

LOCATION, PRODUCTIVITY, AND TRADE

KEVIN WISEMAN*

JOB MARKET PAPER

ABSTRACT. Does trade liberalization increase competition? Location models provide a natural tool for answering this question, but, because of technical problems, their use in the trade literature has been limited. This paper develops a novel location model which is robust to arbitrary differences in productivity. In an otherwise standard general equilibrium trade environment, the model remains tractable under consumer preferences which are dramatically more general than those in traditional location models. I find that firms set variable markups which are increasing in firm productivity but which decline under a trade reform. Variable markups offer several improvements over the standard Dixit-Stiglitz framework with CES preferences, including the presence of small firms and measured gains in GDP under a trade reform. In this model trade liberalization reduces the total number of firms and increases their isolation in product space in a range of parameters that best explains firm data.

JEL Classification: F12, F15, L11, L16

Keywords: Location, Firm Heterogeneity, Variety, Trade Reform

Date: November 27, 2009.

I would like to thank Tim Kehoe, Fabrizio Perri, and Kim-Sau Chung for their help and encouragement. I am indebted to Costas Arkolakis, Ellen McGrattan, Larry Jones, Tom Holmes, and Chris Phelan for their helpful comments and suggestions. I would also like to thank participants in the Workshop On Dynamic Macroeconomics, Universidad de Vigo and the Trade and Development workshop at the University of Minnesota for many helpful comments. All remaining errors are mine.

* University of Minnesota and the Federal Reserve Bank of Minneapolis, wise0078@umn.edu.

1. INTRODUCTION

How does trade reform affect the competitiveness of a country's markets? Location models allow us to look at a competitive margin not captured by other models – product differentiation. Will firms respond to trade reform by further differentiating their products, or will new importers force them to compete against very similar goods? Location models offer a language to describe this competitive margin as well as other margins, such as price markups and average productivity in a unified framework.

In this paper I develop a location model with the possibility of sequential entry which mitigates the existence problems in more standard models. In equilibrium firms pick prices and locations to prevent high-cost firms from entering. Prices are the solution to a simple maximization problem which depends only on a cutoff marginal cost. This firm behavior keeps the model tractable under arbitrary heterogeneity in firm costs and a broad class of utility specifications.

In a general equilibrium trade model this market structure generates variable markups which offer several improvements over the standard Dixit-Stiglitz model with CES preferences. Without fixed costs, the model permits arbitrarily small exporters. Because markups vary with market conditions, measured GDP increases under a trade reform while it is constant in the standard model. Last, the model allows us to look at product similarity before and after a trade reform, a margin exclusive to location models. I find that in a range of parameters consistent with firm statistics the number of goods available in the domestic market will decline under a trade reform. The entry of new exporters is outweighed by the exit of domestic firms. In fact, the market appears *less* competitive after a trade reform from the perspective of product differentiation because these fewer firms have greater average distance between them.

In spite of the promise of location models, their use in the trade literature has been limited because they cannot handle firms with significantly different costs. Trade necessarily involves differences between domestic and foreign firms. Any model with countries in between autarky and complete market integration requires extra costs, e.g. trade barriers and transportation costs, associated with imported goods. Much recent work has also considered the role of within-country firm heterogeneity and the effect of trade reforms. These models reflect the wide variation in firm productivity in the data and consider the exit of unproductive domestic firms under a trade reform. In models with an active exit margin it is the small, marginally productive firms whose behavior determines changes in aggregate productivity.

In a seminal contribution, Vogel (2008) solves a location model with two stages where firms simultaneously pick locations and then simultaneously pick prices. He shows that there is a unique equilibrium with tractable policies for firms as long as firms do not vary too much in productivity¹. How big can these differences be? It is difficult to characterize the maximum difference in firm costs, but we can consider some illuminating examples. Among order-independent equilibria in which any firm may be next to any other, no firm can be more than twice as large as any other or the equilibrium will not exist.² This limitation prevents Vogel's model from realizing the degree of heterogeneity seen in empirical work, such as the widely cited Bartelsman and Doms (2000) finding that narrowly specified industries feature 85-15 ratios of firm productivities in the range of 2-4. Further, in order to have an active exit margin a heterogeneous firms model must have small firms earning small profits who exit as the market becomes marginally more competitive. These are exactly the types of firms which break down the equilibrium Vogel identifies.

In this paper I take a different tack. I describe a location and price setting game where firms may enter sequentially. In this sense the model follows in the tradition of sequential

¹The equilibrium is unique up to a rotation along the circle or permutation in firm order. All equilibria yield identical firm prices, demands, and profits.

²To express this difference in terms of productivity would require more notation, but a quick calculation is given in the appendix. Putting more structure on the order of firms will relax this result but it is difficult to say how much.

entry location models such as Hay (1976) and Prescott and Visscher (1977). As in these models, in equilibrium firms choose prices and locations in order to prevent subsequent entry. In the context of heterogeneous firms this means choosing prices and locations so that no firms with marginal costs above a cutoff level $\bar{a}$ may enter.

This setup allows me to dramatically generalize preferences and the notion of locations. The classic location model has consumers inelastically demanding a single unit of a good and facing linear distance costs to obtain the good. Without this setup, prices in equilibrium depend on the locations and marginal costs of all other firms in the market making the equilibrium highly sensitive to order and thus intractable. When firms are concerned with excluding further entry, equilibrium variables are tractable under much more general conditions including elastic demand, concave or convex distance costs, and marginal or fixed distance costs. This flexibility makes the model more attractive to interpret in the style of Lancaster (1979) where elastic demand and convex distance costs are motivated by appeal to preferences over product characteristic space.

A trade reform increases competitiveness on a number of dimensions. The maximum marginal cost of operating firms, $\bar{a}$, declines with a reduction in tariffs so that the least productive firms can no longer profitably operate. For a flexibly parameterized utility function, firms charge variable markups which are functions of $\bar{a}/a$, the ratio of the cutoff marginal cost to their marginal cost. The most productive firms charge the biggest markups while the highest cost firms charge vanishing markups. Individual firm markups decline under a trade reform, but the high-cost low-markup firms exit so that the simple average of firms' markups before and after a trade reform are the same. The space of product types is limited, so the similarity between a firm's product and its nearest competitor's moves with the total number of firms in the industry. The number of available products declines, thus products become more isolated in characteristics space when tariffs are reduced if individual demand is more than unit elastic. This demand specification is desirable because it also implies that firm

output is unbounded, and the firms with more output hire more workers, which is consistent with firm revenue data.

Compared to a benchmark Dixit-Stiglitz model with CES preferences and heterogeneous firms, this model offers several improvements which bring it in line with empirical evidence on trade. In the benchmark model, the set of operating firms is limited by fixed costs which imply a minimum scale of production. Variable markups limit the set of operating firms without fixed costs in this model and permit the small-scale exporters we see in the data. Constant markups imply that measured GDP and hence TFP do not change under a trade liberalization in spite of increases in average firm productivity. Variable markups allow the model to see the measured gains in GDP found in studies like Treffer (2004). This location model shares these features with recent models such as those in Melitz and Ottaviano (2008) and Simonovska (2009) which achieve variable markups by imposing non-homothetic preferences but arrives at this result as an endogenous feature of the market structure.

Section 2 discusses the model in the context of other models in the location and trade literatures. Section 3 sets out and characterizes the model. Section 4 describes a general equilibrium trade model whose industries have the location model as a market structure and explores improvements on the benchmark model and novel implications. Section 5 concludes.

2. RELATED LITERATURE

The location model presented here and its extension to a general equilibrium trade context are related to papers in several literatures which help highlight the forces at work in the model and its contribution to the field. In this section I discuss some predecessor sequential entry location models, some modern location models, the state of the trade literature using location models, and modern heterogeneous firm models with variable markups.

2.1. Sequential Entry Location Models. The original model of Hotelling (1929), extended to include price competition, has an existence problem formulated most clearly by d’Aspremont, Gabszewicz, and Thisse (1979). This problem is usually resolved either by

using quadratic distance costs as recommended by d’Aspremont et al., or by modifying Hotelling’s original line segment to be the circumference of a circle as in Salop (1979) or both. Even on a circle, however, differences in firm costs lead to a non-existence problem under linear distance costs while the resulting mixed strategy equilibria have proven difficult to characterize. Quadratic distance costs models retain pure strategy equilibria for greater degrees of firm heterogeneity, but each firm’s price becomes dependent on the entire sequence of firm prices and locations. These models are intractable and permit many equilibria where the same firm charges dramatically different prices.

This non-existence problem is one of the original motivations of the sequential entry literature papers, principally Hay (1976) and Prescott and Visscher (1977). In these papers identical firms choose locations in a pre-specified order and charge an exogenous common price until no firms can profitably enter. These models address a concern about the standard Nash equilibrium concept. When conjecturing a deviation from an equilibrium, a firm calculates its profits with other firms following on-equilibrium actions, not their best response to the deviation. In sequential entry any deviation induces a subgame where subsequent firms react to the deviation. In the location setting this makes deviations less profitable and recovers pure strategy equilibria.

Besides this methodological contribution, both Hay and Prescott and Visscher find starkly different behavior from that of standard location models. In their fixed-price, identical firm world, firms act to prevent subsequent entry, playing against outside firms rather than their neighboring operating firms. I find the same results in a similar model which includes both price competition and heterogeneous firms in spite of the fact that all operating firms enter simultaneously. The threat of subsequent entry by outside firms is sufficient to sustain the same behavior.

Rothschild (1976, 1979) also solves a sequential entry model but finds somewhat different results due to the fact that all firms operate. Eaton and Wooders (1985) solve the standard location model for the smallest operating number of firms (and hence largest profit) such that

no other firms can profitably enter. Their paper identifies a particular number of firms in equilibrium but otherwise shares all the features of the standard heterogeneous firms model and becomes intractable under firm heterogeneity.

2.2. Location Models with Heterogeneous Firms. Several recent papers in addition to Vogel (2008) have considered a location model with heterogeneous firms. In a separate paper, Vogel (2009) motivates a price-discrimination location model with concerns about the limits on heterogeneity in his earlier model. In this model firms simultaneously choose locations and then simultaneously choose prices at each location – firms perfectly price discriminate.³ This set up permits an arbitrary degree of firm heterogeneity but makes considerable sacrifices in terms of the model’s applicability. Firms price according to the marginal cost plus distance cost of their nearest competitor which means that the consumers closest to a firm are charged the highest prices. In the context of a product characteristic space, where a consumer’s location represents his ideal type of good, location is more reasonably regarded as private information. Even in a geographic context, it is not common to see a firm charging what is essentially a negative shipping cost.

Syverson (2004) and Aghion and Schankerman (2004) also write down location models with heterogeneous firms, but to make the model workable they fix the set of equidistant locations around the entire circle while firms pick their prices knowing only the average market price, not the specific prices of their neighbors. In an appendix, Syverson shows that it is possible to solve for a Nash equilibrium in prices, but that these prices are the solution to a system of equations as large as the number of firms and depend on firm order. Alderighi and Piga (2008) achieve a similar result, and both papers require bounds on heterogeneity which are more stringent than those in Vogel (2008).

2.3. Location and Trade. The difficulty that location models have handling firm heterogeneity has restricted their use in international trade to a few special examples. Lancaster

³The robustness of this model permits an entry stage before locations are chosen, but this feature is tangential to this discussion.

(1979) solves a symmetric two-country model with identical firms but with importers facing an extra marginal cost to bring the product to market. This setup can only be tractably managed in the case with alternating foreign and domestic firms which cannot be extended to non-symmetric countries, or where all domestic firms are bunched together and separated from all foreign firms. The first specification, also used in Helpman (1981) and Schmitt (1990), requires the same number of imported goods as domestically produced goods and places limits on import costs. A model with N countries would require that $1/N$ firms in the market are domestic producers. Hummels and Lugovskyy (2009) share the second approach of complete separation between foreign and domestically produced goods.

A separate literature focuses on a two-firm case where monopolists in each country become duopolists under free trade. Originally proposed by Shachmurove and Spiegel (1995) with linear distance costs, the conjectured equilibrium did not exist. This case was studied for quadratic distance costs by Tharakan and Thisse (2002) and extended to include intermediate inputs and outsourcing by Egger and Egger (2007). Gopinath, Gourinchas, Hsieh, and Li (2009) motivate an empirical study of costs at the border with a similar geographically-set location model in which a circle is divided into two countries and a large number of firms are located at constant intervals around the circle. Here trade occurs only to serve consumers near the border and all firms in the same country are identical while variations across countries in the cost of intermediate goods drives prices.

