Lee Kuan Yew School of Public Policy

Asia Competitiveness Institute

ACI Research Paper #02-2022

Trade Policy in a Ricardian World with International Knowledge Diffusion

Chi ZHANG

April 2022

Please cite this article as:

Zhang, Chi, "Trade Policy in a Ricardian World with International Knowledge Diffusion", Research Paper #02-2022, *Asia Competitiveness Institute Research Paper Series (April 2022).*

Trade Policy in a Ricardian World with International Knowledge Diffusion[*]

Zhang Chi[†]

April 5, 2022

[*]Zhang Chi is grateful to Lin Ma and Davin Chor for their continued support and encouragement. I wish to thank Deng Liuchun, Li Bingjing, Liu Chen, Albert Hu, Mario Larch, along with seminar participants at the National University of Singapore and various conferences for very helpful comments. All remaining errors are mine.

[†]Asia Competitiveness Institute, National University of Singapore, chizhang@nus.edu.sg.
Zhang Chi acknowledges the financial support from NUS Economics department research scholarship during her Ph.D candidature.

Abstract

This paper studies the optimal trade policy in the presence of international knowledge diffusion. Trade policies can affect welfare through the pattern of comparative advantages, terms of trade, and the growth of technology driven by knowledge diffusion. Compared with a static Ricardian model, knowledge diffusion incentivizes countries to impose higher tariffs on relatively under-developed trading partners, such that Home country will not only trade more with but also have a higher probability of learning from advanced economies, which leads to faster economic growth and higher welfare. Quantitative simulations show that optimal tariff schedule is in favor of the relatively more developed trading partners in the short run. The ability to impose discriminatory tariffs is quantitatively important for developing countries, while developed countries can achieve similar welfare even under the restriction of the GATT/WTO "non-discrimination principle". As countries grow more alike, the incentive to intervene diffusion process diminishes. In the absence of retaliation and coalition, the long-run optimal tariff is positive due to the terms-of-trade argument.

Keywords— Trade Policy, Tariff, Knowledge Diffusion, GATT/WTO, non-discrimination principle.

1 Introduction

Despite the prevailing stand in trade theories that free trade is beneficial for economic growth, trade protectionism has a long history in reality. Even with the series of trade liberalization measures brought by the General Agreement on Tariffs and Trade (GATT) and its successor, the World Trade Organization (WTO), effective applied tariffs in most bilateral relationships are far from zero. This paper provides a justification for the existence of trade policies — governments have incentives to intervene in international trade since it can influence economic growth through international knowledge diffusion. Technology diffusion is an important channel of productivity growth, especially for countries in the catch-up phase. Trade of intermediate goods serves as an important channel of such international knowledge diffusion (Keller, 2002; Keller and Yeaple, 2009; Sjöholm, 1996). Rapid economic growth in South Korea, China, and India during the past few decades underscores the crucial role of trade openness in driving technology and economic growth. Hence, it is important to understand the linkage between international trade and knowledge diffusion and perhaps, more importantly, the implications on the corresponding implications on policy arrangements.

This paper also sheds light on whether laissez-faire policy is optimal. To maximize production efficiency by preserving the natural pattern of comparative advantage, GATT/WTO requires its members not to discriminate between importing origins. Nevertheless, de facto discrimination is rampant in reality. Regional trade agreements like The European Union and The North American Free Trade Agreement are examples of exceptions to the "Most-Favored Nation" treatment. I show that, although the "non-discrimination" principle is optimal in a static perspective, when dynamic knowledge diffusion is taken into consideration, countries will set trade policy discriminatively to reorient trade towards technologically more advanced trading partners for faster economic growth.

I build on a heterogeneous-firm Ricardian model (Eaton and Kortum, 2002) with international knowledge diffusion (Buera and Oberfield, 2020). While goods are traded and producers interact, technology is transmitted. International knowledge diffusion is linked to international trade in that the probability of learning from an origin is positively correlated with trade volume. Thus, a country's trade pattern can alter its law of motion of technology and trade policies can influence economic growth through adjusting trade barriers. In order to enjoy a higher growth rate by engaging more with advanced economies, countries tend to discriminate against relatively under-developed trading partners when setting trade policies. This implies that the non-discrimination principle advocated by WTO is suboptimal, as it requires countries to impose a uniform tariff on all trading partners.

Curse of dimensionality is a problem of this model in the quantitative exercise since the number of dimensions grows exponentially as the number of countries or the number of time periods expands. To maintain the feasibility of quantification analysis, I simulate a world with three countries for 100 periods and focusing on the trade policy imposed by only one country - Home country. Home country has two trading partners. In order to illustrate the discriminatory nature of optimal tariff due to the consideration of knowledge diffusion, I assume that the two trading partners are homogeneous in all dimensions except for initial technology levels. The technology levels in the two trading partners are calibrated to be the 10th and 90th percentile of world technology level in 2010.[1] To study the optimal trade policies of countries in different development levels, I consider three cases of Home country — 5th percentile if Home is under-developed, 50th percentile if middle-income, and 95th percentile if most-developed.

This paper shows that optimal tariff is discriminatory based on trading partners' development levels. The discriminatory nature of optimal tariff is a result of the dynamic interaction between international trade and knowledge diffusion. In comparison with non-discriminatory tariff schedules, the ability to tax imports based on their origins enables countries to obtain a higher living standard. With free trade being the baseline, an under-developed country can promote welfare by 2.14% if discrimination is allowed, compared with 1.98% achievable by a non-discriminatory tariff schedule. Welfare gains from trade policy increase by 7.87% in this case. However, welfare gains from trade policy is only 0.08% for a developed Home country, indicating that the WTO Most-Favored-Nation (MFN) principle is plausible for developed countries yet costly

[1] Fundamental technology levels are computed using the method proposed by in Costinot et al. (2012).

for under-developed countries.

In the long run, countries' technology levels converge in the absence of external shocks. Thus the incentive to influence diffusion weakens. As a result, trade policies converge to balanced growth path policies, which is equivalent to the trade policy preferred by a myopic social planner under the parameters assumed in this paper's benchmark simulations. The optimal trade policy in the long term is non-discriminatory. Moreover, without the concern of retaliation or coalition, countries will charge positive tariffs on imports even in the long run. By doing so, countries can lift their own terms of trade and shift the distribution of welfare gains from trade across countries in favor of themselves.

Literature Review

This paper is related to the literature on optimal trade policy, which usually relies on the terms-of-trade argument to justify the positive import tariffs of large economies (Torrens, 1833; Bickerdike, 1907; Broda et al., 2008; Bagwell and Staiger, 1999, 2001, 2016). Our model echos this stream of literature in that a myopic social planner who does not consider knowledge diffusion would set positive tariffs even if it causes efficiency losses. Costinot et al. (2014) also studies terms of trade in a dynamic model, but it focuses on policies on capital controls.

Another line of literature studies trade policies in the presence of market failure (Gros, 1987; Demidova and Rodríguez-Clare, 2009; Felbermayr et al., 2013; Ossa, 2011, 2014; Caliendo et al., 2021). The key difference between this paper and the literature is that market failure in these models mainly arises from monopoly power, and as a result, tariff generates efficiency gains. In this paper, global production allocation is most efficient under free trade, yet trade policy can serve as an instrument to internalize the externality brought by the dynamic interaction between international trade and diffusion. It is one of the contributions of this paper that, in contrast to the uniform tariff schedules proposed by the literature, the optimal trade policy in this paper is discriminatory in order to correct for trade intensity between different trading partners.

This paper is also one of the few trade policy studies that include more than two countries in the model. Bagwell et al. (2020) and Bagwell et al. (2021) also emphasize the importance of including more than two countries and embracing the possibility of tariff discrimination. However, the results of this paper are contrary to theirs because we focus on different types of externality. In their papers, "non-discrimination" principle helps countries to escape from the term-of-trade driven Prisoners' dilemma, whereas in my paper, "non-discrimination" principle restrict countries to learn from the best.

The model in this paper follows Eaton and Kortum (2002) and Buera and Oberfield (2020) closely. Firms are heterogeneous and face Ricardian selection from international competition. The probability of diffusion is related to trade intensity, so trade policies influence economic growth through exerting (alleviating) trade barriers. In comparison with Buera and Oberfield (2020), the focus of this paper is to explore the externality between trade and knowledge diffusion and to study the implied optimal trade policy. Sampson (2016) and Perla et al. (2021) also study knowledge diffusion in an open economy, but they build on Melitz model and assume that knowledge diffusion can happen only domestically. The implied trade policy in such models are interesting for future studies.

Firm heterogeneity is an important element of the model, since it allows countries to learn from each other. In contrast, the North-South product cycle model only allow developed country to innovate and developing countries to imitate (Helpman, 1993; Borota, 2012; Sampson, 2021).

The rest of the paper is organized as follows. Section 2 sets up the model. The balanced growth path of this economy is described in Section 3. Then using the parameters assumed in Section 4, Section 5 lays out the main results of this paper. Sensitivity checks and extensions are discussed in Section 6. Section 7 concludes.

2 Model

This section describes the structural model of this paper. Instantaneous equilibrium is a heterogeneous-firm Ricardian framework. The dynamic component is modeled using the flow of ideas, which leads to technology development and economic growth.

Consider a world with N countries. In each country, a benevolent social planner chooses trade policy to maximize the total lifetime utility of people living in this country. Suppose there is one industry in the economy, whose final product is produced by assembling a continuum of intermediate goods. Countries can trade intermediate goods subject to iceberg trade cost and tariff, whereas the final product is not tradable. Intermediate good markets are perfectly competitive. Time is infinite and continuous.

2.1 Individual Consumption

Agents in a country are homogeneous. Each agent is endowed with one unit of labor, which is inelastically supplied to the production of intermediate goods. Labor can move freely within a country, and receive nationwide uniform wage $w_n(t)$. Given the market price of final product and wage, individuals in country n maximize utility according to the following problem:

$$\begin{aligned} \max_{C_n(t)} \quad & U_n = \int_{t=0}^{\infty} e^{-\rho t} \log(C_n(t))dt \\ s.t. \quad & P_n(t)C_n(t) = w_n(t) + \frac{R_n(t) + D_n(t)}{L_n(t)}, \end{aligned} \tag{1}$$

where $P_n(t)$ and $C_n(t)$ respectively represent the price and individual demand of final good at time t. The size of the population at time t is indicated by $L_n(t)$. Tariff revenue (subsidy spending) $R_n(t)$ and trade deficit (surplus) $D_n(t)$ are rebated to (taxed on) the consumers as lump-sum transfer. Contemporaneous utility is discounted by factor ρ.

2.2 Production

Given the technology levels at a specific point of time t, the production structure is static. Time subscript t is suppressed where there is no confusion.

2.2.1 Final Good

The final product is produced using a continuum of intermediate goods indexed by $\omega \in [0, 1]$ under the constant-elasticity-of-substitution (CES) production function:

$$Q_n = \left(\int_0^1 s_n(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}}, \tag{2}$$

where Q_n is the total output of the final good in country n and $s_n(\omega)$ is the demand of intermediate good ω in country n. In this model, the non-tradable final good can not be saved or stored. Therefore, total consumption in the society $C_n(t)L_n(t)$ is equal to $Q_n(t)$ in equilibrium.

