# Optimal Tariff with Firm Heterogeneity, Variable Markups, and FDI*

Ziran Ding
University of Washington

*Version: 1.2*

*Date: April 10, 2020*

**Abstract**

This paper studies the welfare implication of tariff and optimal tariffs in an environment features firm heterogeneity, variable markups and FDI. The findings can be broadly summarized in three aspects. First, the quadratic quasi-linear preference generates multiple externalities in this economy, causing market outcome to differ from the socially optimum outcome systematically. Permitting FDI lowers the domestic cutoff levels and reduces the misallocation in the economy. Second, free trade is not always socially optimal. If the domestic cutoff is sufficiently high, an additional firm entry can improve social welfare. In this case, a positive import tariff is welfare-improving because it encourages firm entry. Third, both positively and normatively, the interaction of variable markup and FDI generates novel trade policy insights that are absent if consumers possess CES preference.

**Keywords:** Optimal tariff, Firm heterogeneity, Misallocation, Variable markup, Foreign direct investment

# 1 Introduction

What is the welfare implication of protectionist trade policy in an environment that features variable markups and foreign direct investment (henceforth, FDI)? On the one hand, protectionism

*I'm indebted to Theo Eicher, Fabio Ghironi and Mu-Jeung Yang for guidance, encouragement and continuous support. For helpful discussions, I thank Kristian Behrens, Keith Head, Antonio Rodriguez-Lopez, Povilas Lastauskas and seminar participants at the Bank of Lithuania, UW International and Macro MTI Brownbag. All errors are my own.

may hurt consumer welfare in the presence of variable markups if protection results in higher market concentration. This dilemma has been a concern since Adam Smith, and it has received increasing attention in recent years[1]. On the other hand, in a highly-integrated global market, foreign firms can avoid import tariffs by locating production within the destination market. Such 'tariff-jumping' activities can diminish the market power of domestic producers, thereby substantially mitigating the welfare consequences of the original protectionist trade policy.

The goal of this paper is to utilize the theoretical framework developed in Ding [2020a] to study the welfare implication of tariff and optimal tariffs. I first follow Nocco et al. [2014] to compare the market allocation with the socially optimum allocation. I find with the free-entry condition, the market outcome in the monopolistically competitive sector[2] is not efficient in several dimensions: (i) The selection is too weak in domestic and export cutoff, but too strong in FDI cutoff. (ii) The market oversupplies high-cost varieties and undersupplies low-cost varieties. (iii) It may feature excessive (insufficient) entry and oversupply (undersupply) the total number of varieties. These market failures stem from several externalities: (i) On the supply side, both the markup-pricing and business-stealing effect tend to create too many varieties in the economy. (ii) On the demand side, the 'love of variety' from the quadratic quasi-linear preference tends to create insufficient varieties in the economy. (iii) With variable markup, firm heterogeneity becomes another source of inefficiency: as the demand becomes more inelastic with consumption, low-cost firms charge higher markups, and do not fully transmit their cost advantage to prices. This behavior leaves inefficiently large room for entry and also allows high-cost firms to be inefficiently large. These externalities collectively result in inefficiencies in the market outcome. If the market selection is too weak, then an increase in tariff can improve market selection, reducing the welfare gap between market allocation and the planner's allocation. In contrast, with constant elasticity of substitution (henceforth, CES) preference, the market outcome coincides with the socially optimum outcome.

Several interesting policy implications stand out from the analysis. First, similar to Cole and Davies [2011], I find that free trade is not always socially optimal. The intuition, however, is different.

[1]Outside of the academic literature, increasing market concentration has received significant attention, e.g., *A lapse in concentration* (The Economist, September 2016), CEA [2016]. In the academic literature, see Asker et al. [2017], De Loecker and Eeckhout [2017] for recent evidence.

[2]Since tariff is only imposed in this sector, here I focus exclusively on the policy implication in this sector. There is, indeed, an inter-sector misallocation between the numéraire good sector and the monopolistically competitive sector. Interesting readers should refer to Nocco et al. [2014].

Cole and Davies [2011] find the socially optimal tariff is always to subsidize trade. They reach this conclusion because, in their model, trade liberalization can foster competition and eliminate the least productive firms, increasing aggregate productivity. Here I find whether imposing a tariff is socially optimal depends on the level of market selection. If the domestic cutoff is sufficiently high, which means the selection is too weak, then an additional firm entry can increase social welfare. In this case, a positive tariff is socially optimal because it encourages firm entry[3]. In addition, I find that the degree of firm heterogeneity, which is governed by the Pareto shape parameter $k$ and the upper bound of cost draw $c_M$, affects the welfare implication of tariff. For example, when the domestic cutoff is sufficiently high, an increase in firm heterogeneity (through an increase in $c_M$ or a decrease in $k$) is socially inefficient because it reduces the positive externality generated through the firm entry, dampening the welfare impact of the tariff.

Second, the Nash tariff is lower than the socially optimal tariff. This result can be analyzed from two perspectives. On the one hand, when Home country sets its uncooperative tariff level, it focuses exclusively on its own tariff revenue and consumer surplus, ignoring the impact on Foreign tariff revenue and consumers. Therefore, Home country will set a higher tariff than the one social planner would choose. On the other hand, different from the intuition in Cole and Davies [2011], a higher Nash tariff level could also arise from Home country's incentive to manipulate the terms of trade. Due to the presence of variable markup, Home's import price varies with its tariff level. Furthermore, the profits of Foreign exporters and multinationals are all affected by Home's tariff level. An increase in Home's import tariff thus generates a terms of trade gain at the cost of Foreign's terms of trade deterioration. This channel is absent in Cole and Davies [2011] due to the constant markup implied by the CES preference.

Third, variable markups, especially their interactions with FDI, yield novel insights on trade policy. While recent studies on the welfare implication of trade liberalization[4] emphasize the importance of variable markup, the insight here is that the role of FDI should not be ignored. A decrease in Home's import tariff makes it easier for the most productive Foreign domestic firms to export, increasing the number of Foreign exporters serving the Home market, and creating downward pressure on the Home average markup. At the same time, the reduction of tariffs also

[3]In Ding [2020a], I established that entry is increasing in tariff, as shown in Proposition 7.

[4]For example, Edmond et al. [2015] and Arkolakis et al. [2018].

makes it less desirable for the least productive Foreign multinationals to pursue FDI, decreasing the number of Foreign FDI firms and generating upward pressure on the Home average markup. If the initial protection level is sufficiently high, the decrease of multinational firms can dominate the increase of exporters, driving up the average markup in the Home market, and generating a negative pro-competitive effect.

Another insight comes from the interaction of variable markup and FDI related to the Nash tariff level. In Cole and Davies [2011], the presence of FDI effectively reduces the tariff base, mitigating the tariff competition between the two countries. In the current framework, FDI's impact on the tariff level depends on a particular parameter that governs the variable markups: the shape of Pareto distribution, $k$. A decrease in $k$ means an increase in the degree of firm heterogeneity, which has two impacts on the economy: changing the equilibrium cutoff levels and altering the relative distribution of firms with different marginal costs. When FDI is an option, the easiness of doing FDI (measured by $\varphi$) interacts with $k$, jointly affecting the Nash tariff level: if the degree of firm heterogeneity is big (smaller $k$), reducing the FDI barrier can lower the Nash tariff level; if the degree of firm heterogeneity is small (larger $k$), however, promoting FDI can increase the Nash tariff level. Another parameter that also impacts the degree of firm heterogeneity, $c_M$, only affects the equilibrium cutoffs proportionally, without shifting the relative distribution of firms. Therefore, its interaction with FDI does not change the Nash tariff level.

The remainder of this paper is organized as follows. Section 2 compares the market allocation with the socially optimum allocation. Section 3 studies whether free trade is socially optimal in this economy. Section 4 contrasts the socially optimal tariff level with the Nash tariff level, and also compares the Nash tariff level with and without FDI. Section 5 is dedicated to the role of variable markup in this economy. The policy implication of its interaction with FDI is also discussed in this section. Section 6 concludes. All tables and graphs to which this paper refers are included in the appendix.

## 2 Social Optimum vs. Market Outcome

In this section, I first derive the socially optimum outcome in the framework introduced in Ding [2020a]. Then I compare the socially optimum outcome with the market outcome derived in Ding [2020a]. Then I analyze the forces in the economy that result in the departure of market outcome from socially optimum outcome.

The social planner's problem can be described as the following. Since the quadratic quasi-linear utility implies transferable utility, social welfare can be expressed as the sum of all the representative consumers' utilities. Following Nocco et al. [2014], the planner chooses the number of entrants($N_E^H, N_E^F$), and production level for homogeneous good and heterogeneous good($q_0^H, q_0^F, q_i^H, q_i^F$) to maximize social welfare subject to aggregate resource budget constraint:

$$\max_{\{N_E^H, q_0^H, q_i^H, N_E^F, q_0^F, q_i^F\}} \mathbb{W} \equiv \mathbb{U}_{\mathbb{H}} + \mathbb{U}_{\mathbb{F}}$$

$$\text{s.t. } q_0^H + q_0^F + f\left(N_E^H + N_E^F\right) + N_E^H \int_0^{c_M} \left[c q_D^H(c) + \tau^F c q_X^H(c) + \varphi^F c q_{FDI}^H(c)\right] dG(c)$$
$$+ N_E^F \int_0^{c_M} \left[c q_D^F(c) + \tau^H c q_X^F(c) + \varphi^H c q_{FDI}^F(c)\right] dG(c) = 2 + \bar{q}_0^H + \bar{q}_0^F$$

where $q_0^H + q_0^F$ stands for the supply of homogeneous good in both countries, $f\left(N_E^H + N_E^F\right)$ the sunk entry cost in the monopolistically competitive sector in $H$ and $F$, $N_E^H \int_0^{c_M} [c q_D^H(c) + \tau^F c q_X^H(c) + \varphi^F c q_{FDI}^H(c)] dG(c)$ the supply of differentiated varieties in the Home country, and the last term on the left-hand side of the constraint gives the supply for differentiated varieties in the Foreign country. On the right-hand side, we have the endowment of labor and homogeneous good in both countries. The differences between socially optimum outcome and market outcome can be summarized in the following proposition:

**Proposition 1.** *In the current framework, compared to the socially optimum, the market outcome differs in several dimensions:*

*(A) Marginal cost cutoffs*

*(i) the Home domestic market selection is weaker than the socially optimum selection*

$$c_D^{HM} > c_D^{HO}$$

*(ii) the Home exporter market selection is weaker than the socially optimum selection*

$$c_X^{HM} > c_X^{HO}$$

*(ii) the Home FDI market selection is stronger than the socially optimum selection*

$$c_{FDI}^{HM} < c_{FDI}^{HO}$$

*(B) Intensive margin*

*(i) Home's domestic producers undersupply varieties with low marginal production cost*[5]*,*

$$q_D^{HM} < q_D^{HO} \text{ if } c < \left[2 - (2/\triangle_F)^{1/(k+2)}\right] c_D^{HO}$$

*(ii) Home's exporters undersupply varieties with low marginal cost*

$$q_X^{HM} < q_X^{HO} \text{ if } c < \left[2 - (2/\triangle_F)^{1/(k+2)}\right] \frac{c_D^{FO}}{\tau^F(2 - t^F)}$$

*(iii) Home's FDI firms also undersupply varieties with low marginal production cost*

$$q_{FDI}^{HM} < q_{FDI}^{HO} \text{ if } c < \left[2 - (2/\triangle_F)^{1/(k+2)}\right] \frac{c_D^{FO}}{\varphi^F}$$

*(C) Extensive margin*

*Depending on the of domestic cutoff levels ($c_D^{HO}$,$c_D^{FO}$) in the socially optimum, the market outcome does not always yield the same level of the total number of varieties ($N^H, N^F$) and the number of entrants ($N_E^H, N_E^F$) as those in the socially optimum.*

**Proof.** See Appendix C.1. □

[5]Please refer to the Appendix for a detailed expression of $\triangle_F$.

This proposition says that the market outcome differs from that of the socially optimum in several dimensions. For the Home country, the market selection is weaker than the planner's selection in both the domestic producers' and exporters' market, but is stronger than the planner's selection in the FDI margin, as illustrated in **Figure 1**. Intuitively, compared to the planner's choice, Home country produces too many varieties in the domestic market, imports too many varieties from abroad, and does not have enough varieties from Foreign FDI firms. On the intensive margin, the more productive domestic producers (i.e. firms with low marginal costs) produce less than the socially optimal level. The same results hold true among the exporters and FDI firms, indicating the distribution of products is skewed too much toward high cost varieties. On the extensive margin, the number of entrants and the equilibrium number of varieties are also different from those of the socially optimum level. The market produces too many products from firms with large marginal costs, but not enough from firms with low marginal costs.

Since the trade policy study in this paper exclusively focuses on the monopolistically competitive sector, the discussion of inter-sector inefficiency is omitted here[6]. The inefficiencies in the market outcome originate from multiple externalities in this economy. Two inefficiencies occur with quadratic quasi-linear preference: (i) The consumers display the 'love of variety' feature, which, however, firms do not consider when making entry decisions. As a result, there are not enough varieties in the economy. (ii) Firms can charge variable markups, and thus firm heterogeneity becomes another source of inefficiency in this economy. Since markup decreases in marginal cost, the low marginal cost firms (more productive) are inefficiently small and high marginal cost firms (less productive) are inefficiently large in the market outcome. The monopoly power in the differentiated-good sector allows a firm to price over its marginal cost. Under the free-entry condition, this externality tends to create too many varieties. The new entrant will take up the market share of existing firms, and this business-stealing effect also tends to create too many varieties. All these externalities work together to generate market failures in the current economy.

According to Dhingra and Morrow [2019], when monopolistic competition is combined with CES preference, the market outcome coincides with the socially optimum outcome. The externalities mentioned above exactly cancel each other out[7]. The current economy, however, deviates from this

[6]Interesting reader could refer to Nocco et al. [2014] for further discussion.

[7]Note, firm heterogeneity does not create externality since all firms charge identical markup under CES preference.

benchmark due to the quadratic quasi-linear preference. The forces that generate externalities do not cancel each other out, and firm heterogeneity becomes a new source of inefficiency in the economy. The market outcome hence differs from the first-best outcome systematically. As discussed in Ding [2020a], tariff can affect not only the cutoff levels, entrants, and number of varieties in the equilibrium, but also social welfare. I will come back to this issue again in Section 4.

