A new test for monopoly with limited cost data

Charles C. Moul
Miami University

August 2011

Working Paper # - 2011-03
A new test for monopoly with limited cost data

Charles C. Moul
Miami University
August 2011

ABSTRACT: I introduce a monopoly test that requires only limited, if any, cost information. The test’s intuition is that demand estimates in isolation will be similar to demand estimates under joint estimation with cost if the monopoly null hypothesis is correct. These estimates, though, may diverge if the monopoly null hypothesis is false. The test’s power increases with the extent of inside-outside segmentation and the market’s true number of firms. Simulations indicate that the test has substantial power using duopoly data when plausible levels of inside-outside segmentation are present.

Keywords: market structure test, monopoly

I. Introduction

The detailed data required for many economic applications often necessitate that variables be collected at the firm rather than market level, but this implies that the analyst may be unaware of critical aspects of the firm’s environment. A similar issue can arise in cases of uncertain market definition. The number, identity, and price-quantity outcomes of competitors are all likely to be incomplete or missing altogether. Even in the absence of market-level information, though, a firm’s actions may reveal key attributes of the competitive environment. I consider what may be learned by comparing demand estimates derived alone and those jointly estimated with cost under a maintained market structure hypothesis.

My market structure test’s intuition is somewhat similar to that of a Hausman specification test (1978). If market structure is correctly stated, demand estimates in isolation will be consistent but inefficient while demand estimated jointly with cost will be consistent and efficient. The difference between these estimates will be statistically negligible. If, however, market structure is misstated, both estimates will be inconsistent, though in likely different ways. The difference between these estimates can thus provide a novel test for the hypothesized market structure itself.

While nominally applicable to testing any market structure or conduct hypothesis, the statistical power of the test comes primarily with respect to rejecting the null hypothesis of monopoly. This particular null is somewhat unorthodox but not unprecedented. Panzar and Rosse (1987) employ firm-level revenue functions and their comparative statics to construct a statistic.
that may reject the monopoly hypothesis as well as other market structures. That test’s applicability is somewhat limited as it requires the observation of all factor prices. While my proposed test is assisted by cost information, all relevant factor prices need not be observed, and the test can even be executed with no cost information when the firm sells a portfolio of goods. This contrasts with Bresnahan (1982) which identifies conduct in linear demand and cost from cost shifters and demand rotators.

As Monte Carlo simulations illustrate, my proposed test is fairly good at correctly rejecting monopoly when the market exhibits some degree of competition and when consumers are unlikely to substitute from a particular brand to making no purchase (i.e., substantial inside-outside segmentation). By contrast, it shows little power when the data are simulated from a collusive oligopoly. A rejection of the monopoly null therefore implies not only that other firms are in the market but that those firms place competitive pressure on the observed firm.

II. Model and Identification

I utilize the discrete-choice framework formalized by McFadden (1978) and popularized by Berry (1994). For expositional convenience, I will assume that firms are retailers selling the same set of products, though this can readily be adapted to other contexts. Consumers maximize utility by choosing among B brands sold by N symmetric retailers (inside options); they may also choose to make no purchase (outside option). Denoting the choice of a specific brand at a specific retailer as choice j and the outside option as choice 0, consumer i’s utility is specified as u_{ij} = X_j \beta - \alpha p_j + \xi_j + \epsilon_{ij}. Here \(X\) and \(\xi\) denote observed and (mean valuation of) unobserved product characteristics, \(p\) denotes prices, \((\beta, \alpha)\) are parameters to be estimated, and \(\epsilon\) are consumer idiosyncratic preferences. The extent of inside-outside segmentation depends upon \(\epsilon\)’s distribution, which I characterize as a special case of McFadden’s Generalized Extreme Value (GEV) model. This reduces to the nested logit using retailer-brand market shares conditional on purchase. While the details can be found in the appendix, the parameter \(\sigma\) captures the extent of inside-outside segmentation, with \(\sigma = 0\) implying no segmentation and \(\sigma \rightarrow 1\) implying near-total segmentation. This nested logit demand specification can be transformed to yield (for firm 1 selling product j)

\[
\ln(s_{1j}) - \ln(1 - N \sum_{k=1}^{B} s_{1k}) = \sigma \ln \left( \frac{s_{1j}}{N} \right) + \delta_{1j} = \sigma \ln \left( \frac{s_{1j}}{N} \right) + X_j \beta - \alpha p_{1j} + \xi_j
\]

where \(s\) denotes unconditional purchase probabilities, \(\tilde{s}\) denotes observed purchase probabilities conditional on making some purchase at retailer 1 (so \(\tilde{s}_{1j} = \frac{s_{1j}}{\sum_{k} s_{1k}}\)), and \(\delta\) denotes the mean utility of the choice. The mean utility of no purchase \(\delta_0\) is normalized to zero.

