Product Variety and Competition in the Retail Market for Eyeglasses

Randal Watson*
r-watson@northwestern.edu

March 14, 2003

Abstract

I analyze an original dataset on the display inventories of several hundred eyewear retailers to study how firms' product-range choices depend on separation from rivals in geographically-differentiated markets. A two-stage estimation approach is used to model firms' initial location decisions and their subsequent choices of product variety, borrowing methodologies from Seim (2002) and Mazzeo (2000). Per-firm variety varies non-linearly with the degree of local competition. Holding fixed the total number of rivals in a market, a retailer stocks the widest variety when it is near several other competitors. Its product range is somewhat smaller if it faces no local competition, and substantially smaller if it faces four or more nearby rivals. This suggests that business-stealing eventually dominates any clustering effects when there is intense competition in a neighbourhood.

*For comments and guidance at various stages of this project I am grateful to Michael Mazzeo, Robert Porter, and Asher Wolinsky. I also had helpful conversations with David Barth, Avi Goldfarb, Ithai Lurie and Robert Vigfusson, and seminar participants at Northwestern. Partial financial support is gratefully acknowledged from a Northwestern University Graduate School Graduate Research Grant, and from the University's Centre for the Study of Industrial Organization. All errors herein are my responsibility.

1 Introduction

This paper is an empirical examination of how firms compete in the variety of products that they offer to consumers. Consider a retail market in which each store sells many horizontally differentiated varieties of a single class of good, for example, music CD's, books, clothes, or video rentals. Consumers in such markets typically have idiosyncratic preferences over the different available styles of the good. They may need to search across multiple retailers to find the outlet that sells their preferred combination of style and price, in which case they are naturally drawn to sellers with a broader range of available varieties. A store's choice of product variety is then a strategic variable, depending endogenously on the variety choices of its competitors. Thus a music store manager choosing whether to add CD's to his stock weighs the costs of additional inventory and display space against the increased probability that customers find a good match for their musical tastes here, rather than at a rival outlet elsewhere.

When consumer preferences are not directly observable, the choice of optimal inventory size in such situations may become a matter of (costly) speculation. For example the Blockbuster video rental chain reportedly spent 50 million dollars in the late 1990's on a marketing experiment that drastically increased inventories at outlets in six test markets. Management guessed that extra video tapes in stores might substantially raise revenues by matching more customers with their most desired movies. Results from the pilot project subsequently encouraged Blockbuster to implement a new business model based on expanded retail inventories and revenue sharing with movie studios.[1]

In that particular case the key to improving the store-level availability of good matches to consumer tastes may have been the depth of inventory in popular movie titles. However the breadth of retail inventory is no doubt also an important element in such calculations. Throughout this paper I focus on this breadth variable, measured as the number of different styles of a good on display at each outlet. I use the terms 'product variety' and 'product range' to denote this measure of inventory coverage.[2]

How then does a retail manager adjust his product range if a new rival opens next door? What if the rival is three miles away? Does the incumbent's response depend upon the number of other competitors already in

[1] Redstone (2001).

[2] Note that terms like product range are not meant to connote distances in an explicit space of product characteristics. Rather they refer to the number of different product lines carried by a retailer.

place? Using original data the present study aims to provide answers to these questions in a particular context: the retailing of eyeglasses. I delineate a sample of geographic markets in the Midwestern U.S. and develop a two-stage econometric model of competition amongst eyewear sellers. In the first stage the entry behaviour of sellers in each market is modeled using Seim's (2002) framework for endogenous location choice. The second stage then shows how sellers' product- range choices depend upon the resulting configurations of competitors in each market. As in Mazzeo (2000), estimates from the first stage provide corrections for the endogeneity of seller locations in the second stage.

Eyeglasses were chosen for analysis firstly because they are usually (but not always) sold at businesses dedicated to eyecare.[3] The potential statistical interference from a store's other lines of business is thereby minimized; this interference could be a problem if, for example, books or CD's were under study. Second, eyewear sellers typically stock hundreds of different styles of spectacle frames, reflecting heterogeneity in consumer tastes for colour, shape and construction. A measure of the number of different frame styles in a seller's display inventory can then be used as an indicator of product variety. Third, consumer behaviour in this market can be thought of in terms of sequential search. Any prospective buyer of eyeglasses may need (because of the infrequency of purchase) to search across stores for information on prices and styles. This suggests a link to theories of spatial competition that are based on consumer search.[4]

These models are often concerned with the agglomeration phenomenon: why firms sometimes cluster in the product space. Heterogeneity in consumer preferences (which is the reason for product variety) strengthens the agglomeration incentive: consumers like clusters of firms in part because they facilitate comparisons between different styles of a good. Assuming one variety per firm, Wolinsky (1983) and Konishi (1999) show that this preference for a wide product selection can induce producers to cluster at a particular point in the product space, notwithstanding the intense competition there. A model with multi- variety firms is Anderson and de Palma (1992); however it has no search or spatial elements and is therefore not directly concerned with agglomeration. To link clustering with competition in product variety it is thus at present necessary to rely on informal argu-

[3] Previous empirical studies which analyze other aspects of the eyecare industry include Benham (1972), Kwoka (1984), and Haas-Wilson (1989).

[4] See Wolinsky (1983), Dudey (1990), Konishi (1999), and Fujita and Thisse (2002, Ch. 7), among others.

ments.[5] It is suggested that clustering among multi-variety firms might still be observed in theory, depending upon specifics of the market such as the costs of search and the distribution of consumer preferences.

The empirical evidence presented below is consistent with a limited degree of clustering. When the distance to rival sellers is reduced the baseline effect is for the profits and product variety of a firm to fall. This may be interpreted as reflecting loss of customers to the closer competition. However these responses are significantly less negative when the incumbent faces relatively little local competition, with few proximate competitors. In fact, all else equal, the widest per-seller variety in a market might be at stores which have a few rivals nearby, rather than at local monopolies. Groups of nearby sellers make the customers there 'more choosy', leading each seller to boost variety. Moreover the increased variety and lower search costs at a retail cluster could induce consumers to switch purchases to this location from elsewhere in the market, giving firms a further incentive to expand their product ranges. When just a few competitors are grouped together, the data suggest that these factors at least ameliorate, and may dominate, the baseline business-stealing effect noted above.

Several other recent papers (e.g., Thomadsen (1999), Manuszak (2000), Davis (2001)) have looked at the relationship between retail competition and geographic differentiation. These studies take firms' locations as given. Thanks to the entry model of Seim (2002) the following analysis is able to address spatial competition in a model with endogenous firm locations. As far as I am aware this combination of Seim's strategic location framework with information about firms' post-entry interactions is novel in this literature. Collecting the data needed to effect this combination was a non-trivial task, involving visits to several hundred widely separated eyewear sellers. While time-consuming, this approach gives the researcher much greater control over sources of measurement error than would be possible with, for example, a simple mail survey.

Empirical interest in competition in product variety is of a fairly recent genesis. Berry and Waldfogel (2001) examine variety choices in radio broadcasting. They ask how the product range in local radio markets – measured as the number of different programming formats on air – is affected by mergers. A reduced-form study in the marketing literature is Bayus and Putsis (2000), who analyze the determinants of additions to, and deletions

[5] Watson (in progress) incorporates multiproduct firms into the search-based monopolistic competition framework of Wolinsky (1986). Firms' strategic choices of product ranges can also be thought of as competition in product availability – on the latter see Dana (2001).

from, the product lines of personal-computer makers. Draganska and Jain (2001) is a more structural marketing analysis. Studying the retailing of yoghurt, they introduce choices of product range (number of flavors) into a differentiated-goods model of interfirm competition.[6]

In contrast to the above studies, the aim in the present paper is to elicit the relationship between a firm's chosen product range and its differentiation from its competitors. I model explicitly the determination of each firm's location in the product space, and then relate a measure of its product range to its distance from rival sellers. This approach gives a richer picture of the competitive effects than if competition is measured by a market-level quantity, e.g., a concentration ratio, or total number of entrants, as in Berry-Waldfogel or Bayus-Putsis. My results suggest that distance from rivals is indeed an important determinant not only of firms' profits but also of their product ranges.

An examination of the market-level aggregates in my data highlights the advantages of relating firm behaviour to within-market differentiation. Figure 1 graphs for each market the number of inhabitants per eyewear seller against total market population. There is a clear positive relationship.[7] Following the arguments in Bresnahan and Reiss (1991) this might reflect variations in competitive behaviour across markets, with larger towns sustaining more intense competition amongst sellers. Such variations in the intensity of competition are not immediately obvious from the market-level data on product variety. Figure 2 plots market population against the mean logarithm of the number of eyeglass styles per seller.[8] Allowing for heteroscedasticity, the trend in the figure is fairly flat. The results to follow nevertheless reveal significant relationships between product variety and competition when these are compared across different locations within a market.

The next section introduces an intuitive analysis of the behaviour of firms and consumers in a theoretical eyewear market. This discussion is not rigorous; rather it is intended as a useful reference point for interpreting the results of the subsequent product-variety regressions. Section 3 outlines

[6] See also Israelevich (2002). A simulation approach based on data from a small cross-section of video rental stores is de Palma et al. (1994).

[7] In a linear regression an extra one thousand inhabitants raises people-per-seller by two hundred – the effect is significant at the 1% level and remains significant when per capita income is included in the regression.

[8] An earlier version of the paper plotted market-level variety as an average of levels, rather than logs. The measurement errors in my variety counts are likely to be multiplicative, rather than additive, so the mean logarithm, i.e., the log of the geometric average, is a more appropriate measure.

the application of the entry model, discussing in particular problems of multiple equilibria. A model of post-entry competition in product ranges is expounded in Section 4, and the data are explained in Section 5. Estimates of firm-level profit and variety equations are in Section 6. That section also considers firms' predicted responses to alternative configurations of seller locations. Section 7 concludes.

2 An informal framework

The medium-sized towns that appear in this study are each thought of as a finite collection of points scattered unevenly over a circle in 2-dimensional space. Consumers with heterogeneous demographic characteristics are distributed irregularly through this set of locations, along with sellers of eyeglasses and other points of retail interest like shopping malls. It may be difficult to derive equilibrium predictions for product-variety competition in a rigorous theory for such settings.[9] Instead I use the intuition of simpler theoretical frameworks to inform the interpretation of the empirical results to follow.

Assume that different varieties (styles) of eyeglass frames are a horizontally differentiated good, in the sense that consumers have an idiosyncratic taste v for each variety. Each consumer demands at most one pair of eyeglasses. Let the taste (or customer-specific quality) v be identically and independently distributed across both varieties and consumers. Thus there is no vertical product differentiation: consumers do not agree on a common quality ranking of all available varieties. Furthermore a consumer who dislikes any given variety at a particular seller is no less inclined to appreciate the other styles available at that outlet.

I hypothesize that the distances between firms affect their product-range choices, because these distances are related to consumer travel costs. It is convenient to think of this relationship in search-theoretic terms. Assume that consumers *ex ante* do not know their tastes for the varieties stocked by any seller; nor do they observe *ex ante* any seller's prices. Instead they must learn tastes and prices in a process of sequential search, visiting sellers in turn until they find an acceptable combination of price and (customer-specific) quality. The cost of searching at a seller is the cost of traveling to that location, which is increasing in distance.[10]

[9] De Palma *et al.* (1994) resort to simulation techniques to analyze equilibria in a particular example of this kind of market.

[10] Assume that search within any outlet is also sequential.

For simplicity assume that each seller charges a single price for all varieties in its stock, and that in equilibrium all firms at a given location behave symmetrically with respect to both prices and number of styles on offer. Thus prices and variety-per-firm may vary across locations in a market but not across firms at the same location. Consumers *ex ante* observe the numbers of sellers at each site in town and can use this information to infer the prices and variety-per-firm that would result in equilibrium at each location. Based on this inference they choose which location to visit next in their search sequence. On arriving at a location with multiple sellers they choose their initial point of call there randomly.

