Public Monopoly and Economic Efficiency: Evidence from the Pennsylvania Liquor Control Board's Entry Decisions

Katja Seim
The Wharton School
University of Pennsylvania

Joel Waldfogel
The Wharton School
University of Pennsylvania and NBER

February 1, 2010

Abstract

When fixed costs are present, markets can deliver an excessive number of products in inefficient product space locations under free entry. In this paper, we study product provision under an alternative form of market organization, state-run monopoly, using data from the Pennsylvania Liquor Control Board (PLCB), the state's monopolist wine and spirits retailer. We use information on store location choices, prices, wholesale cost, and sales to uncover the goals implicit in its entry decisions.

We estimate a model of demand for liquor as a function of the price, distance to stores, and other demographic characteristics. In counterfactual analyses, we calculate configurations of stores that 1) maximize welfare, 2) maximize profit, and 3) mimic free entry. We then compare the actual PLCB store network to these analogs of theoretical benchmarks.

We find that the PLCB's implicit goal is better characterized by welfare than profit maximization. The state has roughly 3 times the number of stores that would maximize profit and about a fifth more than would maximize welfare. We also find that atomistic location decisions would dissipate between 6.9 and 11.5% of welfare from excessive clustering and entry of stores arising from the failure to internalize business stealing.

I. Introduction

It is well known in theory that when fixed costs are present, markets can deliver the wrong number of products under free entry. When the welfare benefit of a product exceeds its costs, but the revenue available to the seller does not, then markets will deliver inefficiently too few products. When the revenue available to a product covers the costs of multiple entrants, the market can deliver too many.[1] Additional entry benefits consumers from putting downward pressure on prices as well as – with differentiated products – offering additional variety. But the private benefit of entry includes both new business and business diverted from existing firms, so the private benefit of entry generally exceeds the social benefit, leading to excessive entry as well as entry in the wrong locations. Thus, nothing about market processes guarantees that the benefits of additional entry will overcome the additional fixed or sunk costs required for the operation of additional entrants. In the few instances when they have been studied empirically, the magnitude of misallocation from free entry has been large. Berry and Waldfogel (1999) document that free entry in radio broadcasting results in roughly three times too many firms, relative to either profit or welfare maximization.[2] When products are differentiated, for example by location, then free entry can generate not only the wrong number of products but also the wrong configuration of product positions; that is, locations.

The possibility of market entry processes generating the wrong firm configuration places attention on possible alternative arrangements, such as entry regulation or government operation of industries. Market entry may have theoretical and even empirically documented shortcomings; we need to know how the possible alternatives function, however, to move beyond mere academic criticism of market entry processes.

While state-run enterprises are increasingly rare around the world, the Commonwealth of Pennsylvania has a state-run monopoly on the sale of wine and spirits, the Pennsylvania Liquor Control Board (PLCB). (Beer is sold in separate, privately owned stores.) In this paper we try to uncover the goals implicit in its entry decisions. Does the PLCB behave like a monopolist, maximizing profits? Or does it maximize welfare? And how does its entry pattern compare with

---

[1] Spence (1976), Dixit & Stiglitz (1977), Mankiw & Whinston (1986).

[2] Hsieh and Moretti (2003) study a related phenomenon, excess entry into real estate markets. The friction causing excess entry in their context, however, is a collusive price rather than a supply side non-convexity.

what would occur under free entry? To study these questions we develop a simple model of demand for a product located in geographic space. In our model – as in the empirical context that we study – prices are fixed. Consumers are distributed across the state, and they benefit, in the form of lower effective prices, from proximity to outlets. Additional outlets, however, are costly to run. The state's problem is to choose a set of store locations – and a corresponding number of stores operating – to best serve its objective. Given the state's choice of locations for stores, our goal is to infer its objective.

We estimate a model of demand for liquor as a function of the price, distance to stores, and other demographic characteristics. The distance parameter and the price elasticity are the two key parameters that together allow us to quantify the consumer benefit from greater proximity. We observe the fixed retail price as well as the wholesale cost, allowing direct calculation of producer surplus. The model allows us to calculate demand, consumer surplus, and producer surplus for any configuration of stores. In particular, we calculate configurations of stores that 1) maximize welfare, that 2) maximize profit, and that 3) mimic free entry. We then compare the actual PLCB store network to these analogs of theoretical benchmarks.

Compared to other work that uses a model to simulate the number of outlets operating under different entry regimes (monopoly, welfare maximization, and free entry), we study both the number and location of outlets under each regime. Because prices are uniform across stores in the context we study, we cannot measure or directly incorporate effects of price competition. Instead, we compare regimes, given fixed markups. In the US, entry into liquor retailing is regulated widely through the number of licenses per municipality, so while free entry is of general academic interest, it is not a relevant counterfactual for this industry. Even under regulatory limits on the total number of stores in an area, firms are free to choose their locations. The analog to free entry relevant in this context is the possibility that a given number of atomistic firms that do not internalize business stealing enter in the wrong locations. By our choice of the price-cost markup in conjunction with the store fixed cost, we can induce a free-entry equilibrium with any target number of stores. We can compare the locations chosen under free entry with the locations of equal-sized configurations chosen in ways that internalize business stealing. We can ask, for example, whether stores would cluster excessively – like Hotelling (1929) duopolists – under free entry.

Our paper proceeds in six sections. First, we present a simple theoretical model that illustrates the differing entry patters arising under profit-maximization, welfare-maximization, and competition. We then proceed to a description of the PA liquor retailing system, current policy discussions, and a discussion of the data used in the study. Section 3 describes the data and documents the relationships of interest for the study. Section 4 presents estimates of a simple model of liquor demand at particular stores. We then turn to exercises made possible by the model. Section 5 presents simulation results based on profit and welfare maximization: (i) we compare the incremental benefits of existing stores to store operation costs for an indication of whether the current system is too large or too small; (ii) we derive the optimal geographic configuration of stores – for profit and welfare maximization – in five counties of varying sizes; (iii) and we develop a tractable algorithm for calculating optimal configurations in larger areas, and we show that it approximates the exact solution well. Section 6 analyzes the welfare properties of free entry. We first present an algorithm determining approximate free entry equilibria in small (and medium-sized) pieces of geography. We then run this algorithm on 5 counties in Pennsylvania. We calculate the number of stores operating under free entry as well under free entry regulated to target either the actual number of stores operating or the model's welfare maximizing number. Comparison of welfare under free entry and welfare maximization, for a given network size, provides an estimate of the loss from unregulated atomistic store location. Section 7 turns to the whole-state problem. We first use the algorithms to determine profit and welfare-maximizing configurations for the whole state. We then turn to the constrained problem of which of its existing outlets the state should operate if maximizing profit or total welfare. Section 8 asks a related question: what welfare weights are implicitly attached to types of consumers by existing entry patterns?

We find that the PLCB's implicit goal is better characterized by welfare than profit maximization. The state has roughly three times the number of stores that would maximize profit and about a fifth more than would maximize welfare, but – judging by results from select county markets – fewer than would operate with free entry. We also find that atomistic location decisions would dissipate between 6.9 and 11.5% of welfare from excessive clustering and entry of stores arising from the failure to internalize business stealing.

II. A Model Illustrating Welfare Effects of Entry with Space

Consider the following simple model with fixed prices and transport costs. Consumers are uniform on [0, 100]. The price of the product is \$10. Marginal costs are \$0, and fixed costs are \$240. Consumers derive $U = 60 - p -$ distance.

If one firm operates, it will garner the most business locating at 50. Because $U > 0$ throughout the space when the store locates at 50, all 100 consumers buy, so revenue is \$1000 (and profits are \$760). Consumer surplus is 2500.[3] If more firms operate, all consumers will continue to buy, so revenue and producer surplus remain \$1000, although profits fall by the fixed cost with each entrant. Consumer surplus rises with entry because stores get closer to more consumers who therefore incur smaller travel costs. That profits fall with entry while consumer surplus rises creates the tension between the interests of consumers and producers.

With J outlets spaced symmetrically at $d_0 + 2d_0 i$, i=0,…,J-1 and $d_0 = \frac{100}{2J}$, consumer surplus can be expressed as a function of the distance traveled by consumers at the midpoint between two stores, d_0:

$$CS = J(d_0^2) + 100(60 - 10 - d_0). \tag{1}$$

In this setup, profits are maximized at \$760 with J=1. Total welfare is maximized with J=3. Free entry dissipates the entire producer surplus (up to an integer constraint) and results in 4 firms operating.

While simple, this model shows important consequences of different ownership arrangements. A monopolist seeking to maximize profits would operate too few outlets. Free entry, on the other hand, would result in too many. In addition to the welfare losses from the wrong number of firms, there are also possible losses from firms choosing the wrong locations. For example, the well-known Hotelling duopolists cluster at the center of the space under free entry even though welfare maximization dictates spreading. The Hotelling result is fragile, and it is in general difficult to analytically predict effects of ownership structure on product mix. A number of studies examine the positive effects of ownership structure – and internalization of

[3] Consumer surplus is represented by two right triangles that are length 50 on a leg.

business stealing – on product spreading.[4] We seek here to examine both effects on location and spacing as well as to measure their welfare impacts.

We can also contrast the incentives to operate an additional outlet under three objectives – monopoly, competition, and welfare-maximizing planning – using a demand curve. Given the fixed cost of operating a store, each objective gives rise to a different number and configuration of stores. Increasing the number of stores by one decreases the distance traveled by some consumers; in our simple model, d_0 falls from $\frac{100}{2N}$ to $\frac{100}{2(N+1)}$. The new store's proximity reduces the effective price, $p + d_0$. The monopolist does not add this store because there are no market expansion effects and no additional revenue is generated. It's clear that the benevolent planner operates more stores because he or she adds a store when the increment to consumer surplus exceeds fixed costs. We can also say *a priori* that free entry will result in more stores than profit-maximizing monopoly since an incremental store is viable as long as its equal share of revenue exceeds the fixed cost.

III. Background and Data

Background

Pennsylvania is the largest of 18 "control" states that control the sale and distribution of alcohol at the wholesale and retail levels in different forms to protect the public from the negative externalities of alcohol consumption. The states take different approaches to liquor retailing, however. Some states operate fully private retailing operations, but limit the supply of licenses; others turn over the operation of state stores to private enterprises; yet others continue to operate state-run stores that in some cases compete with privately-run stores. Pennsylvania operates a privatized system for the sale of beer, but acts as a state monopolist in selling wine and liquor through a system of state-run stores. Relative to other states, Pennsylvania has a very low store density. Figure 1 shows histograms of the number of stores per 1000 people, separately for liquor control states and for the remaining states, differentiating between the number of stores selling alcohol (spirits, wine, and beer) and the number of stores selling spirits for the subset of states in which legislation restricts the sale of spirits to stand-alone liquor stores.

[4] See Steiner (1952), Berry and Waldfogel (2001), Sweeting (2006).

Since Pennsylvania's regulation of the sale of wines and spirits is tighter than the sale of beer, the latter may represent the more relevant comparison. On both metrics, however, Pennsylvania's store density is unusually low; in the median state 0.76 stores and 0.19 spirits stores operate per 1000 people compared to Pennsylvania's 0.16 stores and 0.05 wine and spirits stores per 1000 people.

Figure 2 shows the stores' locations around the state. Many are in or near the two large metropolitan areas in the state, Philadelphia and Pittsburgh. Others are clustered around the smaller areas such as Wilkes-Barre, Erie, Lancaster, Harrisburg, and York. Others stand alone in less densely populated areas. We can calculate each store's ambient demographics, based on Census tracts for which the store is closest. By this method, for the first week of the year, the average ambient population of a store was 18,154. The inter-quartile range ran from 12,842 to 24,781. Assuming that all population resides at Census tract centroids, the average (median) distance to the nearest store was 3.2 (2.4) kilometers, with an interquartile range of 1.0 to 3.6 km.

While privatization is not currently under discussion, the PLCB has taken strides toward improving the system's efficiency and consumer friendliness. In the 1970s the stores consisted of a front counter where a customer would request a bottle of liquor from the back room. In the last few years the stores have added chilled rooms for fine wines in many stores. According to an August 2007 press release, the agency has recently tried to operate more "like a business" and hired a CEO to oversee day to day operations. "Governor Rendell has given this agency a mandate to operate like a business, and that means getting costs under control." One strategy the PLCB has recently pursued is reducing the number of stores operating. According to PLCB Chairman Patrick Stapleton, "We took important steps last year toward that end, starting with reducing the number of stores in our system to 630 from 646 a year earlier. This year, PLCB 75 [the Board's strategic plan] will continue to make improvements to our operations that will have a direct impact on our bottom line, and in how Pennsylvanians interact with us."[5]

External observers agree that store closings are an avenue to greater profitability for the system. A private equity lawyer quoted in the Pittsburgh Gazette speculated that private owners would achieve greater profitability by closing stores in remote locations rather than by replacing

---

[5] PA LIQUOR CONTROL BOARD REPORTS RECORD RETURN FOR FISCAL 2006-07, August 28, 2007. Press release. http://www.lcb.state.pa.us/webapp/agency/press/press_detail.asp?press_no=07-18&psearch=&offset=4, accessed October 17, 2008.

union workers with minimum wage clerks. At the same time, the Independent State Store Union criticized the PLCB for being too focused on profit in reducing the number of stores, suggesting that small, rural communities are hurt in selection and availability from a move toward profit maximization.[6] There is also speculation in the press that political lobbying and considerations play a significant role in store closing and transfer decisions, countering the stated profit motives of the board.[7]

As of the first week of 2005, the PLCB operated 624 stores, representative of the store count patterns between 2001 and 2005 depicted in Table 1. In 2005 and 2006, the PLCB opened a number of outlets within the premises of grocery stores, accounting for 68.42% of the entering stores. Both new and closing stores tend to be smaller on average; average monthly revenue over the six-month period subsequent to store opening and prior to store closing amount to 89.36% and 45.86% of the statewide average. These new outlets reduce travel times to the closest store; the average distance between a new outlet and its next closest store is 1.37 km (median 1.38 km), compared to an average distance between an existing store and its next closest store of 3.08 km (median 1.90 km). Closing stores occupy similar locations within the larger PLCB store network to existing stores, with an average distance to the next closest store of 3.06 km (median 1.50 km), suggesting that the PLCB is not closing primarily rural stores. The PLCB further operates seven "outlet" stores near the borders with neighboring states. As of the first week of 2005, 65 stores, or 10.37%, are designated premium-collection stores that are larger in size and carry a wider variety of products than the remaining locations. Lastly, approximately 25% of stores are open part of the day on Sunday.

Our analysis below investigates how close the PLCB comes to replicating a monopolist's decision, relative to the welfare-optimizing solution. We rely on the usual inputs in calculating store profitability and discuss how these are determined by the PLCB in turn.

Retail Prices. The PLCB charges an identical retail price for a particular product in all of its stores and uses a simple mark-up rule to determine the price, applying a 30% markup and an

[6] Pittsburgh Post Gazette, "Is Time Right to Sell State Stores?" June 12, 2007. http://www.post-gazette.com/pg/07163/793365-85.stm, accessed October 17, 2008.
[7] Pittsburgh Post Gazette, "LCB works in curious ways" January 1, 2008. http://www.post-gazette.com/pg/08028/852743-85.stm, accessed October 17, 2008.

18% liquor tax to the wholesale price.[8,9] It determines the appropriate wholesale price to use based on a FIFO system. The PLCB engages in some promotional activity, using manufacturers' coupons and running monthly sales (28-day period beginning on the Monday closest to the end of the month).

