ISSN 1415-4765

TEXTO PARA DISCUSSÃO Nº 793

ESTIMATION OF THE BRAZILIAN CONSUMER DEMAND SYSTEM *

Seki Asano **
Eduardo P. S. Fiuza***

Rio de Janeiro, maio de 2001

* This work has been undertaken under a cooperation agreement between IPEA and Japan International Cooperation Agency (Jica), whose support is gratefully acknowledged. We thank Eustáquio Reis and Octávio Tourinho for comments, support and encouragement. We are also grateful to Rita Sbano, Fátima Louro, Eulina Nunes, Lúcia Pinto, Edílson Silva, Nézio Pontes and Márcia Quintslr of IBGE for their patient assistance in providing the required data and documentation and clarifying their obscure points. Viviane Silva, Paola Dias and Márcia Pimentel's help on SAS programming is highly appreciated. Erica Novello and Adrian Heringer's assistance is gratefully acknowledged. Remaining errors are of our exclusive responsibility.

** From Tokyo Metropolitan University, Department of Economics, Tokyo, Japan.

*** From Instituto de Pesquisa Econômica Aplicada (IPEA), Diretoria de Estudos Macroeconômicos (DIMAC), Rio de Janeiro, Brasil.

SUMMARY

RESUMO

Neste estudo, estimamos o sistema de demanda de consumo brasileiro através de dados de dispêndios das famílias, que cobrem todas as categorias de consumo, e de um novo conjunto de índices regionais de custo de vida. As fontes de dados de dispêndios são as duas Pesquisas de Orçamentos Familiares (POF) do IBGE, realizadas em 1986/87 e 1995/96, que coletaram dados das nove áreas metropolitanas, Distrito Federal e município de Goiânia, e que serviram de base para as atualizações do Sistema Nacional de Índices de Preços ao Consumidor. Os índices de preços correspondentes foram construídos a partir de preços detalhados de bens e serviços, tornados públicos pelo IBGE para cada uma das regiões pesquisadas. Até onde sabemos, este é o primeiro estudo desse gênero com base nas POFs. O trabalho destaca-se da literatura existente por: *a*) contar com variações de preços tanto ao longo do tempo como entre regiões, o que nos permite estimar elasticidades de preços com alta precisão; *b*) observar grandes variações na renda (dispêndio total), o que é raramente disponível em dados agregados; e *c*) poder usar controles para fatores específicos de tempo, explorando a estrutura de painel do banco de dados.

Ao contrário do que costuma acontecer em estudos empíricos de sistemas de demanda, os resultados da estimação mostram consistência com a teoria da demanda, e as elasticidades estimadas são próximas ao esperado pelo senso econômico comum. O sistema estimado servirá como base microeconômica para avaliar várias questões relacionadas à política econômica.

Palavras-chaves: Sistema de demanda de consumo, Almost Ideal Demand System, dispêndios das famílias brasileiras.

ABSTRACT

In this study we estimate the Brazilian consumer demand system through family expenditure data, which cover all consumption categories. The model is estimated from family-level expenditures on seven consumption categories, and a new set of regional cost-of-living indexes. The sources for expenditures are the national expenditure surveys conducted in 1986/87 and 1995/96, which collected data from 11 metropolitan areas. Corresponding price indexes were constructed from detailed commodity prices, also from each metropolitan area. The salient features of our study are: *a*) price variations come from both time and regional differences, which allows us to estimate price elasticities with high precision; *b*) we have large variations in income (total expenditures) which is rarely available in aggregated data; and *c*) we can control for time specific factors by exploiting the panel structure of the data set.

Contrary to the rule of thumb in empirical studies of demand systems, the estimation results display consistency with demand theory, and elasticity estimates are close to economic common sense. The estimated system will serve as a micro-economic basis for evaluating various policy-related issues.

Keywords: Consumer demand System, Almost Ideal Demand System, Brazilian family expenditures.
JEL: C33, C81, D12.

1 - INTRODUCTION

This study estimates the Brazilian consumer demand system based on family expenditure data, for all the consumption categories, and their corresponding price indexes. The data sources for expenditures are the national expenditure surveys conducted in 1986/87 and 1995/96 (named Pesquisa de Orçamentos Familiares, or simply POF) by Instituto Brasileiro de Geografia e Estatística (IBGE). The sources for price indexes are the monthly national survey of consumer prices.

The next section reviews previous estimations of consumer demand systems in Brazil. Section 3 introduces our model, an extension of the Almost Ideal Demand System. Section 4 describes the data utilized. Section 5 displays the results and Section 6 summarizes the conclusions. An Appendix describes the construction of our regional price indexes, another contribution of this work.

2 - PREVIOUS STUDIES

Few studies exist on estimation of consumer demand in Brazil. None of them used POF, but rather on older surveys. The methodology adopted varied widely.

Medeiros (1978) estimated Engel elasticities for food and education in the City of São Paulo using a local Family Expenditure Survey conducted in 1971/72 by the University of São Paulo (henceforth called POF-USP). He modeled demand with a Box-Cox transformation for both dependent variable (expenditure on food or on education) and explanatory variable (total expenditure), allowing for more flexibility of the functional specification. He restricted the Box-Cox parameter to be equal for both sides (estimated λ's were –0.85 for food and from 0.03 to 0.14 for education, depending on inclusion or not of additional explanatory variables).

Estudo Nacional de Despesas Familiares (Endef), a comprehensive survey undertaken from August 1974 to August 1975 in all metropolitan and urban areas, and rural areas in the Southern, Southeastern and Northeastern regions, was the data source for a large number of studies. Rossi (1982) modeled demand using a Lorenz curve for concentration, following Kakwani (1977*a* and *b*, 1978), and applied it to Endef data from the City of Rio de Janeiro. He also compared the results with estimates obtained by a Box-Cox transformation, and found them very similar, except for the tails of the income distribution, where Kakwani's method allows more flexibility (including non-monotonic paths) but ends up coming out with unreasonable patterns. Rossi (1983*b*) retrieved these estimates and compare them with the ones obtained for São Paulo using POF-USP and with all metropolitan areas using Endef. Box-Cox estimates are also reported for São Paulo. The sum of residuals is lower for Kakwani's method, but it is worth remarking that this superiority is mostly due to three expenditure groups: clothing, recreation and health care.

Hoffmann (1983 and 1988) proposed an alternative framework, a piecewise linear regression, and obtained estimates for Rio de Janeiro not far from Rossi's. Although his fit was better — an elementary result, since he only added more variables, as Rossi (1983*a*) points out —, it is worth noting that it came at expense of finding different break points for each commodity run. Another caveat is the omission of standard error. Both Rossi's and Hoffmann's sets of elasticities add up to one approximately, but neither one makes use of demographic variables.

Another approach was undertaken by Rossi and Neves (1987); by transforming a logit to a linear specification and using the same explanatory variables, he was able to estimate each equation by OLS. The explanatory variables were income (total expenditure), prices and demographic variables (age of head, family size). As opposed to the previous articles, the estimation makes use of all regions covered by Endef. The following elasticities are reported for nine income classes: income and family size. Fit is reasonably well-done, but unfortunately standard errors are also missing. He found income-inelastic demands for food and elastic demands for transportation and education and reading were found above 1. Elasticities with respect to family size suggest that, as opposed to our results below, economies of scale are only present in food, transportation and education, indicating that these expenditures on these items increase at the expenses of the category housing and other expenses. Price elasticities are not reported.

The same approach of logit transformation to OLS had been chosen by Cipriano and Brandt (1983) for agricultural products (food and tobacco), but the authors added some original variables, such as: state population, regional dummies and an income inequality index. Income elasticity estimates are suspiciously too low, though — probably due to omission of other goods — and standard errors were also not reported.

Simões and Brandt (1981) ran an expanded linear expenditures system (Eles) using all expenditure categories, but reported both price and income elasticities only for food and tobacco. Parameter estimates attained 10% significance. Unfortunately additional explanatory variables were simply not listed. The adding-up criterion was met.

Thomas, Strauss and Barbosa (1989) ran a generalized version of Almost Ideal Demand System (Aids) on Endef data, also on all regions covered. Because of the authors' concern on the use of the estimates for agricultural policymaking the system is very disaggregate, containing some detailed food commodity items as dependent variables. However, disaggregation is limited, mainly by the difficulty to handle non-zero expenditures, i.e., commodities which are never purchased or their purchase was not recalled by the household during the survey. Prices are estimated indirectly by the ratio expenditure/quantity (Endef collected quantities of all commodities), but to prevent endogeneity and minimize the effect of outliers (mainly due to measurement errors) the regional median prices are used. The system they eventually ran is the following:

$$w_i = \beta_{0i} + \beta_{1i}\ln(x^*) + \beta_{2i}[\ln(x^*)]^2 + \sum_j \gamma_{ij} \cdot \ln(p_j) + \delta_i \cdot \ln(n) + \\ + \sum_d \delta_{id} \cdot (n_d / n) + \phi_i \cdot z + \varepsilon_i \quad (1)$$

where:

x^* is total expenditure;

w_i is the share of item i in total expenditure;

p_j is price of good j;

n is the household size;

n_{id} is number of household members in each of eight age groups; and

z is a vector of household characteristics: head and spouse's education, head's gender dummy and existence or not of spouse.

Regression estimates provide not only expenditure[1] elasticities (a proxy for Engel elasticities) but also price elasticities.

