cost and price distributions. The following Results recap or extend BEJK's analytical results that turn out to be useful for our analysis.

BEJK Result 1 The probability that country i provides a good at the lowest cost in country n is:

$$\pi_{ni} = \frac{T_i(w_i\tau_{ni})^{-\theta}}{\sum_{k=1}^{2} T_k(w_k\tau_{nk})^{-\theta}} = \frac{T_i(w_i\tau_{ni})^{-\theta}}{\Phi_n}.$$

Since there is a continuum of goods, π_{ni} is also the fraction of goods that the consumers at n purchases from i.

BEJK Result 2 In a country n, the probability of buying a good with price lower than p is independent of where the good is purchased from. Thus, π_{ni} is also the share of expenditure of consumers in n on the goods from i:

$$X_{ni} = \pi_{ni} Y_n,$$

where X_{ni} is the total sales from country i to n, and Y_n denote the total (nominal) income in n. Let X_n denote the total sales of producers in n, and note that the above equation differs from the formula in BEJK, $X_{ni} = \pi_{ni} X_n$, because in our model total revenue from firms producing differentiated products is not equal to total income due to the international royalty payments.

BEJK Result 3 Under $\theta + 1 > \sigma$, the price index in country n is:

$$P_n = \eta \Phi_n^{-\frac{1}{\theta}},$$

where $\eta \equiv \left\{ \left[1 + \frac{(\sigma-1}{1+\theta-\sigma} \left(\frac{\sigma}{\sigma-1} \right)^{-\theta} \right] \Gamma\left(\frac{\theta+1-\sigma}{\theta} \right) \right\}^{\frac{1}{1-\sigma}}$ with Γ denoting the gamma function.

BEJK Result 4 A fraction $\theta/(1+\theta)$ of revenue goes to variable cost.

2.4 Labor Markets

We assume two types of labor in the North, entrepreneurial type and research type, both share a common skill distribution. By occupational choice, the entrepreneurial type endogenously sort into entrepreneurs and production workers whereas the research type as

R&D labor and production workers. The South is populated with labor of the entrepreneurial type only. Labor markets in both countries are perfectly competitive.

In the North, the shares of the two types of labor are exogenously given, with ψ being the entrepreneurial type and $1-\psi$ the research type; the population of the two types are denoted as $N_1^M = \psi N_1$ and $N_1^R = (1-\psi) N_1$, respectively. Each individual in each type has an entrepreneurial/research skill a, and the common distribution of a is denoted as G. The skill-augmented workforce is thus measured in skill units, or equivalently, labor efficiency units. Being a production worker is always an option for each individual, whose wages w_i are independent of individual skill. Assume that a follows the Pareto distribution, $G(a) = 1 - a^{-k}$, with $k > 1$ and the minimum skill normalized to one.

By BEJK Result 4 and the fact that part of the profit is paid to the GPT firm, the expected payoff for one with skill a to become an entrepreneur is $a \times v_{i\nu} = a\frac{1-\beta}{1+\theta}\frac{X_{i\nu}}{M_{i\nu}}$. An individual of the entrepreneurial type in country i at innovation time ν choose to become an entrepreneur if and only if $a \times v_{i\nu} \geq w_{i\nu}$. Note that by definition, $M_{iv} = N_i^M \int_{a_{i\nu}^M}^{\infty} adG(a)$, where $a_{i\nu}^M$ is the cutoff of occupational choice. With the Pareto G, it is readily verified that the entrepreneurial activity $M_{i\nu}$ is given by,

$$M_{iv} = \left(\frac{kN_i^M}{k-1}\right)^{\frac{1}{k}} \left(\frac{1-\beta}{1+\theta}\frac{X_{i\nu}}{w_{i\nu}}\right)^{\frac{k-1}{k}}. \tag{3}$$

When sales-wage ratio $X_{i\nu}/w_{i\nu}$ increases, entrepreneurship also increases; this is a property that is absent in the BEJK model.

We now shift our attention to the research type, whose occupational choice depends critically on the value of innovation. As the aggregate profit earned by all differentiated-product firms is $\frac{X_{1\nu}+X_{2\nu}}{1+\theta}$ at any point in time between ν-th and $\nu+1$-th innovations, the value of an innovation is, in a similar fashion to AH, specified as:

$$\begin{aligned} V_{\nu+1} &= \beta \int_0^\infty \left[\int_0^t \frac{(X_{1\nu+1} + X_{2\nu+1}) e^{-r\tau}}{1+\theta} d\tau\right] \lambda(R_{\nu+1}) e^{-\lambda(R_{\nu+1})t} dt \\ &= \frac{\beta (X_{1\nu+1} + X_{2\nu+1})}{(1+\theta)[r + \lambda(R_{\nu+1})]}, \end{aligned} \tag{4}$$

where r is the real interest rate, and the Poisson arrival rate λ is an increasing function of R&D labor R_ν hired for innovating a new GPT. Here, the value of innovation is the sum of the GPT profits over time properly discounted by both the real interest rate and the rate of creative destruction, $\lambda(R_{\nu+1})$. Taking $V_{\nu+1}$ and the wages for researchers in the North

w_ν^R as given,[10] a GPT innovator solves:

$$\max_{R_\nu} \lambda\left(R_\nu\right) V_{\nu+1} - w_\nu^R R_\nu.$$

Whenever an innovation arrives, the innovator becomes the GPT monopoly. With (4), the first-order condition is:

$$MB \equiv \lambda'\left(R_\nu\right) V_{\nu+1} = \frac{\beta\left(X_{1\nu+1} + X_{2\nu+1}\right)}{1+\theta} \frac{\lambda'\left(R_\nu\right)}{r + \lambda\left(R_{\nu+1}\right)} = w_\nu^R \equiv MC, \tag{5}$$

where the left-hand side is the marginal benefit of R_ν and the right-hand side is the marginal cost. We restrict $\lambda'' < 0$ so that the second-order condition $\lambda''\left(R_\nu\right) V_{\nu+1} < 0$ holds.

A research-type individual with skill a at innovation time ν chooses to be a researcher if and only if $a \times w_\nu^R \geq w_{1\nu}$. By definition, $R_\nu = N_1^R \int_{a_\nu^R}^{\infty} adG\left(a\right)$, where a_ν^R is the cutoff of occupational choice. Solving for a_ν^R, $a_\nu^R w_\nu^R = w_{1\nu}$ can be rewritten as:

$$\frac{w_\nu^R}{w_{1\nu}} = \left(\frac{k-1}{k} \frac{R_\nu}{N_1^R}\right)^{\frac{1}{k-1}}. \tag{6}$$

This captures the occupational-choice effect – the higher the wages for a researcher relative to those for a production worker, the larger the proportion of individuals choosing to be researchers.

Using (6) and rewriting the marginal cost and benefit in unit of the North's labor, (5) becomes:

$$\frac{\beta\left(X_{1\nu+1} + X_{2\nu+1}\right)}{\left(1+\theta\right) w_{1\nu}} \frac{\lambda'\left(R_\nu\right)}{r + \lambda\left(R_{\nu+1}\right)} = \left(\frac{k-1}{k} \frac{R_\nu}{N_1^R}\right)^{\frac{1}{k-1}}. \tag{7}$$

The left-hand side of (7) contains two components – the AH component $\frac{\lambda'(R_\nu)}{r+\lambda(R_{\nu+1})}$ that captures the effect of arrival, interest, and creative-destruction rates on the marginal benefit, and a multiplier $\frac{\beta(X_{1\nu+1}+X_{2\nu+1})}{(1+\theta)w_{1\nu}}$ due to royalty revenue that is proportional to the world economy's aggregate revenue, in unit of the North's labor. Clearly, a higher multiplier increases the incentive to do R&D, and this multiplier term, which we call *R&D multiplier*, is the key to link trade and growth as we will explain in Section 3.2. The right-hand side of (7) implies an increasing marginal cost of GPT R&D due to the occupational-choice effect.

The labor market clearing condition for production workers (L_i) requires that labor demand be equal to labor supply determined by the occupational choices. Utilizing the BEJK Result 4, $\frac{\theta}{1+\theta} X_{1\nu} = w_{1\nu} L_1$, we have $\frac{\theta}{1+\theta} \frac{X_{1\nu}}{w_{1\nu}} = N_1^M G\left(a_{1\nu}^M\right) + N_1^R G\left(a_\nu^R\right)$ and $\frac{\theta}{1+\theta} \frac{X_{2\nu}}{w_{2\nu}} =$

[10] As there is a continuum of GPT innovators, each innovator is an atomless agent, who takes both the expected GPT monopoly revenue $\beta\left(X_{1\nu+1} + X_{2\nu+1}\right)/\left(1+\theta\right)$ and the expected rate of creative destruction $\lambda\left(R_{\nu+1}\right)$ as given.

$N_2 G\left(a_{1\nu}^M\right)$, which can be combined with the definitions of M_{iv} and R_ν to yield two useful equilibrium conditions for production workers:

$$\frac{\theta}{1+\theta}\frac{X_{1\nu}}{w_{1\nu}} = N_1^M\left[1-\left(\frac{k}{k-1}\frac{N_1^M}{M_{1v}}\right)^{\frac{-k}{k-1}}\right]+N_1^R\left[1-\left(\frac{k}{k-1}\frac{N_1^R}{R_v}\right)^{\frac{-k}{k-1}}\right], \quad (8)$$

$$\frac{\theta}{1+\theta}\frac{X_{2\nu}}{w_{2\nu}} = N_2\left[1-\left(\frac{k}{k-1}\frac{N_2}{M_{2v}}\right)^{\frac{-k}{k-1}}\right]. \quad (9)$$

2.5 Goods Markets

The goods markets differ from those in the BEJK model because of royalty payments $\frac{\beta}{1+\theta}X_{2\nu}$ from the South to the North. That is, the national incomes in the North and the South are, respectively, $Y_{1\nu} = X_{1\nu} + \frac{\beta}{1+\theta}X_{2\nu}$ and $Y_{2\nu} = X_{2\nu} - \frac{\beta}{1+\theta}X_{2\nu}$. These together imply that $X_{12,\nu} = X_{21,\nu} + \frac{\beta}{1+\theta}X_{2\nu}$. Denote the wage gap between the North and South by $w_\nu \equiv w_{1\nu}/w_{2\nu}$, the gap in the initial knowledge stock for differentiated products by $t_0 \equiv \frac{t_{10}}{t_{20}}$, and the gap in entrepreneurial activities by $m_\nu \equiv M_{1\nu}/M_{2\nu}$. In Appendix B, we derive the following equilibrium condition linking relative output scales to the wage and the entrepreneurial activity ratios, w_ν and m_ν:

$$\frac{X_{1\nu}}{X_{2\nu}} = m_\nu w_\nu^{-\theta}\left[\frac{m_\nu t_0 w_v^{-\theta}+\tau^{-\theta}}{m_\nu t_0\left(w_v\tau\right)^{-\theta}+1}\frac{1+\theta-\beta}{1+\theta}+\frac{\beta\tau^\theta}{1+\theta}\right]. \quad (10)$$

Because GPT innovation is the sole driver of economic growth, total revenues and wages evolve according to

$$\frac{X_{i,\nu+1}}{X_{i,\nu}} = \frac{Y_{i,\nu+1}}{Y_{i\nu}} = \frac{w_{i,\nu+1}}{w_{i,\nu}} = \frac{w_{\nu+1}^R}{w_\nu^R} = 1+g\left(R_\nu\right) = 1+\lambda\left(R_\nu\right)\ln\left(\gamma\right). \quad (11)$$

The evolution of these key variables all depend on R&D devoted to GPT. This ensures a well-defined concept of steady-state growth at a common rate, to which we now turn.

