Innovation and Distribution: A General Equilibrium Model of Manufacturing and Retailing

Bart J. Bronnenberg*

January 19, 2017

Abstract

I propose a general equilibrium model of competition in manufacturing and retailing. Relative to the counterfactual of direct sales by manufacturers, the retail sector increases manufacturing entry and produced variety. Although double marginalization in the sales channel raises prices and hurts consumers in quantity, in equilibrium it increases consumed variety and convenience, both valued positively. Consistent with observations, the equilibrium predicts that the size of the retail sector is a constant fraction of the total economy across nations of differing size and wealth.

JEL: D11, D21, L13, L81, M31

*Professor of Marketing, Tilburg University, The Netherlands and Research Fellow at CEPR, London, United Kingdom. I thank Dan Ackerberg, Jeff Campbell, Paul Ellickson, Matthew Gentzkow, Leslie Marx, Sanjog Misra and John Sutton for input and discussion. I am also grateful to seminar participants at Humboldt University, London School of Economics, MIT, Ohio State University, Stanford University, Tilburg University and UCLA for comments and questions. Finally, I thank the Netherlands Foundation for Scientific Research (NWO Vici Grant) for financial support. All errors are mine.

1 INTRODUCTION

The distributive trades, largely consisting of retailing and wholesaling, constitute a large sector of the economy, accounting for 21% of the gross domestic product (GDP) of a typical country.[1] The added value, function, and size of the distribution sector have been a focus of debate and discussion since at least Shaw (1912), and continues to attract attention in economics (Anderson and van Wincoop, 2004; Hortaçsu and Syverson, 2015) and marketing (Shaw, 1990; Srinivasan and Hanssens, 2009).

The distributive trades have historically been considered part of the marketing function of the firm. As Braithwaite and Dobbs (1932, p.5) write:

> "The high level of marketing costs [...] is primarily one of the necessary corollaries of our present industrial organization. The main features of that organization are large-scale specialized production, a high degree of geographical concentration of industry, and consequently a wide separation of producer from consumer. The resulting economies in production costs are familiar to all, but they are offset in part, though by no means entirely, by the enormously increased cost of distribution which accompany them."

However, a framework intended to evaluate how the market trades off economies of scale in manufacturing and the costs of retailing and wholesaling has heretofore remained largely absent from the economic literature.[2] Instead, the literature has studied the provision of manufactured product variety ("Chamberlinian" variety) and store density ("Hotelling" variety) mostly in isolation of each other (see, e.g., Lancaster, 1990).

The central issue studied in this paper is then how the market allocates scarce resources to the costly production of manufactured variety and to the costly entry of retailers. Additional economic questions about the distributive trades emerge from the identity that manufactured and distributed variety depend on each other. For instance, whereas a typical supermarket in the United States carried 9,000 items in 1975, it carries 47,000 items today (Consumer Reports, 2014). Does this expansion in retail assortment affect variety in manufacturing?[3] Relatedly, if retailers raise

[1] Using country level data from the UN National Accounts database, this share is computed as the total 2010 expenditure for retail, wholesale, and transportation & communication (ISIC G + H + I) relative to 2010 GDP. Reported is the GDP weighted average of the share across all countries. The unweighted average is 22%.

[2] The literature on trade (e.g., Krugman, 1979) considers the role of transportation cost on scale economies in manufacturing.

[3] Indeed, the *Directorate General Competition* of the European Commission recently investigated whether compe-

consumed variety and leisure to consumers relative to a counterfactual of manufacturers selling directly, should the economic debate about double marginalization focus almost completely on hurting welfare through price and quantity?

To provide some answers to these questions, I propose a free-entry model of manufacturers and retailers selling to variety loving consumers. Manufacturers, producing a single variety, compete on wholesale price and supply stores. Retailers compete on assortment variety and on final prices. Using this model, I study how intra-sector competition and inter-sector dependence affects product variety, store density, and prices. I find the following.

First, under very mild conditions, a manufacturer finds it unprofitable to sell directly to distant consumers who can buy a variety of substitute goods at a nearby retailer. In this case, retailers distribute the full assortment of varieties produced by manufacturers. Relative to the counterfactual of direct sales by manufacturers, the retail sector raises manufactured variety in two ways. Retailers lower the consumer's transaction cost associated with buying more variety, and thereby raise demand for it. This gives retailers an incentive to compete with assortment –consistent with the expansion of supermarket assortment noted above– ultimately resulting in more entry in manufacturing. Next, and perhaps more fundamentally, the distributive trades reduce competition among manufacturers. Distribution makes a manufacturer's variety available to all consumers in the market. This reduces the incentive to compete over the extensive consumer margin relative to the counterfactual of selling directly to nearby consumers. Left to compete only over the intensive consumer margin, manufacturers command higher margins than when selling directly to consumers, and more of them enter.

Second, the welfare associated with profit maximizing manufacturers and retailers, each setting their own margins, is higher than the social planner's welfare maximum with manufacturers selling directly to consumers. Hence, the presence of double margins does not hurt welfare per se. Welfare is surprisingly close, however not equal, to the welfare maximum in the two-sector model of manufacturing and retailing. Double marginalization in my model raises prices relative to the planner's choices. But, the effect of reduced quantities on consumer welfare is buffered by increased product variety and increased convenience.

tition among retailers hinders or enables product innovation and manufactured variety (European Commission, 2014). Lacking data on a counterfactual distribution system, the investigation remains however largely inconclusive.

Third, the proposed model endogenizes pricing power from entry primitives. In particular, margins in both sectors are tightly linked to what restricts consumption variety more: the setup cost of manufacturing a variety or the store-entry cost of distributing it. For instance, if opening a store with a large assortment is expensive, retailers carry few varieties. Manufacturer margins are now eroded from intra-sector competition over access to retailers' shelves. Conversely, if starting a plant is expensive, few manufacturers will enter and retailers will not be able to procure enough varieties. Manufacturers, assured of distribution, now charge high wholesale prices. However, not being able to compete on assortment, retailer margins will be suppressed from stronger competition on price.

Finally, circling back to the debate about the prominence of the distributive trades, the model predicts that its size is a constant fraction of the economy, e.g., that it is the same fraction of GDP in the United States as it is in, say, Venezuela. Next, using consensus estimates in the literature about the value of variety to consumers, the model rationalizes the observed size of the distributive trades at approximately 20% of GDP. Finally, using data from the UN National Accounts Data Base, the analysis also provides empirical support for the prediction that the relative size of the distributive trades does not depend on wealth, consumer travel costs, or firm entry costs.

The remainder of the paper is organized as follows. Section 2 discusses several streams of related literature and how the paper contributes to these. Section 3 contains the model of consumers, retailers, and manufacturers. Section 4 presents the general equilibrium and the comparative statics. Section 5 discusses the role of the retail sector and uses the equilibrium to make predictions about its size. Section 6 concludes.

2 LITERATURE

My paper draws from the literature in several ways. There is a large literature on love of variety (Dixit and Stiglitz, 1977; Lancaster, 1971; Spence, 1976) and satiation (Hartmann, 2006; Kahn, Kalwani, and Morisson, 1986; McAlister, 1982). I contribute to this literature, by proposing there are demand side factors such as time spent purchasing and other inconveniences that limit the demand for variety. My consumer model additionally draws from the literature on labor supply and home production (Becker, 1965; Gronau, 1977; Pryor, 1977), which I apply to answer how much effort consumers invest in shopping for variety. The complete model provides a basis for welfare analysis of retailing as a solution to address the consumers' time costs of using the market.

Next, retail power has been studied in the context of information advantages, retailer size (Chen, 2003; Dobson and Waterson, 1997), and retail mergers (Inderst and Shaffer, 2007; Staelin and Lee, 2014). Part of this literature has considered that upstream variety is lowered because of downstream retailer power or shelf space constraints (e.g., Inderst and Shaffer, 2007; Marx and Shaffer, 2010). Complementary to this, I find that upstream variety is lowered when the retailer optimally chooses a smaller assortment. With a low need for variety in downstream competition, retailers stock the varieties with the lowest wholesale price, thereby eroding the pricing power of manufacturers. Vertical power has also been studied in the context of price bargaining. Using a generalized Nash bargaining game, Crawford and Yurukoglu (2012) and Draganska, Klapper, and Villas-Boas (2010) measure vertical power in a retail setting. Similar analyses have been presented elsewhere. My study adds to this literature by discussing how the division of channel surplus and bargaining power are outcome variables rooted in entry primitives and technological advantages.

Additionally, the literature in industrial organization on vertical restraints discusses how retailing affects extensive versus intensive margin competition among manufacturers. In particular, Rey and Tirole (1986) argue the use of exclusive sales territories as a means to reduce competition among manufacturers. My argument is complementary and focuses on how the retail system reduces the manufacturer's geographic extensive consumer margin by offering market coverage. There is a related literature in marketing on how retailing softens manufacturer competition when products are closer substitutes (McGuire and Staelin, 1983; Moorthy, 1988). My model predicts more entry in manufacturing independently of the degree of product substitution.

Lastly, there is a large literature on contracting and vertical channels in economics and marketing. A central theme in this literature is double marginalization. A (very small) selection of this literature includes Rey and Tirole (1986), Iyer (1998), and Villas-Boas (2007). I contribute to this literature by showing that double marginalization in the free-entry equilibrium, although it leads to higher prices, also generally leads to more entry in both sectors, i.e., more manufactured variety and more convenience.

3 The Model

3.1 Overview

A brief overview of my model is as follows. Consumers with a mass s are uniformly dispersed on a unit circle, and derive utility from quantities of goods, the variety of these goods, and leisure. Product variety is costly to obtain for consumers because it involves travel to retailers, or counterfactually to direct-selling manufacturers.

The distributive trades consist of one type of agent, retailers, who sell an assortment of varieties to consumers. Retailers compete on final prices and on assortment variety, defined as the mass of unique varieties that each of them sells. They enter dispersed on the circle and their entry cost rises in the amount of variety they choose to sell.

Manufacturers produce a single unique variety and compete on wholesale prices, taking downstream retailer actions into account. They enter dispersed on the circle[4] at a fixed setup cost. Manufacturers have the option to either sell through retailers or sell directly to consumers (in this case they set final prices). Entry in both sectors takes place until profits are driven to zero in each.

Finally, the model accounts for the possibility that entry and variety in either sector potentially pose a constraint on decisions in the other. I am interested in pure strategy equilibria with symmetric retailers and symmetric manufacturers.[5]

Relative to the modeling tradition in spatial competition and Hotelling variety (e.g., Salop, 1979), my model endogenizes store-entry costs and features consumers with a love of variety. Relative to the modeling tradition in monopolistic competition and Chamberlinian variety (e.g., Dixit and Stiglitz, 1977; Spence, 1976), my model features transaction costs associated with purchasing and distributing variety. These costs limit downstream demand for and the availability of variety, respectively.

[4]The assumption of manufacturer dispersion is not essential to the results. It allows me to derive demand for variety for the counterfactual of manufacturers selling directly to consumers in a simple way using the same demand primitives as in the full two-sector model.

[5]Although the symmetric case is limiting it simplifies the analysis, and –to paraphrase Dixit and Stiglitz (1977)– even the symmetric case yields interesting results.

3.2 Demand

Agents in my model have a single resource, time, which they use for three activities: (1) earning income at a wage rate of $w = 1$ (which is the numeraire), (2) purchasing consumption goods, and (3) leisure. Income Y (and time worked) is assumed fixed and is available to buy quantities of goods. The time remaining after work, T, is available for leisure and for using the market, i.e., purchasing varieties, at the cost of time lost in travel. These assumptions recognize that time spent outside the labor market can be different from leisure (see, e.g., Becker, 1965; Gronau, 1977) and that purchasing activity is a form of non-market labor. This idea implies that demand for leisure and Marshallian demand for goods are related through choices about purchasing effort.

The consumer derives utility from a composite good $X(V)$ consisting of quantities $x(v)$ of varieties v contained in an endogenous purchase set V, and from leisure time $L(V)$ left after shopping for this purchase set V. Utilities for the composite good and for leisure are combined in a Cobb-Douglas utility function:

$$U(X(V), L(V)) = X(V)^{1-\rho} \times L(V)^{\rho}, \tag{1}$$

with ρ equal to the preference for leisure. Next, the utility for the composite good $X(V)$ follows a CES specification over a continuum of varieties, with elasticity of substitution $\sigma > 1$, i.e.,

$$X(V) = \left(\int_{v \in V} x(v)^{\frac{\sigma-1}{\sigma}} dv \right)^{\frac{\sigma}{\sigma-1}}. \tag{2}$$

This utility function is positive in quantities $x(v)$ and in variety, i.e., the mass of V.

The utility for leisure $L(V)$ is leisure itself. If $\tau(V)$ is the travel cost associated with buying the varieties in V then $L(V) = T - \tau(V)$. Assuming a retailer is within reach, consumers can travel there and buy each variety at zero marginal travel cost. On the other hand, when no retailer is within reach, the consumer can visit multiple manufacturers in a single trip and buy variety at a non-zero marginal cost of travel to the next manufacturer. Thus, leisure is the time remaining after traveling the shortest distance to a store, or counterfactually the shortest path connecting a set of manufacturers in absence of a store. The time it takes to make a round trip is t per unit of distance.

The consumer's problem can now be formalized as making choices about quantities, X, the purchase set V, and leisure L, subject to separate income and time constraints,

$$\max_{X,V,L} U\left(X\left(V\right), L\left(V\right)\right), \text{ s.t.} T = \tau\left(V\right) + L\left(V\right) \text{ and } Y = \int_{v \in V} p\left(v\right) x\left(v\right) dv,$$

in which the price for a unit of a variety v is $p\left(v\right)$. Unless stated differently, each consumer is fully informed about all varieties, travel costs, and prices.