This framework is an embellishment of the search-theoretic model of monopolistic competition introduced in Wolinsky (1986).[11] The analysis in that study focuses on firms which each sell a single variety, and which are symmetrically situated - the cost of switching to search at a new seller is the same constant everywhere. Watson (in progress) attempts to extend the model to incorporate sellers with multiple varieties, while retaining the assumption of symmetric locations. Results from that extension suggest that on the demand side a firm's equilibrium choice of variety is driven by two factors.[12] First, holding the actions of all other firms constant, a seller's product range is increasing in the number of consumers who visit the store. Extra visitors raise the probability of sale (and therefore the marginal revenue) at a given price of any variety in stock. Second, a seller's product range is decreasing in consumer search costs, holding fixed the number of visitors to this seller. Higher search costs reduce the reservation value of anyone currently searching at this seller: they make visitors 'less choosy' about the seller's inventory. In the present context the search cost that enters the seller's variety optimization is the distance the customer would have to travel to visit the next alternative seller. As this distance increases so rises the consumer's opportunity cost of rejecting all vareties at the current seller and switching to the alternative store.

Consider then a general setting in which a particular seller A has sole occupancy of one of the town's locations and several other sellers, say four or more, are distributed (perhaps in asymmetric clusters) across the other sites in the market.[13] Suppose that an entirely new seller B enters the market at the same location as A. For simplicity we will assume that this incremental change in the number of sellers does not affect the number of consumers in

[11] See also Anderson and Renault (1999).

[12] On the cost side I assume that retailers at all locations face the same constant marginal costs of stocking eyeglass frames.

[13] In the data the number of sellers per market ranges from 4 to 23, with a mean of 13.

the market who are searching for spectacles.

Three effects may then be imagined in the move to a new equilibrium. Half of A's previous visitors now go first to B instead; some of this group stop their search at B, and therefore A gets fewer visitors:

a. (*business-stealing within location*) the incumbent gets a smaller share of the existing number of visitors to this location, and so reduces its product variety.

On the other hand visitors to A now have a better alternative option because there is a new competitor in close proximity (meaning lower costs of switching to the next seller). Therefore:

b. (*reduced search costs*) visitors are 'more choosy' about the varieties on display and so the incumbent raises its product variety.

Finally it seems reasonable to suppose that the entry of B will overall make this location a more desirable destination for shoppers, as the heightened competition feeds into lower prices and/or an increase in the *aggregate* variety of frames available here. In this case the entry of B may attract new visitors to this location from elsewhere in town:

c. (*business-diversion across locations*) the location gets more visitors overall, and so the incumbent raises its product variety.

The net effect of this change ('plus (b) plus (c) minus (a)') will depend upon the specifics of the market.[14] An empirical finding of reduced variety in response to new entry suggests that (a) dominates the combined effect of (b) and (c), and *vice versa.* If no variety response is apparent it may be that none of the above effects is of significant magnitude, or that the effects are significant but cancel each other out. As an alternative scenario suppose that B entered the market at some other location in town. Then there is no intra-location business-stealing, but visitors could be diverted to B's new location, meaning lower variety at A. Since there is a new seller in town those who do visit A may still get a lower cost of switching to a rival, which would mean greater variety at A. A finding of higher variety in response to new entry elsewhere would then suggest the dominance of this search-cost effect (b).

[14] B's entry also affects the variety at sellers elsewhere in town, which in turn may affect the search behaviour of visitors to A/B. I am treating such effects as second-order issues.

3 The entry model

In an important development in the literature on competition in differentiated products, Seim (2002) has introduced an empirical model of producer entry that admits considerably larger choice sets than have hitherto been tractable in structural frameworks. The key insight in her work is to model entry as a simultaneous, one-stage game of *incomplete* information among potential entrants. Empirical analyses based on games of complete information restrict agents to choosing from a few possible product types.[15] Complete information precludes bigger choice sets because of the need to check a large number of entry configurations in order to find the equilibria of the game. With incomplete information this problem is eliminated because a distribution of equilibrium outcomes can usually be found numerically as the fixed point of a contraction mapping.

Following Seim, let there be M distinct geographic markets, each partitioned into K_m cells, $m = 1, \ldots, M$. A firm entering any market m chooses one of the K_m cells therein as its business location. The post-entry realized profits of firm i locating in the k-th cell of market m are:

$$\Pi_{mki} = \mathbf{X}_{mk}\beta + f(\mathbf{C}_{mk}, \theta) + \xi_m + \epsilon_{mki} \tag{1}$$

Here $\mathbf{X}_{mk}$ is a vector of exogenous demographic and locational characteristics of cell mk; for example, median age of residents, population within one mile, number of shopping malls within one mile. Parameters in β capture the effects of such variables on firm profits. The terms ξ_m and ϵ_{mki} are profit components unobserved by the econometrician; the former is a component common to all entrants in market m, while the latter is specific to entrant i and cell mk. At present the model does not admit tract-specific profit effects that are common to all firms at a location but unobservable to the econometrician.[16]

The second term on the right-hand side of equation (1) is a general function capturing the effect of competition on firm i's profits, with parameters θ. Here $\mathbf{C}_{mk}$ is a vector classifying firm i's competitors in the market according to their distance from cell mk. If, for example, there are two distance classifications (or 'distance bands') then $\mathbf{C}_{mk} = (C^1_{mk}, C^2_{mk})$, where

[15] See Mazzeo (2002).

[16] Note also that in its present form the entry model does not allow for any observable differences (e.g., type of premises) between firms who locate in the same cell in a market. In principle it might be possible to allow for such differences by expanding the set of location options to distinguish different storefront types in the same cell. Whether this extension is econometrically tractable remains to be seen.

C^1_{mk} represents the number of nearby competitors (e.g., within one mile), and C^2_{mk} represents a more distant group (e.g., everyone more than a mile away).

Seim's original estimations make f linear in $\mathbf{C}_{mk}$:

$$f = C^1_{mk}\theta_1 + C^2_{mk}\theta_2 \ . \tag{2}$$

This specification[17] attributes the same profit effect to all rivals in a distance band, regardless of how many other rivals are as close or closer to firm i. I have found it necessary to consider a more general formulation which admits additional profit effects for the distance from the firm to its closest rival or rivals. A model with only two distance categories measures the distance to the closest rival just with a pair of dummy variables (in which $\mathcal{I}(.)$ is the indicator function):

a. $\mathcal{I}(C^1_{mk} > 0) \rightarrow$ 'closest rival is in band 1'

b. $\mathcal{I}(C^1_{mk} = 0, C^2_{mk} > 0) \rightarrow$ 'closest rival is in band 2'

Similar pairs of dummies are defined to indicate the location of the second and third-closest competitors. In each case one of the pair must be dropped from the regression to avoid collinearity. Furthermore I restrict the additional profit effect for each of these three closest competitors to be equal, leading to:

$$f = C^1_{mk}\theta_1 + C^2_{mk}\theta_2 + \theta_3\{\mathcal{I}(C^1_{mk} > 0) + \mathcal{I}(C^1_{mk} > 1) + \mathcal{I}(C^1_{mk} > 2)\} \ . \tag{3}$$

Consider the *additional* profit effect (relative to θ_1) that the first, second, and third-closest rivals would each exert if it were in band 1. Compare this with the additional profit effect (relative to θ_2) that these rivals exert if they are in band 2. Only the *difference* between these additional effects is identified: it is θ_3. Suppose for example that a firm that initially operates alone at a particular location in a market is subsequently joined there by other sellers, relocating from elsewhere in town. Each of the first three such relocators causes the incumbent's profits to change by $\theta_1 + \theta_3 - \theta_2$. The θ_3 effect appears here because each of these early relocators reduces the distance from the incumbent to one of its closest rivals. For any subsequent relocation the profit effect is just $\theta_1 - \theta_2$. If these nearby rivals were entirely new entrants to the market then the profit effects would be $\theta_1 + \theta_3$, and θ_1, respectively. Graphically the specification in (3) can be represented as in

[17] I use two distance bands, rather than the three in Seim's work.

Figure 3 (which supposes, for clarity, that $\theta_1 < 0$, $\theta_3 > 0$, $\theta_1 + \theta_3 < 0$ and $\theta_2 < 0$.) The solid dots represent the total profit effect of a firm's band-1 rivals ($= C^1_{mk}\theta_1 + \theta_3\{\mathcal{I}(C^1_{mk} > 0) + \mathcal{I}(C^1_{mk} > 1) + \mathcal{I}(C^1_{mk} > 2)\}$) while the hollow dots represent the effects of competitors in band 2 ($= C^2_{mk}\theta_2$). No kink appears in the curve for the band-2 profit effects because of the omitted dummies mentioned above.[18]

Firms in each market play a one-shot simultaneous entry game, in which each firm's strategy is a type-contingent choice of a cell mk, $k = 0, \ldots, K_m$.[19] Here the 'zero' cell represents a decision not to enter the market at all. A firm's privately observed type in this game is its vector of profit components $\epsilon_{mi} = (\epsilon_{m1i}, \ldots, \epsilon_{mKi})$, plus an idiosyncratic payoff to not entering, ϵ_{m0i}. Other firms $j \neq i$ only know a distribution $F(.)$ for $(\epsilon_{m0i}, \epsilon_{mi})$; this distribution is assumed to be known to the econometrician, and to be independent and identical across all players. Each potential entrant at market m knows that there are $N^{max}_m - 1$ other potential entrants here. The profit component ξ_m is observed *ex ante* by all potential entrants, and is assumed to be distributed independently of ϵ_{mi}.

Given its privately observed type, and knowledge of other elements of the model, each firm chooses the location that maximizes expected profits:[20]

$$\hat{\Pi}_{ki} = \mathbf{X}_k\beta + \hat{f}(\mathbf{C}_k, \theta) + \xi + \epsilon_{ki} \ , \tag{4}$$

where a hat denotes an expectation formed by a firm over the possible outcomes of $\mathbf{C}_k$. Expected profits for $k = 0$ are normalized to 0. Using (3) we get

$$\hat{f}_k = \hat{C}^1_k\theta_1 + \hat{C}^2_k\theta_2 + \theta_3\{\Pr(C^1_k > 0) + \Pr(C^1_k > 1) + \Pr(C^1_k > 2)\} \ . \tag{5}$$

Assume an equilibrium in symmetric strategies and an independent multivariate extreme-value distribution for F. Then any player's *ex ante* probability of a cell choice k, $k = 1, \ldots, K$, is:

$$\frac{\exp[\mathbf{X}_k\beta + \hat{f}_k + \xi]}{1 + \sum_{l=1,\ldots,K}\exp[\mathbf{X}_l\beta + \hat{f}_l + \xi]} \tag{6}$$

[18] Strictly speaking this curve should be represented as a function of two variables, C^1_{mk} and C^2_{mk}. In practice the omitted dummies mean that it is separable in C^1_{mk} and C^2_{mk}.

[19] As a check on the robustness of the results it would also be of interest to consider a *sequential* entry process. In this kind of process each firm, on making its entry decision, might observe the location choices of all previous entrants, but have incomplete information about the location-specific profits of subsequent firms in the entry sequence.

[20] Henceforth the subscript m will be dropped where convenient.

Each of these probabilities depends on $\hat{f}_k$, which in turn depends through (5) on the cell location probabilities of all the other players, which, by symmetry, are the same as (6). To evaluate $\hat{f}_k$ for each $k = 1, \ldots, K$ players need to know N^{max}. Given this number of potential entrants, it would in principle be possible to solve for the equilibrium probabilities of the whole game with a fixed-point procedure.[21] However as N^{max} is essentially not observable in the present context I follow Seim in adopting a simpler approach which partially separates the entry decision from the location decision and focuses on finding equilibrium location behaviour conditional on the *observed* number of entrants N_m in each market.