Marginal Cost. The PLCB negotiates wholesale prices directly with its suppliers. A new product's wholesale price remains fixed for one year after introduction. For established products, the PLCB re-negotiates over cost increases on a quarterly basis rotating through product categories over the course of its four-week long reporting period. Each reporting period, the wholesale price of a subset of products is adjusted, translating into changes in the retail price. In contrast to sales periods, reporting periods begin on a Thursday, usually in the middle of the month.

Fixed Cost. We consider as fixed costs the PLCB's cost to run stores (personnel, rental cost of property, storage, etc.) and warehousing. The stores had revenue of $1.47 billion in fiscal year 2004-05, equivalent to 118 million bottles of liquor at the 2005 average price of $12.38 (and at an average wholesale price of $7.31). Of the $384.3 million in variable profit and taxes, $245.9 million was used to run stores and warehousing. This corresponds to $394,051 per store in annual fixed costs (or $1,110 per day), given 624 stores in operation as of 1/2005. The remainder was given to the state.[10]

Data

The basic data for the study is a store-level panel obtained from the PLCB under the Pennsylvania Right-to-Know Law.[11] It contains daily information on quantities sold and gross receipts at the UPC and store-level over the period 7/2000 to 6/2006. In addition, we received information on the wholesale cost of each product that is constant across stores and varies over time according to the reporting periods described above. The retail price does not necessarily reflect contemporaneous wholesale cost, but instead depends on the wholesale cost of any

[8] The specific pricing rule is: $retail\ price = (wholesale\ price(1.3) + bottle\ fee)(1.18)$, where the bottle fee amounts typically to $1 and the PLCB rounds the resulting retail price to end in the nearest nine cents. In addition, the consumer pays a 6% PA sales tax.
[9] The PLCB sells certain products at outlet stores that are unavailable in the remaining stores, but does not discount the remaining products. These products tend to be larger-sized bottles or multi-packs.
[10] PLCB's fiscal year 04-05 summary, http://www.lcb.state.pa.us/plcb/cwp/view.asp?a=1334&q=557420&tx=1, accessed 11/5/2008.
[11] 65 P.S. §§ 66.1 *et seq*., as amended.

inventoried product. We geocode the stores' street addresses to assign them to a geographic location, which we link to data on population and demographic characteristics for nearby consumers based on information from the 2000 Census. Because stores open and close during the year, the characteristics of stores' ambient consumers change over time.

We aggregate our data across products to the level of either the day or the week and conduct our analysis for 2005, the latest full year of data. This periodicity accounts for the strong seasonality inherent in liquor sales, which are disguised in more aggregate definitions. Averaging across 32,509 store weeks in 2005, stores sell an average of 2,674 bottles per week. Figure 3 exhibits the strong seasonal pattern to sales, with a trough after New Years (week 1) and peaks at July Fourth (week 26), Thanksgiving (47), and Christmas through New Year's Eve (50-52).

Because we treat liquor as a single quantity in our analysis below, we also need a single price. The PLCB passes on changes in wholesale prices by changing its prices at the beginning of its reporting periods. The first fully contained reporting period in 2005 began on January 12. Prices also change due to sales that begin at the end of each month. Our data are by store, product, and day, so we observe the day that a product's price changes. Since we deduce a store's availability of a product from a sale, our data records the date at which a product was first sold at a new price in any of the PLCB stores. The price changes are clustered heavily at the beginning of a new reporting period and the end of the month: across all products, 38.15% of price changes occur within three days from the beginning of new reporting periods and another 33.88% occur within three days from the beginning of sales periods, in the case where these do not overlap with the beginning of a new reporting period. 90.26% of price changes occur within one week from the beginning of a new reporting or sales period, reflecting that not all products have daily sales in at least one PLCB store. We observe some price changes at other times, which may indicate the use of a manufacturers' promotion.

PLCB stores carry literally thousands of products. The PLCB reviews its product assortment on a category-by-category basis once or twice a year depending on the category, replacing slow-selling or declining products by new products. We calculate a state-wide price index, shown in Figure 4, that is a weighted sum of the system-wide product prices in each week, where the weights are each product's share of 2005 sales. The price series resembles a step function, reflecting the discrete changes in prices due to either sales or wholesale cost changes discussed above. Because of the different bundles purchased at different stores, there is substantial variation across stores in average prices paid locally. A store-specific price index based on each store's sales of each product during 2005 illustrates such differences: for the median week, the average value of the store-specific price index is $12.07 (median $12.05), with a standard deviation of $0.91, and prices of $11.25 and $13.05 in the 10^{th} and 90^{th} percentiles, respectively.

These differences in price indices reflect differences in preferences across consumers, more so than differences in the availability of products across stores. Absent inventory information at the store and product level, we can derive product availability from observed sales only and treat a product as being available in a store if it sold at least once during a given week. Of the 100 best selling products statewide in 2005, the median store carried 98.0% in its median week, while a store at the fifth percentile carried 72.0% of the products. Similarly, of the 1000 best selling products statewide in 2005, the median store carried 82.03% in its median week, while a store at the fifth percentile carried 44.2% of the products. The product availability at premium stores is somewhat better than the average, with the median premium store carrying all of the top 100 products and 95.1% of the top 1000 products. These statistics suggest that most stores carry most products, however, providing some support for our assumption below that differences in product availability do not drive customers' store choices to a significant degree.

Descriptive Evidence

Our goal is to model consumption as a function of demographic characteristics, the configuration of stores, and price. In this section we explore these relationships as a step toward estimating a model of demand. We first examine the relationship between prices and demand, via regressions of log quantities on measures of log prices as well as flexible time effects. Our price measures vary only across time and not place, in line with the PLCB's uniform pricing policy, and thus do not allow fully flexible time dummies. The pricing series in Figure 4 suggests that the PLCB puts products on sale during periods of high demand. The pricing series thus reflects the anticipation by the PLCB of temporary sales spikes during the upcoming month. We address this potential source of endogeneity of the price series in a number of ways.

First, we employ a price index that removes price declines due to potentially endogenous sales. We call this the list price and build a statewide, constant weight, price index based on it. The first two columns of Table 2 report the results of log-log regression of weekly quantity on the state-bundle list price index with selected holiday week dummies and, in column (2) store fixed effects. This results in a price elasticity of approximately –1. Columns (3) and (4) repeat this exercise with a more flexible quadratic seasonality specification, allowing seasonal effects to differ in the last six weeks of the year from the trend of the earlier months. Because we include flexible week dummies, the price elasticity is identified from the co-variation in quantity and the price index after accounting for the constant-across-store change in sales experienced at the same time.

Second, we use a price index based on the actual price series, including price reductions due to sales. We control for unusual spikes or declines in demand around holidays by including time dummies for holiday weeks or days. These time effects address endogeneity concerns to the extent that they control for the relevant temporary changes in demand that the PLCB anticipates in choosing its sale prices. We use weekly data, as in columns (1) through (4), as well as daily data that allow finer controls for seasonality. The resulting price elasticities are similar in magnitude, ranging from –0.86 to –1.94.

Last, we limit our sample to days immediately prior to and immediately following price changes. If the PLCB chooses sales in response to other anticipated non-holiday demand spikes in addition to the ones we control for, the price elasticities in columns (5) through (8) would still be biased. The immediate response in sales to a price change provides clean identification provided the underlying demand does not change significantly between the day before and the day of a price change and provided that customers do not anticipate the sale. Columns (11) and (12) depict estimated elasticities for this subsample of the daily data, again in line with the results above.

Table 2 thus indicates that demand for liquor is higher when the price is lower, with a price elasticity between –0.9 and –1.9. This is in line with estimates from the large empirical literature attempting to measure the elasticity of demand for liquor. Cook and Moore (1999) review the literature on demand for alcohol, most of which use state-level time series data. According to Chaloupka, Grossman, and Safer (2002), "An extensive review of the economic literature on alcohol demand concluded that based on studies using aggregate data (i.e., data that report the amount of alcohol consumed by large groups of people), the price elasticities of demand for beer, wine, and distilled spirits are –0.3, –1.0, and –1.5, respectively (Leung and Phelps 1993)." Furthermore, Chaloupka, Grossman, and Safer report that analyses using individual data "suggest that alcohol demand may be even more responsive to price than these estimates indicate."

The second relationship of interest is between ambient population and quantity demanded. We assign the state's population to stores based on the distance between each store and the Census tracts that make up Pennsylvania. Figure 5 depicts average weekly sales against the population of the resulting catchment areas. Liquor stores with more ambient population sell more bottles, on average.

Table 3 explores this relationship more systematically with multiple regression. The first column reports a regression of average weekly bottles on ambient population and the distance to nearby stores, using the cross section of stores operating in the first week of the year. The population coefficient is positive: each additional 1000 people nearby add 49 bottles per week. The farther people live from the store, the lower is demand. Column (2) repeats the exercise using log sales as the dependent variable. Columns (3) and (4) return to panel data with regressions of log quantities on log statewide bundle prices, selected week dummies, the district pop, and the average distance. Note, again, that because of openings and closings, a store's population and distance vary over time, so in columns (4) and (6), with store fixed effects, identification comes from the changes. Here, again, higher population is related to higher demand, greater average distance reduces demand, and the price elasticity is about –0.88.

Since our data are at the store, rather than the consumer, level, we cannot explore directly the extent to which people who live further from a store choose to make fewer, but larger, shopping trips and store the product more. However, we do not observe a larger response to a price decline for stores that serve a more distant population. In regressions of the absolute or percentage change in daily sales on the day (or days) following a price decline, stores with a higher average distance to consumers experience a lower, generally statistically insignificant, increase in sales due to the drop in prices, conditioning on the amount of the price change and other demographics of the catchment area. If storage were important, we would expect a larger response with consumers in more distant areas stocking up more intensively.

We also descriptively explored the sensitivity of the results in Table 3 to some of the salient features of the Pennsylvania liquor market. First, we re-estimated specifications (3) and (4) excluding holiday weeks (Thanksgiving week and last three weeks of December) from the sample, to test whether the base results are driven by differences in willingness to pay for liquor or travel to the store in these unusual weeks. We obtained very similar results with this limited sample. Second, we explored whether demand is systematically more elastic near Pennsylvania's borders due to the easier access to alternative shopping sources. Excluding all border counties yields a more elastic demand estimate, while estimates excluding only Philadelphia and Pittsburgh suggest less elastic demand. In short, the descriptive regressions did not yield clear suggestions about the nature of systematic variation in the elasticity of demand.

Table 2 and Table 3 provide clear evidence for the mechanisms that underlie our story: having more potential customers nearby raises demand, as does their proximity to their nearest store. Higher prices reduce demand, via the demand curve. We now turn to a simple model to estimate these effects, while allowing us to predict sales under alternative store configurations.

IV. A Simple Model of Demand with Travel Cost

We seek a model that, for any set of store locations, can indicate both the demand and producer and consumer surplus from consumption by individuals in each piece of geography. The key behavioral relationships that the model must describe are a) the sensitivity of demand to consumers' distance to stores and b) the price elasticity of demand, which allows us to attach a dollar value to proximity and, in turn, to calculate the change in consumer surplus from an additional store. We use a discrete-choice model of demand for liquor at each of the PLCB's current stores to address these requirements.

Demand and Distance

There are J stores located around the state. We assume that consumers patronize the store nearest their residence. This assumption, which would arise endogenously if stores were identical in selection, given that pricing is identical across stores, divides the state into N catchment areas containing all of the population nearest to each store. We make this assumption, as well as several others, to facilitate the determination of optimal store configurations using integer programming, discussed below.

There is no price variation across stores, so the store-level demand relates the quantity sold at a store *j* in store location s (Q_{jt}) on a given day *t* to characteristics of consumers in location s's catchment area for a given level of the state-wide price. We could directly relate Q_{jt} to, say, population in area s and other demand shifters, such as the percent black or median income in the area. Table 3 reports such regressions, but they cannot be used to predict sales under a counterfactual set of stores or a counterfactual set of locations. This goal, instead, requires a model of demand at the level of geography where consumers reside. We use a discrete-choice framework to model the consumer's decision to purchase liquor. We assume that while consumers are spread across the state, they have the same utility function for liquor except for differences from different demographic attributes, including distance traveled to the store, and from unobserved tastes for the store closest to them. We model consumer *i*'s conditional indirect utility from traveling to store *j* to purchase a bottle of liquor as:

$$V_{ijt} = X_i'\beta_x - \beta_d d_{r(i)s(j)} - \beta_p p_t + \varepsilon_{ijt} = \bar{V}_{ijt} + \varepsilon_{ijt} \qquad (2)$$

In equation (2), X_i is a vector of the consumer's race and income and seasonal effects. Since we do not observe individual level attributes or purchase decisions, we aggregate consumers to demographic groups *m* whose attributes we denote by X_m. d_{rs} denotes the distance of a consumer in location *r* from store *j* in location *s*. Lastly, ε_{ijt} denote unobserved utility shifters that we assume to be distributed extreme value, resulting in Logit probabilities of the purchase incidence.

We discretize consumer locations and aggregate over the decisions of consumers in all of the locations that make up a store's catchment area. We model demand at the level of the Census tract. While the use of Census tracts – relative to finer divisions of the state such as Census block groups – introduces some measurement error into the distances consumers travel, it yields a more manageable set of 3,123 potential store locations for the simulations that follow. We consider as the potential market the population of each Census tract over the age of 21 who we further subdivide into black (*B*) and other-race residents (*O*).

Since we assume that the consumer travels to the closest store, a consumer chooses to purchase one bottle of liquor from that store provided that his utility exceeds the utility of the outside good. We normalize its value to zero. The purchase incidence of a consumer of type *m* in a particular location *r* is thus given by:

$$s_{mjt} = \frac{\exp(X_m'\beta_x - \beta_d d_{rs} - \beta_p p_t)}{1 + \exp(X_m'\beta_x - \beta_d d_{rs} - \beta_p p_t)} \tag{3}$$

where we normalize the utility from not purchasing to zero.

We calculate total demand for liquor at each location as the weighted probability of purchase scaled up by the total consumers in the location, using as weights the mass of consumers of a particular type. Total demand for liquor at store *j* is then the sum of the location-specific demands across all locations for whom store *j* is closest. We denote the set of these locations as store *j*'s catchment area, C_{jt}. Formally, C_{jt} is the set of locations *r* such that store location $s(j)$ is the closest liquor store to location *r* or $C_{jt}{:}\,\{d_{rs(j)} = \min_{s'} d_{rs'} \forall s' = 1, \dots, S, \forall r = 1, \dots, R\}$. Due to store closures and openings, catchment areas vary over time. Total demand equals:

$$\hat{Q}_{jt} = \left(\sum_{r \in C_{jt}} \sum_{m=\{B,O\}} s_{mjt}(X_m, d_{rs}, p_t | \beta) \frac{M_{mrt}}{M_{jt}}\right) M_{jt} = s_{jt} M_{jt} \tag{4}$$

where M_{mrt} denotes the total population of type *m* in location *r* and M_{jt} the total size of the market in the store's catchment area.