Alves, Disch and Evenson (1982) estimated demand for food by running share equations in logarithmic prices (and including income and squared income) within a SUR model. Prices also came from Endef itself and were also averaged on a regional basis. Both price and income elasticities were reported, and they appear sensible. The authors note that results were better for a subset of nine geographical regions. Estimates for non-food goods and services were harmed by lack of independent price data and a consequent untested assumption of separability from food.

3 - THE MODEL

3.1 - AI Demand System

The model used in the estimation is based on Almost Ideal Demand System (Aids), proposed by Deaton and Muellbauer (1980), which allows a flexible approximation to general preference structure. Aids specifies the log expenditure function as:

$$\ln e(p,v) = \alpha_0 + \sum_i \alpha_i \ln p_i + \frac{1}{2}\sum_i \sum_j \gamma_{ij} \ln p_i \ln p_j + v\beta_0 \prod_i p_i^{\beta_i} \qquad (2)$$

where p is the price vector, v is utility, p_i is the i-th price, and α_i, β_i, γ_{ij} are the parameters. Note that one can assign arbitrary values to utility v. Deaton and Muellbauer suggested assigning zero to utility at subsistence, and one to utility at the bliss point. Then we can interpret α_0 as the log expenditure at subsistence level when all the prices are normalized at one.

[1] The authors instrumented total expenditure with non-labor income.

The linear homogeneity of the expenditure function with respect to the price vector requires the following constraints:

$$\sum_i \alpha_i = 1,\ \sum_i \beta_i = 0,\ \sum_i \gamma_{ij} = \sum_j \gamma_{ij} = 0 \tag{3}$$

We denote the number of commodity groups by M. So the summation in (3) runs through one to M. The share of the ith expenditure group, w_i, is given by:

$$w_i = \alpha_i + \sum_j \gamma_{ij} \ln p_j + \beta_i \ln(Y/P)\,,\ \ i = 1,.., M \tag{4}$$

where Y is the total expenditure, P is the cost of living index given by:

$$\ln P = \alpha_0 + \sum_i \alpha_i \ln p_i + \frac{1}{2}\sum_i \sum_j \gamma_{ij} \ln p_i \ln p_j \tag{5}$$

which is a non-linear function of prices (p_j's). From (2) we can see that ln P is the log of the income required to attain subsistence utility.

The Hicks substitution matrix is given by:

$$S = [S_{ij}] = [\{\gamma_{ij} + \beta_i \beta_j \ln(Y/P) - w_i \delta_{ij} + w_i w_j\} Y/(p_i p_j)] \tag{6}$$

where, δ_{ij} is Kronecker's delta (δ_{ij} = 1 if $i = j$, δ_{ij} = 0 if not). Note that symmetry of the substitution matrix implies symmetry of γ_{ij} ($\gamma_{ij} = \gamma_{ji}$). The negative semi-definiteness of the substitution matrix can be examined by calculating the eigenvalues of (6). Also, the expenditure elasticities are given by:

$$\eta_i = 1 + \beta_i / w_i \tag{7}$$

It follows that if β_i is negative the i th group is a necessity, and if β_i is positive it is a luxury.

3.2 - Estimation Procedure

The model allows for taste variation due to demographic factors, such as age of the household head, family size, education etc. We denote these factors by Z. Then, in the context of our data structure, the model eventually used for estimation is written as:

$$w_{imlt} = \alpha_i + \sum_j \gamma_{ij} \ln p_{jlt} + \beta_i \ln(Y_{lmt} / P_{lt}) + \sum_k \omega_{ik} Z_{klmt} + \varepsilon_{ilmt} \tag{8}$$

where, additional subscripts l, m, and t represent, l: region (l = 1,...,11); m: individual or cohort in each region; t : time period; and ε_{ilmt} is the disturbance term.

Note that prices are common within the same region l and time period t, and total expenditure (ln$Y$) and control variables Z (may) vary across l, m, and t.

The disturbance term has a variance component structure. Namely we write ε_{ilmt} as,

$$\varepsilon_{ilmt} = \lambda_{it} + \upsilon_{ilmt} \tag{9}$$

where, λ_{it} is the time specific factor which uniformly affects all the regions in a given year but changes over time, and υ_{ilmt}'s are other white noise random factors.

It is well known that when time effects are correlated with the explanatory variables, usual OLS and GLS estimators will be biased. If that is the case, we should correct for the bias by introducing time specific dummies, which is the so-called fixed effect specification. We employ this specification, so that the covariance matrix of disturbance terms will have the following structure:

$$\begin{gathered} \text{Cov}(\varepsilon_{ilmt}, \varepsilon_{i'l'm't'}) = \sigma_{ii'}, \text{ if } l = l', \; m = m', \; t = t' \\ \text{Cov}(\varepsilon_{ilmt}, \varepsilon_{i'l'm't'}) = 0, \text{ if not.} \end{gathered} \tag{10}$$

The resulting system in (8) and (9) is nonlinear in parameters, with fixed time effects. Although it is a common practice to estimate the system by replacing ln P by the Stone's index, ln $P^* = \Sigma_i w_i \ln p_i$, and apply OLS, we estimate the system by fully non-linear maximum likelihood.[2]

4 - DATA

The data sources for expenditures are POF surveys run in 1986/87 and 1997/98, which collected households' expenditure on consumption goods/services from 11 metropolitan areas.[3] These surveys were undertaken after two major stabilization plans took place in Brazil: the Cruzado Plan (February/1986) and the Real Plan (July/1994) in order to update the weighting structure of the National Consumer Price Index System (see Appendix). They were supposed to reflect, therefore, household consumption behaviors in low inflation environments, but this applies to the last POF only: the Cruzado Plan and four other stabilization plans thereafter failed; inflation rates continued escalating until the successful Real Plan was launched. In fact, the average inflation rate during the collection period of the first POF (1987/88) was 11.70% a month, as opposed to 1.68% during the second POF

[2] See Buse (1994 and 1998), Alston, Forter and Green (1994), and Pashardes (1993) for discussion of bias caused by this approximation. Asano (1997) is one of the few studies which estimated an AI demand system as a nonlinear system by maximum likelihood.

[3] The metropolitan areas are: 1. Rio de Janeiro; 2. Porto Alegre; 3. Belo Horizonte; 4. Recife; 5. São Paulo; 7. Belém; 8. Fortaleza; 9. Salvador; 10. Curitiba. In addition, 6. Brasília-DF and 11. the municipality of Goiânia are also surveyed. For the sake of simplicity, henceforth we will call all of them metropolitan areas.

(1995/96). The first POF had started in early 1987, but was extended six months so as not to cover periods when supply shortages and black market premiums had arisen due to price freezes; for the period eventually utilized the first six months were discarded.

Family level monthly expenditures on various detailed commodity groups are available in the original POF survey. We aggregate them into seven broad categories: 1. food; 2. housing; 3. furniture and appliances; 4. clothing; 5. transportation and communication; 6. health and personal care; and 7. personal expenses, education and reading. Corresponding price indexes were constructed for the seven categories as described on the Appendix. The price indexes allow for comparisons both across-time and across-region.

For expenditure data, after sorting families by their per capita expenditure, we focus on a more aggregate level. Sample families were assigned to 20 per capita expenditure cohorts (5 percent quantiles) for each region and each time period, and average expenditures within cohorts were estimated. This grouping will keep estimation in a manageable order. The resulting sample size is 440 (two years, 11 regions, and 20 cohorts per region-year) which provides a sample large enough to estimate the demand system of seven expenditure categories, with high accuracy.

4.1 - Sub-Sample Used for Estimation

The total numbers of observations in the original survey, were 12,568 in 1987, and 14,551 in 1996 — after selecting families with income between one and 40 minimum wages only, so as to be consistent with the price index's target population. Our estimation is based on a sub-sample of families, which satisfy the following conditions:

1. male head;
2. married head;
3. family size less than eight (inclusive);
4. household head is older than 18 (inclusive) and younger than 60 (inclusive); and
5. spouse is older than 16 (inclusive) and younger than 60 (inclusive).

Some of the observations are dropped from the lower and higher ends of income, total expenditure, and per capita expenditure distributions. Also observations with implausible values and those with missing values are excluded. The resulting sample sizes are 6,874 in 1987, and 7,427 in 1996. Table 1 shows the number of samples dropped by screening. Table 2 shows the number of original samples retained for estimation for each region/year.

Table 1
Selection of Sample Households (Exclusion Criteria)

Year	1987	1996
Total Expenditure Lowest 1%	125	145
Total Expenditure Highest 1%	126	146
Per Capita Expenditure Lowest 2%	248	291
Per Capita Expenditure Highest 5%	629	728
Food Share Less than 2%	97	544
Transportation Negative	157	188
Income Lowest 1%	125	143
Income Highest 1%	126	146
Head Female	2,776	3,776
Spouse not Female (Including Single)	3,587	4,831
Education of Head Missing	5	28
Head Age > 60	1,710	2,479
Head Age <18	14	22
Spouse Age >60	580	791
Spouse Age<16 (Including no Spouse)	3,428	4,564
Family Size > 8	650	344
Single Household	643	1,068
Survivor	6,874	7,427
Original Observations	12,568	14,551

Table 2
Number of Observations, by Region

Region	Year	
	1987	1996
1. Rio de Janeiro	631	736
2. Porto Alegre	567	608
3. Belo Horizonte	551	696
4. Recife	664	843
5. São Paulo	816	587
6. Brasília	379	414
7. Belém	487	651
8. Fortaleza	816	931
9. Salvador	576	691
10. Curitiba	742	589
11. Goiânia	645	681

4.2 - Aggregation

For each region we created aggregated shares and total expenditures for samples classified by 20^{th} quantile values. Aggregation proceeds as follows: first, the original observations are sorted by per capita expenditure. Then the sample weights and family size are used in simulating the distribution of the population per capita expenditures in the given region/year. Then, the resulting distribution is divided into 20 equal sized cohorts classified by magnitude of per capita expenditure. The average shares and total expenditures are calculated for each cohort.

