Measuring the effect of local competition on prices in a context of spatial differentiation

Jacint Balaguer[a]
José C. Pernías[a]
[a]*Department of Economics, Universitat Jaume I, Spain*

Abstract

This paper examines the problem of estimation bias when the effect of local competition on prices is investigated. The hotel industry was chosen as a representative case of industries that offer a spatially differentiated product. It is shown that, in this context, including spatial variables as controls can be especially useful for reducing bias in the cross-sectional estimator. We also show how our results are robust to alternative size definitions for geographically relevant markets.

Keywords: Retail prices, spatial differentiation, local competition, cross-sectional estimator.

JEL codes: L11, L81, D43, C21

Corresponding author:
Jacint Balaguer
Department of Economics
Universitat Jaume I
Av. Vicente Sos Baynat s/n, 12071 Castelló, Spain
Phone: +34 964 728612
Fax: +34 964 728591
E-mail: jacint.balaguer@uji.es

1. Introduction

Although the traditional monopolistic competition approach still plays a central role in textbooks on microeconomics to explain how a change in the number of firms affects market prices, nowadays there is a considerable array of theoretical models that offer alternative predictions about the phenomenon.[1] In response to this situation, interest in the topic is turning increasingly toward the empirical testing of existing competitive models with real datasets, which are frequently obtained from different market structures. Empirical economists therefore often find themselves faced with the fact that the number of firms is highly correlated with unobserved but nonetheless relevant variables. This is the case where firms are more likely to be located in areas where consumer demand is higher. In this context, particularly careful analysis is required to shed light on the effect that the number of competitors has on prices from the conventional regression-based reduced form specifications.

In many study cases, location is an important attribute for consumers and fortunately researchers know how sellers are spatially differentiated with respect to their potential competitors. Therefore, the introduction of spatial variables may help to alleviate potential estimation bias by capturing heterogeneity from demands and costs. A representative case, where location has been acknowledged as the most essential attribute for consumers, is provided by the hotel industry (e.g., Bull, 1994). Evidence suggests that, in this industry, costs play a secondary role in hotel room pricing behavior (see, for example, Kotas, 1986). However, distribution of unobserved consumer demand over space has previously been recognized as a relevant problem in the study of hotel pricing behavior. This is the case of one particular paper by Mazzeo (2002a)[2], where the author introduced variables associated with the number of travelers in each motel location in order

[1] Some modern models offer predictions that are the contrary to those of the traditional monopolistic approach as regards the effect of competitors on price level. In this sense, price-increasing competition is seen to be unexceptional from the theoretical perspective (e.g., Rosenthal, 1980; Janssen and Moraga-González, 2004; Chen and Riordan, 2008). Evidence about the phenomenon remains scarce at the present time (i.e., Ward et al., 2002; Thomadsen, 2007; Melzer and Morgan, 2009).

[2] Mazzeo (2002a) and Mazzeo (2002b) also deal with the empirical treatment of endogeneity in the motel case when quantities or product choice are analyzed, respectively.

to prevent estimation bias when cross-sectional data is used for different markets (along U.S. interstate highways).

To explore the benefits of control variables, in our case we focused on local competition in the metropolitan area of Madrid. This choice made it is easy to introduce information about the distance of sellers from certain points of interest for consumers in the city, as well as discrete variables associated to some geographical areas. The empirical strategy to define relevant geographical markets follows several previous studies based on local competition on the retail gasoline industry. Hence, as in Barron et al. (2004) and in Lewis (2008), we will address the issue of how competitors placed within a certain radius around the seller's location affect their prices. Moreover, some extensions are performed in the present work. On the one hand, we will use the methodology recently developed by Kelejian and Prucha (2007) to prevent effects caused by heteroskedasticity and spatially autocorrelated disturbance terms. This procedure will be employed because unobserved effects related to location may be affecting the prices of closely located hotels and, hence, the regression disturbances are spatially dependent. On the other hand, since it is difficult to define the relevant market for each hotel, we will also explore whether our main results are robust to different radii.

2. Data, estimation and results

We consider a dataset of hotels inside the metropolitan area of Madrid[3] (see Data appendix for variable description and data sources). Pricing behavior will be analyzed separately for a standard room on each of four consecutive weekdays, where there is likely to be a swap-over between leisure and business consumers. Associated to each sort of consumer preference, individual regressions will then allow us to capture specific intensities of demands across the metropolitan sub-areas.