# 3 Free Trade and Its Welfare Implication

In this section, I study the classical question in the trade policy literature: is free trade socially optimal in the current economy? Here the free trade is referring to zero net tariff value. To answer this question, I set $t^H = t^F = 1$ and study the joint welfare in $H$ and $F$:

$$\mathbb{W} \equiv \mathbb{U}^H|_{t^H=1} + \mathbb{U}^F|_{t^F=1}$$

The results can be summarized in the following proposition:

**Proposition 2.** *Free trade is, in general, not socially optimal. If $H$ and $F$ start with free trade ($t^H = t^F = 1$), then a small symmetric increment in import tariff raises social welfare if and only if $\tilde{c}_D > \alpha/2$, lowers social welfare if and only if $\tilde{c}_D < \alpha/2$ and has no effect on social welfare if and only if $\tilde{c}_D = \alpha/2$, where $\tilde{c}_D$ is the domestic cutoff under symmetry when net tariff is zero.*

**Proof.** See Appendix C.2. □

Interestingly, in this economy, free trade is not always socially optimal. On the one hand, if $\tilde{c}_D$, the domestic marginal cost cutoff level under symmetry with free trade, is sufficiently high, then a small increase in import tariff will raise the social welfare. This result means that if domestic cutoff is sufficiently high, the market selection too weak, then there are not enough firms competing in the economy. Therefore, the social planner should increase tariff to encourage firm entry. On the other hand, if the domestic cutoff is too low, the market selection too strong, there will be too many firms competing in this economy. In this situation, the social planner should discourage entry by reducing tariff (i.e., subsidize trade). If the domestic marginal cost cutoff level, however, is exactly equal to the threshold value, then free trade is socially optimal. The parameter $\alpha$ here stands for the relative demand of consumers toward differentiated varieties. So whether free trade is socially optimal crucially depends on whether the market outcome could meet consumer's demand on the varieties.

To gain more intuition of this proposition, and to elaborate on the externalities mentioned in Proposition 1, I rewrite the social planner problem as the following:

$$\mathbb{W} = \max_{\{N_E^H, N_E^F\}} I^H + \frac{\alpha - c_D^H}{2\eta}\left(\alpha - \frac{k+1}{k+2}c_D^H\right) + I^F + \frac{\alpha - c_D^F}{2\eta}\left(\alpha - \frac{k+1}{k+2}c_D^F\right)$$

Following Mankiw and Whinston [1986], here I consider a constrained-optimum problem faced by a social planner, who cannot affect the market outcome for any given number of firms. This is particularly relevant under the current heterogeneous firm framework since the first-best outcome cannot be reached due to the presence of externalities in this economy, as discussed in Section 2. On top of that, one should also keep in mind that the presence of a numéraire good adds an extra distortion to the model. There is no markup in the numéraire-good sector, but in the differentiated-good sector, producers charge prices above their marginal costs due to their monopoly power. As pointed out by Bhagwati [1969], the presence of distortions can result in the breakdown of Pareto-optimality of laissez-faire.

The planner chooses the optimal level of entry to maximize social welfare. Under free entry condition, firms make entry decisions irrespective of the externalities that they generate on consumers and other firms, so the market entry level might not be socially desirable. Notice, since wage in the economy equals to one, and tariff revenue equals to zero at free trade, maximizing the income is equivalent to maximizing aggregate profits. After imposing symmetry and $t = 1$, the above social welfare function $\mathbb{W}$ can be rewritten as:

$$\max_{\{N_E\}} \mathbb{W} \equiv \Pi + \underbrace{\frac{\alpha - c_D}{2\eta}\left(\alpha - \frac{k+1}{k+2}c_D\right)}_{\text{Consumer Surplus}}$$

where $\Pi = N_E \times (\bar{\pi} - f_E)$. Consumers take the number of entrants as given and maximize their utilities. The above consumer surplus can be rewritten in terms of the optimized choice of variety $i$ as follow:

$$CS = \frac{1}{2}\gamma \int_{i \in \Omega^H} (\hat{q}_i)^2 \, di + \frac{1}{2}\eta \left(\int_{i \in \Omega^H} \hat{q}_i di\right)^2$$

According to Ottaviano et al. [2002], the first term corresponds to the sum of consumer surplus at each variety $i$, and the second term reflects the variety effect that brings to consumer surplus. To understand the role of the entry in this economy and its impact on the welfare, one can follow Bagwell and Lee [2015] to rewrite the above equation in the following way:

$$CS = \underbrace{N_E \times \frac{\gamma}{2} \int_0^{\tilde{c}_D} \left(q_D\left(c\right)\right)^2 dG\left(c\right)}_{N_E \times \text{Average CS for a variety}} + \underbrace{\frac{\left(\alpha - \tilde{c}_D\right)}{2\eta} \left[\alpha - \frac{\left(k+1\right)\left(1+\tau^{-k}\right)+1}{\left(k+2\right)\left(1+\tau^{-k}\right)} \tilde{c}_D\right]}_{\text{Variety Effect (VE)}}$$

where equation (31) in Ding [2020a] is utilized to express $CS$ in terms of $c_D^H$ to obtain the exact expression of variety effect. The $\tilde{c}_D$ is the domestic cutoff level under symmetry. Therefore the social planner's problem can be further rewritten as

$$\max_{\{N_E\}} \mathbb{W} \equiv N_E \times \text{Avg.CS} + \text{VE} + \Pi$$

The first order condition related to entry will generate the following condition:

$$\underbrace{\text{Avg. CS} + N_E \frac{\partial \text{Avg. CS}}{\partial N_E} + \frac{\partial \text{VE}}{\partial N_E} + N_E \frac{\partial \bar{\pi}}{\partial N_E}}_{\neq 0 \text{ due quadratic quasi-linear preference}} + \underbrace{\bar{\pi} - f_E}_{\text{Free entry}} = 0 \tag{1}$$

The free entry will only take care of the last item, and that is why it is not guaranteed to deliver the socially desirable level of entry. According to the seminal work[8] by Spence [1976] and Mankiw and Whinston [1986], the first term is positive, representing the average consumer surplus gain from a new variety following entry. The second term is negative, representing the average consumer surplus loss for existing varieties when a new variety becomes available (substitution effect). The third item is positive, representing the variety effect/benefit from a new variety. Lastly, the fourth item is negative, which represents the business-stealing effect since it measures how the new entrant affects the average profit of existing firms. These four items added up together give the externality of firms' entry. In the Appendix C.3, I show that this externality effect is positive when $\tilde{c}_D > \alpha/2$, which means firm entry increases social welfare. In this case, positive import tariff will increase social

[8]For recent related discussions under heterogeneous firms framework, see Dhingra and Morrow [2019], Weinberger [2015], Bagwell and Lee [2015] and Behrens et al. [2018].

welfare by encouraging entry. When $\tilde{c}_D < \alpha/2$, however, the sum of these four terms is negative, indicating that firm entry decreases social welfare. In this case, positive import tariff decreases social welfare by introducing more entry. The optimal thing to do in this case is to subsidize trade and discourage firm entry. Only when $\tilde{c}_D = \alpha/2$, the market entry level coincides with the socially optimum entry level.

The result here is different from that in Cole and Davies [2011], where the authors find that the socially optimal tariff in their setting is always a subsidy. The intuition is that opening up to trade will expose domestic firms to the Foreign competition, driving out the least productive firms and reallocating resources to the more productive firms. When trade barrier is a choice variable, the social planner will have an additional incentive to promote trade since trade-liberalization can boost aggregate productivity. In the current setup, their conclusion only holds when $\tilde{c}_D < \alpha/2$. The fundamental reason for this difference is the deviation from CES preference. Under CES preference, the first four terms in equation (1) always add up to zero. As demonstrated in Dhingra and Morrow [2019], free-entry delivers the first-best outcome. In the current framework, the presence of quadratic quasi-linear preference creates variable elasticity across the varieties, generating multiple externalities in the economy, causing the sum of those four terms to deviate from zero.

More importantly, firm heterogeneity, which does not produce any externality under the CES preference, now becomes a source of inefficiency in this economy. To be more specific, firm heterogeneity is governed by two parameters here: $c_M$ and $k$. $c_M$ represents the upper bound of the marginal cost distribution. Larger $c_M$ indicates larger region that marginal cost can be drawn from, leading to an increase in firm heterogeneity. $k$ governs the shape of Pareto distribution. When $k$ equals to 1, the marginal cost follows uniform distribution, and different marginal costs can be drawn with equal probability. As $k$ approaches infinity, the marginal cost distribution becomes degenerate at $c_M$. Therefore, an increase in $k$ means the distribution of firms is skewed toward less productive firms, reducing the degree of firm heterogeneity.

For example, suppose $\tilde{c}_D > \alpha/2$, which means the sum of the first four terms in equation (1) is positive, then an additional entry creates positive externality to the society. In this situation, an increase in $c_M$ or a decrease in $k$ (both represent an increase in firm heterogeneity) will reduce the aggregate externality of firm entry. Therefore, an increase in firm heterogeneity is socially inefficient because it holds back the positive externality of firm entry. However, these two dimensions work quite differently. An increase in $c_M$ will reduce the average consumer surplus, the absolute value of substitution effect, the variety effect, and the absolute value of business-stealing effect. The positive terms (average CS and VE) dominate the negative terms (substitution and business-stealing effect), and therefore externality decreases as $c_M$ increases. A decrease in $k$ has the exact opposite impacts on these four terms. The negative terms dominate the positive terms, and therefore externality decreases as $k$ decreases. The fundamental reason behind this observation can be seen from the solutions of cutoffs and the expressions of these four terms[9]. $c_M$ only affects the share of valid varieties on the market ($\tilde{c}_D/c_M$), but $k$ affects both the share of valid varieties and the average profits. As we will see in the next section, this impact on cutoffs generates important welfare implications of tariff.

# 4 Socially Optimum Tariff and Nash Tariff

In this section, I first follow the discussion in Section 2 to study the welfare implication of tariff, with an emphasis on the comparison between market outcome and socially optimum outcome. Then I derive the socially optimal tariff and the Nash tariff in this economy, and investigate the forces that result in the differences between them. Toward the end of this section, I compare the Nash tariff with FDI with the Nash tariff without FDI.

## 4.1 Welfare Implication of Tariff

Imposing symmetry and utilizing the setup in Appendix C.1 and the model setup in Ding [2020a], one can write the social welfare in market outcome and socially optimum outcome as functions of the corresponding cutoffs:

$$\mathbb{W}^M = 1 + \bar{q}_0 + \frac{\alpha - c_D^M}{2\eta}\left(\alpha - \frac{k+1}{k+2}c_D^M\right)$$
$$\mathbb{W}^O = 1 + \bar{q}_0 + \frac{1}{2\eta}\left(\alpha - c_D^O\right)^2$$

[9]Please refer to Appendix C.3.

Therefore, the welfare implication of a tariff change can be summarized in the following proposition:

**Proposition 3.** *(i) Imposing symmetry, if the domestic cutoff level is sufficiently high, i.e.,* $c_D^M > \alpha(A+B)/2B$[10]*, then a bilateral increase of tariff is welfare-improving; if the domestic cutoff level is sufficiently low, i.e.,* $c_D^M < \alpha(A+B)/2B$*, then a bilateral increase of tariff is welfare-deteriorating. (ii) Removing symmetry, then a unilateral increase of Home's import tariff is welfare-improving to the Home country, but welfare-deteriorating to the Foreign country.*

**Proof.** See Appendix C.4. □

The first part of this proposition is directly built on Proposition 1 and 2. On the one hand, when the domestic cutoff is sufficiently high, the market selection is too weak. In this case, the market outcome does not have enough varieties. According to Proposition 7 in Ding [2020a], an increase in import tariff can increase the level of entry, foster market selection, reduce the equilibrium domestic cutoff level, and therefore improve the social welfare. On the other hand, if the domestic cutoff is too low, which means the market selection is too strong, then an increase in import tariff will only make the market selection stronger, deteriorating the social welfare. From another perspective, the reduction in domestic cutoff reduces the prices charged by all the firms, but more so for the more productive firms that charge higher markups. This effect reduces the distortions created by all the externalities mentioned in Proposition 1, thereby improving social welfare.

The second part of this proposition is more straightforward. Based on Proposition 2 and 7 in Ding [2020a], an increase in Home's import tariff increases firm entry, resulting in more equilibrium varieties and lower domestic cutoff. These impacts raise the welfare in the Home country. For the Foreign country, Home's import tariff decreases firm entry, reducing the equilibrium number of varieties and producing a higher domestic cutoff. These impacts deteriorate Foreign country's welfare. With these understandings about the welfare implications of tariff in mind, now I turn to the study of optimal tariffs in this economy.

## 4.2 Socially Optimal Tariff vs. Nash Tariff

In this subsection, I derive both the socially optimum tariff and Nash tariff. For the socially optimum tariff, due to the symmetric nature of Home and Foreign, I assume the social planner puts

[10] Please refer Appendix C.4 for the exact expression of A and B.

identical weight on the welfare of each country. Thus, the socially optimal tariff maximizes the sum of the two countries' consumer utilities:

$$\max_{\{t^H,t^F\}} \mathbb{W} \equiv \mathbb{U}^H + \mathbb{U}^F = I^H + CS^H + I^F + CS^F$$

Recall that the income $I^l \equiv w^l + (t^l - 1) \times IM^l + \Pi^l$ for $l \in \{H, F\}$, and the equilibrium wages in both countries are equal to one. It is straightforward to verify that the optimal level of $t^H$ satisfies the following condition:

$$\frac{\partial \mathbb{W}}{\partial t^H} = \underbrace{IM^H + (t^H - 1) \times \frac{\partial IM^H}{\partial t^H}}_{\text{Effect on } H\text{'s tariff revenue}} + \underbrace{\left(t^F - 1\right) \times \frac{\partial IM^F}{\partial t^H}}_{\text{Effect on } F\text{'s tariff revenue}} + \underbrace{\frac{\partial CS^H}{\partial t^H} + \frac{\partial CS^F}{\partial t^H}}_{\text{Effect on } CS^H \text{ and } CS^F} = 0 \tag{2}$$

From the socially optimum perspective, the social planner needs to consider Home import tariff's impact on the tariff revenue and consumer surplus in both countries. In contrast, the Nash tariff level for $H$ only focuses on the tariff revenue and consumer surplus of its own country. It is defined as follow:

$$\max_{\{t^H\}} \mathbb{U}^H = I^H + CS^H$$

It is easy to show that the optimal non-cooperative tariff level should satisfy:

$$\frac{\partial \mathbb{U}^H}{\partial t^H} = \underbrace{IM^H + (t^H - 1) \times \frac{\partial IM^H}{\partial t^H}}_{\text{Effect on } H\text{'s tariff revenue}} + \underbrace{\frac{\partial CS^H}{\partial t^H}}_{\text{Effect on } CS^H} = 0 \tag{3}$$

Combining equation (2) and (3), and if we focus on one of the cases in Proposition 2, $\tilde{c}_D > \alpha/2$, which implies the socially optimal import tariff is greater than one, it then can be shown that the Nash tariff level for country $H$ satisfies the following proposition:

**Proposition 4.** *When the symmetric domestic cost cutoff is sufficiently high ($\tilde{c}_D > \alpha/2$), the Nash tariff ($t_N$) is higher than the socially optimal tariff ($t_S$).*

**Proof.** See Appendix C.5. □

This finding confirms the Proposition 2 in Cole and Davies [2011], but it is established in the Melitz and Ottaviano [2008] framework with the presence of FDI. We can obtain the intuitions from several different angles. To understand the incentive of Home's import tariff, let us investigate $F$'s free-entry condition, which can be expressed in a similar fashion as in equation (22) in Ding [2020a]:

$$\underbrace{\left(c_D^F\right)^{k+2}}_{\uparrow \text{in } t^H} + \underbrace{\Phi_1^H}_{\downarrow \text{in } t^H} \underbrace{\left(c_D^H\right)^{k+2}}_{\downarrow \text{in } t^H} + \underbrace{\Phi_2^H}_{\uparrow \text{in } t^H} \underbrace{\left(c_D^H\right)^{k+2}}_{\downarrow \text{in } t^H} = \gamma\phi$$

where the first term on the left represents the expected profit of being a domestic producer in $F$, the second term is the expected profit of being an exporter in $F$, and the third term is the expected profit of being multinational firm in $F$. When $H$ country sets its tariff, according to Proposition 4-6 in Ding [2020a], $F$'s exporter cutoff level decreases in $t^H$. Based on the proof in Ding [2020a], $\Phi_1^H$ decreases in $t^H$, and $\Phi_2^H$ increases in $t^H$. So if $t^H$ increases, the expected profit of an exporter in $F$ goes down. In fact, the sum of the expected profit of exporter and multinational in $F$ also goes down. When $H$ sets its unilateral optimal tariff, it ignores the impact of its tariff on $F$'s exporter and multinational. Therefore, the tariff level that solves (3) will be negative when evaluated at (2), implying $H$ will set a higher tariff than the social planner would choose.