2.2.2 Intermediate Goods

Intermediate goods are produced under a linear production function, using labor as the sole input. Domestic supply of variety $\omega \in [0,1]$ is given by the following equation:

$$q_n(\omega) = z_n(\omega)l_n(\omega). \tag{3}$$

Hicks-neutral productivity $z_n(\omega)$ is variety-specific; labor allocation is represented by $l_n(\omega)$. Marginal cost of producing variety ω in country n is $\frac{w_n}{z_n(\omega)}$. Intermediate good markets are assumed to be perfectly competitive.

2.2.3 Productivity Draws

As in a standard Eaton-Kortum model (Eaton and Kortum, 2002; Caliendo and Parro, 2015), productivity of each variety ω in country n is independently drawn from a Frechet distribution with the following cumulative distribution function (CDF):

$$\Pr\{z_n(\omega) < z\} \equiv F_n(z) = \exp\left[-\lambda_n z^{-\theta}\right]. \tag{4}$$

By the property of Frechet distribution, λ_n governs the unconditional mean of $z_n(\omega)$ and represents the level of fundamental technology level in a country. The key difference between this model and the traditional Eaton-Kortum model is that fundamental technology $\lambda_n(t)$ evolves endogenously in this paper. The law of motion of technology is illustrated in section 2.5. θ is associated with the variance of productivity draws, and greater value of θ implies weaker comparative advantage forces.

2.3 International Trade

Denote by $d_{ni}(t)$ the trade cost between an importing country n and an exporting country i, and $\tau_{ni}(t)$ the trade policy. The composite trade barrier is given by $\kappa_{ni}(t) =$

$d_{ni}(t) \times (1 + \tau_{ni}(t))$ and producers of the final good choose the lowest-price provider for each variety ω. That is, the price of variety ω is given by

$$p_n(\omega, t) = \min_i \left\{ \frac{w_i(t)\kappa_{ni}(t)}{z_i(\omega, t)} \right\}. \tag{5}$$

Assume there is no trade cost or tariff for domestic trade, $d_{nn} = 1$ and $\tau_{nn} = 0$. Further assume the triangle inequality holds, $\kappa_{ni'}(t)\kappa_{i'i}(t) \geq \kappa_{ni}(t)\ \forall i'$, such that entrepots like Singapore and Hong Kong are ruled out.

2.4 Market Clearing and Instantaneous Equilibrium

Denote $\pi_{ni}(t)$ as the fraction of country n's income spent on intermediate goods from country i and $P_n(t)$ as the price index in country n. By the CES production function of final good, price index is an aggregate of intermediate good prices, $P_n(t) = \left[\int_0^1 p_n(\omega, t)^{1-\sigma} d\omega\right]^{\frac{1}{1-\sigma}}$. By the law of large numbers and the probabilistic structure of productivity draws, bilateral trade share and price index are functions of multilateral resistance:[2]

$$\pi_{ni}(t) = \frac{\lambda_i(t)\left(w_i(t)\kappa_{ni}(t)\right)^{-\theta}}{\sum_h \lambda_h(t)\left(w_h(t)\kappa_{nh}(t)\right)^{-\theta}}, \tag{6}$$

$$P_n(t) = A\left[\sum_i \lambda_i(t)\left(w_i(t)\kappa_{ni}(t)\right)^{-\theta}\right]^{-\frac{1}{\theta}}. \tag{7}$$

Let $I_n(t) \equiv P_n(t)C_n(t)$ denote the total expenditure of individuals in country n. National budget constraint is

$$I_n(t) = w_n(t)L_n(t) + R_n(t) + D_n(t), \tag{8}$$

where $R_n(t) = \sum_{i=1}^N \tau_{ni}(t)X_n(t)\frac{\pi_{ni}(t)}{1+\tau_{ni}(t)}$ is the tariff revenue (import subsidy) collected (spent) by the government and $D_n(t)$ is the trade deficit financed through international borrowing and lending. Trade deficit (surplus) is assumed to be exogenously determined in this model. In each period t, global financial market is cleared $\sum_n D_n(t) = 0\,\forall t$.

[2] $A = \Gamma\left(1 + \frac{1-\sigma}{\theta}\right)^{\frac{1}{1-\sigma}}$ is a constant, where $\Gamma(\cdot)$ is the Gamma function defined as $\Gamma(z) = \int_0^\infty x^{z-1}e^{-x}dx$ for $z > 0$.

Labor is allocated to the production of intermediate goods. Total labor income in country n equals the net income of domestic intermediate good producers:[3]

$$w_n(t)L_n(t) = \sum_{i=1}^{N} X_i(t)\frac{\pi_{in}(t)}{1+\tau_{in}(t)}. \tag{9}$$

2.5 International Knowledge Diffusion

Idea flow model in this paper follows Buera and Oberfield (2020) closely. Technology advancement is a result of international knowledge diffusion. In the aggregate, trade pattern not only determines the intensity of diffusion, but also moderates the quality of ideas through Ricardian selection.

From a micro perspective, all (potential) producers in country n's intermediate good markets can interact with one of their peers drawn from a mass of firms with productivity distribution $G_{nt}(z')$. At each instant t, a producer with productivity z in country n can produce an insight with probability $\eta_n(t)$. An insight is a new technology featured with productivity bz'^{β}, which is composed of a diffusion shock b randomly drawn from a Pareto distribution and the productivity level of the counterpart z'. β stands for the strength of diffusion and there is incomplete pass-through of the counterpart's productivity represented by the assumption $\beta \in [0, 1)$. Upon arrival, the insight is adopted by the producer if $bz'^{\beta} > z$.

The Best-practice producer within a country is the producer with the highest productivity draw for a given variety. Frontier technology is characterized by distribution $F_{nt}(z)$ — the probability that the best-practice producer of an intermediate variety has productivity less than z. By the law of large numbers, $F_{nt}(z)$ also represents the fraction of goods for which domestic best-practice producer's productivity is below z. Frontier technology evolves according to the following formula:

$$1-F_{nt+\Delta}(z) = \underbrace{[1-F_{nt}(z)]}_{\text{scenario 1}} + \underbrace{F_{nt}(z)\left[\int_t^{t+\Delta}\int_0^{\infty}\eta_n(\mu)\left(1-H\left(\frac{z}{z'^{\beta}}\right)\right)dG_{n\mu}(z')d\mu\right]}_{\text{scenario 2}}. \tag{10}$$

[3] Equation (9) is equivalent to the labor market clearing condition.

The best-practice producer of an intermediate good in country n has productivity greater than z at time $t+\Delta$ is either due to that it had a productivity higher than z at time t (scenario 1 in equation 10) or because it received an insight with quality better than z during the Δ time frame (scenario 2 in equation 10).

Rearranging equation (10) and taking the limit as $\Delta \to 0$, we obtain the following law of motion for $F_{nt}(z)$:

$$\begin{aligned} \frac{d \ln \left(F_{nt}(z)\right)}{dt} &= \lim_{\Delta \to 0} \frac{F_{nt+\Delta}(z) - F_{nt}(z)}{\Delta F_{nt}(z)} d\tau \\ &= -\int_0^{\infty} \eta_n(t) \left(1 - H\left(\frac{z}{z'^{\beta}}\right)\right) dG_{nt}(z'). \end{aligned} \tag{11}$$

The fundamental technology level $\lambda_n(t)$ in country n corresponds to $\int_0^t \eta_n(\mu) \int_0^{\infty} x^{\beta\theta} dG_{n\mu}(x) d\mu$ if initial frontier technology follows a Frechet distribution (Buera and Oberfield, 2020). To see this, note that the randomness factor b follows CDF $H(b) = 1 - b^{-\theta}$, thus Equation (11) can be written as $\frac{d\ln(F_t(z))}{dt} = -\dot{\lambda}_n(t) z^{-\theta}$. Since initial frontier technology follows a Frechet distribution, $F_{n0}(z) = e^{-\lambda_n(0) z^{-\theta}}$, the distribution of productivity draws $F_{nt}(z)$ remain Frechet for all periods t and $\lambda_n(t)$ represent the shape parameter of frontier technology distribution at time t in country n.

Further assume producers in country n learn from international providers that sell to country n with equal probability, $G_{nt}(z') = \sum_i \int_{\{\omega \in \Omega_{nit} | z_i(\omega,t) < z'\}} \mathbf{1} d\omega$, where $\Omega_{nit} \equiv \{\omega \in [0,1] | p_{ni}(\omega,t) = \min_{i'} p_{ni'}(\omega,t)\}$ denotes the set of intermediate goods observed to be imported from country i by n. The source distribution of ideas implies that the probability of a producer in country n interacting with a producer in country i is related to the probability of trade.

Under the aforementioned assumptions, the shape parameter λ evolves according to:

$$\dot{\lambda}_n(t) = \eta_n(t) \Gamma(1-\beta) \sum_i \pi_{ni}(t) \left(\frac{\lambda_i(t)}{\pi_{ni}(t)}\right)^{\beta}. \tag{12}$$

Producers can draw insights domestically and internationally. Import shares $\pi_{ni}(t)$ serve as the weights of different origins in the diffusion process. That is to say, the more important trading partner i is as an import origin to country n, the more likely home country n is going to engage with and learn from country i's producers. Meanwhile, the quality of ideas is determined by $\frac{\lambda_i(t)}{\pi_{ni}(t)}$.[4] This implies that trade functions as a selection mechanism when producers search for new insights. Only the relatively more productive and comparatively advantaged firms in i can export and spread out knowledge, and smaller trade share implies stricter selection criteria. Quality of ideas is discounted by $1-\beta$ proportion in log-terms during the diffusion process.

Innovation is not modeled endogenously in this model, but we can reckon the path of insight arrival rate $\eta_n(t)$ as the technology growth resulted from innovation, which exogenously depends on country level R&D intensity. In the extreme case where $\beta = 0$, diffusion is shut down and technology development is entirely determined by insight arrival rates (Zhang, 2021).

2.6 Social Welfare and Optimal Trade Policy

Given the response of individuals and firms, the government in country n chooses tariff schedule $\{\tau_{ni}(t)\}_{i\neq n} \in \mathcal{R}^{N-1}$ with respect to all trading partners for each period t in order to maximize the total domestic social welfare, suggesting the maximization problem:[5]

$$
\begin{aligned}
\max_{\{\tau_i(t)\}_{i\neq n}} \quad & L_n \int_{t=0}^{\infty} e^{-\rho t} \log(C_n(t))dt \\
s.t. \quad & \text{consumers maximize individual utility (equation (1)),} \\
& \text{firms maximize profit (equation (2) to (7)),} \\
& \text{markets are cleared (equation (8) to (9)),} \\
& \text{technology evolves according to knowledge diffusion (equation (12)).}
\end{aligned}
$$

Given tariff schedules in other countries, the problem can be re-written in a recursive form:

$$
V\left(\lambda_n(t), \{\lambda_i(t)\}_{i\neq n}\right) = \max_{\{\tau_{ni}\}_{i\neq n}} \left\{\log(C_n) + \rho V\left(\lambda_n(t+\Delta), \{\lambda_i(t+\Delta)\}_{i\neq n}\right)\right\}.
$$

[4]The quality of ideas can be rewritten as $\frac{\lambda_i(t)}{\pi_{ni}(t)} = \left(\frac{Aw_i(t)\kappa_{ni}(t)}{P_n(t)}\right)^{\theta}$. Standing on the perspective of a particular country n, $w_i(t)\kappa_{ni}(t)$ is a sufficient statistic to evaluate the quality of ideas from different origins.