At a symmetric equilibrium players are now assumed to formulate $\hat{f}_k$ given a known number of actual entrants into the market. Given these expectations $\hat{f}_k$ any player's *ex ante* 'probability' of entry would be

$$\Pr(\text{in}) = \frac{\exp(\xi) \sum_{l=1,\ldots,K} \exp[\mathbf{X}_l \beta + \hat{f}_l]}{1 + \exp(\xi) \sum_{l=1,\ldots,K} \exp[\mathbf{X}_l \beta + \hat{f}_l]} \ . \tag{7}$$

It is assumed that ξ_m adjusts so that in each market

$$\Pr(\text{in}) = \frac{N_m}{N_m^{max}} \ . \tag{8}$$

Equation (8) implicitly defines N_m as a function of the other variables because the $\hat{f}_k$'s in the probability on the left-hand side are now conditioned on N_m. It can be seen through a re-arrangement of (7) and (8) that N_m is then determined by

$$\ln(N_m) - \ln(N_m^{max} - N_m) = \ln\left(\sum_{l=1,\ldots,K} \exp[\mathbf{X}_l \beta + \hat{f}_l] \right) + \xi_m \ . \tag{9}$$

Since N_m^{max} is not observed by the econometrician it is suggested that different values be tried, e.g., $N_m^{max} = 2N_m$ (implying $\Pr(in) = \frac{1}{2}$ in every market) or $N_m^{max} = 50$.

The simplification inherent in (8) and (9) estranges the statistical and game-theoretic structures somewhat. According to (9) agents will know the actual number of entrants N_m once they observe ξ_m and N_m^{max}. But if they make simultaneous entry decisions then they should only know a distribution for N_m, rather than its final realized value. In this respect equation (9) can be thought of as an *ad hoc* function determining the number of entrants,

[21] To do this ξ would have to be made a function of observable variables.

who then play a game of simultaneous location choices within the market conditional on this N_m. For the type-contingent location strategy of an entrant in this game write $s_i : R^K \to \{1, \ldots, K\}$. To find a symmetric equilibrium s^* for a given number of entrants we look for a set of probabilities $\mathbf{P}^* = (p_1^*, \ldots, p_K^*)$ such that

$$p_k^* = \frac{\exp[\mathbf{X}_k \beta + \hat{f}_k]}{\sum_{l=1,\ldots,K} \exp[\mathbf{X}_l \beta + \hat{f}_l]} \ . \tag{10}$$

Write Ω_k^d for the set of cells which are in band d relative to k, $d = 1, 2$, and write $q_k^d = \sum_{l \in \Omega_k^d} p_l^*$. To be consistent with (5) $\hat{f}_k$ must be such that

$$\begin{aligned} \hat{f}_k \quad = \quad & (N-1)(q_k^1 \theta_1 + q_k^2 \theta_2) + \theta_3 \Big\{ 3(1 - (q_k^2)^{N-1}) \\ & - 2(N-1) q_k^1 (q_k^2)^{N-2} \\ & - \frac{(N-1)(N-2)}{2} (q_k^1)^2 (q_k^2)^{N-3} \Big\} \ . \end{aligned} \tag{11}$$

To interpret equation (11) note that q_k^d is the probability that any given rival locates in band d relative to cell k, $d = 1, 2$. Then $\hat{C}_K^d$, the expected number of rivals in band d, is $(N-1) q_k^d$. The probability that at least one rival will be found in band 1 is

$$\Pr(C_k^1 > 0) = 1 - (q_k^2)^{N-1} \ , \tag{12}$$

while the probability that there will be at least two rivals in band 1 is

$$\begin{aligned} \Pr(C_k^1 > 1) \quad & = \quad \Pr(C_k^1 > 0) - \Pr(C_k^1 = 1) \\ & = \quad \Pr(C_k^1 > 0) - (N-1) q_k^1 (q_k^2)^{N-2} \ . \end{aligned} \tag{13}$$

For three or more rivals the probability is

$$\begin{aligned} \Pr(C_k^1 > 2) \quad & = \quad \Pr(C_k^1 > 1) - \Pr(C_k^1 = 2) \\ & = \quad \Pr(C_k^1 > 1) - \frac{(N-1)(N-2)}{2} (q_k^1)^2 (q_k^2)^{N-3} \ . \end{aligned} \tag{14}$$

The sum of terms (12), (13) and (14) gives the expression in braces in (11).

Equations (10) and (11) define a fixed point in the $K-1$ dimensional unit simplex which can be found numerically, given the data and values of the parameters. The fixed point shows an equilibrium distribution of entrants across locations in the market, assuming that they take the actual number of entrants as given when deriving their optimal type-contingent strategies. To complete the statistical model it is assumed that ξ is distributed $N(\mu, \sigma^2)$, independent of the ϵ's. The log- likelihood of the observed data is then

$$\ln L = \sum_m \sum_k N_{mk} \ln p^*_{mk}(N_m, \mathbf{X}) \\ - M \ln \sigma - \frac{1}{2} M \ln 2\pi - \sum_m \frac{1}{2\sigma^2}(\xi - \mu)^2 \ . \tag{15}$$

The first term on the right-hand side is the contribution from the observed distribution of cell locations conditional on N_m, as in (10) and (11), while the later terms represent the contribution from the variation in N_m across the different markets, as described in equation (9).

Note that any variables in $\mathbf{X}$ that are constant across all cells in a market drop out of the probabilities in (10). Their parameters are instead identified (up to scale) by the variation in N_m. Such variables could be used as instruments if, in extensions of the model, it proved possible to allow for stochastic dependence between ξ and ϵ_{mi}. If instruments are not available it would appear that independence of ξ and ϵ_{mi} would be necessary for identification.

The other foundations for the identification of the model concern the uniqueness of the number of entrants in each market, and of their equilibrium distribution in each market given $(N_m, \mathbf{X})$. Seim reports a numerical analysis of the fixed-point relationship (10) which suggests that uniqueness of $\mathbf{P}^*$ is fairly robust to different values for θ_1 and θ_2 in (2), given the amount of observed heterogeneity in the exogenous location characteristics in $\mathbf{X}$. A particular concern in the present analysis is whether these uniqueness properties break down with the introduction of the 'closest-rivals' parameter θ_3 in (3). In particular if θ_3 is positive (as the estimates below imply) we might get agglomeration effects in entry behaviour, with accompanying problems of multiple equilibria.

I first looked over the observed range of $\mathbf{X}$ for combinations of parameter values which yield multiple solutions to the fixed-point relationship in (10). The results indicate that it is at the least necessary to assume $\theta_1 \leq \theta_2$ to guarantee uniqueness of $\mathbf{P}^*$. Given this restriction no further multiple fixed points were found even when θ_3 is set at its estimated (significantly positive) level of 0.8.[22] Thus identification of the model requires that the baseline effect of competitors on profits is more deleterious when they are

[22]Some multiple equilibria were found when $\theta_3 = 0.8$ and $\theta_1 = \theta_2 = 0$. However the difference between any pair of probabilities in these multiple equilibria was always less than 0.002, meaning that they have no practical impact on the estimation results.

closer to the incumbent, which seems intuitively reasonable. However given this restriction uniqueness of the solution to (10) may still be sustained even if the opposite relationship holds (presumably up to some maximum value for θ_3) for the first few nearby rivals.

It is harder to guarantee uniqueness of the number of entrants implied by any N_m, $\mathbf{X}$, and ξ_m in (9). Note that the first term on the right-hand side of (9) is an 'inclusive value': it can be interpreted as the expected maximum profit which an entrant would realize if it first decided whether or not to enter, given $N_m - 1$ rivals, and only then observed ϵ_{mi} and chose its preferred location. This term is not in general monotonically decreasing in N_m at the estimated parameter values (apparently because of the positive estimate for θ_3). To ensure that the observed N_m is the only value consistent with (9) it was consequently necessary to adopt the 'fixed N^{max}' version of equation (8) (setting $N^{max} = 30$), rather than the suggested alternative which sets $\Pr(\text{in}) = \frac{1}{2}$.

Further uniqueness problems arose in (9) arose when positive estimates were obtained for θ_2 under some specifications. A positive value for θ_2 essentially means that the baseline effect of an extra competitor in the market is to raise everyone's profits, rather than reduce them. Not surprisingly such profit functions give rise to multiple values for N_m.[23] To counter these effects I interpreted a positive estimate of θ_2 as implying a true value of zero, and repeated the estimation with the restriction $\theta_2 = 0$: given a fixed N^{max} the resulting N_m predicted by (9) was then unique in each market.

It is undesirable that the coherency of the model should depend on such restrictions and on a particular specification of $\Pr(\text{in})$ in (8). The fixed-N^{max} version of (8) would be appropriate for cases in which the same set of chain stores comprised the potential entrants in every market. Certainly chain stores are an important presence in eyewear retailing. However the majority of sellers are not affiliated with nationwide chains, which might argue in favor of a model where the set of potential entrants varies with market size. As a check on the robustness of the results to changes in the rule determining N_m I also run a 'limited-information' version of (15). To motivate this version note that the assumption of independence of ξ_m and ϵ_{mi} implies that the selection of N_m via (9) is exogenous with respect to the location choices of the actual entrants. Hence one could just maximize the likelihood of firms' location choices within each market, conditional on the

[23]Note that a positive θ_2 can create multiple equilibria even when $\theta_3 = 0$, although the problem is exacerbated when $\theta_3 > 0$.

observed N_m, i.e.,

$$\ln L = \sum_m \sum_k N_{mk} \ln p^*_{mk}(N_m, \mathbf{X}) \ , \tag{16}$$

where the p^*_{mk} are defined by (10). Under the independence assumption this procedure yields consistent estimates of those parameters in β which are identified by the within-market variation in the data. These estimates can then be compared with the results from the full-information likelihood (15).[24]

4 Competition in product variety

The second period of the model analyzes the competition in product variety between rival eyewear sellers in each market. Theory suggests that both the total number of entrants and their locations in the market are likely to be important factors in this competition. Since these factors are determined in the entry process, correlation between the unobservables across periods could lead to a problem of endogenous market configurations in the second period. I allow for this possibility using corrections for sample selection similar to those introduced to the entry literature by Mazzeo (2000).

Let V_{mki} denote the equilibrium number of varieties of a good sold by the i-th firm when it is located in cell k of market m. In the present study V_{mki} is measured by a simple count of the number of spectacle frames on display at each eyewear seller. It is assumed that the logarithm of V_{mki} is determined by the following reduced-form relationship, conditional on the earlier location choices of N_m observed entrants:

$$\ln V_{mki} = \mathbf{Z}_{mk}\alpha + g(\mathbf{C}_{mk}, \psi) + \omega_{mki} \ , \tag{17}$$

for $m = 1, \ldots, M$, $k = 1, \ldots, K_m$, $i = 1, \ldots, N_m$. Here $\mathbf{Z}_{mk}$ is a vector of exogenous location characteristics for firm i's current cell mk, comprising some or all of the characteristics in $\mathbf{X}_{mk}$. The effects of competition on variety are captured by g, which has the same form as the function f in (3). Unobserved firm-specific determinants of equilibrium variety are captured by the random term ω_{mki}.

Note that the characteristics in $\mathbf{Z}_{mk}$ do not include any variables that distinguish sellers within a cell, e.g., chain affiliation, or location inside or

[24] Of course the assumption of independence of N_m and ϵ_{mi} could be criticized, but relaxing this assumption is a separate issue from the consideration of different functional forms for the determination of N_m in (9).

outside a particular department store or shopping center. Such specific characteristics are ignored in (17) because they were not modeled at the entry stage – including them now could lead to a selection bias. In future work it may be possible (e.g., with nested-logit error structures) to incorporate chains, or detailed storefront characteristics, into the Seim entry model. For the present I maintain as a necessary fiction the assumption that potential entrants are all symmetric *ex ante*, and symmetric conditional on cell location *ex post*.