We assume the following regression model:

$$p_i = \alpha + X_i^{'}\beta + S_i^{'}\delta + C_i^{'}\gamma + \varepsilon_i \qquad (1)$$

[3] We decided to use the definition of this metropolitan area proposed by García Ballesteros and Sanz Berzal (2002), which comprises the city of Madrid and another 26 municipalities.

where p_i is the logarithm of the lodging price set by hotel i (for a standard double room); X_i is a vector of lodging-specific characteristics; S_i includes hotel-location variables to control heterogeneity in local demands (i.e. certain metropolitan sub-areas are more attractive than others due to proximity to cultural settlements, airport, etc.); C_i represents the number of competitors around hotel i; and, lastly, ε_i is a regression disturbance term. In the case that control variables for local demands to be omitted, the positive covariance between firm agglomeration and local demand intensity would introduce a positive bias in estimates of γ.

We perform OLS with standard errors robust to heteroskedasticity and spatial autocorrelation. Both complications can be taken into account with the consistent covariance matrix estimator developed by Kelejian and Prucha (2007).[4] The estimates related to each of the consecutive days for both the complete specification and an alternative specification that omits the S_i variables are reported. We are particularly interested in the estimates associated to C_i variables. More specifically, we split competitors into those that have the same category as the hotel under consideration (*CloseComp*) and those with a different category (*OtherComp*).

In Table 1 we show results obtained assuming as relevant geographical market for each hotel in 200-meter radius (which is approximately the average nearest-neighbor distance in the central districts). That is, all hotels which lie within a circle with this radius dimension are considered to be competitors of the hotel located at the center of the circle. Results for lodging-specific characteristics are quite similar in the two models. Most of the point estimates associated with the official hotel category are significant, thus indicating that the level of retail prices rises as new services are included and consumers are guaranteed better facilities. An exception is the inclusion of breakfast in room service. So, the weakly significant negative coefficient obtained for Thursday in the complete model could be because, in the presence of an excess capacity, sellers implement an aggressive strategy to capture consumers. In accordance with the higher

[4] In our application we use a Parzen kernel but, following the more recent analysis conducted by Lambert et al. (2008), with a larger bandwidth parameter than the one suggested by these authors. For the distance measure, d_{ij}, we use the Euclidean distance between each pair of hotels.

reputation of the AC and NH chains, the associated coefficients are all positive and, in most cases, strongly significant.[5] Moreover, average price is also affected by the size of the hotels. If (unobserved) hotel production were positively associated with hotel size, this last result would capture the existence of economies of scale.

[Insert Table 1 here]

In general the geographic variables that are included have a significant effect on prices and are jointly significant (as suggested by the Wald-test). Specific results suggest that intensity of local demand is lower in the suburbs. Moreover, it is also lower the farther the hotel is from two focal points in the metropolitan area. The distance from the airport is strongly significant for every weekday, while distance from the city center has a significant effect only on weekend days (where leisure consumers are probably more relevant). Furthermore, it is not surprising that high economic activity is associated to more local demand for hotel lodging, but only on midweek days (where there are more business consumers).

In accordance with the specification model that omits the S_i variables, we obtain the puzzling result that closest competitors (*CloseComp*) are not relevant, while competitors with a product from a different category (*OtherComp*) would have a significant influence on pricing behavior. Moreover, if we test with sellers from different categories, we can show how the results could be erroneously interpreted as evidence in favor of price-increasing competition. In the results yielded by the complete specification model, the opposite effect of unobserved intensities of demand seems to have been largely eliminated. In fact, from sellers that provide the same quality product, we find that greater competitive pressure clearly implies a fall in average retail prices. Unsurprisingly, the effect is smaller when the competitor belongs to a different category.

As indicated in the introduction, since we recognize the difficulty involved in defining the local relevant market for hotels, we further explore whether our findings are robust to increases in radii. In Table 2 we present the estimates related to closest competitors (*CloseComp*) and to different category competitors

[5] The "Key Audience Research" carried out by the Ipsos agency in 2008 on surveys among journalists indicated that the most highly valued chains in Spain are AC and NH.