A non-cooperative firm sets its products’ prices to maximize its own profits, fully exploiting all portfolio effects:

\[
\max_p \pi = \sum_{k=1}^{B} (p_{1k} - c_k) M s_{1k}
\]

where \(M\) denotes the market size of potential consumers and \(c\) denotes marginal cost. Profit-maximizing prices must then satisfy the following system of B first-order conditions:

\[
s_{1j} + \sum_{k=1}^{B} (p_{1k} - c_k) \frac{\partial s_{1k}}{\partial \delta_{1j}} (-\alpha) = 0
\]
Given the nested logit specification assumed above and imposing symmetry, profit-maximizing markups \((m_{1k} = p_{1k} - c_k)\) are

\[
m_{1j} = \frac{1}{\alpha \left( \left( \frac{1}{1 - \sigma} \right) - \left( \frac{\sigma}{1 - \sigma} \right) \left( \frac{1}{N} \right) - \sum_{k=1}^{B} s_{1k} \right)}
\]

Given these demand assumptions, markups are equalized across a retailer’s brands and are (reassuringly) decreasing in \(N\). Marginal costs can then be recovered by subtracting markups from observed prices. To ensure strictly positive marginal costs, I specify \(c_k = \exp(X_k \gamma_X + W_k \gamma_W + U_k)\).

I am interested in the two cases of \(N = 1\) and \(N > 1\). Under the null hypothesis of monopoly, the demand and cost regressions are

\[
\ln(s_{1j}) - \ln(1 - \sum_{k=1}^{B} s_{1k}) = \sigma \ln(\hat{s}_{1j}) + X_{1j} \beta - \alpha p_{1j} + \xi_{1j}
\]

\[
\ln\left( p_{1j} - \frac{1}{\alpha(1 - \sum_{k=1}^{B} s_{1k})} \right) = X_j \gamma_X + W_j \gamma_W + U_j
\]

The proper regressions for a symmetric \(N\)-firm competing oligopoly are

\[
\ln(s_{1j}) - \ln(1 - N \sum_{k=1}^{B} s_{1k}) = \sigma \ln\left( \frac{s_{1j}}{N} \right) + X_{1j} \beta - \alpha p_{1j} + \xi_{1j}
\]

\[
\ln\left( p_{1j} - \frac{1}{\alpha \left( \left( \frac{1}{1 - \sigma} \right) - \left( \frac{\sigma}{1 - \sigma} \right) \left( \frac{1}{N} \right) - \sum_{k=1}^{B} s_{1k} \right)} \right) = X_j \gamma_X + W_j \gamma_W + U_j
\]

When \(N > 1\), the monopoly hypothesis forces inferred markups to be too large and inferred marginal costs to be too small and perhaps even negative. Joint estimation under the monopoly hypothesis accommodates the pricing equation by inflating \(\alpha\) (compared to the demand-only case). As own-price elasticities are \(\eta_k = -\alpha p_k \left( \left( \frac{1}{1 - \sigma} \right) - \left( \frac{\sigma}{1 - \sigma} \right) \frac{\hat{s}_k}{N} - s_k \right)\), joint estimation then maintains the elasticity implications by deflating \(\sigma\).

My proposed test is useful to the extent that estimated parameters differ across demand-only and joint estimation. The pricing equation cleanly illustrates, as the monopoly pricing equation and the generalized pricing equation are the same when \(\frac{N - \sigma}{N - N \sigma} = 1\). This condition is satisfied for any \(\sigma\) when \(N = 1\) and for any \(N\) when \(\sigma = 0\). The test’s power will therefore increase with the extent of inside-outside segmentation (\(\sigma\)) and with the actual number of firms in the market (\(N\)).

