Econometric Estimation of PCAIDS Models

by Germán Coloma (CEMA University, Buenos Aires, Argentina)[*]

Abstract

This paper presents a version of the proportionally calibrated almost ideal demand system (PCAIDS) model, useful for merger simulations, which can be econometrically estimated using price data for two firms in a market. The model is then applied to a database of the Argentine gasoline market, and its results are compared to the ones obtained with other alternative specifications.

JEL Classification: C51, C33, L71.

Keywords: PCAIDS, demand estimation, merger simulations, Argentine gasoline market.

1. Introduction

The applied economic literature generally uses demand models to estimate different aspects of the behavior of a market, group of markets or set of consumers' decisions. One demand model which is considerably popular is the so-called "almost ideal demand system" (AIDS), originally proposed by Deaton and Muellbauer (1980).

The estimations using AIDS are usually econometric, and require data on prices and expenditure shares (or revenue shares) for a number of goods or varieties of a certain good. One area in which these models have been extensively used is the analysis of the effects of a horizontal merger, especially when that merger involves differentiated products[1]. This use is particularly important to analyze the probable competitive effects of a proposed merger, since that analysis is generally an important input in the process of approving that merger by an antitrust authority.

The objective of the use of demand models such as AIDS in the process of evaluating a proposed merger is generally to perform a merger simulation. Merger simulations imply a procedure of forecasting the impact of a horizontal merger on prices and quantities in a certain market, and one of the most important factors that determine that impact is the cross elasticity between the products supplied by the merging firms. That elasticity, together with other own-price and cross elasticities, emerge as the result of an econometric estimation, if we are able to gather enough data to perform it.

In many merger contexts, however, the analyst does not have all the data necessary to

[*] The views and opinions expressed in this publication are those of the author and are not necessarily those of CEMA University. I thank comments by Andrés Chambouleyron and Pablo Spiller.

[1] In that area, the first important contribution is Hausman, Leonard and Zona (1994). Other developments along

perform a complete econometric demand estimation. This has generated the need to develop merger simulation models that keep some of the virtues of the econometric demand models but have less data requirements. One alternative based on AIDS is the so-called "proportionally calibrated almost ideal demand system" (PCAIDS), developed by Epstein and Rubinfeld (2002). This model has the same logic than AIDS but it incorporates a number of restrictions that make that all the elasticity values depend on a single parameter and on a set of revenue market shares. Making use of that property, a PCAIDS merger simulation can be performed using an estimation of the aggregate market price elasticity, an estimation of the own-price elasticity of a single product variety, and a vector of market shares for the different varieties that are sold in the market.

As we see, PCAIDS models do not require the econometric estimation of a complete demand system, but simply an estimation of an aggregate demand and an estimation of the elasticity of a certain variety, together with a single observation of market shares. Compared to the full AIDS model, the PCAIDS model has the advantage that it requires fewer data, and that it also imposes restrictions that guarantee that the estimated coefficients have the "right" signs and relative magnitudes[2]. PCAIDS estimations, however, are sensitive to the assumptions embedded in the additional restrictions imposed. Moreover, if more information is available, a PCAIDS model is not as efficient as the full AIDS model, since it is unable to incorporate that information to the estimation of the coefficients[3].

The objective of this paper is to develop an intermediate alternative between the full AIDS model and the PCAIDS model, that can be econometrically estimated but does not require as much data as the full AIDS model. In particular, we will be concerned with a case, which is likely in the context of a proposed merger evaluation, in which we have data on prices for the merging firms but not for the other firms in the market. We will also assume that we have data on market shares, in the form of series with the same number of observations than the price series. With those data, and some algebraic transformations, we will be able to use the standard econometric AIDS estimation methodology together with the restrictions imposed by the logic of PCAIDS, obtaining efficient estimators for the PCAIDS model parameters.

---

the same lines include Hausman and Leonard (1997) and Werden (1997).

[2] The logic of PCAIDS, for example, implies that if own-price elasticities are negative, then cross elasticities are always positive and smaller in absolute value, as well as other desirable properties of demand parameters.

[3] For example, if we have data on the evolution of market shares along time or about differences among market shares in different geographic regions or groups of buyers, this information is not used in a typical PCAIDS merger simulation, since that simulation is exclusively based on a "photograph" of average market shares at a certain point in time.