2.4. Variable Markups. The endogenous generation of variable markups drives many of the results in the trade section of this paper. The model presented here shares this feature and many of its implications with trade models such as Melitz and Ottaviano (2008), Simonovska (2009), and Behrens and Murata (2009) which modify the Dixit-Stiglitz setup so that consumers have non-homothetic preferences. The model presented here generates these results endogenously, in some cases generating identical demand functions, for example, the preferences in section 3 will generate the linear demand seen in Melitz and Ottaviano.

Other trade papers achieve variable markups through market structure, but at sacrifice significant advantages. Bernard, Eaton, Jensen, and Kortum (2003) solve a model where firms in each industry engage in Bertrand competition, but where the number of firms is fixed and equal to the number of industries. Atkeson and Burstein (2008) offer a similar model with Cournot competition, but cannot characterize firm behavior analytically. Further, goods in the same industry in both of these models are perfect substitutes, so the models cannot address changes in the degree of substitutability between goods.

3. LOCATION MODEL

The central model of this paper takes the standard players and payoffs from traditional location models but changes the timing of firm actions. Firms may enter in one of many stages with other firms possibly entering subsequently. This change in timing has two major advantages. First, it solves an existence problem. It had pure strategy equilibria where none exist in a traditional model. Second, it simplifies firm behavior. Much like sequential entry models with fixed prices and identical costs, firms set prices with respect to the costs of the outside firms rather than the prices of their neighbors. Because all firms play against the same outside firms, the equilibrium is order-independent and highly tractable.

The conventional location model features two stages: a simultaneous choice of locations followed by simultaneous choice in prices. A subgame perfect Nash equilibrium (SPNE) in this model requires a Nash equilibrium in the second stage which is the source of non-existence for excessive firm cost heterogeneity. The Nash equilibrium concept is a demanding one in the sense that equilibria must be immune to deviations against which no other firm can respond. In a location setting, a firm may consider a deviation in which it lowers its price and captures all of a neighboring firm's demand. The firm calculates the profit from this deviation assuming that its neighbor charges its equilibrium price. Sequential entry reduces

this problem because a deviation induces a subgame in which firms who have not entered can still respond to the firm's deviation. In the location context, this reduces the payoffs to conjectured deviations and recovers equilibria that cannot be supported in one-shot games.

The model is set in a now standard location environment. Consumers are uniformly distributed around the circumference of a circle. Firms are located at points around the circle and charge a mill-price p_j where j indexes the firm. Firms produce output with labor input subject to a fixed cost f and a constant marginal cost a_j .

The structure of consumer preferences vary across location models. To fix ideas and to allow for easy comparison with other work in the literature, I focus first on a common demand specification (cf. Hotelling 1934, Vogel 2008). Consumers inelastically demand one unit of a good. A consumer i who is distance x_{ij} along the circumference from firm j experiences a linear distance cost ηx_{ij} when buying from firm j . The consumer picks which firm to buy from by solving

$$(3.1) \quad \max_j \quad u - \eta x_{ij} - p_j$$

where u is the utility from consuming one unit of the good and is assumed large enough that all consumers buy a product in equilibrium. Thus the consumer ranks firms according to their effective price $p_j + \eta x_{ij}$. Section 3.5 summarizes the modifications required in the more general model.

3.1. Actions and Timing. This model differs from most location models in the timing of firms' decisions. A firm may enter in any stage s with $s \in \{0, \dots, \infty\}$. All firms entering in a given stage simultaneously pick a location and a price which cannot be changed in subsequent stages. Thus an action in this game is a triple (s_j, p_j, x_j) : a stage, a price, and a location. A history, $h^s \in \mathcal{H}^s$, is a list of all actions taken through stage s with $\mathcal{H}^\infty$ being the set of all infinite histories. A strategy is a mapping from the set of all histories to actions

which is measurable with respect to stage s information. An equilibrium of this game is a standard SPNE.

3.2. Payoffs. Each firm with positive output serves the consumers at its location. As we consider consumers further from the firm's location, the consumer's effective price rises linearly at rate η until we reach a consumer who faces the same effective price from firms on either side of them. A firm's demand is the distance between its indifferent consumers on either side of its location times ϕ , the density of consumers on the circle. Figure 1 represents this demand determination where the x axis is a section of the circumference of the circle and the y axis is the effective price. In each stage all firms who have already entered produce

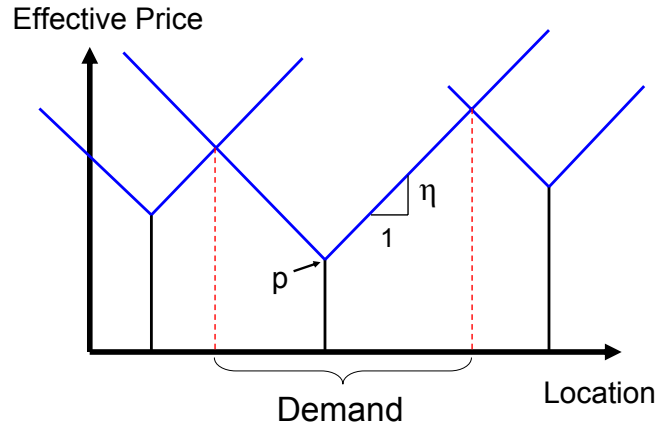


Figure 1: Demand Determination

to satisfy demand and earn profits. Firms discount profits at rate $1/R$ where R is the gross interest rate.

Finally, there is a cost, γ , which a firm pays when it undercuts another firm. In some states of the world a firm may serve 0 consumers because all consumers see another firm offering a lower effective price. If, for example, a new firm is located very close to the center firm as in figure 1 but charges $p_1 \ll p_0$, all of the consumers around the center firm 0 will prefer to buy from firm 1 as illustrated in figure 2. Firm 0 is undercut by firm 1 so firm 1

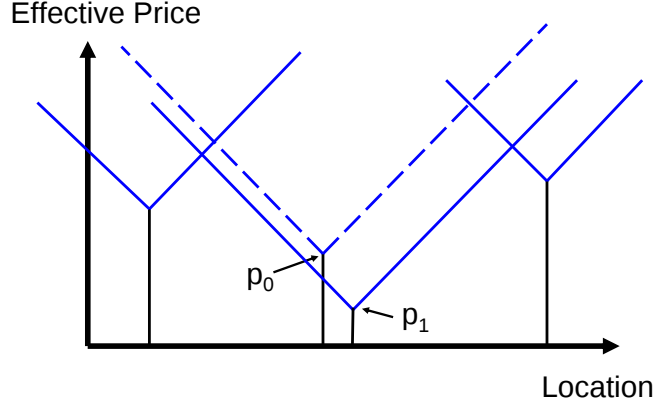


Figure 2: Undercutting – consumers at firm 0's location buy from firm 1.

will have to pay a small cost γ . This cost may be thought of as representing an un-modelled inventory dumping, legal challenge, retaliation, or other end-of-life firm behavior. γ is needed to rule out behavior which is marginally more profitable than equilibrium behavior when the number of firms is small. A firm which steals some of another firm's consumers but not those at its exact location does not incur this cost. We will see that as the number of firms becomes large, γ can be taken arbitrarily close to zero.

To summarize, a firm j 's discounted total profit is

$$\pi_j = \frac{R-1}{R} \sum_{s=s_j}^{\infty} \left(\frac{1}{R^s} (p_j - a_j) \cdot D \right) - (f + \gamma \chi_j) / R^{s_j}$$

where D is the firm's demand and χ_j is the number of firms undercut by firm j upon entry in stage s_j . Both D and χ_j are determined in equilibrium.

3.3. Equilibrium. I study an equilibrium in which firm behavior is directed at preventing subsequent entry. In this section, I describe on-path behavior characterized by a few critical

variables, state an existence theorem for this equilibrium, and sketch its proof. A more detailed specification of strategies and fuller proof are given in the appendix.

All operating firms enter simultaneously in the first stage in arbitrary order along the circumference of the circle. The equilibrium is characterized by a cutoff firm with marginal cost $\bar{a}$ who can profitably enter at any location where the consumer faces an effective price of $\bar{p} = \bar{a} + 2\sqrt{\eta f}$. At that location, with the optimal price, the cutoff firm can just cover the fixed cost of investment. Thus, in equilibrium, every consumer on the circle faces a minimum effective price less than or equal to $\bar{p}$. Any firm with marginal cost $a_j \geq \bar{a}$ cannot profitably enter in the first or any subsequent stage.

Almost all entering firms' prices are given by a simple maximization problem. Firms pick prices as if they will serve all consumers who see the effective price of their good as less than $\bar{p}$. That is, firms are sure that no other firms will price them out of consumers facing an effective price less than $\bar{p}$, but that firms will steal all consumers facing an effective price above $\bar{p}$. Given the linear distance cost, for any price p this implies a demand of $2\phi \cdot \frac{\bar{p}-p}{\eta}$. Firms prices solve

$$(3.2) \quad \max_p \quad (p - a_j) \cdot 2\phi \cdot \frac{\bar{p} - p}{\eta}$$

with optimal price $p^* = (\bar{p} + a_j)/2$ serving all consumers along a line segment of length $d(\bar{a}, a_j) = (\bar{p} - a_j)/\eta$. Conditional on the cutoff marginal cost $\bar{a}$, this is the best a firm can do in any equilibrium at any price since additional demand would imply consumers with an effective price above $\bar{p}$ which would induce subsequent entry by the cutoff firm.

The cutoff level firm cost $\bar{a}$ is chosen to keep out as many firms as possible consistent with individual rationality. The maximization problem for each firm implies a length of line segment served $d(\bar{a}, a_j)$. For these demands to be consistent we need

$$(3.3) \quad \sum_{a_j < \bar{a}} d(\bar{a}, a_j) = L$$

where L is the circumference of the circle.

In practice, it will not in general be possible to satisfy this equation because the left hand side of (3.3) is discontinuous in $\bar{a}$, and there are only a finite number of possible $\bar{a}$'s, the set of firm marginal costs. Instead, $\bar{a}$ is the smallest marginal cost such that $\sum_{a_j < \bar{a}} d(\bar{a}, a_j) \geq L$. To make these demands consistent, the one or two operating firms with the highest marginal costs will serve a smaller segment of the circle. Locations are then specified so that the boundary consumer between every pair of firms faces effective price $\bar{p}$.

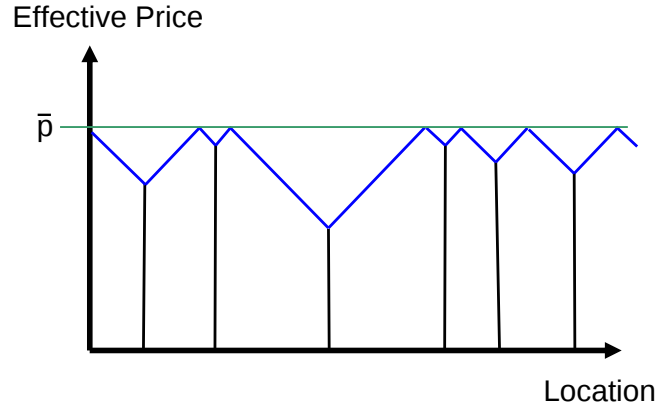


Figure 3: Firm Prices and Locations in Equilibrium

Theorem 3.1. *For any set of firm marginal costs $A := \{a_1, \dots, a_N\}$ there is an (f, R, γ) such that there is an equilibrium of this game in which all firms with $a_j < \bar{a}$ enter in the first stage, almost all set prices according to the maximization problem 3.2, and no other firms enter.*

A more detailed statement of the equilibrium and a more mechanical proof of this proposition are given in the appendix, but I discuss the important possible deviations here. Prices

set by the firms in this model are lower than in standard models with one-stage price setting. If we eliminated subsequent entrants, a firm would find it profitable to raise prices from $(\bar{p} + a_j)/2$ to $(3\bar{p} + a_j)/4$. This higher markup more than offsets the loss in demand from setting a higher price. Entry makes this deviation worse than equilibrium behavior because subsequent firms steal all customers under this deviation who face an effective price higher than $\bar{p}$. Serving only consumers facing an effective price lower than $\bar{p}$ with the higher price $(3\bar{p} + a_j)/4$ is suboptimal since the price does not solve 3.2.

