Spatial Differentiation in Retail Markets for Gasoline*

Jean-François Houde
Department of Economics, Queen's University

Job Market Paper

February 14, 2006

Abstract

This paper studies an empirical model of spatial competition. The main feature of my approach is to formally specify commuting paths as the "locations" of consumers in a Hotelling-type model of spatial competition. This modeling choice is motivated by the fact that consumers are moving across the market when consuming the product. Although this feature is perhaps more relevant for gasoline markets, this also applies to most retail markets since consumers are not immobile. The consequence of this behavior is that competition is not fully localized as in the standard address-model. In particular, the substitution patterns between stations depend in an intuitive way on the structure of the road network and the direction of traffic flows. Another feature of the model is that consumers' available options are directly linked with their commuting behavior. Consumers who commute more encounter more stations, observe more prices, and therefore may pay less on average; a result observed empirically (Yatchew and No [2001]). The demand-side of the model is estimated by combining a model of traffic allocation with econometric techniques used to estimate models of demand for differentiated products (Berry, Levinsohn and Pakes [1995]). The empirical distribution of commuters is computed with a shortest-path (or Dijkstra) algorithm, combining detailed data on the Québec City road network with aggregate Origin-Destination commuting probabilities. The model's parameters are then estimated using a unique panel data-set on the Québec City gasoline market from 1991 to 2001. The empirical analysis of the model clearly shows that adding commuting behavior is important for explaining the distribution of sales across the city. In particular, gasoline sales are poorly correlated with the population distribution (address-model) but more closely related to the distribution of commuting. The results of the structural model also reveal important differences between the home-location and the commuting models. Contrary to the commuting model, the estimated transportation cost is unrealistically low under the home-location model. The evaluations of market power under the two models are also different. The home-location model leads to markup estimates that are close to 50% higher than the commuting model. The estimates of the structural model are also used to assess the relative market power of major retail chains over independents. The results indicate that the role of independents in this market is marginal, since on average they are operating low-value and high-cost stations.

*Preliminary and Incomplete. I am grateful for the advice and support of Chris Ferrall and Susumu Imai. I have benefited from the comments and discussions of Jason Allen, Beverly Lapham, Jeremy Lise, Sumon Majumdar, Shannon Seitz, and seminar participants at Queen's, Concordia, UQAM, HEC-Montréal, Université de Montréal, University of Pennsylvania, Bank of Canada, Simon Fraser University, New York University, New York Fed, and University of Wisconsin. Funding from the Social Sciences and Humanities Research Council of Canada (SSHRC), and the *Fonds Québecois de la Recherche sur la Nature et les Technologies* (FCAR) are gratefully acknowledged.
Correspondence: (phone) +1-613-533-6000 ext. 75532, (email) houdejf@econ.queensu.ca

Keywords: Spatial differentiation; Retail markets; Transportation; Market power.

1 Introduction

This paper studies an empirical model of spatial competition. The main feature of my approach is to formally specify commuting paths as the “locations” of consumers in a Hotelling-type model of spatial competition. This modeling choice is motivated by the observation that consumers can be located at more than one point in the product space when deciding which station to purchase gasoline. This characteristic is central to the analysis of gasoline markets, because consumers are by definition moving when consuming the product. Although this feature is perhaps more relevant for gasoline markets, this also applies to most retail markets since consumers are not immobile when choosing where to go shopping. Formally, I construct a new Hotelling-type address model which defines consumer locations as their home-to-work commuting path, rather than their home or workplace separately. The demand-side of the model is estimated using a unique panel dataset on the Québec City gasoline market between 1991 and 2001, following the techniques developed by Berry [4] and Berry, Levinsohn and Pakes (BLP) [5]. The model is then used to measure retailer mark-ups and evaluate the market power of major retail chains.

Understanding the sources of market power in retail markets has important policy implications. Over the years, many provincial and state governments have attempted to limit the extent of market power of larger retailing chains. In gasoline markets, these concerns have led to the adoption of price floor regulations and contract restrictions[1]. These policies are typically aimed at protecting independent retailers from alleged predatory behavior by major chains. For instance, in 1997 the Québec provincial government established a price floor regulation, after the occurrence of major price wars which spread throughout the Province. The design of the regulation is similar to an anti-dumping trade policy, and allows firms to sue their local competitors if they fix prices below a lower bound published every week by the Government.

A well specified model of demand is the key ingredient to evaluate the usefulness of these policies. In particular, locating consumers incorrectly could lead to incoherent measures of the elasticity of substitution between store locations, and invalid predictions regarding firm mark-ups

[1] In 1996, 12 US States and 4 Canadian Provinces had price floor regulations targeted at gasoline retail prices, and 7 US States had contract restrictions preventing major Oil companies to be vertically integrated in the retail market[31].

in counter-factual experiments market structure. In the standard address model consumers are located at single points in the product space (e.g. home), and decide where to buy a product by weighting the value of price differences with the relative travel cost of each product from this central point. In gasoline markets, since consumers are not using the product at home, it is not obvious why a station located anywhere along a common commuting path should be valued differently than a station close to home.

To take this feature of the market into account, I define the location of consumers to be the set of intersections on a road network representing the shortest driving path to commute from home to work (or school). By defining a location in this way, consumers choose where to buy gasoline by trading off price differences with the relative travel costs of deviating from their main commuting path. For a specific type of commuter, two gasoline stations at the two ends of the city can be close substitutes if they happen to be on her daily route. The estimated model therefore generates substitution patterns that are very different from the ones generated by a single location model.

The model can reproduce documented features of gasoline markets. For instance, Yatchew and No [42] show using micro data on gasoline consumption that households who consume more gasoline also tend to pay lower prices, so that prices are endogenous in a household level demand equation. In the model, long distance commuters will naturally encounter more stations along their driving path, and therefore may pay lower prices on average. On the other hand, a standard address model would have difficulty matching this fact without resorting to heterogeneous time costs, or underestimating the transportation cost value.

Furthermore, from the firm's point of view, competition is not fully localized since consumers can substitute stations far from each other but located along a common path. Consequently, even if consumers are not willing to deviate far from their path to shop for gasoline, price differences between two regions of a city are unlikely to be persistent if there is enough frequent commuters between the two regions. This is consistent with the observation that there is significantly less price dispersion within gasoline markets than in most other retail markets. For example the statistics reported in Lach [22] and reproduced in Table 1, suggest that the dispersion of gasoline prices is half as large as most grocery products. Gasoline price dispersion is also similar in magnitude to refrigerators for which search costs are potentially quite small in magnitude relative to the price.

The estimation of the model is performed in two steps. First, I compute the empirical distribu-

Table 1: Standard deviations of residual prices (in logs)

Refrigerator†	Flour†	Coffee†	Chicken†	Regular Gas	Premium Gas
0.0323	0.0436	0.0585	0.0728	0.0248	0.0313

† Reported in Lach [22].

tion of consumers across the road network using a deterministic route choice model, common in the transportation demand literature (Oppenheim [29]). More specifically, the model predicts traffic on each segment of the road network, conditional on the distribution of origin-destination (OD) commuting pairs, assuming that individuals are choosing the optimal shortest driving time route. The empirical distribution of commuters is obtained by combining the results of an OD survey conducted by the Québec Ministry of Transportation, with aggregate population tables from the three most recent Canadian Censuses and the Canadian Monthly Labour Force Survey. The set of optimal routes is computed with a shortest path algorithm common in the Geographic Information System (GIS) literature, using detailed data on the Québec City road network.

Conditional on the empirical distribution of commuting paths, the preference parameters are estimated by Generalized Method of Moments (GMM), following the techniques developed by Berry [4], Berry, Levinsohn and Pakes [5], and Nevo [27]. In particular, the estimation methodology deals explicitly with the endogeneity of prices using an Instrumental Variable approach, and by controlling for fixed unobserved heterogeneity at the gasoline station level.

Estimation of the demand-side parameters is performed using a unique panel dataset of gasoline stations in the Québec City area, spanning 46 bi-monthly periods between $1991 - 2001$. The time-period chosen is characterized by important changes in the structure of the North-American gasoline retail industry, associated with massive exit of stations and entry of new categories of retailers (e.g. large stations with a convenience stores). Also, for more than half of the sample periods, the market is subject to the previously discussed price floor regulation. These two characteristics of the sample represent important sources of exogenous variation in the data, changing substantially the choice set of consumers over time, and reducing the correlation of prices with respect to unobserved station characteristics.

The results can be summarized as follows. First, the model based on commuting behavior is shown to fit the observed distribution of sales more closely than the standard home-address model. In particular, stations with high market shares are not the ones located close to home. This leads to a correlation close to zero (or even negative) between the number of people living in a neighborhood of stations and their sales. The results of structural model also reveal important differences between the home-location model and the commuting location model. The commuting model sucessfully explains the fact that consumers who consumer more gasoline pay lower price on average; a result obtained by Yatchew and No [42]. More importantly, the estimated transportation cost is unrealistically low under the home-location model, leading to negative shopping cost (i.e. consumers value positively the time necessary to shop for gaoline). The commuting model on the other hand provides an estimate of the shopping cost which is high, and in the range of current estimates reported the literature. The market power evaluation under the two models are also very different. The home-location model provides profit margins that are on average 45% higher than the commuting model. The role of independents is found to be marginal however, because they are on average much less efficient (i.e. low quality products and higher costs). Keeping an artificially high number of independent stations in the market, through a price floor regulation for instance, can therefore lead to higher prices overall.

This paper is part of a large literature in empirical industrial organization devoted to the estimation of discrete choice models of demand using firm level data (see Bresnahan [9] for an early example, and Berry [4] and BLP [5] for further developments of the estimation methodology). Recently, the methodology proposed by BLP has been extended in several directions to evaluate market responses to policy changes. Examples of this approach include the evaluation of international trade policies (e.g. BLP [6] and Brambilla [8]), the analysis of mergers (e.g. Nevo [26] and Dubé [15]), and the valuation of new goods (e.g. Petrin [30]). See Ackerberg, Benkard, Berry and Pakes [1] for an extensive review of this literature.

Recently researchers have studied empirically markets for spatially differentiated products. Closely related to the techniques used in this paper, Manuszak [25] and Davis [11] extend the BLP methodology to estimate an address-model applied respectively to Hawaiian gasoline markets and the US movie theater industry. Thomadsen [41] also estimate an address model using data on prices and store characteristics in the fast-food industry, imposing the equilibrium conditions of a Bertrand pricing game with spatial differentiation[2]. Although the econometric techniques used in

[2]Other papers looking empirically at markets with spatial differentition include: Smith [39] studies demand for grocery products using micro-data from a UK household survey Pinkse, Slade and Brett [33] and Pinkse and Slade

these papers are similar, on the consumer side the model is different. Specificaly, I consider the possibility that consumers have multiple locations through their commuting behavior. The theoretical literature on spatial competition has extended the Hotelling model in many directions, including recently the possibility of multiple dimensions of differentiation on the firm side (see Manez and Waterson [24], and Anderson, de Palma and Thisse [2] for reviews of this literature). To the best of my knowledge however, I am not aware of any other papers who study spatial competition in markets where consumers are located at more than one point in space. In a different context, Lagos [23] develops an endogenous matching model based on the market for taxi-cabs, in which the distribution of consumers is given by an origin-destination probability matrix similar to the one used in this paper.

The rest of the paper is organized as follows. Section 2 presents the model. Section 3 presents the estimation method and the identification strategy. Section 4 discusses the data used to construct the distribution of commuters, and the gasoline stations used to estimate the model. Section 5 presents the empirical results, including the policy experiments conducted using the estimated parameters. I conclude the paper and discuss extensions in section 6. Further computational and data construction details are placed in the Appendices.