Another alternative hypothesis to the monopoly null is that the firm is part of a symmetric \(N\)-firm cartel characterized by the following demand and cost regressions:

\[
\ln(s_{1j}) - \ln(1 - N \sum_{k=1}^{B} s_{1k}) = \sigma \ln\left( \frac{s_{1j}}{N} \right) + X_{1j} \beta - \alpha p_{1j} + \xi_{1j}
\]

\[
\ln\left( p_{1j} - \frac{1}{\alpha (1 - N \sum_{k=1}^{B} s_{1k})} \right) = X_j \gamma_X + W_j \gamma_W + U_j
\]

As before, markups are equalized across brands, but, unlike the non-cooperative oligopoly case, the segmentation parameter \(\sigma\) appears nowhere in the pricing equation. Identification for my test is correspondingly weak, and it is unlikely that the monopoly null will be rejected when the true market structure is a collusive oligopoly.
III. Monte Carlo results

I evaluate the power of my proposed test by simulating datasets of firms setting profit-maximizing prices and resulting consumer choices. Each of the 100 simulated samples consists of 250 periods (markets) and thus roughly corresponds to observing a firm’s weekly data for five years. I let each retailer sell two brands, each of which is described by two observable characteristics and two observable factor prices.

The following is an extension of Berry (1994). Let each consumer’s utility in each market for the inside and outside goods be respectively given by

\[ u_{ij} = \beta_0 + \sum_k \beta_{xk} x_{kj} + \sigma_{xd} \xi_j - \alpha p_j + \epsilon_{ij} \]

and \( u_{i0} = \epsilon_{i0} \), with \( \epsilon_{ij} \) in both cases being the appropriate GEV error which depends on the parameter \( \sigma \). When compared to the \( \beta \) parameters, the parameter \( \sigma_{xd} \) conveys the relative importance of unobserved characteristics. I constrain marginal cost to be positive and specify it as \( c_j = \gamma_0 + \sum_k \gamma_{xk} x_{kj} + \sigma_{x} \xi_j + \sum_k \gamma_{wk} w_{kj} + \sigma_{\omega} \omega_j \). The variables \( x \) and \( \xi \) then correspond to observed and unobserved product characteristics that affect both demand and cost, while the variables \( w \) and \( \omega \) correspond to observed and unobserved variables that affect cost without affecting demand. Coefficients on unobserved variables (\( \sigma_{xd}, \sigma_{x}, \sigma_{\omega} \)) are parameters that capture the effect of the unobservables; \( \beta_0, \beta_{xk}, \alpha, \sigma, \gamma_0, \gamma_{xk}, \) and \( \gamma_{wk} \) are parameters to be estimated.

I create the exogenous data (\( x_{kj}, \xi_j, w_{kj}, \omega_j; k = 1, 2 \)) as independent standard normal random variables. Conditional on the inside-outside parameter being \( \sigma = 0, 0.5 \) and \( 0.75 \), I chose the true parameter values by somewhat ad hoc experiments. The primary goals were to create strictly positive unconditional purchase probabilities that averaged less than 0.05 and sufficient variance in purchase probabilities and prices. I furthermore chose \( \beta \) values relative to \( \sigma_d \) so that 20% of the total variance in exogenous product characteristics would be attributable to unobserved product characteristics. This level ensures that common econometric issues such as endogeneity bias arise but do not overwhelm the estimation routine. Parameter values can be found in Table 1. I computed prices and purchase probabilities to satisfy the firm’s first order conditions for each market.

Table 1 presents two estimation methods for the demand parameters under the monopoly null under various true market structures. The first estimation method is two-stage least squares where the dependent variable is \( \ln(s_{ij}) - \ln(s_0) \) such that \( s_0 \) and the right-hand side \( \tilde{s}_{ij} \) are constructed conditional on the market structure hypothesis and \( N \). Price and conditional market share are treated as endogenous variables with the brand’s observed cost factors \( w_j \) and the other brand’s characteristics \( x_j \) used as instruments. The second estimation method is the generalized method of moments, where the first stage is a two-stage least squares joint estimation of demand and cost using a simplex search over \( \alpha \) and \( \sigma \). The second stage employs the previous residuals to construct the optimal weighting matrix and then re-estimates to yield the efficient estimates under the null.

