# Spatial differentiation and vertical mergers in retail markets for gasoline

## Web Appendix

Jean-François Houde*

July 8, 2011

## 1 Calculation of the empirical distribution of consumers

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.

### 1.1 Distribution of population across location

The distribution of consumers in period $t$ across origin-destination locations is given by $T^t_{sd}$. I decompose this number in three components: the number of workers ($W^t_s$), full-time students ($S^t_s$), unemployed ($U^t_s$), and the number of outside commuters ($O^t_{sd}$). For workers and students, the probability of commuting between $(s, d)$ is denoted by $\Omega^{W_t}_{sd}$ and $\Omega^{S_t}_{sd}$ respectively. For unemployed individuals this pair is unique (i.e. $s = d$). For outside commuters the pairs of origin/destination correspond to beginning and end points of each highway segment. I denote this set of points by $\mathcal{H}$. The number of commuters between $(s, d)$ is given by:

$$T^t_{sd} = W^t_s \Omega^{W_t}_{sd} + S^t_s \Omega^{S_t}_{sd} + U^t_s \mathcal{I}(s = d) + O_t \mathcal{I}((s, d) \in \mathcal{H}), \quad (1)$$

I approximate the number of outside commuters $O_t$ by the average number of occupied hotel room in the city in period $t$. This due to the fact the transportation survey is related only to local commuters, and I do not have access to any traffic count data on the highway network.

*Department of Economics, University of Wisconsin-Madison, houdejf@ssc.wisc.edu.

The main issue 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:

$$w_{iT}^{DA}(X) = \frac{X_{iT}^{DA}}{\sum_j X_{jT}^{DA}} \tag{2}$$

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:

$$w_{it}^{DA}(X) = \frac{X_{ct(i)t}^{CT}}{X_{ct(i)T}^{CT}} w_{iT}^{DA}(X) \tag{3}$$

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 3 so that they sum to one, the predicted value for $X_{it}^{DA}$ is obtained by:

$$X_{it}^{DA} = \frac{w_{it}^{DA}(X)}{\sum_{j \in cma(i)} w_{jt}^{DA}(X)} X_{cma(i)t}^{CMA} \tag{4}$$

where $cma(i)$ is the CMA indicator of region $i$, and $X_{cma(i)t}^{CMA}$ 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).

## 1.2 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 $27,839$. 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 $\omega_{ij}^t$ 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[1]. Combining this information with the aggregate OD probabilities, we can compute the number of commuters $T_{ij}$ for the DA pair $(i, j)$ using the following formula:

$$T_{ij} = \sum_{p=\{\text{work,study}\}} \omega_{taz(i),taz(j)} \frac{Y_j^p}{\sum_{j' \in taz(j)} Y_j^p} X_{it}^p, \tag{5}$$

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.

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

## 2 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 Hall (2003) 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 $r_0 = s$.

# 3 Computation of the GMM estimator

In this section I describe the details involved in the estimation and statistical inference of the parameter vector $\theta$, and various functions of those parameters (e.g. elasticities, willingness to travel, markups, etc). In particular, I discuss three elements of the procedure which are specific to my problem: (i) estimation of income distribution, (ii) simulation of market shares, (ii) the inversion procedure, (iii) the construction of the weighting matrix.

## 3.1 Income distribution

I obtained data on the distribution of income from the most recent census (2001). For each 1,200 Dissemination Areas (i.e. residential location $L$), the aggregate census tables give the average individual annual income, along with the median and standard-deviation. I use these three moments to estimate a parametric distribution of income that is location specific. In particular I assume that individual income for location $l$ is distributed according to a Weibull distribution with parameters $(\lambda_l, \mu_l)$. The Weibull distribution is attractive in this context because it has positive support and the mean, median and variance all have closed-form expressions. For each location $l$ the parameters of the income distribution are estimated by minimizing the square difference between the three predicted moments and the observed ones from the census:

$$(\hat{\lambda}_l, \hat{\mu}_l) = \arg\min_{\lambda,\mu} \begin{pmatrix} \bar{y}_l - \lambda\Gamma\left(1+\frac{1}{\mu}\right) \\ y_{0.5,l} - \lambda\left(\ln(2)\right)^{1/\mu} \\ \sigma^2_{y_l} - \lambda^2\Gamma\left(1+\frac{2}{\mu}\right) - \left(\lambda\Gamma\left(1+\frac{1}{\mu}\right)\right)^2 \end{pmatrix}^2 \tag{6}$$

where $\Gamma(\cdot)$ is the Gamma function. Using this estimated income distribution it is then straightforward to sample income to construct the predicted demand of each store. I also assume that the income of outside commuters is distributed according to the average distribution (i.e. average mean, median and standard-deviation).