As noted, ω_{mki} in (17) could be correlated with the ξ_m and $\{\epsilon_{mi}\}_{i=1,\ldots,N_m}$ that determined the configuration of a firm's rivals in the first period. To account for this I restrict ω_{mki} to be correlated only with ξ_m and ϵ_{mki}. Any other correlation is assumed to be zero. In particular a firm i's ω_{mki} is independent of the ϵ_{mli} for any other cell l, $l \neq k$. It is also independent of the ϵ's for any other firm.

These assumptions imply that, conditional on ξ_m, ω_{mki} is uncorrelated with the location $s_j^*(\epsilon_{mj})$ chosen by any other firm in the first period. To see this note that ξ_m determines N_m and therefore s^*. Given s^*, firms' location choices, determined by s^* and their respective draws ϵ_{mj}, $j = 1, \ldots, N_m$, are by assumption uncorrelated. This motivates the following endogeneity correction.

Let A_{ki} be the set of ϵ_i such that $s_i^*(\epsilon_i) = k$, i.e., such that firm i chose cell k given ξ. Rewrite (17) as:[25]

$$\ln V_{ki} = \mathbf{Z}_k\alpha + g(\mathbf{C}_k, \psi) + \mathrm{E}[\omega_{ki} \mid \xi,\ \epsilon_i \in A_{ki}] + \eta_{ki}\ , \qquad (18)$$

where

$$\eta_{ki} \equiv \omega_{ki} - \mathrm{E}[\omega_{ki} \mid \xi,\ \epsilon_i \in A_{ki}]$$

and so

$$\mathrm{E}[\eta_{ki} \mid \mathbf{C}_k] = \mathrm{E}[\eta_{ki} \mid \xi,\ \{\epsilon_i \in A_{ki}\}_{i=1,\ldots,N}] = 0.$$

I assume that ω_{ki} is distributed independently across firms conditional on ξ. For the conditional expectation in (18) write

$$\begin{aligned} \mathrm{E}[\omega_{ki} \mid \xi,\ \epsilon_i \in A_{ki}] &= \mathrm{E}[\mathrm{E}[\omega_{ki} \mid \xi, \epsilon_i] \mid \xi,\ \epsilon_i \in A_{ki}] \\ &= \mathrm{E}[\mathrm{E}[\omega_{ki} \mid \xi, \epsilon_{ki}] \mid \xi,\ \epsilon_i \in A_{ki}]\ , \end{aligned}$$

where the second equality follows from the fact that ω_{ki} and ϵ_{li} are uncorrelated, $l \neq k$. Adopting a linear specification for the inner expectation we

[25]Market subscripts are dropped again and the conditioning on $\mathbf{Z}_k$ is left implicit.

get

$$\begin{aligned} \mathrm{E}[\omega_{ki} \mid \xi,\ \epsilon_i \in A_{ki}] &= \mathrm{E}[a\epsilon_{ki} + b\xi \mid \xi,\ \epsilon_i \in A_{ki}] \\ &= a\mathrm{E}[\epsilon_{ki} \mid \xi,\ \epsilon_i \in A_{ki}] + b\xi\ , \end{aligned} \tag{19}$$

where a and b are parameters. The conditional expectation for the error in a multinomial logit model is known to be:[26]

$$\mathrm{E}[\epsilon_{ki} \mid \xi,\ \epsilon_i \in A_{ki}] = \gamma - \ln\, p_k^*\ ,$$

where $\gamma \approx 0.577215$. After substitution for (19) equation (18) then becomes

$$\ln V_{ki} = \mathbf{Z}_k \alpha + g(\mathbf{C}_k, \psi) + a(\gamma - \ln\, p_k^*) + b\xi + \eta_{ki}\ . \tag{20}$$

To operationalize (20) I use the consistent estimates of ξ and p_k^* obtained from the first-period estimation. By construction the error η in (20) is mean-independent of the explanatory variables.

5 Data

The area of this study covers 44 medium-sized geographic markets in six Midwestern states.[27] Similar criteria to those in Seim (2002) are used to define the set of markets. A town or group of towns was initially included in the sample if it comprised a continuous built-up area with total population in the range 25,000–200,000, located entirely within (one or more of) the six states. I dropped a market from this set if any one of its principal business centers was less than 20 miles from a business center in a separate built-up area of population 25,000 or greater. Thus the sample excludes markets close to big metropolises in favour of regional centers at least 20 miles from the next major town.[28]

Of the six states in the sample only Ohio, with nine markets, mandates the licensing of opticians. This interstate variation in professional regulation is captured by a dummy for Ohio. Intuitively we expect that the licensing of opticians should increase the cost of a labor input essential to the sale of eyeglasses, and therefore reduce the profitability of entry for the business owner.

[26] See, e.g., Dubin and McFadden (1984).

[27] Illinois, Indiana, Iowa, Michigan, Minnesota, Ohio. These states were chosen for convenience of access, rather than as a random sample. The econometric inferences below are confined to the behaviour of businesses in these states only.

[28] The Rand McNally Marketing Atlas (.) was used to define built-up areas and locate business centers.

Sellers of eyeglasses within the sample markets were initially located from listings in telephone directories.[29] All businesses in these listings were then telephoned to confirm location and current operation. This survey also covered the practices of ophthalmology MD's, which sometimes have eyewear shops.

There are 572 sellers in the whole sample. Table 1 breaks this total down by category of outlet. A majority of sellers (467) operate out of premises dedicated to eyecare, i.e., in their own shopfront or professional office. Within this group of specialist sellers there is variation in the qualifications of the eyecare provider. About 19% of this group (or 15% of all sellers) are in the offices of ophthalmologists – MD's who diagnose and treat any kind of eye condition. In the towns under study glasses are sold by 77% of ophthalmology offices (excluding some retinal specialists who don't do examinations for eyewear prescriptions). Most specialist sellers provide eyecare through an optometrist, who diagnoses eye conditions and tests eye function. All optometrist's offices sell spectacles. Third, in a small number of outlets the principal eyecare provider is an optician, who just fills prescriptions for eyeglasses written by optometrists and ophthalmologists. They comprise less than 5% of all sellers.

These specialist eyecare businesses can be classified not only by qualifications but also by the nature of their chain affiliation – see Table 2. About 20% of specialist sellers operate under the name of a wide-area chain, defined here as a brand with a presence in four or more markets in the sample. The more ubiquitous chain stores (e.g., Lenscrafters, Pearle Vision) are affiliated with upstream manufacturers, either as franchises or through direct ownership, and (in this sample at least) have a strong tendency to locate in or next to shopping malls. In markets of the sizes under consideration these large eyecare chains typically do not have more than one outlet in any market. However multiple branches per market are sometimes observed amongst more localized operations. Overall 45 out of the 467 specialist sellers are 'multiples', outlets sharing a market with another branch of the same trade name. A further four (i.e., two pairs) of the 467 share a market with a sister specialist outlet operating under a different identity. (That is, an eyewear company owns specialist outlets with different brand names, and the brands' geographic territories overlap slightly.)

In addition to the eyecare specialists there is a minority (18%) of sellers who are located inside large department stores (or 'discount' stores). Six such stores have eyewear shops in some or all of their outlets in the study

[29] American Business Disk, also www.yahoo.yp.com.

area; all of these offer eye exams through an optometrist.[30] The management arrangements of these in-store outlets can be complicated. A host department store might operate some of its optical shops directly and lease others to outside contractors.[31] In one case such an outside contractor combines department-store leasing with a substantial presence in the specialist eyewear retail market. Cole National Corporation, the owner of the Pearle Vision specialist retail brand, also runs most of the optical shops in Sears and Target stores.[32] To further complicate matters a small number (about 5%) of the Sears shops are managed by a different contractor, which also happens to have most of the optical leases in JC Penney stores. On the other hand Walmart and ShopKo directly operate all or most of their eyewear shops.[33]

These summaries indicate considerable heterogeneity amongst eyewear sellers, not all of which can be accounted for in the entry model. The location of any department store is treated as an exogenous cell characteristic. This seems reasonable given the range of goods unrelated to eyewear that are sold in such emporiums. Moreover for some of these department stores the locations of their optical shops are also held to be exogenously fixed. Four of the six retail emporiums selling spectacles have optical shops in almost all their outlets in the sample.[34] For these operations the decision to sell eyeglasses seems to be made (perhaps in consultation with contractors) on a company-wide basis, rather than store-by-store. Since the information on store locations reveals almost all the information on the sites of their optical shops, I simply treat the number of such stores in each cell as an exogenous characteristic that may influence the entry decisions of other sellers.

Optical shops in the remaining two department stores[35] only appear in about 46% of their outlets. Their location decisions are therefore treated endogenously along with those of the 467 specialist sellers. I regard these 31 optical shops as separable from the other activities of the two department

[30] Sears, JC Penney, ShopKo, Super Target, and two kinds of Walmart – ordinary Walmart and Walmart Supercenter. Ordinary Targets have eyewear shops in larger cities, but not in markets in this sample.

[31] I am grateful to David Prentice for bringing this phenomenon to my attention. See Prentice and Sibly (1996).

[32] There are 19 Pearle Vision outlets which share a market in the sample with a Sears optical shop. In ten of these cases the Pearle outlet and the Sears store are in the same mall. Below I hold entry decisions for Sears optical shops to be exogenously fixed, so there is no joint modeling of the Sears/Pearle entry decision.

[33] See www.2020mag.com, April 2001 edition.

[34] Sears, Walmart Supercenter, Super Target, ShopKo, hereafter referred to as the 'SWTS' group. Overall 95% of these stores' outlets in the sample have optical shops.

[35] Ordinary Walmarts and JC Penney.

stores, run by freely floating entrepeneurs for whom the store is just one of a number of possible locations in town. The presence of one of these department stores is assumed to exert the same profit influence on all of the endogenous sellers who eventually locate in that cell. This is somewhat unrealistic – at most one of the sellers will be in the store; the others will be outside. However the small number of sellers in these two stores suggests that the bias due to this simplification should not be large.

I make no attempt to model sellers' choices of the number of outlets to establish in each market. Instead the location of each endogenous seller is assumed to result from a profit-maximization problem that is entirely separate from the problem of any other outlet. In view of the relatively minor position of multiple outlets in the set of endogenous sellers (about 10% of the total) it is hoped that this abstraction does not bias the final results too much. Developing a Seim-type entry model that explicitly allows for multiple outlets per market is left as an interesting topic for future work.

Since independent opticians are few in number I regard them as identical to optometrists for estimation purposes. Ophthalmologists are treated in a similar manner to the endogenous department-store sellers. Locations of ophthalmology practices are assumed to be exogenously fixed, but their optical shops are thought of as separable activities whose presence is jointly determined with the locations of other eyewear sellers. This could be an appropriate specification if, for example, ophthalmology locations are mostly affected by the incidence of ocular ailments unrelated to eyeglasses (and unrelated to the ξ's and ϵ's in the model). Attached optical shops may be run by independent managers who make their own entry and location decisions. A variable for the number of MD's in a cell shows the effect of proximity to an ophthamologist on the profits of an eyewear seller. As a simplification the profit effect is again restricted to be the same for all sellers in that cell.

As in Seim's work, the Census' partition of counties into tracts was used to define the set of possible business locations ('cells') in each market. Census tracts are non-overlapping irregular polygons defined to correspond roughly to a neighbourhood or locale. Each tract usually contains three to five thousand inhabitants. The position of each tract is summarized by its population-weighted centroid, and a tract counts in the set of possible locations if this centroid is within ten miles of the geographic center of the market's main town or towns. Using this criterion each market contains 22 cells on average. Locations of eyeglass sellers and other relevant sites were mapped into tracts using the Census' TIGER/Line files.[36] For simplicity all stores and other entities in a tract are assumed to be located at the centroid.

Tract-level demographic and economic variables from the Census help explain the profitability of one location relative to another. Differences across tracts in the number, median age, median income, and proportion-in-college of residents could create local variations in eyeglass demand. On the cost side I use the median household rent as a proxy for the cost of retail space in a tract.