Unlike equilibria in games with one-shot location stages, this equilibrium is immune to global deviations defined as a change in location, x_d , and in price, p_d , such that there is no overlap between the set of consumers served in equilibrium and in the deviation. Define p_L and p_R to be the effective prices of the indifferent consumers at the left and right boundaries of a firm's demand under this deviation. For a fixed price, the firm's deviation profits, $\pi(p_d, p_L, p_R)$, are increasing in p_L and p_R because they imply greater demand, $D = \eta((p_L - p_d) + (p_R - p_d))$. Equilibrium profit is $\pi(p^*, \bar{p}, \bar{p})$, and by (3.2) we have $\pi(p^*, \bar{p}, \bar{p}) \geq \pi(p_d, \bar{p}, \bar{p})$. Since $\bar{p}$ is the highest effective price on the circle in equilibrium, we have $\pi(p_d, \bar{p}, \bar{p}) \geq \pi(p_d, p_L, p_R)$, and no global deviation improves on equilibrium profits.

Local deviations in which a firm's deviation set of consumers overlaps with its equilibrium set of consumers require more work to rule out. It is not true that $p_L \leq \bar{p}$ since it may be that the equilibrium effective price was the one charged by the deviating firm. In this case, the region of the circle whose consumers face an effective price higher than $\bar{p}$ are ripe for poaching. After a local deviation, firms that do not operate on-path enter profitably in this region. They play a subgame equilibrium identical to the one identified here, there will be a new cutoff firm $\tilde{a} > \bar{a}$ and effective price $\tilde{p} > \bar{p}$.

Without γ , local deviations may be profitable on two counts. First, firms enjoy one stage of profits before subsequent entry. The payoff from this stage is always positive but goes to zero in the limit as we take $R \rightarrow 1$ essentially reducing the time between stages. Second, $\bar{p} < \tilde{p} = p_L \vee p_R$ means the deviation may be marginally profitable in every stage. Figures

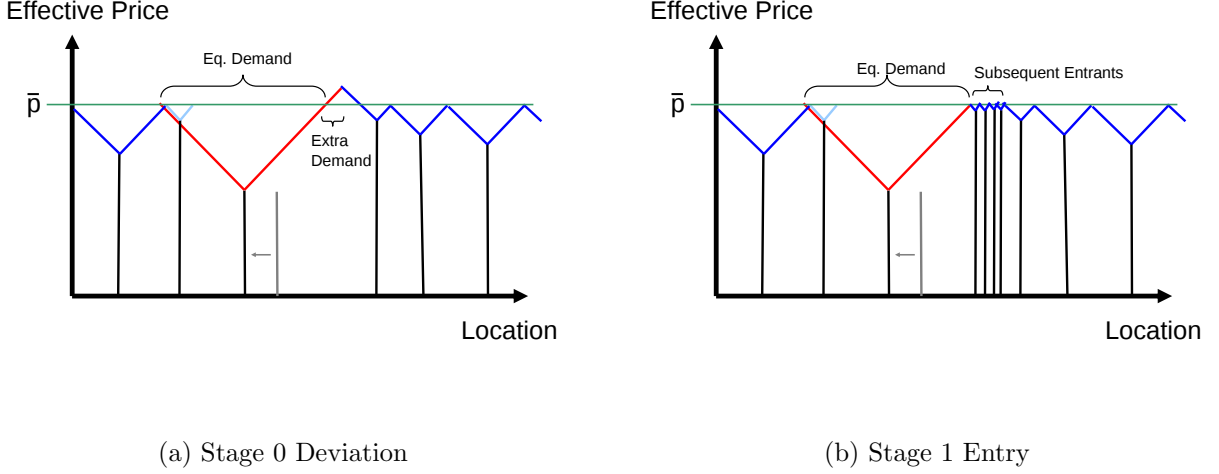


Figure 4: Local Undercutting

4a and 4b present a clear example of this situation. The large firm in the center deviates to just undercut the firm to his left. This yields $p_L = \bar{p}$ while we have $p_R = \tilde{p}$ starting with the second stage. To bound the payoff in this case I need to introduce

$$(3.4) \quad \Delta = \max_j a_j - a_{j-1}$$

the maximum difference between two consecutive marginal costs where the firms are indexed so that $j < k \Rightarrow a_j < a_k$. For any fixed region of consumers who face effective prices above $\bar{p}$ in a deviation, we have that $(\tilde{p} - \bar{p}) \rightarrow 0$ as $\Delta \rightarrow 0$, a fact which comes straight from the subgame equivalent of (3.3), the equation that determines $\tilde{a}$. Thus for Δ small enough and R close enough to 1, the extra profit from this kind of deviation is less than γ .

So far I have only considered first period deviations. What will subsequent firms do if a firm deviates by not entering at all in the first period? The continuation equilibrium following this history is for the firm to enter in the second stage with the same price and location. In fact for any history in which one firm has not entered in m stages the continuation equilibrium is for it to enter in stage $m + 1$ with the same price and location as prescribed for on-path play.

By the recursive nature of the stages, any firm which did not enter and considers entering at another location and price will find it not profitable by the same argument used for first period deviations.

3.4. Limiting Equilibria. For (Δ, f, R) small enough we can find an arbitrarily small γ sufficient to rationalize the equilibrium.

Theorem 3.2. *For any $\gamma > 0$ there is a (Δ, f, R) small enough that the equilibrium described above exists.*

A more detailed proof than the discussion above is available in the appendix. It is possible to further relax this theorem by allowing for arbitrary f , but it involves specifying different pricing rules for small firms to ward off predation and is omitted for simplicity. In section 4, I study a trade model where each industry is in the limiting equilibrium as $(\gamma, \Delta, f, R-1) \rightarrow \vec{0}$ which this theorem guarantees exists as we take a sequence of γ going to zero.

3.5. Extensions for general utility. The extension of this model to a general specification of consumer utility is quite natural. The equilibrium is still characterized by a marginal firm $\bar{a}$, and firm prices are given by the same maximization problem. The determination of firm demands can be expressed using the same graphs as before, except that in the case of elastic individual demand the notion of effective price must be generalized to indirect utility. Consumers rank firms by the total indirect utility they experience from consuming an optimal quantity at their location. We can represent the determination of demand graphically in figure 1 where the vertical axis now represents negative indirect utility.

There will not be a universal cutoff marginal utility equivalent to $\bar{p}$. Because the slopes of the indirect utility curves vary, a marginal firm entering with a fixed price would seize more demand where two flat sections of curves intersected than a place where two steep curves intersected. To make sure that the cutoff firm earns zero profits from entry everywhere, firm demands will adjust slightly from the ones implied by (3.2) with the distance between

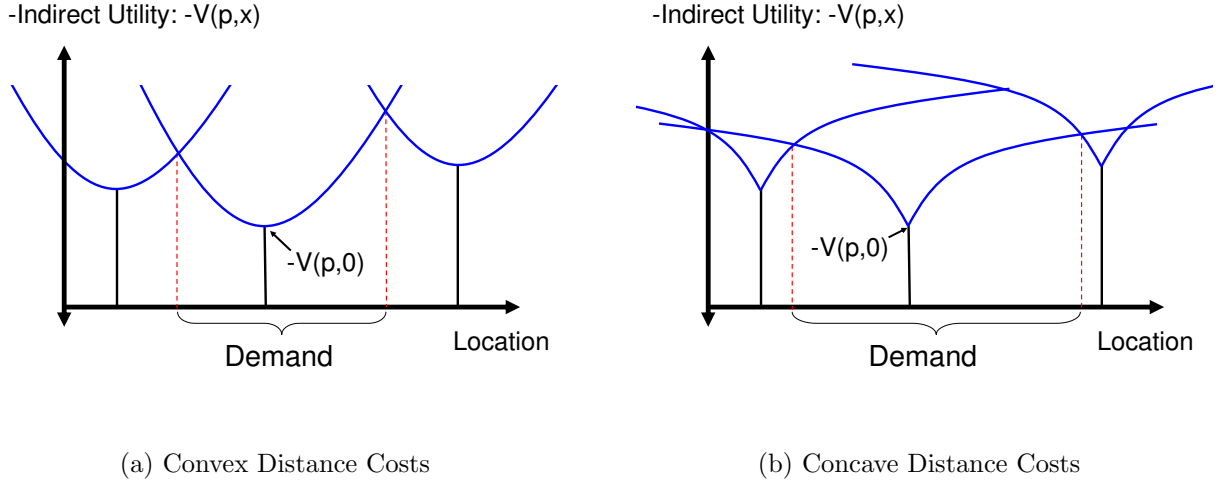


Figure 5: Alternative Distance Costs

firms being increased or decreased until the entry profit is exactly zero. The size of this adjustment goes to 0 with f , so that in the limiting equilibrium studied in section 4 we have elastic demand but a genuinely universal cutoff in indirect utility space.

4. TRADE AND LOCATION

What can this location model tell us about the effects of a trade reform? How does a model with such a market structure compare with the more standard Dixit-Stiglitz framework? I take the location market structure as the building block of a simple two-country, symmetric, general equilibrium trade model in order to investigate the effect of a trade reform.

The model performs better on a number of dimensions than a standard Dixit-Stiglitz model with CES preferences, such as those in Melitz (2003) and Chaney (2008) which I will call the benchmark model. First, although each consumer's *intensive* demand is CES, the firm's demand function is not because there is an *extensive* margin to firm demand as well.

Higher prices reduce the set of consumers. In a Dixit-Stiglitz environment, CES preferences imply constant markups: all firms set prices to a constant percentage above their costs. The location model features variable markups across firms, i.e. high cost firms have smaller profit margins.

Variable markups allow the location model to improve on the benchmark model on several dimensions. First, the benchmark model requires fixed costs to limit the set of exporting firms, implying a minimum scale of export. As discussed in Arkolakis (2008), a significant fraction of exporting firms in French export data are smaller than the standard calibrated minimum size.

The model suggests that markets become more competitive by two widely considered measures, average productivity and firm profit margins, but looks less competitive in terms of the number of goods available in the market and the similarity between products measured as the distance between their locations. The cutoff firm's productivity, which implies the average productivity under a Pareto distribution, declines with a reduction in trade costs. High-cost domestic firms can no longer profitably produce. Firms that are not forced to exit see their profit margins decline. When firm demand is sufficiently elastic, the exit of domestic firms outpaces the entry of foreign firms, and the total number of varieties declines.

4.1. Setup. In each country there is a continuum of industries indexed by $\omega \in [0, 1]$. In each of these industries N firms play the location game from section 3. Consumer i picks firms $j(\omega)$ in each industry and a quantity to purchase from that firm $c(\omega)$, to solve

$$(4.1) \quad \max_{j(\omega), c(\omega)} \int_0^1 \left(\frac{c(\omega)^\rho - 1}{\rho} - \eta x_{ij}(\omega)^\sigma \right) d\omega$$

$$(4.2) \quad \text{s.t.} \quad \int_0^1 p(\omega) c(\omega) d\omega \leq 1 + \Pi + T$$

where labor income is normalized to 1, Π is per-capita firm profits so that all consumers hold equal shares in all firms, and T are the tariff revenues lump-sum transferred back to consumers.

This consumer's utility function is significantly more general than the standard preferences in section 3. Consumers have CES demand and face a fixed distance cost which may be convex, linear, or concave but does not vary with quantity consumed. This is a retail model in the sense that consumers pay a cost to 'get to' the firm's location and then consume independent of that cost. This specification is particularly tractable because all consumers with the same $j(\omega)$ will consume the same $c(\omega)$. Lastly, consumers' locations on each industry's circle are i.i.d., so consumers have the same outside value of expenditure everywhere on the circle.

There is a large number of potential firms in each industry in each country. For tractability, I approximate the firm CDF as $H(a)$, a continuous distribution. Consistent with the empirical literature and common practice, I specify it to be Pareto in productivities ($1/a$) with curvature parameter κ so that $H(a) = (a/a_{max})^\kappa$. I restrict κ to be greater than $\frac{\sigma+\rho}{\sigma(1-\rho)} + 1$ to ensure that aggregate demand is well defined.

Domestic firm prices are the solution to a maximization problem equivalent to problem (3.2) from section 3.3

$$(4.3) \quad \max_p \quad (p - a) \cdot c(p) \cdot d(p)$$

where $c(p)$ is the *intensive* demand by each consumer, and $d(p)$ is the *extensive* demand, the length of the circumference served by the firms.⁴ In a Dixit-Stiglitz context, ρ is restricted to be greater than 0 so that the firm's problem has a solution, but this model can accommodate any curvature in the demand function because the *extensive* margin puts an upper bound on prices charged by firms.

⁴For clarity I will normalize the density of consumers on the circle, ϕ , to 1 in this application. One of (η, ϕ, L) is redundant in this model, but when fixing η , the choice of variable to represent country size will be important for results in an asymmetric model and comparative statics in country size.

