Title: AJAE Appendix for *Evaluating Alternative Policy Responses to Higher World Food Prices: The Case of Increasing Rice Prices in Madagascar*

Authors: David Coady, Paul Dorosh, and Bart Minten

Date: December 1, 2008

Note: The material contained within is supplementary to the article named in the title and published in the American Journal of Agricultural Economics (AJAE)

Many transfer programs contain some element of means testing. Simple means testing links eligibility and transfer size to some estimate of individual or household income. This requires a capacity to verify information provided by applicants regarding the monetary income of various household members from various sources. It is thus more feasible and relevant in economies where economic activity is mostly formalized. It is less attractive in economies where informal economic activity is prominent and verifying income more difficult. This is especially likely to be the case for the lower income groups targeted by safety net programs.

For this reason, a growing number of low- and middle-income countries are using a more statistical *proxy-means approach* that identifies key socio-economic characteristics strongly correlated with economic status, attaches a numerical weight to each characteristic and calculates a score by summing the product of weights and characteristics. The characteristics used should be easily observable and thus verifiable by program officers and not easily manipulated by applicants. This score is used to determine eligibility. Such an approach will undoubtedly result in the exclusion of

genuinely eligible households and leakage of benefits to some less needy households. In addition, its structural nature means that it is more useful for dealing with structural "poverty" as opposed to transitory or short-run poverty. For these reasons, this approach is best combined with other approaches, e.g. with some element of community targeting or self-selection.

The proxy-means approach to targeting starts by identifying household characteristics that are highly correlated with poverty. Each of these characteristics is given a numerical *weight* and this is used to calculate a household *score*. This score is then used as the basis for determining eligibility. For example, consider the case where three household characteristics are used: a binary variable indicating whether or not the house is made of brick (H), a continuous variable capturing the years of education of the head of household (E), and a binary variable indicating whether or not the head of household is female (F). Using a national household survey, one can regress, say, per capita household consumption on these variables. The coefficient estimates (say, h,e and f) are then used to calculate a household score (S) as:

$$S = c + h*H + e*E + f*F$$

where c is a constant in the regression. A high score indicates a high predicted per capita consumption level. If the program is expected to cover the poorest 30 percent of households, then households that fall within the lowest 30 percent of households based on S are deemed eligible to participate in the program.

In practice, the model used to calculate the score (e.g. the types and number of household characteristics used to estimate the weights and the statistical procedure used) can be much more sophisticated than the example given above. But a key ingredient to applying the proxy means approach is access to a national household survey data set that contains information on an acceptable indicator of household welfare (e.g. household consumption) as well as a range of household characteristics that are correlated with poverty.

As an illustration, we construct a simple proxy means algorithm based on information available in the 2001 Madagascar national household survey (ECM2001). We use per capita household consumption as our welfare indicator. This is regressed on a range of household characteristics to derive a set of regression coefficients that are used as weights. The variables included and the results of the regression analysis are presented in table 1. The underlying regression used per capita household consumption as the dependent variable, with independent variables including information on geographic location; gender, age, education and sectoral employment of the head of household; household size and composition; types of housing and materials for walls, floors and ceilings; housing area; source of water and lighting; and possession of various consumer durables. The r-squared for the regression was 0.63 based on 4,955 household observations. By multiplying each of the variables by the associated coefficient and adding across these products, we construct a score for each household – this can be interpreted as predicted household consumption per capita. We then select the poorest 30 percent of households, based on this score, as program beneficiaries.

Because the score is not perfectly correlated with household per capita consumption (our indicator of true welfare), this selection process will not identify the poorest households perfectly. In other words, some poor households (i.e. those in the bottom 30 percent of households based on actual household per capita consumption) will be wrongly excluded (i.e. "errors of omission", "type-I errors" or "undercoverage") while some nonpoor households will be included (i.e. "errors of inclusion", "type-II errors" or "leakage"). To get a sense of the magnitude of these errors, based on the simple model, in table 2 we present the coverage of the program across income groups (i.e. "participation rates") and the distribution of beneficiaries across income groups (i.e. "beneficiary shares").

Under this simulation of proxy-means targeting, 78 percent of the bottom income decile is deemed eligible and 57 percent of the next decile. The participation rates fall for higher income groups, especially for the top 40 percent of households. Therefore, this approach is quite effective at reducing coverage of higher income groups. It is likely that this leakage to the higher income groups could be reduced further by combining the proxy means approach with other approaches. For example, one might introduce an element of self-selection into the application process by requiring households to make an office application, reapply at short regular intervals and possibly locate offices in the poorest rural and urban areas. Alternatively, one might introduce a second qualifying hurdle by having program officials make a home visit and leaving open the possibility that households selected by the proxy means score could be dropped if the visit suggested that they were clearly much better off than predicted by the model. Or one might

introduce more refined geographic targeting of the budget by varying the proportion of households that could be deemed eligible across districts based on known district poverty rates. And, of course, the proxy means model itself may be developed in a more sophisticated manner.