In Cole and Davies [2011], the terms of trade effect is not present for two reasons: (i) Pre-tariff import prices do not change due to the fixed markup over a constant wage. (ii) The quasi-linear utility pushes domestic and overseas income changes onto the numéraire good, leaving the profits from Home exporters or multinationals unaffected by Home's import tariff. In the current setting, neither of these two reasons is valid: (i) Pre-tariff import prices do change due to the variable markups responding to tariff change. (ii) Profits from the Foreign exporters or multinationals do depend on the relevant cutoffs, which are all affected by Home's import tariff level. To understand the terms of trade effect in the current setup, one can easily obtain the following conditions between average prices and their corresponding cutoffs:

$$\bar{p}^H = \frac{2k+1}{2k+2} c_D^H, \quad \bar{p}^F = \frac{2k+1}{2k+2} c_D^F$$

Also notice the $IM$ and $CS$ in equation (2) and (3) are all functions of $c_D^H$ and $c_D^F$, so we can rewrite the welfare function and the first-order conditions in terms of $\bar{p}^H$ and $\bar{p}^F$. Based on Proposition 4 in Ding [2020a], it is easy to verify that an increase in Home's import tariff generates a terms of trade gain for itself at the cost of Foreign's terms of trade deterioration. It is evident that the incentive to manipulate the terms of trade also results in the inefficiency of Nash tariff.

### 4.3 Nash Tariff with and without FDI

In this subsection, I compare the symmetric Nash tariff when FDI is an option with the case when it is not. Due to the quadratic quasi-linear preference and the numéraire good, there is no closed-form analytical solution for the socially optimal tariff and the Nash tariff. Therefore, I numerically compute the tariff levels based on equations (2) and (3) in Mathematica. I follow Behrens et al. [2011] in choosing the parameter values, which are listed in **Table 1**. To focus on the role of FDI, here I fix the parameters that affect the degree of firm heterogeneity ($k$ and $c_M$)[11].

For illustration purpose, I plot the computed Nash tariff level as a function of $\varphi$ and $\alpha$[12]. When the degree of firm heterogeneity is fixed, the Nash tariff levels are plotted in **Figure 2**. The yellow plane separates the space: the area above indicates no FDI activity, and the area below indicates where FDI occurs. Since $\alpha$ is chosen such that the optimal tariff when FDI occurs is greater than one, it then makes sense that the blue plane is in-between the yellow plane and the red plane.

In **Figure 3**, I plot a two-dimensional version of **Figure 2**, and contrast it with the Figure 4 in Cole and Davies [2011]. First of all, given the current parameter choice, the Nash tariff without FDI is always higher than the one with FDI. This confirms the finding in Cole and Davies [2011]. On the one hand, the gain from implementing tariff is smaller due to the tariff-jumping multinationals. On the other hand, based on Corollary 1 in Ding [2020a], the presence of FDI causes the domestic cutoff to have a bigger response to tariff change, which affects the consumer surplus component in equation (3). These two channels collectively result in the lower Nash tariff level when FDI is present.

Second, $\varphi$ is in a similar position as the fixed cost parameter ($\lambda$) in Cole and Davies [2011], but they have notable differences. $\varphi$ does not affect the Nash tariff without FDI because it does not have any impact on exporters when FDI is not an option. The fixed cost parameter affects the Nash tariff level, regardless of the presence of FDI. The reason is that both the fixed costs of export and FDI are affected by $\lambda$. Hence, the change of $\lambda$ will have a direct impact on the tariff level. In the current

[11] I will discuss the role of heterogeneity and its interaction with FDI in Section 5. In that case, $k$ and $c_M$ will be varied.

[12] To focus on one case, here $\alpha$ is chosen to be small enough so that the optimal tariff level will be greater than 1.

framework, the change in $\varphi$ will affect the tariff level only when FDI occurs.

Third, as $\varphi$ increases, the Nash tariff with FDI increases, and gets closer to the Nash Tariff without FDI. One the one hand, if $\varphi$ approaches infinity, then the FDI cutoff will be zero, indicating Foreign firms only access Home country through exports. Hence, the Nash tariff level returns to the Nash tariff without FDI case. On the other hand, when FDI is an option, Nash tariff level increases in $\varphi$. This is similar to the results in Cole and Davies [2011] regarding $\lambda$: a higher $\varphi$ reduces the cutoff of multinational ($c_{FDI}$), causing the least productive multinationals to switch to export, increasing the tariff base, and hence increasing the incentive of imposing a higher tariff. If $\varphi$ is sufficiently high, FDI will occur in the equilibrium, confirming the corner solution finding in Cole and Davies [2011]. As we will see in the next subsection, the interaction of $\varphi$ and the degree of firm heterogeneity generates very interesting policy implications that would have been missing in the CES framework.

# 5 Role of Variable Markup

This subsection is dedicated to the discussion of variable markups. Under quadratic quasi-linear preference, firms with different marginal costs can charge different markups[13]. This feature not only enables the tariff to affect the entire distribution of markups, but also generates misallocation in the economy. As we will see toward the end of this subsection, the interaction between variable markups and FDI unveils important trade policy implications, which would be otherwise absent in the CES framework.

## 5.1 Average Markup

In the current setup, a movement in iceberg trade cost ($\tau$) does not affect the average markup[14] due to the assumption of Pareto cost distribution. However, the *ad valorem* tariff does have the ability to affect the average markup. Note that all the operating firms ($N_D^H$) serve their domestic market, and on top of that, there are Foreign exporters ($N_X^F$) and multinationals ($N_{FDI}^F$). The average

[13]There are other ways to generate variable markups, as discussed in Ding [2020b], so the implications of variable markups discussed here might not be universal. For example, if variable markups is generated through firms engaging in Bertrand competition, then the implication will be certainly different from what we obtain here.

[14]See Melitz and Ottaviano [2008] Section 3.2.

markup of all the firms in country $H$ can be expressed as the following:

$$\bar{m}^H = \frac{1}{N_D^H + N_X^F + N_{FDI}^F} \left[ N_D^H \int_0^{c_D^H} m_D^H(c) \frac{dG(c)}{G(c_D^H)} \right. \\ \left. + N_X^F \int_{c_{FDI}^F}^{c_X^F} m_X^F(c) \frac{dG(c)}{G(c_X^F)} + N_{FDI}^F \int_0^{c_{FDI}^F} m_{FDI}^F(c) \frac{dG(c)}{G(c_{FDI}^F)} \right] \tag{4}$$

To simplify the analysis, here I focus on the symmetric case. This is similar to the bilateral liberalization studied in Section 4.1[15] in Melitz and Ottaviano [2008]. After imposing symmetry, the average markup can be rewritten as follow (for detailed derivation, see Appendix C.6):

$$\bar{m} = \underbrace{\frac{1}{1+(t\tau)^{-k}} \times \frac{2k-1}{2k-2}}_{\text{weighted expected markup in domestic}} + \underbrace{\underbrace{\frac{(t\tau)^{-k} - \xi^k}{1+(t\tau)^{-k}}}_{\text{share of Foreign exporters}} \times \underbrace{t\left\{\frac{1}{2}\left[1-(t\tau\xi)^k\right] + \frac{k}{2k-2}\left[1-(t\tau\xi)^{k-1}\right]\right\}}_{\text{expected markup of Foreign exporter}}}_{\text{weighted expected markup of Foreign exporters}}$$

$$+ \underbrace{\underbrace{\frac{\xi^k}{1+(t\tau)^{-k}}}_{\text{share of foreign FDI}} \times \underbrace{\left(\frac{k}{2k-2}\frac{1}{\varphi\xi} + \frac{1}{2}\right)}_{\text{expected markup of Foreign FDI}}}_{\text{weighted expected markup of Foreign FDI}}$$

Based on this expression, I obtain the following proposition regarding the impact of a tariff change on the average markup:

**Proposition 5.** *If the level of protection is high, the increase of tariff-jumping Foreign multinational firms, which creates downward pressure on average markup, can dominate the decrease of Foreign exporter firms, which creates upward pressure on average markup. The average markup in the economy decreases as protection level increases. Therefore, protectionist trade policy can end up reducing Home market's average markup.*

**Proof.** See Appendix C.6. □

As $t$ increases, the weighted expected markup from domestic firms (the first term) increases. This relation is due to the fact that protection reduces the degree of competition and makes it easier

[15]Note, different from the long-run results established in Ding [2020a], bilateral reduction in tariff delivers the same results as in the short-run case: liberalization increases competition and decreases the domestic cutoff level, making it harder for a firm to survive.

for domestic firms to survive. As a result, the expected markup will increase. The weighted expected markup from Foreign exporters (the second term) will decrease as $t$ increases. This effect is due to two channels: the decreasing share of Foreign exporters (extensive margin, based on Proposition 5 and 6 in Ding [2020a]), and the expected markup, which increases first and then decreases as $t$ increases (intensive margin, based on Proposition 2 in Ding [2020a]). The weighted expected markup of Foreign FDI (the third term) will increase as $t$ increases. This relation also comes from two channels, the increasing share of Foreign FDI (extensive margin, based on Proposition 6 in Ding [2020a]) and the increasing expected markup (intensive margin, based on Proposition 2 in Ding [2020a]). The first and third term will dominate the second term at the beginning, but as $t$ increases, the second term will eventually dominate the other two terms, dragging down the average markup. For illustration purpose, the average markup without FDI and the average markup with FDI are plotted in **Figure 4**.

Edmond et al. [2015] suggest that under certain conditions, a reduction in trade barriers (iceberg-type trade costs) can lead to lower domestic markups (as Home producers lose their market share). Combined with higher markups on imported goods (as Foreign producers gain market share), the overall markup dispersion increases and the misallocation in the economy becomes worse. In this case, the pro-competitive gains from trade would be negative. In the current framework, a similar result is found when FDI is an option: the average markup can go up when the tariff level reduces. As the tariff level drops, although the number of imported varieties increases (hence exerting a downward pressure on average markup), the exiting of multinationals (which reduces the competition in the domestic market and exerts an upward pressure on average markup) also contributes to the increase in average markup.

## 5.2 Misallocation

Misallocation is a byproduct of all the inefficiencies discussed in Section 2. To gain more insight on the role of variable markups, in this subsection, I follow the approaches developed in Arkolakis et al. [2018] and Hsieh and Klenow [2009] to explore the misallocation in the current economy[16].

[16] These are the two most prominent papers on misallocation.

According to Arkolakis et al. [2018], variable markups can create a new source of gain or loss from trade liberalization, depending on whether low-cost firms, which charge high markups and under-supply their varieties, end up growing in size. According to their Appendix A.4, the effect of trade liberalization on the welfare of country $j$ depends on two things: (i) the sign of the covariance of the markup, charged by a firm in country $j$ that produces the variety for market $i$, and (ii) a change in its labor share that is needed to produce the variety for that market. These two forces can be expressed through the following term:

$$\mathrm{cov}\left(m^i(\omega), \frac{dl^i(\omega)}{L^j}\right) \tag{5}$$

where $l^i(\omega)$ is the total employment associated with a production of variety $\omega$ in country $j$ for sales in country $i$. In other words, if this covariance is positive, then trade liberalization has an additional positive effect on welfare in country $j$ through a reduction in misallocation. In their setup, without considering the choice of FDI, equation (5) becomes:

$$\begin{aligned}\mathrm{cov}\left(m^i(\omega), \frac{dl^i(\omega)}{L^j}\right) &= N_D^H \int_0^{c_D^H} \frac{p_D^H(c)}{c} \frac{d\left[c q_D^H(c)\right]}{L^H} \frac{dG(c)}{G\left(c_D^H\right)} \\ &+ N_X^H \int_0^{c_X^H} \frac{p_X^H(c)}{\tau^F c} \frac{d\left[c\tau^F q_X^H(c)\right]}{L^H} \frac{dG(c)}{G\left(c_X^H\right)}\end{aligned}$$

It is important to notice that this covariance is at the firm-level. Therefore, it relates to not only firm's domestic production decision but also its export decision. In their setting, this covariance is negative, so the presence of variable markups reduces the welfare gain from trade. This negative effect is present because a decrease in trade costs makes exporting firms relatively more productive, leading to changes in markups. When demand is log-concave, as in Krugman [1979], higher markups imply incomplete pass-through of changes in marginal costs to prices, lowering the welfare gains from trade.

In the current setup, the covariance term is:

$$\begin{aligned}\operatorname{cov}\left(m^{i}(\omega), \frac{dl^{i}(\omega)}{L^{j}}\right) &= N_{D}^{H}\int_{0}^{c_{D}^{H}} \frac{p_{D}^{H}(c)}{c}\frac{d\left[cq_{D}^{H}(c)\right]}{L^{H}}\frac{dG(c)}{G\left(c_{D}^{H}\right)} \\ &+ N_{X}^{H}\int_{c_{FDI}^{H}}^{c_{X}^{H}} \frac{p_{X}^{H}(c)}{\tau^{F}c}\frac{d\left[c\tau^{F}q_{X}^{H}(c)\right]}{L^{H}}\frac{dG(c)}{G\left(c_{X}^{H}\right)} \\ &+ N_{FDI}^{H}\int_{0}^{c_{FDI}^{H}} \frac{p_{FDI}^{H}(c)}{\varphi^{F}c}\frac{d\left[c\varphi^{F}q_{FDI}^{H}(c)\right]}{L^{H}}\frac{dG(c)}{G\left(c_{FDI}^{H}\right)}\end{aligned}$$

There are several differences compared to Arkolakis et al. [2018]. First, trade liberalization takes the form of a tariff reduction in this economy. Second, the covariance has an additional item due to the choice of FDI, which means now the welfare implication of a change in tariff also depends on firm's FDI activity. Third, as discussed in their June 2012 working paper[17], with quadratic quasi-linear preference, the change in welfare depends on the substitutability between the homogeneous good and the differentiated goods. In the current framework, the substitutability is also affected by FDI. In Appendix C.7, I analytically derive the covariance term under symmetry:

$$\begin{aligned}\operatorname{cov}\left(m^{i}(\omega), \frac{dl^{i}(\omega)}{L^{j}}\right) &= \frac{\left(\alpha - c_{D}^{M}\right)dc_{D}^{M}}{2\eta\left(1+t^{-k}\tau^{-k}\right)}\left\{2k+1+\left[(t\tau)^{-k}-\xi^{k}\right]\right. \\ &\quad \left.\times\left[2k+1-k\left(1-t\tau\xi\right)(t\tau\xi)^{k}-(t\tau\xi)^{k}\right]+\xi^{k}\left(k+k\varphi+1\right)\right\}\end{aligned}$$

And I show that the covariance term is *positive*, indicating a reduction in misallocation through protection. As discussed earlier, when $c_{D}^{M}$ is sufficiently high, an increase in tariff is welfare-improving. The presence of variable markup and FDI results in a positive covariance term between the firm-level markup and change in firm-level employment share. This effect means the welfare gain from protection is even larger due to the reduction in misallocation. Intuitively, an increase in tariff will decrease the relative demand for high-cost varieties, and labor will be reallocated toward the low-cost varieties, which include those produced by Foreign FDI firms. Therefore, misallocation is reduced since the market becomes more concentrated, generating a positive correlation between markups and the labor share, and hence increasing the gains from the change in tariff.

