A ROBUST POVERTY PROFILE FOR BRAZIL USING MULTIPLE DATA SOURCES

AN ANALYSIS OF MARGINAL EFFECTS.

While the cross-tabulations are informative, they have two shortcomings. First, the simple associations between personal characteristics and different measures of poverty are essentially bivariate, and do not control for the effects of other variables. Second, the long tables are not wieldy to test the robustness of the profile with respect to changes in spatial price deflation or in the assumptions about scale economies within households, which was one of the advantages of the methodology proposed in Section 2. We therefore conduct the robustness tests in a 'marginal effect' version of the profile, given by simple transformations of a probit model, regressing the probability of being poor on the relevant household characteristics which were used in the cross-tabulations.[1] In this exercise, poverty statistics are computed from income data in the PPV sample, and all covariates come from the same source.

These profile probit regressions are intended to be merely descriptive, and no inference of causation whatsoever is made. The transformed coefficients should be seen only as estimates of partial correlation coefficients with the probability of being poor. The vector of independent variables X includes the following household variables: regional location; some housing characteristics, access to water, electricity and telephones, and the following attributes of the household head: gender, age, race, years of schooling and labor status. The coefficients β are then transformed into marginal effects of a change in the relevant element of X on the probability of being poor, dF/dx. These are tested for statistical significance using standard errors which are adjusted for the clustering process inherent in the sampling procedure. The marginal effects and their p-values for the preferred regression (with the São Paulo price index, and $\theta = 1$) are reported in Table 4 below.

Table 4: Probit Analysis Results, $z = z^{-} =$ R$ 131.97, $I = I_{+}$, $\eta = 1.0$

Variable	dF/dx	P > ½z½	Variable	dF/dx	P > ½z½
Demographic variables					
Household size	0.0838	0.000	Proportion of HH aged 5-15	0.4635	0.000
{Household size}^{2}	-0.0035	0.002	Proportion of HH aged > 65	0.0050	0.949
Proportion of HH aged < 5	0.7788	0.000			
Characteristics of Household Head					
Age	0.0050	0.204	Mulato dummy	0.0157	0.490

[1] As θ varies, we scale the poverty line up by a factor equal to $\bar{n}^{1-q}$, where $\bar{n}$ is the average household size, so as to keep the overall poverty incidence rate constant for households with the average household size. This allows us to compensate for the pure size effect of the adjustment to the income effect, while preserving the re-rankings which are an important part of the exercise.

{Age}^{2}	-0.0001	0.176	Indigenous dummy	0.1870	0.183
Years of schooling	-0.0229	0.000	Self-employed dummy	0.0970	0.153
Female dummy	-0.0038	0.882	Unemployed / Unpaid	0.0688	0.300
Black dummy	-0.0304	0.445	Employee	-0.0530	0.368
Housing Characteristics and Access to Services					
Dirt floor in house	0.1226	0.011	Piped Water	-0.1129	0.001
# Bedrooms	-0.0676	0.000	Electricity	-0.1374	0.008
Dirt Road outside	0.0178	0.494	Phone	-0.2281	0.000
Favela dummy	0.0648	0.114			
Regional Dummies					
RM Fortaleza	0.3603	0.000	RM B. Horizonte	0.1249	0.002
RM Recife	0.5325	0.000	RM Rio	0.1973	0.000
RM Salvador	0.4889	0.000	SE – Other Urban	0.0909	0.025
NE – Other Urban	0.5367	0.000	SE - Rural	0.1940	0.001
NE - Rural	0.3549	0.000			

Table 4 contains a number of interesting results. First, controlling for the other variables included, household size has a significant positive and concave effect on poverty. Large households do appear likely to be poorer, controlling for other attributes, although the relationship is concave in family size. Similarly, the proportion of children is positively correlated with poverty, and more strongly so for younger children. No such significant correlation is found for the proportion of over-65s in the household. These results are robust not only to different price deflation procedures but also, more interestingly, to changing the household equivalence scale parameter θ to 0.75. In that regression, household size remained positive, concave and significant, and the results for children and the elderly were unchanged. Only when the probit was run for an income vector adjusted by $\theta = 0.50$, did we observe a reversal in the sign of the marginal effect of household size, which then became insignificant. This suggests that, unless there are reasons to suppose that economies of scale within Brazilian households are greater than those implied by a theta in the (0.7, 1.0) range, the stylized fact that larger households are poorer, controlling for other attributes, survives scrutiny. Our findings also suggest that a larger number of children is correlated with a greater probability of being poor, while the same is not true of a larger number of older people.

