Community Targeting for Poverty Reduction in Burkina Faso

Author(s): David Bigman, Stefan Dercon, Dominique Guillaume and Michel Lambotte

Source: *The World Bank Economic Review*, Jan., 2000, Vol. 14, No. 1 (Jan., 2000), pp. 167-193

Published by: Oxford University Press

Stable URL: http://www.jstor.com/stable/3990040

THE WORLD BANK ECONOMIC REVIEW, VOL. 14, NO. 1: 167-93

Community Targeting for Poverty Reduction in Burkina Faso

David Bigman, Stefan Dercon, Dominique Guillaume, and Michel Lambotte

This article develops a method for targeting antipoverty programs and public projects to poor communities in rural and urban areas. The method calls for constructing an extensive data set from a large number of sources and then integrating the entire set into a geographic information system. The data set includes demographic data from the population census; household-level data from a variety of surveys; community-level data on local road infrastructure, public facilities, water points, and so on; and department-level data on agroclimatic conditions. An econometric model that estimates the impact of household-, community-, and department-level variables on household consumption is used to identify the key explanatory variables that determine the standard of living in rural and urban areas. This model is then applied to predict poverty indicators for 3,871 rural and urban communities in Burkina Faso and to map the spatial distribution of poverty in the country. A simulation analysis assesses the effectiveness of village-level targeting based on these predictions. The results show that such targeting is an improvement over regional targeting in that it reduces leakage and undercoverage.

Budgetary and social pressures to improve the impact of health, education, and rural development programs on the poor are a strong impetus to improve targeting. Undifferentiated transfers that cover the entire population, such as general food subsidies, have proved to be beyond the budget constraints of most devel-

David Bigman is with the International Service for National Agricultural Research, the Netherlands; Stefan Dercon is with the Department of Economics at the Catholic University of Leuven, Belgium, and the Centre for the Study of African Economies at Oxford University; Dominique Guillaume is with the African Department at the International Monetary Fund; and Michel Lambotte is with I-Mage, a consulting group based in Belgium. David Bigman was with the Africa Region at the World Bank while this work was done. The authors' e-mail addresses are *d.bigman@cgiar.org, stefan.dercon@econ.kuleuven.ac.be, dguillaume@imf.org*, and *info@i-mage.be*. This study is part of a project on geographical targeting in Burkina Faso. The authors would like to thank the Belgian Trust Fund for providing financial support and the following institutions in Burkina Faso for providing assistance and data: the Institut National de Statistique et de Démographie (INSD), Famine Early Warning System (FEWS), and many other government departments and institutions including Institut Géographique du Burkina Faso, Direction Générale de l'Hydraulique, Direction des Statistiques Agro-pastorales, Office National de l'Elevage et de l'Agriculture, Direction National de la Météorologie, Ministère de l'Enseignement de Base et de l'Alphabétisation, Programme National de Gestion de l'Information sur le Milieu, Programme National de Gestion des Terroirs, and Direction des Marches Publics. The authors are grateful to Bart Capéau, Dirk Van de Gaer, Jesko Hentschel, Peter Lanjouw, Pramila Krishnan, three anonymous referees, and participants in the World Bank workshop Geographical Targeting for Poverty Reduction and Rural Development for useful comments.

oping countries. Further, the benefits of such transfers go disproportionately to the nonpoor.[1] In many developing countries, particularly those in Sub-Saharan Africa, targeting criteria that cover large geographic areas or large population groups also are likely to be ineffective and too costly to implement. A program that is targeted to the entire rural population, for example, may cover the majority of the country's population—both poor and nonpoor.

More accurate targeting requires criteria that can effectively identify eligible recipients. Such criteria can be narrowly defined, for example, at the level of individual households. Or they can be more broad-based, at the level of a region or province, by identifying the geographic areas or population groups that have a higher than average incidence of poverty (van de Walle 1995). Targeting at the household level is very information-intensive and thus very costly. Identifying eligible households requires complex and expensive means testing. But means-testing is only partly successful even in many industrial countries—despite the wide range of household data that are available in those countries—and a large portion of the benefits leak to ineligible households.

Most developing countries do not have the needed information on individual households, particularly poor households, which limits the scope for narrow targeting. As an alternative to direct means testing, standard household income and expenditure surveys, such as the World Bank's Living Standards Measurement Study (LSMS) surveys, can be used to identify more general characteristics of the poor and thereby to determine a set of indicators, such as number of children or place of residence, that can distinguish the poor.