Alternatively, we can also follow the way that Hsieh and Klenow [2009] introduce misallocation.

[17] `http://www.econ.uzh.ch/dam/jcr:00000000-0db7-f8ad-0000-00005b2a6145/Arkolakis_Costas_The_Elusive_Pro_Competitive_Effects_of_Trade.pdf`.

For illustration purpose, here I focus on the exporters. A Home exporter with marginal cost $c$ has the following corresponding $TFPR$:

$$TFPR_X^{HO}(c) \equiv \frac{p_X^{HO}(c)}{c} = \frac{\tau^F c}{c} = \tau^F$$
$$TFPR_X^{HM}(c) \equiv \frac{p_X^{HM}(c)}{c} = \frac{c_D^{FM}/c + \tau^F t^F}{2}$$

In the planner's economy, $TFPR_X^{HO}$ is the same for all the exporters, and there is no misallocation in this case. However, in the market outcome, an exporter with lower marginal cost will have a bigger $TFPR$, implying that the low cost firms are allocated with too little labor. This is consistent with the conclusion in Proposition 1. Hence, misallocation also exists in the market outcome according to the definition in Hsieh and Klenow [2009]. Based on Proposition 4 in Ding [2020a], an increase in $t^H$ will increase $TFPR_X^{HM}$, exacerbating the misallocation among Home's exporters.

**Figure 5** is an attempt to present all the misallocation concepts on the graph. Both panels are demonstrating the responses of cutoffs to an increase in Home's import tariff. The differences between socially optimum outcome and market outcome in Proposition 1 are revealed in the cutoff levels. The positive covariance term in equation (5) can be seen from Panel B: following the increase of $t^H$, $c_{FDI}^{FM}$ and $c_D^{HM}$ both move toward their socially optimum level, indicating that Home's labor are reallocated toward the more productive firms, and hence misallocation is reduced. Finally, the increase of $TFPR_X^{HM}$ can be seen from the widening gap between $c_X^{HO}$ and $c_X^{HM}$.

## 5.3 Interaction of Variable Markup and FDI

In this subsection, I focus on the interaction of variable markups and FDI through the lens of firm heterogeneity. For the same reason in Section 4.3, the analysis in this subsection is based on numerical computation. All the relevant parameter values are chosen from **Table 1**. As discussed in Section 2, the quadratic quasi-linear preference makes firm heterogeneity an additional source of inefficiency in this economy. To see how the interaction of variable markup and FDI affects the Nash tariff level, I focus on the two parameter values that govern the degree of firm heterogeneity: $k$ and $c_M$. It is also helpful to keep in mind that a country cares for tariff revenue and its consumer surplus, as can be seen from the following objective function:

$$\max_{\{t^H\}} \mathbb{U}^H = I^H + CS^H = 1 + TR^H + CS^H$$

The Nash tariff level is implied by the following first order condition:

$$\frac{\partial TR^H}{\partial t^H} + \frac{\partial CS^H}{\partial t^H} = 0 \tag{6}$$

#### 5.3.1 The Role of $k$

The impact of $k$ can be seen from **Figure 6** and **Figure 7**. $k$ governs the shape of Pareto distribution. An increase in $k$ means the distribution of firms is skewed toward less productive firms, reducing the degree of firm heterogeneity.

With the current parameter values, the Nash tariff without FDI decreases in $k$, which means it increases in the degree of firm heterogeneity. As $k$ increases, two things are happening at the same time: (i) The domestic cutoff level is increasing. (ii) The probability of getting the low marginal cost draws are shrinking. The first channel will affect all cutoff levels proportionally, without changing the relative distribution of firms. The second channel changes the relative distribution of firms, and this channel is particularly important when multinational firms are present.

When FDI is an option, the impact of $k$ on Nash tariff level depends on the size of $\varphi$, the parameter that measures the efficiency loss of FDI. First, an increase in $k$ increases domestic cutoff and lowers the consumer surplus, so the second term in equation (6) is positive. This implies that the first term in equation (6) must be negative. When $\varphi$ is small, the Nash tariff increases in $k$, i.e., decreases in firm heterogeneity. Intuitively, when $\varphi$ is small, many firms access the Foreign market through FDI, so the tariff base is relatively small. In this case, an increase in $k$ lowers the average probability of getting a low marginal cost draw. To maintain the first order condition in equation (6), the country needs to charge a higher tariff. This effect can be seen from the bottom panel in **Figure 7**: when $\varphi$ is small, the blue plane (bigger $k$) is above the green plane (smaller $k$).

When $\varphi$ is large, however, the Nash tariff decreases in $k$, i.e., increases in firm heterogeneity. This is because when $\varphi$ is large, many firms will choose to access the Foreign market through export, so the tariff base is relatively big. In this case, although an increase in $k$ lowers the average probability of getting a low marginal cost draw, the sizable tariff base is big enough to maintain the first order condition in equation (6). Therefore, the Nash tariff level is lower. As shown in the bottom panel in **Figure 7**: when $\varphi$ is big, the blue plane (bigger $k$) is below the green plane (smaller $k$).

#### 5.3.2 The Role of $c_M$

The impact of $k$ can be seen from **Figure 8** and **Figure 9**. $c_M$ represents the lower bound of the marginal cost distribution. An increase in $c_M$ expands the region that marginal cost can be drawn from, raising up the degree of firm heterogeneity.

With the current parameter values, when FDI is an option, the Nash tariff decreases in $c_M$, which means it decreases in the degree of firm heterogeneity. To understand the intuition, recall that an increase in $c_M$ expands the lower bound of cost draws, which eventually results in higher cutoff levels[18]. The impact is uniform to all the firms, without changing the relative distributions of firms. This is why in **Figure 9**, the green plane (smaller $c_M$) is always above the blue plane (bigger $c_M$), i.e., the relative position of Nash tariff levels under different $c_M$ remains the same. There is no interaction between variable markup (induced by $c_M$) and FDI (measured by the size of $\varphi$).

In summary, both $k$ and $c_M$ affect the degree of firm heterogeneity in the economy. While both of them affect the equilibrium cutoff levels, $k$ also alters the relative distribution of firms with different marginal costs. When FDI is an option, the freeness of doing FDI interacts with $k$, generating a novel implication for trade policy. If the degree of firm heterogeneity is big (smaller $k$), reducing the FDI barrier (smaller $\varphi$) can effectively lower the Nash tariff level. If the degree of firm heterogeneity is small (bigger $k$), increasing the FDI barrier (bigger $\varphi$) can effectively lower the Nash tariff level.

## 6 Concluding Remarks

This paper studies the welfare implication of tariff and optimal tariffs in the presence of variable markups and FDI. The conclusions can be broadly summarized as follows. First, I find that quadratic quasi-linear preference generates multiple externalities in this economy, causing market outcome to differ from the socially optimum outcome systematically. Permitting FDI lowers the domestic cutoff

[18]This can be easily verified through the closed-form solution of $c_D$, $c_X$, and $c_{FDI}$ in Ding [2020a]

levels and reduces the misallocation in the economy. Second, I find that free trade is not always socially optimal. If the domestic cutoff is sufficiently high, an additional firm entry can improve social welfare. In this case, a positive import tariff is welfare-improving because it encourages firm entry. Third, I find that the interaction of variable markup and FDI generates novel trade policy insights that are absent if consumers are under CES preference.

Given these results, there are several interesting questions to ask. First of all, do the trade policy results still hold under an alternative demand or supply structure that generates variable markups? I suspect that the alternative demand structures will produce similar results, but different supply-side structures may generate different outcomes. Second, the presence of numéraire good is a blessing and a curse. It would be interesting to drop the numéraire good and endogenize the wage as in Arkolakis [2008]. It is possible to obtain a closed-form solution of the Nash tariff, as shown by Demidova [2017], and investigate the trade policy implication regarding the labor market in the presence of FDI. Lastly, the trade policy implications here primarily focus on the import tariff. It would be interesting and relevant to study other forms of trade policy, such as export subsidy or corporate taxes.

This paper provides evidence that the interaction of variable markups and FDI generates interesting trade policy implications. The steady-state analysis, however, might produce very different tariff levels than the actual tariff levels observed in the data. In the light of Larch and Lechthaler [2013], long-run and short-run effects of tariffs may run in opposite directions, implying that an exclusive focus on the steady-state could lead to biased policy conclusions. Carefully disentangling the dynamic effects of tariffs is undoubtedly a fruitful area for future research in the era of globalization .

# References


**Arkolakis, Costas**, "A Generalized Solution of the Monopolistic Competition Model with Heterogeneous Firms and a Linear Demand (Melitz-Ottaviano)," *Working paper, Department of Economics, Yale University*, 2008.

— **, Arnaud Costinot, Dave Donaldson, and Andrés Rodríguez-Clare**, "The elusive pro-competitive effects of trade," *The Review of Economic Studies*, 2018.

**Asker, John, Allan Collard-Wexler, and Jan De Loecker**, "Market Power and Production (Mis) Allocation. A Study of the World Oil Market," *Mimeo, Princeton University*, 2017.

**Bagwell, Kyle and Seung Hoon Lee**, "Trade policy under monopolistic competition with firm selection," *Mimeogr., Stanford University*, 2015.

**Behrens, Kristian, Giordano Mion, and Gianmarco IP Ottaviano**, "Economic integration and industry reallocations: some theory with numbers," *International Handbook on the Economics of Integration: Competition, Spatial Location of Economic Activity and Financial Issues*, 2011, *2*, 169.

— **, — , Yasusada Murata, and Jens Suedekum**, "Quantifying the Gap between Equilibrium and Optimum Under Monopolistic Competition," *CEPR DP 11642*, 2018.

**Bhagwati, Jagdish N**, *The Generalized Theory of Distortions and Welfare*, Cambridge, MIT, 1969.

**CEA**, "Benefits of Competition and Indicators of Market Power," *Council of Economic Advisers Issue Brief*, 2016, p. April 2016.

**Cole, Matthew T and Ronald B Davies**, "Strategic tariffs, tariff jumping, and heterogeneous firms," *European Economic Review*, 2011, *55* (4), 480–496.

**Demidova, Svetlana**, "Trade policies, firm heterogeneity, and variable markups," *Journal of International Economics*, 2017, *108*, 260–273.

**Dhingra, Swati and John Morrow**, "Monopolistic competition and optimum product diversity under firm heterogeneity," *Journal of Political Economy*, 2019, *127* (1), 196–232.

**Ding, Ziran**, "Firm Heterogeneity, Variable Markups, and FDI," *Doctoral Dissertation*, 2020.

— , "Trade Policy with Firm Heterogeneity, Variable Markups, and FDI: A Synopsis of the Literature," *Doctoral Dissertation*, 2020.

**Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu**, "Competition, markups, and the gains from international trade," *American Economic Review*, 2015, *105* (10), 3183–3221.

**Hsieh, Chang-Tai and Peter J Klenow**, "Misallocation and manufacturing TFP in China and India," *The Quarterly Journal of Economics*, 2009, *124* (4), 1403–1448.

**Krugman, Paul R**, "Increasing returns, monopolistic competition, and international trade," *Journal of international Economics*, 1979, *9* (4), 469–479.

**Larch, Mario and Wolfgang Lechthaler**, "Whom to send to Doha? The shortsighted ones!," *Review of Economic Dynamics*, 2013, *16* (4), 634 – 649.

**Loecker, Jan De and Jan Eeckhout**, "The rise of market power and the macroeconomic implications," Technical

Report, National Bureau of Economic Research 2017.

**Mankiw, N Gregory and Michael D Whinston**, "Free entry and social inefficiency," *The RAND Journal of Economics*, 1986, pp. 48–58.

**Melitz, Marc and Giancarlo Ottaviano**, "Market Size, Trade, and Productivity," *Review of Economic Studies*, 2008, *75* (295-316).

**Nocco, Antonella, Gianmarco IP Ottaviano, and Matteo Salto**, "Monopolistic competition and optimum product selection," *American Economic Review*, 2014, *104* (5), 304–09.

**Ottaviano, Gianmarco, Takatoshi Tabuchi, and Jacques-François Thisse**, "Agglomeration and trade revisited," *International Economic Review*, 2002, pp. 409–435.

**Spence, Michael**, "Product selection, fixed costs, and monopolistic competition," *The Review of economic studies*, 1976, *43* (2), 217–235.

**Weinberger, Ariel**, "Markups and misallocation with trade and heterogeneous firms," *FRB Dallas Globalization and Monetary Policy Institute Working Paper No. 251*, 2015.

# A Tables

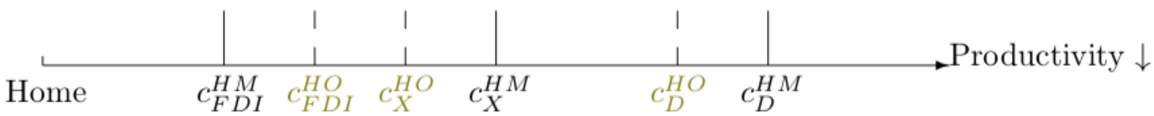
**Table 1:** Paramaterization

| | | |
|---|---|---|
| $\alpha$ | 12 | Relative preferences toward the differentiated varieties |
| $\eta$ | 0.1 | Substitutability among the varieties |
| $c_M$ | 5 | Upper bound of marginal cost draw in Pareto distribution |
| $\gamma$ | 0.6 | Degree of love for variety |
| $\varphi$ | 1.9 | Iceberg-type efficiency loss of FDI |
| $\tau$ | 1.1 | Iceberg-type transportation cost |
| $f_E$ | 0.1 | Fixed cost of entry |

**Note:** These parameters are the baseline values in Behrens et al. [2011]. All of the computations performed in this paper are based on this table. For the comparative statics in Section 4 and 5, some parameters are varied around the benchmark value listed above.