3 Balanced Growth Path

A *dynamic world equilibrium* (DWE) is a sequence of a tuple $\{X_{1\nu}, X_{2\nu}, M_{1\nu}, M_{2\nu}, R_\nu, w_{1\nu}, w_{2\nu}\}$ that satisfies, at any innovation time ν and for a given initial value of w_{10}, three occupational choice conditions with two in the North and one in the South summarized by (3) and (7) (where the last encompasses the first-order condition of GPT innovators), two labor market clearing conditions (8) and (9), the trade condition (10) and a law of motion or growth condition (11).

A *balanced growth path* (BGP) equilibrium is a DWE with a constant, common growth rate g. As seen from (11), along a BGP, wages, total revenues and total income all grow at the common rate $g(R)$, whereas all labor variables are constant.

3.1 Characterization of the BGP

The common growth property of BGP enables us to characterize all growing variables in "great ratios."

To do so, we further define $x \equiv X_1/X_2$ and $\chi \equiv N_1/N_2$. From (3) and (10), we can express the output gap (x) as depending positively on entrepreneurship and wage gaps (m and w),

$$x = \psi^{-\frac{1}{k-1}} \chi^{-\frac{1}{k-1}} w m^{\frac{k}{k-1}}, \tag{12}$$

where entrepreneurship and wage gaps are governed by the trade condition:

$$w^{1+\theta} m^{\frac{1}{k-1}} = \psi^{\frac{1}{k-1}} \chi^{\frac{1}{k-1}} \left[\frac{m t_0 w^{-\theta} + \tau^{-\theta}}{m t_0 (w\tau)^{-\theta} + 1} \frac{1+\theta-\beta}{1+\theta} + \frac{\beta \tau^{\theta}}{1+\theta} \right]. \tag{13}$$

That is, both wage and output gaps are functions of the entrepreneurship gap alone.

Now we turn to labor variables which are all time invariant along the BGP. Using (8), one obtains an entrepreneur-researcher trade-off relationship in the North: $J_1 M_1^{\frac{k}{k-1}} + J_2 R^{\frac{k}{k-1}} = \psi^{\frac{1}{k-1}} N_1^{\frac{k}{k-1}}$, or,

$$M_1(R) = \left(\frac{\psi^{\frac{1}{k-1}} N_1^{\frac{k}{k-1}} - J_2 R^{\frac{k}{k-1}}}{J_1} \right)^{\frac{k-1}{k}}, \tag{14}$$

where $J_1 \equiv \frac{\theta}{1-\beta} \left(\frac{k}{k-1}\right)^{-\frac{1}{k-1}} + \left(\frac{k}{k-1}\right)^{\frac{-k}{k-1}}$ and $J_2 \equiv \left(\frac{k}{k-1}\right)^{-\frac{k}{k-1}} \left(\frac{\psi}{1-\psi}\right)^{\frac{1}{k-1}}$ are constants, measuring the effective shares of entrepreneurial and research activities, respectively. Even though there are two types of individuals in the North, there exists a labor-market link that establishes a negative relation between entrepreneurship M_1 and GPT R&D effort R. An increase in M_1 or R reduces the labor supply for production, but an increase in M_1 increases the labor demand. When M_1 increases, labor demand increases and the labor supply from the entrepreneurial type decreases, and these imply that the short-fall in labor supply needs to be made up by the research type and so R decreases.

We can also manipulate (9) and (3) to derive the revenue-wage ratio in the South,

$$\frac{X_2}{w_2} = \frac{1+\theta}{\theta} N_2 \left[1 - \left(\frac{k}{k-1}\frac{N_2}{M_2}\right)^{-\frac{k}{k-1}}\right] \tag{15}$$

as well as the entrepreneurial activities in the South,

$$M_2 = \left[\frac{\theta}{1-\beta}\left(\frac{k}{k-1}\right)^{-\frac{1}{k-1}} + \left(\frac{k}{k-1}\right)^{-\frac{k}{k-1}}\right]^{\frac{1-k}{k}} N_2 = J_1^{\frac{1-k}{k}} N_2, \tag{16}$$

which are both constant due to the absence of the entrepreneur-researcher trade-off. As a result, the entrepreneurship gap is a function of North R&D activities alone, governed by the following negative relationship: $m(R) = \frac{M_1(R)}{M_2} = \left[\psi^{\frac{1}{k-1}} N_1^{\frac{k}{k-1}} - J_2 R^{\frac{k}{k-1}}\right]^{\frac{k-1}{k}} / N_2$, implying that both wage and output gaps are also functions of R: $w = w(R)$ and $x = x(R)$.

Finally, along the BGP and with (15) and (16), (7) can be written as:

$$\frac{\beta\left(X_1+X_2\right)(1+g)}{(1+\theta)\, w_1} = \frac{\beta k N_2}{\theta k + (1-\beta)(k-1)} \frac{(1+x)(1+g)}{w} = \left(\frac{k-1}{k}\frac{R}{N_1^R}\right)^{\frac{1}{k-1}} \frac{r+\lambda(R)}{\lambda'(R)}, \tag{17}$$

where the left-hand side, $\frac{\beta(X_1+X_2)(1+g)}{(1+\theta)w_1} = \frac{\beta k N_2}{\theta k+(1-\beta)(k-1)}\frac{(1+x)(1+g)}{w}$, is the R&D multiplier on the BGP, and the right-hand side is strictly increasing in R. As x, g, and w are all functions of R, the second equality in (17) demonstrates this is the equation by which the equilibrium R is determined. We further use (15), together with (14), and (16), to derive, aside from a constant,

$$\frac{\beta\left(X_1+X_2\right)(1+g)}{(1+\theta)\, w_1} \propto \left\{N_1\left[1 - \left(\frac{1}{1-\psi}\right)^{\frac{1}{k-1}}\left(\frac{k-1}{k}\frac{R}{N_1}\right)^{\frac{k}{k-1}}\right] + \frac{N_2}{w(R)}\right\}(1+g(R)). \tag{18}$$

Combining this with (17), one can see that had one ignored the effects via $w(R)$ and $g(R)$, a unique level of R&D activity would have been pinned down immediately. Yet, to assess the effect of trade liberalization, the general equilibrium object wage gap w and the growth rate g are both crucial, thereby requiring more thorough analysis with which we now proceed.

3.2 Trade and Sustained Growth

Before establishing the measures of welfare gains from trade, we must explain the mechanism under which trade costs may affect long-run growth. From (18), observe that the R&D multiplier consists of three effects of GPT R&D effort: (i) a labor-reallocation effect as a result of occupational choice via the term in the square bracket, (ii) a growth effect via $g(R)$ reflecting the fact that global aggregate revenue grows with the GPT innovation, and (iii) a general equilibrium wage-gap effect via $w(R)$. In general, how the multiplier changes in R is ambiguous.

Note that for a given R, a reduction in τ only affects the multiplier via the wage gap w, and the right-hand side of (17) is unaffected. Under the North-South structure, $t_0 = t_{10}/t_{20} > 1$, we generally expect that the wage gap $w > 1$.[11] Trade liberalization tends to reduce this gap, as it makes the global economy more integrated and spurs the labor demand in the South relative to that in the North. Thus, trade liberalization tends to increase the R&D multiplier for a given R. Because the right-hand side of (17) is strictly increasing in R, trade liberalization induces an increase in equilibrium R as long as the multiplier decreases in R or does not increase in R faster than the right-hand side of (17) at the initial equilibrium R. This is illustrated in Case (a) and (b) in Figure 1. But is it possible that the multiplier increases in R faster than the right-hand side as illustrated by Case (c) in Figure 1? This is implausible because it can be ruled out by checking whether equilibrium R_ν would increase when there is a transitory positive shock to the arrival rate, written as $\tilde{\lambda} = \lambda_0 \lambda(R_\nu)$ with $\lambda_0 > 1$ during period ν and $\lambda_0 = 1$ from $\nu + 1$ onward. For any given R_ν, such a shock implies that the left-hand side of (17) increases because $w(R_\nu)$ is unaffected and $g(R_\nu)$ increases temporarily for period ν, and that the right-hand side decreases because the rate of creative destruction $\lambda(R_{\nu+1})$ is unaffected but $\lambda'(R_\nu)$ increases. Thus, equilibrium R_ν must decrease in Case (c), but this is implausible. That is, Case (c) can be ruled out by Samuelson's Correspondence Principle.

[Insert Figure 1 here.]

The key intuition here is that *aggregate world revenue and hence aggregate labor demand in unit of the North's labor* $(X_{1\nu+1} + X_{2\nu+1})/w_{1\nu}$ *increase* in face of trade liberalization because

[11] In addition to $t_0 > 1$, the differences in model primitives between the two countries include relative population size χ and the fact that GPT innovation occurs only in the North, the latter of which should also contribute to $w > 1$. Thus, the only possibility for $w > 1$ not to hold is a very large χ. In our calibrated model, $\chi = 2.3948$, which reflects that "North" countries having larger human capital. But this is not large enough to generate $w \leq 1$ in our calibrated model, which entails $w = 1.38$ under benchmark τ and $w = 28.38$ in the autarkic world. See Sections 4.1 and 4.4 and Appendix D for details.

a more integrated economy implies a stronger demand for the South's labor (relative to the North's) which reduces the wage gap (essentially a factor-price-equalization effect of trade). When R_ν increases as a result, the growth rate g becomes higher, which further implies faster growth in $(X_{1\nu+1} + X_{2\nu+1})/w_{1\nu}$ and so the higher growth rate is sustained. Namely, trade amplifies the spiral effect of growth.

3.3 Welfare Gains from Trade

In this subsection, we first derive welfare formula for a given country, and then compare with the static version of the model to investigate how dynamics matter. We also derive formula for the welfare gap between the North and South.