5 - RESULTS

We estimated the parameters of the system by maximum likelihood. Although all the parameters of the system are estimable, we found that the likelihood is very flat with respect to changes in α_0, which is the subsistence income when prices are normalized to one. To bypass this problem we fixed the value of α_0 at 5.5, which corresponds to an annual expenditure amounting to R$ 300 per head in 1996 (at September 1996 prices). Hence the statistical inference in this section is conditional on this assumption. The effects of changes in α_0 on price and expenditure elasticities, however, are of negligible order because of compensating changes in $\hat{\alpha}_i$'s.

5.1 - Coefficients

Table 3 displays the estimates of the parameters from the restricted model in which homogeneity and symmetry constraints are imposed. In addition to prices and total expenditure, we include four explanatory variables to capture demographic factors that shift intercepts of the share equations. They are: a dummy for the North and Northeast regions (the poorest regions in Brazil), the age of the household head (husband), the schooling years of the household head, and the family size.

Table 3
Estimated Coefficients

	Food	Hous	Furn	Clth	Tran	Hlth	Pers Exp
Cnst	0.6768	0.2278	0.1193	0.0989	–0.0922	0.0215	–0.0520
(*t*-val)	(11.51)	(5.44)	(3.57)	(2.94)	(–1.38)	(0.74)	(–1.07)
Hous	–0.0123						
(*t*-val)	(–0.84)						
Furn	–0.0046	0.0008					
(*t*-val)	(–0.37)	(0.12)					
Clth	–0.0153	0.0052	–0.0027				
(*t*-val)	(–0.69)	(0.52)	(–0.37)				
Tran	0.0393	–0.0053	–0.0123	0.0162			
(*t*-val)	(1.82)	(–0.44)	(–1.23)	(1.30)			
Hlth	–0.0149	0.0163	–0.0006	–0.0060	0.0001		
(*t*-val)	(–0.75)	(1.92)	(–0.09)	(–0.23)	(0.01)		
Pers Exp	–0.0040	–0.0135	0.0049	0.0140	–0.0008	0.0120	
(*t*-val)	(–0.24)	(–1.59)	(0.75)	(1.25)	(–0.07)	(1.21)	
lnY	–0.0894	–0.0258	0.0205	0.0115	0.0499	0.0077	0.0255
(*t*-val)	(–9.38)	(–3.88)	(3.85)	(2.08)	(4.71)	(1.64)	(3.35)
D_NE	0.0593	–0.0362	0.0046	0.0020	–0.0199	–0.0045	–0.0054
(*t*-val)	(9.75)	(–8.42)	(1.34)	(0.58)	(–2.91)	(–1.51)	(–1.08)
Age_h	0.0020	–0.0005	–0.0014	–0.0008	0.0006	0.0002	–0.0001
(*t*-val)	(1.50)	(–0.57)	(–1.82)	(–1.03)	(0.39)	(0.37)	(–0.14)
Edu_h	–0.0060	0.0033	–0.0093	–0.0030	0.0058	0.0012	0.0079
(*t*-val)	(–1.70)	(1.32)	(–4.67)	(–1.44)	(1.48)	(0.68)	(2.81)
F_siz	–0.0177	–0.0003	–0.0000	0.0032	0.0018	0.0043	0.0088
(*t*-val)	(–2.40)	(–0.06)	(–0.01)	(0.75)	(0.22)	(1.17)	(1.46)

The dummy variable for metropolitan areas in the North and Northeast regions[4] was introduced after the first runs, when we noticed that food shares in these areas were systematically underestimated after doing all the other controls. Indeed the coefficients of this dummy in the demand equations for food, housing and transportation proved statistically significant. They are negative for housing and transportation, probably capturing omitted variables mentioned in the Appendix: available prices do not distinguish apartment sizes and commuting distance, for example (these areas are smaller, so the expenses on public transportation, parking, etc. are lower and the price per square meter of housing are much lower than in bigger areas. The positive coefficient for food is more difficult to explain; several hypotheses come up to mind: measurement error in consumption versus income (poorer families might spend — or report — more on food during collection out of shame for their ordinarily low levels); measurement error on prices; existence of self-consumption in rural enclaves; other omitted variables (differences in taste, absence of leisure options, etc.).

Age of household head did not prove significant, but schooling years did for furnishings and personal expenses; the former is negative (maybe because the less educated spend more on durable goods such as appliances at the time of stabilization plans) and the latter positive (families with more educated heads tend to invest more in education and reading, for example).

Last but not least, the coefficient for family size turns out to be significant and meaningful: per capita expenditures on food are lower when the family size is larger; this is an evidence of scale economies in the household meal production function.

5.2 - Elasticities

It is rather hard to evaluate results based upon the original parameters. Thus, we examine results based upon the estimates of the price and expenditure elasticities. As shown in Section 2, these elasticities are highly nonlinear in parameters, and depend on the values of prices and total expenditure at which they are evaluated. For both 1987/88 and 1995/96 surveys, elasticities are evaluated at sample mean values of prices and total expenditure. The standard errors for elasticities are obtained by applying Rao's (1973) δ-method.[5] Table 4 presents estimates of elasticities for 1986/87, and Table 5 shows those for 1995/96. Fitted shares are plotted along with observed ones on Figures 1 through 3, for São Paulo, Belo Horizonte and Fortaleza.

[4] Namely: 7. Belém (North); 4. Recife, 8, Fortaleza, 9. Salvador (Northeast).

[5] See Deaton (1997, p.128-129) for a concise description of the δ-method.

Table 4

Expenditure and Price Elasticities — 1987

Eigenvalues							
–0.239	–0.170	–0.128	–0.099	–0.081	–0.073	0.000	

Expd.	Food	Hous	Furn	Clth	Tran	Hlth	Pers Exp
Shares	0.356	0.102	0.093	0.146	0.118	0.089	0.095
Elas'ty	0.749	0.746	1.220	1.079	1.423	1.087	1.267
(s.e.)	(0.033)	(0.080)	(0.052)	(0.037)	(0.077)	(0.053)	(0.073)

Price Elasticities

	Food	Hous	Furn	Clth	Tran	Hlth	Pers Exp
Food	-0.531	0.090	0.062	0.093	0.183	0.040	0.061
(*t*-val)	(-4.901)	-2.344	-1.850	-1.697	-3.142	(0.828)	-1.431
Hous	0.317	-0.787	0.082	0.187	0.020	0.241	-0.060
(*t*-val)	-2.101	(-8.066)	-1.255	-1.604	(0.170)	-2.630	(-0.605)
Furn	0.237	0.090	-0.735	0.126	0.025	0.089	0.168
(*t*-val)	-1.903	-1.382	(-8.643)	-1.546	(0.251)	-1.263	-2.570
Clth	0.227	0.130	0.080	-0.929	0.243	0.050	0.198
(*t*-val)	-1.721	-2.212	-1.704	(-5.081)	-2.947	(0.310)	-3.339
Tran	0.554	0.018	0.020	0.301	-1.121	0.101	0.127
(*t*-val)	-3.175	(0.172)	(0.250)	-2.662	(-5.363)	-1.127	-1.255
Hlth	0.161	0.277	0.093	0.083	0.134	-0.986	0.238
(*t*-val)	(0.768)	-3.596	-1.244	(0.281)	-1.076	(-3.132)	-2.510
Pers Exp	0.229	-0.064	0.164	0.304	0.157	0.221	-1.012
(*t*-val)	-1.332	(-0.598)	-2.197	-2.066	-1.155	-1.976	(-7.492)

Table 5

Expenditure and Price Elasticities — 1996

Eigenvalues							
-0.262	-0.181	-0.128	-0.106	-0.071	-0.050	0.000	

Expd.	Food	Hous	Furn	Clth	Tran	Hlth	Pers Exp
Shares	0.311	0.142	0.065	0.063	0.185	0.080	0.156
Elas'ty	0.712	0.818	1.316	1.184	1.270	1.097	1.164
(s.e.)	(0.030)	(0.043)	(0.104)	(0.122)	(0.069)	(0.068)	(0.055)

Price Elasticities

	Food	Hous	Furn	Clth	Tran	Hlth	Misc
Food	-0.558	0.129	0.029	0.002	0.259	0.024	0.116
(*t*-val)	(-4.284)	-2.622	(0.723)	(0.019)	-3.841	(0.322)	-2.131
Hous	0.283	-0.778	0.057	0.092	0.114	0.189	0.043
(*t*-val)	-3.131	(-11.539)	-1.428	-1.682	-1.496	-3.715	(0.877)
Furn	0.137	0.124	-0.688	0.034	0.053	0.080	0.260
(*t*-val)	(0.726)	-1.270	(-6.332)	(0.294)	(0.335)	(0.783)	-2.369
Clth	0.008	0.208	0.036	-1.112	0.475	-0.010	0.396
(*t*-val)	(0.019)	-1.146	(0.293)	(-1.932)	-2.150	(-0.026)	-1.546
Tran	0.435	0.087	0.019	0.162	-0.967	0.087	0.176
(*t*-val)	-3.819	-1.368	(0.336)	-2.361	(-6.747)	-1.380	-2.506
Hlth	0.092	0.337	0.065	-0.008	0.203	-1.004	0.315
(*t*-val)	(0.343)	-2.423	(0.830)	(-0.026)	-1.556	(-3.581)	-2.088
Pers Exp	0.232	0.040	0.109	0.160	0.209	0.161	-0.910
(*t*-val)	-2.445	(0.872)	-2.841	-2.996	-2.825	-3.047	(-13.653)