Another key determinant of a tract's relative profitability is the density of non-optical businesses in the area. High business density creates a steady flow of potential customers. Seim uses a business-count variable sold by a private- sector vendor to capture this effect. Since this variable is not available to me I use a variety of other measures. As broad indicators of prime retail location I mapped the sites of large shopping malls and shopping centers into Census tracts.[37] To this data I added the positions of all outlets of any department store which is observed to sell eyeglasses somewhere in the sample. Sites of smaller retail centers such as strip malls are harder to locate. Instead I plotted the locations of outlets in each market of four nationwide fast-food chains.[38] The total number of such outlets in each tract is used as a proxy for the density of smaller retail businesses. Variations in this retail index across tracts may also reflect local zoning restrictions, which are otherwise unobserved in the data.

Table 3 lists all the markets in the study area and Table 4 summarizes some market-level characteristics. Summary statistics on explanatory variables for the entry regression are shown in Table 5. Not all tract-level demographics for the 2000 Census had been released at the time of writing. Hence the analysis currently uses 1990 demographics and tract boundaries, with predicted values for 2000 where these are available. As a market-level indicator of a town's general growth prospects I use a weighted average of the total value of annual private housing starts in counties covering that market.[39] More housing construction in a town (for a given population) points to greater future demand and would presumably attract more retailers.

Over the summer of 2001 the author visited each market to conduct an in-person survey of the variety of spectacle frames stocked by sellers. This survey concentrated on obtaining inventory data from a randomly chosen 50% of the sellers in each market. In the 11 Illinois markets the survey

[36] www.census.gov

[37] These were found in National Research Bureau (1998).

[38] Arby's, Burger King, McDonald's, Wendy's.

[39] City and County Data Book, 2000.

aimed for 100% coverage, resulting in a target subsample of 342 out of 572 total sellers. The measure of product variety at each of these businesses is a simple count (usually done by the author) of the number of different frames on display for adult prescription lenses (excluding sun glasses and safety glasses). About 5% of the observations in the frames subsample are missing due to seller non- response. At any such outlet the number of frames is imputed to be the average for similar sellers in the same town.

Tables 6 and 7, and Figure 4, summarize the numbers of eyeglass frames stocked by each kind of seller in the subsample. Specialist eyecare chain stores have the broadest product variety, while department stores typically have low variety. The main distinction in specialists' qualifications, between optometrists and ophthalmologists, does not appear to produce major differences in the variety distribution. Figure 4 suggests that overall the distribution of frame counts is approximately log-normal.

Department-store sellers exhibit not only low variety in absolute terms, but also low variation in display inventories from store to store. The interquartile range in the product variety of department stores as a group is about 27% of their median, compared with 60% for the specialist eyecare outlets. Moreover department-store inventories are even more tightly distributed when viewed at the individual brand level: the two most widespread sellers, Sears and Walmart Supercenter, respectively have proportional interquartile ranges of 20% and 10%. This suggests that department stores, whether for reasons of marketing or cost control, value uniformity of content in their optical shops.

Summary statistics for the explanatory variables in the variety regression are shown in Table 10. The exogenous variables in this table are a subset of those for the entry regression, but here they are averaged over sellers (rather than tracts), which is the unit of observation in the variety model. Statistics are also shown for the variables representing the effects of competition. Sellers on average face 2.4 (non-SWTS) rival firms within one mile. About 80% of sellers have at least one such rival in band 1; 60% have at least two rivals, and 40% have three or more. (The maximum number of non-SWTS rivals in band 1 is seven.)

In addition to prescription spectacles eyewear shops also sell eye examinations, contact lenses, and non-prescription sunglasses (plus other less significant items like non-prescription reading glasses). I have little information about sales of these goods.[40] To abstract away from competition in eye exams I assume that these are sold in fixed proportions to eyeglasses, with

[40] I only know the price each seller charges for an eye examination.

no variation in quality across sellers. The latter assumption is somewhat restrictive, although the former may not be far from the truth. Competition in sales of contact lenses and sunglasses is similarly ignored. Contact lenses might be regarded as a homogeneous good sold under conditions of perfect competition, perhaps reflecting the fact that this good can also be purchased through the Internet. Sunglasses were specifically excluded from the inventory survey because they are carried not just by eyecare specialists but also by kiosks in malls, clothing shops, sellers of sporting goods, and so on. Accounting for (and locating) this variety of retail outlets would complicate the analysis considerably.

6 Results

Table 8 shows estimation results for the first-period entry model, in which the location decisions of the SWTS sellers (i.e., four of the six department stores) in each market are treated as fixed. The likelihood function in (15) is maximized by a numerical optimization method.[41] There are two distance bands, with an inter-band cutoff of one mile. Competition effects are modeled as in (3), with θ_2 set to zero to ensure a unique prediction for the number of entrants in every market. In equation (9) the number N^{max} of potential entrants in each market is set to 30. A dummy for JC Penney is dropped because it is closely correlated with the dummy for the presence of an enclosed shopping mall. Dummies for each state are included as market-level variables to allow for variations in the regulation of eyecare providers – the excluded states are Iowa and Minnesota.[42]

By and large the tract-level characteristics have the expected profit effects. A location's profitability is significantly higher if it has an open-air shopping plaza, a Walmart store, or an ophthalmology practice. The profitability of a location with an enclosed shopping mall is increasing in the area of that mall. Tracts with older residents are more desirable business locations, perhaps reflecting greater demand for eyecare. Proximity to fast-food outlets also significantly raises profits, suggesting that this is a valid proxy for business density. Profitability falls with higher rents in a tract, but this effect is not statistically significant. Population also does not significantly affect profits, suggesting that the other tract-level characteristics

[41]Some computational notes on the optimization are contained in an Appendix.

[42]The small number of markets in Minnesota does not justify a separate dummy for this state.

are adequate controls for the number of potential consumers in a market.[43] Nor are there significant effects for the number of shopping malls elsewhere in town, or for the market-level variables.

Estimates of the competition effects indicate that the number and proximity of a firm's rivals is an important determinant of profitability. Both θ_1 and θ_3 are significantly different from zero, suggesting that, as in Seim's findings, the effects of competition weaken at greater distances. However the direction of these effects depends on the number of competitors already in operation near the incumbent firm. Suppose that a firm initially has no competitors within one mile, but is then successively joined at its location by several new entrants. Its profits are estimated to increase in response to the first three such new entrants, and to fall with subsequent entrants. The exogenous SWTS group of competitors enters the objective functions of the active players as a count of the number of these sellers in each distance band relative to the given tract.[44] Parameters on these variables repeat the pattern of competition effects observed for the endogenous sellers, but without statistical significance.

The contrasting impacts of initial and subsequent nearby rivals are somewhat surprising. Positive competition effects may be evidence of genuine complementarities, or clustering effects, in firm location. However the magnitude of the initial-nearby-rival effect is suspiciously large – it is equal to $\hat{\theta}_3 + \hat{\theta}_1 \approx 0.34$, equivalent to the presence of a small shopping mall for each such competitor. This may reflect misspecification of the model. In particular it may indicate that both the incumbent and the early nearby rivals are responding to some common tract characteristic that is unobserved by the econometrician. That is, the unobservables in ϵ_i could be correlated across firms, contrary to the assumptions above. To allow for this possibility it would be desirable to introduce random effects that are specific to tracts, but common to all firms in a tract. As noted earlier this might require more complicated econometric techniques.

In Table 8 the baseline competition effect for rivals in band 2 was set to zero to rule out positive θ_2 estimates, which lead to multiple equilibria for the total number of entrants in the market.[45] Even with this restriction

[43] I also tried as explanatory variables tract-level measures of the number of hospitals, number of college students, and residents' income but found them to be insignificant.

[44] Because of the small numbers in this group in each market only the closest SWTS rival is distinguished by a dummy variable.

[45] It appears that the regression in Table 8 tries to compensate for the restriction $\theta_2 = 0$ by exaggerating the magnitudes of the state dummies somewhat (although none of them become statistically significant).

the model would only admit the fixed-N^{max} specification of equation (8). It is possible that these restrictions bias the parameter estimates. Table 9 presents a limited-information version of the model which, as discussed in Section 3, provides a cross- check on these estimates. Assuming independence of ξ_m and ϵ_{mi}, this model gives consistent parameter estimates for those variables which are not constant across all tracts in a market. Thus, for example, θ_1 and θ_2 are not separately identified in this 'within-market' model because the total number of rivals is the same at all tracts in a market. Instead only $\theta_1 - \theta_2$ is identified, shown just as θ_1 in the table. Parameters that were statistically significant in Table 8 are also significant in Table 9, and have similar magnitudes.[46] Only the parameter on median rents is newly significant in the limited-information regression; it is negative as expected. These results suggest that the rather special assumptions used to identify the full entry model in Table 8 do not markedly bias the tract-level parameter estimates. In particular the same pattern of competition effects is seen in both regressions.

The principal conclusion on the basis of this entry model is that geographic differentiation matters in competition amongst eyewear sellers. Rival outlets show significantly different baseline effects on profits if they are close by rather than far away. The additional profit effects for a firm's three closest rivals are significantly different if these firms are in band 1 rather than band 2. Not surprisingly, for sufficient nearby competition the profit effects are eventually negative, although we cannot reject the hypothesis that there is some positive externality, perhaps due to clustering, from the first few competitors.

Table 11 shows OLS regression results for the second-period variety equation (20). Tract and market characteristics are as in the model of location choice – some of these independent variables showed little significance and were dropped. Corrections for the endogeneity of market structure enter through ξ and the logit correction $\gamma - \ln p_k^*$, derived from the entry model. The competition effects $f(.)$ in equation (20) are specified as in (3), with ψ replacing θ.

As noted in Table 6, department stores show relatively little variation in product variety across outlets. This uniformity of content is consistent with the uniformity across markets in entry behaviour seen previously for four of the six such stores. Since variety decisions for these sellers seem to be fixed at the corporate level, the sample for the variety regressions only includes

[46] A Hausman-type test for specification error would be appropriate here. This will appear in a later draft.

the non-SWTS firms. Competition effects in $f(.)$ are calculated with respect to these endogenous sellers. (Two department stores which were treated as endogenous at the entry stage are treated as endogenous here too.) The numbers of SWTS sellers in bands 1 and 2 enter the variety regressions as exogenous variables, with a dummy for the effect of the closest such seller. In Illinois markets there is a product-range observation for all sellers in the endogenous category; in markets elsewhere the coverage is a random 50%. A dummy for Illinois accounts for any bias due to the oversampling in this state.

To interpret the effects of competition in this model consider three exogenous changes in market configuration. First, a new entrant to the market may commence operation close to the current firm. Alternatively the new entrant could set up at a more distant location in the market. Lastly a new nearby rival could be an existing seller relocating from a more distant location. The outcomes of each of these experiments may depend on how many rivals are already operating close to the seller in question. In Section 2 it was suggested that the direction of an incumbent's variety response to such changes would depend on the following underlying effects (indicated along with their signs):

a. (*new entry nearby*) in-tract business stealing ($-$), consumer search costs ($+$), cross-tract business diversion ($+$)

b. (*new entry far away*) in-tract business stealing (0), consumer search costs ($+$), cross-tract business diversion ($-$)

c. (*relocation from far away*) in-tract business stealing ($-$), consumer search costs ($+$), cross-tract business diversion ($+$)

Note that experiment (c) is equal to (a) minus (b): somebody leaves a far location and joins a nearby location. Statistically it is not redundant to consider (c) as a separate experiment because the significance of the variety response will depend on the covariance between the parameters representing experiments (a) and (b).

As a group the independent variables in the regression are significant at the 1% level. Shopping malls appear to significantly raise sellers' product variety, *ceteris paribus*. Sellers near a medium-sized mall of 500,000 square feet would have about 13% more variety than if they were near a small mall of 200,000 square feet. (With a p-value of 0.102 this effect is more or less statistically significant.) A more distant mall is still estimated to raise a seller's variety by about 12%. The dummy for Ohio is significantly negative, perhaps reflecting the regulation of opticians in that state. Most of the other exogenous variables which affect profits significantly in the entry model do not have noticeable effects here.