On the BGP, welfare is measured by $U_n = \int_0^\infty C_{n0} e^{gt} e^{-\rho t} dt = \int_0^\infty C_{n0}^{\xi} e^{-[\rho - \lambda(R)\ln(\gamma)]t} dt$. To ensure finite lifetime utility, we impose a modified Brock-Gale condition in a fashion similar to that in Bond, Wang and Yip (1996): $\rho > \max_R \lambda(R) \ln(\gamma)$. This condition will be met if the maximal arrival rate is not too high and the step size of the ladder, γ, not too large, which will be checked in our quantitative analysis.

In country n, the life time utility on a balanced growth path starting from $t = 0$ is:

$$\begin{aligned} U_n &= \int_0^\infty Q_{n0} e^{-[\rho - \lambda(R)\ln(\gamma)]t} dt = \frac{Y_{n0}/N_n}{P_{n0}} \int_0^\infty e^{-[\rho - \lambda(R)\ln(\gamma)]t} dt \\ &= \frac{1}{\rho - g} \frac{Y_{n0}}{N_n} \frac{1}{P_{n0}} = \frac{1}{\rho - \lambda(R)\ln(\gamma)} \frac{Y_{n0}}{N_n} \frac{1}{P_{n0}}. \end{aligned} \tag{19}$$

We first focus on the North. In Appendix C, we show that

$$\begin{aligned} \frac{U_1'}{U_1} &= \frac{\rho - g}{\rho - g'} \times \frac{Y_{10}'/N_1'}{Y_{10}/N_1} \times \frac{P_{10}}{P_{10}'} \\ &= \underbrace{\frac{\rho - \lambda(R)\ln(\gamma)}{\rho - \lambda(R')\ln(\gamma)}}_{\text{GR}} \times \underbrace{\frac{\left(x' + \frac{\beta}{1+\theta}\right) w'^{-1}}{\left(x + \frac{\beta}{1+\theta}\right) w^{-1}}}_{\text{IG}} \times \underbrace{\left(\frac{\frac{T_{10}}{T_{20}} + \left(\frac{\tau}{w}\right)^{-\theta}}{\frac{T_{10}}{T_{20}} + \left(\frac{\tau'}{w'}\right)^{-\theta}} \right)^{-\frac{1}{\theta}}}_{\text{ACR}}, \end{aligned} \tag{20}$$

where the terms in the last line is defined by each multiplicative term in the previous line. The first term that operates solely through the change in growth rate, and is hence called the *growth-rate (GR) effect*. If the relative wage (w) and output scale (x) are held fixed in face of a trade shock, GR is the only dynamic channel affecting welfare. The second term is derived from the relative change in per capita nominal income Y_{10}/N_1. In the ACR framework where balanced trade is imposed, this term does not usually appear. It appears here because trade is not balanced due to the royalty payments from the South to the North. This term is therefore called *income gains (IG) effect*. The price-index effect is named *ACR* because it equals exactly the ACR formula $(\pi'_{11}/\pi_{11})^{-\frac{1}{\theta}}$, as shown in Appendix C. As the BEJK model is within the ACR framework, it is not surprising that this term appears in the formula. Needless to say, general equilibrium objects w and x are intertwined with GPT R&D effort R, and hence there may be dynamic gains from trade that operate through these general equilibrium effects and affecting both the IG and ACR terms.

In our quantitative analysis, we are interested in the above decomposition, and the results will be examined in a convenient way by rewriting (20) as:

$$1 = \frac{\ln \text{GR}}{\ln(\text{Total Gains})} + \frac{\ln \text{IG}}{\ln(\text{Total Gains})} + \frac{\ln \text{ACR}}{\ln(\text{Total Gains})}. \tag{21}$$

As just mentioned, the GR effect does not reflect the true dynamic gains from trade. To account for the true dynamic gains from trade, we compare our model with a purely static version of the model in which $\kappa = 0$. Obviously, both the equilibrium R and growth rate are zero under such a parameter constraint. From (14) and (16), it is immediate that all entrepreneurship variables become constant (which is a also feature in the BEJK model). The fact that growth rate $g = 0$ in the static model implies that the GR effect disappears. To isolate the contribution of dynamics on the total gains from trade (called dynamic gains from trade), one could write:

$$\frac{U_1'}{U_1} = \frac{\frac{U_1'}{U_1}}{\left.\frac{U_1'}{U_1}\right|_{\text{static}}} \times \left.\frac{U_1'}{U_1}\right|_{\text{static}} = \text{GR} \times \frac{\text{IG} \times \text{ACR}}{\text{IG}|_{\text{static}} \times \text{ACR}|_{\text{static}}} \times \left(\text{IG}|_{\text{static}} \times \text{ACR}|_{\text{static}}\right), \tag{22}$$

where the sub-label of static indicates the corresponding value under the static model. The true dynamic gains are thus the first two multiplicative terms, which reflect the GR effect in the dynamic model and the general equilibrium effects of trade on the IG and ACR components that are not captured by the static model.

Equivalently, we can write (22) as:

$$\frac{\ln\left(\text{GR} \times \frac{\text{IG}\times\text{ACR}}{\text{IG}|_{\text{static}}\times\text{ACR}|_{\text{static}}}\right)}{\ln\left(\frac{U_1'}{U_1}\right)} = 1 - \frac{\ln\left(\text{IG}|_{\text{static}} \times \text{ACR}|_{\text{static}}\right)}{\ln\left(\frac{U_1'}{U_1}\right)}, \tag{23}$$

and one can simply compute the share of the dynamic gains by calculating the right-hand side of the above equation after computing the gains from trade in the static model.

Next, we consider the South's welfare by considering the welfare gap between the two countries $u \equiv U_1/U_2$, which, as shown in Appendix C, is given by,

$$u \equiv \frac{U_1}{U_2} = \chi^{-1} \frac{1+\theta}{1+\theta-\beta} \left(x + \frac{\beta}{1+\theta} \right) \left(\frac{\frac{T_{10}}{T_{20}} (w\tau)^{-\theta} + 1}{\frac{T_{10}}{T_{20}} w^{-\theta} + \tau^{-\theta}} \right)^{-\frac{1}{\theta}}, \tag{24}$$

That is, under common growth, the welfare gap is simply a combination of income and price gaps. We next turn to out quantitative analysis of the model.

4 Quantitative Analysis

4.1 Calibration

In our quantitative implementation, we specify the arrival rate as: $\lambda(R) = \kappa R^{\epsilon}$, with $\kappa > 0$, $\epsilon \in (0, 1)$, a form commonly used in the innovation and growth literature. We calibrate our model economy to match the world economy. An important task in the calibration is to reasonably define the North and South countries. Our approach is to take the Penn World Tables (PWT9.0) data and select those countries with at least two-thirds of the U.S. real GDP per worker as the North.[12] The countries under consideration are based on availability in other datasets including the World Development Indicators (WDI) and the Global Report of Global Entrepreneurship Monitor (GEM), which help calibrating the number of entrepreneurs, using Early-Stage Entrepreneurial Activity (TEA) in the GEM Global Report together with the WDI.[13]

The key targets to match are (i) the long-run economic growth rate, (ii) the North-South ratios of TFPs and manufacturing outputs, and (iii) the shares of entrepreneurs and researchers in the North. While the full calibration procedure is relegated to Appendix D, we summarize the results in Table 1 here.

[Insert Table 1 here.]

4.2 Welfare Gains from Trade

As a common practice, we measure welfare gains from trade based on one of the following two scenarios: (i) compared with autarky: $\tau = 2$ (benchmark) compared to $\tau \to \infty$

[12] The GDP measure used is output at constant PPP in 2011 million US$ (rgdpo). All data are averaged over 2000-2014.

[13] Early-Stage Entrepreneurial Activity is the fraction of surveyed individuals (18-64 years old) who are involved in a nascent firm or new firm or both (if doing both, still counted as one active person).

(autarky); (ii) compared with free trade: $\tau = 1$ (free trade) compared to $\tau = 2$ (benchmark). We first show the results of total gains from trade, and decompose them into the GR, IG, and the ACR effects. Then, we conduct three counterfactuals. As mentioned in Section 3, the GR effect in the simple decomposition is not necessarily the dynamic component of the total gains, if the general equilibrium effects are non-negligible. Thus, we compare with the static model in which $\kappa = 0$ (zero arrival rate) to find the "true" dynamic share of the gains from trade. The second counter-factual is to investigate the role of skill distribution by shutting this down. In particular, we compare with the model in which labor is homogeneous, and so researchers, entrepreneurs, and production workers are of the same type of labor. The third counter-factual is to investigate the role of occupational choice by comparing with a model in which entrepreneurial activities are fixed at their equilibrium values.

4.2.1 Simple decomposition

We first conduct a simple decomposition by the three factors in (21) to understand the separate effects through growth rate, income gains, and price index (i.e., the ACR statistics). The results are summarized in Table 2.

[Insert Table 2 here.]

To begin, we examine gains from trade as compared with autarky. The growth rate improves from 1.8511% to 1.8977%, increasing by 0.047 percentage points. The growth-rate effect, income-gains effect, and the ACR statistic are 1.0422, 1.0063, and 1.0044, respectively.

In a model with balanced trade and single input (which are typical assumptions in the trade literature), the nominal income effect does not exist (by choosing the numéraire). But in this model where trade is imbalanced due to the royalty payment, the North benefits from trade through this monopoly rent, thereby yielding a positive income gain. Interestingly, this IG effect is even larger than the ACR statistic compared to autarky. With positive income gains and the large share of the growth-rate effect via GR, our results indicate that the ACR statistic understates total gains from trade by a large amount.

Also note that the ACR effect is much larger in the second scenario than the first. The intuition is that by Shephard's Lemma, the ACR statistic reflects essentially a direct effect of trade costs on prices, which operates mostly through imported goods. Compared with the first scenario, trade costs are lower in the second one, and thus demand for imported goods tend to be larger, thus amplifying the direct price effect. In addition, the

South's wages also amplifies the direct effect of trade cost on prices because they reflect the marginal costs in the South. Again, trade costs are lower in the second scenario, in which case the South's wages tend to be closer to the North's, which amplifies the direct price effect and results in a larger ACR statistic.

To understand why the GR effect is larger in the first scenario, recall from (17) and (18) that for a given R, the right-hand side of (17) is fixed, whereas the left-hand side changes due to change in the relative wage w. With a large trade cost w is necessarily high (as can be seen in Section 4.4 below). Moving from autarky to the benchmark trade costs thus reduces w drastically and results in a sizable increase in R compared with the second scenario. In other words, when trade barriers are high, there are weaker incentives to conduct GPT innovation due to an overall smaller market size. This illustrates why reducing trade barriers from autarky would have a large effect on R&D activity.