I will now discuss the different demand variables –quantity, variety, and leisure– when (1) varieties are only sold by retailers, (2) varieties are only sold by manufacturers, and (3) a single variety is sold by a manufacturer and all others by retailers.

All varieties are sold by retailers. If stores are symmetric except for location, a consumer chooses the nearest one, located at a distance d from her residence. In Appendix A.1, it is shown that quantity demand for each single variety v that is part of the retailer's assortment is,

$$x\left(v\right) = \bar{x} \times \left(\frac{p\left(v\right)}{p}\right)^{-\sigma}, \ v \in V_N, \tag{3}$$

where V_N is the retailers' assortment, with mass N, and $\bar{x}$ is the average per-variety quantity $\frac{Y}{Np}$. Next, $p\left(v\right)$ and p are the retail prices of the variety v and of all other varieties, respectively.

Consumers buy all varieties in the store because their utility rises in variety and the transaction cost for the marginal variety in the store is 0.

The travel involved in purchasing a mass of N varieties is a round trip to the store located at a distance of d. Leisure after purchasing all varieties at this store is

$$L = T - td. \tag{4}$$

All varieties are sold by manufacturers. Counterfactually, in a market without retailers, consumers travel to manufacturers who sell directly to them. Buying more variety is now costly at the margin because it involves more consumer travel. Quantity demand for variety v at an off-factory price of $p(v)$, is

$$x(v) = \begin{cases} \bar{x} \times \left(\frac{p(v)}{p}\right)^{-\sigma} & \text{if} \quad v \in V_\delta \\ \\ 0 & \text{if} \quad v \notin V_\delta \end{cases} \tag{5}$$

with, as before, $\bar{x}$ representing average quantities, but now at off-factory prices and over a different purchase set, and p representing the price charged by other manufacturers.

The set of varieties purchased, V_δ, is determined by the consumer trade-off between variety and leisure. With M manufacturers dispersed uniformly, the marginal travel cost per variety is constant and equal to t times the inter-manufacturer distance $\frac{1}{M}$. In Appendix A.2, it is shown that the optimal purchase set for a given consumer i equals all manufacturers located on a circle segment that includes i's residence and has a length of

$$\delta = \frac{T}{t}\left(\frac{1-\rho}{1+\rho\sigma-2\rho}\right), \tag{6}$$

with $\delta \leq 1$. The consumer makes longer shopping trips when her time resource T is higher, and when the preference for leisure ρ, the elasticity of substitution σ, and the travel cost t are lower. The optimal trip length δ does not depend on the density of manufacturers. The demand for variety is the total number of manufacturers located in this segment, δM. The average purchase quantity $\bar{x}$ in equation (5) is equal to $\frac{Y}{\delta M p}$.

Because of travel costs, consumers buy a local subset of varieties. The introduction of travel costs therefore generalizes the representative consumer approach in Dixit and Stiglitz (1977) to one where different consumers prefer different products and one that is more in line with Hart (1985) and Wolinsky (1986).

Leisure after shopping is equal to $L = T - \delta t$. Making the substitution for δ, leisure simplifies to

$$L = T\rho\left(\frac{\sigma-1}{1+\rho\sigma-2\rho}\right). \tag{7}$$

A single variety is sold by a manufacturer but all others by retailers. The final demand scenario considers that a manufacturer can bypass the retailer and sell directly to consumers. This opens up the possibility of many hybrid channel configurations of direct and intermediated sellers. Instead of considering all of these, I develop a simple existence condition for an equilibrium where manufacturers all sell through retailers and next demonstrate the mildness of this condition. For this condition, I derive the demand for a single manufacturer who sells directly to consumers when all others are selling through the retailer.[6]

Characterizing demand in this case starts with observing that consumers will not travel a non-marginal distance to buy a marginal variety, i.e., that a consumer will not detour to buy directly from a single manufacturer. This means consumers will only buy directly from a manufacturer if it is located on the way to the nearest retailer. Per consequence, such a manufacturer derives its entire demand from a fraction of the clientele of its single closest retailer. Assuming the retailer sells a mass of N varieties at a price of p, the consumer's quantity demand for the direct seller is,

$$x = \begin{cases} \bar{x} \times \left(\frac{p_m}{p}\right)^{-\sigma} & \text{if } m \text{ is on the route to the retailer} \\ \\ 0 & \text{else} \end{cases}, \tag{8}$$

with $\bar{x}$ representing average quantities $\frac{Y}{Np}$, and p_m is the price of the manufacturer selling directly to consumers. Leisure is the same as in the case of visiting a single retailer, i.e., equation (4).

3.3 Supply Technology

Retailers In the tradition of the literature on Hotelling variety (e.g., Salop, 1979) each retailer pays a fixed entry cost to open a store. Larger assortments often involve contracting with more manufacturers, building and maintaining a bigger size store with more shelf-space, etc. Therefore, I allow this cost to rise in my model with the mass of varieties that the retailer distributes. Specifically, I assume that the entry cost, $g(N)$, for a retailer with a mass N varieties in its assortment is

$$g(N) = \gamma N^k, \tag{9}$$

where k increases the marginal cost of adding variety and γ is a cost shifter ($\{k, \gamma\} > 0$). I am particularly interested in the case where $k = 1$ although some solutions exist for $k \neq 1$. The linear

[6] An additional reason for this strategy is that consumers face ancillary transaction costs buying directly from manufacturers, e.g., negotiation, freight, and check-out costs. Such costs, although not considered here, strengthen the appeal to consumers of retailers selling a mass of varieties at a single check-out cost and weaken interest in direct sellers. A final reason is analytical simplicity.

case $k = 1$ is the closest parallel to the counterfactual cost structure for distribution where consumers spend time traveling to spatially dispersed manufacturers and purchase variety at linearly increasing travel costs.

Retailers are discrete. This is to change as little as possible from Salop (1979) and to introduce horizontal differentiation between two retailers. I ignore the integer problem.

Manufacturers In the tradition of the monopolistic competition literature on Chamberlinian variety (e.g., Dixit and Stiglitz, 1977; Spence, 1976), I assume each manufacturer enters at a fixed entry cost f and produces a unit of its variety at a constant marginal cost of c. The entry cost of the manufacturer includes a transportation technology to stock all retailers at no additional cost. The assumption that retailing is costly and transportation is inexpensive is motivated by their observed shares of the total expenditure on the distributive trades.[7]

I also assume that manufacturers themselves do not distribute their single variety to individual consumers because this would be too costly to the manufacturer in terms of total transaction cost. Therefore, in the counterfactual scenario where manufacturers sell direct, they do so "off-factory" and consumers bear the cost of purchasing in the form of travel.

An alternative counterfactual assumption is that manufacturers use the service of distributors to deliver the variety at a fixed shipping charge paid by consumers. As in my counterfactual assumption, this makes variety costly at the margin. Charging consumers for the delivery of a variety is therefore similar to consumers traveling to individual manufacturers, in the sense of both limiting the demand for variety.[8]

[7] Across all countries, transportation, storage and communication (ISIC I) represented 7.7% of GDP in 2010, with only minor part of this share being actual freight costs. In contrast, 13.1% of GDP is due to retailing, wholesaling and restaurants (ISIC G + H). To this number needs to be added storage (which is part of ISIC I) to represent the full cost of the distributive trades that is not related to transportation. The Bureau of Economic Analysis (BEA) reports industries for the US in greater detail. The total value added from the US transportation sector (including transportation of people) in 2010 was $425B or 2.7% of GDP (https://www.bea.gov/industry/gdpbyind_data.htm). It is noted that my assumption contrasts with work in international trade where the exact opposite is assumed, namely that transactions between firms and consumers are costless and transportation is expensive. Arkolakis (2010), on the other hand, presents an exception and considers the cost of informing consumers through a Butters technology in addition to transportation.

[8] For completeness, the counterfactual of direct sellers who charge per variety has been considered in Bronnenberg (2015).

3.4 Decisions and Timing

Decisions by retailers and manufacturers are determined in a two-stage game. In the first stage, manufacturers simultaneously set wholesale prices $w(v)$ taking downstream retailer decisions about prices and assortment into account. In accordance with the Robinson-Patman act of 1936, they do not price discriminate across retailers. They enter until demand sharing and margin erosion drives their profit to zero. In the second stage, retailers take the actions of manufacturers as given. They simultaneously (1) decide how many varieties N to include into their assortment and (2) set a multiplier π, that is common to varieties, such that retail prices are $p(v) = \pi w(v)$. Retailers enter until their profits are driven down to zero.[9] These assumptions specify intra-sector competition among retailers and manufacturers.

The general equilibrium $\mathscr{E}$ has five elements: the multiplier on wholesale price π, assortment variety N, and store density R in the retail sector plus the wholesale price w, and product variety M in the manufacturing sector. Inter-sector relations between manufacturers and retailers emerge from the fact that in a closed economy N and M are related (see below).

The conditions for general equilibrium are: (1) manufacturers and retailers maximize profits, (2) free-entry takes place until the zero-profit condition is met in each sector, (3) all agents are either employed in the manufacturing sector or the retailing sector.[10]

3.5 Retailer and Manufacturer Problems

Retailers Substituting optimal quantities (3) in equations (1) and (2), a consumer has the following indirect utility when purchasing from a retailer located at a distance d,

$$U(\pi, N, d) = \left(\frac{Y}{\pi W(N)}\right)^{1-\rho} (T - td)^{\rho}, \tag{10}$$

where $W(N)$ is a wholesale price index, equal to $\left(\int_{v \in V_N} w(v)^{1-\sigma} dv\right)^{\frac{1}{1-\sigma}}$, that rises in each wholesale price $w(v)$ and drops in total assortment size N.

Consider that a consumer chooses between a focal store located at a distance d and a competing

[9]This is supported by e.g., Ailawadi (2001) and Messinger and Narasimhan (1995), who report low profit levels for manufacturers and retailers in many industries.

[10]These conditions are the same as in Krugman (1979), but applied to two-sectors.

store located at a distance of $\frac{1}{R}-d$. The trade area of a retailer is solved from the condition that the store's customer who lives farthest away is indifferent to visiting the competitor with a wholesale price multiplier π_c and carries an assortment N_c,

$$U(\pi,N,d)=U\left(\pi_c,N_c,\frac{1}{R}-d\right). \tag{11}$$

A retailer's clientele is the mass of consumers contained in its trade area, $2d(\pi,N)$ (this notation for the trade area is shorthand and hides that it also depends on π_c, N_c, and R). Details are in Appendix A.4.

Each retailer maximizes its own profits with respect to its wholesale price multiplier π and assortment N, taking as given the pricing and assortment decisions of competing retailers and wholesale pricing decisions in the manufacturing sector. The entry decisions of manufacturers enter as a constraint on the retailer's assortment, i.e., if a mass M of manufacturers enters, assortment breadth N can maximally be M.

The retailer's problem can be stated as

$$\max_{\pi,N} 2d(\pi,N)\times s\times Y\times\frac{\pi-1}{\pi}-g(N), \text{ such that } N\leq M, \tag{12}$$

which expresses that a retailer earns a fraction $\frac{\pi-1}{\pi}$ of total expenditure Y from consumers with a population density of s in its trade area, $2d(\pi,N)$.

Since the consumer's utility (10) for a retailer drops in the wholesale price of any variety that is part of its assortment, the retailer becomes more appealing to consumers if it procures the least expensive varieties. Retailers will therefore stock the N least expensive varieties in their stores.

Solving the retailer's problem (12) and using symmetry gives an expression for the price multiplier $\pi(R)$, and assortment choices $N(R)$, conditional on the retailer density R. Next, the retailer density R is determined as the number of retailers per unit of distance that can recoup the fixed cost $g(N)=\gamma N$.

Manufacturers The manufacturer problem is to maximize profits with respect to the wholesale price $w(v)$ of its variety, taking into account that the retailer sets a downstream price of $\pi w(v)$ and selects the N least expensive varieties.

Therefore, if a manufacturer wants to sell through all retailers (and thus to all mass s consumers), its wholesale price must be low enough to be among the N least expensive suppliers. The manufacturer's problem then is,

$$\max_{w(v)} s \times x(v) \times (w(v) - c) - f, \quad \text{such that } w(v) \leq w_N, \tag{13}$$

where $x(v)$ is individual quantity demand at a price of $\pi w(v)$ and w_N is the Nth lowest wholesale price. The constraint in the manufacturer problem implies that downstream limits in distribution capacity constrain upstream wholesale pricing by manufacturers. The constraint does not bind at any wholesale price, $w_N \equiv \infty$, when there are fewer than N varieties produced in the manufacturing sector,

4 THE GENERAL EQUILIBRIUM

4.1 Preliminaries

It helps the exposition to start with two quantities that play a recurring role. Define M^* as

$$M^* = \left(\frac{sY}{\sigma f}\right)\left(\frac{\sigma - 1}{\sigma}\right). \tag{14}$$

Discussed below, M^* is the free-entry mass of manufactured variety, when (1) manufacturers charge monopoly prices, and (2) retailers are unconstrained in choosing variety. Because manufacturing entry is fueled by profits and profits are highest at monopoly prices, M^* can be thought of as a measure of the manufacturing sector's capacity to produce variety, given the actions of retailers.

Next, define N^* as

$$N^* = \left(\frac{sY}{\sigma\gamma}\right)\left(\frac{2T}{t}\right)\left(\frac{1-\rho}{1+2\rho\sigma - 3\rho}\right). \tag{15}$$

Below, it is shown that N^* is the mass of variety that retailers include in their assortment, provided they are not constrained by insufficient production of variety.