### 3.2 Simulation of market shares

Demand at each store is obtained by aggregating three sources of heterogeneity: (i) origin location ($s$), (ii) destination location ($d$), and (iii) income ($y$). Given the large number of possible combination of origin/destination and income values, I perform this aggregation through simulation.

I use the following sequential procedure to sample $S$ types from the commuting and income distribution. This procedure is performed separately for each year using the same random numbers.

1. Sample origin location $s_i$ from the unconditional $DA$ population (including outside commuters).

2. Sample income $y_i$ from conditional income distribution $F(y|s_i)$.

3. Sample destination location $d_i$ from commuting probabilities $\Omega^t_{s_i,d}$.

In the empirical application I use 3000 simulated draws.

Finally the simulated demand at each store need to be rescaled to match the observed sales' volume. In particular let $\hat{M}_t = \sum_i \bar{q}(s_i, d_i)$ be the simulated market size (i.e. if all simulated types buy gasoline). The simulated market shares are then give by:

$$s_{jt}(\delta_t|\theta) = \frac{1}{S}\sum_i \bar{q}(s_i, d_i) P_{jt}(s_i, d_i, y_i|\delta_t, \theta). \tag{7}$$

The observed market shares on the other hand are constructed by dividing the observed sales volume by the theoretical market size $M_t = \sum_{s,d} T^t_{s,d}$.

### 3.3 Inversion algorithm

In order to evaluate the objective function at a given parameter vector $\theta$, it is necessary to invert the following system of non-linear equations:

$$\delta(\theta)_{jt} \to \ln s_{jt}(\delta_t|\theta) = \ln \hat{s}_{jt} \tag{8}$$

where $s_{jt}(\delta_t|\theta)$ is the model predicted market share at store $j$ in period $t$, and $\hat{s}_{jt}$ is the observed share. The complexity of the inversion procedure depends in general on the number of consumer types (e.g. the number of simulated consumers or the number of pairs $(s, d)$) and on the number of products available.

In addition, I use a Broyden's root-finding algorithm to evaluate equation 8 (see Miranda and Fackler Miranda & Fackler (2002) for more details). This procedure is proven to converge significantly faster than the standard contraction mapping algorithm proposed by Berry et al. Berry, et al. (1995). Letting $f(\delta^k) = \ln s_{jt}(\delta_t|\theta) - \ln \hat{s}_{jt}$ and $I_{J_t}$ denotes the identity matrix of dimension $J_t$, the algorithm takes the following steps to find $\{\delta(\theta)_{jt}\}_{j=1,\ldots,J_t}$:

1. Set the starting value for the pseudo-jacobian matrix $B^0 = I_{J_t}$ and $\delta^0_{jt}$.
2. For iteration $k \geq 1$:
   (a) Update the vector of mean qualities:
   $$\delta^k_{jt} = \delta^{k-1}_{jt} - B^{k-1} f(\delta^{k-1}_t)$$
   (b) Update the pseudo-jacobian matrix:
   $$B^k = \begin{cases} B^{k-1} + (s-u)s'B^{k-1} * (s'u)^{-1} & \text{if } ||f(\delta^k_t)|| > ||f(\delta^{k-1}_t)|| \\ I_{J_t} & \text{Otherwise.} \end{cases}$$
   where $s = -B^{k-1} f(\delta^{k-1}_t)$ and $u = B^{k-1}\left[f(\delta^k_t) - f(\delta^{k-1}_t)\right]$.
3. Stop if $||f(\delta^k_t)|| \leq \epsilon$, repeat step 2 otherwise.

Note that contrary to standard quasi-newton algorithms, this procedure is guaranteed to convergence as $k \to \infty$. To see this, note that by fixing $B = I_{J_t}$ the procedure is equivalent to the contraction mapping algorithm of Berry et al. Berry et al. (1995). Step 2b uses this property of the problem by reverting to the contraction mapping algorithm each time the algorithm tends to diverge. More importantly, the algorithm typically converges in less than 10 iterations, compare to more than 100 for the contraction mapping algorithm. In fact, the algorithm typically requires almost as few iterations to converge as the Newton algorithm without requiring any matrix inversion, nor evaluating the Jacobian matrix.