4.2.2 Comparison with the static model

The simple decomposition exercises do not reflect the *true dynamic gains* from trade because the GPT innovation also affects wage and hence output gaps that affect the IG and the ACR measures. To rectify this problem, we obtain pure static gains via a counterfactual exercise by shutting down the sole growth driver from GPT innovation. Specifically, we do this by setting $\kappa = 0$ (and hence $R = 0$ in equilibrium), under which not only GR = 1 by construction, but IG and ACR are also adjusted to exclude the R&D effect – so these two components are generally different from those reported in Table 2. By subtracting pure static gains from the total gains, we obtain the true dynamic gains from trade and the dynamic share can then be calculated by (23). The results are summarized in Table 3.

[Insert Table 3 here.]

The gains from trade in the counterfactual static model under the two scenarios are 1.146% and 1.620% , respectively. These numbers are to be compared with the gains in the dynamic model, 5.336% and 3.184%, as reported above. Hence, the gains in the dynamic model are 4.66 times larger than the gains in static model in the first scenario and 1.97 times larger in the second. The dynamic shares in the two scenarios are 78.1% and 48.7%, respectively. It turns out the dynamic shares are rather similar to the growth-rate effects in Table 2, suggesting that the general equilibrium effect of GPT innovation on gains from trade via wage and output gaps are modest. The intuition behind why the dynamic (static) share is larger (smaller) in the first scenario compared with the second are the same as those explained in Section 4.2.1 for why the GR (ACR) effect is larger (smaller) in the first scenario.

4.2.3 Entrepreneurial occupational choice

We now consider a counterfactual in which entrepreneurial occupational choice is eliminated. We do so by fixing M_1, M_2, and m at their benchmark values. Under this given set of equilibrium values M_1, M_2, and m, we use (13) to solve for w, use (12) to solve for x, and then use (17) to solve for R. The welfare results are summarized in Table 4.

[Insert Table 4 here.]

Here, we see only very slight changes compared with Table 3. To understand this, observe that the multiplier in the absence of occupational choice can be written as:

$$\frac{\beta\left(X_1+X_2\right)(1+g)}{(1+\theta)\,w_1} \propto \left[(\psi N_1)_1^{-\frac{1}{k-1}}\, M_1^{\frac{k}{k-1}} + N_2^{-\frac{1}{k-1}} M_2^{\frac{k}{k-1}}/w(R)\right](1+g(R))\,. \qquad (25)$$

Following a similar analysis in Section 3.2 and for a given R, when trade liberalization reduces the wage gap, the multiplier also increases. Thus, qualitatively trade still promotes growth and dynamic gains in the same way as the benchmark model. However, quantitatively, shutting down entrepreneurial occupational choice turns out to have only negligible effects. Comparing (18) and (25) shows that the difference lies in how much an increase in GPT R&D R reduces the entrepreneurial activities in the North. We find that M_1 at $\tau = \infty, 2, 1$ are 0.97494, 0.974661, and 0.974541, respectively, only slightly lowered despite a large reduction of trade cost from autarky to free trade. This suggests that the positive effect of trade liberalization on entrepreneurship is essentially offset by the negative effect of labor reallocation that discourages entrepreneurship.

The above finding also helps explain why we have always obtained a positive income gains effect along the BGP. Observe from (3) that the sales-wage ratio X_i/w_i is proportional to $(M_i)^{\frac{k}{k-1}}$. Also observe that the North's income in unit of North labor implies:

$$\frac{Y_1}{w_1} = \frac{1+\theta-\beta}{1+\theta}\frac{X_1}{w_1} + \frac{\beta\left(X_1+X_2\right)(1+g)}{(1+\theta)\,w_1}.$$

That is, it can be decomposed into the R&D multiplier and a term that is proportional to the sales-wage ratio. Trade liberalization induces only a small fall in M_1 and hence in the sales-wage ratio, but generates a sizable amplifying multiplier effect, thus leading to a positive income gains effect.

4.2.4 Homogeneous labor

Finally, we consider the quantitative impact of labor heterogeneity. As shown in Appendix E, the model mechanism of trade and growth as described in Section 3.2 is generally preserved if we instead assume homogeneous labor so that the resource constraint becomes $R_i + M_i + L_i = N_i$, where R_i, M_i, and L_i are the number of researchers, entrepreneurs, and production workers, respectively. Removing the skill distribution and making these three occupations perfect substitutes greatly simplifies the model. To highlight the role of heterogeneous labor, the counterfactual analyses are conducted under the same parameters. The results of gains from trade are summarized in Table 5.

[Insert Table 5 here.]

Compared to the benchmark model, both the total gains and the dynamic share in the homogeneous model are around 25% and 10% lower, respectively. Consistent with these results, GR falls and the ACR effect increases in both scenarios. The IG effects are similar to the benchmark case. These results are intuitive. The GPT R&D labor is quite a small fraction of the population; in the presence of a fat-tailed skill distribution, the marginal researcher is still quite talented. Lower trade barriers incentivize the GPT innovators to tap these talented people more and produce more intertemporal spillovers.

4.3 Comparative Statics

nest we consider how the total gains from trade and dynamic share change when some key parameters change. In particular, we are interested in the comparative statics with respect to: (i) a fatter tail of the productivity distribution (a lower θ), (ii) a fatter tail of the distribution of worker skill (a lower k), (iii) an improvement in the efficacy of R&D (a higher κ), (iv) stronger bargaining power of the GPT firm (a higher β), and (v) a higher initial relative knowledge stock in the North (a higher t_0). We focus on scenario (i), i.e., comparing the benchmark equilibrium with autarky. The results are shown in Table 6.

[Insert Table 6 here.]

A smaller θ implies larger productivity dispersion among firms as well as higher average productivity. Larger productivity dispersion means larger scope of comparative advantage and higher productivity increases the importance of trade. Thus, it is not surprising to find that the total gains from trade increases. Moreover, a smaller θ implies that the share of profit in aggregate sales, $1/(1+\theta)$, becomes larger, and this increases the R&D multiplier, and hence R and the growth rate as well. This dynamic effect adds further to the total gains from trade. However, the dynamic share actually decreases, implying the former two effects dominate the dynamic effect.

A smaller k implies larger skill dispersion and better top researchers. This is obviously conducive to the GPT innovation and leads to higher growth rate. The increase in innovation effort R induced by trade liberalization can be larger with a smaller k, and this leads to a larger dynamic share of total gains. The total gains from trade increases because of a larger dynamic component and an overall more capable population.

A higher κ implies a higher arrival rate of GPT innovation for a given R, and this incentivizes the GPT R&D despite a creative destruction effect, which is a standard outcome in quality ladders models. It is also intuitive that this increases both the total gains from trade and the dynamic share. A larger β also encourages the GPT R&D and increases the total gains from trade. However, the dynamic share decreases because of a stronger income-gains (IG) effect.

From (17) and (18), we see that t_0 affects only the right-hand side of (17) directly via the wage gap w in (18). A higher t_0 implies a larger wage gap w and hence a smaller R&D multiplier, resulting a lower growth rate. With a higher t_0, the production activities are more concentrated in the North, and hence the gains from reducing trade barriers becomes smaller. However, as a higher t_0 makes the North-South structure more uneven in a static sense, the dynamic channel of gains from trade highlighted by (17) and (18) becomes more important.

4.4 North-South Welfare Gap

This subsection investigates the welfare gap between the North and South countries, i.e., $u = U_1/U_2 = \frac{Y_1/N_1}{Y_2/N_2}\frac{P_2}{P_1} \equiv \frac{y_1}{y_2}\frac{P_2}{P_1} \equiv y/\left(\frac{P_1}{P_2}\right)$, using the formula derived in (24). The results of the welfare gap u and the two sub-components – the (per capita) nominal income gap y and the price gap (P_1/P_2) are shown in Table 7. To further understand the changes here, we also report the corresponding wage gap w.

[Insert Table 7 here.]

Table 7 shows that trade liberalization reduces the North-South welfare (or real income) gap, and the reduction is especially sharp from benchmark $\tau = 2$ to free trade. This may seem a little counter-intuitive because one may tend to think that the effect of trade opening from autarky is stronger. This is, indeed, the case for nominal income gap (y), and changes in y closely follows changes in w. However, the large reduction in nominal income gap or wage gap is largely offset by the large reduction in the price gap, thereby resulting in small changes in the welfare or real income gap.

The price gap in the autarkic world (taking $\tau \to \infty$) is given by $\frac{P_1}{P_2} = \left(\frac{t_0 w^{-\theta} + \tau^{-\theta}}{t_0 (w\tau)^{-\theta} + 1}\right)^{-\frac{1}{\theta}} \to t_0^{-\frac{1}{\theta}} w$, which consists of two components: (i) the technology advantage in the North (the calibrated value of $t_0^{-\frac{1}{\theta}} \approx 0.53$ suggests the price gap is close to one-half); (ii) the wage gap that reflects the higher marginal cost of production in the North (the value of w in autarky at the calibrated model is 28.38, which suggests a much larger price gap even with technology gap accounted). When the economy moves from the autarkic equilibrium to the benchmark economy, there is a strong general equilibrium effect that sharply reduces the wage gap. Due to this large drop in the wage gap, the technology gap becomes the dominant force in the benchmark economy (with $\tau = 2$), thus resulting in a price gap $P_1/P_2 < 1$, which, in turn, helps maintain the high welfare gap under the benchmark. As an economy moves toward free trade, the effect of the technology gap on prices vanishes, subsequently leading to the law of one price with $P_1/P_2 = 1$. This, along with reduced wage and nominal income gaps, then yields essentially a negligible welfare gap in a frictionless world.

4.5 Sensitivity Analysis

We now perform sensitivity analysis. In particular, we check our findings when we alter each of the following four preset parameters one-by-one: (i) set $\theta = 2.74$ as in Simonovska and Waugh (2014) or $\theta = 4.12$ as in Head and Mayer (2014), (ii) set $k = 1.515$ as in Gabaix, Lasry, Lions, and Moll (2016) and $k = 1.667$ as in Jones (2015), (iii) set $\psi = 0.8$ (more potential researchers) or $\psi = 0.9$ (fewer potential researchers), and (iv) set $\tau = 1.5$ (lower trade costs) or $\tau = 2.5$ (higher trade costs). In each case, we recalibrate the model economy and then recompute the welfare gains and the counter-factual static gains.[14] The results are summarized in Table 8.

[Insert Table 8 here.]

First of all, we find tail indices of firm productivity (θ) and worker skill (k) essential for the importance of the total gains from trade and dynamic share. As the productivity

[14]Note that the directions of changes of total gains from trade and dynamic shares might not be exactly the same as those comparative statics reported in Table 6 since the exercises here are based on *re-calibrated* models.

distribution becomes more dispersed (θ changes from its benchmark value of 4 to 2.75), i.e., trade elasticity [in absolute value] becomes smaller), the total gains from trade slightly decrease from the benchmark of 5.34% to 5.11% when comparing with autarky, and the dynamic share reduces sharply from 78.1% to 16.5%. In fact, a smaller θ implies more productivity dispersion among firms, and this implies that the static gains from trade are more important, explaining the smaller share of dynamic gains. The results are the opposite when θ increases to 4.12.