M^* and N^* are functions of parameters only. Given their dependence on manufacturing entry cost f and retailer setup costs γ, the quantities M^* and N^* can be interpreted as technology constraints on manufactured and distributed variety, respectively. For instance, using equations (14) and (15), a free-entry manufacturing sector is capable of producing more variety than retailers want ($M^* > N^*$), when the ratio of manufacturer entry costs to retailer set-up cost, $\frac{f}{\gamma}$, is small enough,

$$\frac{f}{\gamma} < \left(\frac{t}{2T}\right)\left(\frac{1+2\rho\sigma-3\rho}{1-\rho}\right)\left(\frac{\sigma-1}{\sigma}\right). \tag{16}$$

4.2 The Equilibrium

I now proceed by first presenting a condition on the ratio $\frac{f}{\gamma}$ such that an equilibrium exists where manufacturers all sell exclusively through retailers.

Claim 1. *If the cost of starting a manufacturing plant f is sufficiently larger than the cost γ of including a single variety in the assortment of a single store, then a manufacturer can not profitably sell its variety directly to consumers in the presence of a nearby retailer selling an assortment of similar varieties.*

Proof. See Appendix A.3. □

To quantitatively assess this condition, the proof develops a simple result, which is not necessary but guarantees that the claim holds, on exactly how large f needs to be in relation to γ. This result depends only on retailer profitability π, and the elasticity of substitution σ. Using ranges of estimates on both quantities reported in the literature, it is shown that the claim is true if the cost of starting a manufacturing plant is more than 5 times that of including its variety in the assortment of a single store. This being a very mild condition on entry costs,[11] I assume it holds in practice and limit attention to the case where all manufacturers sell through retailers exclusively.

Under this existence condition, the equilibrium depends on the relation between M^* and N^*. This results in three cases, which may in turn be thought of as limits on distribution, $M^* > N^*$, on product innovation, $M^* < N^*$, or on both $M^* = N^*$. Consider the three cases in turn.

Case 1: distribution is scarce, $N^* < M^*$. When distribution is scarce, retailers can procure more manufactured variety than they care to stock in their stores. From the manufacturer perspec-

[11] Even if opening a full-size store were to cost as much as \$5MM, then with almost 50,000 varieties stocked, the average cost to distribute a single variety is \$100. It seems implausible that the cost of starting manufacturing operations for each of these varieties is less than a five-fold of this, i.e., \$500.

tive, this case can be interpreted as the presence of limits on shelf space. In this case, the following proposition holds.

Proposition 1. *When $N^* < M^*$,*

1. *retail prices are $p = \pi^* w$, with wholesale prices w defined below, and*

$$\pi^* = \frac{\sigma}{\sigma - 1}. \tag{17}$$

2. *Retailers carry an unconstrained assortment which contains a mass of variety equal to*

$$N = N^*. \tag{18}$$

3. *The density of retailers (Hotelling variety) is equal to*

$$R^* = \left(\frac{t}{2T}\right)\left(\frac{1 + 2\rho\sigma - 3\rho}{1 - \rho}\right). \tag{19}$$

4. *Manufacturer wholesale prices are equal to*

$$w = c\left(\frac{(\sigma - 1)}{(\sigma - 1) - \frac{f}{R^*\gamma}}\right). \tag{20}$$

5. *The mass of manufacturers (Chamberlinian variety) that enters is equal to the retailer assortment mass N^**

$$M = N^*. \tag{21}$$

Proof. See appendix. □

To interpret this proposition, first, retailers charge a markup (equation 17) over wholesale price which depends only on the elasticity of substitution and not, as in Salop (1979) on travel cost. It is of some interest to explain the difference. There are two opposing effects of consumer travel cost on retail pricing. To start, two retailers at a fixed distance are more differentiated when travel costs rise. Next, travel costs also stimulate store entry. The first effect softens and the second intensifies price competition. In Salop (1979), the first effect dominates and thus he finds that travel costs increase margins. In my endogenous entry-cost model, stores have the option to enter more densely at a lower cost by choosing limited assortment. Therefore, the second effect on price of entry is more prominent than in Salop (1979) and cancels out the spatial differentiation effect.

Second, the mass of assortment chosen by retailers (equation 15 and 18) depends in equilibrium negatively on travel cost t, the elasticity of substitution σ, preference for leisure ρ, and the cost to build assortment γ. The higher the elasticity of substitution σ of the varieties sold by a retailer, the lower the consumer's preference for variety. Because of this, the assortment mass that maximizes retailer profits is small. As another example, when the consumer's preference for leisure ρ increases, consumers prefer less time in the market. The equilibrium adjusts with more retailer entry. Consequently, retailers' entry cost needs to be low and therefore their assortment commensurately small. Assortment mass further depends positively on the population density s, and on consumer resources, i.e., time T and income Y.

Third, using expressions (15) and (19), total entry costs in the retail sector (which is the cost of opening a store γN^* times the density of stores R^*) equal $\frac{sY}{\sigma}$. In the free-entry equilibrium, this also equals the added value of the entire retail sector.

Fourth, recalling distribution is limited ($N^* < M^*$), manufacturers' wholesale prices (equation 20) are impacted by limited downstream distribution. Since retailers source the N^* varieties from the manufacturers with the lowest wholesale price, the latter are Bertrand competitors as long as there is a positive difference between entry in the manufacturing sector, M, and the mass of retail assortment N^*. The smaller the store size N^*, the larger the store density R^*, and the lower manufacturer wholesale prices in equation (20).

Finally, the mass of entering manufacturers M (equation 21) is equal to the mass of the retailer assortment N^* given that no manufacturer can sell directly to consumers.

Case 2: variety is scarce, $N^* > M^*$. When starting a manufacturing plant is expensive relative to distribution, retailers can not procure their preferred level of assortment variety. In this case, the following proposition holds.

Proposition 2. *When $N^* > M^*$*

1. *Retailer prices are equal to $p = \pi w^*$ with*

$$\pi = 1 + \frac{4\left(\frac{\rho}{1-\rho}\right)}{\sqrt{1+16\frac{T}{t}\frac{\sigma f}{\gamma}\left(\frac{\rho}{1-\rho}\right)}-1}. \tag{22}$$

2. *The retailer assortment is constrained by the variety produced by the manufacturing sector*

$$N = M. \tag{23}$$

3. *The density of retailers (Hotelling variety) is equal to*

$$R = \frac{t}{4T}\left(1+\sqrt{1+16\frac{T}{t}\frac{\sigma f}{\gamma}\left(\frac{\rho}{1-\rho}\right)}\right). \tag{24}$$

4. *Manufacturer wholesale prices are equal to*

$$w^* = c\left(\frac{\sigma}{\sigma-1}\right). \tag{25}$$

5. *The mass of entering manufacturers (Chamberlinian variety) is equal to*

$$M = \frac{sY}{\sigma f}\frac{1}{\pi}. \tag{26}$$

Proof. See appendix. □

When there is insufficient production of variety from the perspective of the retailer, retail prices (equation 22) rise in consumer travel costs t, as in Salop (1979). Store size adjustment is reduced from the constraint on variety in the retailer problem (12) which binds. With the constraint working against full adjustment in store entry, travel cost now causes more spatial differentiation and prices rise. Retail prices also increase in preference for leisure ρ, and in retailer cost of providing assortment γ. They drop in the consumer time resource T, the elasticity of substitution σ, and in the entry costs of manufacturers f.

The appendix shows that retailers enter in larger numbers than R^*. They do so because their entry costs are lower from upstream scarcity in manufactured variety. Retailer entry also increases in travel cost t, the elasticity of substitution σ, preference for leisure ρ, and manufacturing entry cost f. On the other hand, it drops with the consumer time endowment T and the retailer's marginal cost of distributing variety γ.

Manufacturers, being assured of distribution at any price, charge monopoly margins (equation 25). Entry in the manufacturing sector (equation 26) is determined from the condition that all manufacturers can just recoup their fixed entry costs f.

Case 3: coordinated in variety, $N^* = M^*$ When the optimal mass of retail assortment is equal to the variety producing capacity of the manufacturing sector, the following result holds,

Proposition 3. *When* $N^* = M^*$ *the equilibrium consists of* $\mathscr{E} = \{\pi^*, N,^* R^*, w^*, M^*\}$ *with* π^* *through* M^* *defined above.*

Proof. See appendix. □

This is a knife-edge case that occurs when equation (16) holds in equality. In this case, manufacturers and retailers both have the same percentage margins, $p = \frac{\sigma}{\sigma-1}w$ and $w = \frac{\sigma}{\sigma-1}c$.

The three cases combined. First, the maximum distance a consumer will travel is $\frac{T}{t}$. The marginal consumer lives $\frac{1}{2R}$ from the nearest store. Hence, as long as $R > \frac{t}{2R}$ the retail sector covers the market. It is easily verified that this is the case in each of the three cases above. Thus, the retail sector covers the market.

Next, the retailers' assortment is equal the supply of variety or vice versa, i.e., $N = M$, across all three cases of the equilibrium.

Finally, it can be verified using the results in Appendix A.4 that all equilibrium outcomes, i.e., product variety $M\,(=N)$, store density R, the retail price multiplier π, and the wholesale price w, are continuous in model parameters. In other words, the three cases of the general equilibrium "connect."

4.3 Comparative Statics

Figures 1-2 show how manufacturing and retailing activity adjust to changes in selected model parameters. In both figures, the left panel shows product variety, M, and store density, R, whereas the right panel shows the percentage retail margins, $\frac{p-w}{p}$, and wholesale margins, $\frac{w-c}{w}$. The shaded area in the figures captures the region in which retailers want larger assortments than they can source from manufacturers, $N^* > M^*$. The non-shaded area depicts the opposite case, $N^* < M^*$, where a manufacturing sector can make more variety than retailers want. The border separating the two areas defines the case where $N^* = M^*$. This section focuses on effect of the cost parameters, i.e., travel cost t and retailer cost γ. Appendix B provides comparative statics for other parameters.

— Insert Figure 1 and 2 here. —

Figure 1 shows that as travel costs t go up, store density R increases and product variety M decreases. Therefore the retail sector's response to less costly travel by retailers is to provide low-density "big box" stores with large assortments. This constitutes an intuitive margin of adjustment, absent from Salop (1979). The example in the introduction about the expansion of assortment variety in stores is consistent with such a trend (see Bhatnagar and Ratchford 2004 for additional support). The initial shallow decrease in manufacturing entry (shaded area) is the result of rising retail prices p in response to increasing consumer travel costs, leaving a decreasing fraction $\frac{w}{p}$ of total consumption expenditure sY for the manufacturing sector. Still, all manufacturers are distributed in this region and can ask monopolist margins. The kink in manufacturing entry at $M^* = N^*$ is caused by this no longer being true. At some level of travel cost t, further increases in store density and associated reductions in retail assortment size cause manufacturer competition on wholesale prices for "shelf space".

Moving to the right panel, the manufacturer percentage profit margin $\frac{w-c}{w}$ is high, as long as manufactured variety is less than the retailer's profit maximizing mass of assortment (shaded region). All manufacturers are assured distribution at monopoly wholesale prices. At the same time, as travel costs go up, retail margins $\frac{p-w}{p}$ rise from increased spatial differentiation. At increasing travel costs, more stores open with ever smaller assortments each, ultimately forcing manufacturers to cut wholesale prices in order to be among the least expensive ones (non-shaded region). Now, retail prices no longer rise in travel costs, because more stores open with smaller assortments.

The total size of the retail sector increases in consumer travel as long as manufactured variety is scarce (shaded region). This can be seen from the increase in the retail percentage margin and the realization that in a free entry equilibrium, the total entry cost of the retail sector will be equal to variable profits in that sector, $sY\frac{p-w}{p}$.

Figure 2 shows the equilibrium as a function of the retailers' cost to stock variety, γ. Product variety and store density decrease in the cost of distribution. Indeed, when it becomes more expensive to create retail assortment, fewer and smaller stores enter, leaving also less room for entry in manufacturing.

In the right-hand graph, retail margins $\frac{p-w}{p}$ are small at low costs of distribution. Not being able to find all the variety they want when large assortments are cheap to set up, retailers compete harder on price to attract consumers (shaded region). Manufacturers now charge high wholesale margins

$\frac{w-c}{w}$, because they are guaranteed distribution. However, as the cost of distributing variety γ rises, the reverse happens (non-shaded region). Manufacturer margins are suppressed from competition to be among the N^* cheapest suppliers and not all of them will have access to retail shelves.

5 An Investigation of the Retail Sector

Turning to the research questions posed in the introduction, I use the equilibrium to first investigate how retailing impacts variety in manufacturing. Next, I analyze how double marginalization impacts consumer welfare. Then, pricing power of manufacturers and retailers is shown to arise in equilibrium as a function of entry cost primitives. Finally, to study the market's allocation of resources to the retail sector, the equilibrium is used to make predictions about the size of the national retail sector, which are then compared to observations.

5.1 The Retail Sector and Manufacturing Entry

How does distribution impact manufacturing entry and the production of variety? The general equilibrium model implies that the retail sector raises manufacturing entry and the production of variety in two complementary ways. First, high distribution costs limit entry by manufacturers through the optimal decisions by retailers to stock a small assortment of varieties. The sharp decline in manufacturing entry M in the non-shaded area of Figure 2 illustrates this impact of retail costs on manufacturer entry. Therefore, lower distribution costs γ will result in (a) retail assortments with more variety and (b) more entry in manufacturing.

There is ample empirical evidence that the retail productivity has risen and the cost of distributing variety has fallen in the past decade (Bronnenberg and Ellickson, 2015; Hortaçsu and Syverson, 2015). Consistent with the prediction above, data bear out that (a) retail assortments are larger than ever before (Consumer Reports, 2014; Pantzar, 1992), and (b) manufacturers produce more variety to stock stores' shelves. For instance, the USDA estimates that, between 1992 and 2010, the number of newly launched food and beverage products has increased from 10K to 22K per year and the number of newly launched non-food products from 6K to 26K per year.[12] The Directorate General Competition of the EU identified a similar mechanism for manufacturing entry stating

[12] See https://www.ers.usda.gov/topics/food-markets-prices/processing-marketing/new-products/.

that "the availability of sufficient retail distribution stimulates innovative efforts, whereas limited distribution has the opposite effect because the investment in innovation can not be recouped" (European Commission, 2014). However, further empirical support is needed for this effect.