2 The Model

I model competition in the market as a static oligopolistic pricing game with spatial differentiation. A market corresponds to a Census Metropolitan Area (CMA), divided into a set of non-overlapping areas $\mathcal{L} = \{1, ..., L\}$. The geography of the market is further characterized by a road network $G = (N, A)$, where N is a set of street intersections (or nodes), and A is the set of road segments (or arcs). The cost of traveling along the network is given by the function $C : A \rightarrow \Re_+$ giving the travel time associated with every arc.

There are F brands in the market organized as a network of retail stores indexed by $j \in \{1, ..., J\}$. Each store sells two grades of gasoline indexed by $g \in \{regular, premium\}$[3]. The mean

[32] estimate a model of spatial competition applied respectively to the wholesale gasoline and beer pints markets, and finally the papers by Hasting [18] and Hastings and Gilbert [19] use a natural experiment approach (namely an observed vertical merger in California) to measure market power and the impact of vertical integration in markets for gasoline.

[3]Note that most gasoline stations are offering three or four grades of gasoline (i.e. regular, midgrade, premium, and premium plus). However, in order to reduce the complexity of the model, I combine the two highest and the two lowest quality grades together so that every station is offering two grades.

value of a product pair, denoted by δ_{jg}, is modeled as a linear function of observed characteristics X_{jg}, posted price p_{jg}, and an index ξ_{jg} of unobserved (to the econometrician) quality:

$$\delta_{jg} = X_{jg}\beta - \alpha p_{jg} + \xi_{jg}. \tag{1}$$

Letting $Q_{jg}(\delta(p))$ be the demand function for product pair (j, g), brand f sets prices as a multi-product oligopolist:

$$\max_{\{p_{jg}\}j\in\mathcal{F}_f} \sum_{j\in\mathcal{F}_f}\sum_{g} Q_{jg}(\delta(p))(p_{jg} - w_{jg}), \tag{2}$$

where $\mathcal{F}_f$ is the set of stations selling brand f, and w_{jg} is the constant marginal cost of selling grade g. The Betrand-Nash equilibrium price vector solves the following set of first-order conditions:

$$Q(\delta(p)) + \Delta(p)(p - w) = 0, \tag{3}$$

where Δ is a $2 \cdot J \times 2 \cdot J$ matrix given by:

$$\Delta_{kl} = \begin{cases} \frac{\partial Q_k(\delta(p))}{\partial p_l} & \text{if } l \in \mathcal{F}_{f_k} \\ 0 & \text{otherwise} \end{cases}. \tag{4}$$

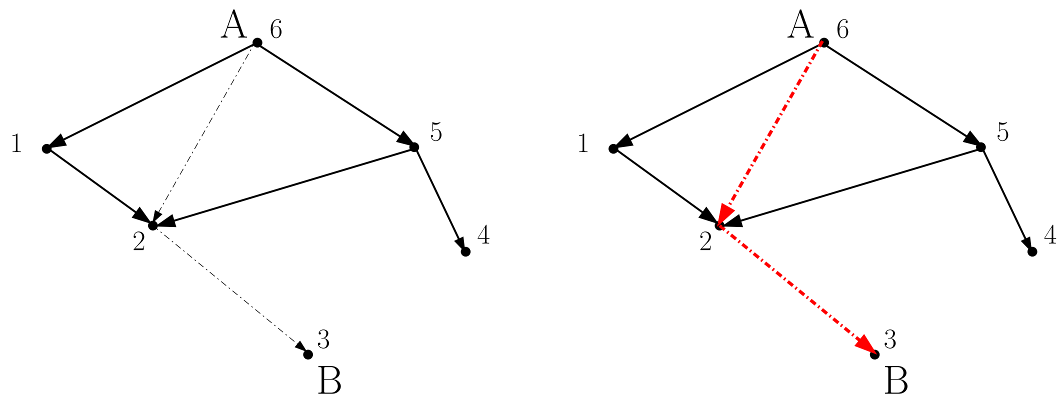
Note that this equilibrium condition on prices will **not** be imposed while estimating the model, as in BLP [5] for instance. However, it will be used to evaluate firms' mark-ups, and simulate counter-factual market outcomes under different market structure assumptions[4].

Consumers in the market are characterized by a pair $(s, d) \in \mathcal{L}^2$ of home and main occupation locations (i.e. work or study). The distribution of individuals across origin-destination (OD) pairs is given by T_{sd}[5]. That is, T_{sd} is the number of people who commute everyday between locations s and d for work or study. Consumers choose optimally the commuting route between (s, d), in the sense of minimizing the travel time between their home and main occupation locations. This decision is modeled using a deterministic route choice model (Oppenhein [29]). The model predicts a single path, denoted by $r(s, d)$, which abstracts from any congestion or unobserved preferences considerations[6]. The optimal route between every pair of locations is calculated using a version of the Dijkstra's Shortest Path Tree (SPT) algorithm (Shekhar and Chawla [36]).

[4]The motivation for not imposing explicitly the equilibrium condition in the estimation, is that several authors have noticed that gasoline markets exhibit alternating periods of cooperation and price wars which are incompatible with the static model described here (e.g. Slade [37],[38], Borenstein and Shepard [7]). Imposing the equilibrium condition on prices could therefore introduce mispecification bias in the estimation.

[5]Although the value of T_{sd} is changing from period to period, time subscripts have been removed to reduce the notational burden. Appendix A describes the calculation of T_{sd}.

[6]The no-congestion assumption is realistic in the Québec City area, since the population is spread over a large territory and the road network is well developed. It has been used also by Thériault et al [40] to study the distribution of commuting trips in the Québec City metropolitan area using similar data.



(a) Network Representation (b) Shortest Path from A to B

Source	Dest.	Distance (c_{sd})
1	2	$\sqrt{8}$
2	3	$\sqrt{10}$
5	2	$\sqrt{10}$
5	4	$\sqrt{8}$
6	1	$\sqrt{15}$
6	2	$\sqrt{5}$
6	5	$\sqrt{5}$

(c) Node and Arc Relations

Figure 1: An Example of the Shortest Path on a Small Nework

Figure 1 illustrates the calculation of a shortest path between a pair of locations (A, B) on a fictitious network. The procedure iterates on the travel time between the starting node and every other nodes on the network. In the example, the optimal path is given by $r(A, B) = \{6, 2, 3\}$, and the travel time is given by $t(r(A, B)) = \sqrt{5} + \sqrt{10}$. On a real road network, the algorithm also records the distance traveled for every paths, which I denote by $m(r(s, d))$. More details on the computation of the shortest paths are provided in Appendix B.

To simplify the notation, consumers are indexed by their optimal route $r = r(s, d)$ whenever it is not confusing to do so. Using these paths, I define two distance functions relating commuters with station locations. In particular, $d_0(j, r)$ measures the Euclidean distance between the starting node of path r and station j, and $d_{min}(j, r)$ denotes the minimum Euclidean distance from any nodes in r to station j.

Given the location of consumers on the road network, I model the gasoline purchase decision as a discrete choice over a set of grades and store locations (j, g), plus a composite outside alternative $j = 0$ which includes options such as walking, public transportation, taxi, car pooling, or not commuting at all. Each consumer is assumed to purchase a fixed quantity of gasoline equivalent units of transportation every day[7]. However, in contrast to most of the literature on discrete choice models of product differentiation I introduce heterogeneous gasoline consumption using information on commuting distances. This fixed quantity for a consumer of type r is given by:

$$\bar{q}(r) = c_0 + c_1 m(r), \tag{5}$$

where c_0 represents the average daily transportation needs of consumers, and c_1 is the average quantity per kilometer. This feature of the model implies that consumers who have longer commutes buy more gasoline.

The indirect utility of purchasing product pair (j, g) for consumer i commuting along path r is given by:

$$u_{ijg}(r) = \delta_{jg} - D(j, r|\lambda) + v_{ig} + \epsilon_{ijg}, \tag{6}$$

where ϵ_{ijg} is an individual specific utility shock associated with product (j, g) distributed independently according to a double-exponential distribution (i.e. logit form), v_{ig} is consumer i valuation for grade of quality g, and $D(j, r|\lambda)$ is the dis-utility of distance for a consumer of type r buying from store j. Since there is a natural quality ordering between the two grades of gasoline, $v_{i,reg}$ is fixed to zero for every consumer, while $v_{i,prem} = v_i$ is distributed independently across the population according to an exponential distribution with mean μ. This distribution assumption ensures that the model is consistent with the vertical differentiation between grades of gasoline. That is, every consumers would strictly prefer the highest quality grade everything else equal[8].

The function $D(j, r|\lambda)$ measures the cost of deviating from path r to buy from store j. I consider a function which is linearly increasing in $d_{min}(j, r)$ (i.e. the minimum euclidean distance from path) within a certain band or choice set $\mathcal{CS}(r)$, and is going to infinity outside this set. In the empirical application, the choice set is defined such that consumers choose among the 100 closest store-locations around their path[9]. I also consider a more general functional form which allows

[7] An alternative representation of the decision could allow consumers to make a mixed continuous and discrete choice over how much and where to buy gasoline. For instance Smith [39] and Berkowitz et al. [3] study empirically both the intensive and extensive margin of the decision with micro data on household consumption of grocery products and gasoline respectively, using a model derived from Dubin and McFadden [16]. I chose to use pure discrete choice model here because of the nature of the data available (i.e. aggregate sales at the station level).

[8] An alternative assumption, like the GEV model used by Manuszak [25], does not have this property since no restrictions are imposed on the quality ordering of grades.

[9] Smith0 [39] similarly assumes that consumers choose among the 30 closest supermarket around their home. This assumption, albeit arbitrary, reduces the impact of the logistic error on the substitution patterns.

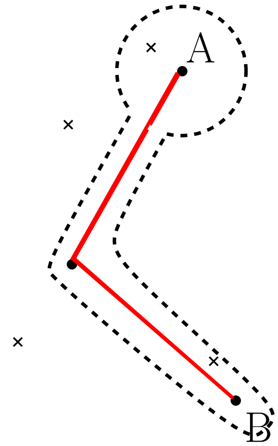


Figure 2: A buffer around path $r(A, B)$ with a home bias, and four gasoline stations (two inside, and two outside)

the distance cost to be different (lower) for stores located close to home. That is, everything else constant, consumers are indifferent between stores located at equal distance from their commuting path, and are willing to travel larger distances for stations located in a neighborhood of their home[10]. Figure 2 represents graphically this function as an iso-cost curve drawn around the path relating areas A and B. Letting b_0 be the limit of this neighborhood, the distance function is defined formally as:

$$D(r, j|\lambda) = \begin{cases} \lambda_1 d_{min}(j, r) \cdot \mathcal{I}(d_0(j, r) \leq b_0) + \lambda_2 d_{min}(j, r) \cdot \mathcal{I}(d_0(j, r) > b_0) & \text{if } j \in \mathcal{CS}(r) \\ +\infty & \text{otherwise} \end{cases} \tag{7}$$

where $\mathcal{I}(\cdot)$ is an indicator function.

The commuting behavior of consumers also has an impact on the value of the outside option. For instance, assuming away congestion problems, the value of alternative modes of transportation is typically consider higher for people located close to their workplace. To take this feature into account, without detailed information on the availability and quality of public transportation services, I model the value of using an alternative mode of transportation as a function of the time necessary to meet the transportation needs $\bar{q}(r)$. The indirect utility of the outside option is given

[10]Stations located close to home are treated differently because consumers can combine gasoline trips with other shopping trips, so that the disutility of distance is lower around their home. In addition, it is reasonable to think that most people are not willing to deviate importantly from the middle of their commuting path to buy gasoline, since this could involve additional driving costs (e.g. getting lost). In later version of the paper, different specifications of the distance function will be evaluated empirically.

by:

$$u_{i0g}(r) = -\lambda_0 p_0(t_0 + t(r)) + v_{ig} + \epsilon_{i0g}, \tag{8}$$

where p_0 is a price index of outside transportation modes (e.g. public transportation, taxi), and t_0 is obtained by converting the minimum transportation needs into units of time[11].