Foreign firms enter the game as any other firm except that they pay an *ad valorem* tariff τ on the final value of imported goods. They are free to pick any location on the domestic circle for their industry, so their problem is just

$$(4.4) \quad \max_p \quad (p(1 - \tau) - a) \cdot c(p) \cdot d(p)$$

or

$$(4.5) \quad \max_p \quad (1 - \tau)(p - \frac{a}{1 - \tau}) \cdot c(p) \cdot d(p)$$

Thus a foreign firm with marginal cost a behaves just like a domestic firm with marginal cost $a/(1 - \tau)$. To close the model I impose a balance of trade condition.

4.2. Variable Markups. The firm's pricing rule is an implicit function of a and $\bar{a}$ from the first order condition of the maximization problem (4.3):

$$(4.6) \quad \frac{p - a}{a/\rho - p} = \sigma \left(1 - \left(\frac{p}{\bar{a}} \right)^{\frac{\rho}{1-\rho}} \right)$$

Since the firm's price will always be below $\bar{a}$, the right hand side of this equation takes the same sign as ρ . When $\rho > 0$, p is constrained to be between a , the perfectly competitive price, and a/ρ , the price associated with the Dixit-Stiglitz model and no extensive margin for demand. As σ , which controls the elasticity of extensive demand, goes to zero, p goes to a and we get perfect competition in the limit. As we take $\sigma \rightarrow \infty$ we get Dixit-Stiglitz markups in the limit.

For any p that satisfies this equation for some $(a, \bar{a})$, θp satisfies the equation for $(\theta a, \theta \bar{a})$. Thus the pricing policy is homogeneous of degree one in $(a, \bar{a})$, and we can write it as $p(a, \bar{a}) = a \cdot M(\bar{a}/a)$ for some markup function M of the ratio of the cutoff marginal cost to the firm's marginal cost. In the special case of $\rho \rightarrow 0$, $\sigma = 1$, for example, the first order

condition becomes

$$(4.7) \quad p = a \cdot W\left(\frac{\bar{a}}{a}e\right)$$

where W is the Lambert W function⁵ defined as $y = W(x) \Leftrightarrow x = y \cdot e^y$. Taking the limit as $\rho \rightarrow -\infty$ with $\sigma = 1$ yields the same limiting pricing rule we saw in section 3 for the case of inelastic demand and linear distance costs.

Markups are always variable in this model with higher productivity firms charging lower prices but higher markups. Threshold firms have markups nearly equal to 1, so that prices equal costs and serve almost no consumers. Markups rise and fall with $\bar{a}$ which declines under pro-competitive reforms like a trade liberalization, as we will see below, or some unmodelled process increasing N . Average markups, however, do not move with $\bar{a}$. This result can be demonstrated with a simple change of variable argument. If we set $\hat{a} = a/\bar{a}$, we can write the average markup $p/a = M(\bar{a}/a)$ as

$$(4.8) \quad \frac{1}{H(\bar{a})} \int_0^{\bar{a}} M(\bar{a}/a) \kappa a^{\kappa-1} da = \bar{a}^{-\kappa} \int_0^1 \bar{a}^{\kappa} M(1/\hat{a}) \kappa \hat{a}^{\kappa-1} d\hat{a} = \int_0^1 M(1/\hat{a}) \kappa \hat{a}^{\kappa-1} d\hat{a}$$

which is independent of all endogenous variables. The intuition for this result is that trade reforms reduce the markups of continuing firms but force the exit of the high-cost firms who had the lowest markups. The net effect is ambiguous, and the two effects cancel with a Pareto distribution.

Variable markups help this model explain several empirical facts which cannot be achieved in a standard Dixit-Stiglitz model with CES preferences. Because the benchmark model has CES demand functions, it requires fixed costs to limit the set of operating firms. These fixed costs imply a minimum scale of production to cover them. As discussed in Arkolakis (2008), conventional estimation of these costs imply large minimum scales for exporting while firm

⁵The Lambert W function has broad applications to a number of applied mathematical fields. For a good survey see Corless, Gonnet, Hare, and Jeffrey (1996).

export data finds a significant fraction of firms who export much less. Here the extensive demand d goes to zero as $a \rightarrow \bar{a}$ while intensive demand goes to $c(\bar{a})$, so that total revenues can be made arbitrarily small by taking a close to $\bar{a}$. Variable markups also allow this model to reflect data on GDP under a reform as discussed in section 4.4.

Several recent papers including those by Melitz and Ottaviano (2008), Simonovska (2009), and Behrens and Murata (2009) have departed from CES preferences to generate variable markups and achieve similar improvements over the benchmark model. This model is distinguished by realizing these improvements as an outcome of firm choice in an imperfectly competitive market structure even with CES intensive demand.

4.3. Equilibrium and Trade Liberalization. In equilibrium the cutoff marginal cost $\bar{a}$ will be the same in every industry in each country by symmetry. Domestic firms with $a < \bar{a}$ will operate, as will foreign firms with $a < (1 - \tau)\bar{a}$. Since foreign firms act like domestic firms with higher costs, we only need to consider the distribution of effective marginal costs. If $H(a)$ is the distribution of domestic firms' costs, then $H^*(a) = H(a) \cdot (1 - \tau)^\kappa$ is the distribution of potential importer costs. Thus the domestic market operates as if it was a closed economy where the number of firms is $\tilde{N} = N \cdot (1 + (1 - \tau)^\kappa)$ for output and consumption and $\tilde{N} = N \cdot (1 + (1 - \tau)^{\kappa+1})$ for profits and employment.

The the cutoff productivity $\bar{a}$ is determined by the consumer's budget constraint, (4.2), and the length of the circumference of the circle

$$(4.9) \quad L = N \cdot (1 + (1 - \tau)^\kappa) \int_0^{\bar{a}} d(a, \bar{a}) dF(a)$$

but can be reduced to a single equation

$$(4.10) \quad N(1 + (1 - \tau)^\kappa) \cdot \left(\frac{1 + (1 - \tau)^\kappa}{1 + (1 - \tau)^{\kappa+1}} \right)^{\frac{\rho}{\sigma}} = \bar{a}^{\frac{\rho}{\sigma} - \kappa} \cdot \eta^{\frac{1}{\sigma}} \cdot \Theta$$

where Θ is a complicated integral that depends only on ρ and σ .

A reduction in tariffs increases the left hand side of the equation for most parameters, requiring a decrease in $\bar{a}$ for permissible values of κ ⁶. This is a standard result: a reduction in costs induces firms to begin exporting. How does this affect the total number of firms in the market? Recognizing that the fraction of active firms $H(\bar{a}) = \bar{a}^k$, we can rearrange this equation to obtain an expression for the number of operating firms.

$$(4.11) \quad (1 + (1 - \tau)^\kappa)N \cdot H(\bar{a}) = \bar{a}^{\frac{\rho}{\sigma}} \cdot \eta^{\frac{1}{\sigma}} \cdot \Theta$$

Here the total number of varieties moves with the sign of ρ , increasing when intensive demand is inelastic and decreasing when it is elastic. The intuition here comes from the elasticity of $\bar{a}$ with respect to τ . There are two first order effects of a reduction in tariffs. Exporters increase exports and some firms begin to export. This increase in demand violates goods market clearing and $\bar{a}$ decreases to return to equilibrium. The number of domestic firms who are forced to exit depends on the change in $\bar{a}$. When firm demand is highly inelastic, when ρ is low and σ is high, existing exporters respond weakly to a reduction in tariffs, and domestic exit is limited. When the demand is highly elastic, $\bar{a}$ moves significantly to maintain equilibrium and domestic exit overwhelms the entry of importers.

Elastic intensive demand is a desirable specification because it is consistent with the firm size distribution. When $\rho < 0$, the limiting firm price as costs go to zero can be read from 4.6 setting $a = 0$. We have

$$(4.12) \quad p_{min} = \bar{a}(1/\sigma + 1)^{\frac{1-\rho}{\rho}}$$

⁶The fraction on the left hand side of 4.10 is complicated by the fact that transferring tariff revenue to the consumer adds an ‘optimal tariff’ flavor to the calculation that a more conventional iceberg transportation cost specification would not. Recall that we require $\kappa > (\sigma + \rho)/\sigma(1 - \rho)$ for aggregate demand to be well defined so that $\rho/\sigma - \kappa$ is negative. This requirement also guarantees that the left hand side of 4.10 is decreasing in τ as shown in the appendix

With a price bounded above zero as firm productivity goes to infinity, firm demand will be bounded below a finite number, $D(p_{min})$, while increasing productivity means the employment required to satisfy that limiting demand, $aD(p_{min})$, declines to zero. Thus as $a \rightarrow 0$ the firms increase demand approaching $D(p_{min})$ but reduce employment.

4.4. Measuring Productivity From a Trade Reform. In the benchmark Dixit-Stiglitz model with CES preferences, aggregate productivity gains in terms of firm's average productivities are lost in measured GDP because of the way prices are handled. To make this illustration sharper, and to bring it into line with a more extensive discussion in Gibson (2007), I extend the model to endogenously determine N in the conventional manner.

Briefly, before the location game is played, let there be a productivity draw stage where firms pay a fixed costs f_e to draw a marginal cost from the distribution $H(a)$. The number of firms is associated with the number of draws, and firms will continue to draw until the expected profits (which are declining in N) are equal to f_e . Consumers fund the costs draws and pool risk, so that for each consumer the per capita profit term in his budget constraint is equal to zero. Consistent with common practice, I reinterpret the tariffs as iceberg transportation costs which are not transferred back to households. The budget constraint now simply equates total consumption to wage income normalized to one.

We can aggregate the consumer's budget constraint and substitute the balance of trade condition to replace consumption of imports with production of exports to get an expression for GDP which is fixed

$$(4.13) \quad \int_S p(s)y(s) = 1$$

where I use S to represent the set of available goods, and s a particular good generically so that it may describe both setups. Regardless of the specification of S , the left hand side of this equation is an expression for GDP and is fixed to be equal to the stock available labor times the wage normalized here to be 1.

Crucially, in the Dixit-Stiglitz framework with CES preferences, the price set by a firm with marginal cost a is just a/ρ independent of any other conditions in the economy. When we consider the real effects of a trade reform between period $t - 1$ and t , we consider GDP before and after the reform in base period prices. But from (4.13) we have

$$(4.14) \quad \int_{S_{t-1}} p_{t-1}(s)y_{t-1}(s) \quad = \quad \int_{S_t} p_t(s)y_t(s) \quad = \quad \int_{S_t} p_{t-1}(s)y_t(s)$$

where the first equality comes from (4.13) regardless of how the set of firms changes, and the second is from the constancy of prices. The first equation still holds in the location model, but variable markups imply that the second does not. Lower tariffs imply a lower $\bar{a}$ hence lower prices. We have $p_t(s) < p_{t-1}(s)$ for all s operating in both periods, and this decline is applied to new goods as an accounting convention. Thus

$$(4.15) \quad \int_{S_{t-1}} p_{t-1}(s)y_{t-1}(s) \quad = \quad \int_{S_t} p_t(s)y_t(s) \quad \leq \quad \int_{S_t} p_{t-1}(s)y_t(s)$$

a trade reform is measured as an increase in measured real GDP in the model as it is in empirical studies like Trefler (2004).

In the model TFP and GDP move together since labor, the only other input in production, is fixed. Thus in the standard model, TFP does not change under a trade reform in spite of the fact that average productivity, $1/a$, declines. In a model with variable markups both decline.

5. CONCLUSION

Location models are a rich environment for thinking about pro-competitive policy changes like trade liberalizations. In addition to describing a policy's effect on firm profits and firm exit, they endogenize the firm's choice of how directly to compete with other firms in terms of product similarity. This extra dimensions has typically come at the cost of tractability which has limited the use of location models at high levels of aggregation. This paper presents a

model with a sequential entry flavor which remains highly tractable in a general equilibrium context.

I find that a trade reform has a counterintuitive effect: the number of firms declines and their average separation in product space increases. From the perspective of product differentiation, markets may appear less competitive even though high-cost firms are forced to exit, and surviving firms curtail their profit margins.

International trade models, particularly the heterogeneous firms models of the last decade, have a rich set of facts to confront in trying to explain the pattern of trade. The focus of this paper is an investigation of the immediate implications for competition from a location model, but there are many dimensions on which the model might be further compared to the data.

There are opportunities to put more structure on the model which are worth exploring. Here I allow firms to produce products with different characteristics for domestic and foreign consumption, but one might also consider fixing the products to have the same location on each circle. The uniform distribution of consumers on the circle is a maintained assumption in the location literature, but specifying different densities on different circles is an interesting way to think about home bias in production and to analyze the findings from that literature.