We estimate the parameters of the demand function using maximum likelihood. The parameter estimates maximize the likelihood of observing actual sales in store *j* and period *t*, $\hat{Q}_{jt}$, given data on the locations making up a store's catchment area, their demographics and distance from the store, and the price of liquor. The log-likelihood function is given by:

$$\ln L = -\sum_{t=1}^{T} \sum_{j=1}^{J} I(open_{jt}) \left(Q_{jt} \ln(s_{jt}) + (M_{jt} - Q_{jt}) \ln(1 - s_{jt})\right). \tag{5}$$

where $I(open_{jt})$ is an indicator of whether store *j* is open on day *t*. We identify the parameters from observing variation in the price of liquor over time (β_p), cross-sectional and time-series variation in the layout of catchment areas (β_d) and cross-sectional variation in locations' demographic attributes (β_x).

Demand Model Estimates

To keep the data manageable, we rely on a randomly drawn 10% subset of the daily data, using the list-price prior to sales as our price index for the composite liquor product.[13] As in our descriptive regressions above, we control for the seasonality by including day of the week effects, week dummies for holiday weeks (the week after New Years (week 1), July Fourth (week 26), Thanksgiving (47), and Christmas through New Year's Eve (50-52)), and additional holiday dummies for Memorial Day (May 28, 2005), days close to July 4 (June 30, July 1 – July 3), Labor Day (Sept. 3, 2005), and days around Thanksgiving (November 23 – 26). The price elasticity is thus identified from a response in sales to price changes in otherwise similar days. We allow demand by each demographic group to have its own intercept and shift it with the demographic group's per-capita income and median age. The coefficients of the estimated demand function appear in Table 4.

The estimated price coefficient translates into an average price elasticity of –1.18. The estimated distance parameter of –0.105 indicates that demand declines by 68 cents ($1.10) for every kilometer (mile) driven to the store. A travel cost of $1.10 per 1 mile of one leg of a round trip would translate into an opportunity cost of time of approximately $6.94/hour, after removing $0.405 in monetary costs (as per the IRS 2005 Standard Mileage Rates) and assuming that consumers travel approximately 20 miles per hour. In comparison, the literature suggests estimates of opportunity cost of time ranging from negligible costs (Davis (2006), to $6 (McManus (2007)), $12 (Houde (2008), and $27 (Thomadsen (2005)). Areas with higher median income have higher demand; however base demand does not differ significantly across demographic groups or age.

The second specification in Table 4 approximates cost of travel with a more flexible, quadratic distance specification. The price elasticity under this alternative specification is –1.21 and the majority of the remaining coefficients remain stable. The travel cost implied by the quadratic specification increases slightly in distance. At the mean distance of consumer to store locations, it amounts to 62 cents, similar to the one above, and ranges from 40 cents to 82 cents for the 10^{th} and 90^{th} percentile of distances traveled, respectively.

[13] For the descriptive regressions in table 3, the estimated parameters using the subsample do not differ significantly from the results obtained using the full sample of data.

In the simulations that follow, we employ the base specification with linear travel costs as our preferred specification for out-of-sample predictions. The simulations consider alternative store configurations that significantly increase the distance traveled by consumers. For example, in the 210-store configuration that we consider in section VII, the 90th percentile of distances traveled increases from 8.33 km under the current configuration to 14.35 km. To the extent that the quadratic specification accurately represents demand even at higher distances, our approach would bias our results towards finding fewer stores to be optimal.

Our base specification employs the great-circle distance between locations. Houde (2008) and others propose the use of distance traveled on the shortest route along existing streets as an alternative proxy for the consumer's travel cost. To investigate the sensitivity of our demand estimates to the use of straight line distance, we computed the route distances between all PA Census tracts and the existing stores using MPMileage.

Driving distance is, not surprisingly, systematically larger than straight-line distance. If driving distance were literally proportional to straight-line distance, then our choice of a distance measure would have no effect – up to a scalar - on our results. They are not literally proportional, but they are close to proportional.

We can compare the two measures of the distance from each tract to its nearest store distance measures by regressing driving distance on straight-line distance. First, we use all 3125 tracts, including those whether the nearest store is in the same tract and therefore – according to the measures – at zero km from the population. We also can run a regression using only the 2636 tracts that do not contain a store, and therefore with positive distances to the nearest store. Both regressions confirm both the strong positive relationships between the two distance measures, as well as the good fit (r-squared is 94 percent in the specification using all of the data and 93 percent in the specification using only the positive distances). The coefficient on straight-line distance is about 1.4 in both regressions, indicating that each additional km of straight-line distance adds 1.4 km of driving distance. Both regressions also include positive and statistically significant constant terms. The two specifications indicate that, on top of the aspect of driving distance that is proportional to straight-line distance, driving distance contains an additional 0.2 or 0.3 km, or systematic deviations from proportionality.

However, these deviations are relatively minor: At the lower end of the inter-quartile range of straight-line distance (0.77km), the constant term makes up 17 percent of the implied driving distance (0.17 = 0.226/(0.226 + 1.434*0.77)). At the upper end of the inter-quartile range (4.14km), the constant makes up 3.7 percent. At the median distance of 1.83, the constant makes up 7.9 percent.

We re-estimated our demand model assigning each consumer to its closest store based on driving distance and using driving distance to proxy travel costs. Specification (3) in Table 4 contains the results. The alternative distance metric has little effect on the estimated price sensitivity, which is –1.28 in this case. The estimated distance coefficient is smaller than the straight-line distance parameter, reflecting that driving distance is typically larger than straight-line distance. We estimate an implied travel cost per km of 40 cents for this specification, which is similar to the implied travel cost under the straight-line distance measure (68 cents) scaled down by the factor of proportionality of 1.4 above, or 48 cents.

In the following simulations, we continue to rely on straight-line distance. Some of our counterfactual analyses below require data on the distance between all 3,135 Census tracts in the state. Computing route distances for as large a set of location pairs is computationally challenging. In future work, we plan to investigate further the sensitivity of our results to the use of straight-line distance by conducting some of the less computationally intensive simulations with the alternative travel distance measure.

A final worry with our travel cost estimates is that they result from cross-sectional variation, as well as from variation in store opening hours over the course of the week. Our estimates may be biased to the extent that distance is correlated with other location-specific shifters of alcohol demand that we do not include in estimation. As a robustness check, Specification (3) also includes the number of churches per capita in the tract as a further, location-specific, proxy for the residents' underlying preference for alcohol consumption. This does not appear to affect the parameter estimates significantly, with the possible exception of the coefficient on the black population, which changes signs and becomes significant. Since it is difficult to imagine valid instruments for store locations, we also investigate the sensitivity of one of our simulations to alternative levels of the distance and price coefficients. The results of this robustness check are described below.

Welfare Measures

To evaluate openings or closures of stores and changes in store locations, we need to compute the welfare benefit of alternative store configurations. Our model shows how much the quantity purchased by persons in each location (and, by extension, the quantity sold at each store) changes with the distance to the closest store. Opening a store near location r has two effects on consumer welfare: First, consumers in location r who already purchase face a lower effective price (inclusive of travel). Second, facing the lower effective price, a larger share of consumers in r purchase, generating additional consumer surplus.

For the chosen specification, consumer surplus for consumers in location r is given by:

$$cs_{rst}(X_m, d_{rs}, p_t|\beta) = -\frac{1}{\beta_p}\left(\sum_{m=\{B,O\}} \ln\left(1 + \exp\left(X_m'\beta_x - \beta_d d_{rs} - \beta_p p_t\right)\right) M_{mrt}\right) + \mathrm{C} \quad (6)$$

if the consumers in location r are served by a store j in s (see Small and Rosen (1981)). The consumers in location r generate producer surplus to the store, that is the usual variable profit prior to fixed cost, based on the markup of the retail price p_t over the wholesale price c_t:

$$ps_{rst} = (p_t - c_t)s_{rt}M_{rt}. \quad (7)$$

The total welfare generated by store j in location s is therefore:

$$W_{st} = \sum_{r\in C_{jt}} cs_{rst}(X_m, d_{rs}, p_t|\beta) + ps_{st} - K, \quad (8)$$

where K denotes the daily fixed cost of operating a store.

Comparing Alternative Entry Patterns

To assess the goals underlying the PLCB's store configuration, we derive several benchmark configurations, including the store layout chosen by a profit-maximizing monopolist and a benevolent monopolist. These rely on the $R \times R$ matrix Y of consumer location to store location matches. We define Y_{rs} to be one when demand in location r is served by a store in location s, and zero otherwise. The Y matrix also indicates J, the total number of stores operating, as $\mathrm{trace}(Y) = \sum_{s=1}^{R} Y_{ss}$. We thus assume in our simulations that demand and supply locations are the same and consist of Census tracts.

For a given store configuration, the profits of the system are then:

$$\Pi = \sum_{s=1}^{R} \sum_{r=1}^{R} ps_{rs} Y_{rs} - K \sum_{s=1}^{R} Y_{ss}, \tag{9}$$

The profit-maximizing monopolist's problem is to find the store configuration that maximizes profit, while the benevolent monopolist's problem is to find the configuration that maximizes welfare.

Approaches to Optimization

This problem can be posed as an integer programming problem in the following way.[14] The planner's maximand, either producer surplus less fixed costs or full surplus less fixed costs, may be expressed as follows:

$$\max_Y \Pi = \sum_{s=1}^{R} \sum_{r=1}^{R} ps_{rs} Y_{rs} - K \sum_{s=1}^{R} Y_{ss} \tag{10}$$

subject to

$$\sum_{s=1}^{R} Y_{rs} = 1 \quad \forall r, \tag{11}$$

$$Y_{ss} \geq Y_{rs} \quad \forall r, s, r \neq s, \tag{12}$$

$$Y_{rs} = \{0,1\} \quad \forall r, s. \tag{13}$$

The maximand (10) for this problem includes two parts. The first, $\sum_{r=1}^{R} \sum_{s=1}^{R} ps_{rs} Y_{rs}$, is the producer (or total) surplus that results from a particular configuration of stores. Given a set of store locations and the rule that demand is assigned to its closest locations, producer (or total) welfare is determined. The second part of the maximand is simply the fixed cost of operating the chosen number of stores.

Constraint (11) indicates that each demand location must be assigned to a single store location. Constraint (12) prevents the assignment of demand to locations without a store. Constraint (13) makes the assignment of demand to supply binary: each demand location is either served by a particular supply location, or not.

Solving this optimization problem is difficult because of the sheer number of possible store configurations. That is, there are 2^R possible configurations of stores to evaluate. Operations researchers have developed efficient algorithms, such as "branch and bound," for

[14] Our problem is closely related to facilities location problem analyzed in Perl and Ho (1990). See also Daskin (1995).

solving problems of this sort.[15] We are able to rely on these sophisticated integer programming algorithms to solve problems of moderately large size.

The maximization problem requires values of ps_{rs} as inputs. That is, we need to know the amount of demand that each demand location would contribute to each store in each possible configuration. Because our demand model assigns each demand location to its nearest store – and because we use the same potential locations for demand and supply, ps_{rs} is easily calculated in advance of the optimization. This would not be the case if we allowed consumers to choose not only whether to purchase liquor, but also from which store to purchase in a multinomial choice model of demand. Then a store's demand from any location would depend not simply on the distance between the store and demand locations but rather on the entire configuration of stores. That is, each ps_{rs} would depend on the entire 2^R configuration. Hence, our simple model of the purchase decision facilitates the calculation of optimal store configurations.

Integer programming approaches are strained by the problem of locating stores throughout the state's 3000 tracts. "Greed" algorithms provide intuitive and less computationally burdensome approaches. We implement such an algorithm, which we term "sequential myopic entry" (SME), as follows. Beginning from a first location that maximizes its standalone profits (or welfare), we keep adding stores that maximize incremental profit (or welfare), holding the previous stores' location fixed, until the marginal profit or welfare of the incremental location falls to the fixed cost of an additional store.

Define as the vector of store locations L the main diagonal of matrix Y. The incremental profit associated with adding a store (going from some store configuration L^0 to another configuration L^1, where L^1 has one additional store operating) is:

$$\Delta\Pi(L^1, L^0) = (p - c) \sum_{s=1}^{R} \sum_{r=1}^{R} \left(\hat{Q}_r(d_{rs}) Y_{rs}^1 - \hat{Q}_r(d_{rs}) Y_{rs}^0\right) - K, \tag{14}$$

while the change in total welfare is:

$$\Delta W(L^1, L^0) = \frac{-1}{1+\beta_p} \sum_{s=1}^{R} \sum_{r=1}^{R} (cs_{rs} Y_{rs}^1 - cs_{rs} Y_{rs}^0) + \Delta\Pi. \tag{15}$$

Define L_1^* as the vector containing one store that produces more profit (or welfare) than any other single-store configuration and $L_2^* | L_1^*$ as the vector containing two stores that

[15] We employ LINGO 11.0 to solve these problems.

maximizes profit by adding one store to L_1^*. $L_2^*|L_1^*$, the optimal two-store configuration under sequential myopic entry, maximizes $\Delta\Pi(L_2|L_1, L_1)$. The process continues with calculation of $L_j^*|L_{j-1}^*$ and so on, until the increment to profit falls below zero. The welfare analogue $L_2^*|L_1^{*W}$ maximizes $\Delta\mathrm{W}(L_2|L_1{}^W, L_1{}^W)$.

$L_j^*|L_{j-1}^*$ is not in general the same as the j-store configuration L_j^* that simultaneously maximizes the profit available from j stores. Sequential myopic entry overstates the benefit of each entrant because its marginal benefit is – myopically – predicated on the $j-1$ stores already operating rather than the total number that will ultimately operate. When the last store has been added, the marginal benefits of the infra-marginal stores are smaller than they were when they were marginal. To assess the magnitude of such biases, we compare results under sequential myopic entry with the simultaneous optima for small areas where these can be calculated.[16]

Free Entry

In addition to examining profit and welfare maximizing store configurations, we would also like to explore configurations that would arise under free entry. The usual condition for equilibrium with free entry with symmetric firms is that the *J* firms in *L* are profitable while *J+1* are not.[17] Here, because of the vagaries of geography, equilibrium is more complicated. Each firm (store) must be profitable; there must be no room for further entry; no firm may wish to switch its location.

While the two-firm game has a unique pure strategy Nash equilibrium, we can verify that the simultaneous entry game involving 3 players has no pure strategy Nash equilibria. This raises a question of how to model counterfactual free entry. Three strategies seem sensible.

First, for small numbers of stores and possible store locations, we can calculate each store's profits for every configuration. This allows us to directly calculate each store's best response function mapping other stores' locations to its best choice. We can then look for

---

[16] A natural alternative to SME is the reverse: sequential myopic exit. This is implemented by saturating geography with stores, one in each of R locations, and then removing the store whose withdrawal minimizes $\Delta\,\Pi(L_R|L_{R-1}, L_{R-1})$ or $\Delta\,\mathrm{W}(L_R|L_{R-1}{}^W, L_{R-1}{}^W)$. This procedure is repeated until the incremental benefit of the least beneficial store rises to the fixed cost *K*. The procedure is myopic in the sense that a store gets eliminated based on its marginal benefit when all currently existing stores are assumed to continue to exist even though many will be eliminated later in the procedure.

[17] This is the condition for equilibrium in homogeneous goods entry models such as Bresnahan and Reiss (1991) and Berry (1992). Entry models dealing with product positioning include Mazzeo (2002) and Seim (2006).

configurations on all of the stores' best response functions. As mentioned, this problem does not generally have a pure strategy Nash equilibrium, so this does not seem promising.