Given the extreme value assumption of ϵ_{ijg}, the conditional probability of buying from store j for a consumer commuting along route r takes the familiar multinomial logit form:

$$P_{j|rg} = \frac{\exp\big(\delta_{jg} - D(j, r|\lambda)\big)}{\exp\big(\lambda_0 p_0(t_0 + t(r))\big) + \sum_k \exp\big(\delta_{kg} - D(k, r|\lambda)\big)}. \tag{9}$$

The average value of choosing each grade (sometimes called the inclusive value of nest g in the nested logit context) takes the form:

$$\bar{u}_g(r) = \log\Big[\lambda_0 p_0(t_0 + t(r)) + \sum_k \exp\big(\delta_{kg} - D(k, r|\lambda)\big)\Big].$$

Using the assumption of exponential preference for the premium grade introduced earlier, the probability that consumer r buys the premium quality is given by[12]:

$$P_{g|r} = \Pr\big(v + \bar{u}_{prem}(r) > \bar{u}_{reg}\big) = \exp\Big[-\frac{1}{\mu}\big(\bar{u}_{reg}(r) - \bar{u}_{prem}(r)\big)\Big]. \tag{10}$$

Finally, the predicted demand at the station level is obtained by aggregating the individual choice probabilities over every OD pairs:

$$Q_{jg}(\delta) = \sum_{s,d} \bar{q}\big(r(s,d)\big) P_{j|rg}(\delta) P_{g|r}(\delta) T_{sd}. \tag{11}$$

Note that the substitution patterns generated by $Q_{jg}(\delta)$ will not depend strictly on the relative shares of stations, even with the multinomial logit assumption. In particular, the cross derivative of demand for product (j, g) with respect to price $p_{kg'}$ depends in a complex way on the location of consumers, the structure of the road network, and the distribution of commuting times. In that sense, the model is analogous to the random coefficient logit model advocates by BLP [5], but replaces the parametric distribution assumption with an empirical distribution of consumer heterogeneity as in Manuszak [25] and Davis [11].

[11] Since the average speed limit in urban areas is 50 km/h, the minimum transportation needs is given by: $t_0 = \frac{50km/h}{60}\frac{c_0}{c_1}$.

[12] Recall that the CDF and the PDF of the exponential distribution, defined over the positive support of the real line, take the following form:

$$f(v) = \frac{1}{\mu}\exp(-\frac{1}{\mu}v) \quad \text{and} \quad F(v) = 1 - \exp(-\frac{1}{\mu}v).$$

3 Estimation Strategy

The set of preference parameters to be estimated are given by:

$$\theta = \{\mu, \alpha, \lambda_0, \lambda_1, \beta\},$$

where μ is the mean valuation of the high quality grade, α is the price sensitivity parameter, λ_0 is the travel cost for the outside alternative, λ is the vector of parameters entering the transportation cost function, and β is the vector of parameters entering the station characteristics value equation. Note that the parameters entering quantity equation (i.e. c_0 and c_1) are fixed in the estimation, a common practice in the literature. The dataset used to estimate the preference parameters is an unbalanced panel of observed sales and product characteristics: $Y_t = \left\{\left\{q_{jgt}, p_{jgt}, X_{jgt}\right\}_{j\in J_t}\right\}_{g=reg,prem}$, $t = 1..T$.

The methodology used to estimate the model follows the techniques developed by Berry [4] and Berry, Levinsohn and Pakes [5] to estimate discrete choice models of demand using aggregate data. The key insight in Berry [4] is a method for inverting the demand system in equation (11) to recover the vector δ measuring the mean value of each product defined in equation (1). In particular, given a value for the parameter vector θ, the system of demand equation is inverted numerically to recover the vector $\{\delta_{jgt}\}$ using the following contraction mapping[13]:

$$g_{jgt}(\delta) = \delta_{jgt} + \log(q_{jgt}) - \log\left(Q_{jgt}(\delta|\theta)\right). \tag{12}$$

Since δ is a linear function of product characteristics and prices, a GMM estimator can be use to estimate the model by dealing with the endogeneity of prices (and potential measurement error) without assuming a parametric distribution function for the structural error term ξ_{jgt}.

An endogeneity problem arises for two reasons. First, gasoline prices are known to adjust very frequently, on a weekly or even daily basis (see for instance Noel [28]). This will introduce a measurement error in prices if the adjustment process affects the position of a station in the period-distribution of prices. For instance, the observed sales of a station might have been generated by a low price relative to the other stations in the market, but the observed price at the end of the period could reflect a high relative price.

[13] The computation cost of the inversion step is rapidly increasing in the number of markets (or time periods) and products. In order to increase the speed of convergence, I use a Newton-Raphson (or Broyden's) root-finding algorithm, and parallelized the task so that each processor is inverting the demand system for a subset of the sample periods.

Secondly, since firms and consumers observe the quality index ξ_{jgt} when making their decisions, prices will adjust in the short run to changes in this unobserved product quality. To further understand the sources of this correlation, I decompose the unobserved quality into a permanent and transitory component: $\xi_{jgt} = \bar{\xi}_j + \nu_{jgt}$. Defining the set of observed characteristics as $X_{jgt} = [X^1_{jgt} : X^2_j]$, the expression for δ can be rewritten as:

$$\begin{aligned} \delta_{jgt}(\theta) &= \iota_{jgt} d_j + X^1_{jgt}\beta_1 - \alpha p_{jgt} + \nu_{jgt}, \\ \text{where,} \quad & d_j = X^2_j \beta_2 + \bar{\xi}_j, \end{aligned} \tag{13}$$

where d_j is station j fixed effect, and ι_{jgt} is a dummy variable. The set of time varying station characteristics include the number of pumps for each grade, the number of service islands, the type of service, the type of convenient store associated (if any), dummy variables indicating whether the station offers car-repair and/or car-wash services, a major chain indicator, and a set of time-dummy variables capturing unobserved period-specific variables (e.g. weather). The number of location specific variables is much more restraint, and includes only indicators for the type of area surrounding the station (i.e. commercial, residential, industrial, etc.). Since the observed station characteristics are only associated with the configuration of stations, the fixed component $\bar{\xi}_j$ refers to characteristics of the location associated mainly with the road network and the organization of the city. These include for instance how easy it is to enter the parking lot of the store, on which side of the street the station is located. The second component is associated with temporary changes to the quality of the station affecting all consumers equally (e.g. changes of employees, temporary road repair).

I use two sets of moment conditions to identify the model[14]. Following BLP and Nevo [27], the first set of moment conditions are obtained by combining an Instrumental-Variable (IV) approach, with fixed-effects at the station location level. If $\tilde{w}_{jgt} = w_{jgt} - \frac{1}{n_j}\sum_{tg} w_{jgt}$ denotes the within transformation of a variable w_{jgt}, this first set of moment conditions for grade g are given by:

$$\bar{g}^1_g(\theta) = \frac{1}{J}\sum_j g^1_{jg}(\theta) = \frac{1}{J}\sum_j \sum_t \nu_{jgt}(\theta)\tilde{W}_{jgt} = 0, \tag{14}$$

[14]Manuszak [25] uses a different identification strategy to estimate a similar model of spatial competition in gasoline markets. In particular, he imposes the Nash equilibrium condition on prices, as in the original application of Berry, Levinsohn and Pakes [5]. I chose to use a different identification strategy to avoid the misspecification bias induced by imposing a potentially invalid pricing rule. It is likely to be the case in my application because of the price regulation, and the fact that retail markets for gasoline are characterized by alternating periods of price wars and (tacit) collusion.

where W_{jgt} is a $1 \times l$ vector of instrumental variables. Note that these empirical moment conditions are expressed separately for each grade as in the Seemingly-Unrelated Regression (SUR) specification. Also, the unit of observation is a station physical location, rather than a pair of location and time-period. As in BLP, this specification admit free auto-correlation patterns.

BLP proposed a set of IVs which measure the firm's own and rivals' product characteristics. These variables are valid instruments since they enter naturally the equilibrium pricing rule in any Bertrand game with product differentiation. Various authors have proposed other instruments based on cost shifters or competing markets prices (see for instance Nevo [27]). The BLP instruments are very appealing in the product differentiation context however, because researchers rarely have access to measures of costs with enough cross-sectional and time variation, while observed product characteristics are by definition readily available. In the spatial differentiation context, Davis [11] and Manuszak [25] constructed similar instruments based on the fact that competition is highly localized in most spatial differentiation models. Using this insight, they construct measures of rivals' product characteristics in a neighborhood of each store, defined as distance buffers around the sites. Following these suggestions, I construct IVs which measure stations' neighborhood characteristics. Various definitions of neighborhood are also considered (e.g. distance buffers, common streets index). The exact definition of the instruments is provided in section 4.3.

The presence of important time-invariant unobserved product characteristics can reduce the validity of these instruments. In particular, the exogeneity assumption of the station characteristics with respect to the structural error ξ_{jgt} becomes harder to maintain. In our context for instance, the unobserved quality of a location is likely shared with close-by stations. The instruments measuring the characteristics of neighborhood competitors will then be correlated with the fixed component of the structural error term. The fixed-effect approach get around this problem by assuming instead that the transitory component of the unobserved quality of products is exogenous to the within variations of the stations characteristics $\tilde{X}^1_{jgt}$ and instruments $\tilde{W}_{jgt}$.

This approach is particularly well suited here, because it exploits the richness of the time variation induced by the entry, exit and reconfiguration of stations over the sample period. Over the period studied (i.e. 1991 to 2001), the North-American gasoline retail industry underwent a major reorganization, associated with massive exit and entry of new categories of stations (see Eckert and West [17] and Houde [20]). These changes were mainly due to technological innovations common to most retailing sectors which increased the efficient size of stations (e.g. automatization of the service, better inventory control systems), as well as changes in the preferences of consumers for certain amenities (e.g. decline in the need for car repair shops). Since these changes are not specific to the Québec City market, they provide a valid source of exogenous variation in the instruments, even after controlling for station fixed effects. Without such inter-temporal variation in the market structure, the within transformation of the variables would eliminate all variation in the instruments. Moreover, adding geographic markets in the analysis, as in Nevo [27], would not solve this lack of variation problem. This is because the product fixed effects are by definition specific of each geographic market, contrary to other product differentiation applications in which we observe the same products in different markets[15].

The second set of moment conditions are obtained by matching the proportion of people who are using their car to go to work (or study), with the empirical frequencies from the 2001 OD survey. In particular, letting q^0_{sd} denotes the observed number of car users for the pair of origin and destination traffic zones (s, d), I compute the following S moment conditions:

$$\bar{g}^2_s(\theta) = \frac{1}{D}\sum_d P_{0|sd}(\theta)T_{sd} - q^0_{sd}. \tag{15}$$

As in Petrin [30], this additional set of moment conditions enhance the identification of the non-linear parameters by providing further restrictions on the transportation cost (λ) and the value of the outside alternative (λ_0).