[5]Negative tariff is essentially an import subsidy. I abuse the use of tariff in this paper. In general, τ can be positive or negative.

(13)

When choosing the optimal tariff schedule, Home country government jointly considers its effect on contemporary consumption, domestic technology growth and foreign technology growth.[6]

When countries are on different growth trajectories, discrimination in trade policy can benefit a country's growth prospects through equation (12). In particular, imposing higher tariffs on under-developed countries would divert trade towards more advanced trading partners, leading to a rise in the quality of ideas and better growth potential.

3 Balanced Growth Path

Optimal trade policies on balanced growth paths (BGP) in this model vary with the forms of the BGP. In this section, I show different forms of balanced growth path based on the assumptions on insight arrival rate and initial technology levels, as well as the corresponding trade policies. Optimal trade policies on BGP can shed light on long-run policies in the quantification simulations.

For starters, let us consider the balanced growth paths without government intervention, where the growth rates of fundamental technology are exogenously determined by the growth rates of insight arrival rates, according to

$$\frac{\dot{\lambda}_n(t)}{\lambda_n(t)} = \Gamma(1-\beta)\frac{\eta_n(t)}{\lambda_n(t)^{1-\beta}}\sum_i \pi_{ni}(t)^{1-\beta}\left(\frac{\lambda_i(t)}{\lambda_n(t)}\right)^{\beta}.$$

With different parameter assumptions, the form of balanced growth path can be categorized into three types, as described in proposition 1, 2, and 3.

Proposition 1. *If insight arrival rates are all constant over time, $\dot{\eta_n} = 0\,\forall n$, then the BGP of this model only exists in the limit where all fundamental technology levels approach infinity $\lambda_n \to \infty\,\forall n$ and the BGP growth rate is zero for all countries.*

Proof. Suppose some countries have a growth rate strictly greater than zero. Then for the fastest-growing country, RHS always converges to zero, since $\frac{\eta_n \lambda_i(t)^\beta}{\lambda_n(t)}$ goes to zero

[6] I do not consider maximization of global welfare in this paper because it's ambiguous whether it is incentive compatible.

for any i as time approaches infinity. This contradicts the assumption that growth rate of λ_n stays at a positive value on BGP. Additionally, RHS approaches zeros if and only if $\lambda_n \to \infty$. □

In a symmetric world with constant insight arrival rates, countries converge to the same level of development and thus there is no incentive to intervene in the diffusion process on BGP. As a result, optimal trade policy in this case is identical to the one preferred by a myopic social planner. The benchmark parameterization leads to a BGP of this kind.

Proposition 2. *If insight arrival rates evolve at the same non-zero growth rate, $\dot{\eta}_n = g_\eta \,\forall n$, and the BGP is well-defined, then on the BGP, all countries grow at the same rate $g_\lambda = \frac{1}{1-\beta} g_\eta$. Relative technology levels $\frac{\lambda_i}{\lambda_n}$ and trade patterns are constant on BGP.*

Proof. Prove by contradiction. Suppose g_η's are equal across countries and technology in some countries grow faster than others. Denote country n to be the country with the highest g_{λ_n}. Then in the limit where $t \to \infty$, all countries will buy from country n and $\pi_{in} = 1$ for all i. This implies that all countries will only draw insights from country n, such that $g_{\lambda_n} = \frac{1}{1-\beta} g_\eta$ and

$$g_{\lambda_i} = \Gamma(1-\beta) \frac{\eta_i(t)}{\lambda_i(t)^{1-\beta}} \left(\frac{\lambda_n(t)}{\lambda_i(t)} \right)^\beta .$$

Since $\lim_{t\to\infty} \frac{\lambda_n(t)}{\lambda_i(t)} = \infty$ and $g_{\lambda_i} < g_{\lambda_n} = \frac{1}{1-\beta} g_\eta$, $\lim_{t\to\infty} \frac{\eta_i(t)}{\lambda_i(t)^{1-\beta}} = 0$, implying $g_{\lambda_i} > \frac{1}{1-\beta} g_\eta$. Hence forms a contradiction. □

Lemma 1. *If countries have the same growth rate of $\eta_n(t)$, then BGP is well-defined if $g_\eta = (1-\beta)\Gamma(1-\beta) \frac{\eta_n(0)}{\lambda_n(0)^{1-\beta}} \sum_i \pi_{ni}^{1-\beta} a_{in}^\beta$, where $a_{in} \equiv \frac{\lambda_i(0)}{\lambda_n(0)}$ and $\{a_{in}\}_{i\neq n}$ satisfy the system of equations*

$$\Gamma(1-\beta) \frac{\eta_n(0)}{\lambda_n(0)^{1-\beta}} \sum_i \pi_{ni}^{1-\beta} a_{in}^\beta = \Gamma(1-\beta) \frac{\eta_{n'}(0)}{\lambda_{n'}(0)^{1-\beta}} \sum_i \pi_{n'i}^{1-\beta} \left(a_{in'} \frac{1}{a_{n'n}} \right)^\beta \quad (14)$$

for any $n' \neq n$, where $\pi_{ni} = \frac{1}{\sum_h \frac{a_{hn}}{a_{in}} \left(\frac{w_h(0)\kappa_{nh}}{w_i(0)\kappa_{ni}}\right)^{-\theta}}$ and $w_i(0)$'s are functions of $\{a_{hn}\}_{h\neq n}$, $\{\eta_h(0)\}_{h=1}^N$, $\{\lambda_h(0)\}_{h=1}^N$, $\{\kappa_{nh}\}_{h=1}^N$, $\{L_h\}_{h=1}^N$, $\{D_h\}_{h=1}^N$ and parameters that affect general equilibrium outcomes.

If both initial insight arrival rates $\eta_n(0)$ and growth rate of $\eta_n(t)$ are homogeneous, countries have the same technology levels on BGP. Hence, the myopic optimal tariff schedule is also the optimal trade policy on BGP.

If initial insight arrival rates $\eta_n(0)$ are differentiated while growth rate of $\eta_n(t)$ are equal, countries will face different technology levels on BGP; intervention in knowledge diffusion through trade policies is always effective; optimal tariff schedule is discriminatory even on the BGP.

Proposition 3. *If insight arrival rates grow at different speeds, BGP only exists in the limit where $t \to \infty$. On BGP, all productions are carried out by the fastest-growing country and trade share matrix is dominated by one column. Country n with the highest growth rate of η will grow at $\frac{1}{1-\beta} g_{\eta_n}$, while the rest will grow at $g_{\lambda_i} = g_{\eta_i} + \frac{\beta}{1-\beta} g_{\eta_n} \, \forall i \neq n$.*

Proof. This result directly follows from the above reasoning. □

In the limit where BGP holds, trade matrix is not affected by trade policies in this case. Thus, optimal trade policy is also identical to the one advocated by a myopic social planner.

Analysis of balanced growth path suggests that for the majority of parameter values, BGP only exists in the limit. Thus, I will focus on the transitional dynamics of this model in the following quantification simulations. The simulated optimal trade policy will eventually converge to the one on BGP.

4 Methodology

4.1 Simulation Strategy

Under the previously stated setup with the number of countries being N, the model embeds N^2 number of policy functions. Moreover, optimal tariff schedules are in general time-varying on transitional dynamics. To keep the number of dimensions within computational feasibility, I focus on the trade policy decision of an open economy with two trading partners in the quantitative exercises.[7]

Home country imposes differential tariffs on country A and B, while treating trade policies in A and B as unresponsive.[8] Suppose country B has a higher initial technology level than A, $\lambda_A(0) < \lambda_B(0)$. To examine the optimal trade policy adopted by countries in different status, I consider three scenarios of Home country, while fixing the level of development in country A and B unchanged. Home country can be under-developed, middle-income, or most-developed.

4.2 Parameterization

The level of $\lambda_A(0)$ and $\lambda_B(0)$ are calibrated to be the 10th and 90th percentile fundamental technology level in the world in 2010. To back out the fundamental technology level, I use the method proposed in Costinot et al. (2012).[9] Home country can be the least-developed — $\lambda_{Home}(0)$ corresponds to 5th percentile in the world, middle-income — 50th percentile, or the most-developed — 95th percentile.

Parameter values are displayed in table 1. Discount factor ρ is calibrated to match the long-term economic growth and interest rate in US between 2000 and 2019. The elasticity of substitution across intermediate varieties is set to be 2, but the results are robust to alternative values of σ. The shape parameter of the Frechet distribution θ is chosen to be 4, which is in the range of the estimates on trade elasticity in the literature (Simonovska and Waugh, 2014). Diffusion parameters are in the range of estimates given in Zhang (2021).[10]

[7]Nash equilibrium and cooperative equilibrium with more than two players are difficult to define and solve. Therefore, these two types of games are not considered in this paper.

[8]It's worth clarifying that the aboved assumptions does not mean Home country is a small open economy in the simulations. Wages and price levels in all countries are endogenously determined, which means Home country's trade policy can influence the equilibrium outcomes of country A and B.

[9]From Costinot et al. (2012), fundamental technology can be backed out from observed technology through $\lambda_n(t) = \Gamma\left(\frac{\theta-1}{\theta}\right)^{-\theta} \cdot \pi_{nn}(t) \cdot E\left[z_n(\omega,t)|\omega \in \Omega_{nt}\right]^{\theta}$ in the Eaton-Kortum model.

[10]This is another paper of mine, where I study the dynamic welfare gains from trade in the presence of knowledge diffusion. I also structurally calibrated the parameters governing the

Data on total factor productivity (TFP), domestic absorption, nominal GDP and interest rate are sourced from PWT 10.0.

In the simulated world, I further assume that the three countries are homogeneous except for their initial technology levels. Following conditions are assumed: labor endowments are identical and constant in the three countries; country-level trade is balanced in every period; trade is costless other than the barrier introduced by tariff; country A and country B impose zero tariffs on their trading partners, regardless of the trade policy adopted by Home country.

Table 1: Model Parameterization

Parameter	Description/Moment	Value
ρ	Preference discount factor	0.96
σ	Elasticity of substitution across intermediate varieties	2
θ	Trade elasticity with respect to variable trade cost	4
d_{ni} $(i \neq n)$	iceberg trade cost	1.0
β	diffusion strength (Zhang, 2021)	0.5
η	arrival rate (Zhang, 2021)	0.2
$\lambda_A(0)$	10 percentile level of λ in 2010 (PWT)	1.5
$\lambda_B(0)$	90 percentile level of λ in 2010 (PWT)	123.5
$\lambda_H(0)$	5th, 50th and 95th percentile level of λ in 2010 (PWT)	1.0/25.7/191.7

Note: This table summarizes the parameter values used in the benchmark simulations. Preference discount factor is calibrated to match long-term economic growth and interest rate in US. Elasticity of substitution between intermediate varieties is set to be 2, yet alternative values have little influence on the results. Trade elasticity is picked in the range of estimates in the literature. Diffusion parameters are chosen to match annual growth rate of technology.