### 3.4 Construction of the weighting matrix

The weighting matrix is block-diagonal as in Petrin Petrin (2002), since the two moments are calculated from different samples.[2] In the second stage of the GMM optimization routine, the weighting matrix of the second set of moment conditions is computed using a heteroskedastic-consistent variance-covariance matrix of the micro-moment conditions.

For the first set of moments (i.e. IVs), I compute a weighting matrix which is consistent with both with spatial and time correlations in the empirical moments. In particular following Conley Conley (1999), the variance-covariance matrix of the first set of empirical moments $V_n^1$ is estimated by weighing observations according to their distance in space and time:

$$\hat{V}_n^1 = \frac{1}{n} \sum_t \sum_{l=-3}^{3} K(t, t+l) g_t^1(\hat{\theta}^1)^T D(t, t+l) g_{t+l}^1(\hat{\theta}^1) \tag{9}$$

where $g_t^1$ is the period $t$ matrix $J_t \times L$ of empirical moments evaluated at the first-stage parameters $\hat{\theta}^1$, $K(t, t+l)$ a Bartlett kernel weighting observations at time $t$ and $t+l$, and $D(t, t+l)$ is a $J_t \times J_{t+l}$ matrix of spatial Kernel weights. I use a truncated, spatial Bartlett Kernel that assigns decreasing weights on observations located within 3 KM of each other:

$$D_{j,k}(t, t+l) = \begin{cases} 1 - \frac{d(j,k)}{3} & \text{If } d(j,k) \leq 3 \\ 0 & \text{Otherwise.} \end{cases} \tag{10}$$

The bandwidth of the truncated, temporal Bartlett Kernel is chosen similarly such that observations within 3 bi-monthly periods of each-other receive positive weights. This allows for correlation over observations from a one year interval around a given observation. The exact form is given by:

$$K(t, t+l) = \begin{cases} 1 - \frac{|l|}{3} & \text{If } |l| \leq 3 \\ 0 & \text{Otherwise.} \end{cases} \tag{11}$$

## 4 Description of the players and vertical arrangements

Gasoline retail markets are unique in that firms with very different vertical arrangements are competing for the same consumers. In 2001, the Québec City market was composed of 20 retail chains operating 283 retail stores, five of which were partially vertically integrated.

The ownership structure and the exact nature of the vertical contracts are not publicly available. I obtained a measure of the ownership of stations by matching gasoline station data with the list of underground storage tanks. This list is publicly available through the Québec Department of Energy and Environment. A site is defined as being vertically

[2] See Imbens and Lancaster Imbens & Lancaster (1994) for further details on the estimation of micro-econometric models with macro moment conditions.

Table 1: Distribution of key players presences and market shares in 1991 and 2001

| | Fall 1991 | | | Fall 2001 | | |
|---|---|---|---|---|---|---|
| Brand | Company owned (%) | Market Share | Nb. Stores | Company owned (%) | Market Share | Nb. Stores |
| Ultramar | 0.549 | 0.134 | 51 | 0.621 | 0.204 | 58 |
| Esso | 0.294 | 0.195 | 68 | 0.568 | 0.160 | 37 |
| Irving | 0.606 | 0.129 | 33 | 0.525 | 0.157 | 40 |
| Shell | 0.269 | 0.138 | 52 | 0.615 | 0.129 | 26 |
| Petro-Canada | 0.353 | 0.115 | 51 | 0.72 | 0.113 | 25 |
| Sunoco | 0 | 0.036 | 15 | | | |
| EKO | 0 | 0.026 | 18 | 0 | 0.072 | 31 |
| Olco | 0 | 0.049 | 18 | 0 | 0.043 | 20 |
| Couche-Tard | | | | 0 | 0.032 | 14 |
| Canadian Tire | 0 | 0.025 | 5 | 0 | 0.053 | 6 |
| Other | 0 | 0.154 | 71 | 0 | 0.038 | 26 |

The five vertically intergrated brand are: Ultramar, Esso, Irving, Shell, Petro-Canada. Sunoco stations in Quebec were merged with Ultramar in 1997. Couche-Tard is a convenience-store chains that started operating stations under its own name in 1996. Canadian-Tire is a hardware store chain.