Second, and perhaps more fundamentally, retailing affects manufacturer entry by changing the nature of manufacturer competition over the extensive consumer margin. To see how entry of a retail sector impacts competition among manufacturers, consider what happens to manufacturing entry when there is no retail sector and manufacturers are selling directly to consumers. Equation (6) now characterizes a consumer's optimal trip length δ, and demand for variety is the variety offered by manufacturers contained in a constant span δ of the market around the consumer's residence.

Figure 3 illustrates that this demand for variety makes manufacturers compete fiercely over the extensive consumer margin.[13] Namely, consumers living between firms j and $j-1$ are indifferent between taking trip 1 or trip 2, which both yield the same variety at the same travel costs. The varieties produced by $j-2$ and $j+1$ are therefore perfect substitutes to the mass of consumers between $j-1$ and j, no matter the product differentiation of the goods involved. The manufacturers of varieties $j-2$ and $j+1$ can both win this mass of customers by slightly undercutting their rival on price. Therefore, somewhat surprisingly, the Bertrand argument applies for these two manufacturers despite their differentiation. This is because demand for variety is bounded by travel cost and consumers do not care deeply about the identity of the varieties that make their endogenous purchase set. This incites manufacturer competition over purchase set membership.[14]

— Insert Figure 3 here. —

By the same argument, it pays for any manufacturer to undercut its rivals slightly to capture an extensive consumer margin as long as margins are positive. Adjustment of prices by manufacturers does not impact the amount of travel that consumers optimally do (see Equation 6). Therefore, firms undercut each other until $p = c$. In the free-entry equilibrium, the mass of entering manufacturers M adjusts to a zero-profit condition. Given a set of prices p the mass of free-entry manufacturers is

[13] Figure 3 portrays manufacturers as being discrete for simplicity. The argument below holds equally when the space between firms is filled up by a continuum of others.

[14] The pressure to compete over the extensive consumer margin is robust to several elaborations. It is not required that all consumers like all varieties equally much, as is assumed above. The argument instead necessitates a mass of indifferent consumers, i.e., that *some* consumers like varieties across multiple feasible trips equally much and that *some* consumers love variety.

equal to total market income sY times the profit margin $\frac{p-c}{p}$, divided by the fixed cost of entry f, or

$$M = \frac{sY}{f}\left(\frac{p-c}{p}\right). \tag{27}$$

But with no sustainable price point above marginal costs, only a minimal mass of manufacturers enter. Therefore, the counterfactual of manufacturers selling directly to consumers produces little variety and has no pure strategy equilibrium. The next proposition formalizes this argument.

Corollary 1. *When consumer travel cost are high, in particular,*

$$t > 2T\left(\frac{1-\rho}{1+\rho\sigma-2\rho}\right), \tag{28}$$

there is no pure strategy equilibrium of manufacturers selling directly to consumers.

Proof. See appendix. □

In words, as long as travel causes consumers to buy subsets of varieties (literally: as long as the consumer doesn't circumnavigate the market), integrated manufacturer-sellers compete fiercely over the extensive consumer margin and in doing so eliminate variety from the market. Retailers prevent this by collectively offering full market coverage to all manufacturers, effectively replacing extensive margin competition by manufacturers with trade area competition by retailers. Manufacturers then limit their rivalry to competing over the intensive consumer margin and access to the retailers' shelves. They charge higher prices and more of them enter.

In conclusion, the presence of a retail sector leads to more manufacturing entry because it lowers manufacturer competition over purchase set membership.

Conversely, when manufacturers sell directly to consumers, full information about prices can lead to too much manufacturer competition and precipitate low levels of manufacturing entry. The same argument also sheds new light on co-location of sellers at farmer's markets, shopping malls and trade shows. Traditionally, the argument has been that co-location is a way to attract more consumers by accepting more inter-firm competition. But, this argument does not apply to price competition when consumers like variety. In contrast, when consumers like variety, co-location lowers inter-firm price competition by lowering the cost of purchasing and thereby raising the demand for variety.

5.2 *Double Marginalization and the Social Optimum*

Final prices are higher in the two-sector equilibrium than in the one-sector equilibrium when manufacturers sell directly to consumers. How does double marginalization impact welfare in the general equilibrium model? On the one hand, double marginalization reduces purchase quantities, which harms consumers. On the other hand, higher prices lead to bigger profit margins. In a free-entry setting, this increases entry by manufacturers and retailers, which benefits consumers who like product variety and leisure. Therefore, the answer to the question above is not clear a priori.

To study the issue further, I evaluate how welfare in the general equilibrium compares to the welfare optimum in two scenarios: (1) planned entry in a one-sector economy with manufacturers selling directly to consumers and (2) planned entry in a two-sector economy. The first scenario avoids double marginalization by construction and the welfare it achieves forms a natural benchmark for the two-sector market (equilibrium) outcome.

Appendix A.7 formalizes the two social planner problems. Here, I discuss the main points. The social planner decides on the mass of product variety M, and in the second scenario additionally on store density R. The social planner further sets final prices equal to the marginal cost of production c and funds entry in either sector directly from income, making agents also owners. Disposable income for a given agent (consumer/owner) is therefore income Y depreciated by a $\frac{1}{s}$ share of the total cost of entry. In the one-sector model, the latter is the total setup cost in the manufacturing sector, $M \times f$. In the two-sector model, the total cost of entry in the retail sector, $R \times M\gamma$, is added.

— Insert Figure 4 here. —

Figure 4 shows the social optimum and the equilibrium outcome as a function of retailer distribution costs γ. Panel (A) shows retailer entry. For reference, the store density (retailer entry) R_1 in the one-sector optimum is 0. The general equilibrium contains approximately twice as many stores, R_E, as the social optimum in the two-sector economy, R_2. Retailers compete for the marginal consumer at the edge of their trade area and not for the average consumer (see also Salop, 1979) and this makes too many of them enter.

Panel (B) depicts manufacturing entry. The general equilibrium contains a lower mass of manufacturers M_E than the social planner chooses in a one-sector model M_1. The reason is simple. Consumers prefer leisure over the inconvenience of purchasing varieties that are too distant. Therefore, they only buy a local subset of varieties, but this subset is different for consumers living in different regions of the circle. In order to provide sufficient variety to dispersed consumers, the social planner thus has to set the global mass of entering manufacturers above local demand for variety in the one-sector model, leading to wasteful replication. On the other hand, the general equilibrium generally admits more entry in manufacturing M_E, than the the social planner wants in the two-sector optimum, M_2. But, equilibrium entry is less excessive than in the retail sector, especially as the cost of distribution γ becomes large and retailers choose an assortment with small mass.

Panel (C) shows the levels of welfare associated with the general equilibrium, W_E, along with the social planner outcome in the one-sector model W_1, and the two-sector model W_2. To start, the equilibrium welfare W_E is higher than the welfare optimum W_1 in the one-sector model. This is because consumers have access to more consumed variety and enjoy more leisure when buying from stores that compete on assortment variety than when buying a subset of varieties directly from manufacturers.

Next, the equilibrium welfare level W_E associated with double marginalization is lower than the two-sector welfare optimum W_2, but closely approximates it despite excessive entry in both sectors. The explanation is that the market compensates consumers' quantity-loss from high prices almost completely with more variety and more leisure.

Finally, referring back to Figure 4, technologies such as online retailing that lower the cost of carrying assortment γ, raise welfare levels and narrow the welfare gap with the two-sector social optimum. In the limit, the two are the same, i.e., $\lim_{\gamma \to 0} W_E \to W_2$.

In conclusion, a free-entry market with manufacturers and retailers provides more welfare despite double marginalization than the social planner's implementation of a market with manufacturers selling directly to consumers. However, relative to the social planner's implementation of a two-sector market, the general equilibrium delivers too much entry and lower, yet comparable, welfare.

5.3 Pricing Power and Coordination on Variety

The division of channel surplus and pricing power among manufacturers and retailers have been studied in economics (e.g., Brandow, 1969; Villas-Boas, 2007) and business (e.g., Draganska, Klapper, and Villas-Boas, 2010). It is also a topic of interest to legislators (Federal Trade Commission, 2001), especially in the context of large and powerful retailers demanding lower wholesale prices from manufacturers (e.g., van Heerde, Gijsbrechts, and Pauwels, 2008; Inderst and Shaffer, 2009; Steenkamp and Dekimpe, 1997). In addition to firm size, the ability of retailers to affect wholesale prices has been linked to financial resources, concentration, and proximity to consumers (Brandow, 1969).

The theory proposed here offers a complementary perspective. Consider the following result.

Corollary 2. *Retail margins are larger than manufacturing margins when retailers want fewer varieties than manufacturers can make ($N^* < M^*$). Margins are equal when the retailers' optimal assortment is equal to the manufacturing capacity to produce variety ($N^* = M^*$). Finally, retail margins are smaller than manufacturing margins when the manufacturing capacity to produce variety is limited ($N^* > M^*$). Formally,*

$$\begin{array}{lll} \frac{p-w}{p} > \frac{w-c}{w} & \text{if} & N^* < M^* \\ \frac{p-w}{p} = \frac{w-c}{w} & \text{if} & N^* = M^*. \\ \frac{p-w}{p} < \frac{w-c}{w} & \text{if} & N^* > M^* \end{array}$$

Proof. See appendix. □

Sector specific costs of entry cause the mass of manufactured variety to differ from the mass of assortment variety in distribution when firms in each sector exercise their monopoly power. In general equilibrium, wholesale or retail margins are adjusted to eliminate this difference, which and this is what drives pricing power. The above result implies that the sector with the larger margins is the one that constrains consumed variety. If, at full pricing power, more variety is produced by manufacturers than distributed by retailers, then intra-sector manufacturer competition for access to retail shelves reduces wholesale prices. On the other hand, if retailers want to distribute more variety than is made by manufacturers, then retailers need compete harder on price for customers. The ability to exercise monopoly power therefore arises in my model from the scarcity of variety in manufacturing or distribution, which in turn depends on entry primitives. This argument holds equally with large, powerful, retailers and with a fragmented retail sector, as in my model.

Finally, the combined variable profits of manufacturers and retailers are highest in equilibrium when variety in upstream production is aligned with variety in downstream distribution, i.e., when $N^* = M^*$. This can be easily verified from propositions 1-3 and can also be directly seen from the size of the combined margin in Figures 1 and 2. Although it is entirely possible that $N^* \neq M^*$ in equilibrium, there are likely economic forces that prevent the gap between N^* and M^* from widening without bounds. For instance, if the marginal cost of building assortment γ is high, then retailers distribute a small mass of varieties ($N^* < M^*$). Retailers would be willing to pay for a technology that lowers the cost γ because it allows them stock more varieties and attract additional consumers. This creates a market for retail innovation of the type discussed by Bronnenberg and Ellickson (2015) and Hortaçsu and Syverson (2015) which would narrow the gap between N^* and M^*. Thus, even if distribution is a constraint in the short run, in the long run the market can adjust through invention of retail formats with lower γ. A similar argument can be constructed about the market provision of a manufacturing technology that lowers the cost of producing variety f, in the case where $N^* > M^*$.

5.4 The Size of the Retail Sector

What does the equilibrium predict about how the market allocates resources to the costly provision of store density and manufactured variety? To answer this question, I use the general equilibrium to make predictions about the size of the retail sector. Next, I present some support for these predictions using national account data. To keep the discussion concise, I focus on the case that a retailer can in principle procure sufficient varieties from the manufacturing sector, i.e., $M^* \geq N^*$. This choice is supported by both recent literature (Richards and Patterson, 2004) and antitrust inquiries (Federal Trade Commission, 2001).

Using equations (15) and (19) , the model predicts that the size of the retail sector as a fraction of the total economy is

$$\frac{\text{size retailing sector}}{\text{size of economy}} = \frac{R^* \times N^* \gamma}{sY} = \frac{1}{\sigma}. \tag{29}$$

This result implies two predictions. First, the relative size of the retail sector is determined by the elasticity of substitution. Second, the relative size of the retail sector is independent of other model parameters such as country wealth, consumer travel costs, or firm entry costs.

To show some support for these predictions, I first relate the prediction in equation (29), at consensus estimates of the elasticity of substitution σ, to the observed share of the distributive trades. Next, I show that, as predicted by equation (29), the observed relative size of the distribution sector is not related to proxies for market size sY, travel costs t, and setup costs f or γ,

As in the introduction, for the size of the distributional trades I use the total expenditure on retailing, wholesaling, and transportation (ISIC G + H + I) in the UN National Accounts database. I use the same source to collect (1) GDP as a proxy for market size sY. In addition, I use data from Euromonitor on (2) the national share of online retailing,[15] and view countries with a large share of online retailing as countries with a low average retailer cost of distributing variety γ, (3) the ease-of-doing-business ranking, as a shifter of manufacturer setup costs f in a typical country; and finally (4) the number of cars per capita in a given country, as a proxy for travel costs t.

With respect consensus estimates of the elasticity of substitution, Broda and Weinstein (2006) report that σ typically is 3-4 at low levels of aggregation, whereas Broda and Weinstein (2010) report an average elasticity of substitution in consumer packaged goods of 7.5 across many brand-product groups. Feenstra (1994) reports estimates of σ of 5.83 (knit shirts), 6.29 (athletic shoes), 2.96 (type writers). I take these reported estimates to mean that in a typical industry the elasticity of substitution σ is 5 by crude approximation.

Figure 5 plots the relative size of the distributive trades against the 4 proxies for model parameters discussed above. First, the model predicts that the distribution sector is about 20% of GDP at the average of the estimates for σ reported in the literature. This is consistent with the data presented in the introduction and in Figure 5.

