A Simple Model of Quality Heterogeneity and International Trade*

Elias Dinopoulos
University of Florida

Bulent Unel
Louisiana State University

Current Version: October 15, 2009

Abstract

This paper develops a trade model with firm-specific quality heterogeneity in high-tech markets where firms face the threat of imitation and engage in limit-pricing strategies. The model generates an endogenous distribution of prices and markups. Firms producing high-quality (high-price) products export, whereas firms producing low-quality (low-price) products serve the domestic market in accordance to the Alchian and Allen (1964) conjecture. Trade liberalization raises the average domestic markup and increases the number of products consumed in each country. However, the impact of trade liberalization on export markups depends on its nature: an increase in the number of trading countries raises the average export markup; a reduction in foreign market entry costs reduces the average export markup; and a reduction in per-unit trade costs has an ambiguous effect on the average export markup. Although the presence of endogenous markups renders the laissez-faire equilibrium suboptimal, trade liberalization raises national and global welfare.

An electronic copy of the latest version of the paper is available at
http://bear.cba.ufl.edu/dinopoulos/research.html
http://www.bus.lsu.edu/unel/research.html

JEL Codes: F10, F12, and F13

Keywords: International trade, monopolistic competition, firm heterogeneity, product quality, trade costs, and trade liberalization.

---

*Elias Dinopoulos, Department of Economics, University of Florida, Gainesville, FL 32611. E-mail: elias.dinopoulos@cba.ufl.edu; Tel: (352) 392-8150. Bulent Unel, Department of Economics, Louisiana State University, 2134 Patrick F. Taylor Hall, Baton Rouge, LA 70803. E-mail: bunel@lsu.edu, Tel: (225) 578-3792.

We would like to thank Tetsu Haruyama, Jee-Hyong Park, Jordan Petsas, Ina Simonovska, Alexandre Skiba, and seminar participants at Louisiana State University, University of Florida, and the Midwest International Economics Group Meetings for fruitful comments and suggestions.

1 Introduction

Several empirical studies have documented the presence of substantial firm heterogeneity in narrowly defined product categories.[1] According to these studies, firm heterogeneity takes the form of productivity or product-quality differences among establishments. More productive firms are larger, charge lower prices, and are more likely to engage in exporting. Firms producing higher-quality products charge higher prices and are more likely to engage in exporting.

These empirical findings have motivated the development of a growing strand of theoretical literature that has formally analyzed the nexus between firm heterogeneity and international trade patterns. Melitz (2003) was the first to develop a tractable model of productivity heterogeneity that captured several aforementioned stylized facts. The theoretical framework proposed by Melitz has been extended by subsequent studies along several dimensions.[2] For example, a number of recent papers have developed trade models based on firm-specific quality heterogeneity to study the pattern of export prices and markups.[3]

The starting point of the present paper is based on the observation that the aforementioned studies assume that firms maximize profits by charging the unconstrained monopoly price. Consequently, the equilibrium price-cost margin (markup) is inversely related to the price elasticity of demand whose absolute value must exceed unity. However, it is well known that in markets where imitation is prevalent, such as many R&D–intensive industries, frequently firms face inelastic demand curves and engage in entry-preventing strategies to minimize the threat of imitation. The low sensitivity of quantity demanded to price changes can be traced on advertising, high switching costs, and high consumer-

[1] Important contributions are Clerides et al. (1998), Bernard and Jensen (1999), Aw et al. (2000), Schott (2004), Hummels and Klenow (2005), Hallak (2006), and Manova and Zhang (2009) among many others.

[2] Several studies have focused on productivity heterogeneity to analyze foreign-direct-investment patterns, wage-income distribution, and bilateral-trade volumes. Helpman (2006) provides an insightful survey of the theoretical literature on trade with heterogeneous firms

[3] See Baldwin and Harrigan (2007), Johnson (2009), Verhoogen (2008), Kugler and Verhoogen (2008), Antoniades (2008), and Simonovska (2009). Manova and Zhang (2009) provide an overview of this literature.

search costs. Consider, for instance, the market for pharmaceuticals, where firms engage in R&D to discover new and better drugs. Siminski (2008) argues that the price elasticity of demand for pharmaceuticals among older people in Australia is -0.1, implying that the quantity demanded is not responsive to price changes.[4] How is the pattern of prices, profits, and markups determined in the pharmaceutical industry where the unconstrained monopoly price is undefined? Does trade liberalization work through a channel other than changes in the price elasticity of demand? Does the threat of imitation changes the effects of trade on the reallocation of resources and welfare in these markets?

The present paper fills up this gap by introducing entry-deterring limit-pricing strategies triggered by the threat of imitation in markets with heterogeneous firms. Firm heterogeneity arises from an R&D structure that follows Melitz's (2003) seminal work. We assume that firms face uncertainty with respect to the future level of product quality when they engage in R&D to discover new products. This uncertainty is modeled by an unrestricted distribution of product quality levels. We also assume that the quality of each discovered variety remains constant during the life time of a firm. In other words, firms cannot alter the quality of their products by changing their fixed or marginal costs. Consequently, our framework differs from models of productivity heterogeneity where firms choose deterministically their product quality levels.[5] Although one could think of examples where firms can alter the quality of their product by choosing the quality of inputs (say, better leather to make better quality shoes), our model is more relevant to high-tech industries (say, pharmaceuticals, electronics, chemicals) where the process of quality improvement involves considerable uncertainty.

Following the insights of endogenous growth theory, we introduce Bertrand price com-

[4] Tellis (1988) describes a meta-analysis of econometric studies that generated 367 price elasticities for about 220 brands. He reports (Figure 1, p. 337) that the brand average price elasticity is -1.76 with a standard deviation of 1.74. This finding implies that many brands face inelastic demand curves.

[5] See, for instance, Baldwin and Harrigan (2007), Johnson (2009), Verhoogen (2008), Kugler and Verhoogen (2008), and Antoniades (2008) among others. Our framework can readily be adjusted to study the nexus between productivity heterogeneity and deterministic endogenous quality choice as it will become clear below.

petition between a firm that discovers a new variety and a competitive fringe of imitators.[6] This assumption results in entry-deterring limit prices (as opposed to unconstrained monopoly prices) that are prevalent in high-technology markets. We also depart from the existing literature on heterogeneous firms and trade by assuming that each firm faces a unitary-elastic demand curve, which is derived from a Cobb-Douglas utility function. The assumption of Cobb-Douglas preferences simplifies the analysis considerably without changing the aggregate properties of the model.[7]

In the model, firms producing higher-quality products charge higher prices and markups, enjoy higher profits, and export their products. In contrast, firms producing lower-quality products charge lower prices, earn lower profits, and serve only the domestic market. The model generates an endogenous distribution of prices and markups that allows the study of trade liberalization in a global economy consisting of many structurally identical countries. Trade liberalization can take a variety of forms given the rich structure of our model: a move from autarky to restricted trade; an increase in the size of the global economy measured by the number of trading countries; a reduction in trade costs; and a reduction in foreign-market entry costs. All these different facets of trade liberalization generate the following effects: an increase in the number of varieties available for consumption; a reallocation of resources from low-quality to high-quality products that generates exit of inefficient firms from the domestic market; and an improvement in national and global welfare.

Trade liberalization increases the domestic minimum quality level that enables firms to produce profitably and serve the domestic market (that is, the production "cutoff" quality level) and raises the average price charged by domestic producers and the average domestic markup. However, the effects of trade liberalization on the average export markup depend on its nature: a move from autarky to restricted trade or an increase in the number of

[6]Segerstrom et al. (1990) and Grossman and Helpman (1991), among many others, provide more details on this type of market structure.

[7]Under Dixit and Stiglitz (1977) preferences and the threat of imitation, each firm faces a kinked demand curve. In this case, the profit-maximizing entry deterring strategy can be either a constant monopoly markup or a limit price depending on the quality level. Under the assumption that the elasticity of substitution across varieties is sufficiently low, all firms follow limit-pricing entry-deterring strategies as in our model.

trading countries increases the average export cutoff level, the average export price, and the average export markup; a decline in foreign-market-entry costs reduces the export cutoff quality level, the average export price, and average export markup; and a reduction in per-unit trade costs reduces the export cutoff quality level and the average export price, but it has an ambiguous effect on the average export markup.

Although trade liberalization is welfare improving, the laissez-faire equilibrium is inefficient. We demonstrate the inefficiency of the laissez-faire equilibrium for the case of a closed economy in order to simplify the exposition. This welfare property can be traced to the difference between the socially optimal and the market-equilibrium average markups. The social planner is interested in the welfare of the average consumer which depends on the average quality and consumer surplus, whereas the market is interested in the behavior of the marginal consumer which depends on the product with the marginal quality. This welfare property is consistent with the generalized theory of distortions and welfare and with the insights of quality-ladders growth theory. However, it is not present in models with Dixit and Stiglitz (1977) preferences and heterogeneous firms. For instance, Feenstra and Kee (2008) use a version of Melitz's model to establish that the laissez faire cutoff productivity levels are socially optimal. In contrast, our model generates a socially suboptimal laissez-faire cutoff quality level.

Before discussing the relationship of our model to the growing empirical literature on trade with heterogeneous firms, it is worth for stating two points. First, our model should be viewed as a tractable theoretical framework that can be readily modified or augmented depending on the empirical question analyzed. For example, it is well known that the assumption of symmetric Cobb-Douglas preferences generates identical market shares across brands independently of their price (quality in our model). This implication is clearly unrealistic, but can be readily rectified by introducing quality-based cost heterogeneity into the model (see Section 5). Second, our model is motivated and therefore better fitted to analyze quality heterogeneity in high-tech markets where the threat of imitation and entry-deterring limit-pricing strategies are more common. Our search of the existing empirical literature has revealed that most of these studies rely on data coming from developing countries (e.g., China, Chile, Cote d'Ivoire, India, Mexico, Colombia, Costa Rica, and Slovenia). These countries account for a very small fraction of innovative global R&D and are typically populated by low-tech industries. Thus we view the link between the empirical findings of these studies and our model's predictions suggestive rather than conclusive.

Despite these concerns, our model offers several predictions that enjoy empirical support. For instance, the prediction that each surviving firm charges a price that is proportional (as opposed to be just positively correlated) to its product quality level is consistent with empirical studies that routinely use unit values to measure product quality (Schott, 2004; Hummels and Skiba, 2004; and Hallak, 2006). Furthermore, the prediction that the average quality (and price) of exports is higher than the average quality (and price) of products sold only in the domestic market provides a novel general-equilibrium explanation based on quality sorting of the Alchian and Allen (1964, 74-75) conjecture of "shipping the good apples out." This prediction is also supported by several empirical studies (Hummels and Skiba, 2004; Baldwin and Harrigan, 2007; and Verhoogen, 2008).[8] Interestingly, models of firm-specific productivity heterogeneity predict that more productive exporters charge lower prices.[9]

Section 5 demonstrates how the empirical relevance of our model can be improved by the introduction of quality-based cost heterogeneity. The augmented version of the model predicts that firms with higher-quality products charge higher prices in more distant markets; and enjoy higher markups, larger revenues, and larger market shares. The model also predicts that more firms export to more proximate markets, and that sufficiently high

[8]Feenstra (2004) reports that the average price and quality of Japanese exported cars to the US was higher than the corresponding average price and quality of domestic US cars. For instance, according to Table 8.3 on page 274, in 1979 (two years before the auto VER was imposed on Japanese cars) the unit value of Japanese cars imported into the US was $4,949 compared to $4,186 of small US cars. The corresponding unit-quality values for imported Japanese and domestic US cars were $4,361 and $4,197, respectively.