The baseline impact of new entry nearby on an incumbent seller's product variety is a statistically significant reduction of 8% per entrant. New entry further away is also estimated to reduce variety, albeit without significance. If an entrant is one of the first three rivals to locate near the incumbent then the negative effect on product ranges is ameliorated. Entry by these initial rivals reduces the distance from the incumbent to its closest competitors: this counters the baseline negative effects of new entry with a positive variety impact of 10%. We can reject with 90% confidence the null hypothesis that each of the first three proximate rivals has the same effect on product range as any subsequent such competitor.

Consider then an experiment of type (c): an initially isolated incumbent faces a succession of (non-SWTS) rivals relocating close by. Each of the first three such relocations induces a positive variety response of $\theta_3+\theta_1-\theta_2 \approx 4\%$ from the incumbent. With a p-value of 0.17 the statistical significance of this effect is marginal.[47] But for any subsequent relocation we get a negative response of $\theta_1 - \theta_2 \approx -6\%$, which is significant at the 10% level. That is, we can reject with 90% confidence the null hypothesis that the baseline effect of competition on product variety is unrelated to the distance between firms.

Variety responses to sellers in the SWTS group show a similar pattern of signs, albeit with less statistical significance. The presence of a single SWTS rival (an entire department store) in the neighbourhood of a seller is estimated to raise variety by 17%, with a p-value of 0.11. We can reject with 90% confidence the hypothesis that the first such competitor has the same effect on variety as each subsequent SWTS seller in the neighbourhood. Distant rivals in this group appear to have a slightly positive effect on variety, but this effect has no statistical significance.

Neither of the parameters for the endogeneity corrections in this regression is statistically significant. If the corrections are omitted the estimated magnitudes of the competition effects are essentially unchanged, but with smaller standard errors. The parameter θ_2 now becomes significant at the 1% level, while the t-statistics for the other effects discussed above also tend to increase substantially. In general omitting the endogeneity corrections would not seem to alter the discussion greatly, which suggests that the number of firms in each market, and their locations, are well explained by

[47] This effect had somewhat more significance in an earlier version of the paper, due to the inclusion there of an outlier with very low product variety: a part-time seller.

the observable explanatory variables in the model.

Which of the three effects of competition on product variety discussed in Section 2 can be seen in these results? First, more competition at any location may attract business from other points in the product space. This tentative conclusion follows from the weakly negative sign on θ_2 (which means a weakly negative response to experiment (b) above). But the significantly negative value for θ_1 (representing experiment (a), with several other competitors already present nearby) indicates that an extra firm at a site also steals business from the other sellers there. Both of these effects are consistent with store-level variety choices that depend on the number of consumers visiting the outlet.

A third influence on variety, that of consumer search costs, is potentially more interesting but cannot be separately identified from these estimates. We have seen that relocation experiments of type (c) may yield positive variety responses when few other firms are present at a location. This points to the dominant influence of some combination of lower search costs for consumers and the diversion of business from other locations. Together these effects might outweigh the negative impact on variety of business stealing within the location. Unfortunately, the contribution of search costs to this combination is unknown. In particular even if greater competition at a location has no directly positive effect on variety (through lower search costs) it might still produce lower prices, which would be enough to explain the attraction of some customers from elsewhere in town. It would be possible to identify the search-cost effect in a structural model that took consumer preferences over the space of product characteristics as primitives and explicitly modeled consumers' travel costs.[48] Such an approach would be complicated in the present instance by the absence of data on the prices of eyeglasses or quantities sold at each firm.

Cross-tract business diversion effects are of interest in themselves because they dilute the incentives of firms to differentiate their products. In its strongest form this kind of effect could lead firms to tolerate a certain amount of clustering in equilibrium, instead of each developing de facto monopoly power in an isolated region of the product space. Of course some grouping of sellers is seen in the data because of the desirability of particular locations like shopping malls. Clustering might benefit firms even in the absence of special locational characteristics if it served to concentrate more eyewear customers at a location. The entry-model estimation suggested that a firm's

[48]See Davis (2001), Manuszak (2000) and Thomadsen (1999) for structural models of interfirm competition in geographically differentiated markets.

first few nearby competitors did indeed produce a significantly less negative profit effect than any later rivals. This is consistent with a limited amount of clustering behaviour, but, as previously noted, these profit responses could be a reflection of misspecified tract-level error distributions. Further work on this issue would seem to be warranted. Note that both the entry and variety models indicate that any clustering behaviour is limited to small groups of up to four firms - entry beyond that level has clearly negative effects on profits and product ranges.

As noted previously, the product-range model used here cannot accurately measure the role of chain affiliation in variety choices because chain identity was not modeled at the entry stage. Nevertheless Table 7 suggests that sizeable chain effects might be observed in a more general model which endogenizes this aspect of seller's brand identities. For illustrative purposes Table 12 shows the results of a regression which adds dummies for chain affiliation to the model in Table 11. One dummy is for outlets of Lenscrafters or Pearle Vision, of which there are 20 in the sample of 301 variety observations. A second dummy is for the 41 other chain-affiliated outlets – where a chain is defined as an operation with stores in at least two different markets in the sample.

The results in the table suggest that brand identity is likely an important explanator of a seller's product variety. All else equal, Lenscrafters or Pearle Vision outlets are estimated to have about 70% greater product variety, while for other chains the differential is about 20%. These parameter estimates could be biased if chain identity is endogenous with unobserved features of firms' locations (as seems likely). Note however that adding the chain dummies makes little difference to the estimated magnitudes and significances of the competition effects. Only the dummy for the first nearby SWTS rival loses some significance. It appears that the preceding inferences on competition and product variety might still hold even in a model which controls for sellers' chain affiliation. Amongst the parameters for the locational characteristics it is interesting that the Ohio dummy is considerably less negative when chain dummies are included. This suggests that lower per- seller variety in that state reflects in part fewer chain stores.

7 Conclusions

Estimation results in this paper suggest first that geographic differentiation matters in competition between eyewear sellers. In the entry model the most significant effects on firm profits accrue from competitors in close proximity to a seller. This accords with Seim's (2002) findings for the video rental industry. Distance effects were also observed in the variety regressions, in the predicted impacts of moving rival sellers closer to a given firm.

These differentiation effects are non-uniform in that their directions depend on the current configuration of sellers in a market. The baseline effect is for a new entrant (or a relocating firm) to steal business and profits from nearby sellers, who therefore reduce their product range. However a relocating firm (or group of firms) which joins with a previously isolated seller may cause that incumbent to raise product variety. This seems to reflect at least in part the attraction of customers away from other locations in the market. In the entry model we find that the incumbent's profits are significantly *less* negatively affected by these early nearby rivals.

At the least these results point to the desirability of allowing for non-linear competition effects within each distance category in applications of Seim's model of endogenous location choices. More generally the analysis confirms earlier indications[49] that data on product ranges can be used to study competition in retail markets. This is of interest because information on firms' other choice variables such as prices and quantities is often unavailable for reasons of confidentiality etc. A researcher can more easily observe basic information about the number and types of good sold by each firm.

A pressing task in any extensions of this work would be to explicitly incorporate the behaviour of eyewear chain stores, which tend to have the widest product variety. Above it is assumed that chain identity reveals no information about the distribution across locations of a seller's idiosyncratic profit component. In fact it is likely that chain stores are particularly suited to specific locations, such as shopping malls. This would imply a different set of distributional assumptions for the entry model. If such assumptions could be rendered tractable in the first-stage entry estimation they might yield additional information about the variety choices in the second stage.

Structural approaches to competition in variety would also be of interest. In retail markets these would need to incorporate the geographic differentiation between sellers, following the work of Thomadsen (1999), Manuszak (2000) and Davis (2001). Those authors derived sellers' demand functions taking as primitives consumer preferences over locations and store characteristics.[50] However in both cases the locations of sellers are held to be

[49] Bayus and Putsis (2000), Draganska and Jain (2001).

[50] To some extent the 'number of screens per cinema' variable in Davis (2001) might be thought of as a proxy for product variety. His welfare analysis suggests that the market provides fewer screens-per-theatre than the social optimum.

exogenously determined. It is hoped that the present paper constitutes a small step toward combining detailed analysis of competition in geographically differentiated markets with an endogenous treatment of firms' locations in the product space.

Appendix: Notes on the likelihood maximization

The Seim entry framework falls into the class of nested-fixed-point models (e.g., Rust (1987)). I used a Broyden-Fletcher-Goldfarb-Shanno procedure to maximize the likelihood in (15) numerically, after concentrating out σ, μ, and any parameters in β on market-level variables (state dummies, housing starts) which don't vary across tracts. Gradients of the likelihood were calculated analytically. At each iteration of the optimization routine the fixed point in (10) and (11) is found numerically by successive approximations and then substituted into (15). Convergence of these approximations to the fixed point is aided by 'dampening' (e.g., Judd (1998)). A (potentially faster) Newton method of finding the fixed point was also attempted, but with no success due to poor convergence.

References

Anderson, S.P., & A. de Palma (1992), 'Multiproduct firms: A nested logit approach', *Journal of Industrial Economics*, 40, 3, 261-76.

______ & R. Renault (1999), 'Pricing, product diversity, and search costs: a Bertrand-Chamberlin-Diamond model', *RAND Journal of Economics*, 30, 719-35.

Bayus, B.L. & W.P. Putsis (2000), 'Product proliferation: An empirical analysis of product line determinants and market outcomes', *Marketing Science*, 18, 137-53.

Benham, L. (1972), 'The effect of advertising on the price of eyeglasses', *Journal of Law and Economics*, 15, 337-52.

Berry, S.T. & J. Waldfogel (2001), 'Do mergers increase product variety? Evidence from radio broadcasting', *Quarterly Journal of Economics*, 116, 3, 1009-25.

Bresnahan, T. F. & P.C. Reiss (1991), 'Entry and competition in concentrated markets', *Journal of Political Economy*, 99, 977-1009.

Dana, J. (2001), 'Competition in price and availability when availability is unobservable', *RAND Journal of Economics*, 32, 497-513.

Davis, P., (2001), 'Spatial Competition in Retail Markets: Movie Theaters', mimeo.

De Palma, A., R. Lindsey, B. von Hohenbalken, & D. S. West (1994), 'Spatial price and variety competition in an urban retail market: A nested logit analysis', *International Journal of Industrial Organization*, 12, 331- 357.

Draganska, M. & D. Jain (2001), 'Product-line length decisions in a competitive environment', mimeo.

Dubin, J.A., & D. McFadden (1984), 'An Econometric Analysis of Residential Electric Appliance Holdings and Consumption', *Econometrica*, 52, 2, 345-62.

Dudey, M. (1990), 'Competition by choice: the effect of consumer search on firm location decisions', *American Economic Review*, 80, 1092-104.

Fujita, M. & J.-F. Thisse (2002), *Economics of agglomeration: cities, in-*

dustrial location, and regional growth, Cambridge U.P.

Haas-Wilson, D. (1989), 'Strategic regulatory entry deterrence: An empirical test in the ophthalmic market', *Journal of Health Economics*, 8, 339-52.

Israelevich, G. (2002), 'Assessing product-line decisions with supermarket scanner data', mimeo., Univ. of Chicago GSB.

Judd, K.L. (1998), *Numerical methods in economics*, MIT Press, Cambridge MA.

Konishi, H. (1999), 'Concentration of competing retail stores', mimeo., Boston College.

Kwoka, J.E. (1984), 'Advertising and the price and quality of optometric services', *American Economic Review*, 74, 211-16.

Manuszak, M.D. (2000), 'Firm conduct and product differentiation in the Hawaiian retail gasoline industry', mimeo.