Joining equations 14 and 15, the GMM objective function is given by:

$$\begin{aligned} Q(\theta) &= J\bar{g}(\theta)\Phi^{-1}\bar{g}', \\ \text{where} \quad & \bar{g} = [\bar{g}^1_{reg} : \bar{g}^1_{prem} : \bar{g}^2] \end{aligned} \tag{16}$$

Assuming independence accross observations (i.e. locations), the efficient (heteroskedastic-consistent) weighting matrix Φ is calculated in two steps using the $\Phi = \tilde{W}'\tilde{W}$ initially (see Davidson and MacKinnon [10]). Note also that the weighting matrix is block-diagonal as in Petrin [30], since the two moments are calculated from different sample. See Imbens and Lancaster [21] for further details.

[15] Nevo's application was the ready-to-eat market, for which he observed the same brands in a cross section of cities over time.

4 Description of the Data

In the three following subsections, I describe the sources and the construction of the variables used to calculate the empirical distribution of commuters over time and within the market, and the gasoline station data used to estimate the model.

4.1 Distribution of Commuters

The starting point for the estimation of consumer heterogeneity is an OD survey conducted in 2001 by the Québec government in the Québec city CMA. The results of the survey are available in the form of aggregate OD tables, providing the predicted number of commuters between every pair of Traffic Area Zones (TAZs)[16]. The tables are available for four categories of trips and transportation mode and three periods of the day. I use the results of two OD tables for every transportation modes, representing work and study trips over a full day window[17]. The relevant OD matrices include every transportation modes, since the model includes an outside option. Each OD matrix is then converted into a commuting probability matrix by dividing each row by the total number of trips generated by each origin TAZ.

Additional information on the distribution of employees and schools is used to disaggregate the OD matrices into smaller location areas. Further disaggregation is required because the TAZs are too large to represent accurately consumer locations relative to stores. The final definition of locations is fully nested in the original TAZ definitions. These areas represent a mixture of census tracts (CTs) and dissemination areas (DAs) (i.e. the smallest statistical neighborhood). Further details are provided in Appendix A. Table 2 presents a set statistics describing the distribution of population across location areas. Note that the distribution of area size is highly skewed because the CMA territory includes a set of rural fringes.

To compute OD probabilities for each pair of locations, I combine the original OD matrices with data on the distribution of jobs across DAs obtained from the Canada Business Summary database assembled by MapInfo Canada. In particular, I fix the destination probabilities within each TAZ

[16] The aggregate OD matrices are freely available on the ministry website: http://www1.mtq.gouv.qc.ca/fr/services/documentation/statistiques/enquetes/index.asp. The survey report available on the same website provides further details on the conduct of the survey [12].

[17] Adding leisure and shopping trips is conceptually feasible, but would increase significantly the computation cost of the model. In addition, since consumers are commuting to their workplaces on a daily basis, it is the most appropriate commuting path to characterize preferences for gasoline station locations.

Table 2: Descriptive Statistics of the Location Areas

	Mean	Sd-Dev	Q25	Q50	Q75
Population	1089	1310	385.5	513.8	1076
Workers	629.8	743	227.6	311.3	639.5
Students	117.5	164.8	40.52	62.57	108.4
Area size (km^2)	6.296	28.47	0.1809	0.4991	2.536
Number of areas	501				
Number of OD pairs	18,841				

to be proportional to the number of employees working in each location area. Similarly, the OD matrix for study trips is disaggregated using the distribution Colleges and Universities across the city, obtained from the DMTI Enhanced Point of Interest database. The resultant matrices of commuting **probabilities** (Ω^w, Ω^s) are then fixed for all sample periods, assuming that the spatial distribution of workers and students is stationary over the period studied[18].

To compute the **number** of commuters between every area pair at each time period t, that is T_{sd}^t, I assume that the population distribution within each census tract (in which the location areas are nested) is constant. I need to make this assumption because the definition of dissemination area was created for the 2001 Canadian census, and no information is available prior to that date at this level of aggregation. However, Statistics Canada uses a definition of census tract which is comparable between censuses. By assuming that the population weights within each census tract is constant over time, I can use the average growth rate of the census tract population to interpolate the population weights of each location area throughout the sample. Using these population weights, I then use the monthly Canadian Labour Force survey (available at the CMA level) to predict the month-to-month variations in the relevant population measures of each area (i.e. students, workers, and total adult population). Appendix A describes in greater details the calculation of T_{sd}^t.

4.2 Traffic Distribution

The distribution of commuters across the road network is calculated by combining the empirical distribution of commuters between each OD areas, with the shortest driving path assumption discussed in the model section.

[18]This assumption will be relaxed in the future, by integrating results from two other OD surveys conducted in 1996 and 1991.

Data on the Quebec city CMA street network is coming from the CanMap RouteLogistics database, obtained from DMTI Spatial [13]. This database provides detailed information on the Canadian road and highway network, including turn restrictions, speed limits, segment length, and travel time. Table 3 describes the size of the road network data.

Table 3: Description of the Québec City Road Network

Number of nodes	$23,394$
Number of arcs	$32,167$
Avg. length (meters)	$217,8$
Avg. travel time (minutes)	0.28

A total of $1,200$ shortest path trees (SPT) were computed for every originating DAs, using a version of the Dijkstra algorithm[19]. The commuting time and path length are also obtained as a by-product of the shortest path algorithm.

Business travelers and tourists possibly represent a large fraction of gasoline sales in some regions, especially along expressways surrounding the city. However, the Origin-Destination survey used to compute T_{sd} does not consider this category of commuters. To rectify this problem, I add an "outside commuters" category to the model by adding K new types of consumers located on each segments of the main expressways. Each one is characterized by the set of connected nodes forming one segment of the expressway network of the city[20]. The number of outside commuters in a given period is modeled as a function of the number of tourists, and a measure of the trade intensity in the region. The number of tourists is measured by the number of occupied hotel rooms, and the trade intensity is measured by the real value of manufacturing shipments in the province. The choice of the weights assigned to each variable, w_1 and w_2 respectively, is discussed in section 5.1. Table 4 presents a set of descriptive statistics for these two macro variables used to predict the number of outside commuters. Note that the scale of the shipment variable is expressed in $100,000$ real dollars. The relative value of the weights (w_1, w_2) will therefore adjust accordingly

[19] The computation of each SPT was performed using the matrix programming software Ox [14], and executed in parallel on the HPCVL supercomputer network. The computation of each SPT takes between 2 and 3 CPU hours, so that the total task would take several months if executed on a serial architecture. Instead, the task was performed in parallel, by distributing the calculation of each SPT over a set of processors, reducing the computing time to less than a week.

[20] A segment of expressway is defined as the set of expressway arcs with the same name. In the Québec CMA, there is 10 such segments

Table 4: Descriptive Statistics of Macro Variables

	Mean	Std-Dev.	Min	Max
Public Transportation Price Index (1992 = 100)	106	3.6	98	112
Number of Occupied Hotel Rooms	6261	1853	3449	9778
Manufacturing Shipments in the Province of Québec ($X100,000$ in 1992 \$CA)	19	4.1	11	27

Sources: Monthley Survey of Manufacturing (Statistics Canada), Consumer Price Index (Statistics Canada), Qubéc Ministry of Tourism [34][35]

(i.e. higher value for shipments than for tourists). The table also includes descriptive statistics for the consumer price index of public transportation (i.e. taxi and local public transportation services). This last variable is used to measure the price of the outside option in equation 8.

Figure 3 presents a sample of the commuting paths generated by local commuters (top), and by outside commuters (bottom). The top figure has been computed by choosing one specific dowtown location (located at the center of the graph), and drawing a sample representing 10% of the paths going to this point (reprensented by the dotted lines). The darker paths represent high traffic routes, according to this sample of locations. A quick look at the top-figure reveals that the algorithm accurately finds the fastest driving paths since most routes are straight and exhibit few “detours”. The localization of stations in the last period of the sample is represented by the crosses. As we can see from the picture, the localization of many stations corresponds exactly to some of the main driving paths. Note also that the lines generating the driving paths of outside commuters (bottom figure) are much thicker because expressways have multiple drive-ways and exit-lines.

4.3 Gasoline Station Data

The gasoline station data used to estimate the model are collected by Kent Marketing, the leading survey company for the Canadian gasoline Market. The panel spans 46 bimonthly periods between 1991 and 2001 for every firms in the Québec City market (i.e. once a year between 1991 and 1994, and six times a year between 1995 and 2001). The survey offers very accurate measures of sales and station characteristics since each site is physically visited at the end of the survey period.

Table 5 presents a set of descriptive statistics for the variables which vary by grade. That is, the posted price at the end of the survey period, the quantity of gasoline sold per day, and the number of pumps. Note that for about 10% of the sample, firms refused to participate in the survey for some

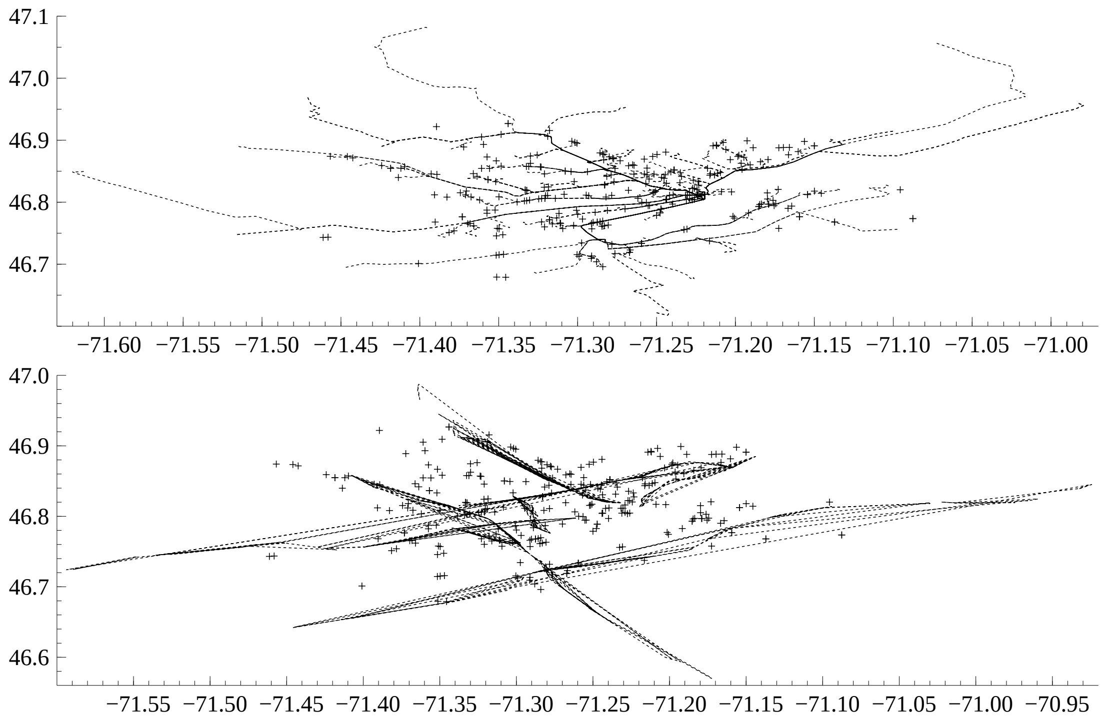


Figure 3: Sample of Commuting Paths for Local (top) and Outside (bottom) Commuters. The axes are the longitude and latitude coordinates of the points.

or every periods. For those observations, the station characteristics (including prices) are accurately measured, but the volume sold is not available. Since the inversion procedure cannot accommodate missing values in the market share variable, I imputed the missing values using a linear regression method. The set of the explanatory variables include the average neighborhood market shares, a polynomial function of the geographic coordinates of the locations, prices, characteristics and lagged sales (for stations who were previously participating in the survey)[21].