— Insert Figure 5 here. —

Second, the individual panels in Figure 5 also show support for the second prediction, i.e., that the relative size of the retail sector is not systematically related to the 4 proxies for selected model parameters. For instance, the first panel shows that the relative size of the distributional trade does not depend on GDP. Indeed, the retail sector is 22% of total economic activity in a country like Great Britain (GRB) with a high GDP, as it is in a country like Guinea-Bissau (GNB) with a low GDP. Similar examples can be constructed for the proxies of γ, f, and t.

While not intended as an empirical test of the theory, these graphs connect the theory, evaluated at publicly available proxies for the model parameters, to observations about the size of the distributional trades.

[15] These and other variables are available from the Euromonitor Passport data base at www.euromonitor.com.

6 Conclusion

This paper presents a tractable general equilibrium model of a complete economy with a manufacturing and a distribution sector. The model consists of consumers, retailers, and manufacturers, each pursuing their own objectives. The proposed model can be used to address questions about how the market allocates resources to the costly provision of product variety and store density.

Substantively, the paper shows that a manufacturer can generally not profitably sell its variety directly to consumers with a love of variety who can buy an assortment of similar varieties at a nearby retailer. Retailers perform two interrelated functions in this case, both of which are central to providing product variety. To consumers, retailers lower the cost of shopping for variety. Without this, the demand for variety would be low. To manufacturers, retailers offers market coverage. Without this, the supply of variety would be low. Therefore, retailers are a catalyst to more production of variety by manufacturers and more consumption of variety by consumers.

Vertical interaction between retailers and manufacturers, and especially their relative pricing power, is impacted by scarcity of variety in manufacturing or distribution. Pricing power is therefore a resulting property of the equilibrium and is determined by entry costs.

For realistic values of the elasticity of substitution, the retail sector is predicted to be approximately 20% of the total economy across nations large and small, rich and poor. Support using the UN National Accounts data base is given.

Several extensions and topics for future research remain. First, the relation between productivity in retail and manufacturing entry deserves further scrutiny and empirical study. Next, the analysis has not addressed the impact of preference heterogeneity and of differences in firm productivity. Finally, the analysis does not accommodate the existence of other transaction costs such as wait time or queuing time in addition to the travel costs considered here. However, such costs will likely strengthen the views espoused in this paper about the benefits of retailing, and not undermine them.

REFERENCES

AILAWADI, K. L. (2001): "The Retail Power-Performance Conundrum: What Have We Learned," *Journal of Retailing*, 77, 299–318.

ANDERSON, J. E., AND E. VAN WINCOOP (2004): "Trade Costs," *Journal of Economic Literature*, 42(3), 691–751.

ARKOLAKIS, C. (2010): "Market Penetration Costs and the New Consumer Margin in International Trade," *Journal of Political Economy*, 118(6), 1151–1199.

BECKER, G. S. (1965): "A Theory of the Allocation of Time," *The Economic Journal*, 75(299), 493–517.

BHATNAGAR, A., AND B. T. RATCHFORD (2004): "A model of retail format competition for non-durable goods," *International Journal of Research in Marketing*, 21, 39–59.

BRAITHWAITE, D., AND D. DOBBS (1932): *The Distribution of Consumable Goods*. Routledge.

BRANDOW, G. E. (1969): "Market Power and Its Sources in the Food Industry," *American Journal of Agricultural Economics*, 51(1), 1–12.

BRODA, C., AND D. E. WEINSTEIN (2006): "Globalization and the Gains from Variety," *Quarterly Journal of Economics*, 121(2), 541–585.

BRODA, C., AND D. E. WEINSTEIN (2010): "Product Creation and Destruction: Evidence and Price Implications," *American Economic Review*, 100, 691–723.

BRONNENBERG, B. J. (2015): "The Provision of Convenience and Variety by the Market," *RAND Journal of Economics*, 46(3), 480–498.

BRONNENBERG, B. J., AND P. B. ELLICKSON (2015): "Adolescence and the Path to Maturity in Global Retail," *Journal of Economic Perspectives*, 29(4), 1130–134.

CHEN, Z. (2003): "Dominant Retailers and the Countervailing-Power Hypothesis," *RAND Journal of Economics*, 34(4), 612–625.

CONSUMER REPORTS (2014): "What to Do When There Are Too Many Product Choices on the Store Shelves?," *Consumer Reports Magazine*, (March).

CRAWFORD, G. S., AND A. YURUKOGLU (2012): "The Welfare Effects of Bundling in Multichannel Television Markets," *American Economic Review*, 102(2), 643–685.

DAMODARAN, A. (2016): "Margins by Sector (US)," Available from http://www.stern.nyu.edu/ adamodar/pc/datasets/margin.xls, last accessed 7/13/2016.

DIXIT, A. K., AND J. E. STIGLITZ (1977): "Monopolistic Competition and Optimum Product Diversity," *American Economic Review*, 67(3), 297–308.

DOBSON, P. W., AND M. WATERSON (1997): "Countervailing Power and Consumer Prices," *The Economic Journal*, 107(441), 418–430.

DRAGANSKA, M., D. KLAPPER, AND S. B. VILLAS-BOAS (2010): "A Larger Slice or a Larger Pie? An Empirical Investigation of Bargaining Power in the Distribution Channel," *Marketing Science*, 29(1), 57–74.

EUROPEAN COMMISSION (2014): "The Economic Impact of Modern Retail on Choice and Innovation in the EU Food Sector," Discussion paper, Directorate General Competition of the

European Commission.

FEDERAL TRADE COMMISSION (2001): "Report on the Federal Trade Commission Workshop on Slotting Allowances and Other Marketing Practices in the Grocery Industry," Discussion paper, FTC.

FEENSTRA (1994): "New Product Varieties and the Measurement of International Prices," *American Economic Review*, 84(1), 157–177.

GRONAU, R. (1977): "Leisure, Home Production, and Work–the Theory of the Allocation of Time Revisited," *Journal of Political Economy*, 85(6), 1099–1123.

HART, O. D. (1985): "Monopolistic Competition in the Spirit of Chamberlin: A General Model," *The Review of Economic Studies*, 52(4), 529–546.

HARTMANN, W. R. (2006): "Intertemporal Effects of Consumption and Their Implications for Demand Elasticity Estimates," *Quantitative Marketing and Economics*, 4, 325–349.

HORTAÇSU, A., AND C. SYVERSON (2015): "The Ongoing Evolution of US Retail: A Format Tug-of-War," *Journal of Economic Perspectives*, 29(4), 89–112.

INDERST, R., AND G. SHAFFER (2007): "Retail Mergers, Buying Power, and Product Variety," *The Economic Journal*, 117(516), 45–67.

——— (2009): "Market Power, Price Discrimination, and Allocative Efficiency in Intermediate-Goods Markets," *The RAND Journal of Economics*, 40(4), 658–672.

IYER, G. (1998): "Coordinating Channels under Price and Nonprice Competition," *Marketing Science*, 17(4), 338–355.

KAHN, B. E., M. U. KALWANI, AND D. G. MORISSON (1986): "Measuring Variety-Seeking and Reinforcement BBehavior Using Panel Data," *Journal of Marketing Research*, 23(2), 89–100.

KRUGMAN, P. R. (1979): "Increasing returns, monopolistic competition, and international trade," *Journal of International Economics*, 9(4), 469–479.

LANCASTER, K. (1971): *Consumer Demand: A New Approach*. New York, Columbia University Press.

LANCASTER, K. (1990): "The Economics of Product Variety: A Survey," *Marketing Science*, 9(3), 189–206.

MARX, L., AND G. SHAFFER (2010): "Slotting Allowances and Scarce Shelf Space," *Journal of Economics and Management Strategy*, 19(3), 575–603.

MCALISTER, L. (1982): "A Dynamic Attribtue Satiation Model of Variety-Seeking Behavior," *Journal of Consumer Research*, 9(2), 141–150.

MCGUIRE, T. W., AND R. STAELIN (1983): "An Industry Equilibrium Analysis of Downstream Vertical Integration," *Marketing Science*, 2(2), 161–191.

MESSINGER, P. R., AND C. NARASIMHAN (1995): "Has Power Shifted in the Grocery Channel?," *Marketing Science*, 14(2), 189–223.

MOORTHY, K. S. (1988): "Strategic Decentralization in Channels," *Marketing Science*, 7(4), 335–355.

PANTZAR, M. (1992): "The Growth of Product Variety – A Myth?," *Journal of Consumer Studies* *and Home Economics*, 16, 345–362.

PRYOR, F. L. (1977): “Some Costs and Benefits of Markets: An Empirical Study,” *The Quarterly Journal of Economics*, 91(1), 81–102.

REY, P., AND J. TIROLE (1986): “The Logic of Vertical Restraints,” *American Economic Review*, 76(5), 921–939.

RICHARDS, T., AND P. PATTERSON (2004): “Slotting Allowances as Real Options: An Alternative Explanation,” *The Journal of Business*, 77(4), 675–696.

SALOP, S. C. (1979): “Monopolistic Competition with Outside Goods,” *The Bell Journal of Economics*, 10(1), 141–156.

SHAW, A. W. (1912): “Some Problems in Market Distribution,” *Quarterly Journal of Economics*, 26(4), 703–765.

SHAW, E. H. (1990): “A Review of Empirical Studies of Aggregate Marketing Costs and Productivity in the United States,” *Journal of the Academy of Marketing Science*, 18(4), 285–292.

SPENCE, M. E. (1976): “Product Selection, Fixed Costs, and Monopolistic Competition,” *The Review of Economic Studies*, 43(2), 217–235.

SRINIVASAN, S., AND D. M. HANSSENS (2009): “Marketing and Firm Value: Metrics, Methods, Findings, and Future Directions,” *Journal of Marketing Research*, 46(3), 293–312.

STAELIN, R., AND E. LEE (2014): *The History of Marketing Science,*vol. 3 of *Now Publishers Series in Business*, chap. Distribution Channels, pp. 261–288. World Scientific, Hannover, MA.

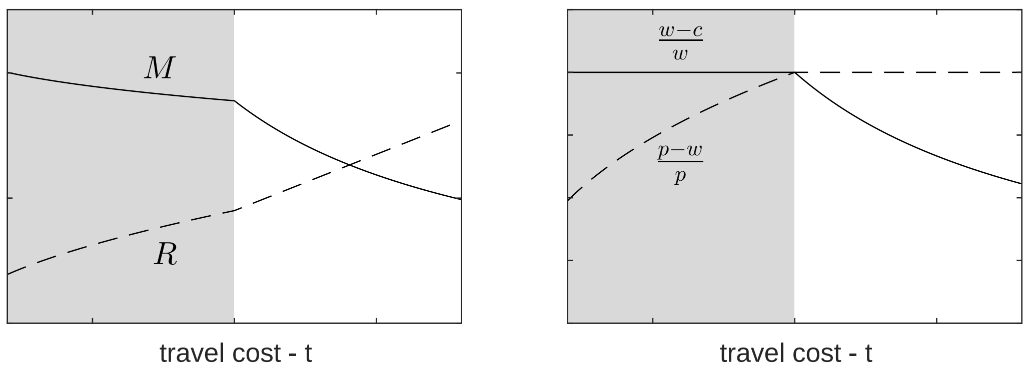
STEENKAMP, J.-B. E., AND M. G. DEKIMPE (1997): “The Increasing Power of Store BBrand: Building Loyalty and Market Share,” *Long*, 30(6), 917–930.

VAN HEERDE, H. J., E. GIJSBRECHTS, AND K. PAUWELS (2008): “Winners and Losers in a Major Price War,” *Journal of Marketing Research*, 45(5), 499–518.

VILLAS-BOAS, S. B. (2007): “Vertical Relationships between Manufacturers and Retailers: Inference with Limited Data,” *The Review of Economic Studies*, 74(2), 625–652.

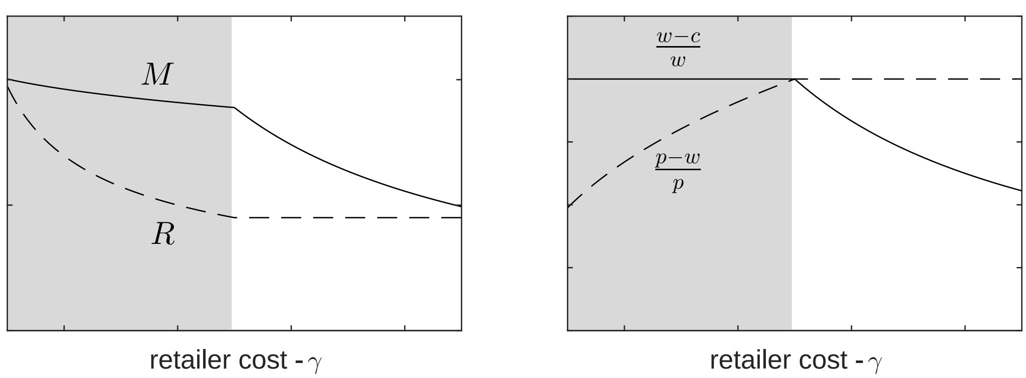
WOLINSKY, A. (1986): “True Monopolistic Competition as a Result of Imperfect Information,” *The Quarterly Journal of Economics*, 101(3), 493–511.

Figure 1: The equilibrium and the consumer cost of travel



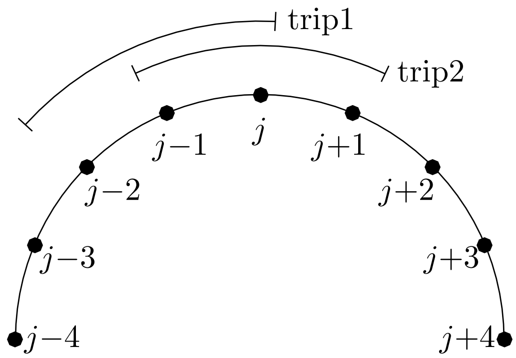
Notes: The shaded area represents the case where variety is scarce, $N^* > M^*$. The non-shaded area represents the opposite case where shelf space is scarce, $N^* < M^*$.