Mazzeo, M.J. (2002), 'Product choice and oligopoly market structure', *RAND Journal of Economics*, 33, 221-42.

______ (2000), 'Competitive Outcomes in Product-Differentiated Oligopoly', forthcoming in *Review of Economics and Statistics.*

National Research Bureau (1998), *Shopping Center Directory: Midwest Volume.*

Prentice, D. & H. Sibly (1996), 'A search theoretic explanation of multi-outlet retailers', *Economic Record*, 72, 359-69.

Rand McNally Marketing Atlas ().

Redstone, S. (2001), *A Passion to Win*, Simon and Schuster, New York.

Rust, J.D. (1987), 'Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher', *Econometrica*, 55, 5, 999-1033.

Seim, K. (2002), 'An empirical model of firms' entry and endogenous product-type choices', mimeo., Stanford GSB.

Thomadsen, R. (1999), 'Price competition in industries with geographic differentiation: the case of fast food', mimeo.

Watson (in progress), 'Multiproduct firms in monopolistic competition', mimeo.

Wolinsky, A. (1983), 'Retail trade concentration due to consumers' imperfect information', *Bell Journal of Economics*, 14, 275-82.

Wolinsky, A. (1986), 'True monopolistic competition as a result of imperfect information', *Quarterly Journal of Economics*, 101, 493-511.

Table 1: Sellers of eyeglasses by category

Category	**Number of sellers**	
	Overall	In frames subsample
Eyecare specialists:		
Optician	18	11
Optometrist	361	212
Ophthalmologist	88	55
Specialists total	467	278
Department store	105	64
Total all categories	**572**	**342**

Note 1. The frames subsample has 100% of sellers in IL, 50% elsewhere.

Table 2: Eyecare specialists, by affiliation

Type	**Number of sellers**	
	Overall	In frames subsample
Lenscrafters, Pearle Vision	40	22
Other wide-area chain	55	25
Local chain or unaffiliated	371	231

Note 1. The frames subsample has 100% of sellers in IL, 50% elsewhere.

Note 2. 'Wide-area' means operating in more than three markets in the sample.

Table 3: List of markets in sample

State	**Markets**
Illinois	Bloomington, Carbondale, Danville, Decatur, De Kalb, Freeport, Galesburg, Kankakee, Quincy, Springfield, Urbana
Indiana	Bloomington, Columbus, Kokomo, Lafayette, Marion, Richmond, Terre Haute
Iowa	Ames, Burlington, Cedar Rapids, Clinton, Dubuque, Fort Dodge, Iowa City, Mason City, Waterloo
Michigan	Battle Creek, Benton Harbor, Holland, Jackson, Muskegon
Minnesota	Mankato, Rochester, St. Cloud
Ohio	Ashtabula, Findlay, Lancaster, Lima, Mansfield, Marion, Newark, Sandusky, Zanesville

Table 4: Market characteristics

Description	Min	Mean	Max
No. of tracts in market	8	22	39
Total popn. of market (thousands)	26	86	165
No. of eyeglass sellers in market	4	13	23
No. of ophthalmology practices in market	0	2.6	6
No. of malls in market $\geq 150,000\text{ft}^2$	0	1.3	2
No. of shopping plazas in market $\geq 150,000\text{ft}^2$	0	2.2	5

Note 1. Population figures are approximate, based on estimates for 2000.

Note 2. 'Ophthalmology' excludes a few MD's who do not do exams for eyeglasses.

Note 3. Areas for malls and plazas are gross leasable areas.

Table 5: Explanatory variables for location-choice model

Description	Mean	St. dev.
Tract-level variables		
Median age, 2000 (years)	37	5.9
Median rent in tract, 1990 ($)	370	76
No. of exogenous (SWTS) dept. stores in band 1	0.12	0.42
No. of exogenous (SWTS) dept. stores in band 2	1.7	1.1
No. of ophthalmology practices in tract	0.12	0.41
No. of fast-food shops in band 1	1.4	1.7
GLA of malls in band 1 (100,000 ft^2)	5.4	2.4
No. of malls in band 2	1.3	0.60
No. of unenclosed plazas in band 1	0.16	0.44
Pop. of tracts in band 1, 2000 (10,000)	0.87	0.53
Pop. of tracts in band 2, 2000 (10,000)	9.1	3.4
Market-level variables		
New private house starts, 2000 ($million)	51	40

Note 1. GLA means gross leasable area.

Note 2. Only malls and plazas with $\geq 150,000\text{ft}^2$ GLA are counted.

Note 3. GLA of malls is averaged only over tracts with a mall in band 1.

Note 4. 'Band 1' refers to the band 0-1 miles, 'band 2' to anything further.

Note 5. 'SWTS' means Sears, Walmart Super, Target, ShopKo.

Table 6: No. of frames per seller, by category

Category	**No. of frames per seller**		
	25%-ile	Median	75%-ile
Eyecare specialists:			
Optician	385	484	613
Optometrist	450	603	809
Ophthalmologist	396	612	793
All specialists	438	600	795
Department stores	385	434	501
All sellers	**410**	**537**	**750**

Note 1. Frames data is for the subsample with 100% of sellers in IL, 50% elsewhere.

Table 7: No. of frames at eyecare specialists, by affiliation

Affiliation	**No. of frames per seller**		
	25%-ile	Median	75%-ile
Lenscrafters, Pearle Vision	836	1099	1411
Other wide-area chain	578	670	835
Local chain or unaffiliated	410	548	738

Note 1. Frames data is for the subsample with 100% of sellers in IL, 50% elsewhere.

Note 2. ‘Wide-area’ means operating in more than three markets in the sample.

Table 8: Location-choice model, with exogenous SWTS sellers

Variable	**Estimate**	**Std. error**
Median age in tract (tens of years)	0.20‡	0.12
Median rent in tract ($100)	-0.11	0.082
Dummy for ordinary Walmart in tract	0.58*	0.27
Dummy for first SWTS seller in band 1	0.23	0.49
No. of SWTS sellers in band 1	-0.14	0.31
No. of SWTS sellers in band 2	-0.016	0.16
No. of ophthalmology practices in tract	1.01*	0.13
No. of fast-food sellers in band 1	0.17*	0.044
Population in band 1 (10,000)	-0.073	0.19
Population in band 2 (10,000)	-0.0019	0.055
Dummy for enclosed mall in band 1	0.014	0.34
GLA of malls in band 1 ($100,000$ ft^2)	0.16*	0.061
No. of malls in band 2	0.0047	0.25
No. of open shopping plazas in band 1	0.23‡	0.13
Base effect of extra rival in band 1 (θ_1)	-0.46*	0.080
Base effect of extra rival in band 2 (θ_2)	(Fixed at 0)	
Extra effect for each initial band-1 rival (θ_3)	0.80*	0.16
Mean of ξ (μ)	-5.2*	0.62
Variance of ξ (σ^2)	0.17*	0.053
Val. of private housing starts in mkt. ($100 mn)	0.50	0.32
Illinois dummy	-0.40	0.28
Ohio dummy	-0.19	0.24
Indiana dummy	0.17	0.36
Michigan dummy	0.17	0.51

(*) significant at 5% level. (‡) significant at 10% level.

Note 1. Standard errors are BHHH estimates.

Note 2. GLA means gross leasable area.

Note 3. ‘SWTS’ means Sears, Walmart Super, Target, ShopKo.

Table 9: Limited-information location model

Variable	Estimate	Std. error
Median age in tract (tens of years)	0.24*	0.087
Median rent in tract ($100)	-0.14‡	0.073
Dummy for ordinary Walmart in tract	0.79*	0.24
Dummy for first SWTS seller in band 1	0.091	0.47
No. of SWTS sellers in band 1	0.0049	0.30
No. of SWTS sellers in band 2		
No. of ophthalmology practices in tract	1.2*	0.11
No. of fast-food sellers in band 1	0.21*	0.037
Population in band 1 (10,000)	0.038	0.17
Population in band 2 (10,000)		
Dummy for enclosed mall in band 1	-0.20	0.56
GLA of malls in band 1 ($100,000$ ft^2)	0.19*	0.064
No. of malls in band 2	-0.19	0.42
No. of open shopping plazas in band 1	0.21‡	0.12
Base effect of extra rival in band 1 (θ_1)	-0.52*	0.074
Base effect of extra rival in band 2 (θ_2)		
Extra effect for each initial band-1 rival (θ_3)	0.75*	0.14
Mean of ξ (μ)		
Variance of ξ (σ^2)		
Val. of private housing starts in mkt. ($100 mn)		
Illinois dummy		
Ohio dummy		
Indiana dummy		
Michigan dummy		

(*) significant at 5% level. (‡) significant at 10% level.

Note 1. Standard errors are BHHH estimates.

Note 2. GLA means gross leasable area.

Note 3. 'SWTS' means Sears, Walmart Super, Target, ShopKo.

Note 4. Omitted parameters are not identified in this model.

Table 10: Explanatory variables for variety regressions

Description	**Mean**	**St. dev.**
Dummy for first SWTS store in band 1	0.34	0.48
Number of SWTS stores in band 1	0.45	0.71
Number of SWTS stores in band 2	1.4	1.1
Dummy for ordinary Walmart store in tract	0.12	0.32
No. of ophthalmology practices in tract	0.58	0.81
Dummy for enclosed mall in band 1	0.37	0.48
GLA of malls in band 1 ($100,000$ ft^2)	6.1	2.1
No. of malls in band 2	0.96	0.67
Pop. of tracts in band 1, 2000 (10,000)	0.91	0.55
Number of rival sellers in band 1	2.4	1.9
Number of rival sellers in band 2	9.0	3.9
(*)Dummy for closest rival in band 1	0.79	0.41
(*)Dummy for 2nd-closest rival in band 1	0.60	0.49
(*)Dummy for 3rd-closest rival in band 1	0.42	0.49
New pvt. house starts in market, 2000 ($million)	57	39

Note 1. GLA means gross leasable area.

Note 2. GLA of malls is averaged only over sellers with a mall in band 1.

Note 3. 'Band 1' refers to the band 0-1 miles, 'band 2' to anything further.

Note 4. 'SWTS' means Sears, Walmart Super, Target, ShopKo.

Note 5. Variables are averaged over the sellers in the variety regression subsample.

Note 6. By assumption asterisked variables all have the same coefficient: ψ_3.

Table 11: OLS variety regressions for non-SWTS sellers

Dependent variable: log(number of frames).

RHS parameters	**Estimate**	**Standard error**
Dummy for ordinary Walmart in tract	-0.074	0.11
Dummy for first SWTS seller in band 1	0.26‡	0.14
No. of SWTS sellers in band 1	-0.087	0.090
No. of SWTS sellers in band 2	0.022	0.034
No. of ophthalmology practices in tract	-0.046	0.073
Population in band 1 (10,000)	-0.043	0.061
Dummy for enclosed mall in band 1	-0.13	0.16
GLA of malls in band 1 ($100,000$ ft^2)	0.044	0.027
No. of malls in band 2	0.12‡	0.062
Base effect of extra rival in band 1 (ψ_1)	-0.080*	0.037
Base effect of extra rival in band 2 (ψ_2)	-0.025	0.020
Extra effect for each initial band-1 rival (ψ_3)	0.098‡	0.051
Market-level endogeneity correction (ξ)	-0.10	0.12
Tract-level (logit) endogeneity correction	-0.032	0.067
Val. of private housing starts in mkt. ($100 mn)	0.046	0.12
Illinois dummy	-0.0054	0.086
Ohio dummy	-0.17‡	0.092
Indiana dummy	0.021	0.084
Michigan dummy	-0.0030	0.11
Constant	6.0*	0.59
R^2	0.14	
F	$F(19, 281) = 2.37, \Pr[> F] = 0.001$	
N	301	

Note 1. Std. errors are not adjusted for the sampling error in the endogeneity corrections.

Note 2. Std. errors are robust.

Note 3. ‘SWTS’ means Sears, Walmart Super, Target, ShopKo.