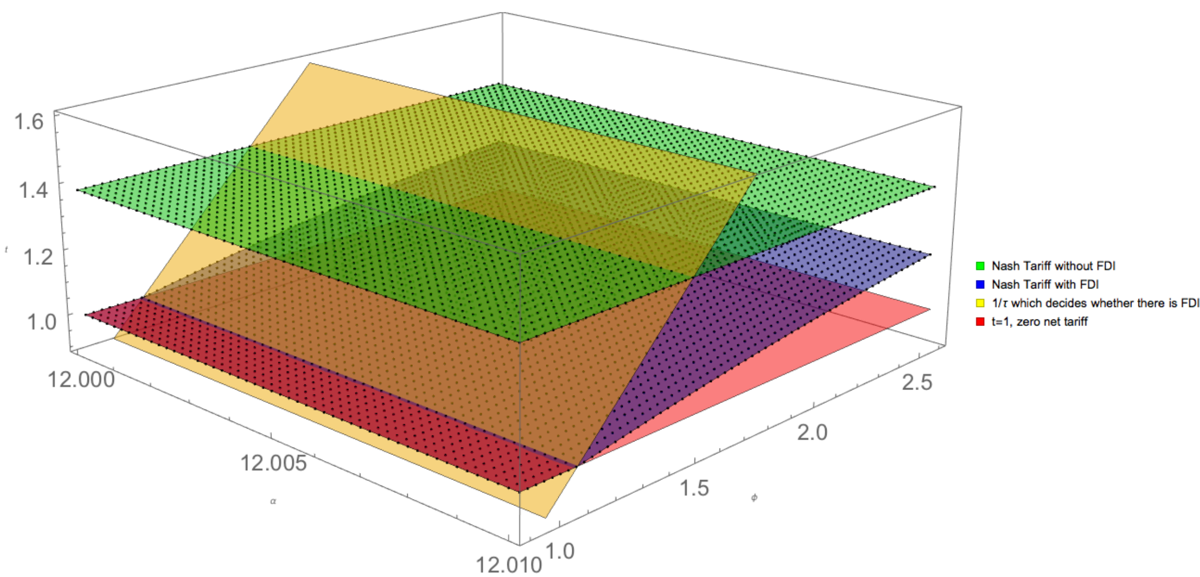
# B Figures

**Figure 1:** Productivity Cutoffs Comparison between Market Outcome and Socially Optimum Outcome



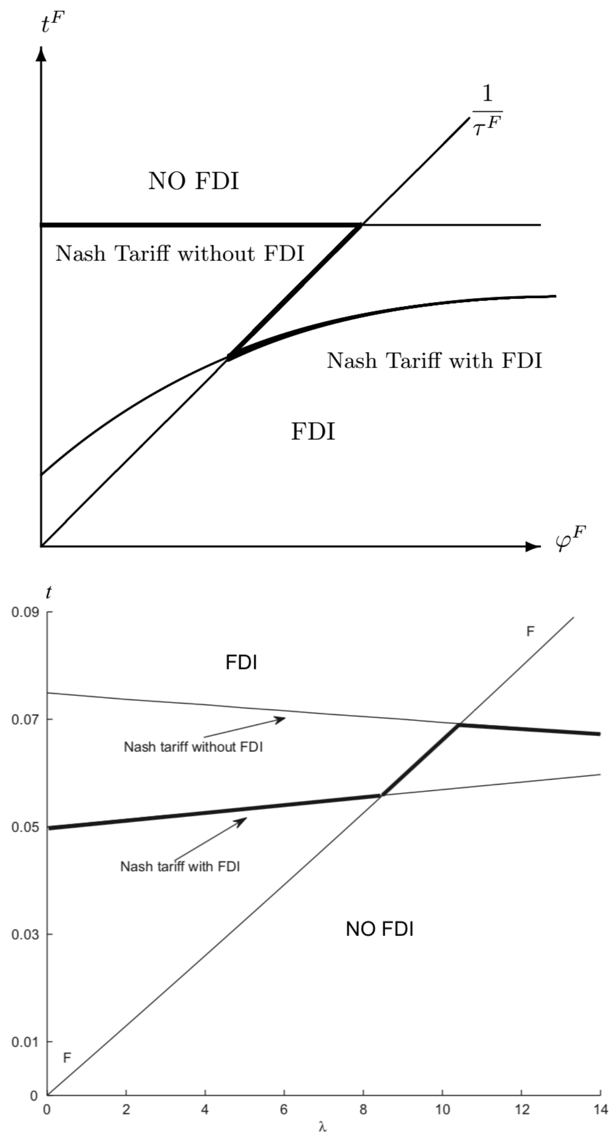

**Note:** This graph shows the marginal cost distributions of all the domestic firms in Home country. The letter $M$ stands for the market outcome, which are marked with the black color. The letter $O$ stands for the socially optimum outcome, which are marked with the olive color. Note, the indexes here are different from those in **Figure 5**.

**Figure 2:** Three-dimensional Nash Tariff with and without FDI



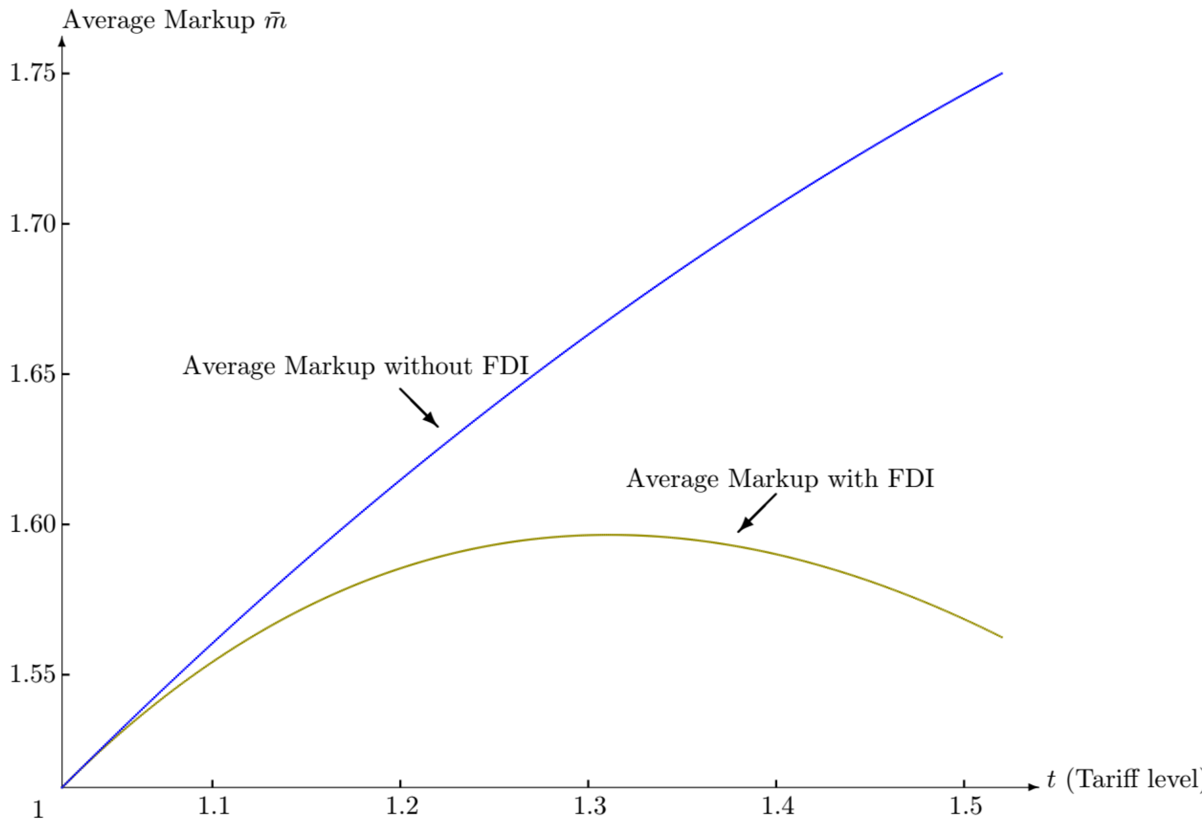

**Note:** The graph is computed based on the parameter values in **Table 1**. The green plane plots the Nash tariff without FDI and the blue plane plots the Nash tariff with FDI. The yellow plane separates the space into FDI region and no FDI region. The red plane indicates zero net tariff.

**Figure 3:** Two-dimensional Nash Tariff with and without FDI



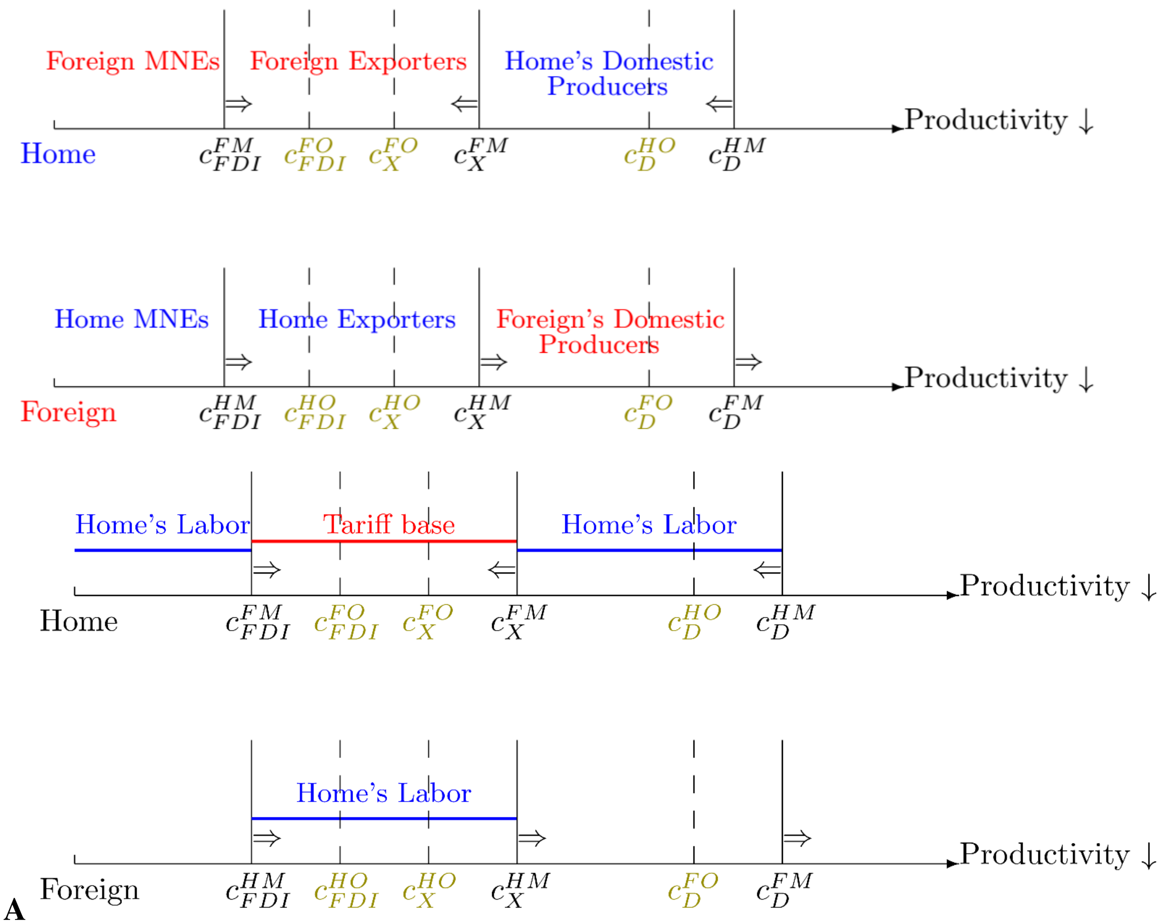

**Note:** The top panel is a two-dimensional graph of **Figure 2**. The vertical axis indicates the tariff level, the horizontal axis indicates the easiness of doing FDI, and the $1/\tau^F$ separates the plane into No FDI region and FDI region. The bottom panel is taken from Cole and Davies [2011]. The vertical axis stands for net tariff, the horizontal axis represents the exogeneous component of the fixed cost for all modes of production, and the FF line represents the combination of $\tau$ and $\lambda$ that induces FDI. The bold lines indicate the path of corner solution.

**Figure 4:** Average Markup with and without FDI




**Note:** All the other parameter values are taken from **Table 1**. The blue curve indicates the weighted average markup without FDI, whereas the green curve indicates the weighted average markup with FDI.

**Figure 5:** Productivity Cutoffs Comparison between Market Outcome and Socially Optimum Outcome when Home's Tariff Increases