Time is discretized into 100 periods. Each period corresponds to one year in reality. Parameters governing the diffusion process β and η are chosen to match annual growth of technology in the data. Discount factor is set to match annual growth of GDP and annual interest rate.

diffusion process at detailed country-industry level.

5 Simulation Results

Section 5.1 shows the main result of this paper — the optimal trade policy in the open-economy world with international knowledge diffusion. Section 5.2 compares the fully flexible optimal trade policy with a restricted non-discriminatory optimal trade policy and studies the economic significance of the discriminatory nature of the tariff schedules for countries in different development levels. Section 5.3 illustrates the choice of a myopic social planner, which provides implications on balanced-growth-path trade policies. Section 5.4 examines the intertemporal tradeoff that the governments face and provides an intuition why optimal trade policy is discriminatory.

5.1 Optimal Tariff Schedule

Taking all factors into consideration, the optimal tariff schedule is displayed in figure 5.1. The three panels correspond to the three scenarios where Home country is under-developed, middle-income, and most-developed. Optimal tariff on country A and B are lines in blue and red respectively. Green line in figure 5.1 displays the difference of tariff imposed on the two trading partners, indicating the level of discrimination in optimal trade policy adopted by Home country.

First of all, optimal tariff schedule is discriminatory. In order to grow faster, Home country impose a higher tariff on the relatively less productive trading partner — country A. By doing so, trade is diverted towards country B and the probability of learning from B is higher. This finding is appealing for a reconsideration of the non-discrimination principle mandated by WTO and many trade agreements. The rationale behind is that, although discrimination in trade policy introduces efficiency loss by distorting comparative advantages, it also induces faster growth by utilizing the externality between international trade and knowledge diffusion.

It is also noteworthy that, in the long run, optimal tariff is uniform. This is a result of the parameter values chosen in the benchmark simulations. As countries converge to the same development level, the incentive to influence diffusion through trade policies diminishes over time and Home country imposes a uniform tariff on imports from A

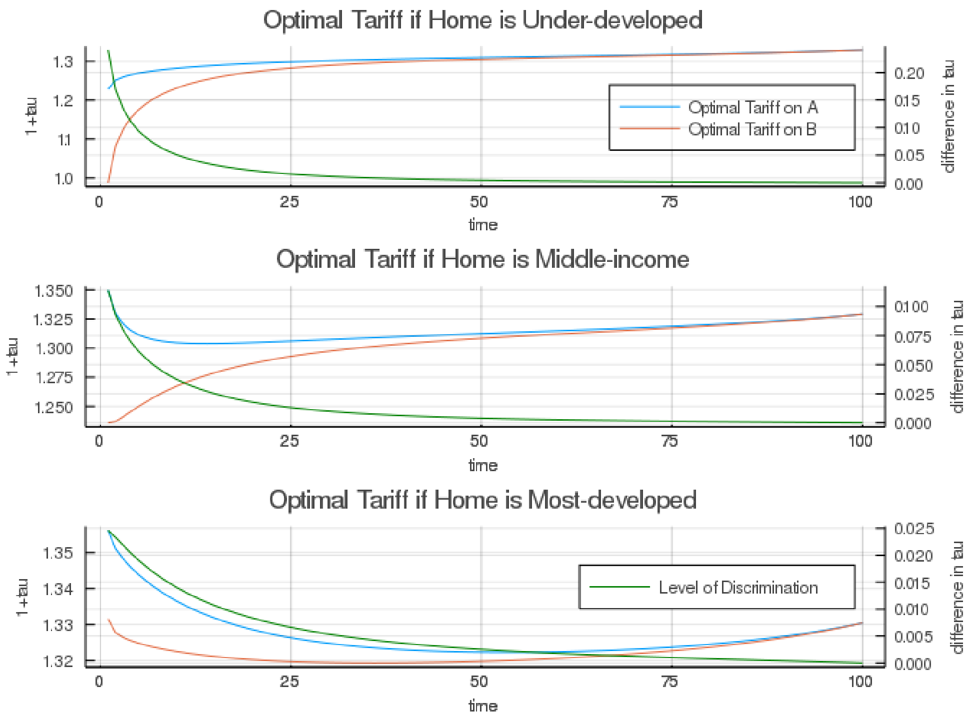


Figure 5.1: Optimal Tariff Schedule and Level of Discrimination

Notes: The figures display the optimal trade policies that maximize social welfare in three scenarios of Home country — under-developed at 5th percentile, middle-income at 50th percentile, and most-developed at 95th percentile. Trading partners country A and country B correspond to 10th and 90th percentile development economy in the world respectively.

and B in the long run.

Long-run tariff converges to the BGP tariff, which is equivalent to the myopic optimal tariff, which is positive in this case.

As for the level of tariff, Home country is willing to impose higher tariff on its trading partners if itself is more advanced technologically. This is because a higher tariff not only diverts trade towards the other trading partner, but also increases its demand on domestic outputs. The more productive its domestic producers are, the less the adjustment costs.

5.2 Comparison between Discriminatory and MFN Trade Policies

Suppose Home country is bounded by trade agreements and has to impose the same level of tariff on all trading partners. Optimal MFN tariff schedule is shown in figure 5.2.

When Home country is under-developed or middle-income, it will impose a lower tariff in the short run to enjoy the benefits of diffusion. When Home country is the leading economy, it will pay more attention to the business-killing effect. Long-term tariffs are positive due to the terms-of-trade argument.

Welfare under this regime is strictly smaller than the welfare generated by a discriminatory optimal tariff schedule. Under the parameters set in benchmark simulations, a discriminatory trade policy promotes welfare of an under-developed country by 2.14% compared to free trade, while a uniform tariff scheme can only do so by 1.98%.

Define welfare gains from trade policy to be the increase in welfare under a certain trade policy compared to free trade. The ability to discriminate imports by origins increase welfare gains from trade policy by 7.87% if Home country is at the 5th percentile. However, the number is 0.08% for a 95th Home country. This implies that, following the non-discrimination principle in trade agreements is reasonable for a developed country, while developing countries will be significantly better off to have full flexibility in making trade policy decisions.

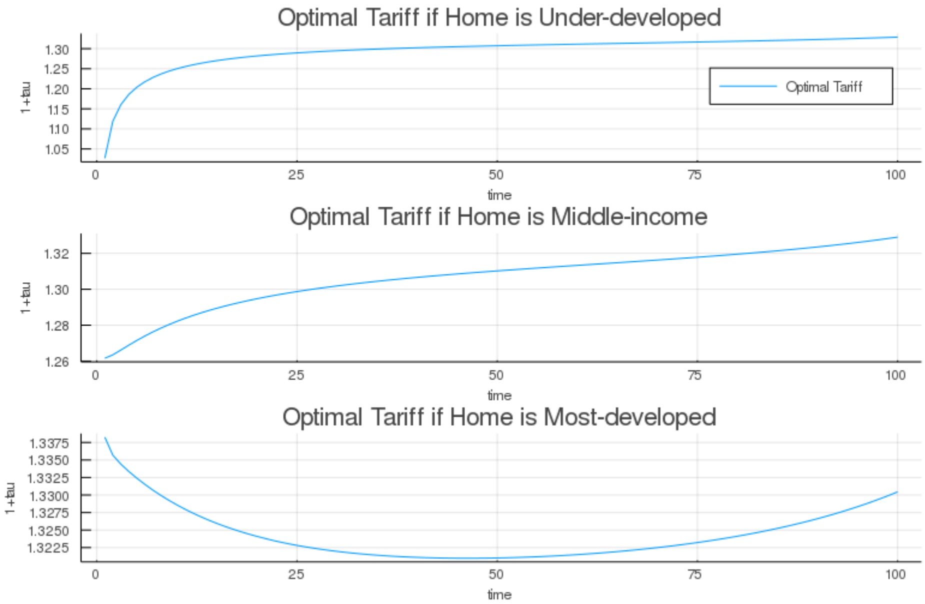


Figure 5.2: Optimal MFN Tariff Schedule

Notes: The figures display the optimal tariff schedules if Home country is bounded by non-discrimination principle. Country A and B receive the same level of tariff from Home in this simulation.

Table 2: Welfare Comparison Of Trade Policies

Home country development level	Welfare Gains non-discriminatory	Welfare Gains discriminatory	Percentage Change
under-developed	1.98%	2.14%	7.87%
middle-income	2.18%	2.21%	1.35%
most-developed	2.20%	2.20%	0.08%

Notes: Welfare gains are defined as welfare gains from trade policies — changes in lifetime utility compared to free trade under different trade policy schemes.

5.3 Myopic Social Planner

As analyzed in Section 3, optimal trade policy in the long run (BGP) for this economy is equivalent to the trade policy preferred by a myopic social planner. The reason is that in the long run, technology differences in the world diminish, economic growth rate is exogenously determined by the growth rate of insight arrival rates, and optimal trade policy does not want to intervene in the diffusion process.

In this section, I study how consumption, real wage and tariff revenue in real terms change as the tariff imposed by Home country varies. Country A and B receive the same level of tariff in this exercise. Since trade is assumed to be balanced, income of a representative agent is made up of wage and transfer financed by tariff revenue $I_n(t) = w_n(t)L_n(t) + R_n(t)$. Thus without loss of generality, the effect of a higher tariff on consumption can be decomposed into two channels — wage and tariff revenue.

Figure 5.3 shows that, an increase in the tariff level leads to a decrease in real wage and an increase in tariff revenue in real terms. A higher tariff undermines the comparative advantage Home country has in the global market, thus global demand of Home country's output shrinks and real wage in Home country drops. That real wage falls as tariff rises is consistent with the welfare formula proposed by Arkolakis et al. (2012). Meanwhile, a unilateral increase in tariff schedule can benefit Home country by improv-

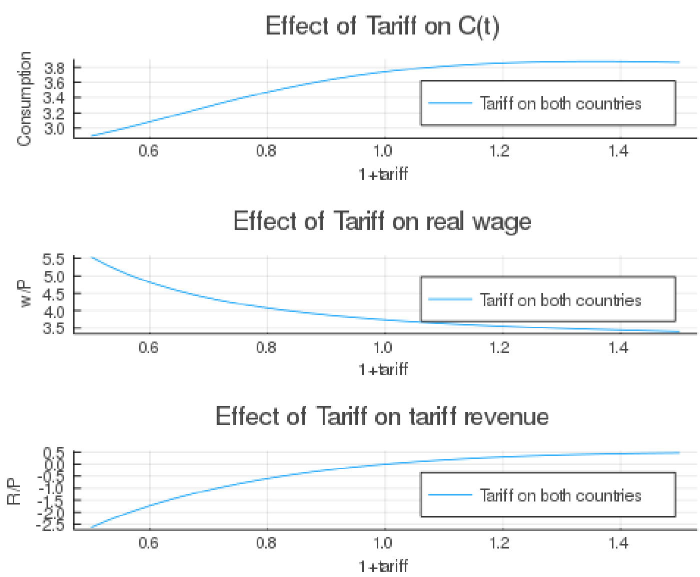


Figure 5.3: The Effects of A Rise in Tariff

Notes: The figures display the changes in Home country's consumption, real wage and tariff revenue in real terms, as Home country imposes a uniform tariff on the two trading partners.

ing its terms of trade and increasing its tariff revenue.[11] In aggregation, consumption exhibits a hump shape with respect to tariff level.