Turning then, to the marginal effects of characteristics of household heads, we find some surprising results. The unsurprising one, of course, is that education is significantly negatively correlated with the probability of being poor (although, even here, the effect is quantitatively much smaller than that of living in a richer area). But apart from education; age, gender, ethnicity and the occupational status of the household head, all turn out to be insignificant correlates of poverty. For age and gender, this is in line with previous findings from decompositions of Generalized Entropy inequality measures (see Ferreira and Litchfield, 2001). It is also confirmed by the tabulation profiles presented in the previous Section.

Race, however, had appeared to account for a significant share of inequality in those static inequality decompositions, and the tabulation profiles show substantial differences between the poverty incidences across households headed by blacks (including 'mulatos'), and whites. Clearly, the insignificance of the race dummy in the probits is a result of controlling for the other attributes included in the regression. While on average, black and indigenous households are substantially more likely to be poor, this

seems to be because of other differences between them and white-headed households, such as education or regional location. This is not to say that there are no grounds for poverty reducing policies which take race into account. Neither can it be interpreted as a verdict on the old sociological debate about whether Brazil's racism is more 'economic' than 'social'. All it does say is that if households headed by non-whites are likelier to be poor, then this is due to their differential access to education, or to their locational choices, or to some other factor, rather than simply because they are non-white.

In terms of housing characteristics and access to services, the direction of causation is almost certainly from poverty to these attributes, rather than the reverse. Our caveat about interpreting these 'marginal effects' merely as descriptive estimates of partial correlation coefficients is particularly pertinent here. The main result is that the poor are indeed significantly less likely to have access to piped water, electricity or, even more markedly, a telephone line. They are also less likely to have many bedrooms, or covered housing floors. The correlations with the nature of the road or street outside, as well as to whether the household is located in a slum ('favela'), turned out to be insignificant, once other factors are taken into account.

Finally, the effect of regional location on the probability of being poor can only be described as dramatic. The reference region (missing dummy) is the metropolitan area of São Paulo. Simply put, the marginal effects reported suggest that living anywhere else is correlated with a greater likelihood of being poor, though the quantitative effects are much larger for the Northeast than within the Southeast. Note that these effects have remained this strongly significant *after* controlling for differences in education, labor status, housing characteristics, etc. The implication is that regional differences in household income, and hence in the vulnerability to poverty, are not only a consequence of different educational attainment levels, demographic differences across regions, or racial make-up. They must be explained by other factors, which deserve continuing investigation.

Table 7 below lists estimates of poverty incidence (headcounts) from the PPV and the PNAD, for the ten sub-regions where the PPV is carried out and is representative. It also presents the (sampling design adjusted) 95% Confidence Interval around each of the PPV estimates. The PNAD headcounts come from the adjusted PNAD distribution described in Section 2, reflecting imputed rent and regional price deflation adjustments. The PPV estimates are presented for three different welfare indicators which can be constructed from the PPV data: the first is the real per capita household consumption expenditure; the second is real per capita household income, calculated from the more detailed income questions in the PPV questionnaire; the third is real per capita income from PPV questions analogous to those in the PNAD questionnaire.

Table 7: Headcount Indices from Different Welfare Concepts and Surveys [#]

PPV Region	PPV Headcount Estimate	95% C. I. lower bound	95% C. I. upper bound	PNAD Headcount Estimate
PPV Welfare Concept 1: Real Per Capita Consumption Expenditure.				
RM Fortaleza	0.1850	0.0117	0.3582	**0.2626***
RM Recife	0.2212	0.1342	0.3082	**0.2768***
RM Salvador	0.1928	0.1431	0.2424	**0.2697**
NE Urban	0.3756	0.2875	0.4638	**0.4011***
NE Rural	0.4981	0.3820	0.6143	**0.6850**
RM B. Horizonte	0.0791	0.0251	0.1332	**0.0856***
RM Rio	0.0304	0.0186	0.0422	**0.0613**
RM São Paulo	0.0375	0.0027	0.0723	**0.0273***
SE Urban	0.0472	0.0197	0.0748	**0.0743***