The set of observed station characteristics include the type of convenience store, a car-repair shop indicator, the number of service islands, the opening hours, the type of service, and an indicator for the availability of a car-wash. A major brand indicator is also added to the set of characteristics to reflect that consumers might view major gasoline brands as higher quality. The majors include five retailing chains who are integrated in the refinery sector: Shell, Esso/Imperial Oil, Ultramar,

[21]Similar imputation methods, using Kernel estimators or polynomials, have been used in similar contexts by Hotz and Miller [?] to permit the use of an inversion procedure. In their context the missing values correspond to empty choice probability cells.

Irving, and Petro-Canada. The sample includes 14,255 observations on each grade, for 429 different gasoline station sites. These stations represent a total of 962 different products, defined as a fixed set of characteristics (i.e. combination of islands, convenience store category, car-wash, repair-shop, and major brand indicator), and a physical location. On average each product is observed for 14 consecutive periods.

Table 5: Descriptive Statistics by Grade of Gasoline

		Volume Sold	Price	Number of Pumps
Regular	Mean	4,910.46	62.09	6.71
	Std-Dev.	3,140.90	7.35	4.71
	Min	33.03	36.90	1.00
	Max	27,527.98	87.40	24.00
Pemium	Mean	515.72	70.08	3.31
	Std-Dev.	390.93	7.20	2.62
	Min	0.02	42.90	1.00
	Max	4,058.89	97.40	24.00

Volumes are expressed in liters per day. Prices are expressed in cents/liter

Figure 4 summarizes the trends in the industry over the sample periods. Figure 4(a) shows that both aggregate prices and demand are very cyclical, and have been steadily increasing over the 11 years. These upward trends reflect the improvements in economic conditions and the city population, and the sharp increase in the world price of oil after 1999. As shown in Figure 4(b), the majors have steadily increased their market share, without expanding their retail network. The next two Figures (i.e. 4(c) and 4(d)) characterize the reorganization of the industry. In particular, the proportion of stations with a convenience store, and the proportion of self-service stations have increased by roughly 30% and 24% respectively. At the same time, the proportion of stations offering car-repair services dropped by 15%. These changes have been induced by the exit of a large number of small-capacity stations and the entry of large capacity ones, as illustrate by Figure 4(d). Note finally that it is mainly the major chains who are responsible for these changes, which explained their larger market share at the end of the period.

Table 6: Description of Station Characteristics

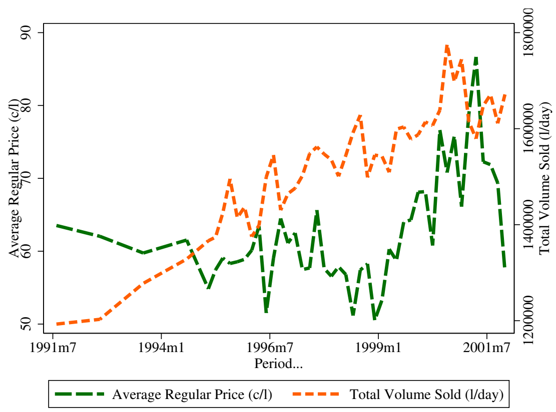
	Mean	Std-Dev.	Min	Max
Convenient Store Size	1.4	1.3	0	3
Repair Shop	.17	.37	0	1
Number of Islands	2.3	1.4	1	8
Opening Hours Category (24 hours = 1)	.4	.49	0	1
Type of Service (Self-Service =1)	.62	.48	0	1
Carwash	.2	.4	0	1
Major Brand	.64	.48	0	1

4.3.1 Instrumental Variables

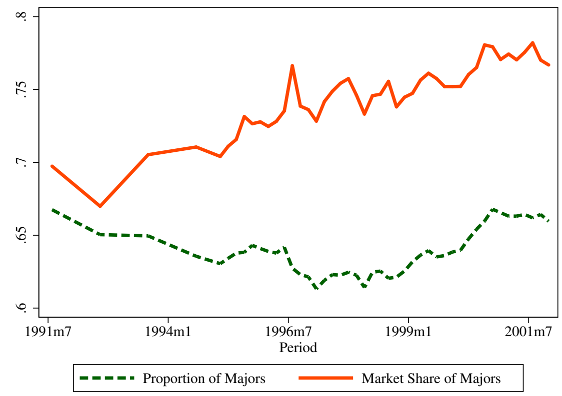
I use 8 BLP-type instruments to correct for the endogeneity of prices. As discussed in the estimation strategy section, these instruments measure the strength of competition in a neighborhood of each station. I use a two definitions of neighborhood which incorporates information on the road network, and a measure of distance between stations. The first definition is such that two stations are considered to be in the same neighborhood if they are located on the same street, and if the driving time between them is less than 5 minutes[22]. The size of each neighborhood is relatively large, and incorporates the notion present in the model that consumers can substitute easily between two stations if they are located along the same commuting path. The second definition constructs three different euclidean distance rings around each location (i.e. 100 meters, 300 meters, and 1 kilometer).

Using these definitions, and following the work of Davis [11] and Manuszack [25], I construct variables characterizing the local market structure of those neighborhood. I choose in particular two physical characteristics of stations which were affected by the reorganization of the industry, to increase the inter-temporal variation of the instruments. For each definition of neighborhood, I calculate the average capacity of stations (measured by the average number of pumps), and the proportion of stations with car-repair shop. In addition, as in Davis [11] and Manuszack [25], I add two measures of local market size, by calculating the number of local commuters within a 300 meters buffer around each station, and the number of people living in a 1 kilometer buffer around each station. The total number of instruments is 10.

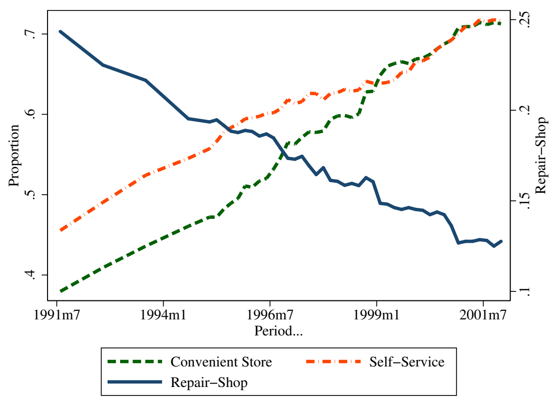
[22]The driving time between stations was computed using a similar shortest path algorithm used to compute the traffic distribution.



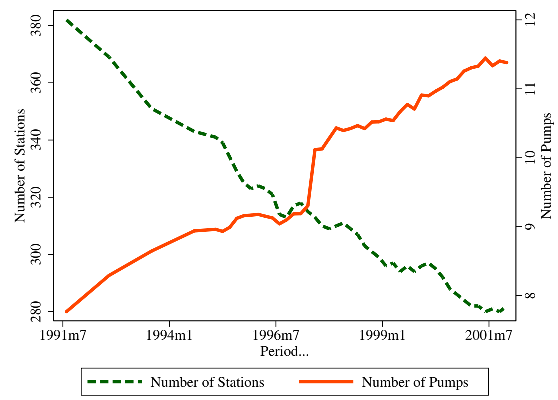
(a) Volume Sold and Prices



(b) Market Share and Proportion of Majors



(c) Changes in the Type of Amenities



(d) Number and Average Capacity of Stations

Figure 4: Trends in the Market

Tables 7 presents a set of descriptive statistics for the instrumental variables. Note that the two variables measuring the size of the local market surrounding each station are standardized by their respective mean. The last two columns also present the coefficients and t-statistics of a fixed-effects regression of prices on station characteristics and instruments. The F-statistic presented a the bottom of the table test the hypothesis that the coefficients associated with the instruments are jointly equal to zero. The results of this regression clearly show that the instruments are correlated with prices, even after controlling for location fixed-effects. Most t-statistics are above 2, and the F-statistic easily reject the null hypothesis that the coefficients are zero. Although the most regression coefficients are difficult to interpret, the effect of the number of people living around

Table 7: Descriptive Statistics of the Instrumental Variables

	Mean	Std-Dev	First-Stage OLS Coef.	t-stat.
Supply-side IVs				
Avg. capacity along street	2.4	4.1	0.049	3.6
Avg. capacity in $(0, 0.1km]$	1.7	3.6	0.26	1.2
Avg. capacity in $(0.1km, 0.3km]$	3.8	4.6	0.03	2.6
Avg. capacity in $(0.3km, 1km]$	6.4	2.7	-0.051	-5
Avg. car-repair along street	0.047	0.18	0.037	2.6
Avg. car-repair in $(0, 0.1km]$	0.045	0.2	-0.027	-0.17
Avg. car-repair in $(0.1km, 0.3km]$	0.081	0.24	-0.68	-3.6
Avg. car-repair in $(0.3m, 1km]$	0.15	0.18	-0.44	-2.1
Demographics				
Home Buffer: $(0, 1km]$	1	0.87	0.72	3.8
Traffic Buffer: $(0, 0.3km]$	1	1	-0.24	-1.4
F-test: $(\beta_{IV} = 0)$			7.32 (0.00)	

each station is positive and significant. This suggest that prices are higher close to highly dense neighborhoods.

5 Empirical Results

The empirical analysis is divided into three sections. I first present a comparison of the predictions of a standard home-address model with the model incorporating commuting. I then estimate a simpler reduced form version of the structural model to assess the validity of the assumptions and the identification strategy. Using estimates of the structural preference parameters, I then analyze the economic predictions of the model. In particular, I investigate the sources and the importance of the market power of major retail chains.

5.1 A First Look at the Data

To evaluate the usefulness of adding commuting into the standard address model, I first construct simple measures of local market sizes based on the two models. I then use these variables to compare the predictions of the two models.

To measure the size of the local markets faced by each station, I calculated the number of people who live within 1 kilometer or commute within 300 meters of the location (weighted by the quantity $\bar{q}(r)$ consumed as defined in equation (5)). By adding outside commuters, measured using the number of tourists and the value of shipments, I constructed four definitions of local markets. The first one considers only the home location buffer, the second considers only the distance from path buffer for local commuters, the third combines the first two as in Figure 2, and the fourth one adds the outside commuters located along the main expressways of the city.

Using these four variables, I calculated the predicted demand at each station, assuming that consumers choose at random one station from their distance buffers, hence abstracting from the effects of prices and station characteristics. Note that the weights assigned to the number of tourists and the value of shipments (i.e. denoted by w_1 and w_2 respectively), have been calibrated to maximize the correlation between the predicted demand defined in this way and the market share of each station. These two parameters will be fixed in this way for the remaining of the paper.