2. Theoretical analysis

Let us assume that we are dealing with a market with N varieties of a product. If we apply the logic of the AIDS model to the analysis of the demand of those varieties, then we can write the following market share equations (S_i), as functions of the natural logarithms of prices (P_i), and other demand-shifting variables (Y):

$$S_1 = a_{10} + a_{11}\cdot\ln(P_1) + a_{12}\cdot\ln(P_2) + ... + a_{1N}\cdot\ln(P_N) + a_{1Y}\cdot Y \quad (1)\ ;$$

$$S_2 = a_{20} + a_{21}\cdot\ln(P_1) + a_{22}\cdot\ln(P_2) + ... + a_{2N}\cdot\ln(P_N) + a_{2Y}\cdot Y \quad (2)\ ;$$

... ... ;

$$S_N = a_{N0} + a_{N1}\cdot\ln(P_1) + a_{N2}\cdot\ln(P_2) + ... + a_{NN}\cdot\ln(P_N) + a_{NY}\cdot Y \quad (3)\ .$$

The coefficients a_{ij} of a model like that have a correspondence with the price elasticities of the demand for the different product varieties. That correspondence is the following:

$$\eta_{ii} = -1 + \frac{a_{ii}}{S_i} + S_i\cdot(\eta+1) \quad ; \qquad \eta_{ij} = \frac{a_{ij}}{S_i} + S_j\cdot(\eta+1) \quad (4)\ ;$$

where η is the aggregate demand elasticity of the product under analysis, η_{ii} is the own-price elasticity of the ith variety, η_{ij} is the cross price elasticity of the ith variety with respect to the price of the jth variety, and market shares are defined so that their sum always adds up to one (that is, $\Sigma S_i = 1$, in every observation).

In a PCAIDS model, all these relationships also hold, together with some other assumptions about the relationships of the a_{ij} coefficients among themselves. These additional assumptions are essentially three: proportionality, adding-up and homogeneity. Under proportionality, we assume that sales are diverted away from a product variety according to the relative market shares of the other varieties. This implies that the cross-price parameters have the following relationship with the own-price parameters:

$$a_{ij} = \frac{-S_i}{1-S_j}\cdot a_{jj} \quad (5)\ .$$

Let us we now introduce an adding-up property, which implies that:

$$a_{ii} = -\sum_{j\neq i} a_{ij} \quad (6)\ .$$

Combining (5) and (6), we therefore have that:

$$a_{jj} = \frac{S_j \cdot (1 - S_j)}{S_i \cdot (1 - S_i)} \cdot a_{ii} \qquad (7)\ .$$

Finally, the homogeneity assumption implies that "$a_{ij} = a_{ji}$", and the matrix of price coefficients implicit in equations (1)-(3) can therefore be written in the following way:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1N} \\ a_{21} & a_{22} & \dots & a_{2N} \\ \dots & \dots & \dots & \dots \\ a_{N1} & a_{N2} & \dots & a_{NN} \end{bmatrix} = \begin{bmatrix} a_{11} & \frac{-S_2 \cdot a_{11}}{1 - S_1} & \dots & \frac{-S_N \cdot a_{11}}{1 - S_1} \\ \frac{-S_2 \cdot a_{11}}{1 - S_1} & \frac{S_2 \cdot (1 - S_2) \cdot a_{11}}{S_1 \cdot (1 - S_1)} & \dots & \frac{-S_2 \cdot S_N \cdot a_{11}}{S_1 \cdot (1 - S_1)} \\ \dots & \dots & \dots & \dots \\ \frac{-S_N \cdot a_{11}}{1 - S_1} & \frac{-S_2 \cdot S_N \cdot a_{11}}{S_1 \cdot (1 - S_1)} & \dots & \frac{S_N \cdot (1 - S_N) \cdot a_{11}}{S_1 \cdot (1 - S_1)} \end{bmatrix} \qquad (8)\ ;$$

whose elements consist of the same parameter (a_{11}) multiplied by different functions of the product varieties' market shares.