When the skill distribution becomes fatter (k lowers to 1.515), the total gains when comparing with autarky is about 8.14%, and the dynamic share about 89.4%; both are substantially larger than the benchmark case. The increase in R induced by trade liberalization can be much larger if the underlying skill distribution has a fatter tail, and this explains the larger gains when k is smaller. The results are the opposite when k increases to 1.667. An interesting contrast here is that the dynamic share becomes smaller when the tail of the productivity distribution becomes fatter, whereas it becomes larger when the tail of the skill distribution becomes fatter. Taken together, the welfare results are quite sensitive to how fat-tailed the productivity and skill distributions are.

The welfare results are clearly insensitive to varying ψ, the population share of the entrepreneurial type. This is consistent with the observation in Section 4.2.3 that the effects of entrepreneurial occupational choice are quite small. Finally, we pick the benchmark value of $\tau = 2$ in the midrange of empirically estimated bilateral trade costs, $(1.5, 2.5)$. The first takeaway is that the total gains from trade and dynamic share are somewhat sensitive to the choice of τ to represent the current economy, much less so than the abovementioned tail indices but more so than ψ. Second, the larger (smaller) gains from autarky to $\tau = 1.5$ (2.5) than the benchmark value are straightforward results of gains from trade. Third, that the dynamic share for $\tau = 1.5$ (2.5) is smaller (larger) than the benchmark case can be understood with the reasoning for the difference between the two scenarios in Tables 2 and 3. Namely, when trade liberalizes from autarky to some τ, the larger (smaller) the τ, the more the GR (ACR) effect operates as explained in Section 4.2.1.

5 Concluding Remarks

This paper develops a general equilibrium model of international trade and endogenous growth via innovation. There are two types of innovation with the first being the upgrading of general purpose technology and the second being the entrepreneurial activities that develop differentiated products. We highlight the channel through which trade liberalization increases the GPT R&D and hence the long run economic growth. It is fundamentally a market-size effect as an R&D multiplier.

The welfare formula is a product of a growth-rate effect, an income-gains effect, and the ACR statistic. When compared with autarky, the total gains from trade are 5.34%, and dynamics accounts for 78% of the gains. Whereas the total gains from trade are not particularly large compared with other studies on dynamic gains from trade, the dynamic share found in this model is higher than most other studies except Perla et al. (2019). We also find that the dynamic share is well approximated by the growth-rate effect, suggesting that shutting down dynamics does not entail large general equilibrium effects on wages and revenues. When trade is liberalized from a benchmark trade cost to frictionless trade, the dynamic share becomes smaller at 48.7%, accompanied by a sizable ACR share. This indicates that the dynamic effects are stronger when trade costs are larger, but the ACR effect is stronger when trade costs are smaller.

In the counter-factual in which labor is homogenous, the total gains from trade and dynamics falls about 25% and 10%, respectively, highlighting the importance of taking worker heterogeneity into account. We also find that fatter tails of firm-productivity and worker-skill distributions increase the total gains from trade, but a fatter tail of skill (productivity) distribution leads to a larger (smaller) dynamic share. We also find that quantitatively, the welfare results are quite sensitive to these tail indices. We have also found large effects of trade liberalization on narrowing the North-South welfare gap.

It is clear that GPT innovation in our model can be combined with other trade models in which the profits/rents of firms can be shared with the GPT firms in a similar fashion. Hence, our model can be generalized to incorporate process or product innovation or other forms of technology diffusion, as considered by Perla et al. (2019), Sampson (2016), Impullitti and Licandro (2017), and Hsieh et al. (2019). It is an interesting avenue of future research because it enables us to decompose dynamic welfare and productivity gains from trade into various technology-related mechanisms.

Reference

Aghion, P., & Howitt, P. (1992). A Model of Growth through Creative Destruction. *Econometrica*, 60(2): 323-351.

Alessandria, G., H. Choi, and K. Ruhl (2018). "Trade Adjustment Dynamics and the Welfare Gains from Trade," NBER Working Paper 20663.

Anant, T.C.A., E. Dinopoulos, and P. Segerstrom (1990), "A Schumpeterian Model of the Product Life Cycle," *American Economic Review,* 80(5): 1077-1091.

Arkolakis, C., A. Costinot, D. Donaldson, and A. Rodríguez-Clare (2019), "The Elusive Pro-Competitive Effects of Trade," *Review of Economic Studies*, 86(1): 46-80.

Arkolakis, C., A. Costinot, and A. Rodriguez-Clare. (2012). "New trade models, same old gains?," *American Economic Review*, 102(1): 94-130.

Baldwin, R. (1992), "Measurable Dynamic Gains from Trade," *Journal of Political Economy*, 10(1): 162-174.

Berkes, E., K. Manysheva, and M. Mestieri (2019), "World Innovation: Evidence from 100 years of Patent Data," Working Paper, Northwestern University.

Bernard, A. B., J. Eaton, J. B. Jensen and S. Kortum (2003), "Plants and Productivity in International Trade," *American Economic Review,* 93(4): 1268-1290.

Blaum, J., C. Lelarge, and M. Peters (2018), "The Gains from Input Trade with Heterogeneous Importers," *American Economic Journal: Macroeconomics,* 10(4): 77-127.

Bloom, N., M. Draca, and J. van Reenen (2016), "Trade Induced Technical Change? The Impact of Chinese Imports on Innovation, IT and Productivity, *Review of Economic Studies*, 83(1): 87–117.

Brooks, W. J. and P. S. Pujolas (2019), "Gains from Trade with Variable Trade Elasticities," *International Economic Review*, forthcoming.

Buera, F. J., and E. Oberfield (2019), "The Global Diffusion of Ideas," *Econometrica*, forthcoming.

Caliendo, L., R. C. Feenstra, J. Romalis, and A. M. Taylor (2017), "Tariff Reductions, Entry, and Welfare: Theory and Evidence for the Last Two Decades," NBER Working Paper No. 21768.

Caliendo, L. and F. Parro (2015), "Estimates of the Trade and Welfare Effects of NAFTA," *Review of Economic Studies*, 82(1): 1-44.

Chang, P.-L., and W. Jin (2018), "The GATT/WTO Welfare Effects: 1950-2015," SMU Economics and Statistics Working Paper No. 02-2018.

di Giovanni, J., A. A. Levchenko and J. Zhang (2014), "The Global Welfare Impact of China: Trade Integration and Technological Change," *American Economic Journal: Macro-*

economics, 6(3): 153-183.

Eaton, J. and S. Kortum (2002), “Technology, Geography, and Trade,” *Econometrica*, 70(5): 1741–1779.

Edmond, C., V. Midrigan, and D. Y. Xu (2015), “Competition, Markups, and the Gains from International Trade,” *American Economic Review,* 105(10): 3183–3221.

Egger, P. H. and N. Loumeau (2018), “The Economic Geography of Innovation,” working paper, ETH Zürich.

Fajgelbaum, P. D., and Khandelwal, A. K. (2016), “Measuring the unequal gains from trade,” *Quarterly Journal of Economics*, 131(3): 1113-1180.

Feenstra, R. C., and Weinstein, D. E. (2016). “Globalization, Markups and U.S. Welfare,” *Journal of Political Economy*, forthcoming.

Gabaix, X., J.-M. Lasry, P.-L. Lions, B. Moll (2016), “The Dynamics of Inequality,” Econometrica, 84 (6) :2071-2111.

Greenaway, D., W. Morgan, and P. Wright (2002), “Trade Liberalisation and Growth in Developing Countries,” *Journal of Development Economics*, 67(2): 229–244.

Halpern, L., M. Koren, and A. Szeidl (2015), “Imported Inputs and Productivity,” *American Economic Review*, 105(12): 3660–3703.

Head, K. and T. Mayer (2014), “Gravity Equations: Workhorse,Toolkit, and Cookbook," in *Handbook of International Economics*: 131-195. Elsevier.

Helpman, E. and M. Trajtenberg (1998), “A Time to Sow and a Time to Reap: Growth Based on General Purpose Technologies,” in Helpman, E. (ed.), *General Purpose Technologies and Economic Growth*. Cambridge: MIT Press.

Holmes, T. J., W.-T. Hsu, and S. Lee (2014), “Allocative Efficiency, Mark-ups, and the Welfare Gains from Trade," *Journal of International Economics*, 94(2): 195–206.

Hsieh, C.-T., P. J. Klenow and I. Nath (2019), “A Global View of Creative Destruction," working paper, University of Chicago.

Hsieh, C. T., and Ossa, R. (2016), “A global view of productivity growth in China,” *Journal of International Economics*, 102: 209-224.

Impullitti, G. and O. Licandro (2017), “Trade, Firm Selection, and Innovation: the Competition Channel,” *Economic Journal*, 128 (608): 189-229.

Jones, C. I. (2015), “Pareto and Piketty: The Macroeconomics of Top Income and Wealth Inequality,” *Journal of Economic Perspectives* 29(1): 29-46

Melitz, M. J. (2003), “The impact of trade on intra-industry reallocations and aggregate industry productivity,” *Econometrica*, 71(6): 1695-1725.

Mokyr, J. (1992), The Lever of Riches: Technological Creativity and Economic Progress, Oxford University Press.

Mokyr, J. (2005), "Long-Term Economic Growth and the History of Technology," in Handbook of Economic Growth, Elsevier B.V., 1113-1180.

Perla, J., C. Tonetti, and M. Waugh (2019), "Equilibrium Technology Diffusion, Trade, and Growth," NBER Working Paper No. 20881.

Ravikumar, B., A. M. Santacreu, and M. Sposi (2018), "Capital Accumulation and Dynamic Gains from Trade," the Federal Reserve Bank of St. Louis Working Paper 2017-005D.

Simonovska, I., and Waugh, M. E. (2014a), "The elasticity of trade: Estimates and evidence," *Journal of International Economics*, 92(1): 34-50.

Simonovska, I., and Waugh, M. E. (2014b), "Trade models, trade elasticities, and the gains from trade" NBER Working Paper w20495.

Table 1: Calibration

Panel A: Preset Parameters

Parameter	Value	Meaning	Notes
ρ	0.03	time preference rate	These three parameter values are standard in the R&D-based endogenous growth literature.
r	0.05	real interest rate	
γ	1.014	GPT innovation ladder size	
θ	4	Fréchet shape parameter of productivity draws	Simonovska and Waugh (2014a). Sensitivity analysis will be performed.
τ	2	trade cost	In the ballpark of estimated bilateral trade costs in the literature. Sensitivity analysis will be performed.
ψ	0.85	The fraction of the entrepreneurial type in the population	To reflect low R&D labor share in the population. Sensitivity analysis will be performed.