Second, we can employ a sequential myopic algorithm analogous to those introduced above, although some adaptation is needed for free entry. First, find the location that maximizes a lone store's revenue. If this location is profitable, it remains. The second store locates at the location that generates the most profit for the second store, given the location of the first store. That is, the second store locates at its best response, evaluated given the first store's location. If either is unprofitable, it is withdrawn. Then another store locates at the most profitable available location, and so on. The process ends when there is no location for profitable entry, and each existing store is profitable. This deviates from Nash equilibrium because the stores, while profitable, might be more profitable if they switched locations. Only the last entrant is necessarily on its best response function. Still, the algorithm shows – approximately – how many stores free entry could support. (That said, if inframarginal stores moved and diverted business from competitors, then the market might not be able to support as many as the algorithm suggests. It is also possible, though, that location switches could expand the market, allowing more stores to fit.)

So far, the algorithm described stops as soon as two conditions hold: a) all existing outlets are stand-alone profitable and b) there is no room for profitable entry. But we can take this algorithm a step farther by continuing to see whether it "converges" to a small set of possible configurations satisfying a) and b).

This algorithm is clearly neither fully rational nor – as a result – fully optimal. When stores enter, while they find the best location, given existing entry, they do not anticipate how subsequent entry will affect the profitability of locations they choose. Entrepreneurs are therefore somewhat sentient. They enter where it is currently most profitable and continue operating until they are rendered unprofitable by other entry that is, by definition, unforeseen by them. Still, it seems reasonable to expect, if J simultaneously operating stores are profitable, that the free entry equilibrium has at least J stores. Even this simple algorithm is somewhat computationally challenging since in each iteration, we must check the profitability of each store (rather than just the entire system).

Our use of myopic algorithms to characterize the number of stores operating under free entry is reminiscent of the approach of evolutionary game theory to equilibrium. In evolutionary game theory, agents are assumed to be myopic and inertial. They respond weakly to incentives, changing to better strategies when opportunities to change arise. These processes are allowed to unfold over multiple periods, giving rise to characterizations of equilibrium. See Samuelson (1997) or Fudenberg and Levine (1998) for extensive discussions of evolutionary game theory.

V. Results

Incremental Benefits of Existing PLCB Stores

We can get some insight into the relationship between the current configuration and optimality by comparing the marginal daily benefit of each of the actual stores with our estimate of daily fixed costs *K*. In a profit-maximizing system, for example, each store's increment to profit would equal – or, with integer constraints, exceed – *K*. If the marginal increments are systematically lower, then there are too many stores. Figure 6 depicts each store's incremental contribution to profit (calculated as the difference in the system's profitability with and without the particular store under our demand estimates), each store's incremental contribution to total welfare, and each store's stand-alone variable profit. While the daily cost of operating a store is $1,110, the mean daily incremental producer surplus is $446, and the median is $340. The mean – and most of the distribution of – incremental profitabilities being below *K* clearly implies that the state has more stores than would maximize profit.

While the state does not appear to be maximizing profit by its store configuration, it is possible that the state seeks to maximize welfare. This would be implemented – without integer problems – if each store's incremental total surplus equaled the cost of operation. The mean incremental welfare is $1,048, and the median is $806. These are much closer to our rough estimate of daily *K*. This simple analysis, which relies only on marginal analysis, indicates that the way the state operates its stores is closer to welfare maximization than to profit maximization.

We can also compare the stores' total, rather than incremental, producer surplus to fixed cost. Under the current system, the mean total producer surplus is $2,122, and the median is $1,892. Because the marginal stores' variable profits substantially exceed *K*, the system appears to have fewer stores than would operate under free entry. The mean incremental benefit of a store in the current configuration is, of course, a statistical estimate that is a function of the demand model parameter estimates. The marginal analysis indicates that the state operates too many stores for profit maximization but fewer than would operate under free entry. The marginal benefits under welfare maximization are closest to the fixed cost estimate. While the analysis indicates directions of welfare or profit increasing reform, the marginal analysis does not directly inform the correct system size for profit or welfare maximization.[18]

Profit and Welfare Maximizing Entry in Five Counties

Taken literally, our demand model can indicate the configuration of stores that maximizes an outcome of interest such as profit or welfare. We begin with distinct pieces of geography – counties within Pennsylvania - for illustrative purposes. We choose five counties that satisfy three conditions: they contain an interior population center and a low-density periphery, they are not too large for our integer-programming algorithms (effectively, less than half a million in population), and they are not on the state border.

For illustrative purposes, we begin with Blair County. Blair County, in the south central part of Pennsylvania fully contains Altoona, PA and is otherwise not urbanized. The county has 129,144 people in 34 Census tracts, and the population is concentrated away from the periphery of the county. The county covers 527 square miles. The PLCB operated six stores in the county during the sample period. Figure 7 shows both the locations of the Blair County stores, as well as the population density of Census tracts in Blair County.

The four other counties are Berks (population 373,638), which contains Reading; Lancaster County (population 470,658), which contains the city of Lancaster; Lycoming (population 120,044) with a population concentration in Williamsport; and Schuylkill (population 157,342), which contains Pottsville.

Simulations require particular values for model parameters. We set p=\$12.38, marginal cost c=\$7.31, and K = \$1,110 (per day). Parameter values for demand are taken from our estimated model. We have two possible maximands for the social planner: profit and social welfare. We begin with profit, and we find optimal configurations using integer programming.

[18] Our calculation of marginal benefits of existing stores recalls the literature on tax reform, where local elasticity estimates allow determination of welfare-increasing directions of tax reform. See Ahmad and Stern (1984).

Table 5 reports the number of stores operating, along with the quantity sold, profits, and welfare, under different possible store configuration, including the actual network (characterized by locating outlets at centroids of tracts containing actual stores), the profit maximizing configuration, and the welfare maximizing configuration. In Blair County, the actual system includes six stores and generates daily profit of $3,456 and daiy total welfare of $110,766. Profit maximization would be achieved with two stores, which together generate $6,720 in daily profit while generating $112,427 in daily welfare. A six-store configuration maximizes welfare, producing $4,998 in daily profit and $114,365 in daily welfare.

The state currently operates six stores in Blair County. This is more than the model's profit-maximizing number of two stores (although in real life, the county's stores get some demand from outside the county, whereas they do not in the Blair-only simulation). The state's actual configuration – see Figure 7 – is slightly more geographically compressed around Altoona than the welfare-maximizing configuration. Based on entry decisions in Blair County, the state's objectives are better described by welfare maximization than profit maximization. We calculate the free entry equilibrium using the sequential algorithm below.

Results for the five counties indicate that the profit maximizing number of stores is always substantially below the welfare maximizing number, usually by a factor of close to three. The actual number of stores is generally much closer to the welfare maximizing number than the profit maximizing number.[19]

Comparing Sequential and Simultaneous Entry

How do the simultaneous solutions calculated via integer programming compare with solutions calculated with myopic ("greedy add") algorithms. The question is of interest because the sophisticated algorithms not implementable for large problems, such as the problem of choosing the optimal configuration for the entire state (which contains over 3000 Census tracts

[19] These simulations assume that the fixed store operating costs are identical across tracts and across counties. We also ran the simulations for the five county markets assuming that a given store's fixed costs are proportional to its tract's median rent, as a proxy for local rental costs, and that fixed costs continue to sum to the PLCB's total store operating costs across stores. The constant fixed cost measure systematically overstated predicted costs for Schuylkill county based on median rents, resulting in fewer stores being optimal than under the variable fixed cost setup. Allowing for location-specific fixed costs also led to some location adjustments within markets. In aggregate, however, the results are largely unaffected by this modification, in terms of the overall welfare and total profit associated with the welfare maximizing configurations and the optimal number of stores.

and – currently – over 600 stores). Table 6 compares the profit and welfare maximizing configurations calculated via SME.

Table 6 reports quantity, net welfare, and profit for profit-maximizing and welfare-maximizing configurations determined via SME. The last three columns report the percent deviation between the SME and the simultaneously optimal values (from Table 5). The myopic algorithm produces similar answers. For example, the maximum profit is within two percent in all five counties, and the maximum welfare is within one percent.

Robustness to Alternative Demand Estimates

The simulations in this paper, by their nature, rest solidly on two estimated parameters, the price elasticity of demand and the distance parameter. This section explores the sensitivity of the paper's basic results to different values of these parameters. Because some of the simulation exercises we undertake are computationally intensive, we focus here on the incremental exercises, comparing the average incremental benefit of a store to average fixed costs, that are easy to compute.

To explore different values of the price elasticity and distance parameters, we re-estimate the basic demand model, with specifications embodying each of the following constraints: 1) twice the baseline price coefficient, 2) twice the baseline distance parameter, 3) half the baseline price coefficient, and 4) half the baseline distance parameter.

Figure 8 shows the resulting distributions of incremental producer surplus, incremental total surplus, and variable profit. Table 7 reports means and, below them, medians. Under all of the alternative parameter values, the mean incremental PS remains below the average FC (of $1,117), indicating that even with rather different parameter estimates, the current configuration has more stores than would maximize profit.

The results for incremental total surplus are rather different. With a higher distance parameter or a lower price elasticity, the average incremental total surplus would be higher than the average fixed costs, indicating that the system has too few stores than would maximize welfare. A smaller distance parameter or a higher price elasticity suggests the opposite conclusion, that the current system is too large to maximize welfare. Broadly, then, the sensitivity analysis supports the baseline analysis. The range of estimates does not rule out the current configuration as welfare-maximizing, while it does appear to rule out the current configuration as profit maximizing.

VI. Free Entry and Social Inefficiency

In our context – and in our model – prices and markups are fixed. The price-reducing mechanism usually operating with free entry is absent, so the model will likely generate more stores than would actually operate if entry were truly unregulated. Hence, the unregulated free entry configuration from the model should be viewed either as an upper bound on free entry N, or as simulations of a fixed-price regime, as might operate if the state regulated prices with an optimal Pigouvian tax.

Table 5 reports features of the free entry configurations in each of the five counties. Unconstrained entry would allow 30 stores in Berks County, twice the actual number. Welfare under free entry falls short of maximized welfare by 11 percent of revenue. The extent of inefficiency from free entry varies across the remaining four counties: relative to maximum welfare, free entry produces 7, 7, 10, and 12 percent less, respectively, in Blair, Lancaster, Lycoming, and Schuylkill counties. At the same time, the results suggest that free entry dissipates between 28 and 58% of monopoly profits.

The welfare losses from free entry have two distinct components. First, free entry can produce the wrong number of outlets. We see that here. Free entry generates too many outlets in three counties and too few in the other two. Second, even given a number of stores, free entry can put outlets in the wrong locations. Separate from the effects of network size, we can simulate the pure effect of location choice under free entry by comparing free entry configurations to other systems of equal size. One natural benchmark is the actual number of stores. We run our free entry algorithm under a range of hypothetical taxes (additional fixed costs) to get free entry configurations with the actual number of stores.[20]

Table 8 compares welfare-maximizing and free entry configurations with the actual number of stores in each county. The welfare maximizing configuration with the actual number of stores is determined with an integer program that finds maximum welfare for a configuration

[20] In these simulations, while the condition for profitability – and therefore our equilibrium – involves a comparison of revenue with tax-inclusive fixed cost, the welfare calculation is predicted only on actual fixed costs. Thus, the taxes are losses to the firms and gains to taxpayers and are neutral in welfare calculations.

of given size. Holding network size fixed free entry produces welfare 5, 14, 4, 13, and 5 percent below maximum welfare in the five counties, respectively. Future work will investigate the sources of the heterogeneity in welfare losses, focusing on differences in the geographic concentration of demand across counties.

While our procedure for determining the free entry configuration is intuitively reasonable, its properties are not known. Recall that there exists no pure strategy Nash equilibria for the entry game. Our goal is to determine the efficiency properties of plausible free entry equilibria, and we are concerned that our iterative procedure converges to an arbitrary configuration, rather than one representative of what might prevail under free entry. To explore this, we tried alternative starting values. That is, while our baseline procedure begins by locating the most profitable standalone store, we would alternatively begin by locating a store in any – as opposed to the myopically best – location.

Using Blair County as a test case, we ran our free entry algorithm locating a first store in each of the county's 34 tracts, producing a distribution of configurations resulting from the algorithm. In 25 of 34 cases, the algorithm converges to the same configuration, with output of 2324 and net welfare of 112,321. The full range of output values runs from 2306 to 2329, a range of about one percent relative to the average. The full range of net welfare values runs from 112,114 to 112,389, or a range of roughly a quarter of a percent relative to its average. We take this as evidence that our free entry procedure produces characterizations of output and welfare that are representative of what would prevail with free entry.

VII. Whole State Simulations

The similarity of simultaneous and sequential configurations for the five counties provides some justification for using the SME approach for larger possible systems. Given a list of possible store locations and a list of possible sources of demand, a maximand (profit or welfare), and a store operation cost, the model and the SME algorithm can determine a list of stores to operate. For any list of possible store locations, we begin by calculating the revenue (welfare) accruing to a lone store if it were in each of the possible locations. Having found the best location for a single store, we then add a second store in the location than maximizes the increment to profit or welfare. We repeat the process until the incremental benefit of the last store falls to the cost of operating an additional store. We then check whether inframarginal stores remain economic when the marginal store becomes uneconomic. (For computational reasons we cannot implement our free entry algorithm on the whole state).

We locate stores around the state "starting from scratch" using, as above, the geographic center of each of the 3,123 Census tracts as possible store locations. Using SME based on total surplus, the welfare maximizing number of stores is 481, generating daily welfare of $11.50 million and profit of $812 thousand. . SME based on producer surplus gives 210 as the profit-maximizing number of stores. The profit-maximizing system produces $922 thousand in daily profit.

The SME algorithm adds locations that maximize incremental benefit (profit or welfare), but it does so myopically, that is, without cognizance of the locations added subsequently. As a result, the marginal benefit of each store when added will exceed its marginal benefit at the optimum. It is possible that stores added by the algorithm will add less to the maximand than their cost, when their incremental benefits are evaluated in the optimal configuration. In short, SME could lead to more stores than would actually maximize profit. In practice, this is rare in our application: only 3 of the 210 stores a profit-maximizing monopolist would operate under SME are unprofitable ex-post, while 10 of the 481 welfare-maximizing stores under SME generate negative net welfare ex-post.

VIII. Welfare Weights Rationalizing Existing Entry Patterns

Our results so far imply that while the state comes closer to welfare than profit maximization in its choice of number and location of stores, discrepancies remain. A possible explanation lies in the state valuing different types of consumers differently.

To uncover the weight that the planner attaches to different types of households, we begin by assessing the welfare effects of each store given parameter estimates. We withdraw one store from the system and compute the demand, profit, and consumer surplus for each tract in the state. Then we re-introduce the store and calculate quantities and welfare again. The sum of the increases in welfare across the tracts in the affected zone show the increase brought about by the store. Repeat this exercise for each store in the current configuration, one at a time. This provides estimates of ΔW_s for each store s.

Suppose the planner locates stores to maximize welfare given the number of stores operating, based on welfare function $U(W_1(.), \ldots, W_S(.))$. Opening the *i*th store raises social welfare by $\Delta U_i \Delta W_i$. Assuming that each store is equally costly to operate, welfare maximization implies that $\Delta U_i \Delta W_i = c \;\; \forall i$. Because we can calculate ΔW_i, we can observe ΔU_i, up to a factor of proportionality. We can explore the weights that the state attaches to types of people by relating ΔU_i to characteristics of the zone, via a regression.