Figure 1
Observed versus Fitted Expenditure Shares per Group: Belo Horizonte

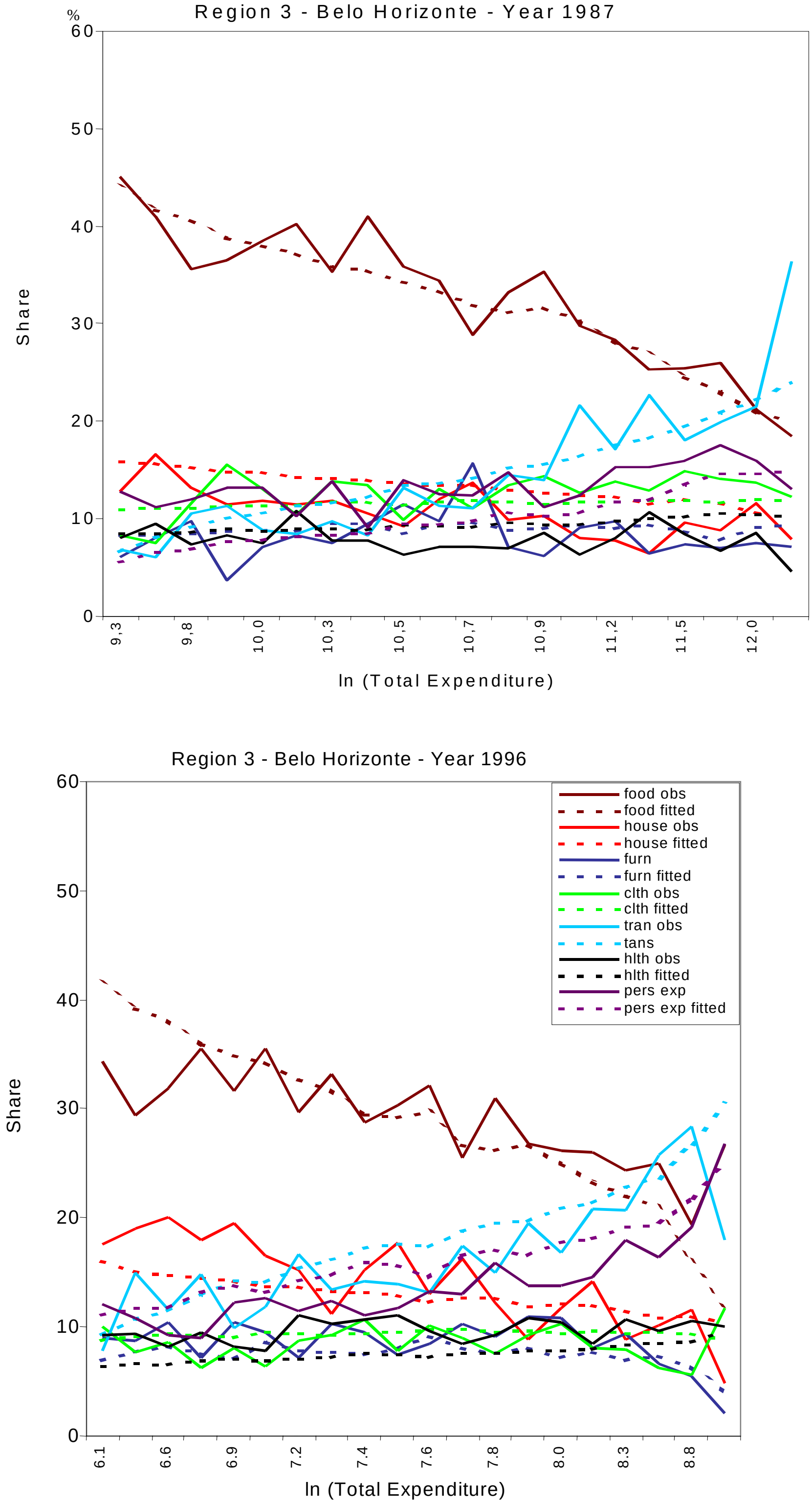
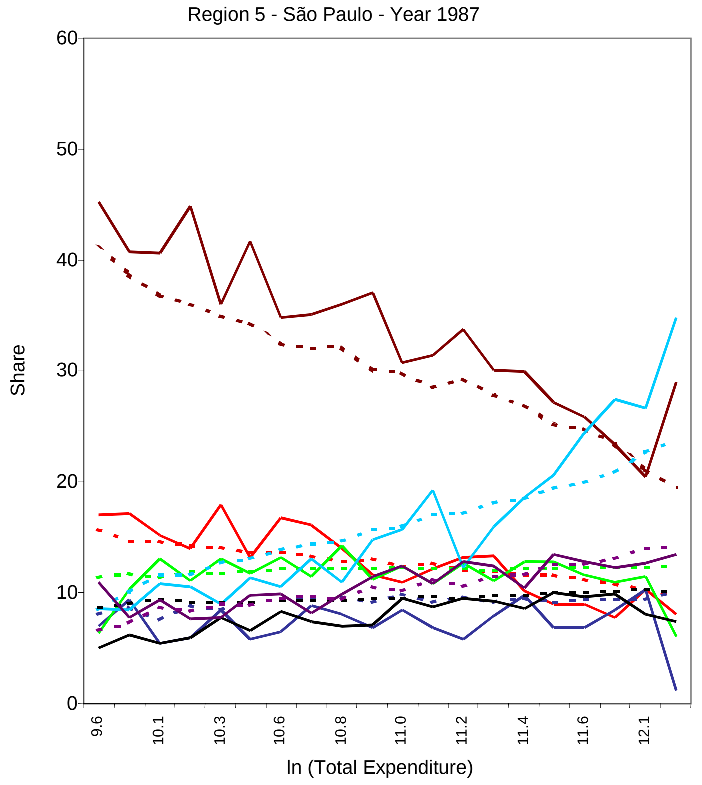


Figure 2

Observed versus Fitted Expenditure Shares per Group: São Paulo



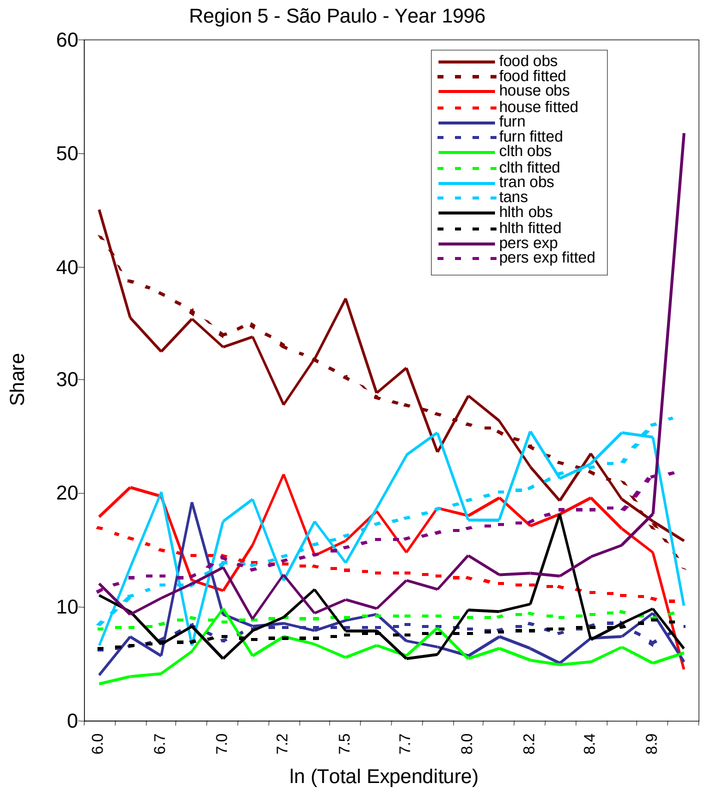
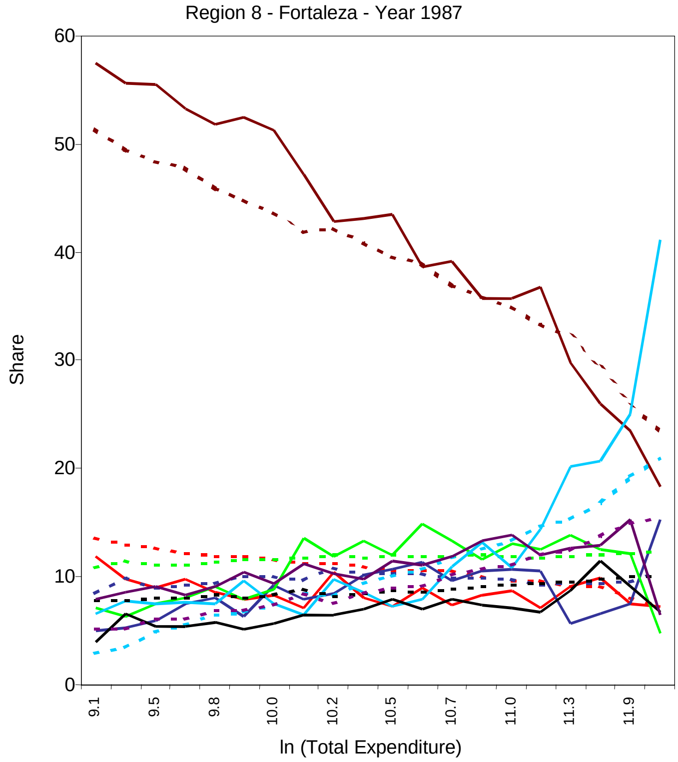
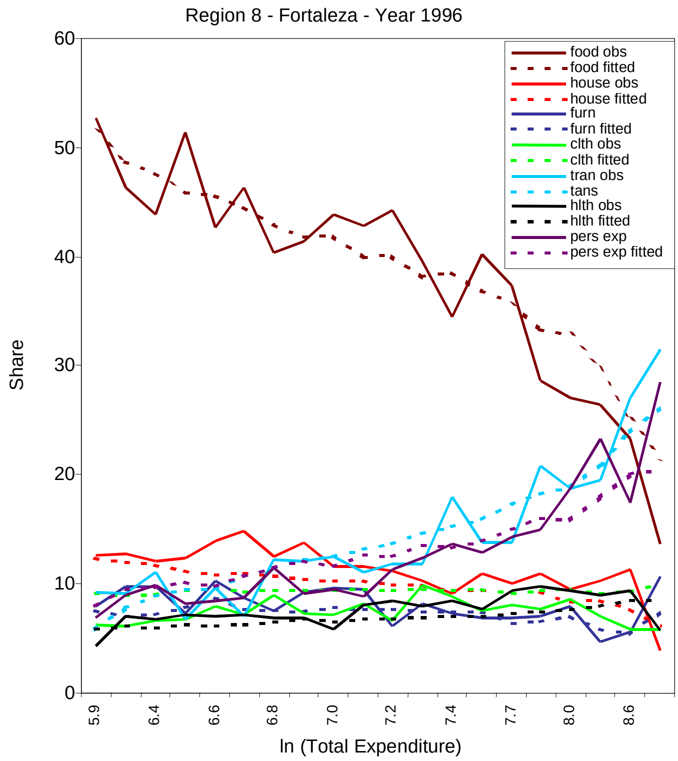