While I have focused on international trade, an animating application for heterogeneous firms models, this location model can also be employed to examine the implications of other pro-competitive reforms. The model has similar implications, for example, associated with lowering firm startup costs. The extra structure of distance costs also permits an examination of the effect of falling transportation costs over time separate from generalized improvements in technology. Plainly, this is a promising environment for future investigation.

REFERENCES

- AGHION, P., AND M. SCHANKERMAN (2004): “On the welfare effects and political economy of competition-enhancing policies,” *Economic Journal*, 114(498), 800–824.
- ALDERIGHI, M., AND C. A. PIGA (2008): “The Circular City with Heterogeneous Firms,” Discussion Paper Series 200803, Department of Economics, Loughborough University.
- ARKOLAKIS, C. (2008): “Market Penetration Costs and the New Consumers Margin in International Trade,” NBER Working Papers 14214, National Bureau of Economic Research, Inc.
- ATKESON, A., AND A. BURSTEIN (2008): “Pricing-to-Market, Trade Costs, and International Relative Prices,” *American Economic Review*, 98(5), 1998–2031.
- BARTELSMAN, E. J., AND M. DOMS (2000): “Understanding Productivity: Lessons from Longitudinal Microdata,” *Journal of Economic Literature*, 38(3), 569–594.
- BEHRENS, K., AND Y. MURATA (2009): “Globalization and Individual Gains from Trade,” Cahiers de recherche 0928, CIRPEE.
- BERNARD, A. B., J. EATON, J. B. JENSEN, AND S. KORTUM (2003): “Plants and Productivity in International Trade,” *American Economic Review*, 93(4), 1268–1290.
- CHANEY, T. (2008): “Distorted Gravity: The Intensive and Extensive Margins of International Trade,” *American Economic Review*, 98(4), 1707–21.
- CORLESS, R. M., G. H. GONNET, D. E. G. HARE, AND D. J. JEFFREY (1996): “On the Lambert W function,” *Computational Mathematics*, 5, 329–359.
- D’ASPREMONT, C., J. J. GABSZEWICZ, AND J.-F. THISSE (1979): “On Hotelling’s “Stability in Competition”,” *Econometrica*, 47(5), 1145–50.
- EATON, B. C., AND M. H. WOODERS (1985): “Sophisticated Entry in a Model of Spatial Competition,” *RAND Journal of Economics*, 16(2), 282–297.
- EGGER, H., AND P. EGGER (2007): “Outsourcing and trade in a spatial world,” *Journal of Urban Economics*, 62(3), 441–470.

- GIBSON, M. J. (2007): “Trade Liberalization, Reallocation, and Productivity,” mimeo.
- GOPINATH, G., P.-O. GOURINCHAS, C.-T. HSIEH, AND N. LI (2009): “Estimating the Border Effect: Some New Evidence,” NBER Working Papers 14938, National Bureau of Economic Research, Inc.
- HAY, D. A. (1976): “Sequential Entry and Entry-Detering Strategies in Spatial Competition,” *Oxford Economic Papers*, 28(2), 240–57.
- HELPMAN, E. (1981): “International trade in the presence of product differentiation, economies of scale and monopolistic competition : A Chamberlin-Heckscher-Ohlin approach,” *Journal of International Economics*, 11(3), 305–340.
- HOTELLING, H. (1929): “Stability in Competition,” *The Economic Journal*, 39, 41–57.
- HUMMELS, D., AND V. LUGOVSKYY (2009): “International Pricing in a Generalized Model of Ideal Variety,” *Journal of Money, Credit and Banking*, 41(s1), 3–33.
- LANCASTER, K. (1979): *Variety, equity and efficiency: Product variety in an industrial society*. Columbia University Press.
- MELITZ, M. J. (2003): “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity,” *Econometrica*, 71(6), 1695–1725.
- MELITZ, M. J., AND G. I. P. OTTAVIANO (2008): “Market Size, Trade, and Productivity,” *Review of Economic Studies*, 75(1), 295–316.
- PRESCOTT, E. C., AND M. VISSCHER (1977): “Sequential Location among Firms with Foresight,” *Bell Journal of Economics*, 8(2), 378–393.
- SALOP, S. C. (1979): “Monopolistic Competition with Outside Goods,” *Bell Journal of Economics*, 10(1), 141–156.
- SCHMITT, N. (1990): “Two-Country Trade Liberalization in an Address Model of Product Differentiation,” *Canadian Journal of Economics*, 23(3), 654–75.
- SHACHMUROVE, Y., AND U. SPIEGEL (1995): “On Nations’ Size and Transportation Costs,” *Review of International Economics*, 3(2), 235–43.
- SIMONOVSKA, I. (2009): “Income Differences and Prices of Tradables,” mimeo.

- SYVERSON, C. (2004): “Market Structure and Productivity: A Concrete Example,” *Journal of Political Economy*, 112(6), 1181–1222.
- THARAKAN, J., AND J.-F. THISSE (2002): “The importance of being small. Or when countries are areas and not points,” *Regional Science and Urban Economics*, 32(3), 381–408.
- TREFLER, D. (2004): “The Long and Short of the Canada-U. S. Free Trade Agreement,” *American Economic Review*, 94(4), 870–895.
- VOGEL, J. (2008): “Spatial Competition with Heterogeneous Firms,” *Journal of Political Economy*, 116(3), 423–466.
- (2009): “Spatial Price Discrimination with Heterogeneous Firms,” NBER Working Papers 14978, National Bureau of Economic Research, Inc.

APPENDIX

APPENDIX A. HETEROGENEITY IN VOGEL (2008)

In Vogel's paper, a firm with cost a_1 in an equilibrium characterized by effective price⁷ $\bar{p}$ will charge price $p_1 = \frac{2}{3}\bar{p} + \frac{1}{3}a_1$ and serve demand $d_1 = \frac{2}{\eta}(\bar{p} - p_1) = \frac{2}{3\eta}(\bar{p} - a_1)$ earning profits

$$\pi_1 = \frac{4}{9\eta}(\bar{p} - a_1)^2$$

How much less productive can another firm be in this equilibrium without the equilibrium breaking down? I consider here the most profitable circumstance for a price stage deviation. The firm is flanked on either side by the highest-cost firms in the distribution, both with marginal costs a_2 and charging $p_2 = \frac{2}{3}\bar{p} + \frac{1}{3}a_2$. p_1 is locally optimal but may not be better than undercutting the adjacent firms. When undercutting, firm 1 does best to charge the least price required for undercutting. I impose the tie breaking rule that consumers who face identical effective prices buy from the undercutting firm to ensure the existence of this least price. Then firm 1's optimal undercutting price is $\hat{p}_1 = p_1 + 2p_2 - 2\bar{p} = \frac{1}{3}a_1 + \frac{2}{3}a_2$, yielding demand $\frac{2}{\eta}((\bar{p} - p_1) + 2(\bar{p} - p_2)) = \frac{2}{3\eta}(3\bar{p} - a_1 - 2a_2)$ and profits

$$\hat{\pi}_1 = \frac{4}{9\eta}(a_2 - a_1)(3\bar{p} - a_1 - 2a_2) = \frac{4}{9\eta}((\bar{p} - a_1) - (\bar{p} - a_2))((\bar{p} - a_1) + 2(\bar{p} - a_2))$$

Thus this price deviation is profitable whenever

$$(\bar{p} - a_1)(\bar{p} - a_2) > 2(\bar{p} - a_2)^2$$

or $\bar{p} - a_1 > 2(\bar{p} - a_2)$. Since demand increases linearly with $\bar{p} - a$ in this model, the equilibrium breaks down due to this deviation whenever the equilibrium calls for one firm to be more than twice as big as another.

This requirement can be relaxed with stronger conditions on firm order, for example, requiring only similarly sized firms to be neighbors. Taken too far this encourages deviations

⁷This maximum effective price is determined differently than in the model presented in this paper.

at the location stage as low cost firms may deviate to regions with only high costs firms. The profits from such deviations are harder to characterize.

APPENDIX B. MODEL PROOFS

I construct an equilibrium characterized by a cutoff firm $\bar{a}$, and model parameters: fixed cost f , undercutting cost γ , maximum firm cost separation Δ , and discount parameter R .

B.1. On Path Equilibrium Allocations. All firms with $a_j < \bar{a}$ enter in the first period. Most firms, which I will call non-marginal firms here, charge prices which solve

$$(B.1) \quad \max_{p_j} (p_j - a_j) \cdot \frac{\bar{p} - a_j}{\eta}$$

where $\bar{p}$ is a function of $\bar{a}$ described below. Thus firms charge $p_j = (\bar{p} - a_j)/2$ and serve a section of the circle of length $d_j = (\bar{p} - a_j)/\eta$.

To construct locations, pick firms in random order. Place the first at any location on the circle. For all other firms, locate the firm distance $(2\bar{p} - p_j - p_k)/\eta$ clockwise from the last located firm where j indexes the firm being located and k indexes the most recently located firm. This guarantees that all firms serve their respective demands and the indifferent consumer between each pair of firms faces an effective price $\bar{p}$ from buying each good.

B.1.1. Determination of $\bar{p}$. Consider a firm with marginal cost $\bar{a}$ entering at a local effective price maximum of $\bar{p}$. The best the firm can do at this location is set $p^* = (\bar{p} + \bar{a})/2$ and earn profits $\pi^* = (\bar{p} - \bar{a})/2 \cdot (\bar{p} - \bar{a})/2\eta - f$ so that this entry is not profitable if $\bar{p} \leq \bar{a} + 2\sqrt{\eta f}$.

B.1.2. Cutoff $\bar{a}$ and Marginal Firms. The cutoff $\bar{a}$ is the smallest marginal cost such that the sum of demands given by (B.1) is greater than or equal to the circumference length L ,

$$(B.2) \quad \sum_{a_j < \bar{a}} (\bar{p} - a_j)/\eta \geq L$$

To make demands consistent with L , some marginal firms serve smaller demand. All but one of these firms charge the higher locally-optimal price $p = (2\bar{p} + a_j)/3$ with smaller demand $2(\bar{p} - a)/3\eta$. Thus each of these firms serves demand $(\bar{p} - a_j)/3\eta$ smaller than they would under (B.1). One firm charges an intermediate price.

The number of marginal firms required depends on how close (B.2) is to holding with equality. Let $a_J = \bar{a}$ where firms are indexed in order of increasing marginal costs. The extra demand associated with all firms being non-marginal, $\sum_{j=1}^J (\bar{p} - a_j)/\eta - L$, can be at most $(J - 2)(a_J - a_{J-1})/\eta + (\bar{p} - a_{J-1})/\eta$. The number of marginal firms depends on K , the largest integer such that reducing all the firms at or above a_K to the higher price $(2\bar{p} + a_j)/3$ would make up for this demand:

$$(B.3) \quad \sum_{j=K}^{J-1} (\bar{p} - a_j)/3\eta \geq \sum_1^{J-1} (\bar{p} - a_j)/\eta - L$$

so that $J - 1 - K$ firms charge the marginal firm price.

Firm K charges an intermediate price so that the demands add up to L exactly. This price satisfies

$$(B.4) \quad \frac{2(\bar{p} - p_K)}{\eta} = L - \sum_{j=1}^{K-1} \frac{\bar{p} - a_j}{\eta} - \sum_{j=K+1}^{J-1} \frac{2(\bar{p} - a_j)}{3\eta}$$

$$(B.5) \quad p_K = \bar{p} - \frac{\eta}{2} \left(L - \sum_{j=1}^{K-1} \frac{\bar{p} - a_j}{\eta} - \sum_{j=K+1}^{J-1} \frac{2(\bar{p} - a_j)}{3\eta} \right)$$

B.1.3. *Summary.* An equilibrium is characterized by a cutoff cost $\bar{a}$ such that

- Between every pair of firms is an indifferent consumer who faces effective price $\bar{p} = \bar{a} + 2\sqrt{\eta f}$
- The $K - 1$ firms with the lowest costs charge $p_j(a_j) = (\bar{p} + a_j)/2$ and serve demand $d_j(a + j) = (\bar{p} - a_j)/\eta$

- The $J - 1 - K$ operating firms with the highest costs charge $p_j^m(a_j) = (2\bar{p} + a_j)/3\eta$ and serve demand $d_j^m(a_j) = 2(\bar{p} - a_j)/3\eta$
- Firm K charges an intermediate price and serves an intermediate demand
- All firms $j \geq J$ do not enter

B.2. Off Path Behavior. If an operating firm deviates from any subgame by not entering in the prescribed period, the continuation game specifies that the deviating firm enter in the following stage at the same price and location.