Figure 2: The equilibrium and the marginal cost of distributing variety



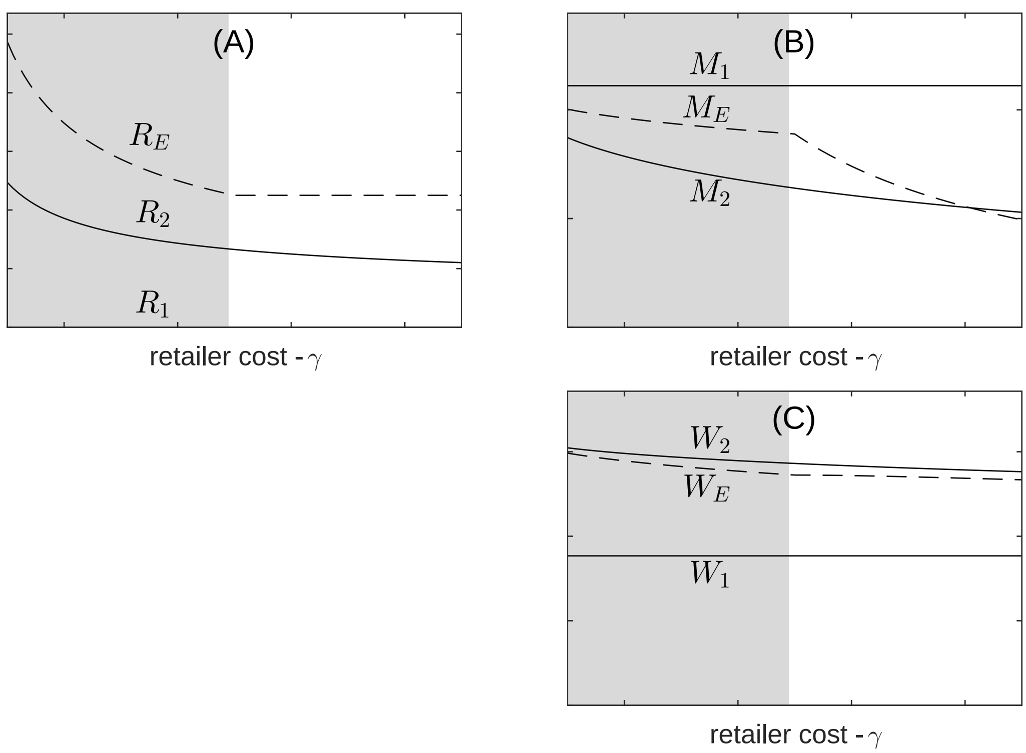
Notes: The shaded area represents the case where variety is scarce, $N^* > M^*$. The non-shaded area represents the opposite case where shelf space is scarce, $N^* < M^*$.

Figure 3: Vigorous competition among manufacturers selling directly to consumers



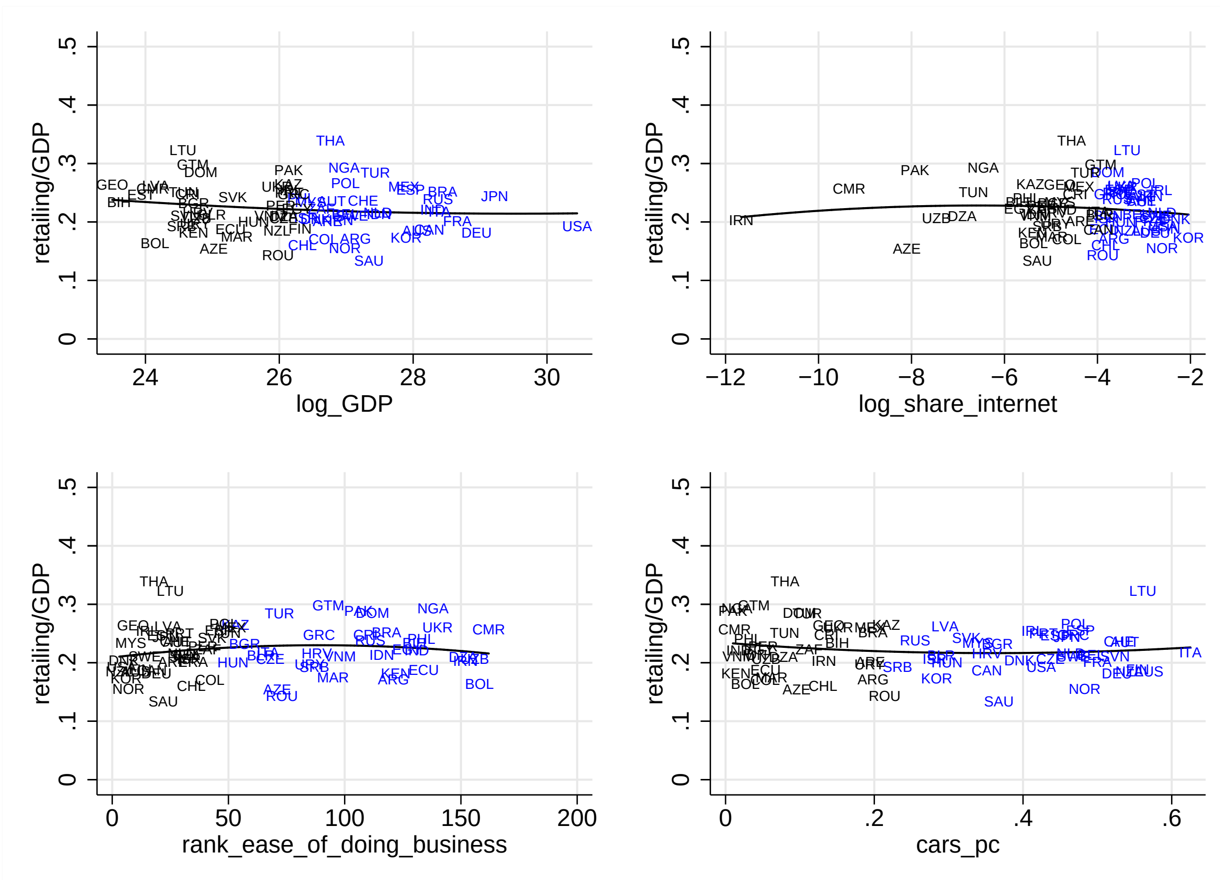
Notes: Varieties $\ldots j-1, j, j+1, \ldots$ are located uniformly on the circle. Consumers are located uniformly on the circle also. Trip 1 and trip 2 are both solutions to the consumer problem and yield the same amount of variety at the same transaction costs for the *mass* of consumers who's residence is located between manufacturer $j-1$ and j. Firms $j-2$ and $j+1$ would therefore fight aggressively for their clientele.

Figure 4: Comparing the equilibrium versus the social optimum



Notes: (i) The figure displays store density (A), product variety (B), and welfare (C). (ii) The subscript E denotes the equilibrium, the subscript 2 denotes the social optimum in the two sector model, and the subscript 1 denotes the social optimum in the "no retailing" one sector model. (iii) The shaded area represents the case where variety is scarce, $N^* > M^*$. The non-shaded area represents the opposite case where shelf space is scarce, $N^* < M^*$.

Figure 5: The distribution sector as a share of the total economy



Notes: The vertical axis gives the total expenditure on the distributive trades (ISIC G + H + G) as a fraction of GDP. Plot symbols are three-letter UN country codes. Data were obtained directly from the National Accounts Data Base at the United Nations.

A PROOFS

A.1 Demand for Quantity

The consumer problem is

$$\max_{X,V,L} U\left(X\left(V\right),L\left(V\right)\right), \text{ s.t. } T=\tau\left(V\right)+L\left(V\right) \text{ and } Y=\int_{v\in V} p\left(v\right)x\left(v\right)dv,$$

The Lagrangian $\mathscr{L}_C$ for the consumer problem is

$$\mathscr{L}_C=\left(\int_{v\in V} x\left(v\right)^{\frac{\sigma-1}{\sigma}}dv\right)^{\frac{\sigma}{\sigma-1}\left(1-\rho\right)}L^{\rho}+\lambda_1\left(T-L\left(V\right)-\tau\left(V\right)\right)+\lambda_2\left(Y-\int_{v\in V} p\left(v\right)x\left(v\right)dv\right).$$

The KKT stability conditions on quantities $x\left(v\right)$ are

$$\left(1-\rho\right)L^{\rho}\left(\int_{v\in V} x\left(v\right)^{\frac{\sigma-1}{\sigma}}dv\right)^{\frac{\sigma}{\sigma-1}\left(1-\rho\right)-1}x\left(v\right)^{\frac{-1}{\sigma}}=\lambda_2 p\left(v\right).$$

Taking ratios for two varieties in the set V

$$\frac{x\left(v_1\right)}{x\left(v_2\right)}=\left(\frac{p\left(v_1\right)}{p\left(v_2\right)}\right)^{-\sigma}.$$

Multiply by $x\left(v_2\right)p\left(v_1\right)$,

$$x\left(v_1\right)p\left(v_1\right)=x\left(v_2\right)p\left(v_2\right)^{\sigma}p\left(v_1\right)^{-\sigma+1},$$

integrate over v_1, and use the complementary slackness condition and $\lambda_2>0$, i.e., that

$$\int_{v\in V} p\left(v\right)x\left(v\right)dv=Y.$$

This will give for each $v_2\in V$ the following expression

$$x\left(v_2\right)=\frac{Y}{\int_{v\in V} p\left(v\right)^{1-\sigma}dv}p\left(v_2\right)^{-\sigma}$$

The effect of $p\left(v_2\right)$ on the price index $\int_{v\in V} p\left(v\right)^{1-\sigma}dv$ is negligible. Therefore, we can write the equation above as

$$x\left(v\right)=A\left(V\right)p\left(v\right)^{-\sigma}, \tag{A.1}$$

where $A\left(V\right)=YP^{\sigma-1}$ is a demand shifter, and $P\left(V\right)$ is a price index $P\left(V\right)=\left(\int_{v\in V} p\left(v\right)^{1-\sigma}dv\right)^{\frac{1}{1-\sigma}}$. This is the usual Marshallian demand as in Dixit-Stiglitz, except it is defined on the subset V.

A.2 *Demand for Variety*

Substitute Marshallian demand (A.1) in the sub-utility for quantities to obtain

$$U(p,V) = Y\left(\int_{v\in V}\left(p(v)^{1-\sigma}\right)dv\right)^{\frac{1}{\sigma-1}}.$$

Combining both parts, indirect utility is

$$U(p,\tau,V) = \left(Y\left(\int_{v\in V} p(v)^{1-\sigma}\,dv\right)^{\frac{1}{\sigma-1}}\right)^{1-\rho}(T-\tau(V))^{\rho} \tag{A.2}$$

Suppose there are M equally spaced and symmetric manufacturers and that the consumer buys a mass δM varieties. Leisure is maximized if these varieties are contiguous on the circle and include the consumer's residence. Total travel distance is then δ each way and the associated travel cost is δt. The indirect utility function (A.2) can be written as,

$$U(p,t,\delta) = \left(\frac{Y}{p}(\delta M)^{\frac{1}{\sigma-1}}\right)^{1-\rho}(T-t\delta)^{\rho}$$

which can be maximized with respect to δ. This yields

$$\delta = \frac{T}{t}\frac{(1-\rho)}{(1+\sigma\rho-2\rho)},$$

which is equal to δ in equation (6) in the text.

A.3 *Proof of Claim* (1)

Consumers do not detour to buy a marginal variety sold directly to consumers. Therefore, the manufacturer's price does not impact the extensive consumer margin. A manufacturer selling directly to consumers then charges the monopoly price $p_m = c\frac{\sigma}{\sigma-1}$.

Also, because consumers do not detour, the best location for a manufacturer who sells directly to consumers is at the location of a retailer. In this case, the manufacturer sells to the retailer's entire clientele. At any other location, the manufacturer sells only to a fraction of that clientele. Thus, depending on its location relative to the retailer, the manufacturer receives at most patronage from a mass of customers equal to $\frac{s}{R}$. Suppose it does.

The retailer charges a price of $p = \pi w$ for each of the mass N varieties it sells. Quantity demand for the candidate direct-selling manufacturer is $x = A(N)\,p_m^{-\sigma}$ per consumer with $A(N) = \frac{Y}{Np^{1-\sigma}}$. The total variable profits for the direct seller with the best location is therefore

$$\underbrace{\tfrac{s}{R}}_{\substack{\text{best case mass}\\ \text{of consumers}}} \times \underbrace{x}_{\substack{\text{quantity}\\ \text{demand}}} \times \underbrace{\frac{p_m-c}{p_m}}_{\substack{\text{percentage}\\ \text{margin}}} = \left(\frac{s}{R}\right)\left(\frac{Y}{N}\right)\left(\frac{p}{p_m}\right)^{\sigma-1}\left(\frac{p_m-c}{p_m}\right) \tag{A.3}$$

Note that RN is the total entry cost of the retail sector divided by γ. In equilibrium, the total entry costs in the retail sector equal gross sector profits $sY\frac{\pi-1}{\pi}$. Thus, $RN = \frac{sY}{\gamma}\frac{\pi-1}{\pi}$. The retail price is $p = \pi w$. Making these substitutions, and recalling $p_m = c\frac{\sigma}{\sigma-1}$, the right hand side of equation A.3 becomes,

$$\text{marginal profits} = \gamma\frac{\pi}{\pi-1}\left(\frac{\pi w}{c}\frac{\sigma-1}{\sigma}\right)^{\sigma-1}\frac{1}{\sigma}.$$

Finally, the direct seller does not enter if these variable profits are less than entry cost f . This yields the following "no-entry" condition for direct sellers.

$$\frac{\pi}{\pi-1}\left(\frac{\pi w}{c}\frac{\sigma-1}{\sigma}\right)^{\sigma-1}\frac{1}{\sigma} < \frac{f}{\gamma}. \tag{A.4}$$

To evaluate this condition, first observe that the other manufacturers maximally charge monopoly prices. Therefore, $\frac{w}{c} \leq \frac{\sigma}{\sigma-1}$. Substituting this in the LHS of the no-entry condition A.4, I obtain

$$\frac{\pi}{\pi-1}\left(\frac{\pi w}{c}\frac{\sigma-1}{\sigma}\right)^{\sigma-1}\frac{1}{\sigma} \leq \frac{\pi^{\sigma}}{\pi-1}\frac{1}{\sigma}.$$

Thus, if

$$\frac{\pi^{\sigma}}{\pi-1}\frac{1}{\sigma} < \frac{f}{\gamma} \Rightarrow \text{A.4 is true}$$

My strategy is next to look up reasonable ranges for the values for σ and π and compute the maximum for the quantity on the LHS.