integrated or company-owned if the owner of the underground tank is one of the five vertically integrated companies. Table 1 summarizes this information for 1991 and 2001. Among stores selling vertically integrated brands (i.e. the majors), 60% were company-owned stores in 2001. The rest are franchisees or lessee stations that own one or more stores, and are in principle responsible of setting their own prices.[3]

## 4.1 Branded stations

Branded stations are selling gasoline under the brand-name of one of the five vertically integrated refineries: Ultramar, Shell, Petro-Canada, Esso, and Irving.[4] Those stations are either owned and operated directly by the company (i.e. company-owned), or owned by the station operator (i.e. lessee-station). The prices and investments at company-owned stores are decided by the upstream company.[5]

[3] Most of the information on the content of vertical arrangements comes from the transcripts of a governmental investigation studying the causes the 2000 price war in Quebec City (i.e. case R-3457-2000). One of the mandate of the regulatory agency is to investigate "major" violations of the price-floor. The investigation lasted four months, and focussed in part on the role played by vertical contracts in causing the price war.

[4] Ultramar also operates two formerly independent chains: Sergaz and Sunoco.

[5] Slade (1998) identifies a third category of branded stations: independently owned stores purchasing gasoline on the spot market (i.e. at their supplier's rack price). While I cannot discard with certainty this type, industry and government reports suggest that the vast majority of independent branded stations operate under lessee contracts in Québec City.

Lessee-station owners sign long-term agreements that are negotiated bilaterally with one of the refiner, and last on average five years. The owner of a lessee-station is responsible of all investments and maintenance costs. In exchange, the upstream firm pays a rent to the owner in (linear) proportion of the amount of gasoline sold in a given period. The value of this rent, expressed in cents per liter, is constant over time and reflects the characteristics of the location and amenities.

There exists two types of lessee contracts: (i) commission, and (ii) traditional. Under a commission contract, the upstream firm keeps the ownership of the gasoline (i.e. the product is on consignment), and is responsible of setting the retail price. The station owner is compensated through fixed commission proportional to the gasoline sold. Under the traditional contract, the lessee-station owner is responsible of setting the retail price, and pays a wholesale price determined at each delivery (weekly or bi-weekly). Importantly, the wholesale price differs across stations, even within the same chain.

Documents obtained from a government inquiry conducted in 2001 reveal that Ultramar uses only commission contracts, while the other brands use traditional contracts. Since 1996, Ultramar has a low-price guarantee marketing policy, which state that it will post the lowest price within a neighborhood its stores (company-owned or lessee). This policy requires full control over the retail price.[6]

Finally, while resale-price maintenance is illegal in Canada, it appears that upstream firms have substantial control over retail prices at traditional lessee stations. Indeed, most companies (including the largest independent chains) offer a price support clause to their lessee retailers that insures the owner of a positive profit margin even if the actual margin falls below a certain level negotiated at beginning of the contract life. This clause generated a great deal of controversy, and is suspected to have caused a major price war in 2000 when a lessee retailer posted a price equal to the floor in order to benefit from a larger volume and fixed margin.[7]

Since retailers and suppliers engage in long-term repeated relationships, this clause can help to enforce a "collusive" equilibrium in which a retailer agrees to eliminate its retail margin, and in exchange is compensated for the risk of earning low margins. In other words, under this implicit contract the upstream supplier is able to enforce a resale-price maintenance equilibrium, without using a more complex non-linear contract. Non-linear contracts are particularly unattractive in gasoline markets since oil prices are very volatile.

[6] According to Slade (1998) commission-based contracts are not present in the U.S. market. Resale-price maintenance laws prevent a refiner to set the retail price at a station not operated by an employee. As a result, commission contracts have disappeared from the U.S. market (they were the norm in the 1970s).

[7] Several witnesses at the government investigation claimed that this retailer negotiated a minimum margin well above 3 cpl several years earlier; during a period in which the average margins were systematically above 5 cpl. According to the key witnesses, most of these contracts were phased out by 2001, and companies after 2000 started negotiating minimum margins around 3 cents.

## 4.2 Independent brands

Independent stations sell gasoline that is either purchased directly from one of the vertically integrated firm, or imported through pipelines from another province or from the USA. Importantly, gasoline sold at independent stations is unbranded, and consumers are not aware of its origin.