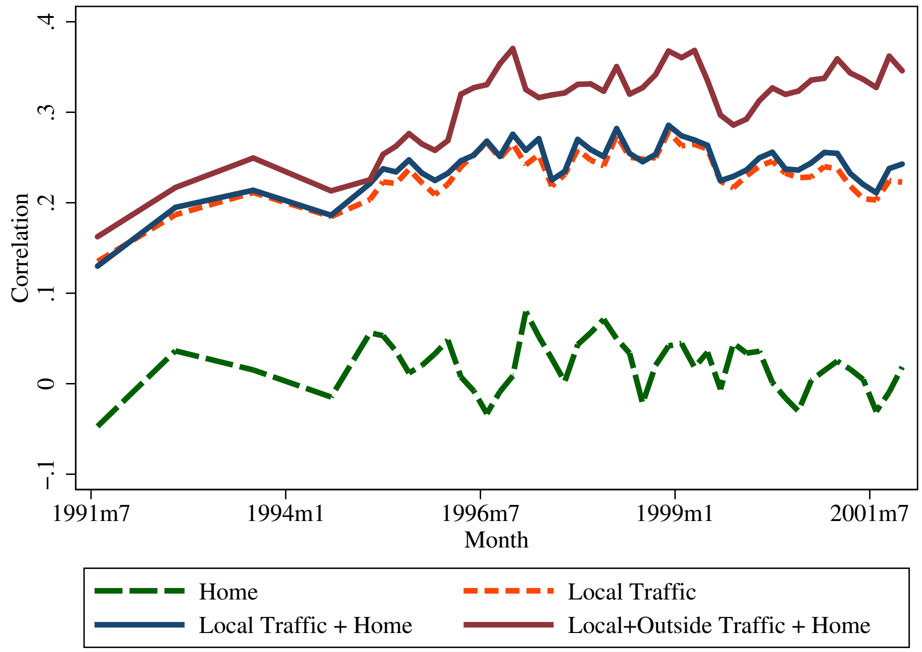
Figure 5 presents the evolution over time of the correlation between the predicted and observed market shares calculated from this simple buffer model. The comparison between the home location model and the commuting model is striking. By locating consumers along their work-commuting path the correlation increases from 0 to 0.25 on average. Furthermore, adding outside commuters increases the average correlation to 0.36. This indicates that stations located close to where people live are not the ones with the highest market shares. Furthermore, a large fraction of gasoline purchases is generated from people's commuting behavior. Another important result from this figure is that the correlation is relatively stable overtime, even though the OD probability matrix used in the computation is held fixed. This suggests that the commuting behavior of workers has not changed dramatically over the sample period.

5.2 Estimates of a Linear Version of the Model

Next I estimate a linear version of the structural demand model. This helps illustrate the validity of the identification strategy, and further compares the predictions of the model with the home-address model. In particular, by eliminating the consumer heterogeneity component of the model, Berry [4] shows that the demand system in equation (11) can be inverted analytically to recover δ_{jt}:

$$\log(q_{jt}) - \log(q_{0t}) = \delta_{jt} = X_{jt} - \alpha p_{jt} + \theta B_{jt} + d_j + \kappa_j + \nu_{jt}, \tag{17}$$

Figure 5: Correlation between predicted demand and observed sales by year



where q_{0t} is the demand of the outside option. It is obtained by assuming that the quantity of gasoline consumed per day is given by equation (5), with $c_0 = 3$ and $c_1 = 0.2$. These values imply that people on average need 25 minutes of transportation for motives other than work commuting. The value of c_1 has been chosen to match the average gasoline consumption of cars by kilometer for a round work-commuting trip (i.e. 0.1 l/km times two).

As in Davis [11], I incorporate the location of consumers in the linear model by adding a measure of local market size in equation (8) (denoted by B_{jt}), using the previously defined distance buffers. This linear model can therefore be seen as a reduced form version of the non-linear model presented earlier, replacing consumer locations by a measure of the local market size around each station.

Table 8 presents the estimated results for the regular grade, under four different specifications. The first two columns present the specifications with the home location buffer, and the two next present the results measuring the local market size as the number of commuters. Columns 1 and 3, and columns 2 and 4 represents a two econometric models: fixed effects solely (FE), and with fixed effects and IVs (GMM+FE).

Table 8: GMM Estimation of the Linear Model for Regular Gasoline

	Home Location		**Commuting Location**	
	FE-OLS	FE-GMM	FE-OLS	FE-GMM
Price ($-\alpha$)	-0.02188	-0.1414	-0.02194	-0.1479
	(-15.48)	(-5.278)	(-15.53)	(-5.3)
Local Market Size (θ)	-0.05614	0.01672	0.1002	0.1056
	(-1.801)	(0.3717)	(3.566)	(3.2)

T-statistics are in brackets. Dependent variable is $\delta_{jt} = \log(q_{jt}) - \log(q_{0t})$. Each specification also include time dummies, and station characteristics. The buffer variables are standardized by their mean.

Looking first at the home-location specifications (columns 1 and 2), the coefficient on the size of the residential neighborhood has a negative or insignificant effect[23]. When measured as the number of commuters (columns 3 and 4) the size of the local market has the anticipated positive and significant sign. Comparing the fixed-effect and the GMM specifications, the price coefficient is negative and significant in the two, but larger in absolute value in the GMM columns.

The results comparing the home location model with the commuting model are in line with the simple unconditional correlations in Figure 5. Table 8 shows that even after controling for prices and station characteristics, the conditional correlation between the residential neighborhood size is either small or negative. This reinforces the conclusion that a model based solely on home location cannot explain the observation that a large fraction of gasoline sales are realized far from densely populated areas, but close to the main commuting paths.

Another result from Table 8 is the evidence of an important endogeneity bias in the price coefficient (α). The second and fourth columns show that the within product transformation and the IVs are successfully correcting for the downward bias (in absolute value) in the price coefficient.

5.3 Estimates of the Structural Model

Table 9 presents the GMM estimates of a restricted version of the commuting and the home location models, considering only regular grade gasoline. The functional form used for the transportation cost also ignore any home-bias (i.e. $b_0 = 0$ in equation 7). The middle rows of Table 9 summarize the value of the parameters which were held fixed in the estimation. As discussed previously, the pa-

[23]The magnitudes of the coefficients are comparable across specifications because the local market sizes are standardized by their mean.

rameters w_1 and w_2 are calibrated using the simple buffer model, which ignores the value of station characteristics. The estimates for the parameters measuring the value of stations' characteristics are provided in Table 10.

Looking first at the estimates of the non-linear parameters in Table 9, the parameters are all estimated very precisely. The three parameters are also quite different under the two models. The the price sensitivity parameter and the distance cost are both larger in column 1 than in column 2. The value of the outside option is also estimated to be twice smaller under the home location. Statistically, the commuting model has a smaller value for the over-identification test (or $J - statistic$). In particular, the p-value for the Sargan-Hansen test is 0.14 for the commuting model, and 0.07 for the home-location model. The orthogonality conditions are therefore more likely to be met under the commuting model.

The impact of these differences lead to starking differences in the models' predictions. The fourth and fifth rows of Table 9 present the price gap required for a consumer to travel one kilometer (or one mile), measured by the ratio λ/α. Under the commuting model, consumers are willing to deviate 1 kilometer from their path if they can save at least 4.79 cents per liter (or 25 US cents per gallon). Under the home location model, this price difference is extremely small: consumers require only 0.416 cents per liter to travel 1 kilometer away from their home (or 2.2 US cents per gallon). The implicit value of an hour corresponding those estimates are $-\$0.51$ for the home-location model, and \$26.86 for the commuting model[24]. In words, the estimates of the home-location model suggest that the monetary gains are lower than the cost of travelling one kilometer, suggesting that consumers are enjoying shopping for gasoline. Under the commuting model, consumers are incurring significant shopping costs, and the implicit value of an hour is higher than the average hourly wage (around $16\$/hour$).

These different conclusions from the two models are easily understood in the light of the results presented in sections 5.1 and 8. Since a large fraction of gasoline sales are not generated close to where people live, the home-location model require an unrealistically small transportation cost to

[24]These results are obtained under the following set of assumptions: (i) a typical gasoline purchase is 25 liters, (ii) the cost of travelling 1 kilometer back and forth is 12.42 cents at the average observed price in the sample (i.e. $0.1242\$/l = 2km \times 0.1l/km \times 0.62\$/l$), and (iii) it takes approximately 0.04 hours to travel two kilometers. The implicit shopping cost (in dollars per hour) is then given by:

$$\$/hour = \frac{\frac{\lambda/100}{\alpha}25 - 0.1242}{0.04}.$$

Table 9: GMM Estimates of the Preference Parameters

	Commuting Model	Single Address Model
Price (α)	0.1963	0.1358
	(0.06113)	(0.05286)
Outside option $\times 100$ (λ_0)	2.108	4.995
	(0.02217)	(0.00991)
Distance cost $\times 100$ (λ)	94.15	5.646
	(1.487)	(0.173)
$\Delta \$CA/litre \Big/ 1km$	0.04795	0.004158
$\Delta \$US/gallon \Big/ 1mile$	0.2586	0.02242
	Fixed Parameters	
Minimum quantity (c_0)	3	
Liters per km (c_1)	0.2	
Shipments (w_1)	1157	
Tourism (w_2)	106.5	
Number of Locations	429.00	429.00
Number of observations	14255.00	14255.00
J-statistic	76.35	81.11
Degrees of freedom	64.00	64.00
P-value	0.14	0.07

Standard errors are in brackets. The parameters are estimated by GMM on regular gasoline only.

explain the large market shares of stations located close to highways or high-traffic intersections. In the commuting model on the other hand, commuters encounter a large number of stations along their path. Consumers therefore have access to a large set of stations without having to incur important transportation costs.

The two models also have different estimates of the store-level price elasticity of demand. In particular, the average elasticity over the periods is -11.857 for the commuting model, and -8.38 for the home-location model. As we will see below, this will lead to different market power evaluations.

I now use the structural parameters to evaluate the source and the importance of the market power of major chains in the market. To do so, I first recover the implicit marginal cost of each station using the equilibrium conditions described in (3). In particular, given a vector of observed

Table 10: GMM estimates Station Characteristics Parameters

	Commuting Model	Single Address Model
Numbe of pumps	0.01365	0.01173
	(0.003789)	(0.003085)
Number of islands	0.01668	0.02121
	(0.02609)	(0.01995)
Self-service	−0.02049	−0.04473
	(0.0622)	(0.04948)
Mixed-service	0.1252	0.09762
	(0.06149)	(0.05075)
Micro conv.-store	0.004593	0.02057
	(0.06358)	(0.03501)
Medium conv.-store	−0.02507	−0.01663
	(0.07724)	(0.05642)
Large conv.-store	−0.02636	0.02337
	(0.1016)	(0.05997)
Car-repair	0.05563	−0.03095
	(0.1025)	(0.08838)
Carwash	0.05312	0.06656
	(0.1001)	(0.0672)
Day time	−0.342	−0.358
	(0.222)	(0.1815)
Extended time	−0.0433	−0.0721
	(0.06286)	(0.05227)
Major brand	0.1699	0.137
	(0.04985)	(0.04281)

Standard errors are in brackets. The parameters are estimated by GMM on regular gasoline only. The parameters are estimated by GMM on regular gasoline only. The unreported components of β include time dummies and area type dummies. The omitted categories are: full service, no convenient store, no repair shop, no car-wash, open 24 hours, commercial area.

prices p_t, and assuming that firms set their prices according to a static Bertrand game, the vector of marginal cost is given by:

$$w_t = p_t + \Delta(p_t)^{-1} Q_t(p) \tag{18}$$

After recovering this vector, I can analyze the equilibrium properties of the model. I chose to concentrate this analysis on one specific period, the spring of 1997, which preceded the introduction of the price floor regulation. This period is interesting because the structure of the market is the same one that has motivated the implementation of regulation aimed at protecting the independent retailers. By evaluating the market power of majors and the role of independents, I can therefore evaluate the validity of the motives which have led the government to regulate prices.

Table 11 compares the estimated average marginal cost for the two groups of stations. First note from this table that independent stations face larger costs and have lower value for consumers on average than major branded stations. These estimates imply that major chains invest more in high value characteristics, and have better locations than independent stations. Since these investments are associated with higher fixed costs, they are not represented in the marginal cost of stations.

I then compute the equilibrium prices under three counter-factual market structures: non co-operative pricing, collusion among major chains, and full collusion. To understand the reaction of firms with respect to consumer driving behavior, I also computed the nash-equilibrium prices using the estimates from home-address model. Table 12 summarizes those results.