Figure 3
Observed versus Fitted Expenditure Shares per Group: Fortaleza





The first rows of Table 4 and 5 show the eigenvalues of the substitution matrix. All the eigenvalues are negative, as expected theoretically. Thus, negativity (i.e., gross substitution) is supported by our data; this is rare in demand studies. Another way of reading these results is looking at the elasticity matrix: diagonal elements (own-price elasticities) are negative and off-diagonal elements (cross-price elasticities) are positive — the categories are substitutes of each other. The standard errors reported are also remarkably low, showing that our estimates are quite reliable.

Total expenditure elasticities (a proxy for income elasticities) indicate that food and housing are the only necessities in Brazilian utility functions, whereas furnishings, clothing, transportation, health care and personal expenses are found to be luxuries. All the own-price elasticities are significantly negative. Among them, those for food, housing and furnishings are significantly less than one (own-price-inelastic), while clothing, transportation and communication, health care and personal expenses display own-price elasticities around 1. Products whose effect may be having an important influence in rendering elastic the demand for health and personal care are: toilet items and cosmetics; and in rendering elastic demand for transportation are: private motor vehicles.[6]

It is also worth noting that the elasticity estimates varied modestly (and not always the variations were statistically significant) between 1987 and 1996, in spite of the multitude of stabilization attempts during this period. High inflation rates like the ones recorded in the 1980s and early 1990s should give rise to a high degree of price dispersion; consumers would lose track of “fair” relative prices and would be expected to be less price-elastic for this reason. Yet price elasticities are significantly lower in 1996 for Furnishings, Transportation and Communication, and Personal Expenses, while the only significant rise was detected for Clothing (would it be due to income increase?). Movements of income elasticities are more ambiguous and none was significantly different from zero.

Comparing our estimates of expenditure elasticities with the ones obtained in previous estimates (see Table 6) is a challenging task, as we have already mentioned how diverse specifications, samples and expenditure classification are. We find ours to be closer to the ones obtained by Rossi (1983*b*), Rossi and Neves (1987) and Hoffmann (1983,1988) than to Thomas, Strauss and Barbosa (1989). This suggests that sample design and categorization may have higher impacts on estimates than the functional form adopted. In particular, we should highlight the proximity of our estimates for health and clothing. Transportation and personal expenses are plausible convex combinations of the subgroups constructed by them, while furnishings is a bit lower. As regards the categories we found as necessities, they are much farther from most of the previous estimates: our estimates for food indicate a much more elastic demand, whereas the opposite occurs to housing — there we find a much less elastic demand than most of the other authors (except Alves, Disch and Evenson, whose figures are not too far from ours).

As noted in Section 2, price elasticities had been reported by Thomas, Strauss and Barbosa (1989), Alves, Disch and Evenson (1982), and Simões and Brandt (1981)

[6] In fact, Andrade and Lisboa’s (2001) tabulations of a recent supplement on health expenditures run by IBGE along with PNAD’s (an annual household survey) 1998 edition suggest that expenditure shares on health care are higher for the lower-income deciles.

Table 6

Expenditure Elasticity Estimates in Previous Studies

Author	Medeiros (1978)	Rossi (1982)	Rossi (1983)	Hoffmann (1983 and 1988)	Cipriano and Brandt (1983)	Rossi (1987)
Methodology	Box-Cox	Lorenz curve	Lorenz curve	Piecewise log-linear	Logit	Logit
Data source	POF-USP 1971/72	Endef 1974/75	Endef 1974/75	Endef 1974/75	Endef 1974/75	Endef 1974/75
Sample	São Paulo (city)	Rio de Janeiro (city)	Brazil (metropolitan areas)	Rio de Janeiro (city)	Brazil	Brazil
Elasticities Estimated	Expenditure	Expenditure (at mean point)	Expenditure (at mean point)	Expenditure (mean)	Expenditure (mean)	Expenditure (at mean point)
Food	Btw. 0.21 and 0.77	0.20	0.54	0.57		0.55
Grains		0.16	0.19	0.25	0.068 (cereals);	
Tubercles		0.43	0.24	0.45	0.028	
Vegetables		0.53	0.63	0.54	0.398 (potherbs); –0.110 (legumes)	
Fruits		0.91	0.94	0.87	-0.325	
Meat and Fish		0.62	0.60	0.63	0.469	
Eggs and Dairy		0.70	0.71	0.68	0.528	
Beverages		0.50	0.53 (bev. and others)	0.52	0.271	
Food Away from Home		0.85	0.83	0.79		
Sugar, Oils, etc.		0.18	0.33 (sugar); 0.31 (oils)	0.31	0.074 (sugar); 0.103 (oils and fats)	
Other Food			-0.01 (legumes)		0.840	
Clothing		1.21	1.15	1.10		1.11
Housing		1.05	1.07	1.12		1.26
Health		1.16	1.14	1.13		1.18 (hygiene)
Education	btw. 1.54 and 1.64	1.70	1.60	1.38		1.71
Recreation		1.46	1.49	1.22		1.60
Tobacco		0.45	0.51	0.46	0.505	
Public Transportation		0.32	0.02	0.57		1.60 (transportation)
Own Vehicle		1.20	1.97	1.49		
Fuel and Transportation			1.68 (long-distance trips)			
Miscelaneous		1.46	1.40	1.44		1.31

(continue)

(*continued)*

Author	Alves, Disch and Evenson (1982)	Simões and Brandt (1983)	Thomas, Strauss and Barbosa (1989)	Our estimates (1996)
Methodology	Serp	Eles	Generalized Aids	Aids
Data Source	Endef 1974/75	Endef 1974/75	Endef 1974/75	POF 1987/88 and 1995/96
Sample	9 regions	Brazil	Brazil	Brazil (metropolitan areas)
Elasticities Estimated			Expenditure (median) and prices	Expenditure (at mean point)
Food				0.726
Grains	–0.0096 (rice); 0.2927 (cereals)	0.1588 (cereals);	0.58 (rice); 0.881 (wheat); –0.49 (corn); – 0.565 (manioc); 0.282 (beans)	
Tubercles	0.3251 (root crops)	0.100	0.705	
Vegetables	0.434	0.4425 (potherbs); 0.014 (legumes)	0.403	
Fruits	0.779	0.630	1.023	
Meat and Fish	0.938 (beef); –0.023 (pork); 0.494 (other meats); 0.453 (fish)	0.429	1.025 (meat); 0.473 (fish)	
Eggs and Dairy	0.3665 (eggs); 0.755 (dairy)	0.513	1.045 (milk); 1.065 (eggs and non-milk dairy)	
Beverages	0.2062 (coffee/tea)	0.357		
Food Away from Home				
Sugar, Oils, etc.	0.1078 (oils); 0.908 (sweets)	0.202 (sweets); 0.216 (oils and fats)	0.266 (sugar); 0.635 (oils and fats)	
Other Food	-0.0095 (legumes)		0.570	
Clothing			1.316	1.130
Housing	0.898		1.009	0.791
Health				1.103
Education	1.09 (education, recreation, culture)			1.17 (personal expenses, education and reading)
Recreation				
Tobacco		0.43		
Public Transportation	1.538 (transportation)			1.296 (transportation)
Own Vehicle				
Fuel and Transportation			1.339	
Miscelaneous	1.054		1.426 (household goods); 1.644 (other goods)	1.304 (furnishings)