[9]See, for instance, Schott (2004) who states that "unit-value patterns are inconsistent with new trade theory models that have producer price varying inversely with producer productivity."

foreign-market-entry costs eliminate bilateral trade flows between distant markets. These predictions are supported by several recent empirical studies (Baldwin and Harrigan, 2007; Kugler and Verhooven, 2008; and Johnson, 2009). Finally, the model's prediction that a reduction in trade costs has an ambiguous effect on the average export markup is also consistent with Harrison's (1994) empirical findings.[10]

Our paper is related to several studies that deliver heterogeneous markups and prices. Bernard, Eaton, Jensen, and Kortum (2003) develop a model of trade with heterogeneous firms, limit prices, and firm-specific Ricardian comparative advantage. Their model delivers an exogenous distribution of markups and relies on an exogenous number of varieties. In contrast, our model generates an endogenous distribution of markups and an endogenous number of varieties.

Melitz and Ottaviano (2008) develop a model of productivity heterogeneity, quasi-linear preferences, and trade between two unequal-size countries. Their model generates an endogenous distribution of markups and addresses similar questions to those addressed in our paper. However, in their model the wage is fixed by the presence of an outside good, production and exporting do not involve fixed costs, the domestic and foreign markets are segmented, and firms charge unconstrained monopoly prices. In Melitz and Ottaviano (2008) trade liberalization intensifies the product market competition by increasing the price elasticity of demand for each variety as in Krugman (1979) without affecting the wage. In contrast, trade liberalization in our model operates through income-induced shifts in the demand for each variety and changes in the real wage. Thus our model offers a unified framework that combines the effects of trade on product and factor markets. This feature is missing from other trade models that generate heterogeneous markups in the context of quasi-linear preferences.[11] Consequently, our findings complement those of Melitz and

[10] Using plant-level data from Cote d'Ivoire, Harrison analyzed changes in markups following a 1985 trade liberalization episode which reduced trade costs. She found that markups, measured as profits over sales, increased in five out of nine sectors and declined in others following the reforms.

[11] For instance, Antoniades (2008) adds the feature of ex-post deterministic endogenous quality choice to the Melitz and Ottaviano (2008) model. The resulting quality heterogeneity is a by product of productivity heterogeneity. In contrast, our model assumes that the process that governs the generation of quality

Ottaviano (2008). In their model, an increase in the foreign market size reduces the average markup and price, whereas in our model an increase in the number of trading countries increases the average markup and price. In addition, our model allows the study of lower foreign or domestic-market entry costs which is missing from their model.

Several recent models have introduced vertical differentiation in the Melitz (2003) model while maintaining the assumption of Dixit and Stiglitz (1977) preferences (Baldwin and Harrigan, 2007; Verhoogen, 2008; and Kugler and Verhoogen, 2008). These models introduce cost heterogeneity based on the assumption that higher quality products entail higher marginal manufacturing costs. These models deliver price heterogeneity but constant markups thanks to Dixit and Stiglitz (1977) preferences.

In summary, this paper makes three novel contributions to the theory of firm heterogeneity and trade. First, it offers a unique and tractable analytical framework of quality-based firm heterogeneity where firms face unitary elastic demand curves, the threat of imitation, and engage in entry preventing limit-pricing strategies. The model delivers heterogeneous prices and markups and allows the study of trade liberalization that is transmitted though the intensity of product market competition and changes in the real wage. This general-equilibrium mechanism is missing from the rest of the literature. Second, the proposed framework offers several predictions on the pattern of prices and markups in markets populated by vertically differentiated products that are missing from the literature. For instance, it demonstrates that the impact of trade liberalization on the average export markup depends on its nature: the average export markup increases in the number of trading countries and in foreign-market-entry costs; and its dependence on per-unit trade costs is ambiguous. Third, consistent with the generalized theory of distortions and welfare, our model demonstrates the sub-optimality of the laissez-faire equilibrium and opens the possibility for studying the nature of welfare improving policies. This important welfare result is also missing from the literature on firm heterogeneity and trade.

heterogeneity is uncertain. In models of endogenous quality choice, trade liberalization has an ambiguous effect on average export prices and markups.

The rest of the paper is organized as follows. Section 2 presents the basic elements of the model and the steady-state equilibrium. Section 3 analyzes the impact of trade liberalization and the effects of a move from autarky to restricted trade. Section 4 describes the model's welfare properties. Section 5 introduces quality based cost heterogeneity and spatial considerations. Section 6 offers some concluding remarks.

2 The Model

In this section, we present the basic elements of the model regarding consumer preferences, structure of production, and firm entry decisions. We consider a global economy consisting of $n+1$ structurally identical countries with $n \geqslant 1$. Each economy has a single industry populated by heterogeneous firms, and labor is the only factor of production. In each country, the aggregate supply of labor, L, is fixed and remains constant over time.

2.1 Consumer Preferences

Consumer preferences are identical across all countries and modeled by the following Cobb-Douglas utility function defined over a continuum of products indexed by ω

$$U = \int_{\omega \in \Omega} \ln \left[\beta \lambda(\omega) \frac{q(\omega)}{L} \right] d\omega, \tag{1}$$

where $\beta > 0$ is a constant, $\lambda(\omega)$ denotes the *time-invariant* product quality, $q(\omega)$ is the aggregate consumption of brand ω, and Ω is the set of varieties available for consumption in a typical country. We focus our analysis on the case where each consumer buys all available varieties, that is, we assume that the non-satiation principle holds. This case arises if parameter β is sufficiently high to ensure that the utility increases monotonically in the mass of varieties consumed.[12]

Maximizing (1) subject to the budget constraint yields the standard Cobb-Douglas demand for a typical variety

[12]The condition $\beta > eL/f_x$, where e is the natural logarithm base and f_x is the fixed foreign market entry cost, guarantees the validity of the non-satiation principle. Section 2.4 provides more details on its derivation. We would like to thank Tetsu Haruyama for pointing this out.

$$q(\omega) = \frac{EL}{p(\omega)M_c}, \tag{2}$$

where E is *per-capita* consumer expenditure, L is the number of consumers in a typical (home or foreign) market, $p(\omega)$ is the corresponding price of brand ω, and M_c is the measure of Ω (i.e., the mass of varieties available for consumption). The market demand for a product increases in aggregate consumer expenditure EL; and decreases in price $p(\omega)$ and the number of available products M_c.

2.2 Production

There is a continuum of firms, each choosing to produce a different product variety. Labor is the only factor of production, with each worker supplying one unit of labor. Production involves both fixed and variable costs: in order to produce q units of output, $\ell = f_p + q$ units of labor are required independently of the level of quality, where f_p denotes the fixed overhead cost of production measured in units of labor. Without any loss of generality, this formulation assumes that the marginal cost of production is equal to the wage of labor.

Firms wishing to export must incur per-unit trade costs and fixed costs as in Melitz (2003). Iceberg trade costs (such as transport costs and tariffs) are modeled in the standard fashion: $\tau > 1$ units of output must be produced at home in order for one unit to arrive at its destination. In addition, exporting involves a fixed foreign-market-entry cost of $F_x > 0$ that does not depend on the firm's quality level or the geographic location of production. This cost covers the costs of setting a distribution system, collecting information about the foreign market demand, product modifications and adjustments to local tastes, and costs based on regulations imposed by governments.[13] The decision to export occurs after the product's quality is revealed.

Each incumbent firm faces a constant probability of death δ in each period. In the present context, this stochastic shock can be interpreted as adverse changes in tastes that eliminate the demand for a particular variety. Consequently, in the steady-state, each firm is

[13] Existence of such market costs of exporting have been well documented by several studies (e.g., Roberts and Tybout, 1997; and Bernard and Jensen, 1999).

indifferent in principle between paying $f_x = \delta F_x$ in each period and the one-time fixed cost F_x in the first period of its existence. Hereafter, we assume that in each period exporters face an overhead fixed cost f_x in addition to the overhead production cost f_p. Firms that serve only the domestic market face just the overhead production cost f_p.

Next, consider the optimal pricing decision of a firm selling a brand of quality λ in its home market. Because each brand is associated with a unique quality level, in what follows we label products based on their quality levels. The aggregate quantity demanded is given by equation (2), which implies that that expenditure per variety, $p(\lambda)q(\lambda)$, is independent of the brand's quality level. Because the elasticity of demand for each variety is unity, a typical firm has an incentive to charge an infinite price and produce an infinitesimally small quantity independently of the product's quality level. To prevent this from happening and to create an endogenous distribution of markups, we follow the spirit of Schumpeterian growth theory and assume that once a product is introduced in a market (domestic or foreign), a generic, lower-quality version of the product can be produced instantaneously by a competitive fringe of firms. The production of each generic product exhibits constant returns to scale with one unit of labor producing one unit of output. We suppose that the generic version of a product cannot be produced in a country unless the original product is sold there. In other words, the technology to produce generic products diffuses internationally through imports.[14] We normalize the quality level of each generic good to one independently of the quality level of the copied product and the location of production.

Denote with $p_d(\lambda)$ and $p_x(\lambda)$ the consumer price prevailing in the domestic and foreign markets respectively, and assume that competition within each product occurs in a Bertrand

[14] Alternatively, one can assume that once a product is developed, its low-quality generic version can be produced by a competitive fringe in all countries, i.e., technology diffuses instantly across all countries. Analysis based on this assumption yields qualitatively the same results, and is available upon request. Moreover, one can also assume that there is no international transfer of technology. This assumption would allow exporters to charge a higher price abroad that would be proportional to the product's quality level adjusted by per-unit trade costs. More precisely, in the absence of international technology transfer, the quality leader charges two limit prices: $p_d(\lambda) = \lambda$ in the domestic market to get rid of the competitive fringe; and (in the absence of a competitive foreign fringe) it charges an export limit price $p_x(\lambda) = \tau\lambda$ which prevents the domestic fringe from exporting. In this case, a reduction in trade costs lowers the average export markup. The rest of the main results hold in this case as well.

fashion. The possibility of costless imitation forces firms to maximize profits by charging a (limit) price no higher than $p_d(\lambda) = p_x(\lambda) = \lambda w$, where w is the common wage rate across all countries, hereafter normalized to unity. This optimal pricing rule drives domestic and foreign imitators out of the market and implies that firms with higher-quality products charge higher prices.[15]

The limit-pricing rule $p_d(\lambda) = p_x(\lambda) = \lambda$ and (2) yield

$$\frac{q(\lambda_2)}{q(\lambda_1)} = \frac{p(\lambda_1)}{p(\lambda_2)} = \frac{\lambda_1}{\lambda_2}, \tag{3}$$

which means that firms with higher-quality products charge higher prices and sell lower quantities.

The per-period profits of exporting firms can be decomposed into two parts: profits earned from domestic sales $\pi_d(\lambda)$, and profits earned from sales in each of n export markets $\pi_x(\lambda)$.

$$\pi_d(\lambda) = [p_d(\lambda) - 1]q_d(\lambda) - f_p = (1 - \lambda^{-1})\frac{EL}{M_c} - f_p, \tag{4}$$

$$\pi_x(\lambda) = [p_x(\lambda) - \tau]q_x(\lambda) - f_x = (1 - \tau\lambda^{-1})\frac{EL}{M_c} - f_x, \tag{5}$$

where the quantities demanded by domestic and foreign consumers $q_d(\lambda)$ and $q_x(\lambda)$ are given by (2) and $p_d(\lambda) = p_x(\lambda) = \lambda$.