---

3 These segmentation levels roughly coincide with the simple logit, Einav’s (2007) movie estimate (\( \sigma = 0.52 \)), Berry and Waldfogel’s (1999) radio listening estimate (\( \sigma = 0.79 \)), and Li and Mou’s (2011) mobile phone estimates (\( \sigma \in (0.75, 0.80) \)).
I encounter no problems in estimation when the hypothesized market structure coincides with the true market structure of the simulated data. When those market structures do not match, however, two problems arise for some samples. First, markets with large unconditional purchase probabilities cannot be scaled up. For example, a true monopoly market with two brands each selling to 0.26 of the market cannot be treated as a duopoly. Second, the Hausman covariance matrix is for some markets of less than full rank. In either case, I discard those samples and draw additional samples to reach 100. I consider three different statistics to test the monopoly null. The Hausman specification test is a natural candidate; I also consider the tests of whether $\sigma$ is the same across the two estimations and whether any of the parameters differs across the two estimations.

Table 1 displays the Monte Carlo results. Moving from top to bottom, I consider three levels of decreasing inside-outside segmentation. From left to right, I consider different market structures and conduct under which the data were simulated. Panel A (high segmentation) shows the empirical means of point estimates and standard errors, as well as the average price and unconditional purchase probability under the three market structures. Scanning this top panel makes several things clear. First, maintaining the monopoly hypothesis when monopoly is the true market structure yields estimates very close to the true parameter values and joint estimation noticeably improves precision. Second, the monopoly hypothesis applied to the colluding duopoly data yields estimates (the intercept excepted) not much different from the true parameters. Furthermore, estimates are quite similar across the two estimations. Last, one can see the divergence between the two sets of estimates in the competing duopoly case. As expected, joint estimation inflates $\alpha$ and depresses $\sigma$ compared to the demand-alone estimation.

The statistical tests reflect these observations. In no samples do the Hausman statistics imply a rejection of the monopoly null when the true market structure is monopoly or collusive duopoly. The monopoly null, however, is rejected with 95% confidence 42% of the time when the data are generated under competing duopolists. Examining the individual parameter rejection rates, one sees that the monopoly null applied to the monopoly and collusive duopoly data generate rejections roughly in line with significance levels. Application to the competing duopoly data, though, generates high levels of rejection. As predicted by the theoretical discussion and shown in Panels B and C, the test’s power falls with lower values of $\sigma$; it has effectively no power in the simple logit case of $\sigma=0$.

IV. Conclusions

The proposed test has the potential to shed unexpected light on market structure and conduct given data from a single firm. While it requires substantial inside-outside segmentation, this requirement appears to be satisfied in several instances of the extant literature.
Appendix

From McFadden (1978), let $G: \mathbb{R}^{NB+1} \rightarrow \mathbb{R}$ be a nonnegative, homogeneous of degree one function with arguments $e^{\delta_k}$ that satisfies certain regularity conditions. Then the unconditional purchase probability of a consumer making choice $j$ is

$$s_j = \frac{e^{\delta_j} \frac{\partial G}{\partial e^{\delta_j}}}{G}$$

where $\delta_j$ is the mean utility of choice $j$. The particular nested logit used in the text employs

$$G = 1 + \left( \sum_{j=1}^{N} \sum_{b=1}^{B} e^{\delta_{jb}/(1-\sigma)} \right)^{1-\sigma}$$

for $N$ retailers and $B$ brands. This implies unconditional purchase probabilities for the observed retailer of

$$s_{1k} = \frac{e^{\delta_{1k}} \left( \sum_{j=1}^{N} \sum_{b=1}^{B} e^{\delta_{jb}/(1-\sigma)} \right)^{-\sigma} \left( e^{\delta_{1k}} \right)^{1-\sigma}}{1 + \left( \sum_{j=1}^{N} \sum_{b=1}^{B} e^{\delta_{jb}/(1-\sigma)} \right)^{1-\sigma}}$$