Table 10109 reports results, using as dependent variable each store's incremental total welfare. In effect, we are inferring the weights that render the state's store location decision welfare maximizing. Implicitly, the state places higher weights on blacks and lower weights on higher-income people and people in more populous areas. How plausible are these estimates? One source of insight into the planner's objectives is the average distance that blacks and whites live from stores. Using black population in each block group as weights, the average distance to the nearest liquor store is 1.3 kilometers, while the analogous average for non-black population is 3.8 kilometers. The stores are sited to be, on average, three times closer to black people. This is consistent with the state's willingness for the stores with black clienteles to have smaller incremental welfare.

IX. Conclusion

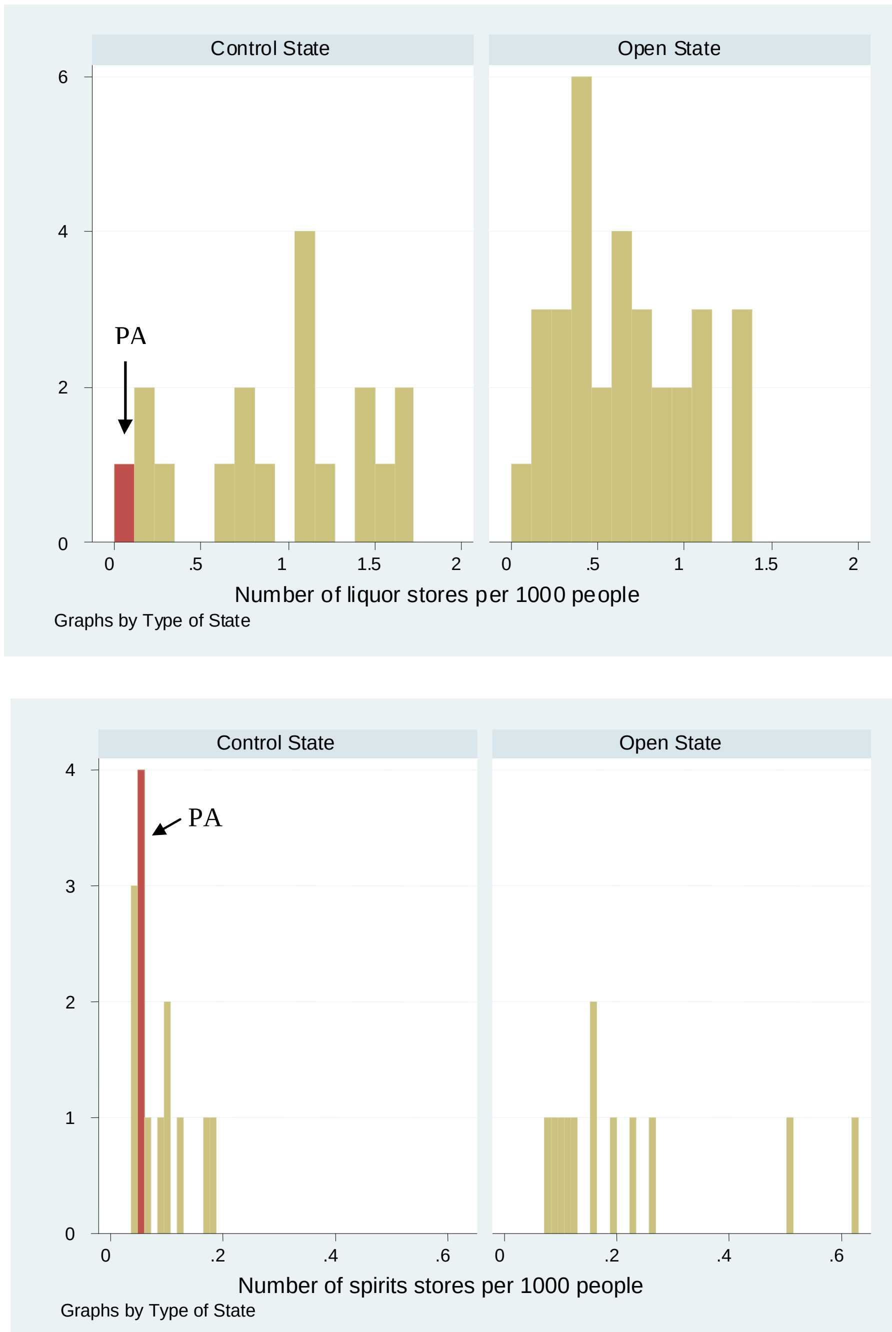
It is well known in theory that free entry results in too many stores, possibly clustered too close together. Monopoly solves some of these problems. A profit-maximizing monopolist internalizes business stealing and so neither operates too many stores nor clusters stores too close together. Even with pricing capped, a monopolist values only additional producer surplus and not consumer surplus created by customers' proximity to stores and therefore operates too few stores. We have studied the state of Pennsylvania's liquor retailing monopoly. We find that it operates more like a benevolent monopoly than a profit maximizing one. That is, it operates far more stores – roughly three times more – than would maximize profit and much closer to the quantity that would maximize welfare.

We have separately analyzed the welfare properties of free entry with geographic differentiation with particular attention to the welfare loss due to wrongly placed products. We find that these losses can be large, ranging from 3.6 to 13.7% of total net welfare in the counties we study. This compares to overall losses in net welfare of 6.9 to 11.5% from free entry overall, which includes welfare losses due to excess entry in addition to incorrect product placement.

A number of important caveats are in order. It should be noted again that the simulations in the paper take prices as given. We also note that our simulations take current store costs as given. If the PLCB employs too many workers or otherwise has excessive costs, then the true value of K is lower than the value we have used, and the optimal number of stores would be larger than we have calculated, under each of the regimes. This is essentially a case study of one government entity, so we cannot say whether results here apply to other entities. Still, we hope that our study provides some insight into the rationales underlying entry decisions in a market run by a government monopolist, and we think additional work on other contexts would be fruitful.

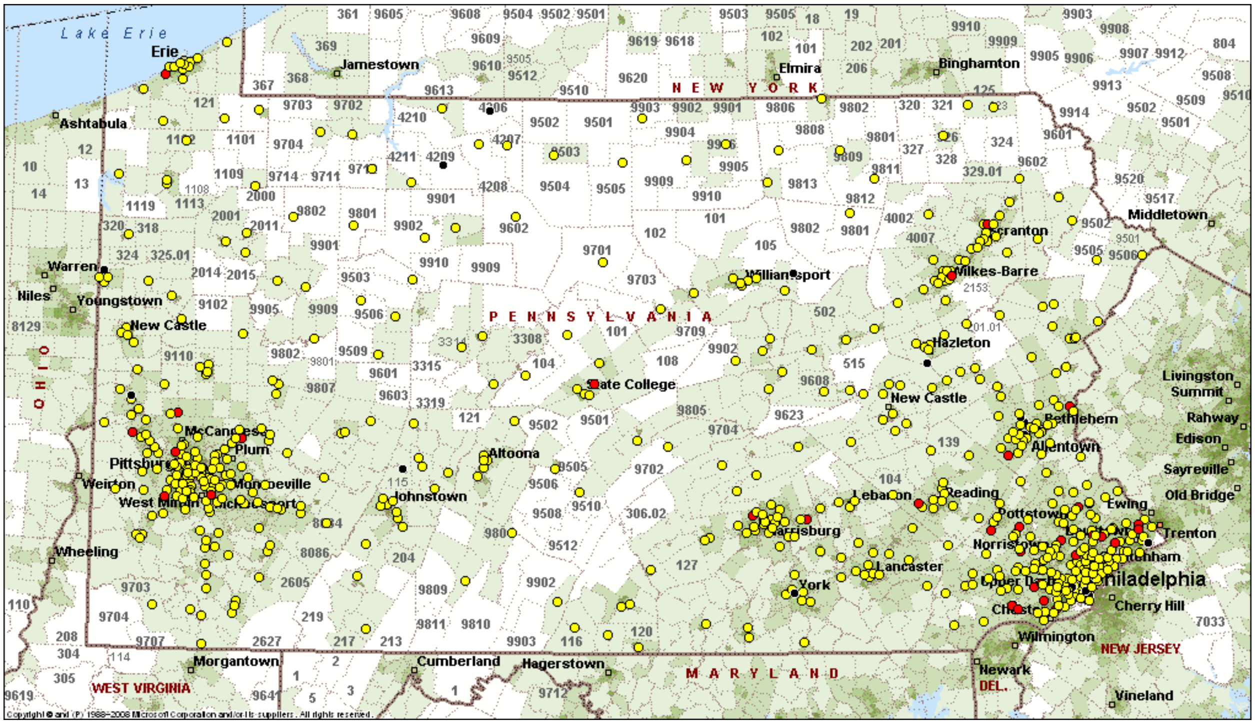
Figures and Tables

Figure 1: Number of Liquor Stores per Capita, US States, 2005



Source: Adams Beverage Group, “Adams Factbook 2006,” Norwalk, CT (2006).

Figure 2: Wine & Spirits Stores in Pennsylvania, 12/2005



Note: Black dots indicate stores that close between 1/2001 and 1/2006, red dots indicate stores that open during the timeframe, and yellow dots all remaining stores.