Table 7

Price Elasticity Estimates in Previous Studies

	Simões and Brandt (1981)	Alves, Disch and Evenson (1982)	Thomas, Strauss and Barbosa (1989)	Our estimates (here we report 1996)
Methodology	Eles	Serp	Generalized Aids	Aids
Data Source	Endef 1974/75	Endef 1974/75	Endef 1974/75	POF 1987/88 and 1995/96
Sample	Brazil	Nine regions*	Brazil	Brazil (metropolitan areas)
Food				-0.549
Grains	–0.155 (cereals);	–0.96 (rice); 0.66 (root crops); +0.64 (cereals)	–3.59 (rice); –1.971 (wheat); –0.114 (corn); +0.280 (manioc);–1.679 (beans)	
Tubercles	–0.087		–1.954	
Vegetables	–0.007 (legumes); –0.440 (potherbs)	–0.49	–0.786	
Fruits	–0.624	–0.74	–0.882	
Meat and Fish	–0.435	–0.74 (pork); –1.34 (beef); –1.25 (other meats); –1.11(fish)	–0.421 (meat); –2.561 (fish)	
Eggs and Dairy	–0.511	–0.77 (dairy); +0.46 (eggs)	–3.371 (milk); –2.470 (eggs and non–milk dairy)	
Beverages	–0.353	–0.30 (coffee/tea)		
Food Away from Home		–0.1538		
Sugar, Oils, etc.	–0.195 (sweets);–0.211	–0.38 (sugar); –0.95 (oils)	–0.004 (sugar); +2.732 (oils and fats)	
Other Food		–0.52 (legumes)	–1.440	
Clothing			–0.583	–1.035
Housing			–0.445	–0.785
Health				–1.014
Education				–0.917 (personal expenses,
Recreation				education and reading)
Tobacco	–0.442			
Public Transportation				–0.998 (transportation)
Own Vehicle				
Fuel and Transportation			–1.096	
Miscelaneous			0.033 (household goods); –0.021 (other goods)	–0.695 (furnishings)

**A missing Appendix should contain a list of the regions selected.*

only. However, only the estimates by Thomas, Strauss and Barbosa cover all consumption categories, and they are very different from ours[7] (see Table 7): we found clothing and housing to be more price-elastic; food items cannot be compared, due to our aggregation, but transportation figures are very close to each other. Again, differences in sampling or commodity grouping may be blamed for, otherwise either the different Aids specification or a change along time may be the causes for such differences. The other estimates were reported for agricultural products only, and in a much more disaggregate level, thus rendering comparisons unfeasible.

Despite the relative stability of our mean elasticities, the mean shares of the categories display some noteworthy changes between the two dates. These changes are due to a combination of income increase[8] and changes in relative prices and demographic variables (see Table 8).

Personal expenses (from 0.095 to 0.156) and Transportation and Communication (from 0.118 to 0.185) increased their shares most, as their expenditure elasticities are quite high and prices increased moderately in the period, as compared to other categories. Housing shares increased moderately, despite the very low expenditure elasticity, and this is certainly due to the fact that the category recorded the highest inflation among all the categories, combined with a very low own-price-elasticity. Almost the opposite occurred in the Furnishings category: expenditure elasticity is high, but prices increased the least and own-price elasticity is also very low, driving shares down. Clothing shares decreased despite the high income-elasticity because their inflation was the second lowest, and the own-price elasticity is high. Health and personal care decreased moderately because the effect of their high inflation combined with a high own-price elasticity dominated the high expenditure elasticity effect. Finally, a low inflation combined with low own-price elasticity and a low expenditure elasticity drove the food share down.

Table 8

Accumulated Variation of Income and Prices by Expenditure Group According to IPCA and RPDI: October 1987-September 1996 (%)

Expenditure Group	RPDI	IPCA
0.General	7.2E+09	12.3E+09
1.Food	5.3E+09	9.9E+09
2.Housing	17.8E+09	21.0E+09
3.Furnishings	3.2E+09	5.6E+09
4.Clothing	4.3E+09	4.5E+09
5.Transportation and Communication	8.1E+09	12.9E+09
6.Health Care	14.0E+09	21.2E+09
7.Personal Expenses, Education and Reading	10.1E+09	21.6E+09
Real Per Capita Income	51.97	–11.45

[7] It is worth noting that the order of size of the cross-price elasticities is, through a visual inspection, not too different from each other.

[8] Using the national general index calculated according to our Regional Price Difference Index (RPDI) methodology, annual per capita real income (Brazilian average) increased from BRL 1,730 to BRL 2,639 at September 1996 prices (a 51.97% accumulated variation, or 4.76% per annum rate) in our sample.

6 - CONCLUSION

This study estimated a complete Brazilian consumer demand system from two family budget surveys (POF/IBGE), conducted in 1986/87 and 1995/96. To the best of our knowledge, it is the first study of this kind based on those data. IBGE's stratified sampling scheme used for the survey provides large variations in total expenditures. We look at seven broad categories of consumption goods and services, which cover all consumption expenditures. To enable the estimation of the system, regional price difference indexes are created. By exploiting inter-temporal and inter-regional differences in prices, we estimated the parameters for the price variables. The result show a striking conformity to micro-economic theory. The estimated price and income (expenditure) elasticities are close to what economic common sense predict, and their standard errors are fairly small, as opposed to earlier estimates from other studies (based on an earlier and more comprehensive survey). Negativity of the Hicks substitution matrix is supported. To our knowledge this is a rare result in empirical studies of demand systems which use flexible functional forms. Also, we found that, despite high inflation in the sample period, Brazilian consumers' preferences, captured in terms of elasticities, showed a quite stable pattern.

The estimated system will serve as a solid micro-economic basis for evaluating various policy related issues, such as the impact of commodity taxation on different goods/services, effects of subsidy to needy families, etc., and for comparing living standards across regions, and across income classes.

APPENDIX

Constructing a Regional Price Difference Index for Brazil

To this date, the only source of cost-of-living difference estimates across regions in Brazil had been the legal minimum Staple Food Basket (whose composition was enacted by a Federal law in 1938), collected by Dieese, a research bureau supported by labor unions, in 16 state capitals. The need of a more comprehensive and up to date household basket of goods and services for estimating a system of consumption demand equations led us then to undertake the construction of a new set of Regional Price Difference Indexes (RPDI) based on a combination of the existing price variations at subitem level, released monthly by Brazilian Geographical and Statistical Institute (IBGE), with special tabulations of nominal prices provided by the same source. Here we describe briefly the main features of the available price data from IBGE and the construction of the RPDI.

A.1 - The National Consumer Price Index System (SNIPC)

The National Consumer Price Index System (SNIPC) was created in 1979[9] and the first national index was released in 1980, based on nine metropolitan areas and the Federal capital, Brasília (in 1989 another city, Goiânia, was added to the sample). It comprises two main sets of indices: the National Consumer Price Index (INPC) and the Broadened Consumer Price Index (IPCA). The sources of weighting structure, the target populations are displayed on Table A.1

Table A.1
SNIPC Structure

Period	January 1979/ May 1989	June 1989/ December 1990	January 1991/ December 1993	January 1994/ July 1999	August 1999/ Present
Number of Subitems	486	461	370	369	512
Source of Weights	Endef 1974/75	POF 1987/88	POF 1987/88	POF 1987/88	POF 1995/96
Target Populations	INPC: 1-5 m.w. IPCA: 1-30 m.w.	INPC: 1-5 m.w. IPCA: 1-40 m.w.	INPC: 1-5 m.w. IPCA: 1-40 m.w.	INPC: 1-5 m.w. IPCA: 1-40 m.w.	INPC: 1-5 m.w. IPCA: 1-40 m.w.
Main Modification		POF introduced	Number of subitems reduced to simplify data collection	Within-subitem average changes from arithmetical to geometrical (17 sub-items excluded and 16 included), so as to abide to circularity property; weights had to be revised back to initial values.	New POF was introduced; two groups and several subitems were subdivided or merged; some codes were renumbered
Our Naming	IPC-1	IPC-2[a]	IPC-2b	IPC-2c	IPC-3

Notes: Endef: National Family Expenditure Study covered 53,000 households in the whole Brazilian urban area and in the rural area of the Northeastern, Southeastern and Southern regions. POFs covered the nine metropolitan areas and Brasilia and Goiânia cities only.

---

[9] Before that, price indices for Brazil had been computed in as many as 13 state capitals by the Ministry of Labor from 1948 to July 1978.

The indexes are obtained the following way: within a defined subitem a set of products[10] is identified for each metropolitan area. Their monthly price variations are then averaged geometrically and non-weighted to obtain the subitem price variation. The subitem *variations* are then averaged to items, subgroups or groups according to Laspeyres's formula. The weights are the expenditure weights provided by POF on a regional basis. The regional price indices thus obtained are then averaged to form the national index. The population of each region represented by the metropolitan area weights this averaging. The only difference between the calculation of IPCA and INPC relies on the weights used in Laspeyres's formula, which were obtained from averaging target populations of different sizes (the former contains the latter). These weights are continuously updated by price variations. Note that this regional weighting scheme allows the indices to compare prices *along time within a given area, but not across areas*; that is why we had to construct the present Regional Price Difference Index (RPDI).