(*OtherComp*) obtained from alternative dimensions for radii (i.e., for 400 and 600 meters, which are the approximate average nearest-neighbor distances in the suburbs and in the surrounding municipalities respectively).[6] Further results are, in general, consistent with the fact that the effect of local competitors on pricing behavior is progressively reduced as the average distance from them is increased. For example, in the complete specification model, a higher density of competitors with similar quality services (*CloseComp*) only reduces room price rates to a significant extent on Wednesdays and Thursdays when a radius of 400 meters is considered. Moreover, parameter estimates associated to different classes of competitors (*OtherComp*) are negative but not statistically significant (at the standard levels).

Interestingly, the set of spatial variables are also statistically significant, which in turn suggests the importance of unobservable heterogeneity in the determination of local prices. In fact, when these control variables are omitted in the regressions, the estimated coefficients related with competitors' density shifts toward a positive value. This allows us to illustrate how the coefficients associated to closest competitors (*CloseComp*) may also be mistakenly interpreted in favor of the price-increasing hypothesis. It is straightforward to show this from the regressions for Friday and Saturday, where these coefficients become significantly positive (in both the radii considered).

[Insert Table 2 here]

In sum, the outcome from this industry suggests that the unobserved distribution of demand can be very relevant in a spatially-differentiated context and could have a strong impact on the cross-sectional estimator. We have shown that in this case the introduction of spatial control variables, like the distance of sellers from some points of interest, is useful for the empirical analysis.

[6] For the sake of simplicity, in Table 2 we only address the effect of the variables *CloseComp* and *OtherComp*. The estimates for the remaining variables are remarkably insensitive to the choice of radius. The estimates for the remaining variables are available from the author upon request.

Acknowledgements

We are grateful to two anonymous referees for their helpful and insightful comments. Financial support from the Spanish Ministry of Science and Technology (ECO2008-06057/ECON), *Valencian Autonomous Government* (GV/2009/57) and *Fundació Caixa Castelló-Bancaixa* (P1 1A2007-19) are also gratefully acknowledged.

Data appendix

Variable definitions

On the one hand, the prices used correspond to lodging rates, including taxes, for one night in a standard double room. On the other hand, we divide characteristics of hotels into three vectors: X_i, C_i, S_i where X_i includes variables related with perceived quality and the number of services offered. One important way of indicating higher quality of an establishment is the official number of gold stars (we use the dummies *3Stars*, *4Stars* and *5Stars* for 3, 4 and 5 gold stars, respectively, where the reference group is 2 gold stars). The dummies *AC*, *NH* and *Tryp-Meliá* correspond to the names of the most important chains that operate in the Madrid area. *Rooms* is the number of rooms (in hundreds). *Breakfast* is a dummy variable that takes a value of one if the reported price includes breakfast. In vector S_i, we introduce two variables reflecting the logarithm of the Euclidean distances of each hotel from two focal points. We take into account the distance from the city center (*DCenter*) and we use the distance from the international airport (*DAirport*), which in turn is also near the Trade Fair Institution. A set of dummies distinguish hotels located in the suburbs (*Suburbs*), which are mainly residential and industrial districts, and in the surrounding municipalities (*Metrop*) that have important cultural and economic attractions. We also control for the level of economic activity by the logarithm of the GDP per capita of each municipality relative to the city of Madrid (*RGDPpc*).

Data sources

The lodging rates were collected on June 25, 2008, from the GTA Hotels website (http://www.gtahotels.com), and correspond to every night from July 9 to July 12 2008 (taxes included). The lodging rates in the sample account for approximately 70% of the total number of hotels, i.e. 217 out of 315. The GDP per capita and locations for all of hotels in the metropolitan area were taken from the Statistical Institute of the Community of Madrid. Characteristics of hotels were obtained from the "Guía Oficial de Hoteles, 2008, Turespaña, Ministerio de Industria, Turismo y Comercio".