If we have data on prices and market shares for all the product varieties, a PCAIDS model can be estimated econometrically by substituting the elements of matrix A that appear in (8) into the system of equations (1)-(3). This implies writing that system of equations in the following way:

$$S_1 = a_{10} + a_{11} \cdot \frac{(1 - S_1) \cdot \ln(P_1) - S_2 \cdot \ln(P_2) - \dots - S_N \cdot \ln(P_N)}{1 - S_1} + a_{1Y} \cdot Y \qquad (9)\ ;$$

$$S_2 = a_{20} + a_{11} \cdot \frac{S_2 \cdot [-S_1 \cdot \ln(P_1) + (1 - S_2) \cdot \ln(P_2) - \dots - S_N \cdot \ln(P_N)]}{S_1 \cdot (1 - S_1)} + a_{2Y} \cdot Y \qquad (10)\ ;$$

... ... ;

$$S_N = a_{N0} + a_{11} \cdot \frac{S_N \cdot [-S_1 \cdot \ln(P_1) - S_2 \cdot \ln(P_2) - \dots + (1 - S_N) \cdot \ln(P_N)]}{S_1 \cdot (1 - S_1)} + a_{NY} \cdot Y \qquad (11)\ ;$$

and estimating it using a technique that imposes the restriction that the coefficient a_{11} must be the same in the N equations.

Let us now assume that we only have information about two prices (P_1 and P_2) and the N market shares. Then the system has to be re-written in order to avoid having the other prices (P_3, P_4, ... P_N) as independent variables. This can be done by inverting the system, letting the natural logarithms of the prices be the dependent variables, and letting the market shares be the independent variables.

Unfortunately, an operation like the one described in the previous paragraph implies inverting matrix A, which –under the restrictions imposed by the PCAIDS model– is singular

and therefore not invertible. Nevertheless, we are still able to eliminate one row and one column of matrix A (let's say, the ones corresponding to $\ln(P_2)$ and S_2) and form a new matrix with dimension (N-1)x(N-1), which we will call matrix B, that is invertible. Performing that operation implies:

$$B^{-1} = \begin{bmatrix} a_{11} & a_{13} & \dots & a_{1N} \\ a_{31} & a_{33} & \dots & a_{3N} \\ \dots & \dots & \dots & \dots \\ a_{N1} & a_{N3} & \dots & a_{NN} \end{bmatrix}^{-1} = \begin{bmatrix} a_{11} & \frac{-S_3 \cdot a_{11}}{1-S_1} & \dots & \frac{-S_N \cdot a_{11}}{1-S_1} \\ \frac{-S_3 \cdot a_{11}}{1-S_1} & \frac{S_3 \cdot (1-S_3) \cdot a_{11}}{S_1 \cdot (1-S_1)} & \dots & \frac{-S_3 \cdot S_N \cdot a_{11}}{S_1 \cdot (1-S_1)} \\ \dots & \dots & \dots & \dots \\ \frac{-S_N \cdot a_{11}}{1-S_1} & \frac{-S_3 \cdot S_N \cdot a_{11}}{S_1 \cdot (1-S_1)} & \dots & \frac{S_N \cdot (1-S_N) \cdot a_{11}}{S_1 \cdot (1-S_1)} \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} \frac{(S_1+S_2) \cdot (1-S_1)}{S_2 \cdot a_{11}} & \frac{S_1 \cdot (1-S_1)}{S_2 \cdot a_{11}} & \dots & \frac{S_1 \cdot (1-S_1)}{S_2 \cdot a_{11}} \\ \frac{S_1 \cdot (1-S_1)}{S_2 \cdot a_{11}} & \frac{(S_1+S_2) \cdot S_1 \cdot (1-S_1)}{S_3 \cdot S_2 \cdot a_{11}} & \dots & \frac{S_1 \cdot (1-S_1)}{S_2 \cdot a_{11}} \\ \dots & \dots & \dots & \dots \\ \frac{S_1 \cdot (1-S_1)}{S_2 \cdot a_{11}} & \frac{S_1 \cdot (1-S_1)}{S_2 \cdot a_{11}} & \dots & \frac{(S_1+S_2) \cdot S_1 \cdot (1-S_1)}{S_N \cdot S_2 \cdot a_{11}} \end{bmatrix} \quad (12) .$$