Panel B: Calibrated Parameters

Parameter	Value	Meaning	Target
k	2	Pareto shape parameter of the ability distribution	Share of income earned by top 1 percent income earners in the US, 0.1903
t_0	12.719	Ratio of initial technology stock between the North and South	Ratio of TFP of the North countries to the South ones, 1.8885
β	0.1222	The bargaining power parameter of the GPT firm	These four parameters (β, χ, κ, and ε) are jointly calibrated from the share of entrepreneurs and and that of researchers in the North population, the manufacturing output ratio between the North and South, and the long-run economic growth rate.
χ	2.3948	Relative labor endowment	
κ	2.510	Efficacy of GPT innovation	
ε	0.2741	Curvature parameter of GPT innovation	

Notes: See Appendix D of the paper for the details of calibration.

Table 2: Simple Decomposition of Total Gains from Trade

Compared with	Change in growth rate (% points)	Gains from trade (% change)	Decomposition of Gains from Trade		
			GR effect	IG effect	ACR effect
(i) Autarky	0.047	5.336	1.042	1.006	1.004
share (%)			79.6	12.0	8.4
(ii) Free Trade	0.017	3.184	1.016	1.003	1.013
share (%)			49.7	8.4	41.9

Table 3: Comparison between Dynamic and Static Models

Compared with	Total gains from trade (% change)	Static gains from trade(% change)	Ratio of total to static gains	Dynamic share (%)
(i) Autarky	5.336	1.146	4.66	78.1
(ii) Free Trade	3.184	1.620	1.97	48.7

Table 4: Gains from Trade in the Model with Homogeneous Labor

Compared with	Total gain from trade (%)	Static gains from trade (%)	dynamic share (%)	GR share (%)	IG share (%)	ACR share (%)
(i) Autarky	4.03	1.17	70.6	75.2	12.8	12.0
(ii) Free Trade	2.83	1.72	38.8	41.5	7.7	50.8

Table 5: Dynamic vs Static Models in the Absence of Occupational Choice

Compared with	Total gains from trade (% change)	Static gains from trade(% change)	Ratio of total to static gains	Dynamic share (%)
(i) Autarky	5.420	1.147	4.72	78.8
(ii) Free Trade	3.219	1.624	1.98	49.6

Table 6: Comparative Statics

	$\theta\downarrow$	$k\downarrow$	$\kappa\uparrow$	$\beta\uparrow$	$t_0\uparrow$
Growth rate	+	+	+	+	-
Total gains from trade	+	+	+	+	-
Dynamic share	-	+	+	-	+

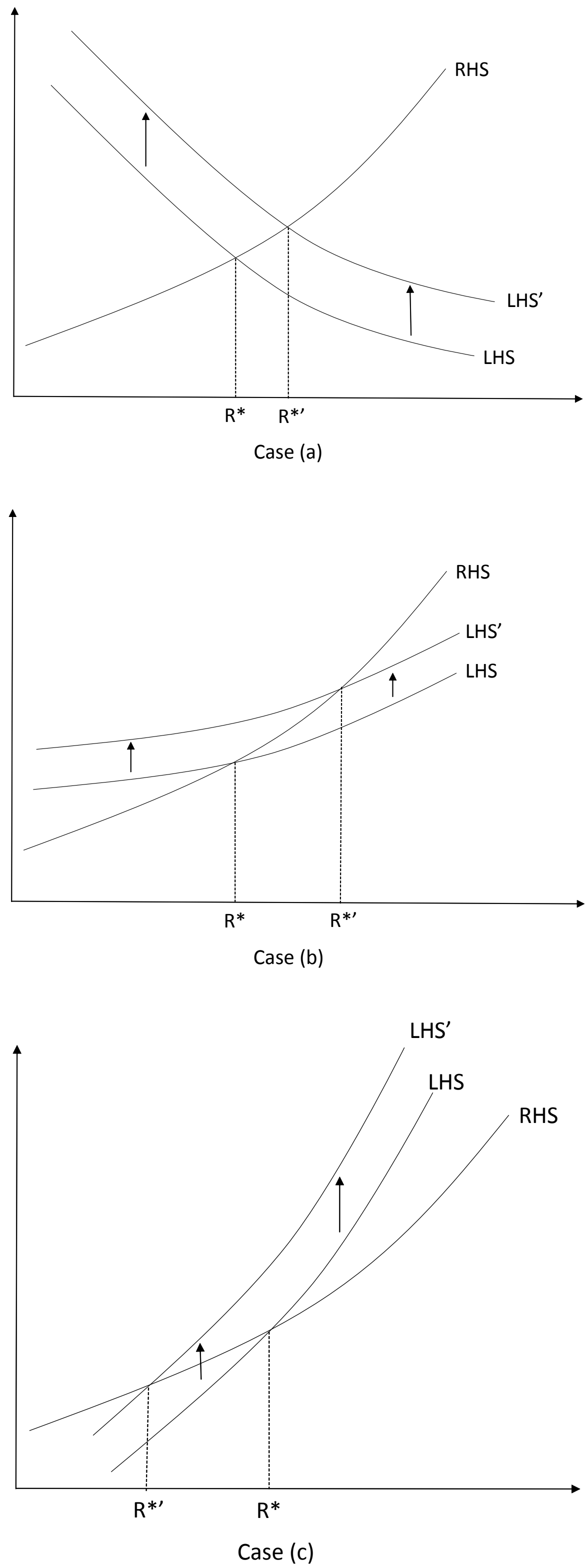
Table 7: North-South Welfare and Income Gaps

	autarky	benchmark	free trade
Welfare Gap (u)	1.93	1.86	1.01
Nominal Income Gap (y)	29.02	1.42	1.01
Price Gap (P_1/P_2)	15.03	0.76	1
Wage Gap (w)	28.38	1.38	0.98

Table 8: Sensitivity Analysis

Compared with	Total gains from trade (% change)	Static gains from trade(% change)	Ratio of total to static gains	Dynamic share (%)
θ=2.74				
(i) Autarky	5.111	4.252	1.20	16.5
(ii) Free Trade	3.309	2.953	1.12	10.6
θ=4.12				
(i) Autarky	6.420	0.936	6.86	85.0
(ii) Free Trade	3.570	1.533	2.33	56.6
k=1.515				
(i) Autarky	8.138	0.835	9.75	89.4
(ii) Free Trade	4.277	1.491	2.87	64.7
k=1.667				
(i) Autarky	4.102	1.664	2.47	58.9
(ii) Free Trade	2.745	1.841	1.49	32.6
ψ=0.8				
(i) Autarky	5.344	1.132	4.72	78.4
(ii) Free Trade	3.145	1.581	1.99	49.3
ψ=0.9				
(i) Autarky	5.343	1.157	4.62	77.9
(ii) Free Trade	3.223	1.658	1.94	48.2
τ=1.5				
(i) Autarky	5.920	1.697	3.49	70.7
(ii) Free Trade	2.120	0.986	2.15	53.2
τ=2.5				
(i) Autarky	5.108	0.909	5.62	81.8
(ii) Free Trade	3.631	1.891	1.92	47.5

Figure 1: Effect of Trade Liberalization on R&D



Appendix

In the appendix, we provide detailed mathematical derivations of the benchmark model and the extensions, as well as detailed calibration strategies.

A Joint Distribution of Top Two Productivities

The c.d.f. of a Fréchet distribution with scaling parameter S and shape parameter θ is given by $F(z) = e^{-Sz^{-\theta}}$. Following standard procedures, the joint distribution of the top two order statistics Z_1 and Z_2 from N draws from the Fréchet distribution is

$$\Pr\left[Z_1 \leq z_1, Z_2 \leq z_2\right] = e^{-NSz_2^{-\theta}} + N\left[e^{-S\left(z_1^{-\theta} - z_2^{-\theta}\right)} - 1\right] e^{-NSz_2^{-\theta}}. \tag{A.1}$$

In our model, $S = 1$ since $F_i^{\text{draw}}(z) = e^{-z^{-\theta}}$ and $N = \widetilde{T}_{i\nu} = M_{i\nu}\gamma^{v}\tilde{t}_{i0}$. Recall from the main text that $\tilde{t}_{i0} = t_{i0}K$ and $T_{i\nu} = \widetilde{T}_{i\nu}/K$. Plug these into (A.1), and we have

$$\begin{aligned}
\Pr\left[Z_1 \leq z_1, Z_2 \leq z_2\right] &= e^{-\widetilde{T}_{i\nu}z_2^{-\theta}} + \widetilde{T}_{i\nu}\left[e^{-\left(z_1^{-\theta} - z_2^{-\theta}\right)} - 1\right] e^{-\widetilde{T}_{i\nu}z_2^{-\theta}} \\
&= e^{-T_{i\nu}Kz_2^{-\theta}} + T_{i\nu}K\left[e^{-\left(z_1^{-\theta} - z_2^{-\theta}\right)} - 1\right] e^{-T_{i\nu}Kz_2^{-\theta}}.
\end{aligned}$$

Consider the joint distribution of the re-scaled top two order statistics $K^{-1/\theta}Z_1$ and $K^{-1/\theta}Z_2$, and we have

$$\begin{aligned}
\Pr\left[K^{-1/\theta}Z_1 \leq z_1, K^{-1/\theta}Z_2 \leq z_2\right] &= \Pr\left[Z_1 \leq K^{1/\theta}z_1, Z_2 \leq K^{1/\theta}z_2\right] \\
&= e^{-T_{i\nu}z_2^{-\theta}} + \frac{e^{-\left(K^{-1}z_1^{-\theta} - K^{-1}z_2^{-\theta}\right)} - 1}{K^{-1}} T_{i\nu}e^{-T_{i\nu}z_2^{-\theta}}.
\end{aligned}$$

As $\lim_{K\to\infty} \frac{e^{-\left(K^{-1}z_1^{-\theta} - K^{-1}z_2^{-\theta}\right)} - 1}{K^{-1}} = z_2^{-\theta} - z_1^{-\theta}$, we have

$$\lim_{K\to\infty} \Pr\left[K^{-1/\theta}Z_1 \leq z_1, K^{-1/\theta}Z_2 \leq z_2\right] = \left[1 + T_{i\nu}\left(z_2^{-\theta} - z_1^{-\theta}\right)\right] e^{-T_{i\nu}z_2^{-\theta}}.$$

B Goods Markets

National income is given by

$$\begin{aligned}
Y_{1\nu} &= (1-\beta)\,\Pi_{1\nu} + w_{1v}L_1 + \frac{\beta}{1+\theta}\left(X_{1\nu} + X_{2\nu}\right), \\
Y_{2\nu} &= (1-\beta)\,\Pi_{2\nu} + w_{2v}L_2.
\end{aligned}$$

That is, there is an international redistribution of royalty payment from South to North whenever $\beta > 0$. Market clearing condition for final goods is

$$Y_{1\nu} = X_{1\nu} + \frac{\beta}{1+\theta} X_{2\nu} = X_{11,\nu} + X_{12,\nu}, \tag{B.2}$$

$$Y_{2\nu} = X_{2\nu} - \frac{\beta}{1+\theta} X_{2\nu} = X_{21,\nu} + X_{22,\nu}. \tag{B.3}$$

Using $X_{1\nu} = X_{11,\nu} + X_{21,\nu}$ and $X_{2\nu} = X_{12,\nu} + X_{22,\nu}$ to combine with the above, we get

$$X_{12,\nu} = X_{21,\nu} + \frac{\beta}{1+\theta} X_{2\nu} \tag{B.4}$$

The lack of balanced trade in final goods is because country 1's royalty income from country 2 makes up its excessive net import.