It is important to emphasize that, as the Melitz model, our approach is isomorphic to a model of process innovations, where firm heterogeneity is derived from productivity differences. The Appendix establishes formally that the latter yields the profit functions described by (4) and (5); an endogenous distribution of heterogeneous markups; and the property that all firms charge the same price, but more productive firms enjoy lower average and marginal costs and higher markups. In summary, the combination of Cobb-Douglas preferences and limit-pricing strategies generates an endogenous distribution of markups

[15] The Appendix shows formally how an augmented version of the consumer utility function (1) can generate the aforementioned optimal limit-pricing rule.

that is missing from models that rely on Dixit and Stiglitz (1977) preferences.[16]

Notice that symmetry of foreign markets implies that the global profit flow generated by exporting equals $n\pi_x(\lambda)$. Because only a fraction of incumbent firms export, a firm producing a good with quality λ earns a per-period profit $\pi(\lambda) = \pi_d(\lambda) + max\{0, n\pi_x(\lambda)\}$. Since each firm faces a constant probability of death δ in each period, the market value of a typical firm is given by

$$\nu(\lambda) = max\left\{0, \sum_{t=0}^{\infty}(1-\delta)^t\pi(\lambda)\right\} = max\left\{0, \frac{1}{\delta}\pi(\lambda)\right\}, \tag{6}$$

where the second equality follows from the fact that each firm's product quality remains constant during its lifetime.

A product of quality λ is produced only if $\pi(\lambda) \geqslant 0$. Therefore, the production cutoff quality level λ_p is determined as follows. Equation (4) implies that $\pi_d(\lambda)$ increases in λ, and that $\pi_d(1) = -f_p < 0$. Thus, for any given values of $f_p > 0$, E, and M_c (which are common to all firms), there exists a production cutoff quality level λ_p, such that all firms producing varieties with quality $\lambda \geqslant \lambda_p$ earn non-negative profits and stay in the market. Firms that know how to produce varieties with quality below the production cutoff quality level λ_p exit the market. Setting (4) equal to zero generates the following *production cutoff quality level*

$$\lambda_p = \frac{1}{1 - M_c f_p/EL} > 1. \tag{7}$$

The production cutoff quality level λ_p depends positively on production fixed costs f_p, and negatively on its market size EL/M_c. This makes sense: in markets with high fixed production costs or low expenditure per variety, only high-quality (high-price) products can earn non-negative profits.

A product with quality λ is exported to all foreign markets, only if $\pi_x(\lambda) \geqslant 0$. Equation (5) implies that, as long as $\lambda > \tau > 1$ holds, $\pi_x(\lambda)$ increases monotonically in λ, and that

[16]This feature of our model complements the productivity-heterogeneity model proposed by Melitz and Ottaviano (2008) which generates variable average markups based on quasi-linear (as opposed to homothetic) preferences.

$\pi_x(\tau) = -f_x < 0$. Therefore, for any $f_x \geqslant 0$, E, and M_c there exists a cutoff quality level $\lambda_x \geqslant \tau$, such that all firms producing products with quality $\lambda \geqslant \lambda_x$ earn non-negative profits from exporting to any foreign market. However, firms that produce varieties with quality $\lambda < \lambda_x$ face strictly negative foreign profits and do not export. Setting (5) equal to zero yields the following *export cutoff quality level*

$$\lambda_x = \frac{\tau}{1 - M_c f_x / EL} \geqslant \tau > 1. \tag{8}$$

The export cutoff quality level λ_x depends positively on factors (such as variable trade costs τ and foreign-market entry costs f_x) that adversely affect profits from exporting; and negatively on factors (such as foreign market size EL/M_c) that have a positive impact on export profits.

Solving (7) for EL/M_c and substituting the resulting expression in (8) yields

$$\lambda_x = \frac{\tau}{1 - (1 - \lambda_p^{-1})(f_x / f_p)}. \tag{9}$$

Equation (9), together with the restrictions imposed by (7) and (8), establishes the dependence of the exporting cutoff quality level λ_x on the production cutoff quality level λ_p and the model's parameters. The requirement that the denominator of (9) must be non-negative implies that $\lambda_p \in (1, k)$, where $k = 1/[1 - f_p/f_x]$. Inspection of (9) also indicates that λ_x is a monotonically increasing function of the production cutoff quality level $\lambda_p \in (1, k)$, the level of per-unit trade costs τ, and the ratio of overhead fixed costs f_x/f_p. In addition, if $f_x = f_p$, then (9) implies that $\lambda_x = \tau\lambda_p > \lambda_p$ for all $\lambda_p \in (1, \infty)$. Therefore, under the parameter restrictions $f_x \geqslant f_p$ and $\tau > 1$, the exporting cutoff quality level λ_x is strictly greater than the production cutoff quality level λ_p. In this case, firms whose product quality level is less than λ_p exit; firms whose product quality level is $\lambda \in [\lambda_p, \lambda_x)$ produce exclusively for the domestic market because they earn non-negative profits from the domestic operations only; and firms producing high-quality products ($\lambda \geqslant \lambda_x$) sell their products in both domestic and all foreign markets.

A sufficient condition for the partition of firms by export status is that the overhead costs of operating in the domestic market f_p must not exceed the overhead costs of entering a foreign market f_x. A similar condition has been derived by Melitz (2003, p. 1709) for the case of Dixit and Stiglitz (1977) preferences. However, in Melitz's model, no level of trade costs τ can generate the aforementioned partitioning in the absence of exporting fixed costs ($f_x = 0$). In contrast, the present model generates this partitioning due to the limit-pricing behavior of firms. Notice that, in the absence of fixed exporting costs, equation (9) yields $\lambda_x = \tau$. In addition, observe that as the level of production fixed costs approaches zero ($f_p \to 0$) equation (7) yields $\lambda_p \to 1$. By continuity, the present model can deliver the partition of firms by export status under sufficiently high trade costs combined with low production fixed costs. This property leads to a novel prediction: a firm facing foreign markets with sufficiently different trade costs will not export to markets with high trade costs (see Section 5 for more details on this issue).

This prediction is consistent with the empirical findings of Schott (2004, figure 2) and Baldwin and Harrigan (2007). The former study uses U.S. product level data and reports that in 1994 about 90 percent of ten-digit HS product categories and almost 80 percent of 4-digit SITC categories exhibited zero imports! The second study argues that trade costs are positively correlated with the absence of U.S. exports in narrowly defined product categories. Models of firm heterogeneity that rely on Dixit and Stiglitz (1977) preferences cannot generate this prediction, thanks to constant markup pricing across all varieties.

The following proposition summarizes the aforementioned analysis.

Proposition 1. *Let λ_p and λ_x denote the production and export cutoff quality levels, respectively; and let $k = 1/[1 - f_p/f_x]$ denote the upper bound of the production cutoff quality level λ_p. Then,*

a. *The export cutoff quality level λ_x is an increasing function of the production cutoff quality λ_p.*

b. *If the production fixed cost does not exceed the foreign-market entry cost (i.e., $f_p \leqslant f_x$),* *then the export cutoff quality level is strictly greater than the production cutoff quality level (i.e.,* $\lambda_x > \lambda_p > 1$*).*

c. *In the absence of foreign market entry cost (i.e.,* $f_x = 0$*), there exists a level of trade cost such that the exporting cutoff quality level is strictly greater than the production cutoff quality level (i.e.,* $\lambda_x = \tau > \lambda_p > 1$*).*

Proposition 1 implies that firms with high-quality products charge higher prices, enjoy higher profits and ship these products abroad in accordance to the Alchian and Allen (1964) conjecture, which has been confirmed empirically by Hammels and Skiba (2004). Our paper develops a general equilibrium model in which trade costs and self-selection among heterogeneous firms lead to the desired result.

2.3 Entry Decision

The determination of the production cutoff quality level depends on entry and exit considerations. We assume that there is a large number of prospective and ex-ante identical entrants. Each entrant faces a fixed entry cost $f_e > 0$, which is measured in units of labor and interpreted as the number of R&D researchers employed by the entrant to discover a new variety. After a firm incurs the fixed entry cost, it draws its quality parameter λ from a common and known distribution $g(\lambda)$ with positive support over $(0, \infty)$ and with continuous cumulative distribution $G(\lambda)$. The properties of $g(\lambda)$ determine the benefits of entry measured by the relevant expected discounted profits.

The ex-ante probability of drawing a quality level λ is governed by the density function $g(\lambda)$ and the ex-ante probability of successful entry $1 - G(\lambda_p)$. As in Melitz (2003), the ex-post distribution of product quality levels μ is the conditional distribution of $g(\lambda)$ on the interval $[\lambda_p, \infty)$:

$$\mu(\lambda) = \begin{cases} \frac{g(\lambda)}{1-G(\lambda_p)} & \text{if } \lambda > \lambda_p \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

The ex-ante probability that an incumbent firm will export is given by

$$\zeta_x = \frac{1 - G(\lambda_x)}{1 - G(\lambda_p)}. \tag{11}$$

In addition, the law of large numbers implies that ζ_x equals the ex-post fraction of incumbent firms that export.

Let M_p denote the mass of varieties (and firms) produced in any country and M_x be the number of varieties that each country exports. Then, we have $M_x = \zeta_x M_p$, which further ensures that $M_c = (1 + n\zeta_x)M_p$ is the the mass of products available for consumption in any country. Armed with the aforementioned probability distributions, one can calculate the aggregate quantities demanded:

$$Q_c = \int_{\lambda_p}^{\infty} q(\lambda) M_p \mu(\lambda) d\lambda + n \int_{\lambda_x}^{\infty} q(\lambda) M_p \mu(\lambda) d\lambda = \frac{M_p}{M_c} \frac{EL}{\tilde{\lambda}_p} + \frac{nM_x}{M_c} \frac{EL}{\tilde{\lambda}_x}, \tag{12}$$

where $\tilde{\lambda}_i \quad (i = p, x)$ is given by

$$\tilde{\lambda}_i \equiv \tilde{\lambda}(\lambda_i) = \left[\frac{1}{1 - G(\lambda_i)} \int_{\lambda_i}^{\infty} \lambda^{-1} g(\lambda) d\lambda \right]^{-1}. \tag{13}$$

That is, $\tilde{\lambda}_p$ is the weighted harmonic mean of the quality levels (and prices) of all produced goods and can be interpreted as the average (expected) quality level. Similarly, $\tilde{\lambda}_x$ is the weighted harmonic mean of the quality levels of a country's exports and can be interpreted as the average export quality.

Because the production cutoff quality level is the minimum quality level of all surviving products, it must be lower than the average quality level. Moreover, an increase in λ_p forces producers with low-quality products to exit the market, which in turn increases the average quality level of all produced varieties. The same intuition applies to the relationship between the export quality cutoff level and the average quality of exports. The following lemma summarizes these properties (see Appendix for proof).

Lemma 1. *The average quality level of all products produced in a typical market is strictly greater and increases in the production cutoff quality level, i.e.,* $\tilde{\lambda}_p > \lambda_p$ *and* $\partial\tilde{\lambda}_p / \partial\lambda_p > 0$.

The average quality level of exports is strictly greater and increases in the export cutoff quality level, i.e., $\tilde{\lambda}_x > \lambda_x$ *and* $\partial\tilde{\lambda}_x/\partial\lambda_x > 0$.