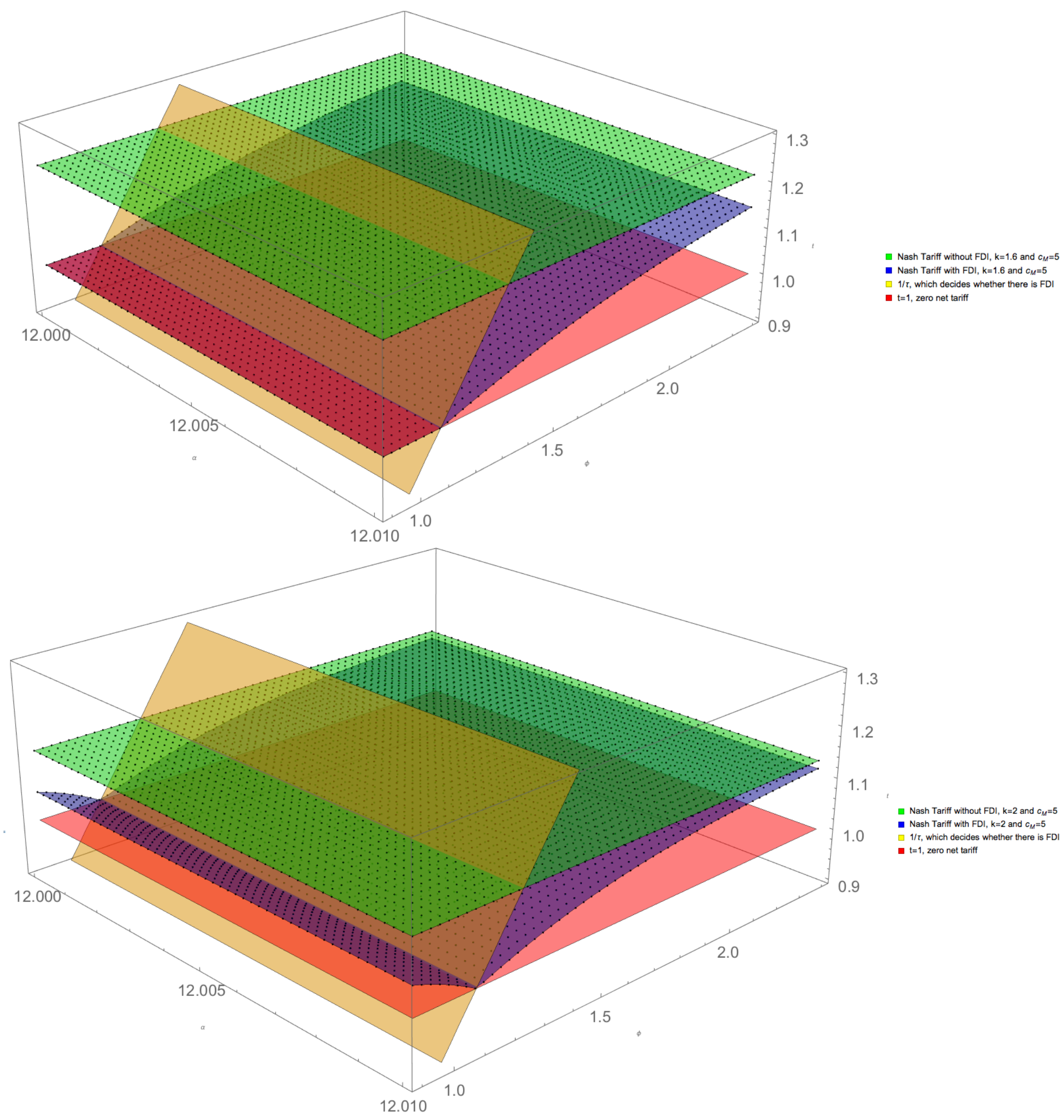



**Panel A**

**Panel B**

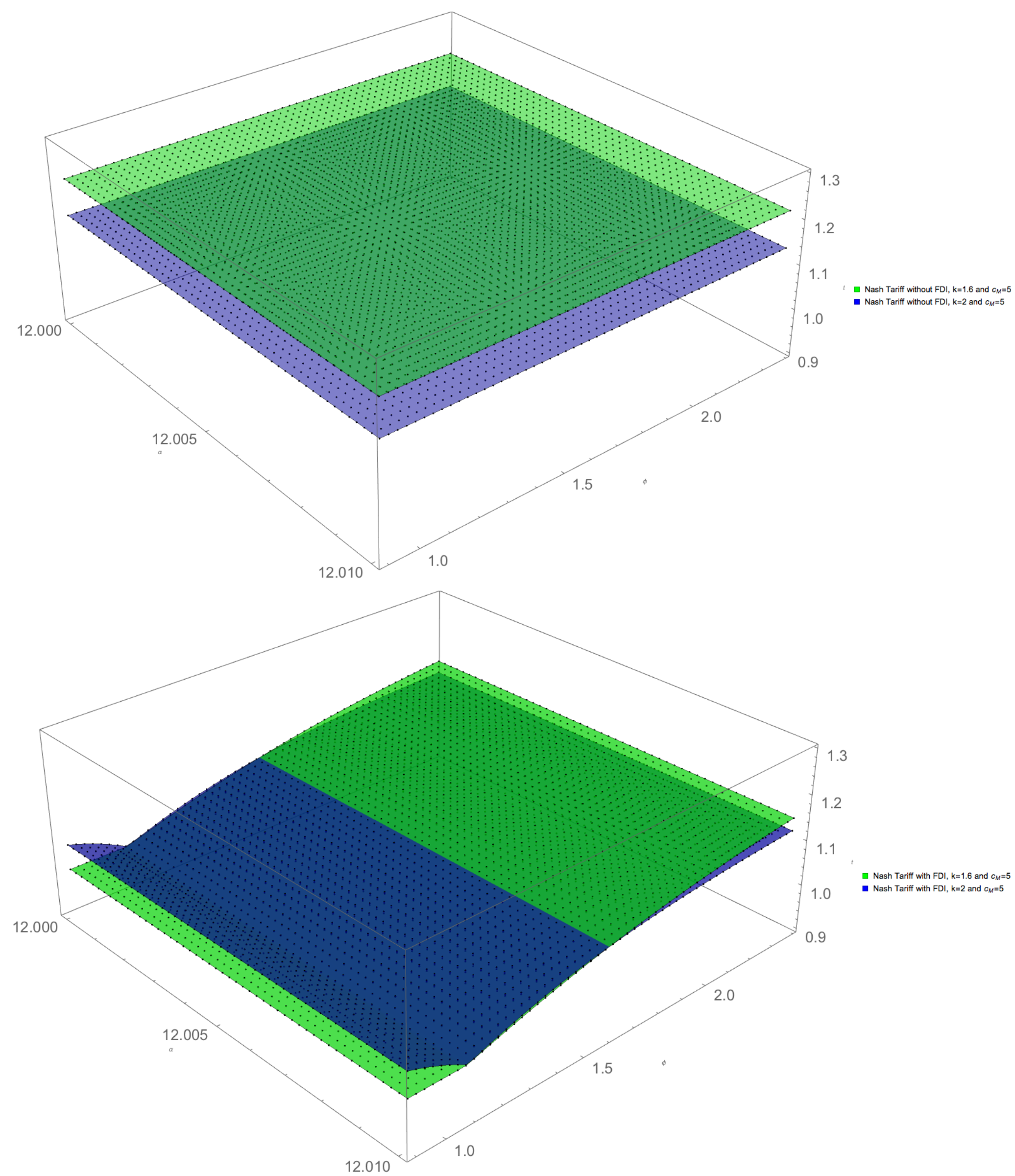
**Note:** Panel A and Panel B are the same plots with different notations. Each axis indicates the varieties that are available within that country. For example, the first axis indicates Home country has varieties from Foreign MNEs, Foreign exporters, and all the Home firms that serve their domestic market. Panel A focuses on the composition of firms, whereas Panel B focuses on the utilization of labor.

**Figure 6:** The Impact of $k$ on the Nash Tariff: Part I



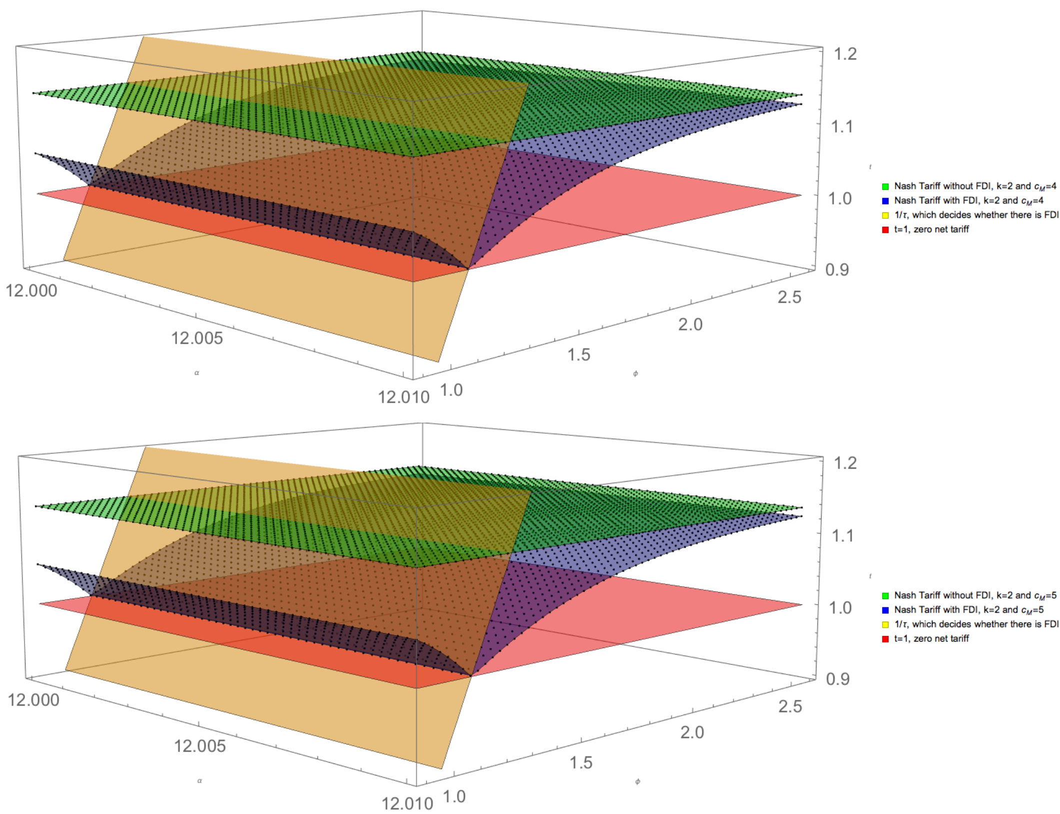

**Note:** The top panel plots the Nash tariff without FDI and the Nash tariff with FDI when $k = 1.6$. The bottom panel plots the Nash tariff without FDI and the Nash tariff with FDI when $k = 2$. The yellow plane separates the space into FDI region and no FDI region. The red plane indicates zero net tariff.

**Figure 7:** The Impact of $k$ on the Nash Tariff: Part II



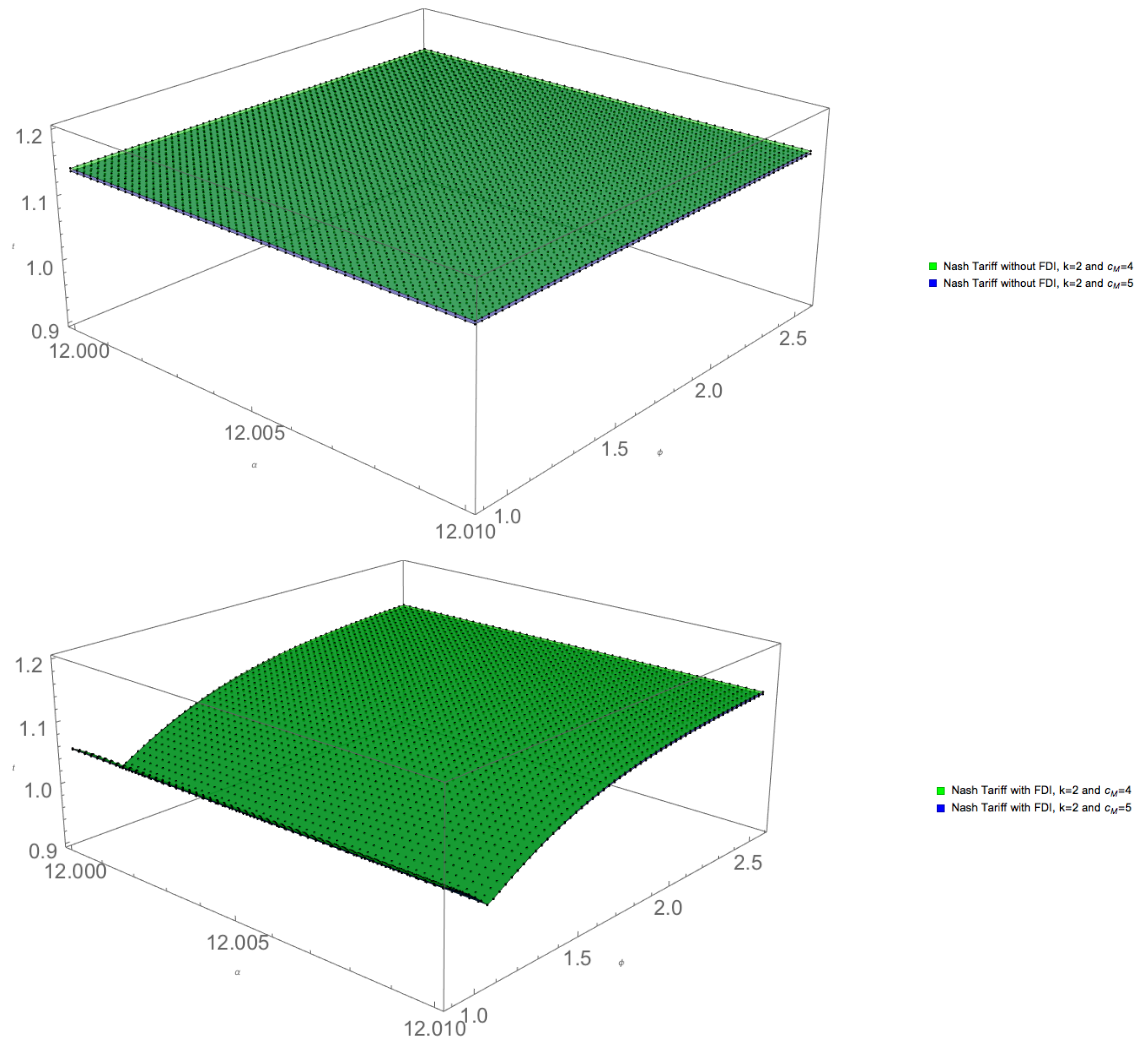

**Note:** These graphs are extracted from **Figure 6**. The top panel plots the Nash tariff level without FDI. The green plane, which has a lower $k$ value, stays entirely above the blue plane. The bottom panel plots the Nash tariff level with FDI. In the region where FDI occurs, the green plane, which has a lower $k$ value, stays above the blue plane when $\varphi$ is big, but stays below the blue plane when $\varphi$ is small. This graph shows the interaction of FDI and $k$ does affect the Nash tariff level.

**Figure 8:** The Impact of $c_M$ on the Nash Tariff: Part I




**Note:** The top panel plots the Nash tariff without FDI and the Nash tariff with FDI when $c_M = 4$. The bottom panel plots the Nash tariff without FDI and the Nash tariff with FDI when $c_M = 5$. The yellow plane separates the space into FDI region and no FDI region. The red plane indicates zero net tariff.

**Figure 9:** The Impact of $c_M$ on the Nash Tariff: Part II




**Note:** These graphs are extracted from **Figure 8**. The top panel plots the Nash tariff level without FDI. The green plane, which has a lower $c_M$ value, stays entirely above the blue plane. The bottom panel plots the Nash tariff level with FDI. In the region where FDI occurs, the green plane, which has a lower $c_M$ value, also stays entirely above the blue plane. This graph shows the interaction of FDI and $c_M$ does not change the relative position of Nash tariff.