Using BEJK Result 2 to rewrite (B.4),

$$\pi_{12} Y_1 = \pi_{21} Y_2 + \frac{\beta}{1+\theta} X_2.$$

Using BEJK Result 1 in conjunction with (B.2) and (B.3), the above equality becomes

$$\frac{T_{2\nu}(w_{2\nu}\tau)^{-\theta}}{\Phi_{1\nu}} \left[X_{1\nu} + \frac{\beta}{1+\theta} X_{2\nu} \right] = \frac{T_{1\nu}(w_{1\nu}\tau)^{-\theta}}{\Phi_{2\nu}} \left[X_{2\nu} - \frac{\beta}{1+\theta} X_{2\nu} \right] + \frac{\beta}{1+\theta} X_2, \tag{B.5}$$

where $\Phi_{1\nu} = T_{1\nu} w_{1v}^{-\theta} + T_{2\nu} (w_{2v}\tau)^{-\theta}$ and $\Phi_{2\nu} = T_{1\nu} (w_{1v}\tau)^{-\theta} + T_{2\nu} w_{2v}^{-\theta}$.

A few algebraic manipulation of (B.5) gives the ratio of total revenues proportional to the ratio of entrepreneurship in skill units:

$$\frac{X_{1\nu}}{X_{2\nu}} = \left[\frac{\Phi_{1\nu}}{\Phi_{2\nu}} \frac{1+\theta-\beta}{1+\theta} + \frac{\beta\tau^{\theta}}{1+\theta} \right] \frac{M_{1\nu}}{M_{2\nu}} w_{\nu}^{-\theta},$$

where $w_\nu \equiv w_{1\nu}/w_{2\nu}$ measures the wage gap. With $T_{i\nu} = M_{i\nu}\gamma^{v} t_{i0}$ and letting $t_0 = \frac{t_{10}}{t_{20}}$ and $m_\nu = \frac{M_{1\nu}}{M_{2\nu}}$, the above becomes

$$\frac{X_{1\nu}}{X_{2\nu}} = m_\nu w_\nu^{-\theta} \left[\frac{m_\nu t_0 w_v^{-\theta} + \tau^{-\theta}}{m_\nu t_0 (w_v \tau)^{-\theta} + 1} \frac{1+\theta-\beta}{1+\theta} + \frac{\beta\tau^{\theta}}{1+\theta} \right].$$

C Welfare Gains from Trade

In country n, the life time utility on a balanced growth path starting from $t = 0$ is

$$\begin{aligned} U_n &= \int_0^\infty Q_{n0} e^{-[\rho-\lambda(R)\ln(\gamma)]t} dt = \frac{Y_{n0}/N_n}{P_{n0}} \int_0^\infty e^{-[\rho-\lambda(R)\ln(\gamma)]t} dt \\ &= \frac{1}{\rho - g} \frac{Y_{n0}}{N_n} \frac{1}{P_{n0}} = \frac{1}{\rho - \lambda(R)\ln(\gamma)} \frac{Y_{n0}}{N_n} \frac{1}{P_{n0}} \end{aligned}$$

We first focus on the North. Using (B.2) and (15), one can rewrite Y_{10}/N_1 as

$$\begin{aligned} \frac{Y_{10}}{N_1} &= \chi^{-1} \left(x + \frac{\beta}{1+\theta} \right) \frac{X_{20}}{N_2} \\ &= \chi^{-1} \left(x + \frac{\beta}{1+\theta} \right) \frac{1+\theta}{\theta} w_{20} \left[1 - \left(\frac{k}{k-1} \frac{N_2}{M_2} \right)^{-\frac{k}{k-1}} \right] \end{aligned} \tag{C.6}$$

Recall from (16) that M_2 is a constant. Using BEJK Result 3 and (19), the relative change in welfare due to a trade shock can be written as

$$\frac{U_1'}{U_1} = \frac{\rho - \lambda(R)\ln(\gamma)}{\rho - \lambda(R')\ln(\gamma)} \frac{\frac{Y_{10}'}{N_1'}}{\frac{Y_{10}}{N_1}} \left(\frac{\frac{T_{10}}{T_{20}} + (\tau \frac{w_{20}}{w_{10}})^{-\theta}}{\frac{T_{10}}{T_{20}} \frac{w_{10}'^{-\theta}}{w_{10}^{-\theta}} + (\tau' \frac{w_{20}'}{w_{10}})^{-\theta}} \right)^{-\frac{1}{\theta}}.$$

Using (C.6) and choosing period 0's labor in country 1 as numeraire, i.e., $w_{10} = w_{10}' = 1$, the above becomes

$$\begin{aligned} \frac{U_1'}{U_1} &= \frac{\rho - \lambda(R)\ln(\gamma)}{\rho - \lambda(R')\ln(\gamma)} \times \frac{\left(x' + \frac{\beta}{1+\theta}\right) w'^{-1}}{\left(x + \frac{\beta}{1+\theta}\right) w^{-1}} \times \left(\frac{\frac{T_{10}}{T_{20}} + (\frac{\tau}{w})^{-\theta}}{\frac{T_{10}}{T_{20}} + (\frac{\tau'}{w'})^{-\theta}} \right)^{-\frac{1}{\theta}} \\ &\equiv GR \times IG \times ACR, \end{aligned}$$

where the last line is defined by the each multiplicative term respectively. Namely, the relative change in welfare in response to a trade shock can be decomposed into a growth-rate effect (GR), an income gains effect (IG), and a price-index effect (ACR). Note that the third term above is exactly the same as the ACR statistic and is henceforth denoted. To see this, observe that with $w_1 = 1$, the domestic consumption share is

$$\pi_{11} = \frac{T_{10}}{T_{10} + T_{20}(\frac{\tau}{w})^{-\theta}},$$

and it is readily verified that the ACR formula $(\pi'_{11}/\pi_{11})^{-\frac{1}{\theta}}$ is exactly the same as the third term.

For the welfare gap between the two countries $u \equiv U_1/U_2$, first note that the growth-rate effect is the same for both the North and South, and hence it will not show up in the gap, and thus

$$u = U_1/U_2 = \frac{Y_1/N_1}{Y_2/N_2}\frac{P_2}{P_1}.$$

From (B.2) and (B.3), we can derive the per capita (nominal) income as:

$$\begin{aligned}\frac{Y_2}{N_2} &= \frac{1+\theta-\beta}{1+\theta}\frac{X_2}{N_2},\\ \frac{Y_1}{N_1} &= \chi^{-1}\left(x+\frac{\beta}{1+\theta}\right)\frac{X_2}{N_2}\end{aligned}$$

Using the above and BEJK Result 3,

$$\begin{aligned}u &\equiv \frac{U_1}{U_2} = \frac{\frac{Y_{10}}{N_1}}{\frac{Y_{20}}{N_2}}\frac{P_{20}}{P_{10}}\\ &= \chi^{-1}\frac{1+\theta}{1+\theta-\beta}\left(x+\frac{\beta}{1+\theta}\right)\left(\frac{\frac{T_{10}}{T_{20}}(w\tau)^{-\theta}+1}{\frac{T_{10}}{T_{20}}w^{-\theta}+\tau^{-\theta}}\right)^{-\frac{1}{\theta}}.\end{aligned}$$

which is a combination of income and price gaps.

D Calibration

We calibrate our model economy to match the world economy. An important task in the calibration is to reasonably define the North and South countries. Our approach is to take the Penn World Tables (PWT) 9.0 data and select those countries with at least 67% of the US real GDP per employment (person) as the North.[15] Then, we take the intersection with the countries in other data sets that are needed for our calibration purpose. Other data sets needed are World Development Indicators (WDI) and the Global Report of Global Entrepreneurship Monitor (GEM). As we proxy the number of entrepreneurs by combining the values of Early-Stage Entrepreneurial Activity (TEA)[16] in the GEM Global

[15] The exact GDP measure we use is output side real GDP at constant PPPs (in mil. 2011 US$; RGDPO). Every measure/variable we use in the calibration is the average during 2000-2014 except for our calibration of the long-run growth rate.

[16] Early-Stage Entrepreneurial Activity is the fraction of surveyed individuals (18-64 years old) who are involved in a nascent firm or new firm or both (if doing both, still counted as one active person).

Report and the WDI, for each country we take the average during 2000-2014 due to the data availability of the GEM report.

Following the R&D-based endogenous growth literature, we set the time preference rate at $\rho = 3\%$, the real interest rate at $r = 5\%$ and GPT innovation ladder size at $\gamma - 1 = 1.4\%$. Following recent trade literature, we take the Fréchet shape parameter (connected to the trade elasticity in the gravity regression) as $\theta = 4$ (Simonovska and Waugh 2014a; sensitivity analysis will be performed). The trade cost parameter is set as $\tau = 2$. This is in the ballpark of most estimated bilateral trade costs in the literature, but it is slightly higher to reflect that the trade costs between the North and South countries are higher than those among North countries. We will thus conduct sensitivity analysis where $\tau = 1.5$ and $\tau = 2.5$**.** Observing much lower R&D labor share, we choose the share of the entrepreneurial type labor to be $\psi = 85\%$ and hence the share of the research type $1 - \psi = 15\%$; again, sensitivity analysis will be conducted on this. In a unit-free environment, we normalize $N_1 = 1$ and initial values $w_{10} = 1$.

Other parameters will be calibrated based on available data moments. We use World Inequality Database (https://wid.world/) to calibrate k. Under the Pareto assumption, top p percents income earners earn a total share of $s = (100/p)^{\frac{1}{k}-1}$. To better reflect the tail, we focus on top 1 percent income earners. Using the share of income earned by top 1 percent income earners in the US and averaging across years during 2000-2014, this share equals to 0.1903. This implies $k = 1.563$, which is taken as our benchmark value of k. We will also perform sensitivity analysis on this parameter.