Given that country A and B are assumed to be identical except for the difference in initial technology level, the optimal myopic tariff schedule on the two trading partners are the same. Suppose the government in Home country is myopic and only maximizes utility in current period, then it would set a tariff close to 33%, where consumption curve peaks. Additionally, in the simulated economy, optimal trade policy approaches myopic trade policy in the long run. We can infer that optimal trade policy for Home country converges to a uniform 33% tariff in the long run.

5.4 Intertemporal Tradeoff

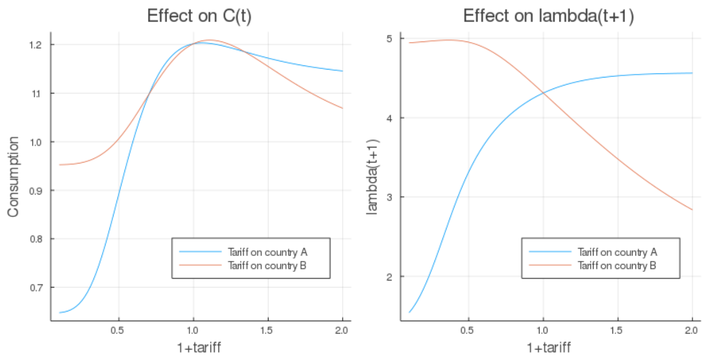
To better understand the discriminatory nature of optimal tariff schedule, I study the intertemporal trade-off social planners face in this section. An increase in tariff can influence welfare through contemporaneous consumption and future technology growth (equation (13)): tariff distort production structure and decrease total output; unilateral increase in tariff can raise a country's terms of trade; a disproportional change of tariff across trading partners will change trade pattern and diffusion process.

As shown in Figure 5.4, by varying tariff on one trading partner while keeping free trade with another, technology of Home country in the next instant exhibits a hump shape. On one hand, taxing imports from relatively less productive trading partners encourage producers in Home country to trade more with relatively more developed countries and enjoy a higher probability of learning from the leading economies. On the other hand, concentration of trade leads to a looser selection of ideas, thus the quality

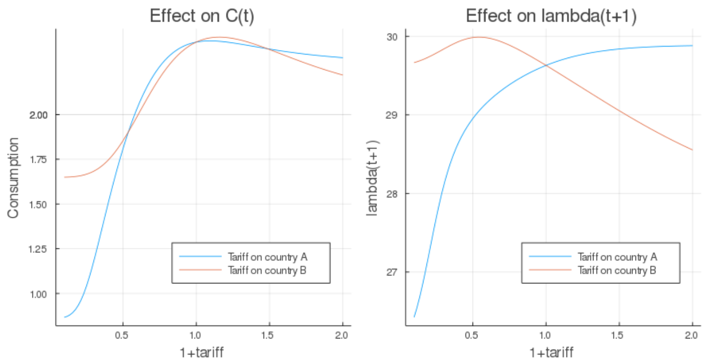
[11] Log-linearization of the tariff revenue formula gives

$$d\ln R_n(t) = \underbrace{\sum_{i=1}^{N} B_{ni}(t) \left[\frac{1}{\tau_{ni}(t)} \underbrace{d\ln(1+\tau_{ni}(t))}_{\text{direct effect}} + \underbrace{d\ln \pi_{ni}(t)}_{\text{cross-substitution effect}} \right]}_{\text{partial equilibrium effect}} + \underbrace{d\ln w_n(t)}_{\text{general equilibrium effect}},$$

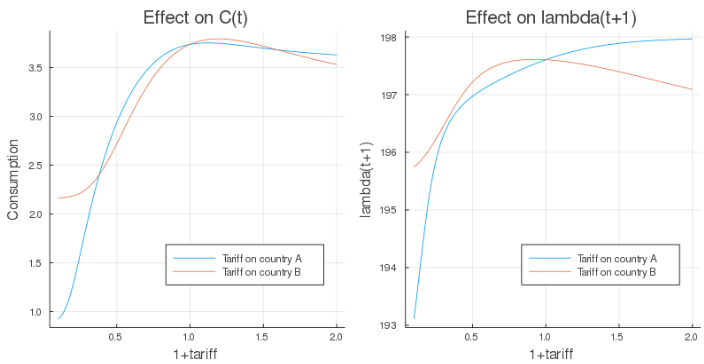
where $B_{ni}(t) = \frac{\pi_{ni}(t)\frac{\tau_{ni}(t)}{\tau_{ni}(t)+1}}{\left(1-\sum_i \pi_{ni}(t)\frac{\tau_{ni}(t)}{1+\tau_{ni}(t)}\right)\left(\sum_i \pi_{ni}(t)\frac{\tau_{ni}(t)}{1+\tau_{ni}(t)}\right)}$. A rise in $\tau_{ni}(t)$ increases per unit income, but reduces trade volumn and general equilibrium wages. Therefore, the effect of tariff on tariff revenue is non-monotonic. See Appendix for derivation.



(a) If Home country is under-developed



(b) If Home country is middle-income



(c) If Home country is most-developed

Figure 5.4: Intertemporal Tradeoff

Notes: The figures display the change in Home country's consumption and next period technology level as Home country imposes an tariff on country A (blue line) or country B (red line) seperately.

of ideas may deteriorate if tariff goes to the extremes.

The peak point of the technology curve by varying tariff on A is always to the right of the peak point by varying tariff on B. Together with the fact that consumption is affected roughly in the same way by tariff on A and B, it is indicated that if Home country tries to influence diffusion through trade policies and foster technology growth, it should impose a higher tariff on country A than B. This result is robust to different initial technology levels of Home country.

6 Extensions

This section tests the sensitivity of benchmark results with respect to change of parameters and inclusion of additional structures.

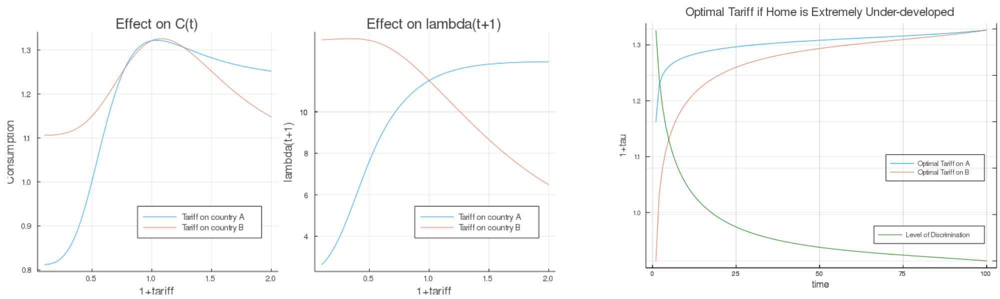
6.1 Extreme Values of Home Country Initial Technology Level

If Home country is extremely lagged behind compared to its trading partners, it can benefit from subsidizing imports from other countries when the gains of knowledge diffusion outweigh the loss in consumption and terms of trade. Higher subsidy will be given to more advanced trading partners, meaning that the optimal tariff schedule is still discriminatory. Figure 6.1 shows the optimal tariff schedule in this case.

Suppose Home country is the extremely more developed than the rest of the world to the extent that the tariff level that maximizes λ_{nt+1} is greater than the tariff level that maximizes C_{nt}. Optimal tariff schedule in this case can be monotonically decreasing, as shown in Figure 6.2.

6.2 Time-varying and Country-specific Arrival Rates of Insight

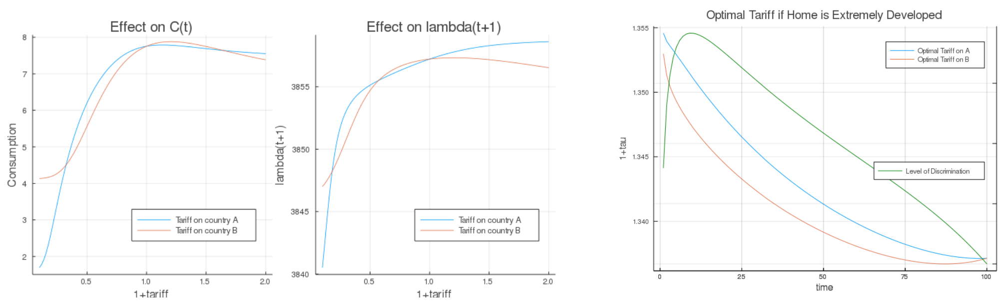
In the benchmark simulations, insight arrival rates are time-invariant and homogeneous across countries. As analyzed in section 3, different assumptions on $\eta_n(t)$ will



(a) Intertemporal Trade off (b) Optimal Tariff

Figure 6.1: Extremely Under-developed Home Country

Note: Home country is assumed to be 10 times less developed than the 5th percentile country



(a) Intertemporal Trade off (b) Optimal Tariff Schedule

Figure 6.2: Extremely Developed Home Country

Note: Home country is assumed to be 10 times more developed than the 95th percentile country lead to different BGP. This section investigates how assumptions of insight arrival rates affect the optimal trade policy.

If all countries share the same non-zero growth rate of $\eta_n(t)$ with the same $\eta_n(0)$, the optimal tariff schedule is the same as the one presented in the benchmark results, as shown in Figure 6.3a.

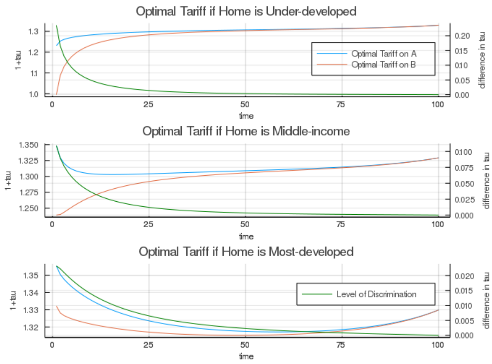
If countries have different initial values of insight arrival rates while growth rate and other parameters remain the same as in the benchmark, $\frac{\lambda_i}{\lambda_n} \neq 1$ for some i on BGP, indicating that Home country has incentives to intervene in knowledge diffusion even on the BGP. In a simulation with finite periods, the optimal trade policy is discriminatory except for the last period. The optimal tariff schedule is shown in Figure 6.3b.

If insight arrival rates grow at different speeds, optimal tariff on BGP is equivalent to myopic optimal tariff because trade share matrix is unresponsive to trade policies and governments only maximize contemporary consumption when making trade policy decisions. If country B has a higher growth rate of η_i than country A, Home country will impose a higher tariff on country A on the transitional dynamics. Figure 6.3c documents the simulated optimal trade policy in this case.

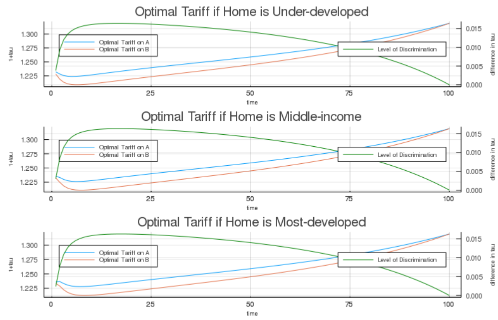
Generally speaking, it is also possible that the optimal policy on BGP is discriminatory even if η's are evolving on the same path. For example, countries with large labor force and demand have greater market power in international relationships, thus they tend to face lower tariff level under the terms-of-trade argument. To tear out the effect of knowledge diffusion on optimal trade policy, I have assumed that countries are homogeneous in labor endowment in the benchmark results.