The ex-ante quantity demanded for each variety is obtained by dividing (12) by M_p and is given by

$$\bar{q} = \tilde{\lambda}_p^{-1} EL/M_c + n\zeta_x \tilde{\lambda}_x^{-1} EL/M_c, \tag{14}$$

where first and second terms correspond to the domestic and foreign demand for a typical variety, respectively. Consequently, the ex-ante per-period profit of a typical prospective entrant is given by

$$\bar{\pi} = \pi_d(\tilde{\lambda}_p) + n\zeta_x \pi_x(\tilde{\lambda}_x) = (1 - \tilde{\lambda}_p^{-1})\frac{EL}{M_c} - f_p + n\zeta_x \left[(1 - \tau\tilde{\lambda}_x^{-1})\frac{EL}{M_c} - f_x\right]. \tag{15}$$

Because the probability of successful entry is $1 - G(\lambda_p)$, the net benefits of entering the domestic market are equal to the expected value of a firm $[1 - G(\lambda_p)]\bar{\nu}$, where $\bar{\nu} = \bar{\pi}/\delta$ is the ex-ante value of a prospective entrant. Setting the benefits of entry equal to the fixed R&D costs yields the free-entry condition

$$[1 - G(\lambda_p)]\frac{\bar{\pi}}{\delta} = f_e, \tag{16}$$

where $\bar{\pi}$ is defined by (15). This concludes the description of the model.

2.4 Steady-State Equilibrium

This section determines the market cutoff quality levels, number of varieties, the average markups, and the welfare level in a typical country. Substituting (9) and (11) in (15) yields an expression for ex-ante profits that depends only on the two cutoff quality levels. Further substitution of $\bar{\pi}$ into the free-entry condition (16) yields the basic steady-state equilibrium condition

$$f_p H(\lambda_p, 1) + n f_x H(\lambda_x, \tau) = \delta f_e, \tag{17}$$

where H is defined as

$$H(\lambda_i, \alpha) \equiv [1 - G(\lambda_i)] \left[\frac{1 - \alpha \tilde{\lambda}_i^{-1}}{1 - \alpha \lambda_i^{-1}} - 1 \right], \tag{18}$$

where $(i, \alpha) \in \{(p, 1), (x, \tau)\}$.

The export cutoff quality level λ_x is an increasing function of λ_p (see equation (9)). In addition, the Appendix proves that H is strictly decreasing in cutoff levels and that the left-hand-side of equation (17) is also decreasing in $\lambda_p \in (1, k)$. The above considerations imply the following result (see Appendix for proof).

Proposition 2. *Let $k = 1/(1 - f_p/f_x)$ and assume that $f_p < f_x$. There exist a unique production and a unique export cutoff quality levels $\lambda_p \in (1, k)$ and $\lambda_x \in (\tau, \infty)$ which satisfy equations (9) and (17) such that $\lambda_p > 1$, $\lambda_x > \tau$, and $\lambda_x > \lambda_p$.*

Once the two cutoff quality levels are determined, one can solve for the values of the remaining endogenous variables. We start with the determination of the mass of products produced in each market. In the steady-state equilibrium, the per-period flow of successful entrants must be equal to flow of incumbents who exit the market because they are hit by a bad shock, i.e., $[1 - G(\lambda_p)]M_e = \delta M_p$, where M_e is the mass of all (as opposed to successful) entrants. Then the aggregate amount of labor employed by prospective entrants is $L_e = M_e f_e = \delta M_p f_e / [1 - G(\lambda_p)] = M_p \bar{\pi}$, where the last equality follows from the free-entry condition (16). Thus the aggregate amount of labor devoted to R&D equals the level of aggregate profits earned by all producers in a typical market.

The aggregate demand for labor in a typical market equals the aggregate supply of labor $L_p + L_e = L_p + \Pi = L$, where L_p denotes the total amount of labor employed in the production of surviving goods and Π is the level of aggregate profits earned by all producers. In addition, the standard GDP identity implies that the total wage bill must be equal to the aggregate expenditure on all goods produced $wL_p + wL_e = EL$. Therefore, per-capita expenditure equals unity due to the choice of labor as the numeraire, i.e., $E = w = 1$. Substituting $E = 1$ in (7) yields the mass of products available for consumption M_c

$$M_c = (1 - \lambda_p^{-1})\frac{L}{f_p}. \tag{19}$$

Substituting[17] the relationship $M_c = (1 + n\zeta_x)M_p$ into (19) yields the mass of varieties produced in each country

$$M_p = \left[\frac{1 - \lambda_p^{-1}}{1 + n\zeta_x}\right]\frac{L}{f_p}. \tag{20}$$

Notice that in the absence of trade ($n = 0$) the mass of varieties produced equals the mass of varieties consumed, i.e., $M_c = M_p$.

As we mentioned in the introduction, the model generates an endogenous distribution of markups. Each incumbent firm charges a price equal to its quality level $p(\lambda) = \lambda$, and therefore its markup measured by the price marginal-cost margin as given by $(p - 1)/p = 1 - \lambda^{-1}$. Subsequently, the aggregate markup over all incumbents equals $PCM_p = \int_{\lambda_p}^{\infty}(1 - \lambda^{-1})M_p\mu(\lambda)d\lambda = M_p(1 - \tilde{\lambda}_p^{-1})$. Similarly, an exporter charges a price $p(\lambda) = \lambda$, incurs a marginal cost τ, and earns a price marginal-cost margin $(p - \tau)/p = 1 - \tau\lambda^{-1}$. Thus the aggregate export markup is $PCM_x = \int_{\lambda_x}^{\infty}(1 - \tau\lambda^{-1})M_p\mu(\lambda)d\lambda = M_x(1 - \tau\tilde{\lambda}_x^{-1})$. Consequently, the average production (domestic) and export markups are given by

$$pcm_p = 1 - \tilde{\lambda}_p^{-1} \qquad \text{and} \qquad pcm_x = 1 - \tau\tilde{\lambda}_x^{-1}, \tag{21}$$

where $\tilde{\lambda}_p$ and $\tilde{\lambda}_x$ are defined by (13). Both average markups increase in the production cutoff quality level λ_p (see Proposition 1), but the average export markup decreases in the level of trade costs τ.

Similar considerations apply to the derivation of the average domestic and export prices. Each surviving firm charges a price λ, and therefore, the aggregate domestic and export

[17] We can now derive a sufficient condition which guarantees that the consumer consumes all available varieties. In principle, each consumer chooses $q(\omega)$ (the quantity of each variety) and M (the number of varieties) to maximize (1) subject to the budget constraint. The first order condition with respect to M yields $\ln[\beta\lambda(M)q(M)/L] \geqslant 1$, which holds with equality if $M < M_c$. Inserting (2) into this condition yields $\ln[\beta\lambda(M)E/\{p(M)M\}] \geqslant 1$. Thus to ensure that the principle of non-satiation holds (i.e., $M = M_c$), we must have that $\ln[\beta\lambda(M_c)E/\{p(M_c)M_c\}] > 1$. Since $p(M_c) = \lambda(M_c)$ and $E = 1$, we must have $\ln(\beta/M_c) > 1$, that is $\beta > eM_c$, where e is the base of the natural logarithm. Using (19) implies that $\beta f_p/eL > 1 - \lambda_p^{-1}$. Moreover, because $\lambda_p \in (1, k)$, the inequality condition holds if $\beta f_p/eL > 1 - k^{-1}$, which in turn yields $\beta > eL/f_x$.

prices are given by $P_p = \int_{\lambda_p}^{\infty} \lambda M_p \mu(\lambda) d\lambda$ and $P_x = \int_{\lambda_x}^{\infty} \lambda M_p \mu(\lambda) d\lambda$, respectively. These expressions imply that the average domestic and export prices are given by $p_p = P_p/M_p = \int_{\lambda_p}^{\infty} \lambda g(\lambda) d\lambda / [1 - G(\lambda_p)]$ and $p_x = P_x/M_x = \int_{\lambda_x}^{\infty} \lambda g(\lambda) d\lambda / [1 - G(\lambda_x)]$, and increase in the corresponding cutoff quality levels (i.e., $\partial p_p / \partial \lambda_p > 0$, $\partial p_x / \partial \lambda_x > 0$).

One can obtain an expression for per capita welfare as follows. Substituting the per-capita quantity demanded $q(\lambda)/L = E/\lambda M_c$ for each variety into the utility function of a typical consumer (1) and performing the integration yields

$$U = M_c \ln \left(\frac{\beta E}{M_c} \right) = M_c \ln \left(\frac{\beta}{M_c} \right), \tag{22}$$

where per-capita expenditure E is set equal to unity due to the choice of labor as the numeraire. Since $\ln(\beta/M_c) > 1$ (see footnote 17), U is always positive. Observe that per-capita welfare depends positively on the mass of varieties consumed M_c and on the expenditure per variety E/M_c.

To unveil the intuition for the welfare expression notice that an increase in M_c has two conflicting welfare effects. First, each consumer becomes better off because she consumes more products. Second, she becomes worse off as her expenditure, which is equal to her wage, spreads among more varieties. Limit pricing renders the consumer indifferent between receiving q units of a generic product with quality 1 and q/λ units of a good with quality λ at a higher price $p(\lambda) = \lambda$, and thus the average quality does not appear directly as an argument in the welfare function. Higher average quality allows firms to reduce the demand for manufacturing labor by charging a higher price and producing less quantity per variety. This, in turn, means that the economy can afford the production of more varieties and enjoy a higher welfare level.

3 The Impact of International Trade

The model is well suited to analyze the general equilibrium effects of trade liberalization measured by an increase in the number of trading partners n, a reduction in per-unit trade costs τ, and a reduction in foreign-market entry costs f_x. These parameters capture a variety of forces including reductions in transportation and communication costs, reductions in trade barriers, and the formation of trading blocks (albeit in a highly stylized fashion given the assumption of structurally identical countries). The impact of trade liberalization is channeled through two interacting general-equilibrium channels: changes in the demand for labor, which are captured by changes in the real wage; and changes in the intensity of product-market competition, which are captured by changes in the average markup.

Formally, the effects of trade liberalization are transmitted through changes in the production cutoff quality level λ_p as are described in Lemma 2 (see Appendix for proof).

Lemma 2. *Trade liberalization, captured by an increase in the number of trading partners ($n \uparrow$), a reduction in per-unit transport costs ($\tau \downarrow$), or a reduction in foreign-market entry costs ($f_x \downarrow$), increases the production cutoff quality level λ_p (i.e., $d\lambda_p/dn > 0$, $d\lambda_p/d\tau < 0$, and $d\lambda_p/df_x < 0$).*

The economic intuition behind Lemma 2 is as follows. For any initial value of the production cutoff quality level λ_p, the export cutoff quality level λ_x and the mass of varieties consumed M_c are fixed (see equations (9) and (19)). Equation (9) and Lemma 1 imply that a decline in τ or f_x increases the average quality of exports $\tilde{\lambda}_x$. Therefore, any form of trade liberalization ($n \uparrow$, $\tau \downarrow$, or $f_x \downarrow$) increases the ex-ante profits $\bar{\pi}$ (see equation (15)) and raises the demand for labor for any wage level (see equation (16) and, in particular, equation (17)). The excess demand for labor induces a reallocation of resources from low-quality products towards high-quality products that translates into a larger mass of products available for consumption M_c. To see this, recall that for any level of expenditure, a firm with a higher quality product charges a higher price, produces less output, and employs less labor than a firm with a lower quality product (see equation (3)). Thus any given aggregate supply of labor can sustain more higher-quality products.

The reallocation of resources from lower to higher-quality products caused by the availability of more varieties intensifies the product-market competition by reducing the demand for each variety as the aggregate consumer income EL is spread among more varieties. Consequently, the zero-profit condition, which defines the production cutoff quality level, becomes negative when evaluated at the initial equilibrium, and requires a higher production cutoff quality level to be restored.

How does trade liberalization affect the export cutoff level λ_x? The following lemma answers this question (see Appendix for proof).