# C Proofs

## C.1 Proof of Proposition 1

The social planner solves the following problem

$$\max_{\{N_E^H, q_0^H, q_i^H, N_E^F, q_0^F, q_i^F\}} \mathbb{W} \equiv \mathbb{U}_H + \mathbb{U}_F$$

$$\begin{aligned} \text{s.t. } & q_0^H + q_0^F + f\left(N_E^H + N_E^F\right) + N_E^H \int_0^{c_M} \left[c q_D^H(c) + \tau^F c q_X^H(c) + \varphi^F c q_{FDI}^H(c)\right] dG(c) \\ & + N_E^F \int_0^{c_M} \left[c q_D^F(c) + \tau^H c q_X^F(c) + \varphi^H c q_{FDI}^F(c)\right] dG(c) = 2 + \bar{q}_0^H + \bar{q}_0^F \end{aligned}$$

Notice, $\mathbb{W} \equiv \mathbb{U}^H + \mathbb{U}^F$ and since labor has been normalized to 1, $\mathbb{U}^H$ is defined as follow

$$\begin{aligned} \mathbb{U}^H \equiv\ & q_0 + \alpha N_E^H \left\{\int \left[q_D^H(c) + q_X^H(c) + q_{FDI}^H(c)\right] dG(c)\right\} \\ & - \frac{\gamma}{2}\left\{N_E^H \int \left(q_D^H(c)\right)^2 dG(c) + N_E^F \int \left[\left(q_X^F(c)\right)^2 + \left(q_{FDI}^F(c)\right)^2\right] dG(c)\right\} \\ & - \frac{\eta}{2}\left\{N_E^H \int q_D^H(c)\, dG(c) + N_E^F \int \left[q_X^F(c) + q_{FDI}^F(c)\right] dG(c)\right\} \end{aligned}$$

The first order conditions with respect to $q_D, q_X, q_{FDI}$ deliver the following results for the Home country:

$$\begin{aligned} q_D^H(c) &= \frac{c_D^{HO} - c}{\gamma}, c_D^{HO} = \alpha - \eta Q^{HO} \\ q_X^H(c) &= \frac{c_X^{HO} - c}{\gamma/\tau^F}, c_X^{HO} = \frac{\alpha - \eta Q^{FO}}{\tau^F} \\ q_{FDI}^H(c) &= \frac{c_{FDI}^{HO} - c}{\gamma/\varphi^F}, c_{FDI}^{HO} = \frac{\alpha - \eta Q^{FO}}{\varphi^F} \end{aligned}$$

The first order condition with respect to $N_E$ delivers the following results

$$Q^{HO} = \frac{N^{HO} + 2N^{FO}}{\gamma + \eta\left(N^{HO} + 2N^{FO}\right)}\left(\alpha - \frac{k}{k+1}\left[\frac{N^{HO} + \left(\tau^H/\varphi^H\right)^{k+1} N^{FO}}{N^{HO} + 2N^{FO}}\right] c_D^{HO}\right)$$

Combine these with the corresponding results for the Foreign country, it's straightforward to obtain Home's domestic cutoff level under the planner's problem

$$c_D^{HO} = \left[\gamma(k+1)(k+2) f c_M^k \frac{1 - O_F}{1 - O_F O_H}\right]^{\frac{1}{k+2}}$$

where $O_F \equiv (\varphi^F)^{-k} + \frac{(k+1)(k+2)}{2}\left[(\tau^F)^{-k} - (\varphi^F)^{-k}\right] - k(k+2)\tau^F\left[(\tau^F)^{-(k+1)} - (\varphi^F)^{-(k+1)}\right] + \frac{k(k+1)(\tau^F)^2}{2}\left[(\tau^F)^{-(k+2)} - (\varphi^F)^{-(k+2)}\right]$. All the rest of the equilibrium variables, such as $N_E^{HO}, N^{HO}$, etc, can be expressed as a function of $c_D^{HO}$ and other parameters. Compare the domestic cutoff from equation (23) in Ding [2020a] with the socially optimum cutoff, it is straightforward to show that:

$$\left(\frac{c_D^{HM}}{c_D^{HO}}\right)^{k+2} = \frac{2}{\frac{1 - O_F}{1 - O_F O_H} / \frac{1 - \Phi_1^F - \Phi_2^F}{1 - \left(\Phi_1^F + \Phi_2^F\right)\left(\Phi_1^H + \Phi_2^H\right)}} \equiv \frac{2}{\triangle_F}$$

The term in the denominator of the above expression is defined as $\triangle_F$. For displaying purpose, I then focus on the comparison of all the market outcomes and socially optimum outcomes in the Home market. It follows from the definition of $c_X^{HM}, c_X^{HO}, c_{FDI}^{HM}, c_{FDI}^{HO}$ that the following three equations must hold:

$$c_D^{HM} - c_D^{HO} = \left[\left(\frac{2}{\triangle_F}\right)^{\frac{1}{k+2}} - 1\right] c_D^{HO}$$

$$c_X^{HM} - c_X^{HO} = \left[\left(\frac{2}{\triangle_F}\right)^{\frac{1}{k+2}} - t^F\right] \frac{c_D^{FO}}{\tau^F t^F}$$

$$c_{FDI}^{HM} - c_{FDI}^{HO} = \left[\left(\frac{2}{\triangle_F}\right)^{\frac{1}{k+2}} - \frac{1}{\varphi^F \xi^F}\right] c_D^{FO} \xi^F$$

It then can be verified that if the tariff level is not sufficiently high, $\triangle_F$ will be less than 1, hence both $\left[(2/\triangle_F)^{\frac{1}{k+2}} - 1\right]$ and $\left[(2/\triangle_F)^{\frac{1}{k+2}} - t^F\right]$ are greater than zero. Notice, $\varphi^F \xi^F < 1$, with the assumption that $\varphi^F > t^F \tau^F$, it is straightforward to show that $\left[(2/\triangle_F)^{\frac{1}{k+2}} - 1/\varphi^F \xi^F\right] < 0$. Therefore, part $(A)$ of proposition 9 is proved.

For part $(B)$, we can just follow the definition of production levels. For example, we know

$$q_D^{HM} = \frac{1}{2\gamma}(c_D^{HM} - c), q_D^{HO} = \frac{1}{\gamma}(c_D^{HO} - c)$$

It can be easily verified that $q_D^{HM} < q_D^{HO}$ if $c < \left[2 - (2/\triangle_F)^{1/(k+2)}\right] c_D^{HO}$. The production levels of $q_X$ and $q_{FDI}$ also follow directly from the comparison of market outcome and socially optimum outcome.

For part $(C)$, one can show that

$$N_H^M = \frac{2\gamma(k+1)}{\eta} \times \frac{\alpha - c_D^{HM}}{c_D^{HM}}, N_H^O = \frac{\gamma(k+1)}{\eta} \times \frac{\alpha - c_D^{HO}}{c_D^{HO}}$$

One can then show that $N_H^M > N_H^O$ if

$$c_D^{HO} < \left[\frac{2}{(2/\triangle_F)^{\frac{1}{k+2}}} - 1\right] \alpha$$

For the level of entrants, one can solve them through the following system of equations:

$$N_H^O = N_E^{HO} \left(\frac{c_D^{HO}}{c_M}\right)^k + N_E^{FO} \left(\frac{c_D^{HO}}{c_M}\right)^k (\tau^H)^{-k}$$

$$N_F^O = N_E^{FO} \left(\frac{c_D^{FO}}{c_M}\right)^k + N_E^{HO} \left(\frac{c_D^{FO}}{c_M}\right)^k (\tau^F)^{-k}$$

And obtain the socially optimum entrant level:

$$N_E^{HO} = \frac{\gamma(k+1)c_M^k}{\eta[1 - (\tau^H \tau^F)^{-k}]} \times \left[\frac{\alpha - c_D^{HO}}{\left(c_D^{HO}\right)^{k+1}} - (\tau^H)^{-k} \frac{\alpha - c_D^{FO}}{\left(c_D^{FO}\right)^{k+1}}\right]$$

Comparing it with the entrant level in the market outcome:

$$N_E^{HM} = \frac{2\,(c_M)^k\,(k+1)\,\gamma}{\eta\left(1 - \delta^H \delta^F\right)} \times \left[\frac{\alpha - c_D^{HM}}{\left(c_D^{HM}\right)^{k+1}} - \delta^H \frac{\alpha - c_D^{FM}}{\left(c_D^{FM}\right)^{k+1}}\right]$$

where $\delta^l = (t^l \tau^l)^{-k}$, for $l \in \{H, F\}$. Together with the fact that $c_D^{HO} = c_D^{FO}\left[(1-O_F)/(1-O_F O_H)\right]^{\frac{1}{k+2}}$, one can then find a similar threshold of $c_D^{HO}$ where $N_E^{HO}$ differs from $N_E^{HM}$. That completes the proof of part $(C)$. $\square$

## C.2 Proof of Proposition 2

**Proof:** Based on equation (28) in Ding [2020a], the social welfare can be rewritten as follow:

$$\mathbb{U}^H + \mathbb{U}^F = I^H + I^F + \frac{\alpha - c_D^H}{2\eta}\left(\alpha - \frac{k+1}{k+2}c_D^H\right) + \frac{\alpha - c_D^F}{2\eta}\left(\alpha - \frac{k+1}{k+2}c_D^F\right)$$

Since consumer receive income from wage (which is equal to 1) and tariff revenue, so the above equation can be rewritten as:

$$\begin{aligned}\mathbb{U}^H + \mathbb{U}^F &= 2 + \left(t^H - 1\right) IM^H + \left(t^F - 1\right) IM^F \\ &+ \frac{\alpha - c_D^H}{2\eta}\left(\alpha - \frac{k+1}{k+2}c_D^H\right) + \frac{\alpha - c_D^F}{2\eta}\left(\alpha - \frac{k+1}{k+2}c_D^F\right)\end{aligned}$$

To see the welfare implication of free trade, I evaluate the first order condition of the above expression with respect to tariff under symmetry when $t^H = t^F = 1$. Since symmetry implies $IM^F = IM^H, \partial IM^F/\partial t^H = \partial IM^H/\partial t^F, c_D^F = c_D^H, \partial c_D^F/\partial t^H = \partial c_D^H/\partial t^F$, therefore :

$$\begin{aligned}\frac{\partial\left(\mathbb{U}^H + \mathbb{U}^F\right)}{\partial t}|_{t^H=t^F=1} &= \underbrace{(t-1)\left(\frac{\partial IM^H}{\partial t^H} + \frac{\partial IM^F}{\partial t^H}\right)}_{\circledast} + IM^H \\ &+ \frac{2(k+1)c_D - (2k+3)\alpha}{2\eta(k+2)}\left(\frac{\partial c_D^H}{\partial t^H} + \frac{\partial c_D^F}{\partial t^H}\right)\end{aligned}$$

Notice, when $t^H = t^F = 1$, $\circledast = 0$. Based on equation (27) and (29) in Ding [2020a],

$$\begin{aligned}IM^H|_{t^H=1} &= N_E|_{t^H=1} \times \frac{(c_D)^{k+2}\tau^{-k}}{2\gamma(k+2)(c_M)^k}|_{t=1} \\ &= \frac{2(c_M)^k(k+1)\gamma\left(1-\tau^{-k}\right)}{\eta\left(1-\tau^{-2k}\right)}\frac{(c_D)^{k+2}\tau^{-k}}{2\gamma(k+2)(c_M)^k}|_{t=1} \\ &= \frac{\tau^{-k}(k+1)}{\eta\left(1+\tau^{-k}\right)(k+2)}(\alpha - c_D)c_D|_{t=1}\end{aligned}$$

Based on the definition of $\Phi^H$ and $\Phi^F$, it is straightforward to show

$$\Phi_1 + \Phi_2 = \tau^{-k}$$

Based on the proof of Proposition 4 in Ding [2020a], it is straightforward to show

$$\begin{aligned}
\frac{\partial c_D^H}{\partial t^H} + \frac{\partial c_D^F}{\partial t^H}|_{t=1} &= \frac{-(k+1)\, c_D^H \tau^{-2k}}{(k+2)\left(1-\tau^{-2k}\right)} + \frac{c_D^F}{k+2}\frac{-(k+1)\,\tau^{-k}}{\left(1-\tau^{-2k}\right)} \\
&= \frac{\tau^{-k}\,(k+1)}{\left(1+\tau^{-k}\right)(k+2)} c_D|_{t=1}
\end{aligned}$$

Therefore, the original first order condition can be rewritten as

$$\begin{aligned}
\frac{\partial\left(\mathbb{U}^H + \mathbb{U}^F\right)}{\partial t}|_{t^H=t^F=1} &= \frac{\tau^{-k}\,(k+1)}{\eta\left(1+\tau^{-k}\right)(k+2)} (\alpha - c_D)\, c_D|_{t=1} \\
&+ \frac{2\,(k+1)\, c_D - (2k+3)\,\alpha}{2\eta\,(k+2)} \frac{\tau^{-k}\,(k+1)}{\left(1+\tau^{-k}\right)(k+2)} c_D|_{t=1} \\
&= \frac{\tau^{-k}\,(k+1)\, c_D}{2\eta\,(k+2)^2\left(1+\tau^{-k}\right)} \left[2\,(k+2)\,(\alpha - c_D) + 2\,(k+1)\, c_D - (2k+3)\,\alpha\right] \\
&= \frac{\tau^{-k}\,(k+1)\, c_D}{2\eta\,(k+2)^2\left(1+\tau^{-k}\right)} \left(-2c_D + \alpha\right)|_{t=1}
\end{aligned}$$

Define $\tilde{c}_D \equiv c_D|_{t=1}$, then this completes the proof of proposition 2. □

## C.3 Proof of Second-Best Social Planner Problem

Based on the definition of average consumer surplus, it can be rewritten in terms of $\tilde{c}_D$:

$$\text{Avg. CS} \equiv \frac{\gamma}{2}\int_0^{\tilde{c}_D} \left(q_D\,(c)\right)^2 dG\,(c) = \frac{\left(\tilde{c}_D\right)^{k+2}}{4\gamma c_M^k\,(k+1)\,(k+2)} > 0$$

Based on equation (26) in Ding [2020a], under symmetry and $t = 1$,

$$N_E = \frac{2\gamma c_M^k\,(k+1)\,(\alpha - \tilde{c}_D)}{\eta\left(\tilde{c}_D\right)^{k+1}\left(1+\tau^{-k}\right)}$$

Then, the variety effect can be defined as the difference between consumer surplus and the sum of average surplus at each variety:

$$\begin{aligned}
\text{VE} &\equiv \text{CS} - N_E \times \text{Avg. CS} \\
&= \frac{\alpha - \tilde{c}_D}{2\eta}\left(\alpha - \frac{k+1}{k+2}\tilde{c}_D\right) - \frac{2\gamma c_M^k (k+1)(\alpha - \tilde{c}_D)}{\eta (\tilde{c}_D)^{k+1}\left(1+\tau^{-k}\right)} \times \frac{(\tilde{c}_D)^{k+2}}{4\gamma c_M^k (k+1)(k+2)} \\
&= \frac{(\alpha - \tilde{c}_D)}{2\eta}\left[\alpha - \frac{(k+1)\left(1+\tau^{-k}\right)+1}{(k+2)\left(1+\tau^{-k}\right)}\tilde{c}_D\right]
\end{aligned}$$

The expected profit of a firm can be derived from equation (20) in Ding [2020a]:

$$\begin{aligned}
\bar{\pi} &= \int_0^{\tilde{c}_D} \pi_D(c)\, dG(c) + \int_0^{\tilde{c}_X} \pi_X(c)\, dG(c) \\
&= \frac{(\tilde{c}_D)^{k+2}}{2\gamma c_M^k (k+1)(k+2)} + \frac{\tau^2 (\tilde{c}_X)^{k+2}}{2\gamma c_M^k (k+1)(k+2)} = \frac{(\tilde{c}_D)^{k+2}\left(1+\tau^{-k}\right)}{2\gamma c_M^k (k+1)(k+2)}
\end{aligned}$$