For a stage 0 price or location deviation, subsequent play depends on the post-deviation length of circle $\tilde{L}$ over which consumers face a higher effective price than $\bar{p}$. If there is no set of consumers facing a higher price, as with a local deviation in which a firm lowers its price, no firms subsequently enter. For $\tilde{L}$ positive, subsequent entry is determined in the same way as the initial equilibrium.

All lowest-cost firms who have not yet entered will do so up to a cutoff firm $\tilde{a}$ given as the smallest marginal cost such that

$$(B.6) \quad \sum_{\bar{a} \leq a_j < \tilde{a}} (\tilde{p} - a_j)/\eta \geq \tilde{L} - \frac{2(\tilde{a} - \bar{a})}{\eta}$$

where $\tilde{p} = \tilde{a} + 2\sqrt{\eta f}$ and the last term reflects the fact that the set of consumer with effective price higher than $\tilde{p}$ is smaller than those above $\bar{p}$. Marginal firms are specified as above. For $\tilde{L}$ smaller than $2(a_{J+1} + 2\sqrt{\eta f} - a_J)/3\eta$ only firm J enters and sets a locally optimal price. Deviations in this subgame induce subgames with entry by firms with $a_j \geq \tilde{a}$ in identical fashion.

In the case where there are two regions with effective prices above $\bar{p}$, arising under a local deviation with a price increase, firms are allotted to both gaps with a common $\tilde{p}$ but with marginal firms picked to fill in both regions exactly.

B.3. Theorem 3.1. We can partition the space of stage 0 deviations into a) global deviations where deviation demand by consumers does not intersect equilibrium demand, b) local

deviations without undercutting where deviation and equilibrium demands intersect, but no firms are undercut by the deviation, and c) local deviations with undercutting where some firm is undercut.

Deviation a) is always unprofitable. Firm demand can be written in terms of the deviation price a firm charges, p_d , and the effective price faced by its boundary consumers on the left and right, p_L and p_R respectively. We have $D = (p_L - p_d + p_R - p_d)/\eta$ and so $\pi(p_d, p_L, p_R) = (p - a)\phi(p_L + p_R - 2p_d)/\eta$ which is strictly increasing in p_L and p_R . Then on-path profits for non-marginal firms are $\pi(p, \bar{p}, \bar{p}) \geq \pi(p_d, \bar{p}, \bar{p})$ by (3.2) and $\pi(p_d, \bar{p}, \bar{p}) \geq \pi(p_d, p_L, p_R)$ since $p_L, p_R < \bar{p}$ by construction of the equilibrium and the non-local deviation.

Marginal firms will not do better under global deviations without undercutting. They may do better under global deviation with undercutting since their price is not optimal by (3.2). In this case the deviation is made unprofitable by γ which needs to be greater than $\frac{5}{9}(\bar{p} - a_K)/\eta$.

Deviation b) is unprofitable when f is big enough or Δ is small enough. f positive implies a minimum size, and $\bar{p}$ chosen so that profits are exactly zero at entry for the marginal firm, means that a marginal reduction in p induces the cutoff firm to enter at the location of the boundary consumer and serve a strictly positive demand.⁸ Thus there is a discontinuity in a firm's profit from a conjectured deviation p_d .

Consider a firm's profit under a subgame in which it deviates and picks $p_d \in [(\bar{p}+a)/2, (3\bar{p}+a)/4]$, where the upper end of the range is the profit maximizing price without subsequent entry, and picks the optimal location with that price. There is a big drop in profits just above the equilibrium price where the $\bar{a}$ firm enters and serves demand $(\bar{p} - \bar{a})/2\eta = \sqrt{f/\eta}$ of which half is taken from the deviating firm. As p_d increases profits change continuously, possibly rising, until a point at which both the deviating firm a_d and the subsequent entrant $\bar{a}$ have $p_L = p_R = \tilde{a} \leq \bar{a} + \Delta$ where $\tilde{a}$ is the marginal cost of the new cutoff firm, the lowest

⁸At the optimal price, this firm is indifferent over locations in a neighborhood around the boundary location so long as its p_L and p_R do not permit subsequent entry. For simplicity I assume the firm always locates so that $p_L = p_R$ when optimal.

cost firm who does not enter. Suppose it enters in between firm $\bar{a}$ and a non-deviating firm so that it steals no demand from a_d . As p_d increases profit again moves smoothly although it is not as steep because firm $\bar{a}$ does not respond as dramatically to a fall in p_d now that it considers the subsequent reaction of firm $\tilde{a}$. The deviating firm's profit $\pi(p_d)$ reaches another kink when these firms have boundary expected prices equal to the next lowest cost firms break-even price, and the slope flattens again, possibly with a discontinuous drop if entry induces a loss in demand. As long as all of these prices are strictly suboptimal after all entry, we can pick R close enough to one so that they are no optimal deviations overall. Since f controls the size of the initial drop in demand and Δ the frequency of kinks at which new firms enter, we can always make f large enough or Δ small enough to prevent these deviations. This analysis holds for marginal firms as well.

In the case of deviation c) the firm undercuts another firm, and for this theorem we can simply pick γ larger than the largest possible value of this deviation. $\square$

B.4. Theorem 3.2. We use γ in eliminating deviations of type a) for marginal firms and deviations of type c) for both types of firms. For deviations of type a) for marginal firms we require $\gamma > \frac{5}{9}(\bar{p} - a_K)/\eta$, so it is enough to show that we can pick Δ and f small enough to generate an a_K arbitrarily close to $\bar{p}$.

The equilibrium picks a_K as close to $\bar{a}$ as possible as long as it satisfies

$$(B.7) \quad \sum_{a_K \leq a_j < \bar{a}} (\bar{p} - a_j)/3\eta \geq \sum_{j=1}^{J-1} (\bar{p} - a_j)/\eta - L$$

The right hand side is bounded above by $(J-2)(a_J - a_{J-1})/\eta + (\bar{p} - a_{J-1})/\eta$ which is less than $(J-2)\Delta/\eta + (2\sqrt{\eta f} + \Delta)/\eta$. The worst case scenario is that firms are very sparsely distributed, so that we have to pick a very low a_K in order to add up enough firms, when every pair of firm costs between $\bar{a}$ and a_K are Δ apart. Then there are at least $(\bar{a} - a_K)/\Delta$ firms losing an average demand of $(\bar{p} - a_k)/6\eta$ so that a_K is bounded below by the largest

value of a_K satisfying

$$(B.8) \quad \frac{\bar{a} - a_K}{\Delta} \cdot \frac{\bar{p} - a_k}{6\eta} \geq \frac{(J-2)\Delta}{\eta} + \frac{2\sqrt{\eta f} + \Delta}{\eta}$$

The right hand side is declining with Δ but increasing with J . We may want to take Δ to zero by taking limits as the number of firms becomes large, and J may become large under these circumstances. However, the right hand side becomes arbitrarily large as Δ goes to zero, so that we can take $a_K \rightarrow \bar{a}$ as long as $(J-1)\Delta^2 \rightarrow 0$. Thus we can take $a_K \rightarrow \bar{p} = \bar{a} + 2\sqrt{\eta f}$ as $f \rightarrow 0$ as well.

For deviations of type c) we can pick R close enough to 1 and Δ close enough to zero that maximum possible profit from a deviation of type c) is smaller than γ . As in figure 4b, the optimal deviation is to maintain the equilibrium price and change location in order to exactly undercut a neighbor, enjoy the extra demand from the hinterland, and enjoy profits $\pi(p, \bar{p}, p_R > \bar{p})$. To calculate this profit we need to consider the subgame equilibrium for additional entrants which determines the new p_R .

When a deviation leaves a region of consumers who face an effective price higher than $\bar{p}$, there is an opportunity of subsequent entry, but the pattern of entry depends on the size of this region. In essence, the subgame is no different than the original game where the entire circumference is the region in question. A line segment bounded by firms charging fixed low prices is equivalent. Thus if the region to fill in is distance m , we determine $\tilde{a}$ as we would $\bar{a}$ to satisfy $m = \sum_j d(a_j, \tilde{a})$ where the sum is over all j such that $a_j \in [\bar{a}, \tilde{a})$. The integer problem is even more dramatic here, but it is resolved the same way. For m very small, $\bar{a}$ just sets the optimal price, and for other cases the last one or two firms serve smaller demands at the prices given by equilibrium.

The deviating firm will earn better operating profits forever for m big enough simply because $\bar{a} < \tilde{a}$. Fixing m , we can bound $\tilde{a}$ using Δ . We have that $m \geq \sum_j d(a_j, \tilde{a} + 2\sqrt{\eta f})$ where $\tilde{a}$ is the largest marginal cost below $\tilde{a}$. Thus we have

$$m \geq \sum_j d(a_j, \tilde{a} + 2\sqrt{\eta f}) \geq \sum_j d(\tilde{a}, \tilde{a} + 2\sqrt{\eta f}) = \sum_j 2\sqrt{f/\eta} \geq \frac{\tilde{a} - \bar{a}}{\Delta} \cdot 2\sqrt{f/\eta}$$

$$\frac{\Delta m}{2\sqrt{f/\eta}} \geq \tilde{a} - \Delta - \bar{a} \quad \Rightarrow \quad \tilde{a} \leq \bar{a} + \Delta(m\sqrt{\eta}/2\sqrt{f} + 1)$$

Thus the firm's deviation profit is the first stage profit which can be made arbitrarily small by choice of R , and the profit in every subsequent stage which is less than or equal to equilibrium profit plus

$$(p - a) \cdot (\tilde{p} - \bar{p})/\eta = (p - a) \cdot (\tilde{a} - \bar{a})/\eta \leq \Delta(m\sqrt{\eta}/2\sqrt{f} + 1)$$

which can be made arbitrarily small by choice of Δ . Thus total surplus deviation profit can be brought below γ . $\square$

APPENDIX C. TRADE

In this note I solve the consumer's and firm's problem in the trade version of the model and characterize the equilibrium. In each country there is a continuum of industries indexed by $\omega \in [0, 1]$. In each of these industries N firms play the location game from section 3.

C.1. Consumers and Firms. Consumers maximize utility over a unit continuum of goods, each of which is supplied in a circular market:

$$\int_0^1 \left(\frac{c(\omega)^\rho - 1}{\rho} - \eta x_{ij}(\omega)^\sigma \right) d\omega \quad \text{s.t.} \quad \int_0^1 p(\omega) c(\omega) d\omega \leq w + \Pi$$

Fixing $j(\omega)$, the first order condition for c_j is

$$c_j(p, \lambda) = (\lambda p_j)^{\frac{1}{\rho-1}}$$

The consumer picks $j(\omega)$ by comparing her indirect utility at offered mill price p_j across firms. By the logic of the maximization problem in the model, consumers will purchase from

a firm with price p as long as they would not do better buying from a firm with price $\bar{a}$ at their own location $V(\bar{a}, 0)$. Because there are a continuum of goods the indirect utility is the utility from optimal consumption less a linear value of outside expenditure λ

$$(C.1) \quad V(p, x_{ij}) = u(c^*(\omega)) - \lambda p c^*(\omega)$$

$$(C.2) \quad = \frac{1}{\rho}((\lambda p_j)^{\frac{\rho}{\rho-1}} - 1) - \eta x_{ij}^\sigma - \lambda(p_j(\lambda p_j)^{\frac{1}{\rho-1}})$$

$$(C.3) \quad = \frac{1-\rho}{\rho}(\lambda p)^{\frac{\rho}{\rho-1}} - \frac{1}{\rho} - \eta x_{ij}^\sigma$$

Thus the maximum distance of a consumer purchasing from a firm charging price p is given by the indifference condition:

$$\frac{1-\rho}{\rho}(\lambda p)^{\frac{\rho}{\rho-1}} - \frac{1}{\rho} - \eta x_{max}^\sigma = \frac{1-\rho}{\rho}(\lambda \bar{a})^{\frac{\rho}{\rho-1}} - \frac{1}{\rho}$$

$$\frac{1-\rho}{\rho} \lambda^{\frac{\rho}{\rho-1}} (p^{\frac{\rho}{\rho-1}} - \bar{a}^{\frac{\rho}{\rho-1}}) = \eta x_{max}^\sigma$$

$$x_{max} = \left(\frac{1-\rho}{\eta \rho} \lambda^{\frac{\rho}{\rho-1}} (p^{\frac{\rho}{\rho-1}} - \bar{a}^{\frac{\rho}{\rho-1}}) \right)^{\frac{1}{\sigma}}$$

A firm serves all consumers up to this maximum distance in both directions so that *extensive* firm demand is

$$d(p; \bar{a}, \lambda) = 2 \left(\frac{(1-\rho)}{\eta \rho} \lambda^{\frac{\rho}{\rho-1}} (p^{\frac{\rho}{\rho-1}} - \bar{a}^{\frac{\rho}{\rho-1}}) \right)^{\frac{1}{\sigma}}$$

Note that although $c(p)$ goes to ∞ as p goes to zero for any parameters, extensive demand is bounded for $\rho < 0$ and unbounded for $\rho > 0$.