Figure 3: Weekly Average Number of Bottles Sold per Store, 2005

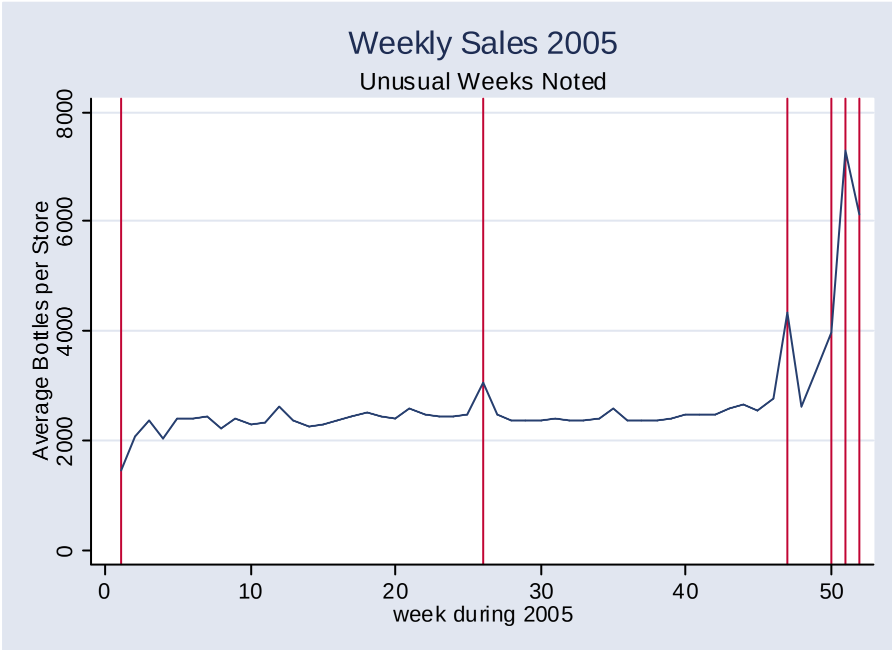


Figure 4: Price Index, 2005

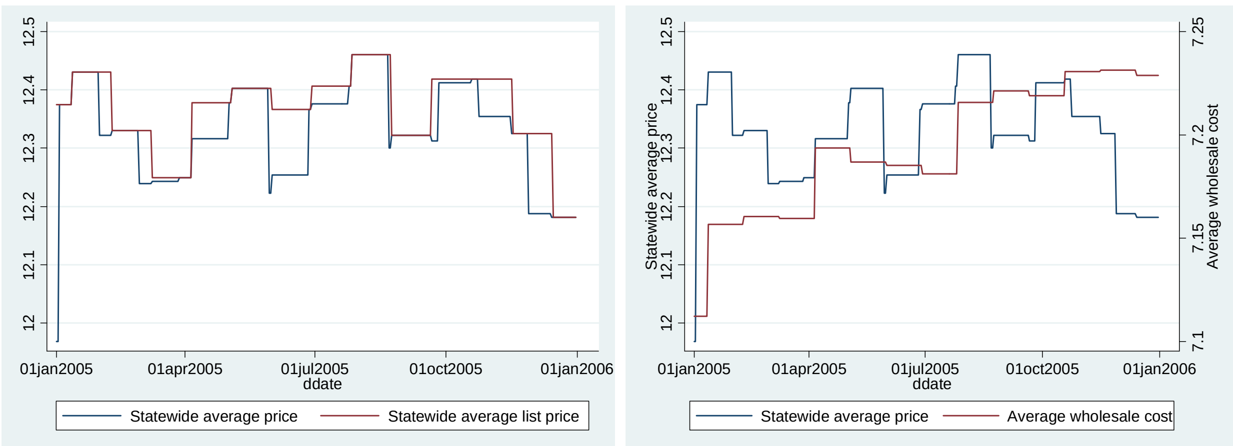


Figure 5: Sales Volume vs. Catchment Area Population

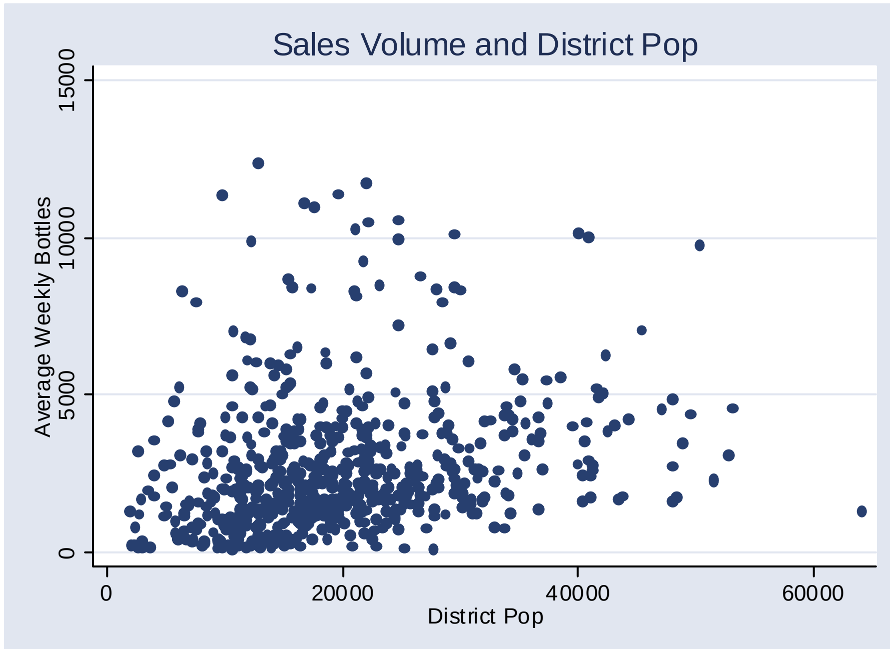


Figure 6: Incremental Benefit of Each Store in Current Configuration

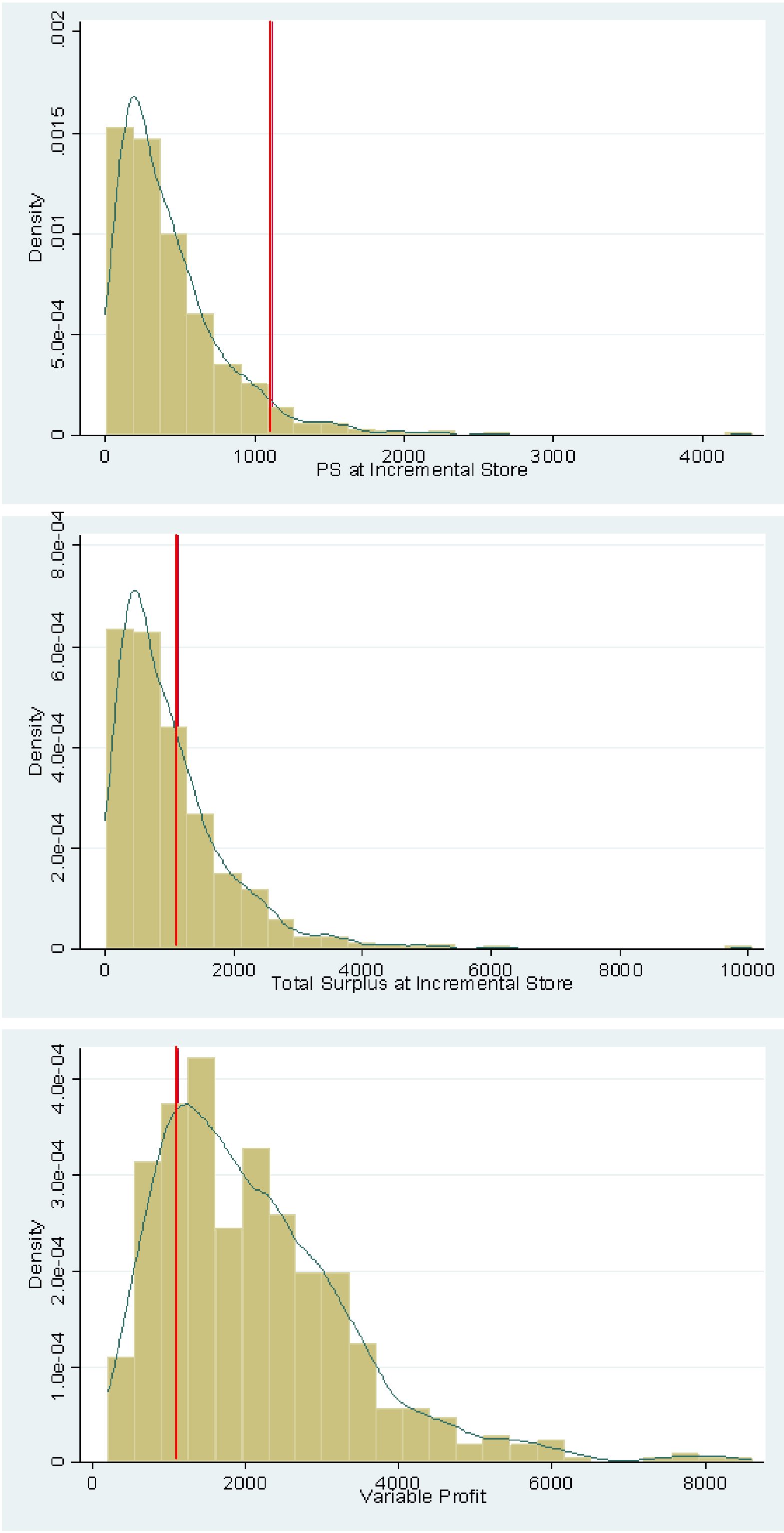


Figure 7: Actual PLCB Stores in Blair County

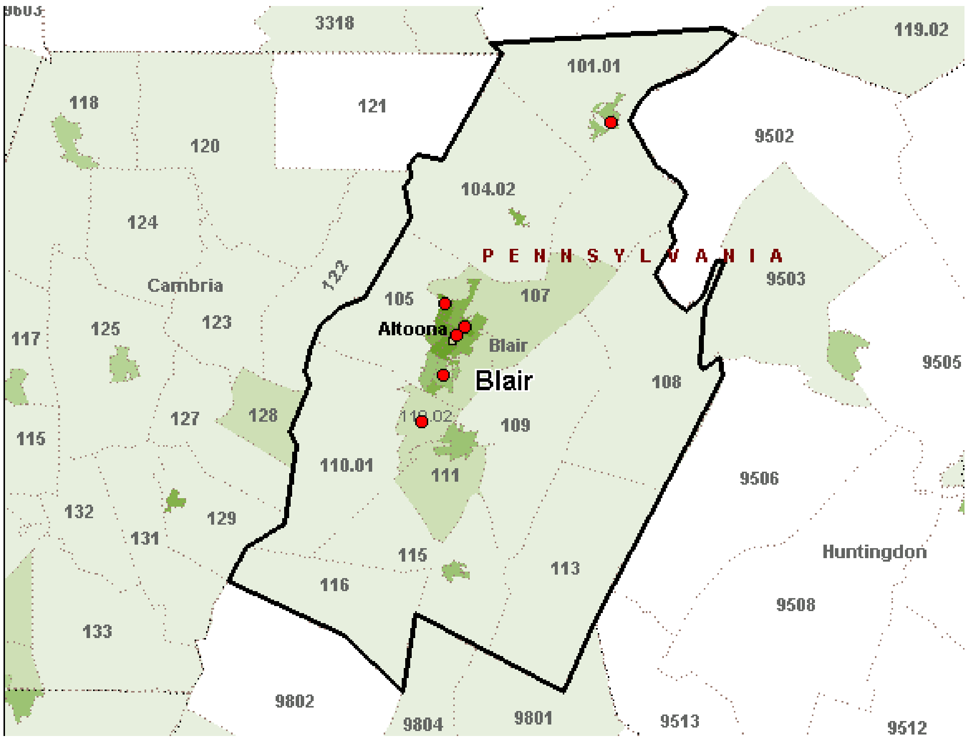


Figure 8: Incremental Welfare Measures under Alternative Demand Specifications

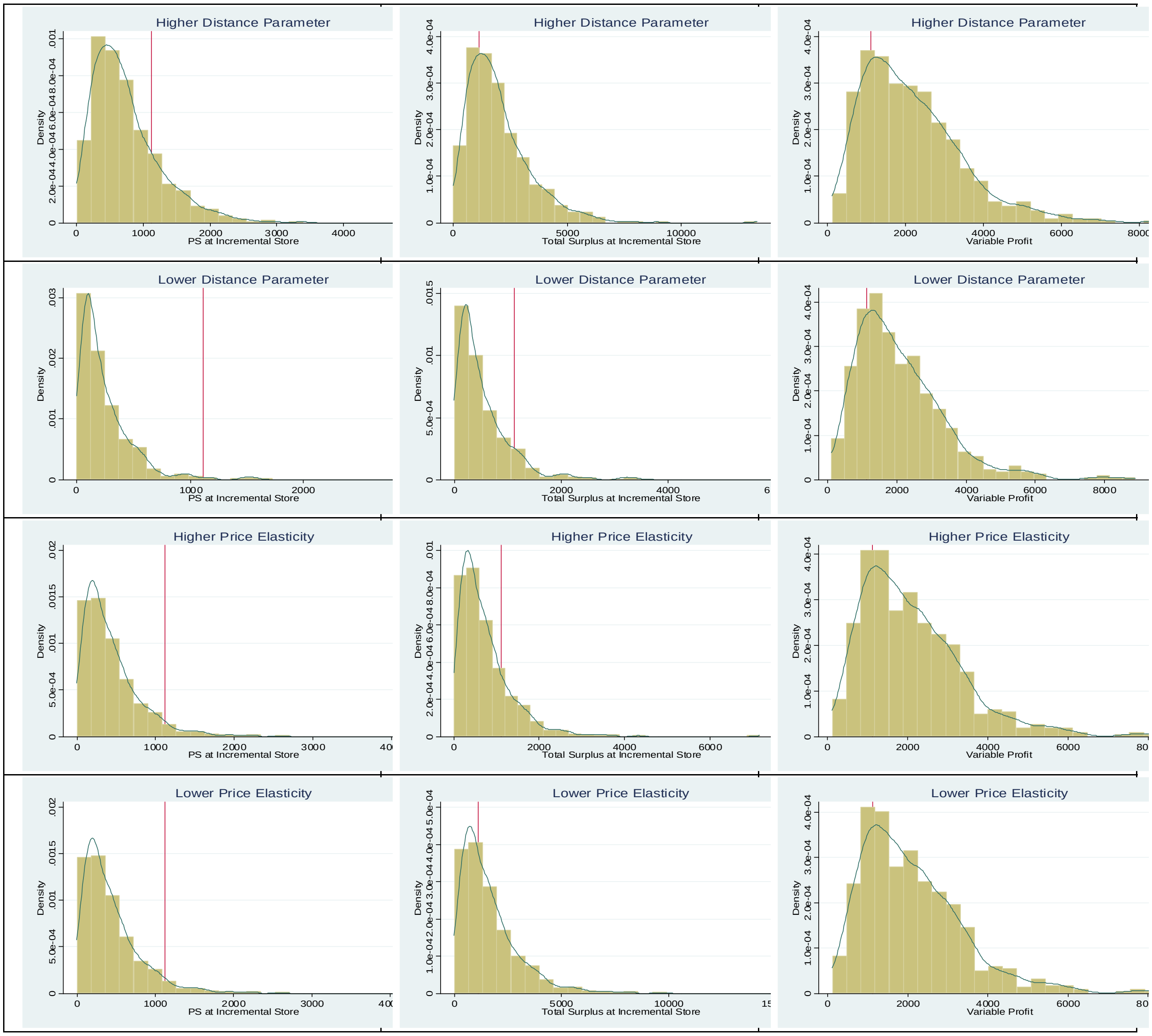
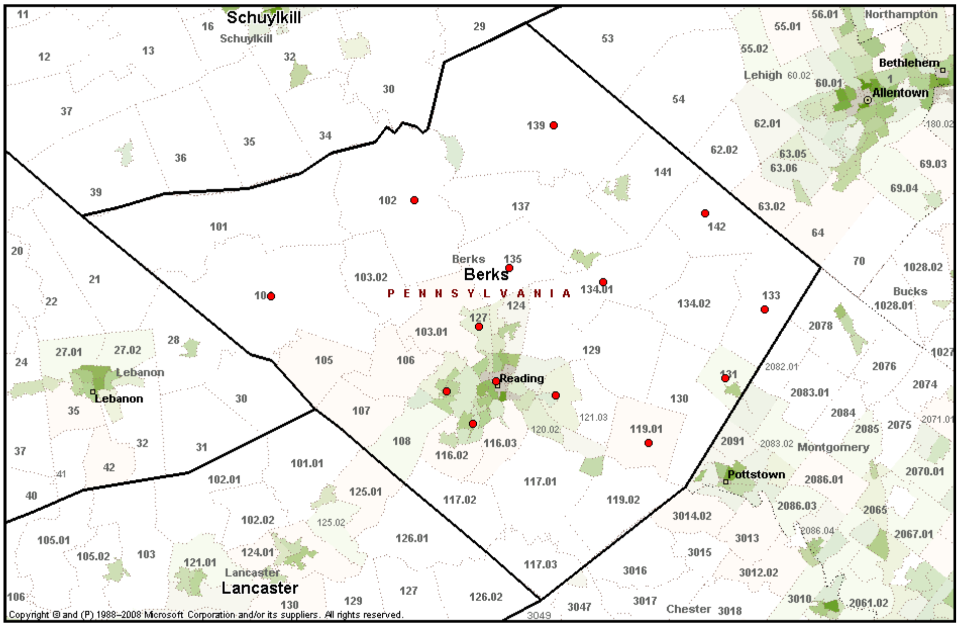


Figure 9: Comparison of 14-Store Configurations in Berks County

(a) Welfare Maximizing 14-Store Configuration



(b) Configuration Chosen by Free Entry Algorithm With *K* chosen so that *J*(*K*) = 14

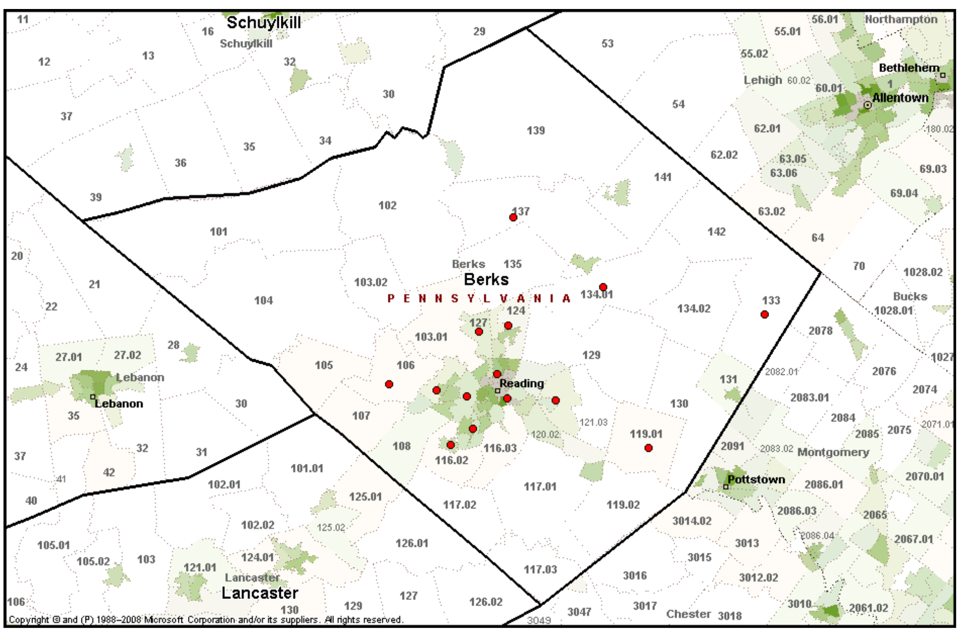
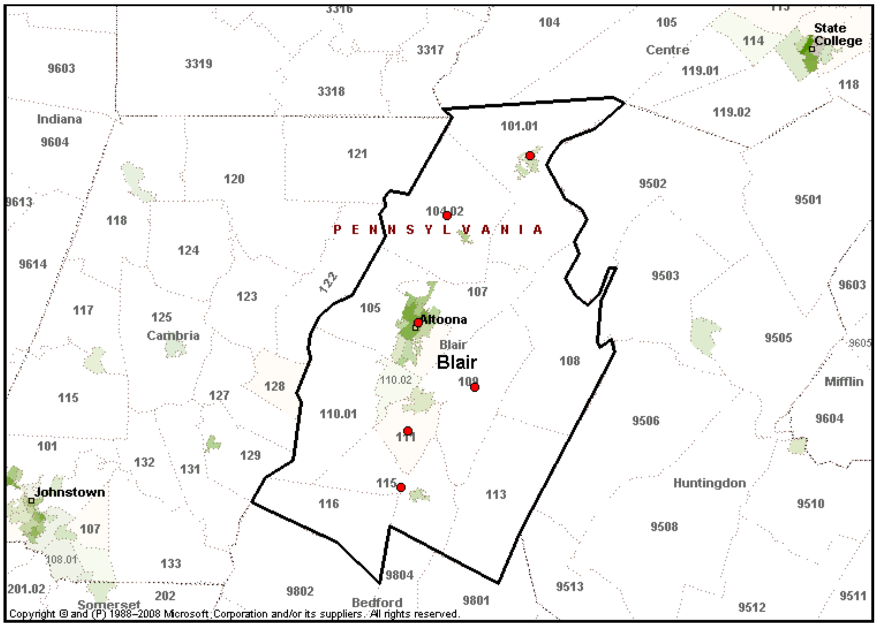