Lemma 3. *An increase in the number of trading partners ($n \uparrow$) increases the export cutoff quality level λ_x; whereas a reduction in per-unit transport costs ($\tau \downarrow$) or a reduction in foreign-market entry costs ($f_x \downarrow$) decreases λ_x (i.e., $d\lambda_x/dn > 0$, $d\lambda_x/d\tau > 0$, and $d\lambda_x/df_x > 0$).*

To get intuition behind these results, first notice that the zero-profit cutoff condition implies that $(1 - \tau\lambda_x^{-1})L/M_c = f_x$, where $E = 1$ is the equilibrium value. Obviously, an increase in n raises the number of available varieties M_c (from Lemma 2 and (19)), this reduces the profits of the marginal exporter, and therefore it requires an increase in the cutoff quality level λ_x to restore profits back to the equilibrium level. In other words, trade liberalization generates an income-based adverse demand effect which reduces the profits of all surviving firms including the marginal firms that earn zero profits and produce products with the cutoff quality level. The zero-profit conditions that define the cutoff quality levels are restored only if the cutoff quality levels rise.

A decline in (variable or fixed) trade costs has two effects: It increases profits from exports (direct supply-based effect) and this requires a decline in the cutoff level to restore profits back to zero; and increases the number of available varieties (the indirect income-based demand effect as described above) which tends to reduce profits and requires an increase in the cutoff quality level to bring profits back to zero. The direct effect dominates (which is desirable for stability purposes), and consequently, lower foreign-market entry costs or lower trade costs generate lower export cutoff quality level λ_x.

The next step of the analysis is to establish the impact of trade liberalization on prices, markups, and welfare. Lemma 2 and equation (21) ensure that any type of trade liberalization increases the production cutoff quality level, and subsequently raises the average domestic price and average domestic markup. The effects of trade liberalization on the average export price depend on the nature of trade liberalization: an increase in the number of trading partners raises the average export price, whereas a reduction in per-unit transport costs or a reduction in foreign-market export costs decreases the average export price (observe that the average export price increases in the export cutoff quality level λ_x, and see Lemma 3). Similar consideration apply to the effects of trade liberalization on the average export markup (see Lemma 2 and equation (21)): an increase in the number of trading partners raises the average export markup; a reduction in foreign-market entry costs reduces the average markup; and a decline in per-unit trade cost has an ambiguous impact on the effective marginal costs of exporting $\tau/\tilde{\lambda}_x$ and the average export markup.[18] In addition, inspection of equation (19) and Lemma 2 establish that any form of trade liberalization increases the production cutoff quality level λ_p and the mass of varieties available for consumption M_c in each country. Finally, differentiating equation (22) with respect to varieties consumed yields $\partial U/\partial M_c = \ln(\beta/M_c) - 1 > 0$, where the inequality follows from the non-satiation assumption that requires a sufficiently large value of parameter β (see 17). Thus trade liberalization raises national and global welfare. The following proposition summarizes these results.

Proposition 3. *Trade liberalization*

a. increases the average domestic price ($p_p \uparrow$) and the average domestic markup ($pcm_p \uparrow$);

b. has a differential effect on the average export price ($p_x \uparrow\downarrow$): an increase in the number of trading partners raises the average export price, whereas a reduction in per-unit transport costs or a reduction in foreign-market export costs decreases the average export price;

[18] Differentiating equation (21) with respect to trade costs τ yields $\partial pcm_x/\partial\tau = -\tilde{\lambda}_x^{-1} + \tau\mu(\lambda_x)(\lambda_x^{-1} - \tilde{\lambda}_x^{-1})d\lambda_x/d\tau$, where $d\lambda_x/d\tau > 0$. The second term of this expression is positive implying an ambiguous effect of trade costs on the average export markup.

c. has a differential effect on the average export markup (pcm_x ↑↓): an increase in the number of trading partners raises the average export markup; a reduction in foreign-market-entry costs lowers the average markup; and a decline in per-unit trade cost has an ambiguous effect on the average export markup;

d. raises the mass of products available for consumption (M_c ↑), and increases the levels of national and global welfare (U ↑).

Next consider an extreme form of trade liberalization: the move from autarky to (restricted) trade. The closed-economy steady-state equilibrium corresponds to the case of no trading partners (i.e., $n = 0$). Equation (17) then implies that, under autarky, the production cutoff quality level λ_p^A is determined by $H(\lambda_p, 1) = \delta f_e / f_p$, and is strictly less than the open-economy cutoff quality level λ_p : the absence of export markets reduces the benefits of entry and shifts labor from the production of higher-quality products towards the production of lower-quality products. Observe that equations (19), (21), and (22) determine the closed-economy values the number of varieties consumed ($M_p^A = M_c^A = [1 - (\lambda_p^A)^{-1}]L/f_p$), the average markups ($pcm_p^A = 1 - (\tilde{\lambda}_p^A)^{-1}$), and the level of welfare ($U^A = M_c^A \ln(\beta / M_c^A)$) which are all functions of the production cutoff quality level. Therefore, a move from autarky to trade has the same qualitative impact as an increase in the number of trading partners n. These effects are summarized in the following proposition.

Proposition 4. *A move from autarky to trade*

a. increases the production cutoff quality level ($\lambda_p > \lambda_p^A$);

b. generates higher markups ($pcm_p > pcm_p^A$);

c. raises the mass of products available for consumption ($M_c > M_c^A$);

d. and has a positive effect on national and global welfare ($U > U^A$).

As in Melitz's (2003) analysis, trade liberalization induces entry of better firms into foreign markets, forces firms with low-quality products to exit, and expands the number of varieties consumed. The model's prediction that trade liberalization increases industry markups in markets with Bertrand competition and vertical product differentiation is based on a novel mechanism that complements the work of Melitz and Ottaviano (2008). In their model, trade liberalization intensifies the product market competition via a change in the price elasticity of demand, whereas in the present model trade expands the mass of varieties and reduces the income spent on each product without affecting the price elasticity of demand. In Melitz and Ottaviano (2008) the effects of trade liberalization on prices and markups are uniform across all facets of trade liberalization: trade reduces both the average price and markup by increasing the price elasticity of demand for a typical variety. In contrast, in the present model the effects of trade liberalization on prices and markups depend on its nature. For instance, an increase in the number of trading partners (e.g., economic integration) raises the average domestic and export markups and prices.

4 Welfare Properties

The presence of quality-heterogeneous firms raises the following welfare question: does the market provide the socially optimal quality cutoff levels, markups, and the mass of available varieties? The constrained optimality (in the Dixit and Stiglitz (1977) sense, where firms make non negative profits) of the cutoff productivity levels has been demonstrated in a multi-sector version of the Melitz model by Feenstra and Kee (2008). This is a surprising result considering the presence of potential welfare distortions caused by imperfect competition, trade costs, and heterogeneous productivity levels.

Contrary to the existing literature on trade with heterogeneous firms, the present model generates the possibility of divergence between the socially optimal and laissez-faire equilibrium due to the endogenous distribution of markups. In order to minimize the algebra, we illustrate this possibility in a closed-economy setting noting that similar considerations apply to an open economy, where trade costs create additional welfare distortions. The social planner maximizes per-capita utility function $U = M_c \ln(\beta E/M_c)$ subject to the resource constraint. Since labor is the only factor of production, the resource constraint requires that at each instant in time labor is fully employed. The aggregate supply of labor L is fixed, whereas the aggregate demand for labor consists of two components: labor employed by potential entrants $L_e = \delta M_c f_e/[1-G(\lambda_p)]$, where the mass of varieties produced equals the mass of varieties consumed ($M_p = M_c$); and labor employed in production. The latter is derived as follows. Since one unit of labor produces one unit of output, each surviving firm hires $f_p + q(\lambda)$ workers, where f_p is the production fixed cost expressed in units of labor and $q(\lambda)$ is the demand for a product of quality λ given by (2). Thus the labor employed in production is $L_p = \int_{\lambda_p}^{\infty}[f_p + q(\lambda)]M_c\mu(\lambda)d\lambda = f_p M_c + \tilde{\lambda}_p^{-1}EL$. Consequently the resource constraint is

$$\frac{\delta M_c f_e}{1-G(\lambda_p)} + f_p M_c + \tilde{\lambda}_p^{-1}EL = L. \tag{23}$$

The social planner maximizes $U = M_c \ln(\beta E/M_c)$ subject to (23) with respect to the production cutoff quality level λ_p, per-capita expenditure E, and the mass of varieties available for consumption M_c. Denoting with ψ the Lagrangian multiplier, one can write the corresponding first-order conditions as follows.

$$(\lambda_p^{-1} - \tilde{\lambda}_p^{-1})(E/M_c)L = \delta f_e/[1-G(\lambda_p)], \tag{24}$$

$$1/\psi = \tilde{\lambda}_p^{-1}(E/M_c)L, \tag{25}$$

$$(1/\psi)[\ln(\beta E/M_c) - 1] - f_p = \delta f_e/[1-G(\lambda_p)]. \tag{26}$$

These conditions determine the socially optimum values of λ_p, E, and M_c. Equation (25) implies that the Lagrangian multiplier must be positive ($\psi > 0$).

We can obtain further insights on the solution to the social planner's problem by substituting (25) into (26) and equating the resulting expression to (24) to obtain

$$\left[\tilde{\lambda}_p^{-1}\ln(\beta E/M_c) - \lambda_p^{-1}\right](E/M_c)L = f_p. \tag{27}$$

Since the average quality level $\tilde{\lambda}_p$ is a function of the cutoff quality level λ_p, equations (24) and (27) determine the socially optimal values of λ_p^S and E^S/M_c^S. Once these two endogenous variables are determined, the resource constraint (23), or the assumption $E^S = 1$, can be used to obtain individual values for E^S and M_c^S.

How does the socially optimum solution compare to the closed-economy market equilibrium? Set the number of trading partners n equal to zero in (15) and substitute the resulting expression for $\bar{\pi}$ and f_p from condition (7) into (16) to obtain the closed-economy free-entry condition which is identical to (24). In other words, for any level of E/M_c the laissez-faire cutoff quality level λ_p^A is socially efficient. This result, which is identical to the ones obtain by Feenstra and Kee (2008), holds because the cutoff quality level does not appear as an argument in the social planner's objective function, but only in the resource constraint. Thus, the social planner chooses the cutoff quality level based on efficient resource allocation considerations. In the case of Dixit and Stiglitz (1977) preferences and exogenous markups, the welfare distortions associated with a marginal increase in the mass of varieties just happen to cancel each other out. However, this is not the case in the present model. To see this, note that the market equilibrium values λ_p^A and E^A/M_c^A are simultaneously determined by the entry condition (24) and the zero-profit condition (7). The latter is reproduced below for illustrative purposes

$$[1 - \lambda_p^{-1}](E/M_c)L = f_p. \tag{28}$$

Comparing the zero-profit condition (28) to the socially-optimal condition (27) reveals the following. The market cutoff quality level depends on profitability considerations that are captured by the marginal markup $1 - \lambda_p^{-1}$. This differs from the efficient "markup" $\tilde{\lambda}_p^{-1} \ln(\beta E/M_c) - \lambda_p^{-1}$ in (27) which depends on the ratio between the marginal utility derived from an additional variety and the average quality of available products. Observe that the social planner and the market assign the same total cost to the product with the cutoff quality level, namely $f_p + \lambda_p^{-1}(E/M_c)L$. However, there is a divergence between the benefits associated with the introduction of a new variety. The market cares about the total revenue earned $\lambda_p[EL/(M_c\lambda_p)] = EL/M_c$, where the term in square brackets equals the quantity demanded and λ_p is the corresponding limit price. In contrast, equation (27) reveals that the social benefits of introducing an additional product equal $\{\ln(\beta E/M_c)\}[EL/M_c\tilde{\lambda}_p]$, where the term in curly brackets is the change in consumer surplus (welfare) associated with the introduction of a new variety (for any given expenditure per variety E/M_c), and the term in square brackets represents the quantity demanded for a product with average (as opposed to marginal) quality.