Table 1. Proxy-Means Weights Based on OLS Regression

Place of Residence	Coefficient	Household characteristics	Corfficient
Fianarantsoa	Base	Ceiling bark	9.71 (0.89)
Toamasina	-12.16 (-3.73)	Ceiling clay	4.69 (0.43)
Mahajanga	-18.33 (-4.79)	Ceiling wood	11.63 (1.08)
Toliara	-5.15 (-1.47)	Ceiling cement	24.35 (1.99)
Antsiranana	-3.24 (-0.91)	Ceiling “natte”	4.91 (0.39)
Rural area	1.42 (0.59)	No ceiling	5.47 (0.51)
		Ceiling other	9.23 (0.80)
Head of Household Characteristics		Ceiling other (missing)	10.22 (0.89)
Male	-2.73 (-1.14)	Interior pluming (main)	11.84 (0.86)
Age	-1.23 (-3.32)	Interior tap (main)	34.03 (3.25)
Age squared	0.01 (2.63)	Water seller (main)	15.11 (1.24)
Education is CEPE	7.46 (2.53)	Water truck (main)	-65.59 (-1.04)
Education is BEPC	8.83 (2.88)	Rainfall (main)	-7.99 (-0.75)
Education is BL	18.08 (1.73)	Neighbour tap (main)	-13.62 (-1.17)
Education is PreBL	13.30 (1.16)	Neighbour well (main)	-5.27 (-0.50)
Education is university	32.08 (7.96)	Exterior tap (main)	12.15 (1.09)
Medium agricultural enterprise	4.17 (0.93)	Public tap (main)	-0.18 (-0.02)
Small agricultural enterprise	9.01 (0.86)	Pump well (main)	-11.71 (-1.07)
Cattle herder/fisherman	-0.81 (-0.14)	Simple well (main)	-4.32 (-0.43)
Agricultural trader	10.94 (2.23)	River/lake/sea (main)	-13.42 (-1.34)
Other trader	11.05 (2.88)	Bottled water (main)	54.19 (1.84)
High skilled	20.50 (4.36)	Other water source (main)	12.58 (1.13)
Low skilled	11.02 (3.40)	Interior pluming (second)	-20.08 (-0.73)
Unskilled	12.62 (2.83)	Interior tap (second)	9.89 (1.56)
Other	-0.13 (-0.03)	Water seller (second)	12.63 (1.61)
		Water truck (second)	-1.21 (-0.02)
Household characteristics		Rainfall (second)	-6.42 (-1.88)
Number of children	-19.75 (-24.75)	Neighbour tap (second)	16.51 (2.03)
Number of adult males	-11.87 (-10.69)	Neighbour well (second)	-2.86 (-0.60)
Number of adult females	-13.37 (-11.83)	Exterior tap (second)	-5.78 (-0.57)
Number of elderly	-10.29 (-4.89)	Public tap (second)	1.28 (0.27)
Apartment	-19.45 (-0.45)	Pump well (second)	-8.80 (-0.95)
Studio	-34.13 (-0.77)	Simple well (second)	7.13 (1.88)
One room	-30.17 (-0.69)	River/lake/sea (second)	-2.03 (-0.62)
Detached house, old	-31.49 (-0.72)	Bottled water (second)	23.35 (2.69)
Detached house, modern	18.16 (0.41)	Other water source second)	2.24 (0.81)
Other house	-12.86 (-0.29)	Electricity	32.48 (0.52)
Other house (missing)	-29.78 (-0.64)	Generator	13.92 (0.22)
Number of rooms/1000	0.23 (0.72)	Kerosene	6.73 (0.11)
House area (log)	7.76 (6.28)	Candles	27.84 (0.44)
Walls bark	-152.25 (-4.25)	Other energy	3.48 (0.06)
Walls clay	-157.25 (-4.40)	Gas stove	48.97 (7.60)
Walls wood	-154.82 (-4.32)	Fridge	27.61 (5.26)
Walls backplate	-149.02 (-3.37)	Radio	14.72 (6.76)
Walls masonite	-198.36 (-4.92)	TV	30.29 (9.81)
Walls tin	-144.89 (-4.04)	Car	57.94 (8.14)
Walls iron	-149.36 (-4.05)	Motorbike	4.78 (0.49)
Walls stone	-165.91 (-4.30)	Bicycle	4.39 (1.61)
Walls brick	-155.28 (-4.35)	Phone	71.01 (11.96)
Walls cement	-147.36 (-4.11)	Computer	-5.10 (-0.30)
Floor clay	216.02 (3.03)		
Floor wood	222.92 (3.13)	Constant	89.25 (0.90)
Floor stone	236.36 (3.31)	Number of observations	4955
Floor cement	229.68 (3.22)	R-squared	0.63

Floor other	224.03 (3.14)	Dependent Variable: Per capita household consumption

Note: t-statistics given in brackets after coefficient estimates.
Source: Based on information available in EPM2001.

Table 2. Performance of Proxy-Means Targeting

Per Capita Consumption Quintiles	*Participation Rates(%)*	*Beneficiary Shares(%)*
Bottom decile	78.3	26.1
2nd decile	57.0	19.0
3rd decile	52.5	17.5
4th decile	36.9	12.3
3rd quintile	27.3	18.2
4th quintile	9.4	6.3
Top quintile	0.8	0.5

Source: Numbers generated based on simulations of proxy means model using ECM2001.

Note: Participation rates are proportion of each income quantile deemed eligible for program based on proxy means algorithm. Beneficiary shares is the proportion of total program beneficiaries falling within each quantile. In all, 30 percent of households were deemed eligible.