Figure 10: Comparison of 6-Store Configurations in Blair County

(a) Welfare Maximizing 6-Store Configuration



(b) Configuration Chosen by Free Entry Algorithm With *K* chosen so that *J*(*K*) = 6

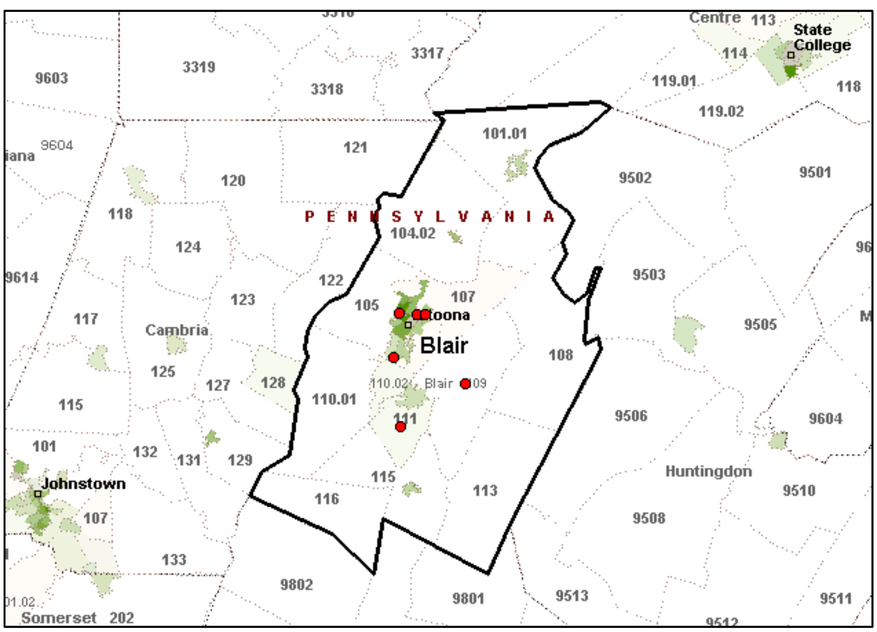
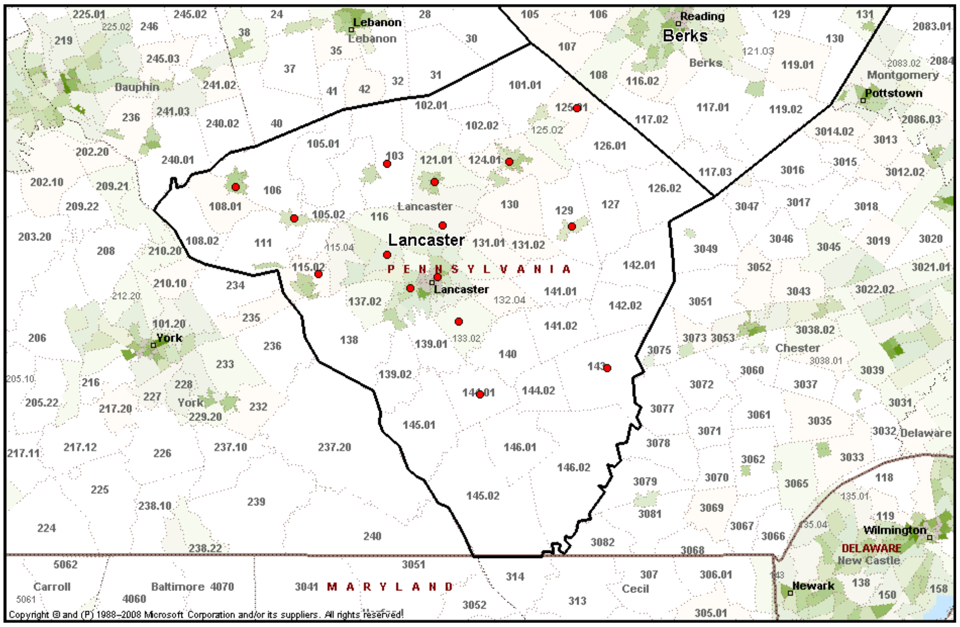


Figure 11: Comparison of 15-Store Configurations in Lancaster County

(a) Welfare Maximizing 15-Store Configuration



(b) Configuration Chosen by Free Entry Algorithm With *K* chosen so that *J*(*K*) = 15

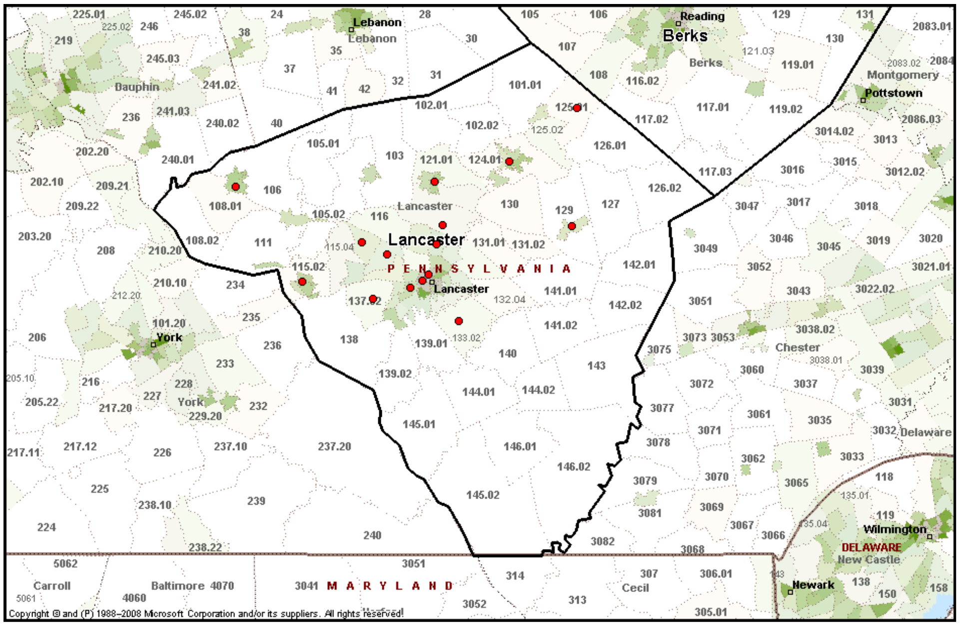
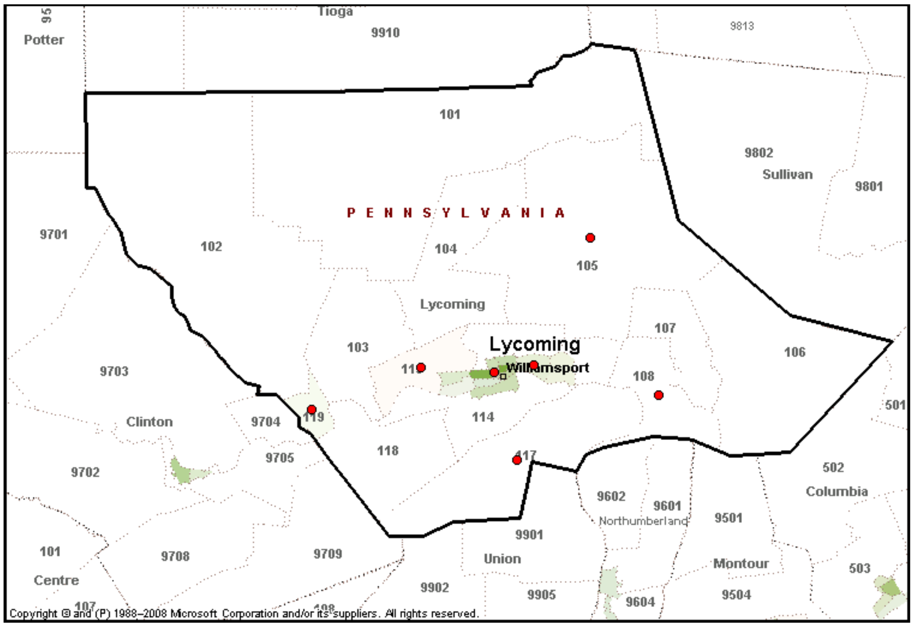


Figure 12: Comparison of 7-Store Configurations in Lycoming County

(a) Welfare Maximizing 7-Store Configuration



(b) Configuration Chosen by Free Entry Algorithm With *J* chosen so that *J*(*K*) = 7

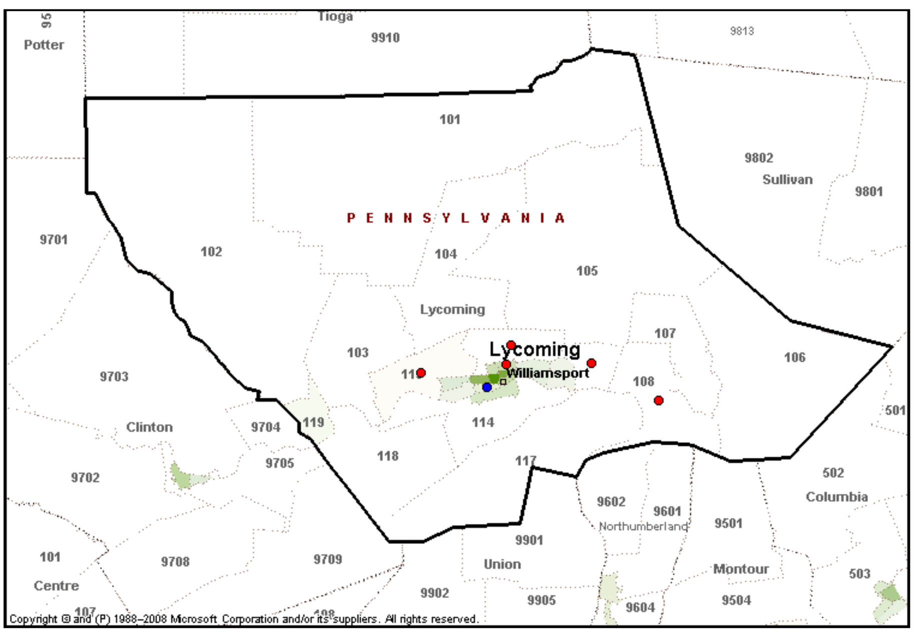
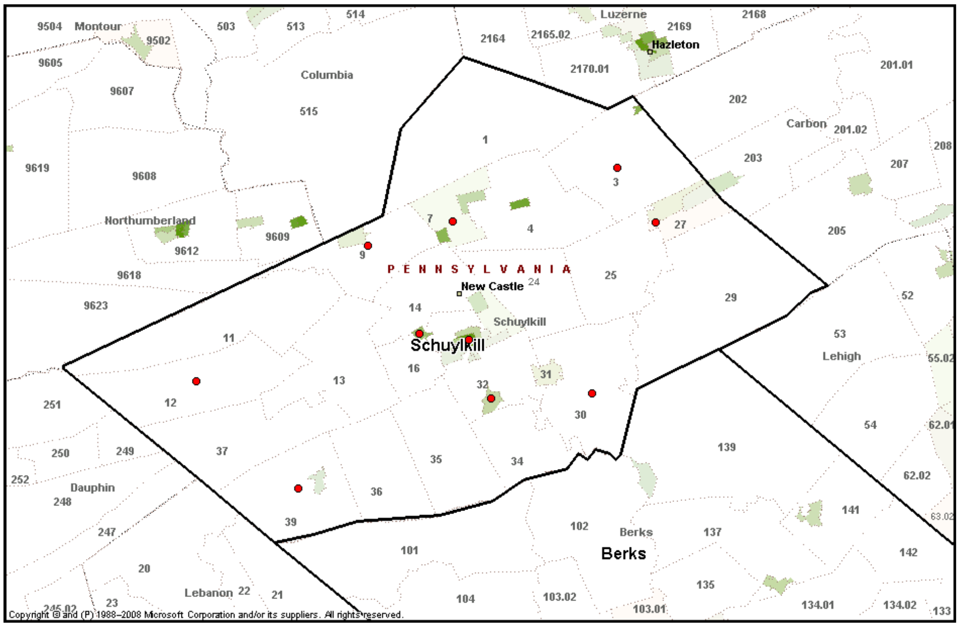


Figure 13: Comparison of 10-Store Configurations in Schuykill County

(a) Welfare Maximizing 10-Store Configuration



(b) Configuration Chosen by Free Entry Algorithm With *K* chosen so that *J*(*K*) = 10

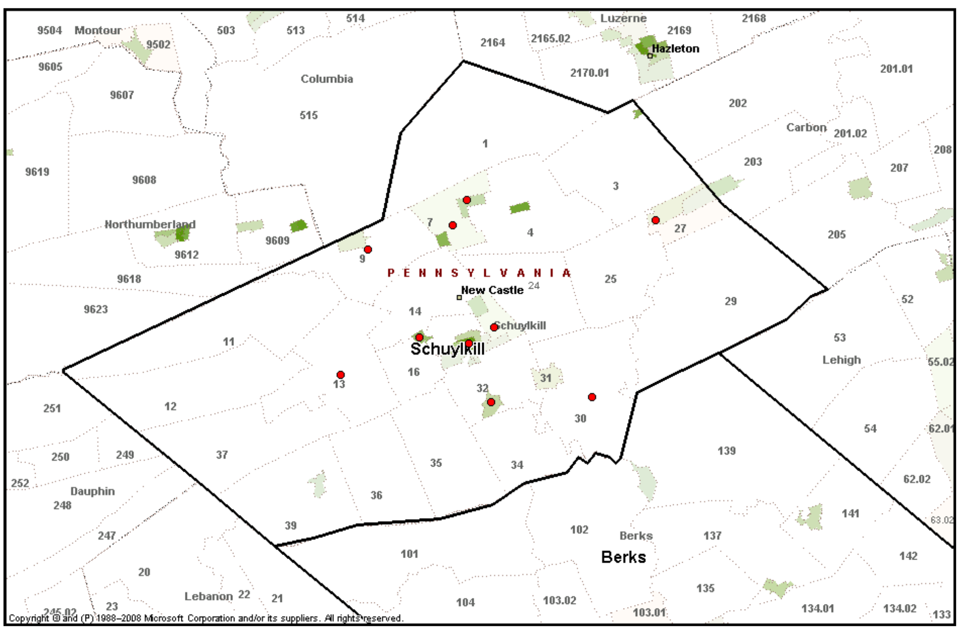


Table 1: Entry and Exit Patterns, PLCB Stores, 2001-2006

	Number of stores operating	Number of store openings, previous 12 m	Number of store closings, previous 12 m
January 2001	625		
January 2002	622	4	7
January 2003	621	2	3
January 2004	620	5	6
January 2005	624	5	1
January 2006	633	14	8

Notes: Store counts, openings, and closings derived from revenue data. We exclude wholesale stores, stores without or with incorrect location information, and stores with persistently low revenue (revenue in the bottom 1 percentile of stores across all periods).