Transactions are typically performed at a spot or rack price, net of a negotiated discount and delivery cost. The discount is based on volume. The Québec City wholesale terminal includes five branded resellers (i.e. Ultramar, Shell, Petro-Canada, Esso and Sunoco), and one independent firm selling imported gasoline. Each firm posts an unbranded-rack price that can change daily, and prices typically exhibit little dispersion. The Sunoco terminal was replaced by the unbranded distributer Olco in 1997, after Ultramar acquired Sunoco's network of stations in the Province.

There exist three groups of independent gasoline stations. The first group includes brands that are partially integrated in the distribution of petroleum products. In Québec City, Olco, Petro-T, Sonic, and Eko together operated 67 stores in Québec City in 2001. According to the companies' website, roughly half of these stores are company-owned, and the rest are independently owned and operated under lessee contracts.

The second group of independent stations are affiliated with a retail chain selling hardware or grocery products. Canadian Tire and Couche Tard are the two largest in Québec. Both own and operate their gasoline stations directly.

The rest are small independent stores who operate only a small number of sites, and purchase gasoline on the spot market. Their presence in the market shrunk considerably since 1991.

# 5 Additional tables

Table 2: Estimates of log-linear demand model under alternative IV specifications

| VARIABLES | (1) FE-OLS | (2) FE-IV | (3) FE-IV | (4) FE-IV | (5) FE-IV |
|---|---|---|---|---|---|
| Price (log) | -1.160$^a$ | -12.30$^a$ | -4.284$^a$ | -5.508$^a$ | -5.237$^a$ |
| | (0.0918) | (2.643) | (1.043) | (0.987) | (0.883) |
| Observations | 14,240 | 14,240 | 14,240 | 14,240 | 14,240 |
| Number of location | 406 | 406 | 406 | 406 | 406 |
| Weak IV (F-stat) | . | 5.582 | 20.74 | 11.08 | 4.910 |
| Over-identification (J-stat) | 0 | 20.51 | 13.62 | 58.66 | 143.4 |
| Degrees of freedom | 0 | 4 | 3 | 8 | 28 |
| P-value | . | 0.000396 | 0.00347 | 8.52e-10 | 0 |

Robust standard errors in parenthesis. a p<0.01, b p<0.05, c p<0.1. All specifications include store and period fixed-effects, time-varying characteristics, and local traffic and population density measures. The weak IV and over-identification tests are the Cragg-Donald and Hansen statistics, calculated using the *ivreg2* command in STATA.

Table 3: Linear parameter estimate from the multi-address model specification

| VARIABLES | Spec. 1 | Spec. 2 | Spec. 3 | Spec. 4 |
|---|---|---|---|---|
| Nb. Pumps | 0.0154 | 0.01271 | 0.01436 | 0.01196 |
| | ( 0.00207) | ( 0.00178) | ( 0.00193) | ( 0.00162) |
| Nb. Islands | 0.01279 | 0.02157 | 0.0179 | 0.03098 |
| | ( 0.0201) | ( 0.0169) | ( 0.0187) | ( 0.0156) |
| Self service | 0.0779 | 0.04931 | 0.0691 | 0.04713 |
| | ( 0.0482) | ( 0.0387) | ( 0.0422) | ( 0.0344) |
| Mixed service | 0.1758 | 0.1424 | 0.1663 | 0.1467 |
| | ( 0.0489) | ( 0.0398) | ( 0.0435) | ( 0.0354) |
| Small conv. store | -0.01122 | -0.007819 | -0.005654 | 0.007675 |
| | ( 0.0302) | ( 0.024) | ( 0.0272) | ( 0.0205) |
| Medium conv. store | -0.008935 | -8.928e-05 | -0.004475 | 0.008824 |
| | ( 0.0431) | ( 0.0365) | ( 0.04) | ( 0.032) |
| Large conv. store | -0.07832 | -0.05516 | -0.06883 | -0.0248 |
| | ( 0.0577) | ( 0.0457) | ( 0.051) | ( 0.0413) |
| Repair shop | 0.05368 | 0.03946 | 0.03533 | 0.0315 |
| | ( 0.0535) | ( 0.0426) | ( 0.0457) | ( 0.0389) |
| Carwash | 0.01011 | 0.0112 | 0.01678 | 0.001914 |
| | ( 0.0521) | ( 0.0339) | ( 0.0391) | ( 0.0303) |
| 12 Hours | -0.3629 | -0.3277 | -0.339 | -0.3406 |
| | ( 0.0998) | ( 0.0818) | ( 0.0911) | ( 0.0761) |
| Extended hours | -0.05168 | -0.07256 | -0.06466 | -0.08701 |
| | ( 0.0313) | ( 0.0258) | ( 0.0288) | ( 0.0219) |
| Observations | 14263 | 14263 | 14263 | 14263 |
| Nb. of stores | 429 | 429 | 429 | 429 |