Profits are

$$\begin{aligned}
\pi(p; a, \bar{a}, \lambda) &= (p - a) \quad c(p, \lambda) \quad d(p; \bar{a}, \lambda) \\
\pi(p; a, \bar{a}, \lambda) &= (p - a) \quad (\lambda p)^{\frac{1}{\rho-1}} \quad 2 \left(\frac{1-\rho}{\eta\rho} \lambda^{\frac{\rho}{\rho-1}} (p^{\frac{\rho}{\rho-1}} - \bar{a}^{\frac{\rho}{\rho-1}}) \right)^{\frac{1}{\sigma}} \\
\pi(p; a, \bar{a}, \lambda) &= \left[\lambda^{\frac{\rho+\sigma}{\sigma(\rho-1)}} \left(\frac{1-\rho}{\eta\rho} \right)^{\frac{1}{\sigma}} 2 \right] (p - a) \quad p^{\frac{1}{\rho-1}} \quad (p^{\frac{\rho}{\rho-1}} - \bar{a}^{\frac{\rho}{\rho-1}})^{\frac{1}{\sigma}}
\end{aligned}$$

So the first order condition is

$$\frac{\pi}{p - a} + \frac{1}{\rho - 1} \frac{\pi}{p} + \frac{1}{\sigma} \frac{\pi}{p^{\frac{\rho}{\rho-1}} - \bar{a}^{\frac{\rho}{\rho-1}}} \frac{\rho}{\rho - 1} p^{\frac{1}{\rho-1}} = 0$$

Multiplying by $(1 - \rho) \cdot p/\pi$ we have

$$\frac{(1 - \rho)p}{p - a} - 1 - \frac{\rho}{\sigma} \frac{1}{1 - \left(\frac{\bar{a}}{p} \right)^{\frac{\rho}{\rho-1}}} = 0$$

$$\text{(C.4)} \quad \frac{p - a}{a/\rho - p} = \sigma \left(1 - \left(\frac{p}{\bar{a}} \right)^{\frac{\rho}{1-\rho}} \right)$$

For $\rho > 0$, the right hand side of this equation is always positive since $p < \bar{a}$. This implies $a < p < a/\rho$.

What happens to firm policies as $a \rightarrow 0$? For $\rho > 0$ we have $p \rightarrow 0$ by $a < p < a/\rho$ and the sandwich theorem. For $\rho < 0$ we can read the answer from (C.4). Setting $a = 0$ we have

$$\begin{aligned}
\frac{p_{min}}{-p_{min}} &= \sigma \left(1 - \left(\frac{p_{min}}{\bar{a}} \right)^{\frac{\rho}{1-\rho}} \right) \\
p_{min} &= \bar{a} \left(\frac{\sigma}{1 + \sigma} \right)^{\frac{\rho-1}{\rho}}
\end{aligned}$$

which is increasing in σ and decreasing in ρ . With price bounded above, both $c(p)$ and $d(p)$ are bounded. Then total demand D is bounded and employment aD goes to zero.

Last, notice that if p solves (C.4) for $(a, \bar{a})$ then θp solves the equations for $(\theta a, \theta \bar{a})$. thus the policy function $p^*(a, \bar{a})$ is homogeneous of degree 1, and we can write it as $aM(\bar{a}/a)$ where $M = p^*(1, \bar{a}/a)$ is a strictly increasing function with $M(1) = 1$ by inspection of (C.4).

C.2. Equilibrium and Trade. By symmetry and the assumption of i.i.d. locations, the total expenditure by the consumer is equal to the average consumer expenditure on any particular circle which I calculate as the extensive-demand-weighted average of consumer expenditure by firm:

$$\int_0^1 p_{j(\omega)} c(\omega) d\omega = N \int_0^{\bar{a}} p(a) c(a) \cdot \frac{d(a)}{L} H(da) + N \int_0^{\bar{a}(1-\tau)} p\left(\frac{a}{1-\tau}\right) c\left(\frac{a}{1-\tau}\right) \cdot \frac{d\left(\frac{a}{1-\tau}\right)}{L} H(da)$$

where I have split the integrals into domestically produced goods and imports, and recognized that foreign firms behave like domestic firms with marginal cost $a/(1-\tau)$.

Normalizing L to one, the consumer's budget constraint is

$$N \int_0^{\bar{a}} p(a) c(a) d(a) H(da) + N \int_0^{\bar{a}(1-\tau)} p\left(\frac{a}{1-\tau}\right) c\left(\frac{a}{1-\tau}\right) d\left(\frac{a}{1-\tau}\right) H(da) = 1 + \Pi + T$$

with

$$(C.5) \quad \Pi = N \int_0^{\bar{a}} (p(a) - a) c(a) d(a) H(da)$$

$$(C.6) \quad + N \int_0^{\bar{a}(1-\tau)} \left(p\left(\frac{a}{1-\tau}\right) (1-\tau) - a \right) c\left(\frac{a}{1-\tau}\right) d\left(\frac{a}{1-\tau}\right) H(da)$$

$$(C.7) \quad T = N \int_0^{\bar{a}(1-\tau)} \tau p\left(\frac{a}{1-\tau}\right) c\left(\frac{a}{1-\tau}\right) d\left(\frac{a}{1-\tau}\right) H(da)$$

Subtracting profits and transfers out of total expenditure, we have

$$N \int_0^{\bar{a}} a c(a) d(a) H(da) + N \int_0^{\bar{a}(1-\tau)} a c\left(\frac{a}{1-\tau}\right) d\left(\frac{a}{1-\tau}\right) H(da) = 1$$

substituting $\hat{a} = a/(1 - \tau)$ we have

$$\begin{aligned}
& N \int_0^{\bar{a}(1-\tau)} ac\left(\frac{a}{1-\tau}\right) d\left(\frac{a}{1-\tau}\right) \kappa a^{\kappa-1} da \\
&= N \int_0^{\bar{a}(1-\tau)} (1-\tau) \frac{a}{1-\tau} c\left(\frac{a}{1-\tau}\right) d\left(\frac{a}{1-\tau}\right) \kappa (1-\tau)^{\kappa-1} \left(\frac{a}{1-\tau}\right)^{\kappa-1} (1-\tau) \frac{da}{1-\tau} \\
&= N(1-\tau)^{\kappa+1} \int_0^{\bar{a}} \hat{a} c(\hat{a}) d(\hat{a}) \kappa \hat{a}^{\kappa-1} d\hat{a}
\end{aligned}$$

and the budget constraint becomes

$$\begin{aligned}
N(1 + (1 - \tau)^{\kappa+1}) \int_0^{\bar{a}} ac(a) d(a) H(da) &= 1 \\
N(1 + (1 - \tau)^{\kappa+1}) \int_0^{\bar{a}} a \cdot (\lambda a M)^{\frac{1}{\rho-1}} 2 \left(\frac{1-\rho}{\rho} \lambda^{\frac{\rho}{\rho-1}} ((aM)^{\frac{\rho}{\rho-1}} - \bar{a}^{\frac{\rho}{\rho-1}}) \right)^{\frac{1}{\sigma}} H(da) &= 1 \\
N(1 + (1 - \tau)^{\kappa+1}) \lambda^{\frac{\sigma+\rho}{\sigma(\rho-1)}} \eta^{-\frac{1}{\sigma}} \int_0^{\bar{a}} a \cdot (aM)^{\frac{1}{\rho-1}} \left(a^{\frac{\rho}{\rho-1}} \Phi(M^{\frac{\rho}{\rho-1}} - \left(\frac{a}{\bar{a}}\right)^{\frac{\rho}{\rho-1}}) \right)^{\frac{1}{\sigma}} \kappa a^{\kappa-1} da &= 1 \\
N(1 + (1 - \tau)^{\kappa+1}) \lambda^{\frac{\sigma+\rho}{\sigma(\rho-1)}} \eta^{-\frac{1}{\sigma}} \bar{a}^{\frac{\rho(\sigma+1)}{\sigma(\rho-1)} - \kappa} \int_0^{\bar{a}} \left(\frac{a}{\bar{a}}\right)^{\frac{\rho(\sigma+1)}{\sigma(\rho-1)}} \cdot M^{\frac{\sigma+\rho}{\sigma(\rho-1)}} \left(\Phi(1 - \left(\frac{aM}{\bar{a}}\right)^{\frac{\rho}{\rho-1}}) \right)^{\frac{1}{\sigma}} \kappa \left(\frac{a}{\bar{a}}\right)^{\kappa-1} \frac{da}{\bar{a}} &= 1
\end{aligned}$$

where $\Phi = 2^\sigma(1 - \rho)/\rho$ and the argument of M is suppressed for space. We can reduce the integral to a function of ρ and σ by substituting $\hat{a} = a/\bar{a}$:

$$\begin{aligned}
(C.8) \quad & \int_0^{\bar{a}} \left(\frac{a}{\bar{a}}\right)^{\frac{\rho(\sigma+1)}{\sigma(\rho-1)}} \cdot M^{\frac{\sigma+\rho}{\sigma(\rho-1)}} \left(\Phi(1 - \left(\frac{aM}{\bar{a}}\right)^{\frac{\rho}{\rho-1}}) \right)^{\frac{1}{\sigma}} \kappa \left(\frac{a}{\bar{a}}\right)^{\kappa-1} \frac{da}{\bar{a}} \\
&= \int_0^1 \hat{a}^{\frac{\rho(\sigma+1)}{\sigma(\rho-1)}} \cdot M(1/\hat{a})^{\frac{\sigma+\rho}{\sigma(\rho-1)}} \left(\Phi(1 - (\hat{a}M(1/\hat{a}))^{\frac{\rho}{\rho-1}}) \right)^{\frac{1}{\sigma}} \kappa \hat{a}^{\kappa-1} d\hat{a} = \Theta_1(\rho, \sigma)
\end{aligned}$$

so that we can write

$$(C.9) \quad N(1 + (1 - \tau)^{\kappa+1}) \lambda^{\frac{\sigma+\rho}{\sigma(\rho-1)}} \eta^{-\frac{1}{\sigma}} \bar{a}^{\frac{\rho(\sigma+1)}{\sigma(\rho-1)} - \kappa} \Theta_1(\rho, \sigma) = 1$$

Note that $\Theta_1(\rho, \sigma)$ depends only on the curvature parameters and not on the endogenous variables λ and $\bar{a}$.