Recall π is the retail multiplier on the wholesale price of goods sold. I use data reported by Damodaran (2016) on the profitability of US retailers in different industries to compute the multiplier π. More specifically, I use earnings before interest, taxes, depreciation and amortization (EBITDA) and sales, general and administrative expenses (SG&A). Using these industry data, π ranges between 1.27 (Automotive) and 1.52 (Building Supply). Using estimated values reported in the literature, σ has a range of 2-7 (see section 5.4). For these ranges of π and σ, the LHS of the no-entry condition A.4 can be bounded at maximally 5.15.

A.4 Proof of Propositions 1,2, and 3

Stage 2: The retailer.
The problem for a representative retailer is as follows

$$\max_{\pi,N} 2d(\pi,N)\, sY\frac{(\pi-1)}{\pi} - g(N) \;\; \text{s.t. } N \leq M,$$

where the mass of entering manufacturers M is fixed. The Lagrangian is

$$\mathscr{L}_R = 2d(\pi,N)\, sY\frac{(\pi-1)}{\pi} - g(N) - \lambda(N-M).$$

The KKT conditions are (1) stability $\frac{\partial \mathscr{L}_R}{\partial \pi} = 0$, $\frac{\partial \mathscr{L}_R}{\partial N} = 0$, (2) primal feasibility, $N \leq M$, (3) dual feasibility, $\lambda \geq 0$, and (4) complementary slackness $\lambda \times (N-M) = 0$.

I first develop the derivatives $\frac{\partial \mathscr{L}_R}{\partial \pi}$ and $\frac{\partial \mathscr{L}_R}{\partial N}$. Competing retailers charge margins π_c, and offer an assortment N_c. Demand for the retailer is determined by the location of the indifferent consumer. Two stores are a distance $\frac{1}{R}$ apart, and a consumer located between them is at a distance d from the focal retailer and at a distance $\frac{1}{R}-d$ from the nearest competitor. Then, the location of the marginal consumer is determined from the indifference condition $U(\pi w, N, d) = U\left(\pi_c w, N_c, \frac{1}{R}-d\right)$,

$$\left(\frac{Y}{\pi W(N)}\right)^{1-\rho}(T-td)^{\rho} = \left(\frac{Y}{\pi_c W(N)}\right)^{1-\rho}\left(T-t\left(\frac{1}{R}-d\right)\right)^{\rho},$$

with the wholesale price index $W(N) = \left(\int_0^N \left(w(v)^{1-\sigma}\right) dv\right)^{\frac{1}{1-\sigma}}$. The solution in d gives the size of the store catching area on each side of the retailer:

$$d(\pi, N) = \frac{\left((\pi W(N))^{\frac{\rho-1}{\rho}} - (\pi_c W(N_c))^{\frac{\rho-1}{\rho}}\right) RT + (\pi_c W(N_c))^{\frac{\rho-1}{\rho}} t}{\left((\pi_c W(N_c))^{\frac{\rho-1}{\rho}} + (\pi W(N))^{\frac{\rho-1}{\rho}}\right) Rt} \tag{A.5}$$

The notation $d(\pi, N)$ used here and in the body of the text, is shorthand for $d(\pi, N, \pi_c, N_c)$. The stability condition for the retail multiple π is

$$d_\pi(\pi-1) + d\frac{1}{\pi} = 0. \tag{A.6}$$

The stability condition for store assortment N is

$$2d_N sY\frac{(\pi-1)}{\pi} - g_N - \lambda = 0. \tag{A.7}$$

Next, take derivatives d_π and d_N of equation (A.5), holding π_c and N_c fixed. Simplify, using that in a symmetric equilibrium, $\pi = \pi_c$ and $N = N_c$. The derivatives are

$$d_\pi = -\left(\frac{1-\rho}{\rho}\right)\left(\frac{T}{t} - \frac{1}{2R}\right)\frac{1}{2\pi}, \tag{A.8}$$

and

$$d_N = -\left(\frac{1-\rho}{\rho}\right)\left(\frac{T}{t} - \frac{1}{2R}\right)\frac{W_N(N)}{2W(N)}. \tag{A.9}$$

Simplify equation (A.9) using that wholesale prices are equal. Now $W(N) = \left(\int_0^n \left(w(v)^{1-\sigma}\right) dv\right)^{\frac{1}{1-\sigma}}$ becomes $W(N) = N^{\frac{1}{1-\sigma}} w$ and the index's derivative at N, becomes $W_N(N) = \frac{1}{1-\sigma} N^{\frac{1}{1-\sigma}-1} w$. This means that we can write equation (A.9) as

$$d_N = \left(\frac{1-\rho}{\rho}\right)\left(\frac{T}{t} - \frac{1}{2R}\right)\left(\frac{1}{\sigma-1}\right)\frac{1}{2N} \tag{A.10}$$

To solve for the retail margin, substitute the derivative (A.8) in the first order condition (A.6).

Combine this with $d = \frac{1}{2R}$. Collecting terms gives

$$\pi = 1 + \left(\frac{\rho}{1-\rho}\right)\frac{2t}{(2RT-t)}. \tag{A.11}$$

This equation holds regardless of whether the retailer's variety constraint binds or not. For the assortment size of the retailer's store N, I obtain the following:

1. I first consider the case that $N < M$. This means that $\lambda = 0$. Substitute equation (A.10) into the first order condition (A.7) to obtain

$$sY\left(\frac{T}{t} - \frac{1}{2R}\right)\left(\frac{1-\rho}{\rho}\right)\left(\frac{\pi-1}{\pi}\right) = N\gamma(\sigma-1). \tag{A.12}$$

   (a) Entry R is determined from the zero profit condition that at the industry level the total variable retail profits must equal R times entry costs γN, or $\left(\frac{\pi-1}{\pi}\right)sY = R\gamma N$. Substitute this relation into Equation (A.12) and rearrange to obtain Equation (19)

$$R^* = \left(\frac{t}{2T}\right)\left(\frac{1+2\rho\sigma-3\rho}{1-\rho}\right).$$

   (b) Using this result for R^* in equation (A.11), the retailer margin simplifies to equation (17)

$$\pi^* = \frac{\sigma}{\sigma-1},$$

   (c) and optimal assortment becomes equation (15)

$$N^* = \left(\frac{sY}{\sigma\gamma}\right)\left(\frac{2T}{t}\right)\left(\frac{1-\rho}{1+2\rho\sigma-3\rho}\right).$$

   R^*, π^*, and N^* are all functions of parameters only.

2. Next, consider that the profit maximizing mass of varieties to the retailer is larger than the mass of entering manufacturers: $N^* > M$.

   (a) In this case, the constraint binds, $\lambda > 0$, and complementary slackness implies that the mass of varieties in the retailer's assortment is equal to the upstream supply of variety: $N = M$.

   (b) The retailer now competes with a constrained assortment, which affects retail markups, $\pi(M)$, through retail entry, $R(M)$. Equation (A.11) still holds.

$$\pi(M) = 1 + \left(\frac{\rho}{1-\rho}\right)\frac{2t}{(2TR(M)-t)} \tag{A.13}$$

   (c) Retailers enter until their profits are driven to 0. The total retail sector earns $\left(\frac{\pi-1}{\pi}\right)sY$. Each retailer gets an equal $\frac{1}{R}$ share of these profits. In equilibrium, this is equal to the cost of entering and providing assortment ,γM. This determines the number of entering retailers as

$$R(M) = \left(\frac{\pi(M) - 1}{\pi(M)} \right) \frac{sY}{\gamma M}. \tag{A.14}$$

Note that this is an implicit equation because π in equation (A.14) depends directly on R. I will solve the explicit expressions below.

(d) The constraint binds if $N^* > M$ which means that

$$2 \frac{sY}{\sigma\gamma} \frac{T}{t} \left(\frac{1-\rho}{1+2\rho\sigma-3\rho} \right) > M$$

3. Finally, $N^* = M$, it is easy to show that case 1 holds.

We can now consider the first stage of the game.

Stage 1: The manufacturer problem

Wholesale prices are set such that a manufacturer maximizes profits subject to the distribution constraint. Then, each manufacturer v's problem is

$$\max_{w(v)} s \times A(M) (\pi w(v))^{-\sigma} \times (w(v) - c) - f, \text{ such that } w(v) \leq w_N$$

where the demand shifter $A(M) = \frac{Y}{Mp^{1-\sigma}}$ and w_N is the wholesale price at mass N in the distribution of wholesale prices. The Lagrangian can be written as

$$\mathscr{L}_M = s \times A(M) (\pi w(v))^{-\sigma} \times (w(v) - c) - f - \mu (w(v) - w_N)$$

The KKT conditions are (1) stability $\frac{\partial \mathscr{L}_M}{\partial w} = 0$, (2) primal feasibility, $w(v) \leq w_N$, (3) dual feasibility, $\mu \geq 0$, and (4) complementary slackness $\mu \times (w(v) - w_N) = 0$.

1. The manufacturer constraint binds when more manufacturers can enter at monopoly prices than retailers want. With many varieties to choose from, the retailer is not constrained and sets π^* and N^*. Further, R^* retailers will enter.

   (a) Now, any manufacturer is required to set a wholesale price that maximizes its profits such that its variety belongs to the N^* least expensive ones. This means that manufacturers undercut each other to the point where only N^* of them enter. Realizing that downstream retail competition is unconstrained and results in π^*, the equilibrium wholesale price solves

$$M = N^* \longleftrightarrow \frac{sY}{\pi^*} \frac{w-c}{wf} = 2 \frac{sY}{\sigma\gamma} \frac{T}{t} \left(\frac{1-\rho}{1+2\rho\sigma-3\rho} \right).$$

   Simplifying this equation results into equation (20)

$$w = c \frac{(\sigma-1) R^*}{(\sigma-1) R^* - \dfrac{f}{\gamma}}.$$

(b) The manufacturer constraint on wholesale prices binds when $w < w^*$ or when the wholesale price for the manufacturer is lower than their monopoly price. This in turn implies equation (16),

$$\frac{f}{\gamma} < \left(\frac{t}{2T}\right)\left(\frac{1+2\rho\sigma-3\rho}{1-\rho}\right)\left(\frac{\sigma-1}{\sigma}\right). \tag{A.15}$$

2. Next, consider the case where the manufacturer is guaranteed distribution. Then, the constraint on wholesale prices does not bind. This is the case when the retailer wants more variety than free-entry manufacturers collectively provide, i.e., when $N^* > M$. In this case, the downstream retailer decisions are governed by the fact that the constraint in the retail sector binds, $\lambda > 0$, i.e., by equations (A.13) and (A.14).

   (a) Given that the constraint in the manufacturing sector does not bind, complementary slackness requires that $\mu = 0$. Recalling constant elasticity demand, profit maximizing wholesale prices are equal to equation (25)

   $$w^* = c\frac{\sigma}{\sigma-1}.$$

   (b) In equilibrium, all manufacturers charge the same wholesale price. Given that all manufacturers charge $w^* = c\frac{\sigma}{\sigma-1}$, or $\frac{w^*-c}{w^*} = \frac{1}{\sigma}$, the free-entry mass of manufacturers who can recoup their fixed entry cost is,

   $$M = \underbrace{\frac{sY}{\pi(M)}}_{\substack{\text{expenditure}\\ \text{mnf'd goods}}} \times \underbrace{\tfrac{1}{\sigma}}_{\substack{\text{\% profit}\\ \text{margin}}} : \underbrace{f}_{\substack{\text{entry}\\ \text{cost}}}. \tag{A.16}$$

   We can use these first stage decisions to simplify the second stage. Using equation (A.16) and substituting it into equation (A.14) gives a quadratic equation with only one positive root. This results in equation (24)

   $$R = \frac{t}{4T}\left(1+\sqrt{1+16\frac{T}{t}\left(\frac{\rho}{1-\rho}\right)\frac{\sigma f}{\gamma}}\right).$$

   Substitution of this expression into (A.11) gives equation (22)

   $$\pi = 1+\left(\frac{\rho}{1-\rho}\right)\frac{4}{\left(\sqrt{1+16\frac{T}{t}\left(\frac{\rho}{1-\rho}\right)\frac{\sigma f}{\gamma}}-1\right)}.$$

Finally, substituting π into (A.16), I obtain equation (26)

$$M = \frac{sY}{\sigma f} \frac{\sqrt{1+16\frac{T}{t}\left(\frac{\rho}{1-\rho}\right)\frac{\sigma f}{\gamma}}-1}{\sqrt{1+16\frac{T}{t}\left(\frac{\rho}{1-\rho}\right)\frac{\sigma f}{\gamma}}-1+\frac{4\rho}{1-\rho}}.$$

(c) The manufacturer constraint does not bind when $M < N^*$ or

$$\frac{sY}{\sigma f} \frac{\sqrt{1+16\frac{T}{t}\left(\frac{\rho}{1-\rho}\right)\frac{\sigma f}{\gamma}}-1}{\sqrt{1+16\frac{T}{t}\left(\frac{\rho}{1-\rho}\right)\frac{\sigma f}{\gamma}}-1+\frac{4\rho}{1-\rho}} < 2\frac{sY}{\sigma\gamma}\frac{T}{t}\left(\frac{1-\rho}{1+2\rho\sigma-3\rho}\right).$$

Write $x = \frac{2Tf}{t\gamma}$, and simplify

$$\frac{\sqrt{1+8x\left(\frac{\rho}{1-\rho}\right)\sigma}-1}{\sqrt{1+8x\left(\frac{\rho}{1-\rho}\right)\sigma}-1+\frac{4\rho}{1-\rho}} < x\left(\frac{1-\rho}{1+2\rho\sigma-3\rho}\right). \tag{A.17}$$

It is easily verified that this inequality has the following two roots,

$$x = \left\{0, \left(\frac{1+2\rho\sigma-3\rho}{1-\rho}\right)\left(\frac{\sigma-1}{\sigma}\right)\right\}.$$

Next, note that the LHS and the RHS of equation (A.17) rises in x. Further, the LHS is strictly concave in x for all $x \geq 0$, and the RHS is linear. A concave and a linear function have at most two points of intersection. Therefore, the two roots above are the only roots. It is then easily verified that the inequality in equation (A.17) holds for all

$$x > \left(\frac{1+2\rho\sigma-3\rho}{1-\rho}\right)\left(\frac{\sigma-1}{\sigma}\right).$$

Finally, this implies the following relation among the parameters in the model

$$\frac{f}{\gamma} > \left(\frac{t}{2T}\right)\left(\frac{1+2\rho\sigma-3\rho}{1-\rho}\right)\left(\frac{\sigma-1}{\sigma}\right). \tag{A.18}$$

Thus, when the ratio of the fixed cost f of manufacturing a variety and the fixed cost $\gamma R^*$of distributing a variety using R^* retailers is larger than $\frac{\sigma-1}{\sigma}$, retailers want more variety than manufacturers can make and the unique equilibrium is as described above.