6.3 Retaliation or Coalition

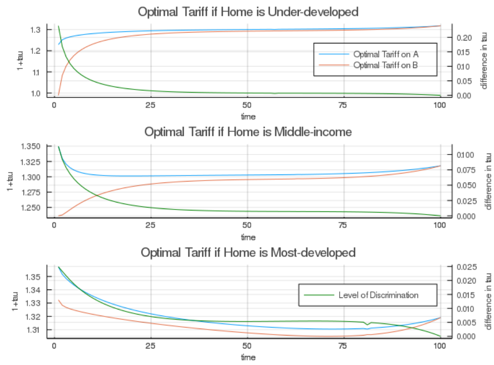
If tariff policies come with a symmetric reciprocation from the corresponding trading partner, Home country will be reluctant to impose any positive tariff on trading partners, because the loss from retaliation outweighs the benefit. However, they might be willing to form trade coalitions. For example, Figure 6.4 shows the intertemporal trade-off and the optimal trade policy if Home country has formed a coalition with country A. In the optimal, Home country will subsidize imports from A and tax imports from



(a) Optimal Tariff with $g_\eta = 1\%$ and $\eta(0) = 0.2$



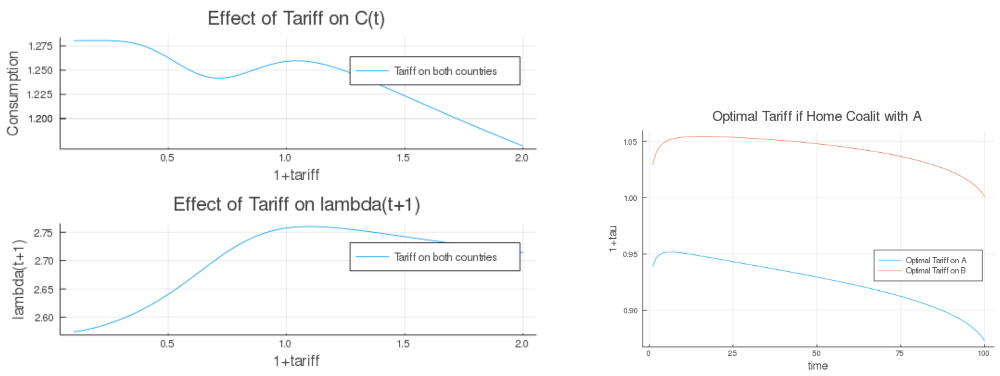
(b) Optimal Tariff with Different η Initial Level



(c) Optimal Tariff with Different η Growth Rate

Figure 6.3: Different Assumptions on $\eta_n(t)$

Notes: The figures display the optimal tariff schedule and level of discrimination when assumptions on insight arrival rates vary.



(a) Intertemporal Trade-off When Coalescing with A

(b) Optimal Tariff When Coalescing with A

Figure 6.4: Retaliation or Coalition

Notes: Panel (a) shows the intertemporal tradeoff of Home country when it ally with country A and Panel (b) shows the optimal trade policy for Home under such arrangement.

B, even though country B is relatively more developed.

7 Conclusion

This paper studies the optimal trade policy in a world where trade is driven by Ricardian forces and ideas can be transmitted along with trade. Since more trade with advanced economies leads to better quality in new insights, countries have incentives to impose discriminatory tariffs on trading partners according to their development levels. The finding that optimal trade policy is discriminatory is robust to a battery of alternative specifications.

This paper calls for a certain level of flexibility in the “non-discrimination” principle for under-developed countries since quantitative simulations show that the ability to impose differential trade policy is significantly welfare improving if Home country is under-developed.

In this paper, optimal tariff in the long run is positive due to the terms-of-trade argument. Optimal trade policy can entail import subsidy if Home country is extremely under-developed compared to its trading partners.

In future studies, the assumption that trade policies are set unilaterally can be relaxed. It would be interesting to examine how trade coalitions and threats of trade war affect optimal trade policy. Additionally, I can compare countries' preferences over different formats of trade agreements and investigate the role of bargaining power in the political economy of trade. Including more industries with different tradability and diffusion strength would be interesting, too. There may be discrimination across industries or need for infant industry protection as a result.

References

Arkolakis, Costas, Arnaud Costinot, and Andres Rodriguez-Clare, "New trade models, same old gains?," *American Economic Review*, 2012, *102* (1), 94–130.

Bagwell, Kyle and Robert W Staiger, "An economic theory of GATT," *American Economic Review*, 1999, *89* (1), 215–248.

_ and _ , "Domestic policies, national sovereignty, and international economic institutions," *The Quarterly Journal of Economics*, 2001, *116* (2), 519–562.

_ and _ , "The Design of Trade Agreements," Working Paper 22087, National Bureau of Economic Research March 2016.

_ , _ , and Ali Yurukoglu, ""Nash-in-Nash" tariff bargaining," *Journal of International Economics*, 2020, *122*, 103263.

_ , _ , and _ , "Quantitative analysis of multiparty tariff negotiations," *Econometrica*, 2021, *89* (4), 1595–1631.

Bickerdike, Charles F, "review of AC Pigou, Preferential and Protective Import Duties," *Economic Journal*, 1907, *17* (65), 98–102.

Borota, Teodora, "Innovation and imitation in a model of North–South trade," *Journal of International Economics*, 2012, *87* (2), 365–376.

Broda, Christian, Nuno Limao, and David E Weinstein, "Optimal tariffs and market power: the evidence," *American Economic Review*, 2008, *98* (5), 2032–65.

Buera, Francisco J and Ezra Oberfield, "The global diffusion of ideas," *Econometrica*, 2020, *88* (1), 83–114.

Caliendo, Lorenzo and Fernando Parro, "Estimates of the Trade and Welfare Effects of NAFTA," *The Review of Economic Studies*, 2015, *82* (1), 1–44.

__ , **Robert C Feenstra, John Romalis, and Alan M Taylor**, "A second-best argument for low optimal tariffs," Technical Report, National Bureau of Economic Research 2021.

Coleman, Chase, Spencer Lyon, Lilia Maliar, and Serguei Maliar, "Matlab, Python, Julia: What to Choose in Economics?," *Computational Economics*, 2020, pp. 1–26.

Costinot, Arnaud, Dave Donaldson, and Ivana Komunjer, "What goods do countries trade? A quantitative exploration of Ricardo's ideas," *The Review of economic studies*, 2012, *79* (2), 581–608.

__ , **Guido Lorenzoni, and Iván Werning**, "A theory of capital controls as dynamic terms-of-trade manipulation," *Journal of Political Economy*, 2014, *122* (1), 77–128.

Demidova, Svetlana and Andrés Rodríguez-Clare, "Trade policy under firm-level heterogeneity in a small economy," *Journal of International Economics*, 2009, *78* (1), 100–112.

Eaton, Jonathan and Samuel Kortum, "Technology, geography, and trade," *Econometrica*, 2002, *70* (5), 1741–1779.

Felbermayr, Gabriel, Benjamin Jung, and Mario Larch, "Optimal tariffs, retaliation, and the welfare loss from tariff wars in the Melitz model," *Journal of International Economics*, 2013, *89* (1), 13–25.

Gros, Daniel, "A note on the optimal tariff, retaliation and the welfare loss from tariff wars in a framework with intra-industry trade," *Journal of international Economics*, 1987, *23* (3-4), 357–367.

Helpman, E, "Innovation, imitation, and intellectual property rights," *Econometrica*, 1993, *61* (6), 1247–1280.

Keller, Wolfgang, "Trade and the Transmission of Technology," *Journal of Economic growth*, 2002, *7* (1), 5–24.

— **and Stephen R Yeaple**, "Multinational enterprises, international trade, and productivity growth: firm-level evidence from the United States," *The Review of Economics and Statistics*, 2009, *91* (4), 821–831.

Ossa, Ralph, "A "new trade" theory of GATT/WTO negotiations," *Journal of Political Economy*, 2011, *119* (1), 122–152.

— , "Trade wars and trade talks with data," *American Economic Review*, 2014, *104* (12), 4104–46.

Perla, Jesse, Christopher Tonetti, and Michael E Waugh, "Equilibrium technology diffusion, trade, and growth," *American Economic Review*, 2021, *111* (1), 73–128.

Sampson, Thomas, "Dynamic selection: an idea flows theory of entry, trade, and growth," *The Quarterly Journal of Economics*, 2016, *131* (1), 315–380.

— , "Technology Transfer in Global Value Chains," 2021.

Simonovska, Ina and Michael E Waugh, "The elasticity of trade: Estimates and evidence," *Journal of international Economics*, 2014, *92* (1), 34–50.

Sjöholm, Fredrik, "International transfer of knowledge: the role of international trade and geographic proximity," *Weltwirtschaftliches Archiv*, 1996, *132* (1), 97–115.

Torrens, Robert, "Letters on commercial policy.," *London: Longman.*, 1833.

Zhang, Chi, "Welfare Gains from Trade Under International Technology Diffusion," *Working Paper*, 2021.

A Appendices

A.1 Derive Price Index and Trade Share Equation

By the assumption on productivity draw distribution, $\Pr\{z_n(\omega,t) < z\} = e^{-\lambda_n(t)z^{-\theta}}$ and perfect competition in goods market $p_{ni}(\omega,t) = \frac{c_i(t)\kappa_{ni}(t)}{z_i(\omega,t)}$,

$$\Pr\left\{p_{ni}(\omega,t) \le p\right\} = 1 - e^{-\lambda_i(t)(c_i(t)\kappa_{ni}(t))^{-\theta}p^{\theta}}.$$

Furthermore, since $p_n(\omega,t) = \min_i\{p_{ni}(\omega,t)\}$,

$$\Pr\left\{p_n(\omega,t) \le p\right\} = 1 - \prod_{i=1}^{N} \Pr\left\{p_{ni}(\omega,t) \ge p\right\} = 1 - e^{-\sum_{i=1}^{N}\left[\lambda_i(t)(c_i(t)\kappa_{ni}(t))^{-\theta}\right]p^{\theta}}.$$

Industry-level price index is given by $P_n(t) = \left\{\int p_n(\omega,t)^{1-\sigma}d\omega\right\}^{\frac{1}{1-\sigma}}$. Additionally, since $p_n(\omega,t)$ follows Frechet distribution, $p_n(\omega,t)^{\theta}$ follows exponential distribution.

$$\Pr\left\{p_n(\omega,t)^{\theta} < p\right\} = \Pr\left\{p_n(\omega,t) \le p^{\theta}\right\} = 1 - e^{-\sum_{i=1}^{N}\left[\lambda_i(t)(c_i(t)\kappa_{ni}(t))^{-\theta}\right]p}$$