In other words, the planner cares about the average (infra-marginal) consumer, whereas the market cares about the marginal consumer. On one hand, because the marginal quality level is less than the average quality level ($\lambda_p < \tilde{\lambda}_p$), the market has a tendency to overstate the benefits of introducing a new variety by charging a lower "price" and producing a higher quantity than the social planner. On the other hand, because $\ln(\beta E/M_c) > 1$ and $\lambda_p > 1$ the ranking between the social and market prices is in general ambiguous. Consequently, there is a divergence between the market and social valuation of the marginal benefits associated with an introduction of an additional variety, and inspection of equations (23), (24), (27) and (28) yields the following proposition.

Proposition 5. *The laissez-faire cutoff quality level λ_p^A, mass of varieties M_c^A, and per-capita expenditure E^A are socially sub-optimal.*

How does the market solution differ from the socially optimum solution? Figure 1 illustrates the market and efficient values of λ_p and E/M_c by plotting the graphs of equations (24), (27), and (28) under the assumption that quality levels are drawn from a Pareto distribution with scale parameter b and shape parameter $\kappa > 0$,

$$G(\lambda) = 1 - \left(\frac{b}{\lambda}\right)^{\kappa} \qquad \text{for } \lambda \geqslant b > 0. \tag{29}$$

Assuming that quality levels follow a Pareto distribution is not essential in our analysis. However, it makes the analysis more tractable. For example, using this distribution function in (13) yields $\tilde{\lambda}_i/\lambda_i = (\kappa + 1)/\kappa$, for $i = p, x$.

Equation (24) defines a direct relationship between E/M_c and λ_p which is illustrated

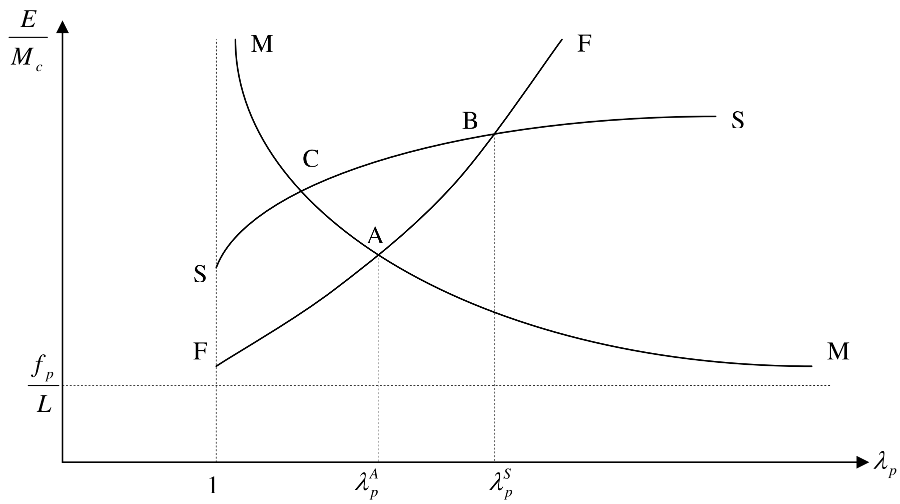

FIGURE 1. Efficient and Market Cutoff Quality Levels

with the convex FF curve in Figure 1.[19] Equation (28) defines an inverse relationship between E/M_c and λ_p, which is illustrated with the negatively-sloped MM curve.[20] The unique intersection between the FF and MM curves at point A determines the market cutoff quality level λ_p^A. The graph of equation (27) is illustrated with the concave SS curve.[21] More interestingly, at any intersection point of the FF and SS curves, the slope of the FF curve is greater than that of the SS curve.[22] Thus, the most the FF and SS curves can intersect each other is once.[23] We assume that the parameters of the model are such that the FF and SS curves intersect at point B, at which the socially optimum cutoff level is greater

[19] With the above cumulative distribution function, equation (24) becomes $E/M_c = (\kappa+1)\delta f_e \lambda_p^{\kappa+1}/(Lb^\kappa)$.

[20] Equation (28) can be written as $E/M_c = f_p/[(1-\lambda_p^{-1})L]$, which implies that E/M_c is decreasing in λ_p, $E/M_c \to \infty$ as $\lambda_p \to 1$, and $E/M_c \to f_p/L$ as $\lambda_p \to \infty$.

[21] With the Pareto distribution defined by (29), equation (27) becomes $[\kappa \ln(\beta E/M_c) - (\kappa+1)](E/M_c) = (\kappa+1)\lambda_p f_p/L$. Totally differentiating this equation yields $d(E/M_c)/d\lambda_p = (\kappa+1)f_p[L\kappa\ln(\beta E/M_c) - L]^{-1} > 0$, where the last inequality follows from the facts that $\kappa\ln(\beta E/M_c) - (\kappa+1) > 0$ and $\kappa > 0$. Differentiation of this slope with respect to λ_p implies that $d^2(E/M_c)/d\lambda_p^2 = -\kappa\,[d(E/M_c)d\lambda_p]^2\,\{(E/M_c)[\kappa\ln(\beta E/M_c) - 1)]\}^{-1} < 0$.

[22] To see this, notice that the slope of the FF curve is given by $d(E/M_c)/d\lambda_p = (\kappa+1)(E/M_c)/\lambda_p$. The slope of the SS curve, on the other hand, can be rewritten as $d(E/M_c)/d\lambda_p = (\kappa+1)(f_p/L)(E/M_c)[(\kappa+1)(f_p/L)\lambda_p + \kappa(E/M_c)]^{-1}$. It then easily follows that at any intersection $(\lambda_p, E/M_c)$, $(\kappa+1)(E/M_c)/\lambda_p > (\kappa+1)(E/M_c)(f_p/L)/[(\kappa+1)(f_p/L)\lambda_p + \kappa(E/M_c)]$.

[23] The slope of the SS curve is generally indeterminate under other distribution functions. In this case, the SS curve may intersect the FF curve more than once, and hence, might give rise to *multiple* socially optimum cutoff quality levels.

than the market cutoff level, i.e., $\lambda_p^S > \lambda_p^A$.[24]

Figure 1 can be used to perform more comparative statics exercises. An increase in production fixed cost f_p does not affect the position of the FF curve, shifts the MM curve to the right raising λ_p^A and E^A/M_c^A; while it shifts the SS curve to the left raising λ_p^S and E^S/M_c^S. Intuitively, an economy facing higher production fixed costs allocates a larger share of resources to the creation of more varieties than to the production of low-quality products. An increase in market size, measured by the labor endowment L, shifts the FF and SS curves to the right and the MM curve to the left (not shown in Figure 1). The magnitude of the FF curve shift equals the magnitude of the MM shift and exceeds the corresponding magnitude of the SS curve shift.[25] Therefore, an increase in market size L lowers E/M_c, raises the efficient cutoff quality level λ_p^S, and does not affect the market cutoff quality level λ_p^A.

5 Cost Heterogeneity

This section augments the simple framework by adding quality-based cost heterogeneity and spatial cost considerations to the original model. We assume that higher-quality varieties require higher manufacturing costs; that trade costs increase with geographic distance; and that the costs of imitating a product increase in the level of product quality. In the modified version of the model, firms with higher-quality products enjoy higher market shares, sell their products in more markets and charge higher prices in far-away markets.

We introduce spatial elements by considering a global economy with four countries

[24] Equation (27) implies that $\ln(\beta E/M_c) > (\kappa+1)/\kappa$, i.e., $E/M_c > \exp(1+1/\kappa)/\beta$. At $\lambda_p = 1$, equation (24) yields $E/M_c = (\kappa+1)\delta(f_e/L)b^{-\kappa}$. Given that $\delta < 1$ and $f_e/L < 1$, unless b is too low and κ is too high, as λ_p approaches 1 the SS curve intercept is higher than the FF curve intercept. In depicting Figure 1, we assume that these conditions are satisfied. Also notice that when k increases the FF curve becomes steeper (assuming that $b \leqslant 1$). In this case, the FF curve may intersect the SS curve at a point to the left of point C in Figure 1, which implies that the socially optimum cutoff level is smaller than the market cutoff level λ_p^A.

[25] Totally differentiating equations (24), (28) and (27) yields the marginal shifts of the FF curve $d(E/M_c)/dL = -(E/M_c)/L$, the MM curve $d(E/M_c)/dL = -(E/M_c)/L$, and the SS curve $d(E/M_c)/dL = -[(E/M_c)/L]\left\{[\tilde{\lambda}_p^{-1}\ln(\beta E/M_c) - \lambda_p^{-1}]/[\tilde{\lambda}_p^{-1} + \tilde{\lambda}_p^{-1}\ln(\beta E/M_c) - \lambda_p^{-1}]\right\}$. The term in curly brackets is positive, less than one, and implies that the magnitude of the SS curve shift is less than that of the FF curve.

located in a circle. Each country faces two adjacent (neighboring) countries and a distant country. Thus, if home is indexed by 0 it has two neighbors (1 and 3), country 1 has two neighbors (1 and 2), and so on. Let $\phi\tau$ denote the level of trade costs between two adjacent countries, and $\phi^2\tau$ denote the level of trade costs between two distant countries, where $\phi > 1$.

Assume that the marginal cost of producing a variety with λ quality is $c(\lambda) = \lambda^\alpha$. In addition, suppose that a firm's ability to imitate a product depends on geographic distance. Specifically, we assume that the marginal costs of imitating a product with quality λ is $c(\lambda) = \lambda^\alpha$, if the imitator resides in the same country where the product is discovered; $\phi\lambda^\alpha$, if a potential imitator resides in an adjacent country; and $\phi^2\lambda^\alpha$, if a potential imitator resides in the distant country.[26] The original model can be recovered by setting $\phi = 1$, $\alpha = 0$, and considering n instead of three trading partners.

Given these assumptions, a home firm that discovers a variety with quality λ follows a limit-pricing strategy and charges a price $p_d(\lambda) = \lambda^{1+\alpha}$ in the domestic market, $p_a(\lambda) = \phi\lambda^{1+\alpha}$ in an adjacent-country market; and $p_f(\lambda) = \phi^2\lambda^{1+\alpha}$ in the distant-country market. Therefore, firms with higher quality products charge higher prices in each market; and conditional on the level of quality, each firm's price increases with distance from the origin of production.

Cost heterogeneity implies that per-period profits of firms that export to all three countries can be decomposed into three parts: profits from domestic sales $\pi_d(\lambda)$; profits from exports into each of the two adjacent countries $\pi_a(\lambda)$; and profits from exports into the distant market $\pi_f(\lambda)$.

[26] We still assume that the technology to produce generic products diffuses internationally through imports. The present way of modeling imitation captures the idea that the degree of technology transfer declines with distance. See, for instance, Keller (2002) for evidence on the local nature of international knowledge spillovers.

$$\pi_d(\lambda) = [p_d(\lambda) - \lambda^\alpha]q_d(\lambda) - f_p = (1 - \lambda^{-1})\frac{EL}{M_c} - f_p, \quad (30)$$

$$\pi_a(\lambda) = [p_a(\lambda) - \phi\tau\lambda^\alpha]q_a(\lambda) - f_a = (1 - \phi\tau\lambda^{-1})\frac{EL}{M_c} - f_a, \quad (31)$$

$$\pi_f(\lambda) = [p_f(\lambda) - \phi^2\tau\lambda^\alpha]q_a(\lambda) - f_f = (1 - \phi^2\tau\lambda^{-1})\frac{EL}{M_c} - f_f, \quad (32)$$

where subscripts a and f denote adjacent and distant (far-away) markets.