Another matrix operation that we will need is to multiply the inverse of matrix B by the opposite of the omitted column of matrix A (that is, the one that corresponds to the variable $\ln(P_2)$). This multiplication gives an N-1 column vector of ones, that is:

$$B_2 = B^{-1} \cdot (-A_2) = \begin{bmatrix} a_{11} & a_{13} & \dots & a_{1N} \\ a_{31} & a_{33} & \dots & a_{3N} \\ \dots & \dots & \dots & \dots \\ a_{N1} & a_{N3} & \dots & a_{NN} \end{bmatrix}^{-1} \cdot \begin{bmatrix} -a_{12} \\ -a_{32} \\ \dots \\ -a_{N2} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \dots \\ 1 \end{bmatrix} \quad (13) .$$

Let us now write the system of equations (1)-(3) in the following way:

$$\ln(P_1) = b_{10} + b_{11} \cdot S_1 + b_{13} \cdot S_3 + \dots + b_{1N} \cdot S_N + b_{12} \cdot \ln(P_2) + b_{1Y} \cdot Y \quad (14) ;$$

$$\ln(P_3) = b_{30} + b_{31} \cdot S_1 + b_{33} \cdot S_3 + \dots + b_{3N} \cdot S_N + b_{32} \cdot \ln(P_2) + b_{3Y} \cdot Y \quad (15) ;$$

$$\dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad ;$$

$$\ln(P_N) = b_{N0} + b_{N1} \cdot S_1 + b_{N3} \cdot S_3 + \dots + b_{NN} \cdot S_N + b_{N2} \cdot \ln(P_2) + b_{NY} \cdot Y \quad (16) ;$$

and substitute (12) and (13) into it. What we obtain is the following:

$$\ln(P_1) = b_{10} + \frac{S_1 \cdot (1-S_1)}{a_{11} \cdot S_2} + \ln(P_2) + b_{1Y} \cdot Y \quad (17) ;$$

$$\ln(P_3) = b_{30} + \frac{S_1 \cdot (1 - S_1)}{a_{11} \cdot S_2} + \ln(P_2) + b_{3Y} \cdot Y \qquad (18)\ ;$$

... ... ;

$$\ln(P_N) = b_{N0} + \frac{S_1 \cdot (1 - S_1)}{a_{11} \cdot S_2} + \ln(P_2) + b_{NY} \cdot Y \qquad (19)\ .$$

If we do not have data on P_3, P_4, ... P_N, we will not be able to estimate equations (18)-(19), but we will still be able to estimate equation (17), for which we only need data on P_1 and P_2. Estimating that equation will give a value for coefficient a_{11}, which we could obtain as a direct estimate if we invert equation (17) and express it in the following way:

$$\frac{S_1 \cdot (1 - S_1)}{S_2} = -a_{11} \cdot b_{10} + a_{11} \cdot \ln\left(\frac{P_1}{P_2}\right) - a_{11} \cdot b_{1Y} \cdot Y \qquad (20)\ .$$

If we interpret equation (20) as representative of the behavior of price-taking buyers, this formulation has an additional advantage, which is the fact that all its dependent variables (P_1, P_2, Y) can be considered as exogenous to those buyers[4]. That does not happen if we estimate equation (17), which has the endogenous expression "$S_1 \cdot (1-S_1)/S_2$" as a dependent variable.

By estimating a_{11} through equation (20) we are able to recover all the parameters of matrix A assumed by the PCAIDS model. If we also estimate an aggregate demand equation of the following form:

$$\ln(Q) = c_0 + \eta \cdot \ln(P_A) + c_Y \cdot Y \qquad (21)\ ;$$

where Q is the aggregate product quantity and P_A is an average price[5], then we can use its estimated value for η, together with the estimated value for a_{11} from equation (20), in order to calculate average price elasticities for every product variety. This is the main virtue of the PCAIDS model, stressed by its proponents[6].

3. An application to the Argentine gasoline market

The Argentine gasoline market is a good example to illustrate the methodology that we have proposed in the previous section, since it is a market in which the Argentine

[4] This advantage does not hold if we estimate equation (20) as part of a system of demand and supply equations, since in that case the two prices (or at least one of them) become endogenous variables.
[5] This price is ideally the weighted average of the prices of the N varieties sold. If we only have data for P_1 and P_2, then it will have to be the weighted average of those two prices only.
[6] See, for example, Epstein and Rubinfeld (2002), or Epstein and Rubinfeld (2004).