Thus,

$$\begin{aligned}
(P_n(t))^{1-\sigma} &= \int_0^1 p_n(\omega,t)^{1-\sigma}d\omega = E\left[p_n(\omega,t)^{1-\sigma}\right] \quad \text{(By LLN)}\\
&= \int_0^{\infty}\left\{\sum_{i=1}^{N}\left[\lambda_i(t)(c_i(t)\kappa_{ni}(t))^{-\theta}\right]\right\} e^{-\sum_{i=1}^{N}\left[\lambda_i(t)(c_i(t)\kappa_{ni}(t))^{-\theta}\right]p}\cdot p^{\frac{1-\sigma}{\theta}}dp\\
&\text{Let } \Phi_n(t) = \sum_{i=1}^{N}\left[\lambda_i(t)(c_i(t)\kappa_{ni}(t))^{-\theta}\right]p,\\
&= \left\{\sum_{i=1}^{N}\left[\lambda_i(t)(c_i(t)\kappa_{ni}(t))^{-\theta}\right]\right\}^{-\frac{1-\sigma}{\theta}} \cdot \int_0^{\infty} e^{-\Phi_n(t)}\cdot(\Phi_n(t))^{\frac{1-\sigma}{\theta}}d\Phi_n(t)\\
&= \left\{\sum_{i=1}^{N}\left[\lambda_i(t)(c_i(t)\kappa_{ni}(t))^{-\theta}\right]\right\}^{-\frac{1-\sigma}{\theta}} \cdot \Gamma\left(\frac{1-\sigma}{\theta}+1\right)\\
\Rightarrow P_n(t) &= \Gamma\left(\frac{1-\sigma}{\theta}+1\right)^{\frac{1}{1-\sigma}} \cdot \left\{\sum_{i=1}^{N}\left[\lambda_i(t)(c_i(t)\kappa_{ni}(t))^{-\theta}\right]\right\}^{-\frac{1}{\theta}}
\end{aligned}$$

As for the expenditure shares, by the law of large numbers,

$$\begin{aligned}
\pi_{ni}(t) = \frac{X_{ni}(t)}{X_n(t)} &= \Pr\left[p_{ni}(\omega,t) \leq \min_{i'\neq i} p_{ni'}(\omega,t)\right] \\
&= \int_0^\infty \prod_{i'\neq i} \Pr\left[p \leq p_{ni'}(\omega,t)\right] \cdot \Pr\left[p_{ni}(\omega,t) = p\right] dp \\
&= \int_0^\infty \prod_{i'\neq i} e^{-\lambda_{i'}(t)(c_{i'}(t)\kappa_{ni'}(t))^{-\theta}p^\theta} \cdot \lambda_i(t)(c_i(t)\kappa_{ni}(t))^{-\theta}\theta p^{\theta-1} e^{-\lambda_i(t)(c_i(t)\kappa_{ni}(t))^{-\theta}p^\theta} dp \\
&= \left[\lambda_i(t)(c_i(t)\kappa_{ni}(t))^{-\theta}\right] \int_0^\infty e^{-\sum_{i'=1}^N \left[\lambda_{i'}(t)(c_{i'}(t)\kappa_{ni'}(t))^{-\theta}\right]p^\theta} dp^\theta \\
&= \frac{\lambda_i(t)(c_i(t)\kappa_{ni}(t))^{-\theta}}{\sum_h \lambda_h(t)(c_h(t)\kappa_{nh}(t))^{-\theta}} \int_0^\infty e^{-x} dx \\
&= \frac{\lambda_i(t)(c_i(t)\kappa_{ni}(t))^{-\theta}}{\sum_h \lambda_h(t)(c_h(t)\kappa_{nh}(t))^{-\theta}}
\end{aligned}$$

A.2 Log-Linearization of the Static Model

A.2.1 Welfare and Instantaneous Utility

From individual's budget constraint, consumption in real terms is

$$C_n(t) = \frac{w_n(t)L_n(t) + R_n(t) + D_n(t)}{P_n(t)L_n(t)}. \tag{A1}$$

Log-linearize equation (A1) gives

$$d\ln C_n(t) = \frac{w_n(t)L_n(t)}{I_n(t)} d\ln w_n(t) + \frac{R_n(t)}{I_n(t)} d\ln R_n(t) - d\ln P_n(t), \tag{A2}$$

where $I_n(t) = w_n(t)L_n(t) + R_n(t) + D_n(t)$ is GDP of country n. $L_n(t)$ and $D_n(t)$ are assumed to be exogenous.

A.2.2 Trade Share and Prices

Under the model setup, expenditure share $\pi_{ni}(t)$ and price index $P_n(t)$ are given by

$$\pi_{ni}(t) = \frac{\lambda_i(t)\left(w_i(t)\kappa_{ni}(t)\right)^{-\theta}}{\sum_h \lambda_h(t)\left(w_h(t)\kappa_{nh}(t)\right)^{-\theta}}, \tag{A3}$$

$$P_n(t) = A\left[\sum_i \lambda_i(t)\left(w_i(t)\kappa_{ni}(t)\right)^{-\theta}\right]^{-\frac{1}{\theta}}. \tag{A4}$$

Log-linearization gives

$$d\ln\pi_{ni}(t) = d\ln\lambda_i(t) - \theta d\ln w_i(t) - \theta d\ln\kappa_{ni}(t) + \theta d\ln P_n(t), \tag{A5}$$

$$d\ln P_n(t) = \sum_h \pi_{nh}(t)\left(d\ln w_h(t) + d\ln\kappa_{nh}(t) - \frac{1}{\theta}d\ln\lambda_h(t)\right). \tag{A6}$$

A.2.3 Tariff Revenue

By definition, tariff revenue is

$$R_n(t) = \sum_{i=1}^{N}\tau_{ni}(t)\frac{\pi_{ni}(t)}{1+\tau_{ni}(t)}I_n(t). \tag{A7}$$

Since $I_n(t) = w_n(t)L_n(t) + R_n(t)$, equation (A7) is equivalent to

$$R_n(t) = \frac{\sum_{i=1}^{N}\tau_{ni}(t)\frac{\pi_{ni}(t)}{1+\tau_{ni}(t)}}{1-\sum_{i=1}^{N}\tau_{ni}(t)\frac{\pi_{ni}(t)}{1+\tau_{ni}(t)}}w_n(t)L_n(t). \tag{A8}$$

Denote $\zeta_{ni}(t) \equiv \tau_{ni}(t)\frac{\pi_{ni}(t)}{1+\tau_{ni}(t)}$, then

$$d\ln R_n(t) = d\ln w_n(t) + \frac{\sum_i\left[\zeta_{ni}(t)\left(d\ln\pi_{ni}(t) + \frac{1}{\tau_{ni}(t)}d\ln(\tau_{ni}(t)+1)\right)\right]}{\left(\sum_{i=1}^{N}\zeta_{ni}(t)\right)\left(1-\sum_{i=1}^{N}\zeta_{ni}(t)\right)} \tag{A9}$$

A.2.4 Marketing Clearing Condition

Labor market clearing condition is equivalent to

$$w_n(t)L_n(t) = \sum_i \frac{\pi_{ni}(t)}{1+\tau_{ni}(t)}\left(w_i(t)L_i(t) + R_i(t)\right) \tag{A10}$$

Log-linearize equation (A10) gives

$$\begin{aligned} d\ln w_n(t) = & \sum_i \frac{\pi_{ni}(t)}{1+\tau_{ni}(t)}\left[\frac{w_i(t)L_i(t)+R_i(t)}{w_n(t)L_n(t)}\left(d\ln\pi_{ni}(t) + d\ln(1+\tau_{ni}(t))\right)\right] + \\ & \sum_i \frac{\pi_{ni}(t)}{1+\tau_{ni}(t)}\left[\frac{w_i(t)L_i(t)}{w_n(t)L_n(t)}d\ln w_i(t) + \frac{R_i(t)}{w_n(t)L_n(t)}d\ln R_i(t)\right] \end{aligned} \tag{A11}$$

A.2.5 Dynamic Transition

Law of motion for technology is

$$\dot{\lambda}_n(t) = \eta_n(t)\Gamma(1-\beta)\sum_i \pi_{ni}(t)\left(\frac{\lambda_i(t)}{\pi_{ni}(t)}\right)^{\beta}. \tag{A12}$$

Log-linearize the growth rate of $\lambda_n(t)$

$$d\ln g_{\lambda_n(t)} = \sum_i \frac{\pi_{ni}(t)^{1-\beta}\lambda_i(t)^{\beta}}{\sum_h \pi_{nh}(t)^{1-\beta}\lambda_h(t)^{\beta}}\left[(1-\beta)d\ln\pi_{ni}(t) + \beta d\ln\lambda_i(t) - d\ln\lambda_n(t)\right] \tag{A13}$$

A.3 Robustness Checks

A.3.1 More than 3 countries

Figure A1 shows the intertemporal tradeoff with 5 trading partners. All countries are homogeneous except for their initial technology level — calibrated to 20th, 40th, 60th, 80th percentile in the world respectively. Optimal tariff is still expected to be higher for less developed trading partners.

A.3.2 Comparative Statics on Parameter Values

For smaller value of discount factor or greater value of intertemporal elasticity of substitution, agents become more risk-averse and value instant gratification more, thus countries are more willing to impose higher tariff on trading partners. (See figure A2 and figure A3.)

If trade elasticity is larger, then by the welfare formula proposed by Arkolakis et al. (2012), the cost of imposing tariff in consumption is greater, thus the magnitude of optimal tariff should be lower. In the meantime, the level of discrimination remain roughly unchanged. (See figure A4.)

For greater value of diffusion strength, countries converge faster, thus the level of discrimination diminishes faster. (See figure A5.)

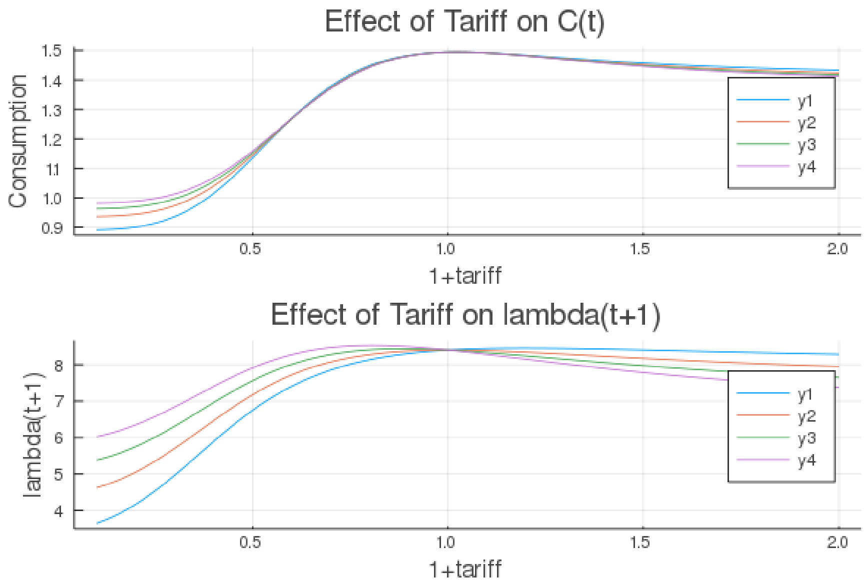


Figure A1: Effect of Tariff on $\lambda_n(t + \Delta)$ with 5 countries

Notes: The figures shows the intertemporal trade-off of Home country with more than three countries in the economy. 'y1' stands for the curve of the most-developed trading partner, while 'y4' stands for the curve of the least-developed trading partner.