The implied conditional purchase probability

$$\tilde{s}_{1k} = \frac{e^{\delta_{1k}/(1-\sigma)}}{\sum_{j=1}^{N} \sum_{b=1}^{B} e^{\delta_{jb}/(1-\sigma)}}$$

can be substituted into the above to yield

$$s_{1k} = \frac{e^{\delta_{1k} \tilde{s}_{1k}}}{1 + \left( \sum_{j=1}^{N} \sum_{b=1}^{B} e^{\delta_{jb}/(1-\sigma)} \right)^{1-\sigma}}$$

which then yields the familiar nested logit equation for demand.
References
Table 1:
Monte Carlo parameter estimates and rejection rates under monopoly null, from 100 random samples

<table>
<thead>
<tr>
<th>Reality</th>
<th>Monopoly</th>
<th>Colluding duopoly</th>
<th>Competing duopoly</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>E(P)</td>
<td>E(s)</td>
<td></td>
</tr>
<tr>
<td>A. High-segmentation ( ( \sigma = 0.75 ))</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.24</td>
<td>0.044</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td>4.26</td>
<td>0.025</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td>3.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parameter</td>
<td>2SLS</td>
<td>Jt GMM</td>
<td>2SLS</td>
</tr>
<tr>
<td>( \sigma )</td>
<td>0.75</td>
<td>0.75</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.03)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>1</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(0.09)</td>
<td>(0.11)</td>
</tr>
<tr>
<td>( \beta_0 )</td>
<td>1</td>
<td>0.95</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>(0.39)</td>
<td>(0.33)</td>
<td>(0.38)</td>
</tr>
<tr>
<td>( \beta_{x1} )</td>
<td>1</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(0.08)</td>
<td>(0.10)</td>
</tr>
<tr>
<td>( \beta_{x2} )</td>
<td>1</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td>(0.08)</td>
<td>(0.10)</td>
</tr>
<tr>
<td>Hausman</td>
<td>0.57</td>
<td>0.55</td>
<td>12.20</td>
</tr>
<tr>
<td>90%/95%/99%</td>
<td>0/0/0</td>
<td>0/0/0</td>
<td>0.48/0.42/0.29</td>
</tr>
<tr>
<td>Individual rejection</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \sigma ): 95%/99%</td>
<td>0.06/0.00</td>
<td>0.07/0.03</td>
<td>0.76/0.68</td>
</tr>
<tr>
<td>any: 95%/99%</td>
<td>0.06/0.02</td>
<td>0.07/0.03</td>
<td>0.77/0.68</td>
</tr>
</tbody>
</table>

B. Mid-segmentation ( \( \sigma = 0.50 \))

|              | 4.44     | 0.43             | 2.15              |
| 90%/95%/99%  | 0/0/0    | 0/0/0            | 0.05/0.02/0.02    |
| Individual rejection |     |                  |                   |
| \( \sigma \): 95%/99% | 0.08/0.01 | 0.09/0.02 | 0.26/0.14        |
| any: 95%/99%  | 0.09/0.01 | 0.09/0.02 | 0.29/0.17        |

C. No segmentation ( \( \sigma = 0.00 \))

|              | 0.70     | 0.68             | 0.60              |
| 90%/95%/99%  | 0/0/0    | 0/0/0            | 0/0/0             |
| Individual rejection |     |                  |                   |
| \( \sigma \): 95%/99% | 0.10/0.03 | 0.10/0.03 | 0.09/0.02        |
| any: 95%/99%  | 0.10/0.03 | 0.11/0.03 | 0.09/0.02        |

Notes: The values given adjacent to coefficients in Panel A are empirical means and (standard errors); other values reflect empirical means of Hausman statistics and fraction of simulated samples for which rejection occurred at relevant confidence level.

Competing duopoly in panel A (B) required 119 (106) simulations to reach 100 acceptable samples.

The utility function is \( u_{ij} = \beta_0 + \beta_{x1} x_{1j} + \beta_{x2} x_{2j} + \sigma_d c_\theta - \alpha_p + \epsilon_{ij}; \sigma_{\theta d} = 2^{-0.5} \) for Panels A and B, \( \sigma_{\theta d} = 1 \) for Panel C.

Marginal cost is \( c_j = \exp(1 + 0.5 x_{1j} + 0.5 x_{2j} + 0.25 c_{\theta j} + 0.25 w_{1j} + 0.25 w_{2j} + 0.25 \omega_j); \) empirical means are 3.14.