References

Barron, J.M., Taylor, B.A., Umbeck, J.R.: Number of sellers, average prices and price dispersion, International Journal of Industrial Organization 22, 1041-1066 (2004)

Bull, A.O.: Pricing a motel's location, International Journal of Contemporary Hospitality Management 6, 10-15 (1994)

Chen, Y., Riordan, M.H.: Price-increasing competition, RAND Journal of Economics 39, 1042-1058 (2008)

García-Ballesteros, A., Sanz-Berzal, B. (eds.): Atlas de la Comunidad de Madrid en el Umbral del Siglo XXI. Comunidad de Madrid y Universidad Complutense (2002)

Janssen, M.C.W., Moraga-González, J.L.: Strategic pricing, consumer search and the number of firms, The Review of Economic Studies 71, 1089-1118 (2004)

Kelejian, H.H., Prucha, I.R.: HAC estimation in a spatial framework, Journal of Econometrics 140, 131-154 (2007)

Kotas, R.: Management accounting for hotels and restaurants. London: Surrey University Press (1986)

Lambert, D.M., Florax, R.J., Cho, S.-H.: Bandwidth selection for spatial HAC and other robust covariance estimators. Working Paper 08-10, Dept. of Agricultural Economics, Purdue University (2008)

Lewis, M.: Retail gasoline price dynamics and local market power, The Journal of Industrial Economics 56, 654-678 (2008)

Mazzeo, M.: Competitive outcomes in product-differentiated oligopoly, Review of Economics and Statistics 84, 719-728 (2002a)

Mazzeo, M.: Product choice and oligopoly market structure, RAND Journal of Economics 33, 221-242 (2002b)

Melzer, B.T., Morgan, D.P.: Price-increasing competition: The curious case of overdraft versus deferred deposit credit, Staff Report 391, Federal Reserve Bank of New York (2009)

Rosenthal, R.: A model in which an increase in the number of sellers leads to a higher price, Econometrica 48, 1575-1580 (1980)

Thomadsen, R.: Product positioning and competition: The role of location in the fast food industry, Marketing Science 26, 792-804 (2007)

Ward, M.B., Shimshack, J.P., Perloff, J.M., Harris, M.J.: Effects of the private-label invasion in food industries, American Journal of Agricultural Economics 84, 962-973 (2002)

Table 1. Estimation results based on radii (*r*) of 200 meters

	Wednesday		Thursday		Friday		Saturday	
Constant	4.3656 ***	4.8470 ***	4.3814 ***	4.8532 ***	4.2882 ***	4.8341 ***	4.3658 ***	4.8446 ***
	[0.0615]	[0.1107]	[0.0684]	[0.1187]	[0.0712]	[0.1089]	[0.0856]	[0.1149]
3stars	0.1562 **	0.1844 ***	0.1353 *	0.1610 **	0.1299 *	0.1483 **	0.0627	0.0824
	[0.0656]	[0.0623]	[0.0713]	[0.0676]	[0.0761]	[0.0743]	[0.0912]	[0.0891]
4stars	0.4123 ***	0.3993 ***	0.4036 ***	0.3889 ***	0.3151 ***	0.3041 ***	0.2456 ***	0.2360 ***
	[0.0799]	[0.0700]	[0.0846]	[0.0747]	[0.0789]	[0.0686]	[0.0920]	[0.0813]
5stars	1.0932 ***	1.0298 ***	1.0630 ***	0.9940 ***	1.0829 ***	0.9844 ***	1.0093 ***	0.9242 ***
	[0.1106]	[0.0937]	[0.1155]	[0.0976]	[0.1293]	[0.1015]	[0.1407]	[0.1186]
Breakfast	-0.0326	-0.0598	-0.0330	-0.0711 *	0.0654 *	0.0153	0.0536 *	0.0197
	[0.0381]	[0.0372]	[0.0384]	[0.0370]	[0.0335]	[0.0334]	[0.0304]	[0.0300]
AC	0.3223 ***	0.3571 ***	0.3155 ***	0.3487 ***	0.0657	0.0911	0.0595	0.0866
	[0.0589]	[0.0608]	[0.0587]	[0.0589]	[0.0760]	[0.0641]	[0.0734]	[0.0614]
NH	0.4492 ***	0.4116 ***	0.4398 ***	0.4061 ***	0.1658 ***	0.1137 ***	0.1645 ***	0.1215 ***
	[0.0531]	[0.0404]	[0.0536]	[0.0408]	[0.0539]	[0.0398]	[0.0522]	[0.0379]
Tryp-Meliá	-0.0078	-0.0196	-0.0436	-0.0598	-0.0126	-0.0228	0.0133	0.0099
	[0.0373]	[0.0457]	[0.0517]	[0.0562]	[0.0361]	[0.0344]	[0.0296]	[0.0338]
Rooms	-0.0484 ***	-0.0555 ***	-0.0439 ***	-0.0498 ***	-0.0477 ***	-0.0518 ***	-0.0457 ***	-0.0509 ***
	[0.0143]	[0.0121]	[0.0152]	[0.0131]	[0.0184]	[0.0156]	[0.0154]	[0.0138]
Metrop		-0.0158		-0.0192		0.0064		-0.0078
		[0.0808]		[0.0791]		[0.0846]		[0.0767]
Suburbs		-0.2419 ***		-0.2307 ***		-0.1967 ***		-0.1612 ***
		[0.0653]		[0.0687]		[0.0623]		[0.0556]
DAirport		-0.1374 ***		-0.1294 ***		-0.1330 ***		-0.1177 ***
		[0.0338]		[0.0356]		[0.0307]		[0.0299]
DCenter		-0.0183		-0.0235		-0.0784 ***		-0.0680 ***
		[0.0241]		[0.0261]		[0.0276]		[0.0207]
RGDPpc		0.2958 ***		0.2884 ***		0.0538		0.0659
		[0.0552]		[0.0554]		[0.0635]		[0.0601]
CloseComp	-0.0097	-0.0173 ***	-0.0096	-0.0177 ***	0.0018	-0.0092 **	0.0010	-0.0088 **
	[0.0060]	[0.0056]	[0.0061]	[0.0057]	[0.0034]	[0.0041]	[0.0035]	[0.0037]
OtherComp	0.0191 ***	0.0021	0.0181 **	0.0001	0.0295 ***	-0.0026	0.0297 ***	0.0011
	[0.0074]	[0.0067]	[0.0076]	[0.0074]	[0.0072]	[0.0091]	[0.0059]	[0.0058]
Mean of dep. var.	4.7444	4.7444	4.7471	4.7471	4.6192	4.6192	4.6295	4.6295
S.D. of dep. var.	0.3748	0.3748	0.3755	0.3755	0.3484	0.3484	0.3319	0.3319
R^2	0.5408	0.6329	0.5298	0.6180	0.4843	0.5778	0.5105	0.5944
Chi-square (5)		86.07 ***		76.92 ***		47.41 ***		61.33 ***