To compute $t_0 \equiv \frac{T_{10}}{T_{20}}$, we utilize the fact that the ratio of the mean productivity between the two countries under Frechét distributions is given by

$$\frac{E_1(z)}{E_2(z)} = \frac{T_{10}^{\frac{1}{\theta}}}{T_{20}^{\frac{1}{\theta}}} = t_0^{\frac{1}{\theta}}.$$

We proxy $E_1(z)/E_2(z)$ by the relative total factor productivity (TFP), which is obtained from the Penn World Table (PWT) 9.0.[17] We obtain $E_1(z)/E_2(z) = 1.8885$, and thus

[17] For the construction of TFPs in Penn World Table, see Feesntra, Inklaar, Timmer (2015). Specifically, for each North/South block, we proxy the TFP by the weighted average of each country's TFP with the weight being the manufacturing GDP. A special feature of the PWT data is that there are one measure of TFP for cross-country comparison (CTFP), where the TFP level of the USA is set to 1 for all years, and another by-country time-series measure (RTFPNA), where the TFP level is calculated relative to the country's 2011 level (hence TFP of each country at 2011 is set to 1). As we need to take average value during 2000-2014, we construct a panel of TFPs in the following way. We calculate a country c's TFP at year t relative to the US' level at 2011:

$$\text{TFP}_{c,t} \equiv \text{CTFP}_{c,t} \times \text{RTFPNA}_{\text{USA},t}.$$

$t_0 = 12.719$.

From WDI, the manufactured output ratio by $x = X_1/X_2 = 3.285$. Since the GPT innovation is conducted in North, the balanced growth rate is computed based on the average growth rate of North countries under the definition above: $g = 1.8977\%$. Turning now to the components of labor, we denote $n_i^R \leq N^R$ as the number (rather than effective units) researchers and n_i^M as the number of entrepreneurs in country i. It is readily verified that

$$n_1^M = N_1^M \left[1 - G\left(a_1^M\right)\right] = \left(\frac{k-1}{k}\right)^{\frac{k}{k-1}} \psi^{-\frac{1}{k-1}} N_1^{-\frac{1}{k-1}} M_{1v}^{\frac{k}{k-1}} \tag{D.7}$$

$$n_2^M = N_2 \left[1 - G\left(a_2^M\right)\right] = \left(\frac{k-1}{k}\right)^{\frac{k}{k-1}} N_2^{\frac{-1}{k-1}} M_2^{\frac{k}{k-1}} \tag{D.8}$$

$$n^R = N_1^R \left[1 - G\left(a_1^R\right)\right] = \left(\frac{k-1}{k}\right)^{\frac{k}{k-1}} (1-\psi)^{-\frac{1}{k-1}} N_1^{-\frac{1}{k-1}} R^{\frac{k}{k-1}}. \tag{D.9}$$

We have reservation about the quality of data of entrepreneur counts in the South (due to large fractions of self-employed who tend to report yes for TEA in the survey) and hence will use only data from North countries. Specifically, we use weighted average TEA (average across countries with the weight being total employment) in the GEM global report to proxy n_1^M/N_1. We further use "Researchers in R&D (per million people)" in WDI to calculate n^R/N_1. These shares are: $n_1^M/N_1 = 0.0730$ and $n^R/N_1 = 0.00357$. Thus, the researcher-entrepreneur ratio turn out to be approximately 1:20 ($\frac{n^R/N_1}{n_1^M/N_1} = \frac{0.00357}{0.0730} = 0.0489$). Combining (14), (16), (D.7), (D.8), and (D.9), one can show that $\frac{n_2^M}{N_2} = \frac{n_1^M/N_1}{1-n^R/N_1}$. Thus, we compute: $\frac{n_2^M}{N_2} = \frac{n_1^M/N_1}{1-n^R/N_1} = 0.07326$. Furthermore, combining (16) and (D.8), we have $\frac{n_2^M}{N_2} = J_1^{-1}\left(\frac{k-1}{k}\right)^{\frac{k}{k-1}}$; with the definition of J_1 given in Section 3.1, β can therefore be solved out:

$$\beta = 1 - \frac{\theta k}{k-1} \frac{n_2^M/N_2}{1 - n_2^M/N_2} = 0.1222,$$

which implies $J_1 = 0.80173$.

We are left to compute four endogenous variables $\{R, M_1, M_2, w\}$ and calibrate the preference parameter ϵ, the technology parameter κ, and relative endowment parameter χ. The seven equations needed are (12)-(17) in the solution algorithm (Section 3.1) plus the law of motion $g = \kappa R^\epsilon \ln(\gamma)$.

Using (14), (16), and (D.9), we can conveniently write $m = M_1/M_2$ as

$$m = \chi \psi^{\frac{1}{k}} \left(1 - \frac{n^R}{N_1}\right)^{\frac{k-1}{k}}. \tag{D.10}$$

From (12), we can write

$$\chi = x^{-(k-1)}\psi^{-1}w^{k-1}m^k, \tag{D.11}$$

and thus (13) and (D.10) can be written as

$$\begin{aligned} w^\theta &= x^{-1}m\left[\frac{mt_0w^{-\theta}+\tau^{-\theta}}{mt_0(w\tau)^{-\theta}+1}\frac{1+\theta-\beta}{1+\theta}+\frac{\beta\tau^\theta}{1+\theta}\right] \\ m &= xw^{-1}\psi^{\frac{1}{k}}\left(1-\frac{n^R}{N_1}\right)^{\frac{-1}{k}}. \end{aligned}$$

Solving jointly the above two equations yields $w = 1.3766$ and $m = 2.1555$. Accordingly, we can compute $\chi = 2.3948$ using (D.11), and $M_2 = J_1^{\frac{1-k}{k}}\chi^{-1}N_1 = 0.4522$, and $M_1 = mJ_1^{\frac{1-k}{k}}\chi^{-1}N_1 = 0.97466$ from (14) and (16).

Using (D.7) and (D.9), we have

$$\frac{R}{M_1} = \left(\frac{1-\psi}{\psi}\right)^{\frac{1}{k}}\left(\frac{n^R/N_1}{n_1^M/N_1}\right)^{\frac{k-1}{k}}, \tag{D.12}$$

and this allows us to compute $R = 0.10834$. Finally, with $\lambda = g/\ln(\gamma) = 1.3649$, we use (17) to calibrate

$$\epsilon = \frac{1-\beta}{\beta}\frac{J_1\chi(r+\lambda)w}{\lambda(1+g)(x+1)}(1-\psi)^{-\frac{1}{k-1}}\left(\frac{R}{N_1}\right)^{\frac{k}{k-1}} = 0.2741,$$

and we use the growth law of motion to obtain

$$\kappa = \frac{g}{R^\epsilon\ln(\gamma)} = 2.510,$$

and this completes the calibration procedure.

E The Model with Homogenous Labor

The model before Section 2.4 is unchanged. So, we start with the labor market. Suppose in contrast to benchmark model, labor is homogenous so that we have

$$R_1 + M_1 + L_1 = N_1, \qquad M_2 + L_2 = N_2.$$

By BEJK Result 4 and the fact that part of the revenue is paid to the GPT firm, we have

$$v_{i\nu} = \frac{1-\beta}{1+\theta}\frac{X_{i\nu}}{M_{iv}}, \qquad w_{i\nu} = \frac{\theta}{1+\theta}\frac{X_{i\nu}}{L_{iv}}.$$

In equilibrium $w_{i\nu} = v_{i\nu}$, and this implies that $M_{iv} = \frac{1-\beta}{\theta}L_{iv}$. In the South, this implies that

$$L_{2\nu} = \frac{\theta}{1-\beta+\theta}N_2, \qquad M_{2\nu} = \frac{1-\beta}{1-\beta+\theta}N_2. \tag{E.13}$$

In the North, we have $N_1 - R_\nu = M_{1v} + L_{1\nu} = \frac{1-\beta+\theta}{\theta}L_{1\nu}$, or alternatively,

$$L_{1\nu} = \frac{\theta}{1-\beta+\theta}(N_1 - R_\nu) \tag{E.14}$$

$$M_{1\nu} = \frac{1-\beta}{1-\beta+\theta}(N_1 - R_\nu) \tag{E.15}$$

The value of an innovation is still the same as $V_{\nu+1} = \frac{\beta(X_{1\nu+1}+X_{2\nu+1})}{(1+\theta)(r+\lambda(R_\nu))}$, but the problem facing a GPT innovator becomes

$$\max_{R_\nu} \lambda(R_\nu)V_{\nu+1} - w_{1\nu}R_\nu,$$

where $w_{1\nu}$ replaces w_ν^R in the benchmark model. Whenever the innovation arrives, the innovator becomes the GPT monopoly. The first-order condition is therefore $\lambda'(R_\nu)V_{\nu+1} = w_{1\nu}$, or

$$\frac{r+\kappa R^\epsilon}{\kappa\epsilon R^{\epsilon-1}} = \frac{\beta}{1+\theta}\frac{X_{1\nu+1}+X_{2\nu+1}}{w_{1\nu}}. \tag{E.16}$$

Again, as explained in Section 3.2, the global sales of differentiated products denominated in the North's wages, $(X_{1\nu+1}+X_{2\nu+1})/w_{1\nu}$, is still the key to determine equilibrium R.

Using (E.13) and (E.14), the labor market clearing conditions are

$$\frac{\theta}{1+\theta}X_{1\nu} = w_{1\nu}L_1 = \frac{\theta}{1-\beta+\theta}(N_1 - R_\nu)w_{1\nu}, \tag{E.17}$$

$$\frac{\theta}{1+\theta}X_{2\nu} = w_{2\nu}L_2 = \frac{\theta}{1-\beta+\theta}N_2 w_{2\nu}. \tag{E.18}$$

Taking the ratio of the above two equations entails

$$x = \chi\left(1 - \frac{R_\nu}{N_1}\right)w. \tag{E.19}$$

Along the balanced growth path and using (E.17) and (E.18), we have

$$\frac{X_{1\nu+1}+X_{2\nu+1}}{w_{1\nu}}=\frac{1+\theta}{1-\beta+\theta}\left(N_1-R+\frac{N_2}{w_\nu}\right)(1+g)\,.$$

Combining the above with (E.16), we obtain

$$\frac{r+\kappa R^{\epsilon}}{\kappa\epsilon R^{\epsilon-1}\left[1+\kappa R^{\epsilon}\ln\left(\gamma\right)\right]}=\frac{\beta}{1-\beta+\theta}\left(N_1-R+\frac{N_2}{w}\right). \tag{E.20}$$

The equations in Section 2.5 remain the same in this model. Combining (E.13) and (E.15) entails

$$m=\chi\left(1-\frac{R}{N_1}\right), \tag{E.21}$$

and from (E.19), we know that $x=mw$, which, when combined with (10), entails

$$w^{1+\theta}=\frac{mt_0w^{-\theta}+\tau^{-\theta}}{mt_0\left(w\tau\right)^{-\theta}+1}\frac{1+\theta-\beta}{1+\theta}+\frac{\beta\tau^{\theta}}{1+\theta}. \tag{E.22}$$

So, a dynamic equilibrium on a balanced growth path can be obtained by solving $\{R,w,m\}$ jointly using (E.20), (E.21), and (E.22).