Department of Energy compiles monthly quantity and price data. That data is available discriminated by brand and by province or urban area. In that market, there are four major national brands (YPF, Shell, Esso and Eg3) and several minor or regional brands. Each of the major national brands has a market share of more than 10%, and they jointly control over 90% of the Argentine gasoline supply[7].

TABLE 1: ARGENTINE GASOLINE MARKET, 1998-2000

Jurisdiction	Quantity	Relative market share (%)				Average gross price ($/liter)			
	(m3)	YPF	Shell	Esso	Eg3	YPF	Shell	Esso	Eg3
Bs Aires Province	5.649.650	41,47	24,48	21,19	12,87	0,9912	1,0305	0,9947	1,0159
Bs Aires City	1.667.876	40,84	35,22	18,37	5,58	0,9674	1,0280	0,9709	1,0076
Catamarca	83.008	53,13	12,41	28,91	5,55	0,9768	1,0234	0,9721	1,0077
Chaco	204.616	47,12	34,08	15,71	3,10	0,9939	1,0319	0,9954	1,0093
Chubut	470.718	41,92	6,28	6,94	44,86	0,5100	0,5251	0,5246	0,5261
Cordoba	1.139.086	52,08	21,85	19,11	6,96	0,9742	1,0276	0,9893	1,0071
Corrientes	195.236	49,35	38,32	9,26	3,07	0,9878	1,0362	1,0010	1,0207
Entre Rios	400.178	48,98	23,44	17,77	9,81	0,9778	1,0227	0,9863	0,9979
La Pampa	220.746	41,76	10,46	7,79	39,99	0,9804	1,0383	0,9816	1,0085
La Rioja	126.055	64,15	16,71	11,56	7,58	0,9899	1,0288	0,9951	1,0267
Neuquen	250.527	71,73	9,06	8,72	10,49	0,9919	1,0509	1,0334	1,0382
Rio Negro	294.810	61,69	9,64	7,52	21,15	0,9884	1,0352	1,0154	1,0337
San Juan	186.054	76,46	10,84	8,16	4,54	0,9811	1,0085	0,9815	1,0062
San Luis	142.701	52,85	5,86	25,50	15,78	0,9813	1,0186	0,9844	1,0038
Santa Fe	1.086.087	38,12	30,17	24,39	7,32	0,9700	1,0296	0,9837	0,9946
Total	12.117.347	44,68	24,34	18,91	12,08	0,9606	1,0221	0,9788	0,9433

Source: Department of Energy of Argentina.

In fifteen out of the twenty-four jurisdictions in which Argentina is divided (twenty-three provinces plus the federal district of Buenos Aires City), the four major national brands are also the four largest gasoline suppliers. These jurisdictions are the city of Buenos Aires and the provinces of Buenos Aires, Catamarca, Chaco, Chubut, Córdoba, Corrientes, Entre Ríos, La Pampa, La Rioja, Neuquén, Río Negro, San Juan, San Luis and Santa Fe[8]. With those jurisdictions we have built a database of 540 observations, that consists of one monthly observation per jurisdiction for the period January 1998-December 2000. The main figures of that database are summarized in table 1. For each observation we have a price and a quantity for each of the four major national gasoline brands, out of which we can calculate series of "relative market shares" (i.e., market shares relative to the group of major national brands) for each brand. We also have some data on characteristics of the jurisdictions (e.g., population and area) and some data on general economic variables (GDP per capita, consumer price index and international oil prices).

As this is an example with a good deal of information, we were able to estimate

[7] These figures correspond to the period 1998-2000. In December 2001 the majority of the shares of Eg3 was bought by the Brazilian firm Petrobras, which since then has been replacing the brand Eg3 by its own one.