According to equations (30)–(32), the flow of profits in each market is an increasing and monotonic function of quality λ. Because $\pi_d(1) = -f_p < 0$, $\pi_d(\phi\tau) = -f_a < 0$, and $\pi_f(\phi^2\tau) = -f_f < 0$, there exist three cutoff quality levels λ_p, λ_a, and λ_f such that $\pi_d(\lambda_p) = \pi_f(\lambda_f) = \pi_f(\lambda_f) = 0$. These ex-post zero-profit conditions lead to the following expressions:

$$\lambda_a = \frac{\phi\tau}{1 - (1 - \lambda_p^{-1})(f_a/f_p)} \qquad \text{and} \qquad \lambda_f = \frac{\phi^2\tau}{1 - (1 - \phi\tau\lambda_a^{-1})(f_f/f_a)}. \quad (33)$$

Several remarks are in order. First, as in the simple framework, the requirement that the denominator of the first equation in (33) must be non-negative implies that $\lambda_p \in (1, k)$, where $k = 1/(1 - f_p/f_a)$. Second, equations in (33) indicate that λ_a and λ_f are increasing monotonically in λ_p. Third, under the standard parameter restrictions $f_f > f_a > f_p > 0$ (i.e., foreign market entry costs increase with distance from home market), $\tau > 1$, and $\phi > 1$, we have $\lambda_f > \lambda_a > \lambda_p > 1$. Therefore the set of firms in each country is partitioned into three groups: firms whose product quality level is less than λ_p exit the market; firms whose product quality level is $\lambda \in [\lambda_p, \lambda_a)$ produce exclusively for the domestic market; firms whose product quality is $\lambda \in [\lambda_a, \lambda_f)$ serve the domestic and each of the two adjacent markets; and firms producing products with $\lambda \geqslant \lambda_f$ sell their products in all markets.

The ex-post distribution of product quality levels μ is given by (10). The probabilities that a surviving firm exports to an adjacent and distant countries respectively are given by

$$\zeta_a = \frac{1 - G(\lambda_a)}{1 - G(\lambda_p)} \qquad \text{and} \qquad \zeta_f = \frac{1 - G(\lambda_f)}{1 - G(\lambda_p)}. \quad (34)$$

Thus, the number (measure) of home firms serving a particular market decreases in distance ϕ.

It is also straightforward to demonstrate that the ex-ante per-period flow of profits of a typical entrant is given by

$$\bar{\pi} = (1 - \tilde{\lambda}_p^{-1})\frac{EL}{M_c} - f_p + 2\zeta_a \left[(1 - \phi\tau\tilde{\lambda}_a^{-1})\frac{EL}{M_c} - f_a\right] + \zeta_f \left[(1 - \phi^2\tau\tilde{\lambda}_f^{-1})\frac{EL}{M_c} - f_f\right]. \quad (35)$$

where $\tilde{\lambda}_i$ represents the weighted harmonic mean of quality levels in each market.

In the presence of cost heterogeneity, the free-entry condition is given by

$$f_p H(\lambda_p, 1) + 2 f_a H(\lambda_a, \phi\tau) + f_f H(\lambda_f, \phi^2\tau) = \delta f_e, \quad (36)$$

where $H(\cdot)$ is defined by (18) and is strictly decreasing in its first argument (see Appendix D). Equations in (33) and (36) determine the unique values of the three cutoff quality levels and can be used to perform the standard comparative statics exercises.

The combination of limit-pricing strategies, Cobb-Douglas preferences and cost heterogeneity generates heterogeneous market shares, prices, and markups consistent with the findings of several recent empirical studies (Baldwin and Harrigan, 2007; Verhoogen, 2008; Kugler and Verhoogen 2008; Johnson, 2009; and Manova and Zheng, 2009). In this modified model, firms with higher-quality products face higher marginal cost of manufacturing; charge higher prices in more distant markets; and enjoy higher markups, larger revenues and market shares.[27] More firms export to more proximate markets, and sufficiently high foreign market entry costs eliminate bilateral trade flows between distant markets.[28]

6 Conclusion

The present paper developed a highly-tractable model of quality heterogeneity and international trade. Firms in our model face Cobb-Douglas preferences and charge prices that are proportional to the quality level of their products as in Schumpeterian growth models.

[27] All our results go through in the case where markets are interpreted as distinct geographic regions within a particular country and trade costs as transportation costs. This interpretation is consistent with the existence of heterogeneous prices, markups and market shares across firms within a particular country.

[28] The second equation in (33) implies that, for a sufficiently high value of foreign market entry cost f_f, the cutoff quality level λ_f becomes infinite. In this case, there is no bilateral trade between two distant countries.

The production side of the model mimics the model of productivity heterogeneity proposed by Melitz (2003). In our model, firms with high-quality products export, firms with intermediate-quality products produce for the domestic market, and firms with low-quality products exit the market. This trade pattern is consistent with the findings of several empirical studies and the Alchian-Allen conjecture. Firms producing high-quality products charge high prices and enjoy high markups, whereas firms producing low-quality products charge low prices and charge low markups.

The model generates several novel insights that complement the existing studies on the effects of trade in the presence of heterogeneous firms. First, the distribution of markups, the mass of varieties produced and consumed, the real wage, and cutoff quality levels are all endogenous. Second, the effects of trade liberalization are channeled through demand-based forces that operate through the expansion of the mass of available varieties, and supply-based forces that operate through changes in the real wage. Trade liberalization (or a move from autarky to restricted trade) benefits more firms producing higher-quality products, forces inefficient firms to exit, and raises the average domestic markup. The effects of trade liberalization on average export markups are in general ambiguous and depend on the nature of trade liberalization. Third, trade liberalization increases the mass of varieties available for consumption and raises national and global welfare. Fourth, despite the positive welfare effect of trade liberalization, there is more room for further welfare improvements because the laissez-faire equilibrium is suboptimal. The main reason for the second-best nature of laissez-faire equilibrium is the ambiguous welfare ranking between the market and socially optimal markups, thanks to the possibility of imitation that forces firms to use entry-deterring, limit-pricing strategies. Therefore, it seems that the present model is better suited to analyze the nexus of trade and firm-specific heterogeneity in global high-tech markets, where entry deterring strategies and the threat of imitation are more prevalent. Finally, our framework can be easily augmented to incorporate cost heterogeneity and spatial elements. In the modified version of the model, firms with higher-quality products enjoy higher market shares, sell their products in more markets and charge higher prices in far-away markets. These predictions are consistent with the findings of several recent empirical studies.

References

Alchian, Armen A. and William R. Allen, *University Economics*, Belmont, CA: Wadsworth Pub. Co., 1964.

Antoniades, Alexis, "Heterogeneous Firms, Quality and Trade," 2008. Mimeo, Columbia University.

Aw, Bee Yan, Sukkyun Chung, and Mark J. Roberts, "Productivity and Turnover in the Export Market: Micro Evidence from the Republic of Korea and Taiwan (China)," *The World Bank Economic Review*, 2000, *14*, 65–90.

Baldwin, Richard E. and James Harrigan, "Zeros, Quality and Space: Trade Theory and Trade Evidence," 2007. NBER Working paper, No:13214.

Bernard, Andrew B. and J. Bradford Jensen, "Exceptional Exporter Performance: Cause, Effect, or Both," *Journal of International Economics*, 1999, *47*, 1–25.

——, Jonathan Eaton, J. Bradford Jensen, and Samuel Kortum, "Plants and Productivity in International Trade," *American Economic Review*, September 2003, *93*, 1268–90.

Clerides, Sofronis K., Saul Lach, and James R. Tybout, "Is Learning by Exporting Important? Micro-Dynamic Evidence from Columbia, Mexico, and Morocco," *Quarterly Journal of Economics*, 1998, *113*, 903–47.

Dixit, Avinash K. and Joseph E. Stiglitz, "Monopolistic Competition and Optimum Product Diversity," *American Economic Review*, 1977, *67*, 297–308.

Feenstra, Robert C., *Advanced International Trade: Theory and Evidence*, Princeton, NJ: Princeton University Press, 2004.

—— and Hiau L. Kee, "Export Variety and Country Productivity: Estimating the Monopolistic Competition Model with Endogenous Productivity," *Journal of International Economics*, 2008, *74*, 500–18.

Grossman, Gene M. and Elhanan Helpman, "Quality Leaders in the Theory of Growth," *Review of Economic Studies*, January 1991, *58*, 43–61.

Hallak, Juan C., "Product Quality and the Direction of Trade," *Journal of International Economics*, 2006, *68*, 238–65.

Harrison, Ann, "Productivity, Imperfect Competition, and Trade Reform: Theory and Evidence," *Journal of International Economics*, 1994, *36*, 53–73.

Helpman, Elhanan, "Trade, FDI, and the Organization of Firms," *Journal of Economic Literature*, 2006, *44*, 589–630.

Hummels, David and Alexandre Skiba, "Shipping the Good Apples Out? An Empirical Confirmation of the Alchian-Allen Conjecture," *Journal of Political Economy*, 2004, *112*, 1384–1402.

—— and Peter J. Klenow, "The Variety and Quality of a Nation's Exports," *American Economic Review*, 2005, *95*, 704–23.

Johnson, Robert J., "Trade and Prices with Heterogeneous Firms," 2009. Mimeo, Dartmouth College.

Keller, Wolfgang, "Geographic Localization and International Technology Diffusion," *American Economic Review*, March 2002, *92*, 120–42.

Krugman, Paul, "Increasing Returns, Monopolistic Competition, and International Trade," *Journal of International Economics*, 1979, *9*, 469–80.

Kugler, Maurice and Eric Verhoogen, "The Quality Complementarity Hypothesis: Theory and Evidence from Columbia," 2008. NBER Working Paper, 14418.

Manova, Kalina and Zhiwei Zhang, "Export Prices and Heterogeneous Firm Models," 2009. NBER Working Paper, 15342.

Melitz, Marc J., "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, November 2003, *71*, 1695–1725.

—— and Gianmarco I. Ottaviano, "Market Size, Trade, and Productivity ," *Review of Economic Studies*, 2008, *75*, 295–316.

Roberts, Mark J. and James R. Tybout, "The Decision to Export in Columbia: An Empirical Model of Entry with Sunk Costs," *American Economic Review*, September 1997, *87*, 545–64.

Schott, Peter K., "Across-Product Versus Within Product Specialization in International Trade," *Quarterly Journal of Economics*, 2004, *119*, 647–78.

Segerstrom, Paul, T. C. A. Anant, and Elias Dinopoulos, "A Schumpeterian Model of the Product Life Cycle," *American Economic Review*, 1990, *80*, 1077–92.

Siminski, Peter, "The Price Elasticity of Demand for Pharmaceuticals amongst High Income Older People in Australia: A Natural Experiment," 2008. University of Wollongong Working Paper, No: 185.

Simonovska, Ina, "Income Differences and Prices of Tradables," 2009. Mimeo, University of California at Davis.

Taylor, Scott M., "Quality Ladders and Ricardian Trade," *Journal of International Economics*, 1993, *34*, 225–43.

Tellis, Gerard J., "The Price Elasticity of Selective Demand: A Meta-Analysis of Econometric Models of Sales," *Journal of Marketing Research*, November 1988, *25*, 331–41.

Verhoogen, Eric A., "Trade, Quality Upgrading, and Wage Inequality in the Mexican Manufacturing Sector," *Quarterly Journal of Economics*, 2008, *123*, 489–530.