Note 4. GLA means gross leasable area.

(*) significant at 5% level. (‡) significant at 10% level.

Table 12: OLS variety regressions for non-SWTS sellers, with chain dummies

Dependent variable: log(number of frames).

RHS parameters	**Estimate**	**Standard error**
Dummy for ordinary Walmart in tract	-0.013	0.099
Dummy for first SWTS seller in band 1	0.18	0.13
No. of SWTS sellers in band 1	-0.079	0.077
No. of SWTS sellers in band 2	0.031	0.031
No. of ophthalmology practices in tract	-0.051	0.070
Population in band 1 (10,000)	-0.028	0.059
Dummy for enclosed mall in band 1	-0.12	0.15
GLA of malls in band 1 ($100,000$ ft^2)	0.028	0.025
No. of malls in band 2	0.11‡	0.060
Base effect of extra rival in band 1 (ψ_1)	-0.077*	0.035
Base effect of extra rival in band 2 (ψ_2)	-0.028	0.020
Extra effect for each initial band-1 rival (ψ_3)	0.089‡	0.049
Market-level endogeneity correction (ξ)	-0.10	0.12
Tract-level (logit) endogeneity correction	-0.046	0.065
Dummy for Lenscrafters or Pearle Vision	0.73*	0.094
Dummy for any other chain	0.18*	0.062
Val. of private housing starts in mkt. ($100 mn)	0.057	0.12
Illinois dummy	-0.0098	0.080
Ohio dummy	-0.093	0.089
Indiana dummy	0.081	0.081
Michigan dummy	0.045	0.094
Constant	6.0*	0.59
R^2	0.27	
F	$F(21, 279) = 4.89, \Pr[> F] = 0.0000$	
N	301	

Note 1. Std. errors are not adjusted for the sampling error in the endogeneity corrections.

Note 2. Std. errors are robust.

Note 3. 'SWTS' means Sears, Walmart Super, Target, ShopKo.

Note 4. GLA means gross leasable area.

(*) significant at 5% level. (‡) significant at 10% level.

Figure 1: Residents per eyewear seller by market population

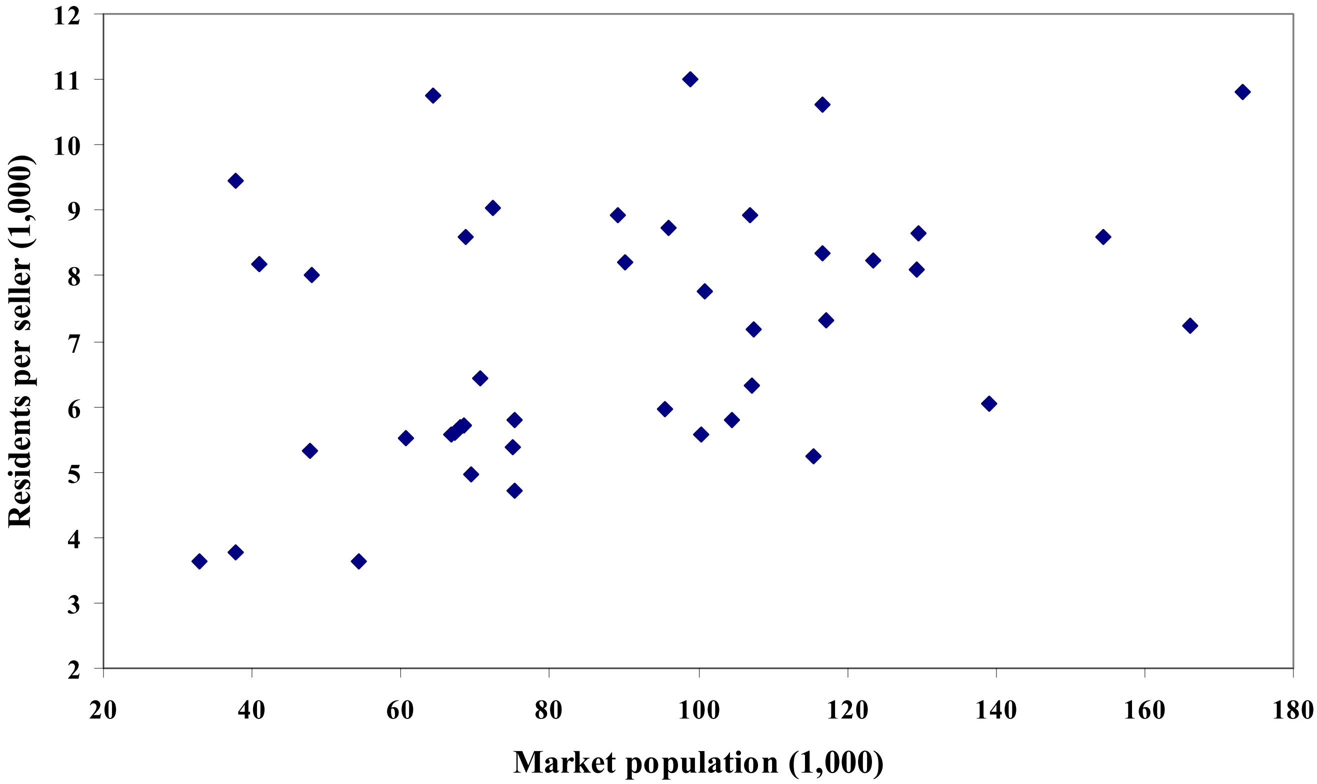
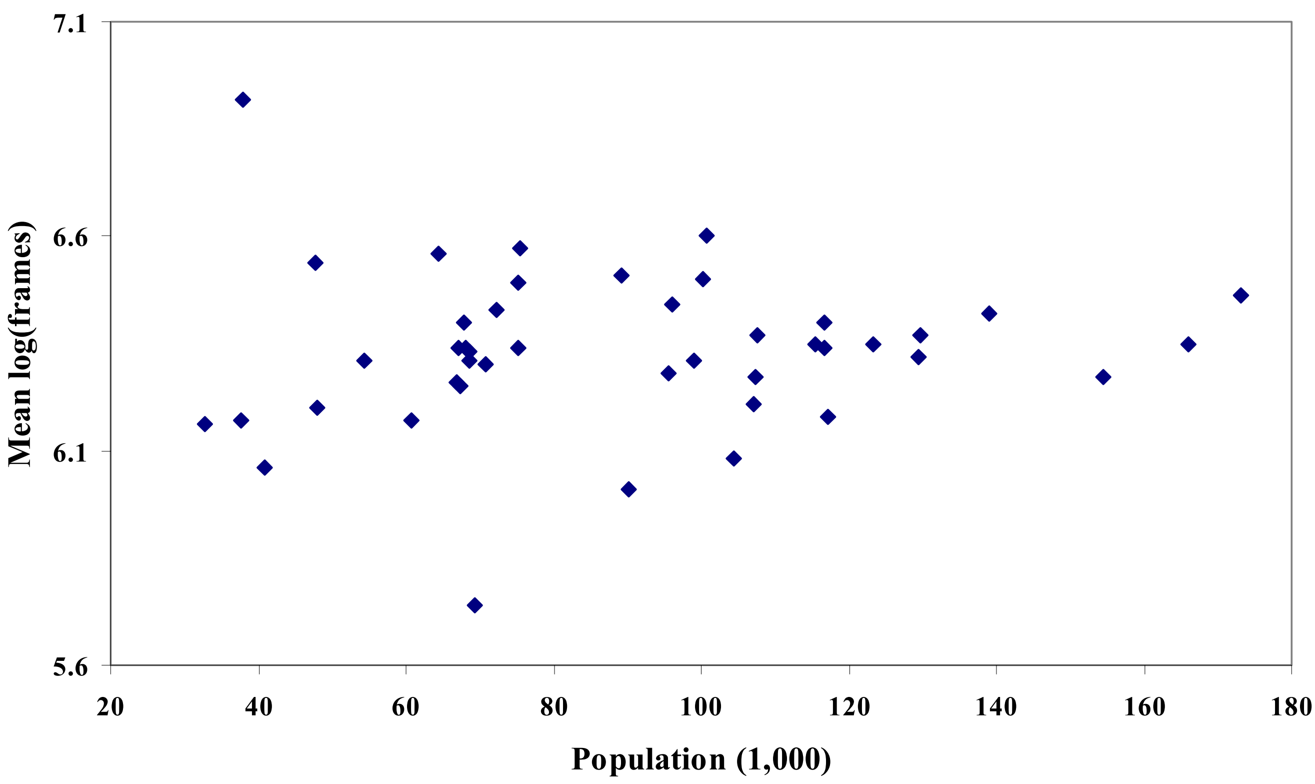


Figure 2: Mean log of frames per seller, by market population



N.B. Averages in Figure 2 are taken over 100% of sellers in Illinois markets, 50% of sellers elsewhere.

Figure 3: Effects of competition on profits

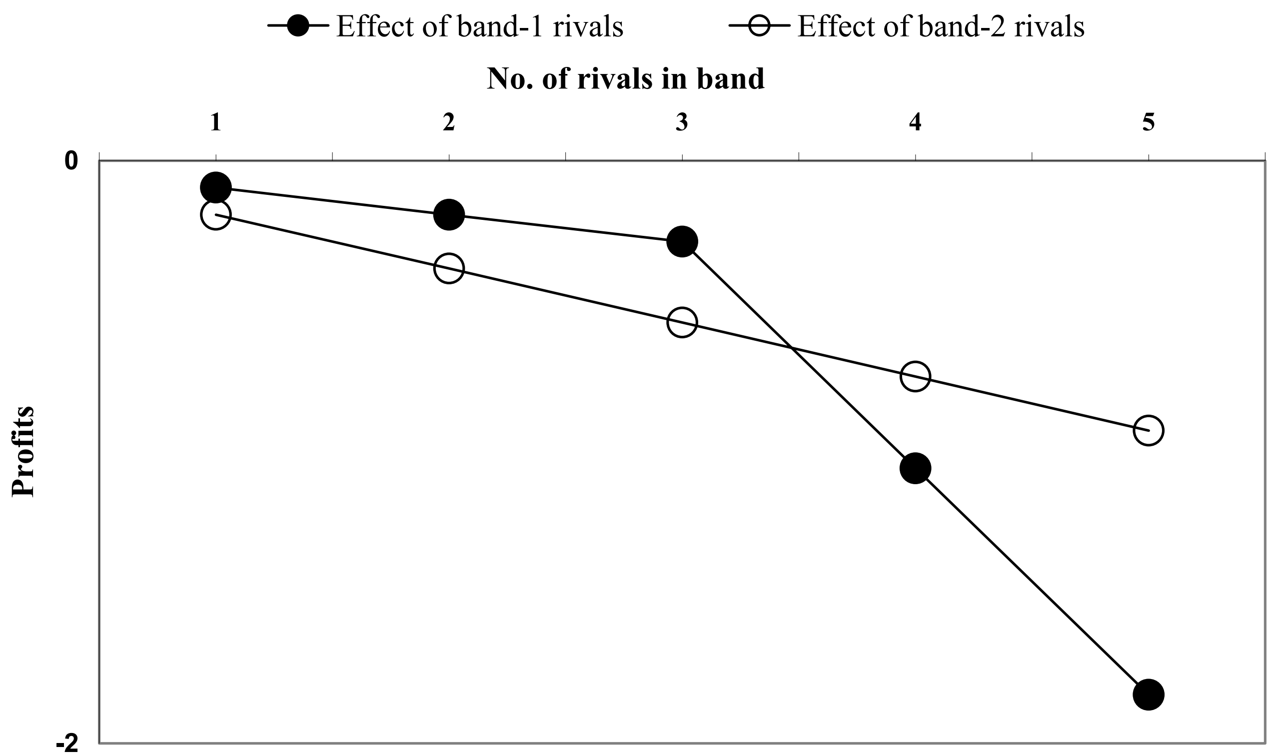


Figure 4: Distribution of sellers, by type and number of frames