[8] In the other nine provinces (Formosa, Jujuy, Mendoza, Misiones, Salta, Santa Cruz, Santiago del Estero, Tierra del Fuego and Tucumán), either there is at least one major national brand that is not present or there is a regional brand which is largest than at least one major national brand.

demands in several alternative ways. The first approach that we used consisted of estimating the following system of equations:

$$\ln(Q_j/Pop) = \Sigma c_i(j)\cdot D_i + c_p\cdot\ln(Pop/Area) + \Sigma c_j\cdot\ln(P_j) + c_g\cdot\ln(GDPpc) \qquad (22)\ ;$$

for "j = YPF, Shell, Esso, Eg3" (i.e., for each major gasoline brand) and for "i = 1, ... 15" (i.e., for each jurisdiction). In that system of equations, Q_j is the gasoline quantity sold, Pop is the population, Area is the area of each jurisdiction, P_j is the gross price for each brand, and GDPpc is an estimation of the gross domestic product per capita in each observation.

Under this specification, the coefficients "c_j" are the price elasticities of the different demands, while "c_g" is the income elasticity. We tried several estimation strategies for these figures, but we got the best results when we used a seemingly unrelated regression (SUR), and when we imposed the restrictions that all the own-price elasticities be the same, all the cross elasticities be the same, and the income elasticities be such that, for each equation, the sum of all elasticities adds up to zero.

The results of this estimation can be compared with the ones obtained using a full PCAIDS model. In this case, such a model can be estimated using the following system of equations:

$$S_j = \Sigma c_i(j)\cdot D_i + c_p\cdot\ln(Pop/Area) + c_j\cdot\frac{(1-S_j)\cdot\ln(P_j) - \Sigma S_k\cdot\ln(P_k)}{1-S_j} + c_g\cdot\ln(GDPpc) \quad (23)\ ;$$

for "j = YPF, Shell, Esso, Eg3" and for "i = 1, ... 15", while imposing the restriction specified by equation (7) on the different "c_j" coefficients. Together with that system of equations, we can also estimate an aggregate gasoline demand like the following:

$$\ln(Q_{tot}/Pop) = \Sigma c_i(j)\cdot D_i + c_p\cdot\ln(Pop/Area) + c_{avg}\cdot\ln(P_{avg}) + c_g\cdot\ln(GDPpc) \qquad (24)\ ;$$

where Q_{tot} is total gasoline quantity and P_{avg} is the average price of gasoline (i.e., the weighted average of the gasoline prices of each major brand, corresponding to each observation).

Under this specification, the coefficients "c_j" are the PCAIDS coefficients, while "c_{avg}" is the aggregate market price elasticity of gasoline. We also tried several estimation strategies for these figures, and we got the best results when we used two-stage least squares (2SLS) for the system and ordinary least squares (OLS) for the aggregate demand.

A third possible strategy consists of estimating three equations (which correspond to three out of the four major national brands), that follow the specification stated by equation (20). In our case those equations can be written in the following way:

$$\frac{S_{ypf} \cdot (1 - S_{ypf})}{S_j} = \Sigma c_i(j) \cdot D_i + c_p \cdot \ln(Pop/Area) + c_{ypf} \cdot \ln(P_{ypf}/P_j) + c_g \cdot \ln(GDPpc) \qquad (25) ;$$

for "j = Shell, Esso, Eg3" and for "i = 1, ... 15", where "c_{ypf}" is the PCAIDS coefficient for YPF. This system of equations has to be estimated together with equation (24), so that we can get a number for the aggregate price elasticity of gasoline.

Using data for all the brands, we estimated equation (25) as a system and got the best results when we applied SUR. We also assumed a case where we had price information for only two brands (YPF and Shell), and estimated only one out of the three equations using OLS.

TABLE 2: ESTIMATION RESULTS

Concept	coefficient	std error	t-statistic	p-value
Estimation 1				
YPF own elasticity	-1,259186	0,193025	-6,523452	0,0000
YPF cross elasticity	0,373186	0,069566	5,364459	0,0000
R squared YPF	0,933065			
Estimation 2				
YPF PCAIDS coeff	-0,310577	0,073157	-4,245359	0,0000
Aggregate elasticity	-0,284380	0,104746	-2,714944	0,0068
R squared YPF	0,927455			
Estimation 3				
YPF PCAIDS coeff	-2,143999	0,931445	-2,301799	0,0215
Aggregate elasticity	-0,284380	0,104746	-2,714944	0,0068
R squared YPF	0,787810			
Estimation 4				
YPF PCAIDS coeff	-2,247674	1,304455	-1,723076	0,0855
Aggregate elasticity	-0,221856	0,103271	-2,148289	0,0322
R squared YPF	0,787813			