Appendix

A. Limit Pricing

Consider the following utility function, which is a generic version of the preferences used routinely in quality-ladders growth models, that describes the tastes of a typical consumer

$$U = \int_{\omega \in \Omega} \ln \left[\beta \frac{q_0}{L} + \beta \lambda(\omega) \frac{q(\omega)}{L} \right] d\omega,$$

where Ω denotes the set of *potential* varieties. Each variety is associated with two quality levels: $q(\lambda(\omega))$ that can be produced by only one firm as in our model, and the low-quality (generic) version q_0 that can be produced by a competitive fringe after the original product has been produced. We assume that the low-quality version can be produced under constant marginal costs equal to the wage and perfect competition. This implies that the generic product commands a price $p_0 = w = 1$. In addition, since both goods adjusted for quality are identical, a consumer spends her income only on the good with the lower quality adjusted price. Thus, if a consumer buys only the generic good q_0, she obtains a utility level equal to $\ln(\beta E/[p_0 M_c L]) = \ln(\beta E/[M_c L])$, whereas if she buys the high-quality good she obtains a utility level equal to $\ln(\beta \lambda E/[p(\lambda) M_c L])$.

Therefore the consumer is indifferent between the two quality versions of the product if and only if $p(\lambda) = \lambda$. If the high-quality producer sets a limit price $p - \varepsilon$, where ε is infinitesimally small, she maximizes profits and drives the competitive fringe out of the market. Following the standard practice of quality-ladders growth models, we assume that even if $\varepsilon = 0$, all consumers buy the high-quality version of each product even if in principle they are indifferent between the two quality versions of each good.

B. Productivity Heterogeneity

Following the spirit of Taylor (1993), suppose that the utility function of the representative consumer is given by

$$U = \int_{\omega \in \Omega} \ln [\beta q(\omega)] \, d\omega.$$

The demand for a typical variety is $q(\omega) = EL/(p(\omega)M_c)$ as in the main text. Assume that firms discover new varieties associated with a productivity level $\lambda(\omega)$, and once the product is developed and sold in the market it can be produced by a competitive fringe with marginal and average costs equal to the wage $w = 1$. The marginal cost of a firm with productivity level $\lambda(\omega)$ is $1/\lambda(\omega)$. Consider the profit flow of a firm with productivity λ which charges the limit price $w - \varepsilon = 1 - \varepsilon$, where $\varepsilon \to 0$, to drive the competitive fringe out of the market. The profit flow is given by

$$\pi(\omega) = p(\omega)q(\omega) - (1/\lambda)q(\omega) - f_p = (1 - \lambda^{-1})\frac{EL}{M_c} - f_p,$$

which is identical to (4) in the main text. All firms charge the same limit price, but more productive firms enjoy lower average and marginal costs and higher markups and profits. In other words, under limit pricing exporters charge the same or higher prices than domestic firms, independently of quality or productivity heterogeneity.

C. Proof of Lemma 1

Substitute $\lambda_p^{-1} > \lambda^{-1}$ in the integral expression of equation (13) to obtain

$$\tilde{\lambda}_p^{-1} < \int_{\lambda_p}^{\infty} \lambda_p^{-1} g(\lambda) d\lambda / [1 - G(\lambda_p)] = \lambda_p^{-1},$$

which yields $\tilde{\lambda}_p > \lambda_p$. Differentiation of (13) yields $\partial\tilde{\lambda}_p/\partial\lambda_p = \tilde{\lambda}_p^2 \mu(\lambda_p)(\lambda_p^{-1} - \tilde{\lambda}_p^{-1}) > 0$. The same logic and calculations apply to $\tilde{\lambda}_x$ defined by (13). *Q.E.D.*

D. Proof of Proposition 2

We will prove that $H(\lambda_i, \alpha)$ defined in (18) is strictly decreasing in λ_i. Let $J(\lambda_i, \alpha) = [1 - G(\lambda_i)](1 - \alpha\tilde{\lambda}_i^{-1})$. Since $\lambda_i > \alpha$, it easily follows that $J(\lambda_i, \alpha) > 0$. Applying the Leibnitz rule in calculus, we have

$$dJ(\lambda_i, \alpha)/d\lambda_i = -(1 - \alpha\lambda_i^{-1})g(\lambda_i).$$

Differentiating $H(\lambda_i, \alpha)$ with respect to λ_i yields the desired result:

$$\frac{\partial H}{\partial \lambda_i} = \frac{(1-\alpha\lambda_i)dJ/d\lambda_i - \alpha\lambda_i^{-2}J}{(1-\alpha\lambda_i^{-1})^2} + g(\lambda_i) = -\alpha\lambda_i^{-2}J/(1-\alpha\lambda_i^{-1})^2 < 0.$$

The property $\partial H/\partial\lambda_i < 0$ in conjunction with equation (9) imply that the left hand side of (17) is strictly decreasing in the production cutoff quality level λ_i.

The full characterization of the solution requires the evaluation of the left-hand-side of (17) when λ_p approaches its boundaries. Equation (9) implies that at the lower bound of the production cutoff quality level ($\lambda_p = 1$) the export cutoff quality level equals the per-unit trade costs ($\lambda_x = \tau$), whereas as λ_p approaches its upper bound ($\lambda_p \to k$), λ_x approaches infinity ($\lambda_x \to \infty$). The following conditions characterize the limiting behavior of and $H(\lambda_p, 1)$ and $H(\lambda_x, \tau)$:

$$\lim_{\lambda_p \to 1} H(\lambda_p, 1) = \infty, \qquad \lim_{\lambda_x \to \tau} H(\lambda_x, \tau) = \infty, \tag{37}$$

$$\lim_{\lambda_p \to k} H(\lambda_p, 1) > 0, \qquad \lim_{\lambda_x \to \infty} H(\lambda_x, \tau) = 0. \tag{38}$$

To prove the claims in (37), first note that $J(\lambda_i, \alpha) = \int_{\lambda_i}^{\infty}[1-\alpha\lambda^{-1}]g(\lambda)d\lambda$. Inserting this expression into $H(\lambda_i, \alpha)$ yields

$$H(\lambda_i, \alpha) = \frac{\int_{\lambda_i}^{\infty}[1-\alpha\lambda^{-1}]g(\lambda)d\lambda}{1-\alpha\lambda_i^{-1}} - [1-G(\lambda_i)].$$

The numerator of the first term on the right hand side is always positive as long as $g(\cdot)$ is continuous at λ_i. The claims in (37) easily follow as we take the limits of both sides as $\lambda_p \to 1$ and $\lambda_x \to \tau$. Also, note that as $\lambda_x \to \infty$, the right hand side of the above equation approaches to 0, i.e., $\lim_{\lambda_x \to \infty} H(\lambda_x, \tau) = 0$. The first claim in (38) immediately follows from $H(k, 1) = [1-G(k)](k^{-1} - \tilde{k}^{-1})/(1-k^{-1}) > 0$ since $k^{-1} > \tilde{k}^{-1}$, where $\tilde{k} \equiv \left[\frac{1}{1-G(k)}\int_k^{\infty}\lambda^{-1}g(\lambda)d\lambda\right]^{-1}$.

In summary, the left-hand-side of (17) is always positive and defines a negatively sloped curve starting at infinity as $\lambda_p \to 1$ and reaching the value $H(k, 1)$ as $\lambda_p \to k$. The right-hand-side of (17) equals to $\delta f_e/f_p$ and it is independent of λ_p. Therefore, the existence of a unique cutoff quality level $\lambda_p > 1$ is guaranteed for a sufficiently high value of $\delta f_e/f_p$. Formally, a sufficient but hardly necessary condition for the existence of the unique equilibrium is that $H(k,1) < \delta f_e/f_p$. It is then obvious from equation (9) that the unique production cutoff quality level λ_p determines the unique export cutoff quality level λ_x such that $\lambda_x > \tau$ and $\lambda_x > \lambda_p$. *Q.E.D.*

E. Proof of Lemma 2

Totally differentiating equations (9) and (17) yields

$$\frac{d\lambda_p}{dn} = -H(\lambda_x,\tau)\left(\frac{f_x}{f_p}\right)\left[\frac{\partial H(\lambda_p,1)}{\partial\lambda_p} + \frac{n}{\tau}\left(\frac{f_x\lambda_x}{f_p\lambda_p}\right)^2 \frac{\partial H(\lambda_x,\tau)}{\partial\lambda_x}\right]^{-1} > 0, \tag{39}$$

$$\frac{d\lambda_p}{d\tau} = -n\left(\frac{f_x}{f_p}\right)\left[\frac{\lambda_x}{\tau}\frac{\partial H(\lambda_x,\tau)}{\partial\lambda_x} + \frac{\partial H(\lambda_x,\tau)}{\partial\tau}\right]\left[\frac{\partial H(\lambda_p,1)}{\partial\lambda_p} + \frac{n}{\tau}\left(\frac{f_x\lambda_x}{f_p\lambda_p}\right)^2 \frac{\partial H(\lambda_x,\tau)}{\partial\lambda_x}\right]^{-1} < 0, \tag{40}$$

$$\frac{d\lambda_p}{df_x} = [1-G(\lambda_x)]\frac{n}{f_p}\left[\frac{\partial H(\lambda_p,1)}{\partial\lambda_p} + \frac{n}{\tau}\left(\frac{f_x\lambda_x}{f_p\lambda_p}\right)^2 \frac{\partial H(\lambda_x,\tau)}{\partial\lambda_x}\right]^{-1} < 0. \tag{41}$$

The sign of the last bracket in each expression is negative due to Proposition 2 which also implies that the sign of the first bracket in (40) is negative. To see this note that the first term of that bracket can be written as

$$\frac{\lambda_x}{\tau}\frac{\partial H(\lambda_x,\tau)}{\partial\lambda_x} = -\frac{[1-G(\lambda_x)](1-\tau\tilde{\lambda}_x^{-1})}{\lambda_x(1-\tau\lambda_x^{-1})^2},$$

and that the second term in the same bracket is

$$\frac{\partial H(\lambda_x,\tau)}{\partial\tau} = [1-G(\lambda_x)]\left[\frac{\lambda_x^{-1}-\tilde{\lambda}_x^{-1}}{(1-\tau\lambda_x^{-1})^2}\right].$$

Adding these two expressions yields

$$\frac{\lambda_x}{\tau}\frac{\partial H(\lambda_x,\tau)}{\partial\lambda_x} + \frac{\partial H(\lambda_x,\tau)}{\partial\tau} = \frac{[1-G(\lambda_x)](\tau-\lambda_x)\tilde{\lambda}_x^{-1}}{\lambda_x(1-\tau\lambda_x^{-1})^2} < 0,$$

where the last equality follows from $\lambda_x > \tau$. Thus, the sign of (40) is negative. *Q.E.D.*

F. Proof of Lemma 3

First, totally differentiating (9) with respect to n yields

$$\frac{d\lambda_x}{dn} = \frac{1}{\tau}\left(\frac{\lambda_x}{\lambda_p}\right)^2 \frac{d\lambda_p}{dn} > 0,$$

where the inequality follows from (39) in the previous section.

Totally differentiating (17), on the other hand, yields

$$\begin{aligned} \frac{d\lambda_x}{d\tau} &= -\left[\frac{\partial H(\lambda_p, 1)}{\partial \lambda_p}\frac{d\lambda_p}{d\tau} + n\frac{f_x}{f_p}\frac{\partial H(\lambda_x, \tau)}{\partial \tau}\right]\left[n\frac{f_x}{f_p}\frac{\partial H(\lambda_x, \tau)}{\partial \lambda_x}\right]^{-1} > 0, \\ \frac{d\lambda_x}{df_x} &= -\left[\frac{\partial H(\lambda_p, 1)}{\partial \lambda_p}\frac{d\lambda_p}{df_x} + \frac{n}{f_p}H(\lambda_x, \tau)\right]\left[n\frac{f_x}{f_p}\frac{\partial H(\lambda_x, \tau)}{\partial \lambda_x}\right]^{-1} > 0. \quad \textit{Q.E.D.} \end{aligned}$$