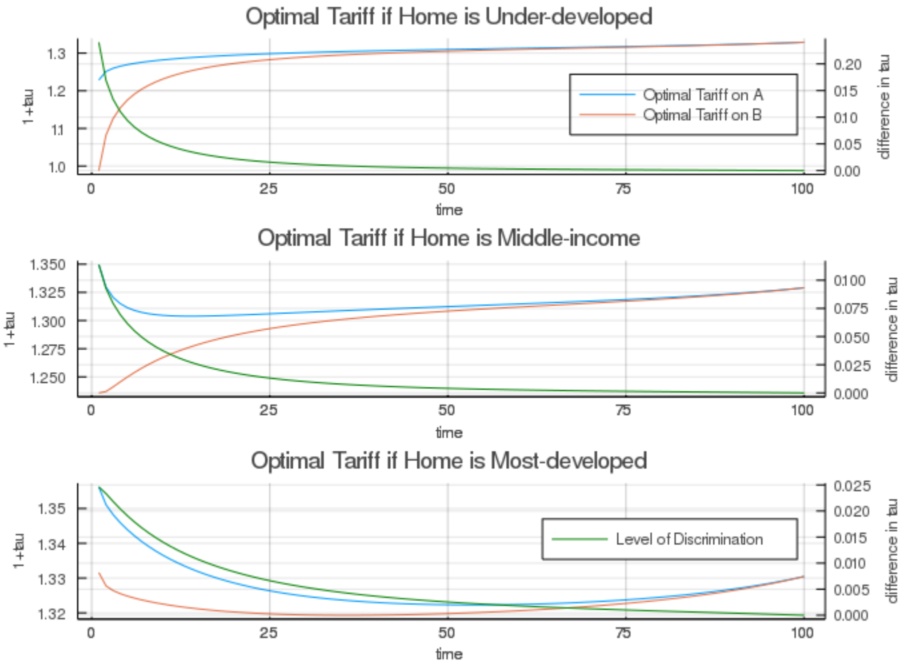


Figure A2: $\rho = 0.6$

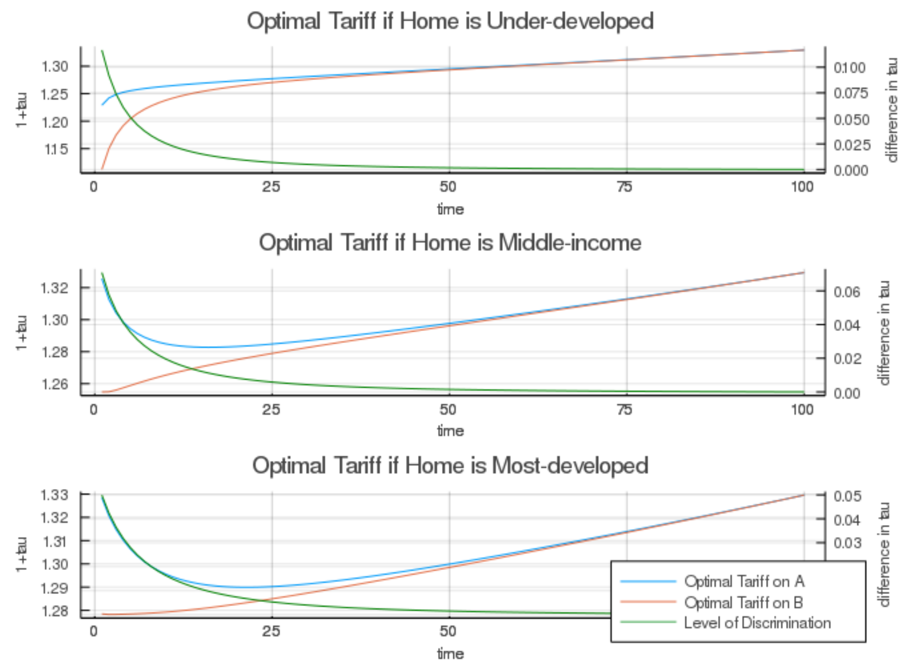


Figure A3: CRRA with intertemporal elasticity of substitution = 2

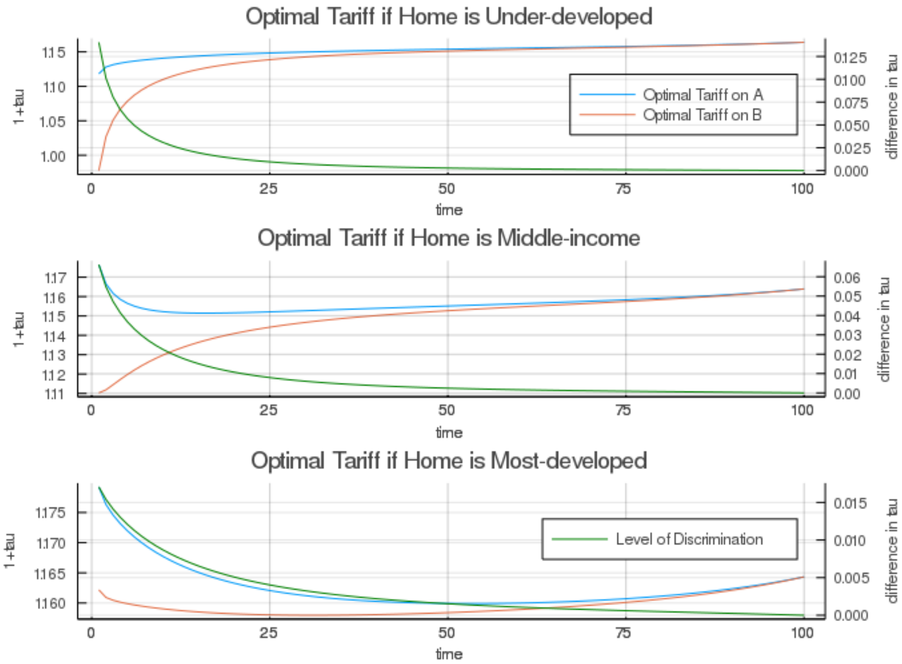


Figure A4: $\theta = 8$

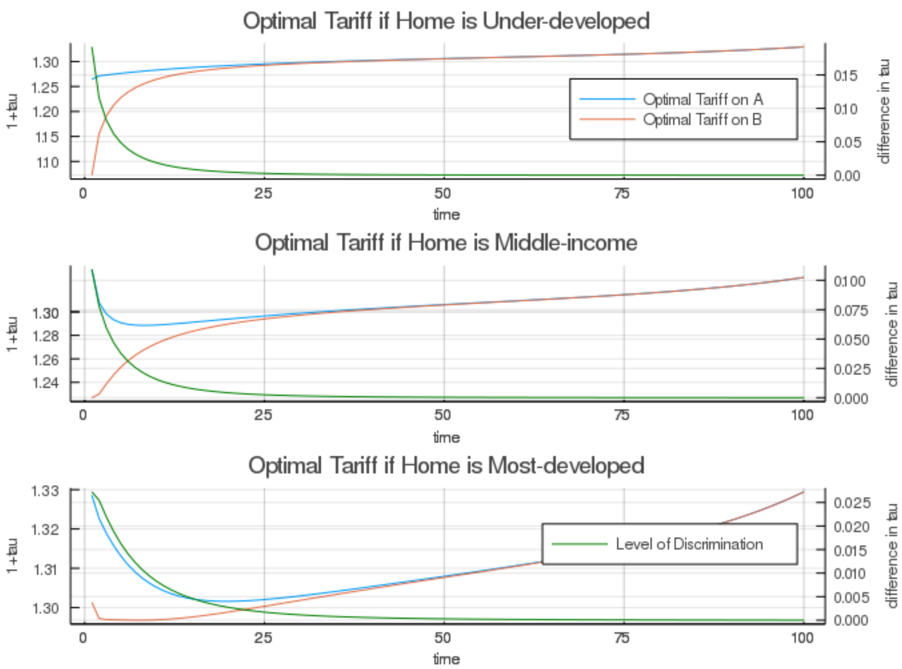


Figure A5: $\beta = 0.7$

A.3.3 Speed of Convergence

In the benchmark result, trade policy converges to BGP policy in 100 periods. This is mainly due to that I am only simulating the model for 100 periods. Figure A6 shows the optimal tariff schedule if the length of periods is set to be 200.

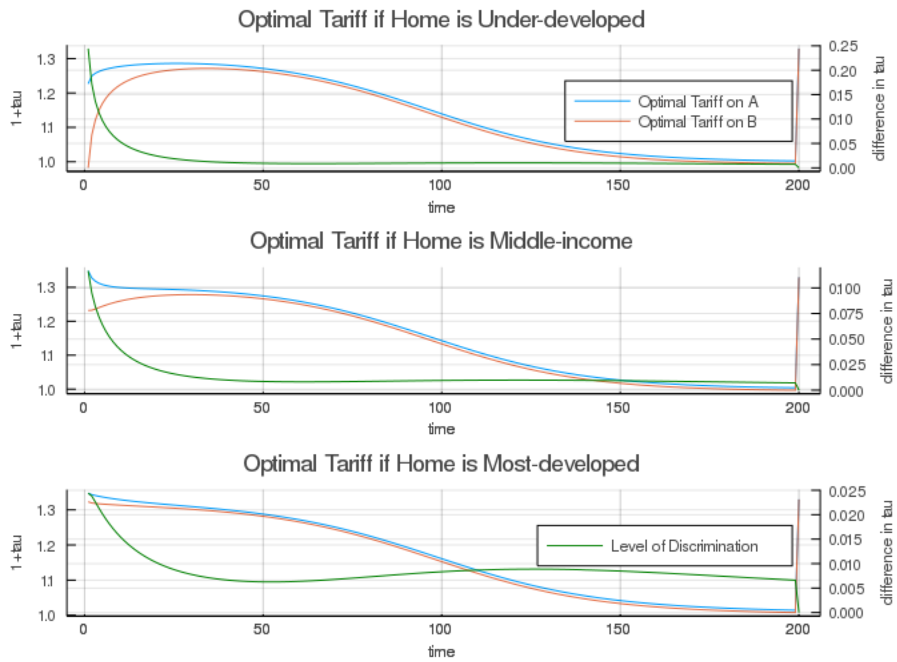


Figure A6: $T = 200$

A.4 Algorithm

Algorithms are written in Julia language, for the sake of speed. Coleman et al. (2020) shows that Julia is around 10 times faster than Python and 30 times faster than Matlab when solving a neoclassical growth model. As country size is set to be 3 and time period 100, I need to solve an optimization problem with 300 choice variables. Running 1 cross-section takes 0.5ms.

Value function iteration methods are not applicable to this paper. This is because BGP of this model is not well-defined. State space is ever expanding. Additionally, the fact that growth rate of the state variable is converging to zero implies that detrending is useless.

To solve the optimization problem, I first create a the choice-variable grid, each variable being discretized with 10 values. Lower bound of tariff is zero; upper bound is set at 50%. (The results are robust to higher values of upper bound.) Objective function is the discounted sum of utility in Home country over 100 periods.

Then I combine a global optimization method (particleswarm) with a local minimizing strategy (gradient descent) to find the maximum. The number of particles in the pso method is set to be 100.