The main results of our four estimations appear on table 2. In it we see that the first two estimation strategies that we used (i.e., the one that directly estimated the price elasticities and the one that used the full PCAIDS model with four equations) gave us the best goodness of fit, with R squared coefficients for the YPF equation of over 0.92. On the contrary, the other two estimations strategies (i.e., the one that used PCAIDS with three equations following equation (25), and the one that only estimated one equation for YPF and Shell) got smaller but similar R squared coefficients and also similar results. We can also see that the PCAIDS YPF coefficient estimated by the second and third estimation strategies are considerably different, but that the estimations obtained under the third and fourth specifications are basically the same. The coefficients for the aggregate elasticity, on the contrary, are the same under estimations 2 and 3, because they have been calculated using the same regression. Under estimation 4 the coefficient is different, because in the regression we

have used a different definition of the average price (using YPF and Shell data only).

The relative equivalence between the first and second estimation strategies can also be observed when we look at the figures that appear on table 3. In that table we have calculated, using equation (4) and numbers from tables 1 and 2, the average estimated elasticities for YPF (both the own-price elasticity and the cross elasticity with respect to Shell) predicted by the different versions of the PCAIDS model (estimations 2, 3 and 4). We have also calculated, inverting equation (4), the implicit PCAIDS coefficient for YPF and the implicit aggregate market elasticity that are consistent with the coefficients obtained under estimation 1. The numbers that appear on table 4 also serve to see that under estimations 3 and 4 the average estimated elasticities are also very similar between themselves, but different from the ones calculated under estimations 1 and 2.

TABLE 3: COMPARISON OF RESULTS

Concept	a(YPF)	Elast(Agr)	Elast(YY)	Elast(YS)
Estimation 1	-0,202434 (*)	-0,217088 (*)	-1,259186	0,373186
Estimation 2	-0,310577	-0,284380	-1,375468 (*)	0,584629 (*)
Estimation 3	-2,143999	-0,284380	-5,479297 (*)	3,007755 (*)
Estimation 4	-2,247674	-0,221856	-5,683424 (*)	3,159993 (*)

(*): Average estimated coefficients.

4. Conclusions

The main conclusions of this paper can be summarized as follows:

a) Not only is the PCAIDS model useful to perform merger simulations in the absence of the data necessary to perform econometric estimations, but also can it be useful to incorporate restrictions to an econometric estimation of an AIDS model, in order to improve the power of that estimation (and, especially, the significance of the estimated elasticity coefficients).

b) The PCAIDS model is also useful to perform econometric estimations when we have enough price data for a few firms but not for all the firms in the market, since its basic coefficient can be estimated using a single equation that depends on price series for two firms only.

c) However, as we show in an example that we prepared with data from the Argentine gasoline market, this last simplification can alter the results, something that is not likely to occur if we are able to estimate a PCAIDS model using price and quantity data for all the

firms that operate in the market.

References

Deaton, Angus and Muellbauer, John (1980). "An Almost Ideal Demand System"; *American Economic Review*, vol 70, pp 312-326.

Epstein, Roy and Rubinfeld, Daniel (2002). "Merger Simulation: A Simplified Approach with New Applications"; *Antitrust Law Journal*, vol 69, pp 883-919.

Epstein, Roy and Rubinfeld, Daniel (2004). "Merger Simulation with Brand-Level Margin Data: Extending PCAIDS with Nests"; *Advances in Economic Analysis and Policy*, vol 4, no 1, article 2.

Hausman, Jerry and Leonard, Gregory (1997). "Economic Analysis of Differentiated Products Mergers Using Real World Data"; *George Mason Law Review*, vol 5, pp 321-346.

Hausman, Jerry; Leonard, Gregory and Zona, Douglas (1994). "Competitive Analysis with Differentiated Products"; *Annales d'Économie et de Statistique*, vol 34, pp 159-180.

Werden, Gregory (1997). "Simulating the Effects of Differentiated Products Mergers: A Practical Alternative to Structural Merger Policy"; *George Mason Law Review*, vol 5, pp 363-386.