Standard errors that are robust to heteroskedasticity and spatial autocorrelation are between brackets (Parzen kernel with bandwidth parameter *dn* = 1650 meters). Significant estimates and statistics at the 10%, 5% or 1% levels are marked with *, ** and ***, respectively.

Table 2. Effect of competitors based on alternative numbers of meters for radii (*r*)

	Wednesday		Thursday		Friday		Saturday	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
r= 400								
CloseComp	-0.0005	-0.0061 **	-0.0008	-0.0066 **	0.0076 *	0.0006	0.0070 *	0.0009
	[0.0033]	[0.0027]	[0.0037]	[0.0030]	[0.0039]	[0.0032]	[0.0039]	[0.0032]
OtherComp	0.0049	-0.0021	0.005	-0.0021	0.0071 *	-0.006	0.0076 **	-0.0034
	[0.0034]	[0.0048]	[0.0038]	[0.0052]	[0.0039]	[0.0066]	[0.0036]	[0.0054]
R^2	0.54	0.64	0.53	0.62	0.49	0.58	0.52	0.60
Chi-square (5)		82.24 ***		75.92 ***		35.15 ***		37.88 ***
r= 600								
CloseComp	0.0014	-0.0034	0.0017	-0.0032	0.0084 **	0.0025	0.0084 **	0.0029
	[0.0032]	[0.0028]	[0.0036]	[0.0033]	[0.0039]	[0.0029]	[0.0041]	[0.0033]
OtherComp	0.0018	-0.0014	0.0015	-0.0019	0.0023	-0.0034	0.002	-0.0033
	[0.0022]	[0.0034]	[0.0027]	[0.0038]	[0.0027]	[0.0042]	[0.0030]	[0.0042]
R^2	0.54	0.63	0.53	0.62	0.51	0.58	0.53	0.60
Chi-square (5)		79.44 ***		71.60 ***		32.24 ***		32.10 ***

Standard errors that are robust to heteroskedasticity and spatial autocorrelation are between brackets (Parzen kernel with bandwidth parameter *dn* = 1650 meters). Significant estimates and statistics at the 10%, 5% or 1% levels are marked with *, ** and ***, respectively. (1) and (2) refer to estimates obtained from specification regressions without spatial control variables and with these variables, respectively.