Throughout the years a number of revisions were introduced into SNIPC regarding the weighting structure and the subitem calculation. They are also summarized by Table A.1 below. Note that revisions were introduced even when the source survey (POF) was the same.

Some subitems require specific treatments. For example, clubs, schools, daycare and other services paid monthly have their dues collected as prices in the previous month. Vehicle and house property tax variations (raised nowadays on a yearly basis) are spread out in a monthly fashion. Subitems belonging to seasonal food items[11] are assigned an annual calendar, whereby some subitems within a given item disappear and others appear each month, while only the whole item's total weight is updated by Laspeyres's formula; the subitem weights must add up to this total weight. For rent, the accumulated variation is used, both for the numerator (variation accumulated from the base period 0 up to time *t*) and the denominator (variation accumulated from the base period 0 up to time *t*-1); the variations are averaged arithmetically and non-weighted. An important caveat is that own-housing rent is *not* imputed. An odd proceeding is the one for domestic servant services: IBGE assumes their price follows the official minimum wage.

As regards public services, two cases exist: *a*) if the subitem is composite of different products, an arithmetical average (weighted by POF shares or revenue shares provided by public utility companies) is used (e.g. local and inter-city buses, postal services, etc.); *b*) if the price or tariff is non-linear [i.e., it depends on consumption level (the usual pattern in public utilities)], case IBGE calculates in each area the mean of the quantity consumed, provided by POF respondents or by the public service suppliers, and calculates the tariff applying to this level, one for

[10] The subitem is the most disaggregate level at which variations are disclosed regularly by IBGE. For example, apple is a subitem of the fruit item, within the food-at-home subgroup, which is a part of food group. However, subitems themselves are made up of products, as defined by IBGE. For example, a 10-door cherry-tree wood wardrobe is a product of the subitem bedroom furniture. Only general and group indexes are calculated .

[11] Namely *a*) tubercles, roots and legumes; *b*) potherbs and vegetables; *c*) fruits.

each target population (INPC and IPCA) — e.g. water and sewage fees; electricity; residential phone; taxi rides.

A.2 - The Regional Price Difference Index (RPDI)

A.2.1. - Definitions and foreign experience

An RPDI, as any cost-of-living index, is nothing but the ratio of the minimum expenditures required to attain a particular indifference curve under two price regimes [Pollak (1989, p.6)]. "Strictly speaking, the cost-of-living index depends only on the comparison prices, the reference prices, and the base indifference curve" (*ibidem*, p.7).

The two price regimes refer naturally to the price vectors in effect in two different states of world, for which the index compares the expenditures of a given "individual" (a representative consumer, described by an indifference curve map). Typically these states of the world refer to different periods and/or different regions. So far, SNIPC, as it has been conceived, only compares pairs of price regimes of different periods in a same area. Thus, each area has its own representative consumer (the area's respective IPCA and INPC average target population), not to be compared to another area.

Making inter-regional comparisons of cost-of-living requires choosing whose preference ordering to use. It follows that we have 11 possible sets of 10 indices each, each of them using a different metropolitan area's preference ordering (basket) as base. Alternatively, one may use the national average basket of goods and services as base.

We are aware thus far of two measures of regional price differences in other countries. One is Japan's Regional Difference Index of Consumer Prices (RDICP) and the other is U.S. Inter-Regional Cost-of Living Index. The former is calculated by the Statistics Bureau of the Management and Coordination Agency (MCA), and compares yearly average prices across 10 districts and 47 major cities or cities where prefectural governments are seated, but takes no account of inflation, so a comparison of these indexes along time requires chaining them with CPI. RDICP has been calculated since 1947. Prices are a subset of the data collected monthly through Retail Price Survey, also undertaken by MCA, over 34,000 establishments and 24,000 households in 167 municipalities for CPI. Some items out of total 580 in CPI are excluded from RDICP because their specifications are not uniform throughout the country. On the other hand, RDICP's weights are always provided by the current year's Family Income and Expenditure Survey (Fies), conducted by MCA every month, as opposed to CPI's weighting structure, which is updated only once every five years by Fies. Fies's target population comprises all Japanese urban and rural households, excluding those mainly engaged in agriculture, forestry and fishery and the one-person households. About 8,000 households are sampled from this population. MCA calculates two series of the index with Laspeyres's formula, one for each base preference for weights: Japan's average and Tokyo area's.

U.S. Inter-Regional Cost-of Living Index is conducted quarterly by American Chamber of Commerce Researchers Association (Accra) an independent professional association of community and economic development researchers. It is a singular index in the sense that data collection is undertaken by voluntary members who follow Accra's standard guidelines either on a quarterly or biannual basis. Important consequences of this arrangement are: *a*) some subjectivity in selecting outlets — no Point-of-Purchase Survey is undertaken; *b*) local researchers have also some freedom of choice in determining how to contact the informers: in person, by mail, e-mail, fax and phone, depending on the item to be collected; *c*) if the local researcher fails to comply with the guidelines, as understood by the regional offices' scrutiny, the city or metropolitan area is excluded from the period's sample, and the average price to be used as base price is therefore calculated on a variable number of areas along time. The target population of the index comprises professional or managerial couples with one child, belonging to the upper quintile of the income distribution. The weighting structure are borrowed from 1992 Current Expenditure Survey, national average. The price set is considerably limited, though: only 60 items. Most of them are quite standardized.

For our own purposes, we use the national average weighting structure observed on POF 1995/96's reference date (September 15, 1996). This criterion is the same used by Japanese RPDI and Accra's Cost-of-Living Index, and has as its main advantage the existence of a base price for each and every subitem used in the already reduced basket.

A.2.2 - Limitations of Brazilian data for calculating RPDI

Brazilian SNIPC's products are defined in each metropolitan area by a survey named Pesquisa de Especificação de Produtos e Serviços (Peps). These products vary from region to region, such that very few subitems are completely comparable across all metropolitan areas. In some cases, products end up classified as different subitems; for example, there are six different bean type and four banana type subitems, none of them present in all areas. This poses a severe challenge for comparing prices across regions.

Ideally, for constructing a Regional Price Difference Index we would need a minimum set of standardized products present in all metropolitan areas with their weights provided by a regular (yearly if possible) household expenditure survey. To circumvent this heterogeneity we had to choose a *core* of comparable products across areas within each subitem. Criteria to choose the products are available from the authors upon request. A caveat is that, since subitem price variations are computed over all products of the subitem, using them to chain RPDI is likely to bias regional comparisons along time. The bias is greater, the farther we move from the date of initial comparison, the less correlated are products excluded with products included in the *core*, and the more products are excluded. The only definite solution would be tracking individual product variations along time or release their nominal prices periodically so that the RPDI could be computed on a regular basis.

What we have so far is a set of nominal prices graciously released by IBGE. Not all prices in this set are contemporaneous, though: food prices had been released earlier by IBGE, for 1987, 1990, 1993, 1995 and 1996. The delivered non-food nominal prices referred to September 1999, the second month of IPC-3 weighting structure. Unfortunately, as IPC-3 has many more subitems than its predecessor, some new subitems (in the index or in some particular region) were missing, because they have delayed to be collected. For this reason we ordered a new set of nominal prices collected in February 2000 in order to fill the gap, and then we deflated them back to the same reference date. In total, 129 food and non-food subitems of the current IPCA weighting structure were used, amounting to 73.07 % of the index. The remaining subitems' weights were redistributed within each item. Details on deflation are provided below.

Other disadvantages of relying on this price data set are apparent: *a*) hedonic adjustment for *housing* (size is lower in smaller metropolitan areas), *clothing, electronic goods, home appliances, computers and automobiles*, etc. is not undertaken; *b*) expenditures on *domestic servants*, which amount to a high share in the personal service subgroup, are treated by IBGE as if no regional difference existed, because it uses the official minimum wage, which is unique in the whole country; now, from outside sources, we know that in smaller metropolitan areas the middle class can still afford a (relatively cheap) monthly maid, whereas in greater metropolitan areas most housemaids are paid on a (much higher) daily basis. The reader should bear these caveats in mind when using the RPDI.

A.3 - Two Deflation Approaches

We calculated RPDI for three instants in time: September 1999, September 1996 and October 1987, respectively the reference dates of the nominal prices and the reference dates of POF 1995/96 and POF 1987/88. All of them are based on September 1999-centered prices (deflated at subitem level) and then deflated back using two alternative approaches.

The first one uses *group subindexes*. Before having access to individual subitem price variations, we experimented deflationing RPDI through group subindexes. The disadvantage is that weights used to calculate group subindexes vary along time. Reference RPDIs (both general and group) were calculated in August 1999 using weights of that time, and the respective subindexes were applied to deflate them back:

$$p_0 = \sum_j w_{j00} \cdot \frac{p_{jr0}}{p_{j00}} \qquad (11)$$

where:

p_0 is the index at reference period $t = 0$;
p_{jr0} is price of good j in metropolitan area r at reference period;
p_{j00} is average price of good j in all metropolitan areas where weight greater than zero is observed in the reference period; and

w_{j00} is the subitem *j* average national weight at reference period $t = 0$, that is, the average share of the subitem expenditure in total (or group or subgroup or item) household expenditure over all metropolitan areas.

$$p_t = p_0 \cdot (1+\pi_{0t}) \qquad (12)$$

where:

π_{0t} is the cumulated variation of the subindex between time *t* and 0.