Robust standard-errors in parenthesis. Each specification also includes location, time and brand fixed-effects. The moment conditions used in each specification are described in Table **??**.

Table 4: Linear parameter estimate from the single-address model specification

| VARIABLES | Spec. 1 | Spec. 2 | Spec. 3 | Spec. 4 |
|---|---|---|---|---|
| Nb. Pumps | 0.01086 | 0.01111 | 0.01029 | 0.01093 |
| | ( 0.00169) | ( 0.00156) | ( 0.0016) | ( 0.00145) |
| Nb. Islands | 0.02118 | 0.02282 | 0.02288 | 0.02077 |
| | ( 0.0167) | ( 0.0151) | ( 0.0158) | ( 0.014) |
| Self service | 0.01357 | 0.0379 | 0.02369 | 0.03871 |
| | ( 0.038) | ( 0.0347) | ( 0.0353) | ( 0.0314) |
| Mixed service | 0.1428 | 0.1606 | 0.149 | 0.1631 |
| | ( 0.0393) | ( 0.0357) | ( 0.0365) | ( 0.0329) |
| Small conv. store | 0.01947 | 0.02065 | 0.01513 | 0.02213 |
| | ( 0.0209) | ( 0.0189) | ( 0.0198) | ( 0.0176) |
| Medium conv. store | 0.02717 | 0.03191 | 0.02245 | 0.02085 |
| | ( 0.0331) | ( 0.0315) | ( 0.0323) | ( 0.0288) |
| Large conv. store | 0.04267 | 0.01843 | 0.03368 | 0.02216 |
| | ( 0.0477) | ( 0.0428) | ( 0.0446) | ( 0.039) |
| Repair shop | 0.02748 | 0.05112 | 0.008157 | 0.02785 |
| | ( 0.0472) | ( 0.043) | ( 0.0434) | ( 0.0383) |
| Carwash | -0.0121 | 0.01846 | 0.001718 | -0.002591 |
| | ( 0.0388) | ( 0.0309) | ( 0.0322) | ( 0.0281) |
| 12 Hours | -0.4143 | -0.4089 | -0.4339 | -0.4245 |
| | ( 0.081) | ( 0.0786) | ( 0.08) | ( 0.0753) |
| Extended hours | -0.1052 | -0.09638 | -0.1056 | -0.1073 |
| | ( 0.0243) | ( 0.0225) | ( 0.0232) | ( 0.0199) |
| Observations | 14263 | 14263 | 14263 | 14263 |
| Nb. of stores | 429 | 429 | 429 | 429 |

Robust standard-errors in parenthesis. Each specification also includes location, time and brand fixed-effects. The moment conditions used in each specification are described in Table ??.

## References


S. Berry, et al. (1995). 'Automobile Prices in Market Equilibrium'. *Econometrica: Journal of the Econometric Society* **63**(4):841–890.

T. G. Conley (1999). 'GMM estimation with cross sectional dependence'. *Journal of Econometrics* **92**(1):1–45.

G. Imbens & T. Lancaster (1994). 'Combining Micro and Macro Data in Microeconometric Models'. *The Review of Economic Studies* **61**(4):655–680.

M. J. Miranda & P. L. Fackler (2002). *Applied computational economics and finance*. MIT Press.

A. Petrin (2002). 'Quantifying the Benefits of New Products: The Case of the Minivan'. *Journal of Political Economy* **110**:705–729.

S. Shekhar & S. Chawla (2003). *Spatial Databases: A Tour*.

M. E. Slade (1998). 'Strategic Motives for Vertical Separation: Evidence from retail gasoline markets'. *Journal of Law, Economics, & Organization* **14**(1):84–113.