3. To complete the proof, when $\frac{f}{\gamma} = \frac{t}{2T}\left(\frac{1+2\rho\sigma-3\rho}{1-\rho}\right)\left(\frac{\sigma-1}{\sigma}\right)$, both constraints bind (just) at the same time, the retailer assortment size N^* is equal to the variety producing capacity of the manufacturing sector M^*, and it is trivial to verify that the equilibrium is $[M^*, w^*, R^*, N^*, \pi^*]$.

A.5 *Proof of Corollary 1*

As long as the consumer does not buy all varieties, competition over the extensive consumer margin will be very strong. Consumer will circumnavigate and buy all varieties when his time budget allows him to can travel as far away as half of the circle. The consumer is can then return by circumnavigation and buy all varieties. The consumer travels a circle segment of length $\frac{T}{t}\frac{(1-\rho)}{(1+\sigma\rho-2\rho)}$. Thus, the consumer will not buy all varieties if

$$\frac{T}{t}\frac{(1-\rho)}{(1+\sigma\rho-2\rho)} < \frac{1}{2}.$$

This immediately produces the result.

A.6 *Proof of Corollary* (2)

1. For the case that $N^* < M^*$, use equation (19) in the right hand side of (16) to get $\frac{f}{\gamma} > \frac{\sigma-1}{\sigma}R^*$. Next, substitute this inequality into equation (20) to obtain

$$w = c\left(\frac{(\sigma-1)R^*}{(\sigma-1)R^* - \dfrac{f}{\gamma}}\right) < c\left(\frac{(\sigma-1)R^*}{(\sigma-1)R^* - \dfrac{\sigma-1}{\sigma}R^*}\right) = c\left(\frac{\sigma}{\sigma-1}\right).$$

   And thus the percentage margin in manufacturing $\frac{w-c}{w}$ is smaller than $\frac{1}{\sigma}$, whereas from equation (17) the percentage margin in retail is $\frac{1}{\sigma}$.

2. For the case that $N^* = M^*$ both percentage margins are $\frac{1}{\sigma}$

3. For the case that $N^* > M^*$, start with the opposite of equation (16), $\frac{f}{\gamma} > \left(\frac{t}{2T}\right)\left(\frac{1+2\rho\sigma-3\rho}{1-\rho}\right)\left(\frac{\sigma-1}{\sigma}\right)$, and substitute this inequality in the equilibrium expression for π, i.e., equation (22)

$$\pi = 1 + \frac{4\left(\frac{\rho}{1-\rho}\right)}{\sqrt{1+16\frac{T}{t}\frac{\sigma f}{\gamma}\left(\frac{\rho}{1-\rho}\right)}-1}.$$

   to obtain

$$\pi < 1 + \frac{4\left(\frac{\rho}{1-\rho}\right)}{\sqrt{1+8\left(\frac{1+2\rho\sigma-3\rho}{1-\rho}\right)(\sigma-1)\left(\frac{\rho}{1-\rho}\right)}-1},$$

   which simplifies to

$$\pi < \frac{\sigma}{\sigma-1}.$$

   In turn, this means that the percentage margin in the retail sector is smaller than $\frac{1}{\sigma}$. Equation (25) implies that the percentage margin in the manufacturing sector is $\frac{1}{\sigma}$.

A.7 *Welfare*

1. First consider the two-sector model. A consumer at location ℓ has utility

$$U_\ell = \left(M^{\frac{\sigma}{\sigma-1}} x\right)^{1-\rho} (T - t\ell)^\rho .$$

Aggregated across all consumers, total welfare is,

$$U = 2Rs \left(M^{\frac{\sigma}{\sigma-1}} x\right)^{1-\rho} \int_0^{\frac{1}{2R}} (T - t\ell)^\rho \, d\ell.$$

Per variety quantity x is net income Ψ divided by Mp. Total utility welfare then is

$$U = 2Rs \left(M^{\frac{1}{\sigma-1}} \frac{\Psi}{p}\right)^{1-\rho} \frac{1}{t(\rho+1)} \left(T^{\rho+1} - \left(T - \frac{t}{2R}\right)^{\rho+1}\right).$$

Net income is personal income minus an equal share of the entry cost of M manufacturers and R retailers $\Psi = Y - \frac{M}{s}(f + R\gamma)$. Further, substituting $p = c$ gives

$$U = 2Rs \left(M^{\frac{1}{\sigma-1}} \frac{sY - M(f + R\gamma)}{sc}\right)^{1-\rho} \frac{1}{t(\rho+1)} \left(T^{\rho+1} - \left(T - \frac{t}{2R}\right)^{\rho+1}\right).$$

The social planner maximizes this with respect to M and R.

2. Next, in absence of retailers, the consumer travels to gather variety and his leisure is $T\rho \frac{\sigma-1}{1+\rho\sigma-2\rho}$. Using the same logic as above, the social planner maximizes

$$U = s \left(M^{\frac{1}{\sigma-1}} \frac{sY - Mf}{sc}\right)^{1-\rho} \left(T\rho \frac{\sigma-1}{1+\rho\sigma-2\rho}\right)^\rho,$$

with respect to M. This has a maximum at the same value of M as the function $M^{\frac{1}{\sigma-1}}(sY - Mf)$ and gives

$$M = \frac{sY}{\sigma f}.$$

B Comparative Statics

B.1 *Details*

Figures (B.1)-(2) are constructed from the equilibria in Equations (18)-(21) and Equations (26)-(23) with the values for parameters as in Table B.1.

The baseline parameter values reported obey equation (16) in equality. This accomplishes that the equilibrium graphs are centered at the case where $N^* = M^*$ when plotting the equilibrium at baseline-centered ranges of the parameters.

parameter	symbol	baseline	range use in graphs
population	s	100	–
income	Y	100	–
time resource	T	100	–
preference for leisure	ρ	0.5	0.25 – 0.75
cost of travel	t	1000	200 – 1800
elasticity of substitution	σ	5	3.5 – 6.5
entry cost of distribution	γ	.5	0.1 – 0.9
entry cost in manufacturing	f	18	9 – 27
marginal cost of production	c	1	–

Table B.1: Parameter values used in the construction of the equilibrium graphs

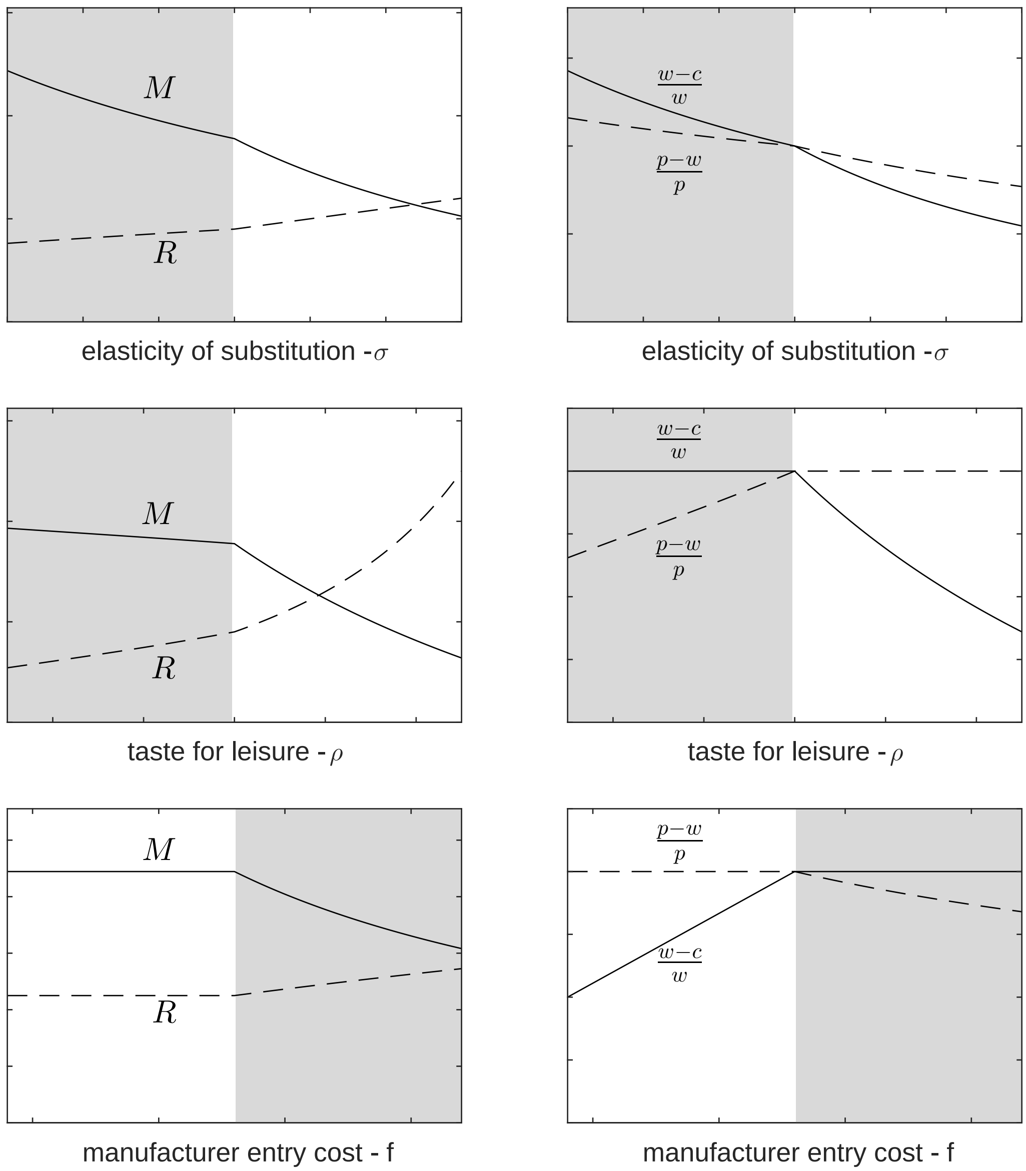
B.2 Additional Graphs

The top panel of figure B.1 depicts the impact of the elasticity of substitution. As varieties become closer substitutes, the manufacturing sector produces fewer varieties in response to retailers providing smaller assortments. In turn, this is because consumers derive less utility from variety and are more attracted by convenience. Store density increases and store size falls. Prices and wholesale prices both fall as products become closer substitutes. The retailer percentage margin, $\frac{p-w}{p}$, is larger than the manufacturer percentage margin, $\frac{w-c}{w}$, if and only if $N^* > M^*$, i.e., distribution of variety is scarce.

The middle panel considers the role of leisure. As the preference for leisure increases, store density R rises. Variety of assortment drops. It is initially slow. As long as manufactured variety is scare ($N^* < M^*$), the adjustment in variety takes place in the manufacturing sector. Although retailers want more variety, their expanding cost leaves less resources for manufacturing entry. However, beyond the point where the retailer can find all the variety it wants ($N^* = M^*$), variety of assortment drops fast because the rising preference for leisure compels retailers to open up smaller "corner stores" with limited assortment. Moving to the right-hand graph, as consumer preference for leisure rises, retailers initially increase their margins as the . However, as the preference for leisure increases further, more stores with smaller assortments enter and eventually manufacturers need to drop prices because at monopoly margins there is too much entry. At the same time, retailers now hold their margin constant (see Equation 17) and charge prices $p = w\frac{\sigma}{\sigma-1}$.

The bottom panel of figure B.1 shows the impact of the manufacturing entry cost. As entry cost rises above the level where it constrains the retailer's assortment, manufacturer entry falls and retailer entry rises. When entry costs are low, the manufacturer margins are low also, driven down by the need to compete away entry in the manufacturing sector.

Figure B.1: Additional comparative statics



Notes: The shaded area represents the case where variety is scarce, $N^* > M^*$. The non-shaded area represents the opposite case where shelf space is scarce, $N^* < M^*$.