Notice when $t = 1, c_{FDI} = 0$ and $\tilde{c}_X = \tilde{c}_D/\tau$. With all these components and the fact that $\tilde{c}_D < \alpha$, equation (1) can now be properly signed:

$$\begin{aligned}
\text{Avg. CS} &= \frac{(\tilde{c}_D)^{k+2}}{4\gamma c_M^k (k+1)(k+2)} > 0 \\
N_E \frac{\partial \text{Avg. CS}}{\partial N_E} &= \frac{(\alpha - \tilde{c}_D)(\tilde{c}_D)^{k+2}}{4\gamma c_M^k (k+1)\left[k\tilde{c}_D - \alpha(k+1)\right]} < 0 \\
\frac{\partial \text{VE}}{\partial N_E} &= \frac{(\tilde{c}_D)^{k+2}\left\{\alpha(k+2)(1+\tau^{-k}) + (\alpha - 2\tilde{c}_D)\left[(k+1)(1+\tau^{-k})+1\right]\right\}}{4\gamma c_M^k (k+1)(k+2)\left[\alpha(k+1) - k\tilde{c}_D\right]} > 0 \\
N_E \frac{\partial \bar{\pi}}{\partial N_E} &= \frac{(\tilde{c}_D)^{k+2}\left(1+\tau^{-k}\right)(\alpha - \tilde{c}_D)}{2\gamma c_M^k (k+1)\left[k\tilde{c}_D - (k+1)\alpha\right]} < 0
\end{aligned}$$

Therefore, the externality of entry equals to

$$\begin{aligned}
&\text{Avg. CS} + N_E \frac{\partial \text{Avg. CS}}{\partial N_E} + \frac{\partial \text{VE}}{\partial N_E} + N_E \frac{\partial \bar{\pi}}{\partial N_E} \\
=& \frac{\left(1+\tau^{-k}\right)(\tilde{c}_D)^{k+2}}{4\gamma c_M^k (k+1)(k+2)\underbrace{\left[k\tilde{c}_D - (k+1)\alpha\right]}_{<0}} \times (\alpha - 2\tilde{c}_D)
\end{aligned}$$

Therefore, the externality will be negative if $\tilde{c}_D < \alpha/2$, will be positive if $\tilde{c}_D > \alpha/2$. □

## C.4 Proof of Proposition 3

**Proof:** Under symmetry, it is straightforward to derive the following welfare expression for market outcome and socially optimum outcome:

$$
\begin{aligned}
\mathbb{W}^M &= 1 + \bar{q}_0 + \frac{\alpha - c_D^M}{2\eta}\left(\alpha - \frac{k+1}{k+2}c_D^M\right) \\
\mathbb{W}^O &= 1 + \bar{q}_0 + \frac{1}{2\eta}\left(\alpha - c_D^O\right)^2
\end{aligned}
$$

And based on the proof of Proposition 1 that $c_D^M = (2/\triangle)^{\frac{1}{k+2}} c_D^O$, where the $\triangle$ is the symmetric version of $\triangle_F$ and $\triangle_H$. To simplify the expression, here I define:

$$
\begin{aligned}
A &\equiv \left(\frac{2}{\triangle}\right)^{\frac{1}{k+2}} \\
B &\equiv \frac{k+1}{k+2}
\end{aligned}
$$

Therefore, the welfare gap between market outcome and socially optimum is:

$$
\mathbb{W}^M - \mathbb{W}^O = \frac{1}{2\eta}\left[c_D^O(2\alpha - AB\alpha - A\alpha) + (c_D^O)^2(A^2B - 1)\right]
$$

It is obvious that the market welfare is smaller than the socially optimum welfare. One can then show the gap is decreasing in $t$:

$$
\frac{\partial(\mathbb{W}^O - \mathbb{W}^M)}{\partial t} = c_D^O \frac{\partial A}{\partial t}\left[\alpha(A+B) - 2ABc_D^O\right]
$$

Given that $\partial A/\partial t > 0$, then the whole expression will be negative if $c_D^O > \alpha(A+B)/2AB$, which is equivalent to $c_D^M > \alpha(A+B)/2B$. Therefore part *(i)* of the proposition is proved.

The second part of the proposition is easy. Notice that:

$$\mathbb{W}^H = 1 + \bar{q}_0^H + \frac{\alpha - c_D^H}{2\eta}\left(\alpha - \frac{k+1}{k+2}c_D^H\right)$$
$$\mathbb{W}^F = 1 + \bar{q}_0^F + \frac{\alpha - c_D^F}{2\eta}\left(\alpha - \frac{k+1}{k+2}c_D^F\right)$$

The rest of the proof directly follows from the Proposition 4 in Ding [2020a]. □

## C.5 Proof of Proposition 4

**Proof:** First, based on the definition of $IM$ in equation (29) in Ding [2020a], it is straightforward to show that $\partial IM^H / \partial t^H < 0$. Now compare equation (2) and (3)

$$t_S^H - 1 = \frac{\frac{\partial CS^H}{\partial t^H} + IM^H + \frac{\partial CS^F}{\partial t^H} + \left(t^F - 1\right) \times \frac{\partial IM^F}{\partial t^H}}{-\frac{\partial IM^H}{\partial t^H}}$$
$$t_N^H - 1 = \frac{\frac{\partial CS^H}{\partial t^H} + IM^H}{-\frac{\partial IM^H}{\partial t^H}}$$

It's straightforward to show that when $\tilde{c}_D > \alpha/2$:

$$\frac{\partial CS^F}{\partial t^H} = \underbrace{\frac{\partial CS^F}{\partial c_D^F}}_{<0} \times \underbrace{\frac{\partial c_D^F}{\partial t^H}}_{>0} < 0$$
$$\left(t^F - 1\right) \times \frac{\partial IM^F}{\partial t^H} = \frac{t^F\left(t^F-1\right)\left(\tau^F\right)^2}{4\gamma\left(k+2\right)\left(c_M\right)^k}\left[2\left(\frac{1}{t^F\tau^F}\right)^{k+2} - \frac{k+2}{\left(t^F\tau^F\right)^2}\left(\xi^F\right)^k + k\left(\xi^F\right)^{k+2}\right]\underbrace{\frac{\partial N_E^H\left(c_D^F\right)^{k+2}}{\partial t^H}}_{>0} > 0$$

The first item indicates the distortion on $F$'s consumption generated by $t^H$, the second item indicates the distortion on $F$'s tariff revenue generated by $t^H$. It's straightforward to show the distortion on consumption is larger than the distortion on tariff revenue. Therefore the sum of these two items is negative, indicating that when the Nash tariff is evaluated at the socially optimum first-order condition, it creates negative externality. Therefore, it must be the case that $t_S^H < t_N^H$. Hence the Nash tariff is higher than the socially optimal tariff. □

## C.6 Proof of Average Markup

Once again, to simplify the proof, I assume symmetry, following Appendix C.7, $c_D = c_D^H = c_D^F$ and

$$\begin{aligned}
N_E &= \frac{2\gamma c_M^k (k+1)(\alpha - c_D)}{\eta (c_D)^{k+1} \left(1 + t^{-k}\tau^{-k}\right)} \\
N_D &= \frac{2\gamma (k+1)(\alpha - c_D)}{\eta \left(1 + t^{-k}\tau^{-k}\right) c_D} \\
N_X &= \frac{2\gamma (k+1)(\alpha - c_D)}{\eta \left(1 + t^{-k}\tau^{-k}\right) c_D} \left[(t\tau)^{-k} - \xi^k\right] \\
N_{FDI} &= \frac{2\gamma (k+1)(\alpha - c_D)}{\eta \left(1 + t^{-k}\tau^{-k}\right) c_D} \xi^k
\end{aligned}$$

Together with equation (6), (10) and (15) in Ding [2020a], the average markup in (3) can be written as follow:

$$\begin{aligned}
\bar{m} &= \frac{1}{N_D + N_X + N_{FDI}} \left[ N_D \int_0^{c_D} m_D(c) \frac{dG(c)}{G(c_D)} \right. \\
&\quad \left. + N_X \int_{c_{FDI}}^{c_X} m_X(c) \frac{dG(c)}{G(c_X)} + N_{FDI} \int_0^{c_{FDI}} m_{FDI}(c) \frac{dG(c)}{G(c_{FDI})} \right] \\
&= \frac{1}{N_D + N_X + N_{FDI}} \left[ N_D \int_0^{c_D} \frac{c_D + c}{2c} \frac{kc^{k-1}}{c_D^k} dc \right. \\
&\quad \left. + N_X \int_{c_{FDI}}^{c_X} \frac{t(c_X + c)}{2c} \frac{kc^{k-1}}{c_X^k} dc + N_{FDI} \int_0^{c_{FDI}} \frac{c_D + \varphi^H c}{2\varphi^H c} \frac{kc^{k-1}}{c_{FDI}^k} dc \right] \\
&= \frac{1}{N_D + N_X + N_{FDI}} \left[ N_D \times \frac{2k-1}{2k-2} + N_X \times t \right. \\
&\quad \left. \times \left( \frac{2k-1}{2k-2} - \frac{k}{2k-2} (t\tau\xi)^{k-1} - \frac{1}{2} (t\tau\xi)^k \right) + N_{FDI} \times \left( \frac{k}{2k-2} \frac{1}{\varphi\xi} + \frac{1}{2} \right) \right]
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{1+(t\tau)^{-k}} \times \frac{2k-1}{2k-2} + \frac{t\left[(t\tau)^{-k} - \xi^k\right]}{1+(t\tau)^{-k}} \times \left[\frac{2k-1}{2k-2} - \frac{k}{2k-2}(t\tau\xi)^{k-1} - \frac{1}{2}(t\tau\xi)^k\right] \\
&+ \frac{\xi^k}{1+(t\tau)^{-k}} \times \left(\frac{k}{2k-2}\frac{1}{\varphi\xi} + \frac{1}{2}\right) \\
&= \underbrace{\frac{1}{1+(t\tau)^{-k}} \times \frac{2k-1}{2k-2}}_{\text{weighted expected markup in domestic}} \\
&+ \underbrace{\frac{1-(t\tau\xi)^k}{1+(t\tau)^{-k}} \times \frac{1}{t^{k-1}\tau^k}\left\{\frac{1}{2}\left[1-(t\tau\xi)^k\right] + \frac{k}{2k-2}\left[1-(t\tau\xi)^{k-1}\right]\right\}}_{\text{weighted expected markup from Foreign exporters}} \\
&+ \underbrace{\frac{\xi^k}{1+(t\tau)^{-k}} \times \left(\frac{k}{2k-2}\frac{1}{\varphi\xi} + \frac{1}{2}\right)}_{\text{weighted expected markup from Foreign FDI}}
\end{aligned}
$$

## C.7 Proof of Covariance Term

Once again, to simplify the analysis, I imposed symmetry. It's clear from equation (26) in Ding [2020a] that

$$N_E = \frac{2\gamma c_M^k (k+1)(\alpha - c_D)}{\eta (c_D)^{k+1}\left(1 + t^{-k}\tau^{-k}\right)}$$

where $c_D = c_D^H = c_D^F$. Given equation (7), (11) and (16) in Ding [2020a], and the following expression for the mass of firms:

$$
\begin{aligned}
N_D &= N_E \times G(c_D) = \frac{2\gamma(k+1)(\alpha - c_D)}{\eta\left(1+t^{-k}\tau^{-k}\right)c_D} \\
N_X &= N_E \times \left[G(c_X) - G(c_{FDI})\right] = \frac{2\gamma(k+1)(\alpha - c_D)}{\eta\left(1+t^{-k}\tau^{-k}\right)c_D}\left[(t\tau)^{-k} - \xi^k\right] \\
N_{FDI} &= N_E \times G(c_{FDI}) = \frac{2\gamma(k+1)(\alpha - c_D)}{\eta\left(1+t^{-k}\tau^{-k}\right)c_D}\xi^k
\end{aligned}
$$

the covariance term can be derived as follow:

$$\begin{aligned}
\operatorname{cov}\left(m^{i}(\omega), \frac{dl^{i}(\omega)}{L^{j}}\right) &= N_D \int_{0}^{c_D} p_D(c)\, d\left[q_D(c)\right] \frac{dG(c)}{G(c_D)} + N_X \int_{c_{FDI}}^{c_X} p_X(c)\, d\left[q_X(c)\right] \frac{dG(c)}{G(c_X)} \\
&+ N_{FDI} \int_{0}^{c_{FDI}} p_{FDI}(c)\, d\left[q_{FDI}(c)\right] \frac{dG(c)}{G(c_{FDI})} \\
&= \frac{2\gamma(k+1)(\alpha - c_D)}{\eta\left(1 + t^{-k}\tau^{-k}\right) c_D} \int_{0}^{c_D} \frac{k d c_D (c_D + c) c^{k-1}}{4\gamma c_D^{k}} dc \\
&+ \frac{2\gamma(k+1)(\alpha - c_D)}{\eta\left(1 + t^{-k}\tau^{-k}\right) c_D} \left[(t\tau)^{-k} - \xi^{k}\right] \int_{c_{FDI}}^{c_X} \frac{t^2\tau^2 k d c_X (c_X + c) c^{k-1}}{4\gamma c_X^{k}} dc \\
&+ \frac{2\gamma(k+1)(\alpha - c_D)}{\eta\left(1 + t^{-k}\tau^{-k}\right) c_D} \xi^{k} \int_{0}^{c_{FDI}} \frac{k d c_D (c_D + \varphi c) c^{k-1}}{4\gamma c_F^{k}} dc \\
&= \frac{(\alpha - c_D)\, dc_D}{2\eta\left(1 + t^{-k}\tau^{-k}\right)} \left\{2k + 1 + \left[(t\tau)^{-k} - \xi^{k}\right]\right. \\
&\left.\times \left[2k + 1 - k(1 - t\tau\xi)(t\tau\xi)^{k} - (t\tau\xi)^{k}\right] + \xi^{k}(k + k\varphi + 1)\right\}
\end{aligned}$$

When $\varphi > t\tau$, it is straightforward to show that $t\tau\xi < 1$. Therefore, the covariance term can be rewritten as

$$\begin{aligned}
\operatorname{cov}\left(m^{i}(\omega), \frac{dl^{i}(\omega)}{L^{j}}\right) = \frac{(\alpha - c_D)\, dc_D}{2\eta\left(1 + t^{-k}\tau^{-k}\right)} &\left\{2k + 1 + (t\tau)^{-k} \underbrace{\left[1 - (t\tau\xi)^{k}\right]}_{>0}\right. \\
&\left.\times \left[k + k(t\tau\xi)^{k+1} + \underbrace{k - k(t\tau\xi)^{k}}_{>0} + \underbrace{1 - (t\tau\xi)^{k}}_{>0}\right] + \xi^{k}(k + k\varphi + 1)\right\}
\end{aligned}$$

Notice, under symmetry

$$\begin{aligned}
\frac{dc_D}{dt} &= \frac{d}{dt}\left[\frac{\gamma\phi}{1 + \Phi_1 + \Phi_2}\right]^{\frac{1}{k+2}} \\
&= \frac{1}{k+2}\left[\frac{\gamma\phi}{1 + \Phi_1 + \Phi_2}\right]^{-\frac{k+1}{k+2}} \times \frac{-\gamma\phi}{(1 + \Phi_1 + \Phi_2)^2} \frac{d(\Phi_1 + \Phi_2)}{dt} \\
&= -\frac{1}{k+2} \times \frac{c_D}{1 + \Phi_1 + \Phi_2} \times \frac{d(\Phi_1 + \Phi_2)}{dt}
\end{aligned}$$

According to the proof of Proposition 4 in Ding [2020a], $d(\Phi_1 + \Phi_2)/dt < 0$, hence

$dc_D/dt > 0$, hence the covariance term is positive for $dt > 0$. $\square$