Looking first at the second columns of Table 12, the estimated price-to-cost margins under the two models are very different. The median PTC is 43% higher when estimated with the home-location model rather than the commuting model. Three factors are affecting this difference. On one hand the average price elasticity at the station level are higher in absolute value under the commuting model, reducing the consumer surplus that firms can possibly extract (positive effect on profits). Moreover, the transportation cost estimates are higher with the commuting model, leading to more differentiation between products, and therefore higher profits under the commuting model. Finally, with the same parameters, the commuting model would naturally lead to more competitive markets, because commuters can more easily arbitrage price differences within the market without incurring transportation cost. The empirical results in Table 12 suggest that the two positive forces are dominating, and more than compensate for the smallest transportation cost in the home-location model. Leading to lower estimated profits under the commuting model.

Table 11: Comparison of Independent and Major Stations

	Majors	Independents
Median prices	19.02	18.74
Median values ($\hat{\delta} + \hat{\alpha}p$)	23.25	22.77
Median marginal costs	16.26	16.33

Prices are expressed in cents/litre. Prices, value and marginal cost are computed using the result of the commuting location model, assuming that prices are set at the brand level.

The evaluations of market power under the two models are therefore different. The estimates of the home-location model lead to higher estimate of the extent of market power in this market.

Moreover, the PTC are estimated to be much smaller than other markets with product differentiation. For instance, Nevo [27] reports median PTC margins closer to 46% in the ready-to-eat cereal, and BLP [5] reports car PTC margins that are in a 15% to 30% range. This suggest that the level of market power at the retail level is small in this market.

A third observation from this table, is that the relative gains from collusion are much higher than the gains from chain-pricing. To separate this profit gain, I define the median gain from collusion for major chains as 9.67 cents/liter (i.e. column four - column one). Decomposing this gain between the gain from chain-pricing and the gain from the collusion, we obtain that 11.6% of this gain is associated with having multi-product firms, 55.6% is due to the collusion between majors, and 32.8% is due to the collusion between independents and majors. Since the proportion of independent stations was 62.2% in 1997, this result suggest that independent retailers do not have a strong impact on the reduction of market power from major chains.

Examining more directly the role of independent retailers in reducing the market power of major chains, I consider two sets of policy experiments. In the first scenario, I create a merger between the independents, so that the market is now formed of 6 major chains of retailers instead of 5. In the second experiment, I consider the extreme example in which all independents simultaneously exit the market.

The results of the two policy experiments are summarized in Table 13. Not surprisingly, the level of prices and profits increase substantially under both scenarios compared to the data. What is surprising is that prices are higher under the merger simulation, than under the massive exit of

Table 12: Median Equilibrium Price-To-Cost Margins of the Home and Commuting Models

All Stations

	Non-Cooperative	Chain Pricing	Major Collusion	Full Collusion
PTC (Commuting) in %	13.03	13.89	18.7	22.75
PTC (Home) in %	18.59	19.84	25.53	28.93

Major Stations

	Non-Cooperative	Chain Pricing	Major Collusion	Full Collusion
PTC (Commuting) in %	13.11	14.23	19.61	22.78
PTC(Home) in %	18.73	20.13	26.18	29

Independent Stations

	Non-Cooperative	Chain Pricing	Major Collusion	Full Collusion
PTC (Commuting) in %	12.86	13.05	13.12	22.69
PTC (Home) in %	18.49	18.63	18.67	28.87

$PTC = 100(1 - mc/p^*)$. The equilibrium margins are calculated using the estimated marginal costs calculated under the assumption that the data are generated from the "Chain-Pricing" market structure. The margins correspond to the spring of 1997.

stations. An explanation for this result is that the independents are on average the least efficient stations, both in terms of their cost and the value of their characteristics. By removing a large fraction of high cost firms, the second scenario increases the average price relative to the data, but less than in the first scenario.

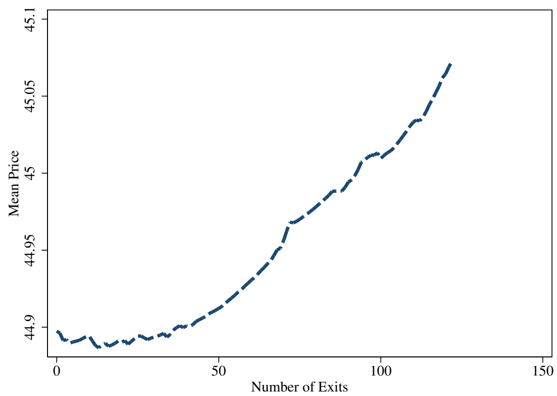
Table 13: Results of Two Policy Experiments

	Data	Policy 1: Merger	Policy 2: Exit
Median price	18.91	19.34	19.06
Median profit margin	2.633	3.003	2.76
Min profit margin	2.392	2.651	2.581
Max profit margin	3.379	5.021	3.776
Nb. Brands	5	6	5
Nb. Stores	310	310	193

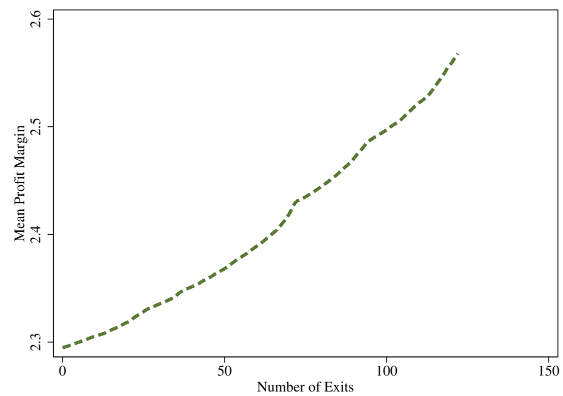
Prices and profits are measured in cents/litre. The first policy experiment refers to the merger of independent retailers. The second policy experiment refers to the exit of all independents.

The results of these two policy experiments confirm that independents are effective at lowering prices by reducing the market power of major chains. However, because they are also less efficient

Figure 6: Counter-Factual Rationalization Among Independent Stations



(a) Effect of the Rationalization on Prices



(b) Effect of the Rationalization on Profits

on average, they create an upward pressure on prices. To illustrate this trade-off between market-power and efficiency, I conducted a gradual counter-factual rationalizing of the industry. More specifically, I ordered every independent in the market according to their profitability, and simulate a sequential exit process from the least profitable independent to the more efficient one. At each step, I computed the equilibrium prices and profit-margins. The results are presented in Figure 6.

As anticipated, while the profit level goes up as more stations exit the market, the level of prices initially goes down due to efficiency gains. In this example, the average equilibrium price first begins to falls as inefficient independents are dropped but begins to go up as concentration reaches a sufficiently high level. From the figure, the average price in the market starts increasing relative to the statu-quo after the 40^{th} station exited the market.

6 Conclusion

In this paper I developed, computed, and estimated a novel model of demand for spatially differentiated products, applied to retail gasoline. My approach contributes to the literature on spatial differentiation by formally modeling commuting paths as the "locations" of consumers. This extension of the standard home-address model generates substitution patterns which depends in an intuitive way on the structure of the road network and the direction of traffic flows.

The methodology combines computing tools from the transportation Geographic Information System literature, and econometric methods developed to estimate discrete choice models of demand. The estimation of the model is performed on a unique panel dataset, which covers an important period in the evolution of the gasoline retail industry. This period is characterized by a large North-American re-organization of retail networks induced by several technological innovations. This feature of the data enables us to identify the structural parameters of the model, even after controlling for important unobserved characteristics of store locations.

The results validate the modeling choices in many ways. In particular, the distribution of gasoline sales within the market is shown to be poorly correlated with the distribution of population, and significantly more so with the distribution of commuters. In fact, this correlation is negative for some specifications studied, suggesting that a large fraction of gasoline demand is generated in areas far from where people are living. Moreover, the preliminary estimates indicate that the substitution patterns predicted by the model are different from the ones generated by a single home-location model. Since the relative substitution of locations feeds directly into predictions of mark-ups and prices, the results present a more reliable evaluation of market power in the market than the standard model.

The market equilibrium simulations performed with the preliminary estimates offer important insights for understanding the sources of market power of retail chains. First, the ability of major chains to exercise market power is tightly linked with their ability to organize their network of stores in ways which maximizes their profits given the driving behavior of consumers. In addition, the five major chains are shown to be more efficient than independent retailers. This difference is reflected by a higher quality of their stores, and lower marginal costs. These differences imply that independent retailers are weak competitors, which reduces their ability to limit the market power of major chains.

The results of the policy experiments suggest that a reduction in the number of independents could be beneficial for consumers, by lowering the average market prices. This has important implications for evaluating the usefulness of protecting independent retailers through price floors or contract restrictions. If those policies reduce the incentive of inefficient firms to exit the market, they will keep the level of prices artificially high (Houde [20]).

A Calculation of the empirical distribution of commuters

The two following subsections describe the methods used to compute the demographic statistics at the residential location level for every periods, and the distribution of individuals across origin-destination pairs.

A.1 Distribution of population across location

The main issue here is to predict the distribution of population between census years, for each census dissemination area (DA). The DAs are the smallest statistical area for which detailed demographic statistics are available. For the Quebec metropolitan area, the average population of each DA is around 500. The DAs were created by Statistic Canada for the 2001 census. In order to compute demographic statistics for previous years, we will use the fact that DAs are geographically nested in the definition of Census Tracts (CT). In particular, we will use a definition of census tract which is common to all three censuses (i.e. 1991, 1996 and 2001).

Let X_{it}^{a} be a variable measured at the level of aggregation $a = \{DA, CT, CMA\}$, for zone i in period t. Two types of weights are used to predict the level of X at the DA level for every periods. First, the distribution of population across DAs for the census year 2001 is obtained directly from the census aggregate tables:

$$w_{iT}^{DA}(X) = \frac{X_{iT}^{DA}}{\sum_j X_{jT}^{DA}} \tag{19}$$

The change in this weight across periods is obtained from the observed average changes at the CT level. In particular the weight of DA i for periods $t < T$ is given by:

$$w_{it}^{DA}(X) = \frac{X_{ct(i)t}^{CT}}{X_{ct(i)T}^{CT}} w_{iT}^{DA}(X) \tag{20}$$

where $ct(i)$ is a function reporting the census tract name of DA i. Assuming that the relevant population distribution within each CT is stable over time, the weight w_{it}^{DA} is an accurate representation of the relative changes in X between year t and $T = 2001$.

In order to get monthly estimates of X, I use the monthly Canadian Labour Force survey. This survey reports estimates of the adult population and the number of workers for the main Census Metropolitan Areas on a monthly basis. Rescaling the weights defined in equation 20 so that they sum to one, the predicted value for X_{it}^{DA} is obtained by:

$$X_{it}^{DA} = \frac{w_{it}^{DA}(X)}{\sum_{j \in cma(i)} w_{jt}^{DA}(X)} X_{cma(i)t}^{CMA} \tag{21}$$

where $cma(i)$ is the CMA indicator of region i, and $X^{CMA}_{cma(i)t}$ is the value of X obtained from the LF survey in period t. The previous calculation is repeated for the three key variables used in the empirical analysis: the population older than 15 years, the number of full time students, and the number of workers (number of full-time and part-time workers who are not full-time students).

A.2 Distribution of commuting trips

In order to compute the number of commuters for each pair of origin and destination zones, I used the aggregate OD matrices from the 2001 Origin-Destination survey performed by the Québec Ministry of Transportation for the Québec city CMA. The sample of individuals surveyed in the fall of 2001 correspond to $27,839$ households or $68,121$. Each individual surveyed was asked questions related to mode of transportation used and destination for four trip purposes: work, leisure, study, and shopping. The micro data generated on each trip surveyed were then aggregated using the 2001 census weights, to generate the predicted number of trips between each pair of traffic area zones (TAZ). In the 2001 survey, each OD matrix included 67 TAZs. The definition of each TAZ represents the agglomeration of one or more census tract.

