Measuring Income in the 1973 PNAD Sample

The purpose of this report is to discuss some general issues in the measurement of income in the 1973 PNAD survey. It is intended as a not-for-publication working paper that will hopefully prove useful as a general reference document.

The PNAD questionnaire contains a number of items that pertain to income. The first of these is "Type of Employment," which assumes the following values:

0. No information

- 1. Employee, private business
- 2. Employee, public business

3. Employer

4. Self-employed worker

5. Unpaid family worker

Respondents who answered "1" or "2" on this question were then asked how much weekly income they received. Responses were tabulated as either money income (so em dinheiro) or income in kind (alem de uma parte em bens). PNAD then translated these responses into cruzeiro values as illustrated in the first panel of Table 1. We use the midpoints of these intervals to estimate weekly income.

Respondents who answered "3" or "4" on the type of employment question were asked to indicate their monthly income. Monthly income is always strictly money income. The second panel of Table 1 shows how this variable was treated.

The upper interval for both the weekly and monthly income distribution is open-ended, so we estimated what we considered a plausible value for the upper category. As we will demonstrate later in this report, these 1. Positive values on this measure indicate clustering to the left with extreme values to the right. Negative values indicate clustering to the right with extreme values to the left. A value of 0 would indicate a normal distribution.

2. Positive values indicate a distribution that is more peaked than a normal distribution, while negative values indicate a flatter distribution. A value of 0 indicates normality.

departures from normality and do violence to regression assumptions, since high values might well tend to dominate income equestions. The usual solution to this issue is to take the logarithm of income. Doing this reduces the skewness to .572 and the kurtosis to .779.

The zero-order correlation between income and ln income is only .761, which (at least in comparison to U.S. samples) is quite low. This might suggest that while we have prohibited high income earners from disproportionately dominating the distribution, we may now be giving undue emphasis to low earners. Since the metric and ln versions of income are only two of a family of power transformations, perhaps something in between the two would be more satisfactory. Two obvious candiates are the square root (i.e., Income^{.5}) and the cube root (Income^{.333}) of income.

The square root of income correlates .936 with income, while the cube root correlates .887 ($r_{inc}^{1/2}_{inc}^{1/3}_{=}$.991). Both specifications, as one would expect, are closer to being normally distributed than is income, but less so than is ln income. Various regression analyses, however, demonstrate conclusively that we gain little if anything either substantively

report no income (these are predominantly in agriculture).

-5-

Table 3 shows the results of including zero-earners in our analyses. We estimate our usual regression equation, and designate our new specification as "Income 0." We also report metric coefficients in addition to betas. Substantively, the two specifications would be interpreted similarly, yet the differences are not trivial. Confining ourselves to metric coefficients for the moment, the first specification makes it seem as if one year of education were worth about three points of occupational status (149.12/51.03=2.92), while the new specification makes the ratio seem closer to 2:1 (114.88/54.78=2.10). Clearly, the decision on what to do with zero-earners is important.

This is even more apparent when we look at the ln form of an income specification that includes zero earners. Since ln(0) is undefined, we arbitrarily assign ln(0)=0. Equation 3 of Table 3 presents the beta coefficients for our usual equation. Basically, the equation implies that variations in education have no effect on income variations at the lower levels of the income distribution. Including zero earners clearly results in income equations that are dominated by respondents to the left of the distribution.¹

1. The problem of zero wage earners will be even more severe for women workers, given the exceptionally high proportion who are unpaid family farm workers (Olson and Tourinho 1981).

Both including and excluding zero earners creates serious problems. At this point I'm not sure what to do about it. Table 1

me weight by AM, MJ, DB, Not 4/ 4/ 24;

Cetter 10 or 31 = 4 pt - 3. 92, et

J. Dece 17, 198

WAGES SALARIOS

In money mly Sõ em dinheiro

> 10 11

12

13

14

15

16

17

18

19

20

21

22

23

24

the dian

WEEKLY .SEMANAL

Ate 3,92

Range

Beaulus a part in Kin

31 -

32

33

34

35

36

37

38

39

40

41

42

43

44

45

Alem de uma parte em bens

5.88	3,92	a	7,84	
11.76	7,85	а	15,68	
17.60	15,69	a	23,52	
12 2.4.1	23,53	a	31,36	
(i i i z 4	31,37	a	47,01	
54.87	47,02	a	62,72	
78.40	62,73	a	94,08	
109.76	94,09	a	125,44	
141.12	125,45	.a	156,80	
172.48	156,81	a	÷ 188,16	
217.02	188,17	a	250,88	
232.24	250,89	a	313,60	
344.96	313,61	a	376,32	
407.68	376,33	a	439,04	
486.08	439,04	a	533,12	
580,17	533,13	a	627,20	
784.00	627,21	a	940,80	
1097,60	940,81	a	1.254,40	
1568.001.	254,41	6	1.881,60	
3000 1.	881,61	е	mais	
Se	m declar	วกวิก		

25	46
26	47
27	48
28	. 49
29	50
30	51
ØI	Ø2

Atē 67,20

.

100,80 67,21 a 134,40 201.60 134,41 a 268,80 403,20 268,81 a 537,60 60 61 63 Mun Phy . MENSAL - continuação

Sõ em dinheiro

	•		
(M2.60 537,0	57 a	806,40	64
1075.20 806,4	1) a	1.344,00	65
1612.80 1.344,0	01 a	1.881,60	66
2084, 80 1.881,0	51 a	2.688,00	67
3360,002.688,0)] a	4.032,00	68
4704,504.032,0)1 a	5.376,00	69
6720.005.376,0)1 a	8.064,00	70
8,064,0	00 e	mais	71
Sem dec	laração)	Øl

-8--

115

	Father's Occupation	Education	Occupation	Job Experience*	Occupational Experience*	L.F. Experience Age	R ²
Income β Income b	034 -6	.233 149	.369 51				. 289
Income β Income b	027 -5	.260 166	.367 51	.169 40	•	· · · · · ·	.317
Income β Income b	032 -6	.255 163	. 381 53	•	.133 28	•	.306
Income β Income b	024 -4	.298 190	.330 46		н	.177 29	.318
Income β Income b	032 -6	.293 188	.317 44		· ·	.191 33	.324

-H

Table 4. Consequences of Using Different Measures of Experience.

*Since the metric for job experience and occupational experience is in months, we multiply the b coefficients by 12.