SE Rural	0.2603	0.1683	0.3523	**0.3539**
PPV Welfare Concept 2: Real Per Capita Income (Constructed).**				
RM Fortaleza	0.1236	0.0149	0.2323	**0.2626**
RM Recife	0.1970	0.1575	0.2365	**0.2768**
RM Salvador	0.1730	0.1413	0.2048	**0.2697**
NE Urban	0.2896	0.2311	0.3481	**0.4011**
NE Rural	0.2241	0.1480	0.3002	**0.6850**
RM B. Horizonte	0.0557	0.0258	0.0855	**0.0856**
RM Rio	0.0553	0.0198	0.0909	**0.0613***
RM São Paulo	0.0227	0.0123	0.0331	**0.0273***
SE Urban	0.0466	0.0202	0.0731	**0.0743**
SE Rural	0.1019	0.0541	0.1497	**0.3539**
PPV Welfare Concept 3: Real Per Capita Income from questions like those in PNAD ***				
RM Fortaleza	0.1060	-0.0182	0.2302	**0.2626**
RM Recife	0.1547	0.1104	0.1989	**0.2768**
RM Salvador	0.1188	0.0978	0.1398	**0.2697**
NE Urban	0.2340	0.1694	0.2986	**0.4011**
NE Rural	0.3935	0.2991	0.4879	**0.6850**
RM B. Horizonte	0.0205	0.0120	0.0321	**0.0856**
RM Rio	0.0247	0.0011	0.0483	**0.0613**
RM São Paulo	0.0105	0.0028	0.0182	**0.0273**
SE Urban	0.0127	0.0017	0.0237	**0.0743**
SE Rural	0.0973	0.0535	0.1410	**0.3539**

Notes: # based on the indigence line z of R$65.07 per month in all cases.

** denotes PNAD headcount estimates which fall within the 95% Confidence Interval for the PPV estimate in each welfare concept category.*

*** This measure of real per capita income is constructed by aggregating for each household the total value of incomes, in cash and kind, reported in response to a large number of separate questions in the PPV questionnaire, and deducting the cost of inputs into household production wherever that is appropriate. The general wisdom is that it provides a more reliable guide to real household income than the single question concept, analogous to that reported in the PNAD.*

**** This measure is also derived from the PPV, but is based on single questions about the incomes of farmers and self-employed workers, like those in the PNAD questionnaire. This concept is thus supposed, ex ante, to be the most comparable with PNAD results.*

Sources: Authors' calculations from the PPV 1996/97 and the adjusted PNAD 1996.

APPENDIX

Table A1: Indigence Profile 1996: Brazil , z = z (R$ 65.07/month), I = L_1, q=1.0[2]

Household Characteristics	Subgroups	f_k	$\mu(y)_k$	P_{0k}	P_{1k}	P_{2k}	s_k

[2] One may argue against the use of imputed rent in indigence measures.

	Total	100.00	283.86	22.59	9.60	5.53	100.00
Region	North	4.84	191.96	30.06	11.80	6.58	6.44
	North-East	29.59	135.37	47.89	22.14	13.28	62.72
	Center-West	6.81	282.75	16.63	5.90	3.08	5.01
	South-East	43.59	380.40	9.19	3.22	1.65	17.73
	South	15.17	325.91	12.08	4.45	2.33	8.11
Location	Metropolitan Core	17.63	498.29	7.47	2.41	1.15	5.83
	Metropolitan Periphery	12.14	300.41	10.07	3.33	1.65	5.41
	Large Urban	18.89	365.02	10.22	3.17	1.46	8.55
	Medium Urban	15.69	271.24	17.58	6.43	3.24	12.21
	Small Urban	15.02	173.80	30.82	12.63	7.03	20.49
	Rural	20.63	106.38	52.03	25.54	15.93	47.52
Dependency Ratio*	1	9.99	630.69	0.53	0.10	0.03	0.23
	1<d=<1.5	14.60	410.76	2.50	0.55	0.20	1.61
	1.5 <d=<2	22.40	326.78	7.60	1.93	0.76	7.54
	2 <d=<3	21.85	211.86	23.44	7.03	2.97	22.67
	3 <d=<4	13.61	184.66	33.36	12.57	6.21	20.09
	d>4	15.31	100.81	58.28	29.31	17.94	39.49
	Other/Not Specified	2.25	37.83	84.12	59.84	48.38	8.37
Housing Status	Own House, Paid, with Own Land	63.76	288.74	22.37	9.51	5.48	63.12
	Own House, Paid without Own Land	5.60	148.08	42.00	20.30	12.60	10.40
	Own House, Still Paying	6.06	440.54	5.26	1.40	0.56	1.41
	Rent	12.23	366.34	10.64	3.50	1.65	5.76
	Ceded	11.70	160.54	35.75	15.55	9.00	18.52
	Other	0.50	172.71	24.75	9.46	5.28	0.55
	Not Specified	0.15	216.01	35.68	14.99	8.65	0.24
Water	Piped	81.59	332.35	13.04	4.53	2.28	47.08
	Not Piped	18.26	67.83	65.19	32.21	20.04	52.68
	Other/Not Specified	0.15	207.79	35.46	15.01	8.75	0.24
Sanitation	Sewerage System	37.84	442.21	5.47	1.59	0.69	9.15
	Concrete Cesspit 1	10.19	388.72	6.26	1.91	0.90	2.82
	Concrete Cesspit 2	12.84	235.26	17.93	6.04	2.97	10.19
	Rudimental Cesspit	22.67	145.50	34.19	13.58	7.35	34.31
	Drain	1.98	112.58	39.99	17.77	10.52	3.51
	River or Lake	2.75	164.73	23.41	8.69	4.48	2.85
	Other	0.19	141.04	43.05	14.88	7.53	0.37
	Not Specified	11.52	57.68	72.16	37.60	24.16	36.79
Electricity	Yes	91.93	303.66	18.25	7.08	3.82	74.26
	No	7.91	55.10	72.87	38.87	25.35	25.50
	Other/Not Specified	0.16	212.15	33.80	14.20	8.19	0.24
Waste Disposal	Collected Directly	63.26	373.41	8.88	2.87	1.35	24.86
	Collected Indirectly	7.36	257.20	21.18	7.60	3.86	6.90
	Burned	14.35	112.50	46.48	20.84	12.31	29.51
	Unused Plot of Land	13.23	79.32	59.44	29.13	18.13	34.81
	Other/Not Specified	1.80	115.39	49.10	21.36	12.50	3.91
Characteristics of the Head	Subgroups	f_k	$\mu(y)_k$	P_{0k}	P_{1k}	P_{2k}	s_k