The second approach uses subitem price variations. In this case, we use September 1996 as reference base for weights. This weighting structure is kept for the three periods:

$$p_{rt} = \sum_j w_{j00} \cdot \frac{p_{jrt}}{p_{j00}} \qquad (13)$$

where:

p_{rt} is the Index for metropolitan area *r* at time *t;*
p_{jrt} is price of good *j* in metropolitan area *r* at time *t;*
p_{j00} is average price of good *j* in all metropolitan areas where weight greater than zero is observed in the reference date; and
w_{j00} is the subitem *j* average national weight at reference period $t = 0$, that is, the average share of the subitem expenditure in total (or group or subgroup or item) household expenditure over all metropolitan areas.

For comparison, Table A.2 displays national average official RPDI subindexes and the "biases" against the official IPCA. Note that the first approach produces figures closer to the official ones, but this is easier to understand, as it takes account of changes in weights along time. The second approach is more rigorous, though, because we are comparing a same basket of goods, with the same weights; that is why we use the output of this approach in our reported estimations of the Consumer Demand System. The difference between the two results is apparent, especially in 1987. RPDI-S stands for RPDI deflated at subitem level, and RPDI-G for RPDI deflated at group level.

Table A.2

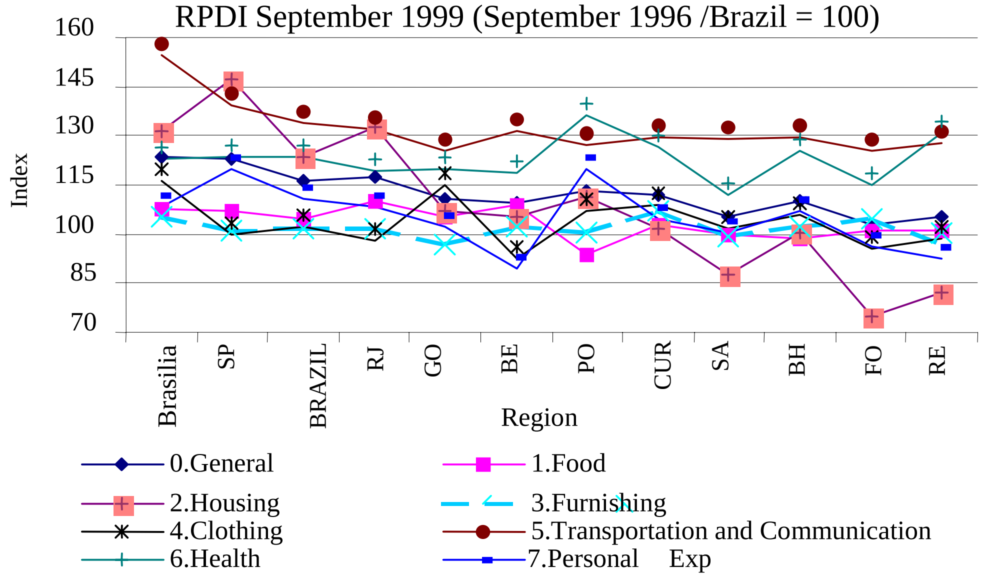
Official IPCA Subindexes — National Index

Expenditure Group	1987	1996	1999	"Bias" RPDI-S (%)	"Bias" RPDI-G (%)
0.General	2,238.94	100.00	114.62	71.63	1.44
1.Food	2,801.38	100.00	106.68	85.13	0.69
2.Housing	1,335.29	100.00	124.58	15.95	–0.18
3.Furnishings	5,083.78	100.00	106.45	69.56	–1.54
4.Clothing	6,330.79	100.00	107.11	0.88	–4.89
5.Transportation and Communication	2,237.64	100.00	139.51	52.03	–6.08
6.Health Care	1,373.08	100.00	127.87	43.21	–4.47
7.Personal Expenses, Education and Reading	1,364.04	100.00	115.87	100.10	6.43

The following figure displays the comparative cost-of-living across regions in September 1999. It allows us to affirm indisputably that São Paulo, Brasília and Rio de Janeiro are the most expensive areas in Brazil as a whole, especially on housing and transportation groups. Earlier indices (not reported here) also show that these ranks have not been altered since 1987. Other areas have alternated themselves in the ranking along time, but also a clear pattern was observed that Northeastern capitals have since long been the cheapest areas.

Figure

RPDI September 1999 (September 1996 /Brazil = 100)



BIBLIOGRAPHY

ACCRA. *Cost-of-living index manual*. 2000, downloaded from http://www.accra.org/edu_prof/pubs/colimanual.pdf.

ALSTON, J. M., FOSTER, K. A., GREEN, R. D. Estimating elasticities with the linear approximate almost ideal demand system: some Monte Carlo results. *Review of Economics and Statistics*, v. 76, n. 2, p. 351-56, 1994.

ALVES, D. C. O., DISCH, A., EVENSON, R. E. *The demand for food in Brazil*. Trabalhos apresentados no IV Encontro Brasileiro de Econometria. Águas de São Pedro: Sociedade Brasileira de Econometria, 1982.

ANDRADE, M. V., LISBOA, M. B. *Gastos com saúde no Brasil*. Rio de Janeiro: EPGE/FGV, 2001, mimeo.

ASANO, S. Joint allocation of leisure and consumption commodities: a Japanese extended consumer demand system 1979-90. *Japanese Economic Review*, v. 48, n. 1, p. 65-80, 1997.

BUSE, A. Evaluating the linearized almost ideal demand system. *American Journal of Agricultural Economics*, v. 76, n. 4, p. 781-93, 1994.

__________. Testing homogeneity in the linearized almost ideal demand system. *American Journal of Agricultural Economics,* v. 80, n. 1, p. 208-220, 1998.

CIPRIANO, J., BRANDT, S. A. *Sistema logit multinomial de demanda agrícola*. Trabalhos apresentados ao V Encontro Brasileiro de Econometria. Belém: Sociedade Brasileira de Econometria, 1983.

DEATON, A. *The analysis of household surveys.* Baltimore, Johns Hopkins U.P, 1997.

DEATON, A., MUELLBAUER, J. An almost ideal demand system. *American Economic Review,* v. 70, n. 3, p. 312-326, 1980.

HOFFMANN, R. Elasticidades de Engel para dispêndios familiares na Cidade do Rio de Janeiro: outro método de estimação. *Pesquisa e Planejamento Econômico,* v. 13, n. 1, p. 267-274, 1983.

__________. Elasticidades de Engel para dispêndios familiares na Cidade do Rio de Janeiro: outro método de estimação — uma correção. *Pesquisa e Planejamento Econômico,* v. 18, n. 1, p. 197, 1988.

IBGE. *Sistema nacional de índices de preços ao consumidor: métodos de cálculo*. Rio de Janeiro: IBGE, 1994 (Série Relatórios Metodológicos, v.14, 2ª ed.).

KAKWANI, N. C. Application of Lorenz curves in economic analysis. *Econometrica,* v. 45, n. 3, p. 719-727, 1997*a*.

——————. On the estimation of Engel elasticities from grouped observation with application to Indonesian data. *Journal of Econometrics,* v. 6, n. 1, p. 1-19, 1977*b*.

——————. A new method of estimating the Lorenz curve: a reply. *Econometrica,* v. 48, n. 4, p. 1.063-1.064, 1978.

MANAGEMENT AND COORDINATION AGENCY. *Annual Report of the Consumer Price Index*. Tokyo: Statistical Bureau, 1998

MEDEIROS, J. A. S. Curvas de Engel e transformação de Box-Cox: uma aplicação aos dispêndios em alimentação e educação na Cidade de São Paulo. *Pesquisa e Planejamento Econômico,* v. 8, n. 3, p. 795-828, 1978.

PASHARDES, P. Bias in estimating the almost ideal demand system with the stone index approximation. *Economic Journal,* v. 103, n. 419, p. 908-915, 1993.

POLLAK, R. A. *The theory of the cost-of-living index.* New York and Oxford: Oxford University Press, 1989.

RAO, C. R. *Linear statistical inference and its applications.* 2nd ed., New York, Wiley, 1973.

ROSSI, J. W. Elasticidades de Engel para dispêndios familiares na Cidade do Rio de Janeiro. *Pesquisa e Planejamento Econômico,* v. 12, n. 2, p. 579-606, 1982.

——————. Elasticidades de Engel para dispêndios familiares na Cidade do Rio de Janeiro: outro método de estimação — comentários. *Pesquisa e Planejamento*

Econômico, v. 13, n. 1, p. 275-284, 1983*a*.

——————. Elasticidades de Engel para dispêndios familiares nas principais capitais brasileiras. *Revista Brasileira de Estatística*, v. 44, n. 176, p. 381-451, 1983*b*.

ROSSI, J. W., NEVES, C. Elasticidades de Engel no Brasil usando um sistema de equações com especificações logit. *Revista Brasileira de Economia*, v. 41, n. 4, p. 425-434, 1987.

SIMÕES, R. C., BRANDT, S. A. *Sistema completo de equações de demanda para o Brasil*. Trabalhos apresentados no III Encontro Brasileiro de Econometria. Olinda: Sociedade Brasileira de Econometria, 1981.

THOMAS, D., STRAUSS, J., BARBOSA, M. M. T. *Estimating the impact of income and price changes on consumption in Brazil.* New Haven (CT), Yale University, 1989 (Yale Economic Growth Center Discussion Paper, 589).