To predict the traffic between each pair of DA locations, I will use two OD matrices: the OD matrix for work trips, and the OD matrix for study trips. Let ω^t_{ij} be the proportion of trips originating from TAZ i going to TAZ j, for purpose $t \in \{\text{work,study}\}$. Since each TAZ includes multiple DAs, I have to assume a distribution of trips within each TAZ. The distribution trips originating from each zone is assumed to be homogeneous across DAs within the same TAZ. This is justified by the lack of additional information, and by the fact that census boundaries are defined such that population within each CTs is as homogeneous as possible. The distribution of destinations zones within each TAZ is, on the other hand, assumed to be proportional to the distribution of employees and schools (Colleges and Universities) respectively. The distribution of employees by DAs is available only for year 2001 from the Canadian Business Summary database compiled by PCensus, while the distribution of schools by DAs is calculated using the DMTI Enhanced Points of Interest database[25]. Combining this information with the aggregate OD probabilities, we can

[25]DMTI Spatial. "Enhanced Points of Interest", version 3.1 [Electronic resource]. Markham, Ontario: DMTI Spatial, 2004.

compute the number of commuters T_{ij} for the DA pair (i, j) using the following formula:

$$T_{ij} = \sum_{p=\{\text{work,study}\}} \omega_{taz(i),taz(j)} \frac{Y_j^p}{\sum_{j' \in taz(j)} Y_j^p} X_{it}^p, \tag{22}$$

where X_{it}^p is the relevant population measure (i.e. workers or full-time students), $taz(i)$ is a function indicating the TAZ name of DA i, and Y_j^p is the number of employees in location j if the trip purpose is work, or the number schools if the trip purpose is study. Note that the previous representation implicitly assume that the geographic distribution of trips is stationary over the sample periods.

Finally, the resultant measures of traffic are aggregated into larger location areas to reduce the computation cost of the model. In particular, I aggregated to the CT level each DA for which either the size (in square kilometer) or the corresponding CT size is smaller than the median DA or CT size.

B Description of the Shortest-Path Algorithm

The set of optimal routes between each pair of origin and destination zones is computed using a version of the Dijkstra's Shortest Path Tree algorithm (see Shekhar and Chawla [36] for an enlightening introduction to this class of algorithm). The road network is represented by a directed graph $G = (N, A)$. Where N is the set of nodes (or intersections), and A is the set of arcs (or street segments). Each segment a is a pair of connected nodes (i, j), ordered according to the direction of the arc. The time cost of traveling along each arc is given by $C = \{c_{ij} | (i, j) \in A\}$. The shortest path algorithm constructs, for every origin nodes s, a shortest path tree (SPT) $\mathcal{P}_s$ which stores the shortest path from s to every other nodes in the network. The procedure is an iterative algorithm which iterates on the cost $t(r, v)$ of traveling from s to any node v until convergences.

At any point during the iteration process, the algorithm keeps track of a list of nodes left to be examined ($frontierSet$), a list of nodes already explored ($exploredSet$), and a function $p_s(v)$ which indicates the parent node in the shortest path from s to v. At each iteration the algorithm removes the lowest cost node from the frontier set, and visit every nodes that are adjacent to this node (i.e. $adjSet(u)$). If the cost of visiting one of these nodes $w \in adjSet(u)$ is lower than the current estimate, the algorithm updates the cost function $t(s, w)$ and the path $p_s(w)$. The valid nodes are then added to the frontier set. The algorithm stops when all nodes in the network have been visited. The pseudo-code below describes the main steps of the SPT calculation.

Algorithm 1 *Shortest path three rooted at node s, on network $G(N, A)$:*

Initialization step:

$$t(s,v) = \begin{cases} \infty & \text{if } v \neq s \\ 0 & \text{otherwise} \end{cases} \qquad frontierSet^0 = \{s\} \qquad exploredSet^0 = \{\emptyset\}$$

Iteration k:

$u = \arg\min_{w \in frontierSet^k} \quad t(s,w)$

$frontierSet^k = frontierSet^{k-1} \backslash \{u\}$

$exploredSet^k = exploredSet^{k-1} \cup \{u\}$

`foreach` $w \in adjList(u)$

`if` $t(s,w) > t(s,u) + c(u,w)$ `then`

{

$t(s,w) = t(s,u) + c(u,w)$

$p_s(w) = u$

`if` $w \ni frontierSet^k \cup exploredSet^k$ `then`

$frontierSet^k = frontierSet^k \cup \{w\}$

}

`if` $frontierSet^k = \emptyset$ `then`

`stop`

`else` $k = k + 1$

The set of routes corresponding to the shortest path tree $\mathcal{P}_s$ are constructed recursively using the function $p_s(v)$. For instance the path $r(s,d) \in \mathcal{P}_s$ is an array of $n_r + 1$ nodes such that the last element is $r_{n_r} = d$, the second-last element is $r_{n_r-1} = p_s(d)$, the k^{th}-last element is $r_{n_r-k+1} = p_s(r_{n_r-k})$, and the first element is $r_0 = s$.

References

[1] Daniel Ackerberg, Lanier Benkard, Steven Berry, and Ariel Pakes, *Econometric tools for analysing market outcomes*, mimeo, Harvard University, May 2005.

[2] Simon Anderson, André de Palma, and Jacques-François Thisse, *Discrete choice theory of product differentiation*, MIT Press, 1992.

[3] Michael Berkowitz, Nancy T. Gallini, Eric J. Miller, and Robert A. Wolfe, *Disaggregate analysis of the demand for gasoline*, The Canadian Journal of Economics **23** (1990), no. 2, 253–275.

[4] Steven Berry, *Estimating discrete choice models of product differentiation*, Rand Journal of Economics **25** (1994), 242–262.

[5] Steven Berry, James Levinsohn, and Ariel Pakes, *Automobile prices in market equilibrium*, Econometrica: Journal of the Econometric Society **63** (1995), no. 4, 841–890.

[6] ______, *Voluntary export restraints on automobiles: Evaluating a trade policy*, American Economic Review **89** (1999), no. 3, 400–430.

[7] Severin Borenstein and Andrea Shepard, *Dynamic pricing in retail gasoline markets*, The Rand Journal of Economics **27** (1996), no. 3, 429–451.

[8] Irene Brambilla, *A customs union with multinational firms: The automobile market in argentina and brazil*, Yale University, October 2004.

[9] Timothy F. Bresnahan, *Competition and collusion in the american automobile industry: The 1955 price war*, The Journal of Industrial Economics **35** (1987), no. 4, The Empirical Renaissance in Industrial Economics, 457–482.

[10] Russell Davidson and James MacKinnon, *Econometric theory and methods*, Oxford University Press, New York, 2004.

[11] Peter Davis, *Spatial competition in retail markets: Movie theaters*, Working Paper, London School of Economics, 2005.

[12] Ministère des Transports du Québec, *Ministère des transports du québec et réseau de transport de la capitale, la mobilité des personnes dans l'agglomération de québec. somaire des résultats de l'enquète origine-destination 2001.*, Tech. report, Gouvernement du Québec, 2002.

[13] DMTI Spatial, Markham, Ontario, *Canmap routelogistics, version 8.2 [electronic resource]*, 2004.

[14] Jurgen A. Doornik, *Ox - an object-oriented matrix programming language*, Timberlake Consultants, 2001.

[15] Jean-Pierre Dube, *Product differentiation and mergers in the carbonated soft drink industry*, Journal of Economics and Management Strategy (Forthcoming).

[16] Jeffrey A. Dubin and Daniel L. McFadden, *An econometric analysis of residential electric appliance holdings and consumption*, Econometrica **52** (1984), no. 2, 345–362.

[17] Andrew Eckert and Douglas S. West, *Rationalization of the retail gasoline station networks in canada*, Review of Industrial Organization **26** (2005), 1–25.

[18] Jusine Hastings, *Vertical relationships and competition in retail gasoline markets: Empirical evidence from contract changes in southern california*, American Economic Review (2004).

[19] Justine Hastings and Richard Gilbert, *Market power, vertical integration, and the wholesale price of gasoline*, University of California Energy Institute Power Working Paper, No. PWP-084, June 2002.

[20] Jean-François Houde, *Entry and exit in a price regulated industry: Gasoline retailling in québec*, April 2005.

[21] Guido Imbens and Tony Lancaster, *Combining micro and macro data in microeconometric models*, The Review of Economic Studies **61** (1994), no. 4, 655–680.

[22] Saul Lach, *Existence and persistence of price dispersion: An empirical analysis*, The Review of Economics and Statistics (2002).

[23] Ricardo Lagos, *An alternative approach to search frictions*, Joural of Political Economy **108** (2000), 5.

[24] J.A. Manez and M. Waterson, *Multiproduct firms and product differentiation: a survey*, working paper, University of Warwick.

[25] Mark D. Manuszak, *The impact of upstream mergers on retail gasoline markets*, Working Paper, Carnegie Mellon University, 2001.

[26] Aviv Nevo, *Mergers with differentiated products: The case of the ready-to-eat ceral industry*, The Rand Journal of Economics **31** (2000), no. 3, 395–421.

[27] ———, *Measuring market power in the ready-to-eat cereal industry*, Econometrica **69** (2001), no. 2, 307.

[28] Michael Noel, *Edgeworth price cycles: Evidence from the toronto retail gasoline market*, University of California at San Diego.

[29] Norbert Oppenheim, *Urban travel demand modelling*, Wiley-Interscience, 1995.

[30] Amil Petrin, *Quantifying the benefits of new products: The case of the minivan*, Journal of Political Economy **110** (2002), 705.

[31] Petro-Canada, *Government regulation and intervention in the downstream oil industry*, Tech. report, Petro-Canada, September 1996, Submitted to the Government of the Province of Québec.

[32] Joris Pinkse and Margaret Slade, *Mergers, brand competition and the price of a pint*, European Economic Review **48** (2004), 617–643.

[33] Joris Pinkse, Margaret E. Slade, and Craig Brett, *Spatial price competition: A semiparametric approach*, Econometrica **70** (2002), no. 3, 1111–1153.

[34] Tourisme Québec, *Le tourisme au québec en 1995*, Tech. report, Tourisme Québec, 1997.

[35] ———, *Le tourisme au québec en 2001*, Tech. report, Tourisme Québec, 2002.

[36] Shashi Shekhar and Sanjay Chawla, *Spatial databases: A tour*, 2003.

[37] Margaret E. Slade, *Interfirm rivalry in a repeated game: An empirical test of tacit collusion*, The Journal of Industrial Economics **35** (1987), no. 4, The Empirical Renaissance in Industrial Economics, 499–516.

[38] ———, *Vancouver's gasoline-price wars: An empirical exercise in uncovering supergame strategies*, The Review of Economic Studies **59** (1992), no. 2, 257–276.

[39] Howard Smith, *Supermarket choice and supermarket competition in market equilibrium*, The Review of Economic Studies **71** (2004), no. 1, 235–263.

[40] Vandersmissen Marie-Helene Lee-Gosselin Martin Theriault, Marius and Denis Leroux, *Modelling commuter trip length and duration within gis: Application to an o-d survey*, Journal of Geography Information and Decision Analysis **3** (1999), no. 1, 40–56.

[41] Raphael Thomadsen, *The effect of ownership structure on prices in geographically differentiated industries*, Rand Journal of Economics (Forthcoming).

[42] Adonis Yatchew and Angela No, *Household gasoline demand in canada*, Econometrica **69** (2001), 1697–1708.