Table 2: Price Elasticity Evidence

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Log State Bundle List Price	-0.9627	-1.0900	-1.3184	-1.3396								
	(1.2303)	(0.1764)**	(1.3557)	(0.1806)**								
Log State Bundle Price					-1.8926	-1.9404	-0.8591	-0.8587	-0.5799	-0.688	-0.7402	-1.2023
					(0.7379)*	(0.1053)**	(0.8224)	(0.1096)**	(0.3896)	(0.1092)**	(0.9977)	(0.2668)**
Constant	9.8875	10.2075	10.6897	10.7466	12.2274	12.3476	9.5368	9.5396	7.3699	6.8425	7.8785	8.253
	(3.0954)**	(0.4438)**	(3.4096)**	(0.4543)**	(1.8567)**	(0.2649)**	(2.0706)**	(0.2758)**	(0.9791)**	(0.2745)**	(2.5089)**	(0.6710)**
Observations	32,509	32,509	32,509	32,509	32,509	32,509	32,509	32,509	191,994	191,994	23,532	23,532
R-squared	0.06	0.77	0.07	0.80	0.06	0.77	0.07	0.80	0.19	0.76	0.07	0.39
Sample	weekly	weekly	weekly	weekly	weekly	weekly	weekly	weekly	daily	daily	daily, price changes	daily, price changes
Holiday weeks	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Quadratic time trend	no	no	yes	yes	no	no	yes	yes	yes	yes	no	no
Holidays									yes	yes	no	no
Day of the Week									yes	yes	yes	yes
Store FE	no	yes	no	yes	no	yes	no	yes	no	yes	no	yes

Notes: Dependent variable is log bottles per time period per store. Regressions of log bottles sold on various measures of the price. Holiday weeks include weeks 1, 26, 47, 50, 51, and 52. We include separate time trends for the period January - October and the holiday period of November - December. Standard errors in parentheses. * significant at 5%; ** significant at 1%. State-bundle prices use a constant bundle for computing the price and vary only by time and not across stores.

Table 3: Demand, Population, and Distance to the Nearest Store

	(1)	(2)	(3)	(4)	(5)	(6)
	Average Weekly Sales per Store	Average Weekly Log Sales per Store	Log Sales per Store	Log Sales per Store	Log Sales per Store	Log Sales per Store
Catchment Area Pop	49.4185	0.0227	0.0241	0.0111	0.0241	0.0111
	(6.3825)**	(0.0025)**	(0.0004)**	(0.0005)**	(0.0004)**	(0.0005)**
Average Distance to Nearest Store	-413.2364	-0.2530	-0.2501	-0.2298	-0.2501	-0.2305
	(107.2327)**	(0.0425)**	(0.0060)**	(0.1247)	(0.0060)**	(0.1247)
Log State Bundle List Price			-1.3604	-1.3733		
			(1.2093)	(0.1794)**		
Log State Bundle Price					-0.8841	-0.8756
					(0.7336)	(0.1088)**
Constant	1,973.24	7.2358	10.4582	10.7347	9.2628	9.4858
	(165.3338)**	(0.0655)**	(3.0413)**	(0.4558)**	(1.8469)**	(0.2813)**
Observations	628	628	32,509	32,509	32,509	32,509
R-squared	0.13	0.20	0.26	0.80	0.26	0.80
Store FE	No	No	No	Yes	No	Yes

Standard errors in parentheses. * significant at 5%; ** significant at 1%.

Table 4: Demand Model Estimates

	(1)	(2)	(3)
State-bundle list price	-0.1533	-0.1565	-0.1657
	(0.0843)**	(0.0778)**	(0.0919)**
Distance	-0.1047	-0.0578	
	(0.0151)***	(0.0401)*	
Distance squared		-0.4225	
		(0.1140)***	
Driving distance			-0.0656
			(0.0097)***
Black	-0.0964	-0.0651	0.0557
	(0.1434)	(0.1249)	(0.0145)***
Median Income	0.0390	0.0387	0.0373
	(0.0037)***	(0.0042)***	(0.0061)***
Median Age	0.0047	0.0039	0.0024
	(0.0111)	(0.0108)	(0.0116)
No Churches per capita			-0.1329
			(0.0975)*
Monday	0.5966	0.5956	0.7088
	(0.0517)***	(0.0686)***	(0.0580)***
Tuesday	0.6542	0.6529	0.7251
	(0.0609)***	(0.0799)***	(0.0746)***
Wednesday	0.7595	0.7579	0.8581
	(0.0518)***	(0.0709)***	(0.0618)***
Thursday	0.9103	0.9088	1.0132
	(0.0517)***	(0.0711)***	(0.0631)***
Friday	1.4668	1.4649	1.5682
	(0.0549)***	(0.0714)***	(0.0584)***
Saturday	1.4920	1.4902	1.5857
	(0.0468)***	(0.0629)***	(0.0583)***
Implied elasticity of demand	-1.1830	-1.2064	-1.2795
Implied travel cost ($) / km	0.6827	0.6167	0.3957

Note: Results based on daily store-level data for a 10% subset (19255 obs) of all store-day observations. Bootstrapped standard errors (50 replications). We also include separate holiday effects for May 28, June 30-July 3, Sept 3, and Nov. 23-26. For specification (2), travel cost per km traveled computed for the mean distance traveled by consumers. Specification (3) uses the shortest travel distance in km along the road network instead of straight line distance.

Table 5: Comparison of Regimes

County	Regime	Number of Stores	Quantity Sold	Total Profit	Net Welfare	Welf rel to max	Loss as % of Exp
Berks	Actual	14	7,072	20,208	333,574	-2.48%	9.7%
(pop: 373,638)	Free Entry	30	8,335	8,726	330,753	-3.31%	11.0%
	Profit Maximization	8	6,715	25,105	335,996	-1.77%	7.3%
	Welfare Maximization	19	8,257	20,625	342,066		0.0%
Blair	Actual	6	2,004	3,456	110,766	-3.15%	14.5%
(pop: 129,144)	Free Entry	8	2,329	2,868	112,389	-1.73%	6.9%
	Profit Maximization	2	1,766	6,720	112,427	-1.70%	8.9%
	Welfare Maximization	6	2,309	4,998	114,365		0.0%
Lancaster	Actual	15	8,173	24,671	404,621	-1.78%	7.2%
(pop: 470,658)	Free Entry	31	9,574	13,891	403,494	-2.05%	7.1%
	Profit Maximization	8	7,668	29,933	406,458	-1.33%	5.8%
	Welfare Maximization	17	8,976	26,506	411,943		0.0%
Lycoming	Actual	7	1,937	1,994	101,091	-1.99%	8.6%
(pop: 120,044)	Free Entry	6	1,825	2,545	100,885	-2.19%	10.0%
	Profit Maximization	2	1,521	5,475	101,759	-1.34%	7.4%
	Welfare Maximization	6	2,017	3,517	103,143		0.0%
Schuylkill	Actual	10	2,746	2,747	141,455	-1.74%	7.4%
(pop: 157,342)	Free Entry	9	2,564	2,937	140,327	-2.53%	11.5%
	Profit Maximization	3	2,054	7,063	141,000	-2.06%	11.7%
	Welfare Maximization	8	2,770	5,101	143,966		0.0%

Table 6: Performance of Myopic Algorithm

county	obj func	No stores	net welf	Q	profit	net welf dev	Q dev	profit dev
Berks	profit max	9	336,722	6,868	24,759	0.22%	2.27%	-1.38%
Blair	profit max	2	112,079	1,737	6,570	-0.31%	-1.68%	-2.24%
Lancaster	profit max	9	406,588	7,771	29,337	0.03%	1.34%	-1.99%
Lycoming	profit max	2	101,738	1,519	5,465	-0.02%	-0.13%	-0.18%
Schuylkill	profit max	3	141,000	2,054	7,063	0.00%	0.00%	0.00%
Berks	welfare max	20	341,545	8,305	19,752	-0.15%	0.59%	-4.23%
Blair	welfare max	5	113,993	2,183	5,481	-0.33%	-5.42%	9.66%
Lancaster	welfare max	18	411,765	9,055	25,789	-0.04%	0.88%	-2.71%
Lycoming	welfare max	6	103,143	2,017	3,517	0.00%	0.00%	0.00%
Schuylkill	welfare max	8	143,936	2,768	5,092	-0.02%	-0.06%	-0.17%

Table 7: Sensitivity of Incremental Welfare Measures

X.	Incremental PS	Incremental TS	Variable Profit	Distance param	Travel Cost / km	Elasticity
High distance	745.6	1978.0	2190.5	-0.21	1.67	-0.97
	619.5	1649.9	1950.1			
Low distance	250.4	542.0	2079.3	-0.05	0.30	-1.37
	169.9	370.1	1823.8			
High elas	443.2	742.3	2108.6	-0.11	0.34	-2.37
	337.7	567.7	1873.4			
Low elas	445.2	1646.8	2118.5	-0.10	1.37	-0.59
	339.2	1258.7	1882.3			

Note: Mean and median incremental welfare measures under alternative parameter estimates: a high distance parameter twice its baseline value, with the other parameters estimated freely, a low distance parameter half its baseline value (other parameters estimated freely), a price elasticity twice its baseline value, and a price elasticity half its baseline value.

Table 8: Regime Comparison for Actual-Sized Configurations

		N	Q	Variable Profit	Net Welfare	Loss as % of Exp
Berks	Welf Max	14	7,699	39,036	341,024	
	Actual	14	7,072	35,855	333,574	8.51%
	Free Entry	14	7,336	37,194	336,765	4.69%
Blair	Welf Max	6	2,309	11,705	114,365	
	Actual	6	2,004	10,160	110,766	14.51%
	Free Entry	6	2,018	10,230	110,942	13.70%
Lancaster	Welf Max	15	8,763	44,428	411,627	
	Actual	15	8,173	41,437	404,621	6.92%
	Free Entry	15	8,444	42,812	407,903	3.56%
Lycoming	Welf Max	7	2,100	10,649	103,026	
	Actual	7	1,937	9,821	101,091	8.07%
	Free Entry	7	1,854	9,402	100,120	12.66%
Schuylkill	Welf Max	10	2,938	14,894	143,708	
	Actual	10	2,746	13,922	141,455	6.63%
	Free Entry	10	2,793	14,159	141,953	5.07%

Table 9: Whole State Summary

System's Size	Method	N	Profit	Net Welfare	Quantity	Average Distance
Profit Maximization	SME(profit)	210	922,137	11,350,117	228,298	5.93
Welfare Maximization	SME(welfare)	481	812,482	11,500,172	266,299	3.88
Actual size		616	618,741	11,244,951	257,849	3.23

Table 10: Correlates of Weights that Render Existing PLCB Entry Optimal

	(1)	(2)	(3)	(4)
% Black	4,241.34	3,397.62	3,549.96	2,175.32
	(1,023.82)**	(1,059.50)**	(1,048.43)**	(1,086.77)*
Avg. Median Inc		-0.0356	-0.0216	-0.0318
		(0.0124)**	(0.0128)	(0.0129)*
Zone Pop			-0.4028	-0.3485
			(0.1043)**	(0.1038)**
Avg Distance to Nearest Store (km)				-291.7368
				(70.4332)**
Constant	1,834.36	3,326.08	4,432.19	5,650.66
	(213.44)**	(561.71)**	(624.97)**	(683.43)**
Observations	616	616	616	616
R-squared	0.03	0.04	0.06	0.09

Standard errors in parentheses. * significant at 5%; ** significant at 1%.

References

Berry, Steven T., “Estimation of a Model of Entry in the Airline Industry.” *Econometrica* 60 (1992): 889-917.

Berry, Steven T. and Joel Waldfogel, “Free Entry and Social Inefficiency in Radio Broadcasting.” *RAND Journal of Economics* 30, 3 (1999): 397-420.

Berry, Steven T. and Joel Waldfogel, "Mergers and Product Variety." *Quarterly Journal of Economics* 116 (2001): 1009-1025.

Bresnahan, Timothy F. and Peter C. Reiss, “Entry and Competition in Concentrated Markets.” *Journal of Political Economy* 99 (1991): 977-1009.

Chaloupka, Frank J., Michael Grossman, and Henry Saffer, “The Effects of Price on Alcohol Consumption and Alcohol-Related Problems.” *Alcohol Research and Health* 6, 1 (2002): 22-34.

Chan, Tat Y., V. Padmanabhan, and P.B. Seetharaman, “An Econometric Model of Location and Pricing in the Gasoline Market.” Journal of Marketing Research XLIV, (2007): 622-635.

Cook, Phillip J. and Michael J. Moore, “Alcohol.” *NBER Working Paper* 6905 (1999).

Daskin, M. S., 1995, Network and Discrete Location: Models, Algorithms and Applications, John Wiley and Sons, Inc., New York.

D'aspremont, Claude, Jean Gabszewicz, and Jacques-François Thisse, “On Hotelling's ‘Stability in Competition’.” *Econometrica* 47, 5 (1979): 1145-1150.

Davis, Peter, “Spatial competition in retail markets: movie theaters.” *RAND Journal of Economics* 37(4) (2006): 964-982.

Dixit, Avinash and Joseph E. Stiglitz, "Monopolistic Competition and Optimum Product Diversity." *American Economic Review* 67 (1977): 297-308.

Fudenberg, Drew and David K. Levine. The Theory of Learning in Games. MIT Press: Cambridge, MA, 1998.

Hausman, Jerry, “Exact Consumer Surplus and Deadweight Loss,” *American Economic Review* 71, 4 (1981): 662-676.

Houde, Jean-Francois, “Spatial differentiation in retail markets for gasoline,” mimeo. University of Wisconsin-Madison (2008).

Hsieh, Chang-Tai and Enrico Moretti, “Can Free Entry be Inefficient? Fixed Commissions and Social Waste in the Real Estate Industry.” *Journal of Political Economy* 111 (2003): 1076-1122.

Mankiw, N. Gregory and Michael D. Whinston, “Free Entry and Social Inefficiency.” *RAND Journal of Economics* 17, 1 (1986): 48-58.

Manning, Willard G., Linda Blumberg, and Lawrence H. Moulton, “The Demand for Alcohol: The Differential Response to Price.” *Journal of Health Economics* 14, 2 (1995): 123-148.

Mazzeo, Michael, “Product Choice and Oligopoly Market Structure.” *RAND Journal of Economics* 33(2) (2002): 241-242.

McManus, Brian, “Nonlinear Pricing in an Oligopoly Market: the Case of Specialty Coffee.” *RAND Journal of Economics* 38(3) (2007): 512-532.

Perl, Jossef and Peng-Kuan Ho, “Public Facilities Location under Elastic Demand.” *Transportation Science* 24(2) (1990): 117-136.

Samuelson, Larry. Evolutionary Games and Equilibrium Selection. MIT Press: Cambridge, MA. 1997.

Seim, Katja, “An Empirical Model of Firm Entry with Endogenous Product-type Choices.” *RAND Journal of Economics* 37(3) (2006): 619-640.

Small, Kenneth and Harvey Rosen, “Applied Welfare Economics with Discrete Choice Models.” *Econometrica* 49(1) 1981: 105-130.

Spence, Michael, "Product Selection, Fixed Costs, and Monopolistic Competition." *Review of Economic Studies* 43, (1976): 217-236.

Steiner, Peter, “Program Patterns and the Workability of Competition in Radio Broadcasting,” *Quarterly Journal of Economics* 66 (1952): 194-223.

Sweeting, Andrew, “Too Much Rock and Roll? Station Ownership, Programming and Listenership in the Music Industry,” mimeo. Duke University (2006).

Thomadsen, Raphael, “The Effect of Ownership Structure on Prices in Geographically Differentiated Industries,” *RAND Journal of Economics*, 36(4) (2005): 908-929.

Willig, Robert D., “Consumer's Surplus without Apology.” *American Economic Review* 66, 4 (1976): 589-597.