Gender	Male	82.26	282.64	23.30	10.06	5.85	84.85
	Female	17.74	289.52	19.30	7.48	4.04	15.15
Race	Indigenous	0.17	168.69	47.20	25.49	17.47	0.36
	White	54.27	384.04	12.66	4.95	2.76	30.39
	Black	45.07	159.79	34.64	15.21	8.86	69.09
	Asian	0.46	671.79	4.53	1.62	0.88	0.09
	Not Specified	0.02	89.60	59.45	28.95	15.01	0.06
Age	0-24	3.97	188.88	27.45	10.90	5.92	4.83
	25 to 44 Years	48.40	268.02	24.59	10.88	6.43	52.66
	45 to 64 Years	36.43	305.75	21.65	9.22	5.32	34.92
	>65 Years	11.20	314.79	15.33	4.89	2.19	7.60
Education	0- 1 Years	21.86	104.48	46.22	21.23	12.70	44.71
	1 to 4 Years	20.03	150.86	32.95	14.37	8.41	29.22
	4 to 8 Years	30.10	230.49	15.78	5.73	2.99	21.03
	8 to12 Years	20.56	394.59	5.44	1.72	0.80	4.95
	> 12 Years	7.45	1077.98	0.30	0.08	0.03	0.10
Immigration Status	Not Immigrant	40.56	258.16	30.23	13.81	8.29	54.26
	0 to 5 Years	7.51	270.34	21.37	8.59	4.77	7.10
	6 to 9 Years	4.25	262.61	20.39	7.86	4.14	3.83
	More Than 10 Years	28.87	295.57	17.94	6.92	3.73	22.93
	Other/Not Specified	18.81	331.48	14.27	5.44	2.96	11.88
Labour Status	Inactive	17.70	279.16	18.71	7.30	4.03	14.65
	Unemployed	2.77	131.51	45.81	22.04	13.82	5.62
	Formal Employees	23.31	292.55	10.96	3.26	1.38	11.30
	Informal Employees	13.30	162.34	36.60	15.28	8.36	21.55
	Self-Employed	27.00	235.64	30.66	14.23	8.60	36.63
	Employer	4.76	781.14	4.95	2.04	1.22	1.04
	Public Servant	8.73	422.27	10.66	3.49	1.59	4.12
	Unpaid	2.39	139.04	46.89	26.93	19.46	4.97
	Other/Not Specified	0.04	124.31	65.55	38.43	28.28	0.13
Employment Tenure	0 Years	20.47	259.16	22.38	9.30	5.36	20.28
	1 Years or More	13.04	215.60	24.23	9.95	5.60	13.98
	1 to 3 Years	14.65	260.42	19.87	7.53	3.96	12.88
	3 to 5 Years	8.23	301.52	18.49	7.06	3.70	6.73
	> 5 Years	43.19	322.23	23.70	10.71	6.39	45.29
	Other/Not Specified	0.42	134.50	44.46	21.45	13.65	0.84
Sector of Occupation	Agriculture#	19.61	117.00	52.44	25.77	16.06	45.52
	Manufacturing	12.15	310.39	15.80	5.96	3.09	8.49
	Construction	8.04	200.47	18.48	6.39	3.20	6.58
	Services	31.50	373.11	11.22	3.64	1.69	15.64
	Public Sector	8.23	443.76	9.62	3.15	1.41	3.50
	Other/Not Specified	20.47	259.15	22.38	9.30	5.36	20.28

Notes: $s_k = \frac{f_k P_{ok}}{P_0}$. Dependency ratio is defined as the number of household members over the number of earners in the household. # Agriculture includes other Primary Sector occupations.