The equation determining $\bar{a}$ ensures that the extensive demands sum to the circumference of the circle

$$N \int_0^{\bar{a}} d(a)H(da) + N \int_0^{a(1-\tau)} d\left(\frac{a}{1-\tau}\right)H(da) = 1$$

And by the same substitution argument this expression can be simplified to

$$\begin{aligned} N(1 + (1 - \tau)^\kappa) \int_0^{\bar{a}} d(a)H(da) &= 1 \\ N(1 + (1 - \tau)^\kappa) \int_0^{\bar{a}} 2 \left(\frac{1 - \rho}{\eta \rho} \lambda^{\frac{\rho}{\rho-1}} ((aM)^{\frac{\rho}{\rho-1}} - \bar{a}^{\frac{\rho}{\rho-1}}) \right)^{\frac{1}{\sigma}} \kappa a^{\kappa-1} da &= 1 \\ N(1 + (1 - \tau)^\kappa) \eta^{-\frac{1}{\sigma}} \lambda^{\frac{\rho}{\sigma(\rho-1)}} \int_0^{\bar{a}} a^{\frac{\rho}{\sigma(\rho-1)}} \left(\Phi(M^{\frac{\rho}{\rho-1}} - \left(\frac{\bar{a}}{a}\right)^{\frac{\rho}{\rho-1}}) \right)^{\frac{1}{\sigma}} \kappa a^{\kappa-1} da &= 1 \end{aligned}$$

Using the same substitution argument as we did for the budget constraint we can write

$$\begin{aligned} \int_0^{\bar{a}} a^{\frac{\rho}{\sigma(\rho-1)}} \left(\Phi(M^{\frac{\rho}{\rho-1}} - \left(\frac{\bar{a}}{a}\right)^{\frac{\rho}{\rho-1}}) \right)^{\frac{1}{\sigma}} \kappa a^{\kappa-1} da &= \\ \bar{a}^{\frac{\rho}{\sigma(\rho-1)}} \bar{a}^\kappa \int_0^{\bar{a}} \left(\frac{a}{\bar{a}}\right)^{\frac{\rho}{\sigma(\rho-1)}} \left(\Phi(M^{\frac{\rho}{\rho-1}} - \left(\frac{\bar{a}}{a}\right)^{\frac{\rho}{\rho-1}}) \right)^{\frac{1}{\sigma}} \kappa \left(\frac{a}{\bar{a}}\right)^{\kappa-1} \frac{da}{\bar{a}} &= \\ \bar{a}^{\frac{\rho}{\sigma(\rho-1)}} \bar{a}^\kappa \int_0^1 \hat{a}^{\frac{\rho}{\sigma(\rho-1)}} \left(\Phi(M(1/\hat{a})^{\frac{\rho}{\rho-1}} - \hat{a}^{\frac{\rho}{1-\rho}}) \right)^{\frac{1}{\sigma}} \kappa \hat{a}^{\kappa-1} d\hat{a} &= \bar{a}^{\frac{\rho}{\sigma(\rho-1)} + \kappa} \Theta_2(\rho, \sigma) \end{aligned}$$

Thus we have reduced the equilibrium to two equations, the simplified budget and extensive demand constraints, and two unknowns, λ and $\bar{a}$.

$$(C.10) \quad \lambda^{\frac{\sigma+\rho}{\sigma(\rho-1)}} \cdot \bar{a}^{\frac{\rho(\sigma+1)}{\sigma(\rho-1)}} \cdot [N \cdot (1 + (1 - \tau)^{\kappa+1}) \cdot \eta^{-\frac{1}{\sigma}} \cdot \bar{a}^\kappa \cdot \Theta_1(\rho, \sigma)] = 1$$

$$(C.11) \quad \lambda^{\frac{\rho}{\sigma(\rho-1)}} \cdot \bar{a}^{\frac{\rho}{\sigma(\rho-1)}} \cdot [N \cdot (1 + (1 - \tau)^\kappa) \cdot \eta^{-\frac{1}{\sigma}} \cdot \bar{a}^\kappa \cdot \Theta_2(\rho, \sigma)] = 1$$

We can solve (C.11) for λ :

$$\begin{aligned}
1 &= \lambda \cdot \bar{a} \cdot [N \cdot (1 + (1 - \tau)^\kappa) \cdot \eta^{-\frac{1}{\sigma}} \cdot \bar{a}^\kappa \cdot \Theta_2(\rho, \sigma)]^{\frac{\sigma(\rho-1)}{\rho}} \\
\lambda &= \bar{a}^{-1} \cdot [N \cdot (1 + (1 - \tau)^\kappa) \cdot \eta^{-\frac{1}{\sigma}} \cdot \bar{a}^\kappa \cdot \Theta_2(\rho, \sigma)]^{\frac{\sigma(1-\rho)}{\rho}} \\
\lambda^{\frac{\sigma+\rho}{\sigma(\rho-1)}} &= \bar{a}^{-\frac{\sigma+\rho}{\sigma(\rho-1)}} \cdot [N \cdot (1 + (1 - \tau)^\kappa) \cdot \eta^{-\frac{1}{\sigma}} \cdot \bar{a}^\kappa \cdot \Theta_2(\rho, \sigma)]^{-\frac{\sigma+\rho}{\rho}}
\end{aligned}$$

Plugging into (C.10) we have

$$\begin{aligned}
\lambda^{\frac{\sigma+\rho}{\sigma(\rho-1)}} \cdot \bar{a}^{\frac{\rho(\sigma+1)}{\sigma(\rho-1)}} \cdot [N \cdot (1 + (1 - \tau)^{\kappa+1}) \cdot \eta^{-\frac{1}{\sigma}} \cdot \bar{a}^\kappa \cdot \Theta_1(\rho, \sigma)] &= 1 \\
\bar{a} \cdot (N \cdot \eta^{-\frac{1}{\sigma}} \cdot \bar{a}^\kappa)^{-\frac{\sigma}{\rho}} \frac{(1 + (1 - \tau)^{\kappa+1})}{(1 + (1 - \tau)^\kappa)^{\frac{\sigma+\rho}{\rho}}} \cdot \Theta_3(\rho, \sigma) &= 1 \\
N^{-\frac{\sigma}{\rho}} \frac{(1 + (1 - \tau)^{\kappa+1})}{(1 + (1 - \tau)^\kappa)^{\frac{\sigma+\rho}{\rho}}} &= \bar{a}^{\frac{\kappa\sigma}{\rho}-1} \cdot \eta^{-\frac{1}{\rho}} \cdot \Theta_3(\rho, \sigma)^{-1} \\
(C.12) \quad N \cdot (1 + (1 - \tau)^\kappa) \cdot \left(\frac{1 + (1 - \tau)^\kappa}{1 + (1 - \tau)^{\kappa+1}} \right)^{\frac{\rho}{\sigma}} &= \bar{a}^{\frac{\rho}{\sigma}-\kappa} \cdot \eta^{\frac{1}{\sigma}} \cdot \Theta_4(\rho, \sigma)
\end{aligned}$$

This equation determines $\bar{a}$ as a function of τ and other parameters. How does $\bar{a}$ change with τ ? We can show that it moves in the same direction for $\rho/\sigma > -1/\tau$, which is true whenever $\rho > -\sigma$ and thus for all elastic intensive demand curves. Define $\varepsilon_{\bar{a}}$ to be the elasticity of $\bar{a}$ with respect to τ , and similarly define

$$\begin{aligned}
\varepsilon_\kappa &= \frac{\partial(1+(1-\tau)^\kappa)}{\partial\tau} \cdot \frac{\tau}{1+(1-\tau)^\kappa} = -\frac{\tau \cdot \kappa \cdot (1-\tau)^{\kappa-1}}{1+(1-\tau)^\kappa} \\
\varepsilon_{\kappa+1} &= \frac{\partial(1+(1-\tau)^{\kappa+1})}{\partial\tau} \cdot \frac{\tau}{1+(1-\tau)^{\kappa+1}} = -\frac{\tau \cdot (\kappa+1) \cdot (1-\tau)^\kappa}{1+(1-\tau)^{\kappa+1}}
\end{aligned}$$

Then taking the elasticity of both sides of (C.12) yields

$$(C.13) \quad \varepsilon_{LHS} = \left(1 + \frac{\rho}{\sigma}\right) \varepsilon_\kappa - \frac{\rho}{\sigma} \varepsilon_{\kappa+1} = \left(\frac{\rho}{\sigma} - \kappa\right) \varepsilon_{\bar{a}}$$

We can calculate ε_{LHS} :

$$\begin{aligned}\varepsilon_{LHS} &= -\left(1 + \frac{\rho}{\sigma}\right) \frac{\tau \cdot \kappa \cdot (1 - \tau)^{\kappa-1}}{(1 + (1 - \tau)^\kappa)} + \frac{\rho}{\sigma} \cdot \frac{\tau \cdot (\kappa + 1) \cdot (1 - \tau)^\kappa}{(1 + (1 - \tau)^{\kappa+1})} \\ \varepsilon_{LHS} &= -\tau \cdot (1 - \tau)^{\kappa-1} \cdot \left(\left(1 + \frac{\rho}{\sigma}\right) \frac{\kappa}{(1 + (1 - \tau)^\kappa)} - \frac{\rho}{\sigma} \cdot \frac{(\kappa + 1)(1 - \tau)^\kappa}{1 + (1 - \tau)^{\kappa+1}} \right)\end{aligned}$$

Which is negative when

$$\begin{aligned}\kappa(1 + \frac{\rho}{\sigma}) &> \frac{\rho}{\sigma}(\kappa + 1)(1 - \tau) \\ \kappa &> \frac{\rho}{\sigma}(1 - \tau) - \kappa \frac{\rho}{\sigma} \tau\end{aligned}$$

which is true whenever $\rho > 0$ since $\kappa > \rho/\sigma$ by specification of κ . For $\rho < 0$ we can rearrange the expression to

$$\begin{aligned}\kappa(1 + \frac{\rho}{\sigma}\tau) &> \frac{\rho}{\sigma}(1 - \tau) \\ \kappa(\frac{\sigma}{\rho} + \tau) &< (1 - \tau) \\ \frac{\sigma}{\rho} &< (1 - \tau)/\kappa - \tau\end{aligned}$$

Which is true when $\kappa \leq (1 - \tau)/\tau$ because the right hand side is positive, otherwise

$$\frac{\rho}{\sigma} > -\frac{1}{\tau} = -\frac{\kappa}{\tau(\kappa + 1) - \tau} > -\frac{\kappa}{\tau(\kappa + 1) - 1}$$

Thus the elasticity is negative whenever $\rho/\sigma > -1/\tau$, and the left hand side of (C.12) is decreasing in τ . Then by (C.12), the cutoff marginal utility $\bar{a}$ is increasing in τ as long as $\kappa > \rho/\sigma$ which we have seen above to be true by our bounds on κ .

Lastly, what is the effect of reducing τ on the total number of firms? We can rearrange (C.12) to get

$$(C.14) \quad N \cdot (1 + (1 - \tau)^\kappa) \cdot \bar{a}^\kappa = \bar{a}^{\frac{\rho}{\sigma}} \cdot \left(\frac{(1 + (1 - \tau)^{\kappa+1})}{1 + (1 - \tau)^\kappa} \right)^{\frac{\rho}{\sigma}} \cdot \eta^{\frac{1}{\sigma}} \cdot \Theta_4(\rho, \sigma)$$

where the left hand side is the expression for the number of operating firms. It is not clear from this expression what happens to the number of firms because $(1 - \tau)^\kappa$ increases with a reduction in τ but $\bar{a}^\kappa$ decreases. We show below that under a reduction in τ the number of operating firms decreases when $\rho > 0$.

From (C.13) we have

$$\begin{aligned} \left(\frac{\rho}{\sigma} - \kappa\right)\varepsilon_{\bar{a}} &= \left(1 + \frac{\rho}{\sigma}\right)\varepsilon_\kappa - \frac{\rho}{\sigma}\varepsilon_{\kappa+1} \\ \varepsilon_{\bar{a}} &= \frac{1 + \frac{\rho}{\sigma}}{\frac{\rho}{\sigma} - \kappa}\varepsilon_\kappa - \frac{\frac{\rho}{\sigma}}{\frac{\rho}{\sigma} - \kappa}\varepsilon_{\kappa+1} \end{aligned}$$

Then the elasticity of the number of operating firms, ε_f , given by the elasticity of the left hand side of (C.14) is

$$\begin{aligned} \varepsilon_f &= \frac{\rho}{\sigma} \left(\varepsilon_{\bar{a}} + \varepsilon_{\kappa+1} - \varepsilon_\kappa \right) \\ &= \frac{\rho}{\sigma} \left(\left(1 - \frac{\frac{\rho}{\sigma}}{\frac{\rho}{\sigma} - \kappa}\right)\varepsilon_{\kappa+1} + \left(\frac{1 + \frac{\rho}{\sigma}}{\frac{\rho}{\sigma} - \kappa} - 1\right)\varepsilon_\kappa \right) \\ &= \frac{\rho}{\sigma} \left(\frac{\kappa}{\kappa - \frac{\rho}{\sigma}} \right) \left(\varepsilon_{\kappa+1} - \left(\frac{\kappa + 1}{\kappa} \right) \varepsilon_\kappa \right) \\ &= \frac{\rho}{\sigma} \left(\frac{\kappa}{\kappa - \frac{\rho}{\sigma}} \right) \left(- \frac{\tau(\kappa + 1)(1 - \tau)^\kappa}{1 + (1 - \tau)^{\kappa+1}} + \left(\frac{\kappa + 1}{\kappa} \right) \frac{\tau\kappa(1 - \tau)^{\kappa-1}}{1 + (1 - \tau)^\kappa} \right) \\ &= \frac{\rho}{\sigma} \left(\frac{\kappa}{\kappa - \frac{\rho}{\sigma}} \right) \tau(\kappa + 1)(1 - \tau)^{\kappa-1} \left(\frac{-(1 - \tau)}{1 + (1 - \tau)^\kappa} + \frac{1}{1 + (1 - \tau)^\kappa} \right) \end{aligned}$$

This expression takes the same sign as ρ . Thus the number of operating firms declines with τ for $\rho > 0$ and increases under a tariff reduction for $\rho < 0$.