With the LSMS data set for Côte d'Ivoire, Glewwe (1991) examines the tradeoff between using a refined and exhaustive set of indicators for narrow targeting and the cost of collecting the information needed to derive these indicators. He concludes that, for Côte d'Ivoire, a fairly limited set of community and household indicators proved to be quite effective in identifying poor households. However, the incentives for households to change or to lie about their characteristics in order to qualify for a program once these indicators are set can significantly reduce its effectiveness and blow up budgetary costs.[2] This problem, together with the high cost of administering a program at the household level and the danger that these eligibility criteria will leave out many of the country's poor, has deterred the governments of most developing countries from targeting social welfare programs to individual households.

Geographical targeting at the level of the province or region may be an effective approach for reaching the poor in countries where there are substantial disparities in living conditions between geographic areas and where administering poverty programs is relatively straightforward because the local administration is already in place. In India the central government disburses funds across states, in part according to the large disparities in poverty among states. The decision to

1. For simulated examples from Latin America see Baker and Grosh (1994).
2. See also Besley and Kanbur (1991).

locate rural development projects in poorer regions has become the center of India's poverty-oriented agricultural development strategy.

However, even in countries where the poor are concentrated in certain states or regions, geographical targeting at the level of large administrative areas is likely to entail considerable leakage to the nonpoor who live in target areas and fail to cover the poor who live in other areas. Although programs targeted at high levels of geographic aggregation are likely to result in less leakage and better coverage than general, nontargeted programs, their effectiveness in terms of poverty reduction tends to be quantitatively small (Ravallion 1993, Baker and Grosh 1994, Ravallion 1996). Targeted programs may also give households an incentive to move to targeted areas, defeating the purpose of the program and raising its costs.

Ravallion (1993) evaluates the costs and effects of geographical targeting at the level of the province in Indonesia. He concludes that, although targeting clearly helps to alleviate poverty, the magnitude of the gain is small. Baker and Grosh (1994) analyze geographical targeting in Republica Bolivariana de Venezuela, Mexico, and Jamaica and conclude that targeting priority regions can be an effective mechanism for transferring benefits to the poor. But with a given budget constraint, poverty reduction is greater the narrower are the target areas. The greatest reduction in poverty is achieved when the target areas are municipalities or villages.

Narrow geographical targeting at the level of the village or the urban community can reduce leakage in countries or regions where the majority of the population in a village or urban community faces similar socioeconomic conditions and living standards. Many of the households in a village or urban community may have similar sources of income, and all households are affected by the same agroclimatic and geographic conditions, including road conditions, distance to the nearest town, and availability of public facilities for services such as health, education, and water supply. Consequently, income inequality often is due, to a considerable degree, to income differences among villages and only to a lesser degree to income differences among individuals within villages.

But targeting at the district or village level requires much more information on the spatial distribution of poverty across districts or villages and on the characteristics of the poor in these areas. In most developing countries information on the standard of living is provided by a household survey. But the sample size of the standard survey is far too small to allow an estimation of the incidence of poverty at the level of the village or the district for the entire country. At present, the LSMS surveys can provide a map of the spatial distribution of poverty only among the country's main regions.

Some countries that rely on geographical targeting establish criteria for targeting that instead are based on more readily available indicators, such as access to public services, the percentage of school-age children who attend school, and the prevalence of certain illnesses associated with malnutrition. All too often, however, these indicators are not sufficiently correlated with the welfare indicators of the local population, which may lead to targeting errors in determining eligibility (Hentschel and others 1998).

The objective of this article is to present a method for geographical targeting at the level of rural villages and urban communities. We construct a large data set from several sources and integrate the data at the level of the village or urban community using geo-referencing. We then organize this database in the form of a geographic information system (GIS). With the GIS database we can generate a mapping of poverty at the level of the community and the province. The database includes several strata of information: demographic and socioeconomic information at the household level (taken from a variety of surveys); village- and community-level information, including demographic information from the population census, distance to urban centers, condition of road infrastructure, availability and quality of public services, and sources of drinking water; and departmental or regional information on agroclimatic and geographic conditions, including the location of the main towns and main transport routes.

In the second step of the analysis we use the GIS database, along with detailed data from a household survey, to construct a prediction model of household welfare. The model includes household-, community-, and department-level variables selected from the GIS database and therefore includes only variables for which mean values are available for all communities in the country.

In the third step we use the predictions of this model to estimate the incidence of poverty and the average level of well-being of the households in a community for all communities in the country. These estimates determine, in turn, the spatial distribution of poverty in the country at the village level. We apply this method to Burkina Faso, using the relatively detailed household data of the Priority Survey.

I. The Data

We collected data for this study from a large number of sources and aggregated them at the level of the village according to the name of the village and its geographic coordinates. Some of the data, most importantly the census data, cover all villages in the country or the entire population; other data cover only a sample of villages and a sample of households within each village (table 1). We could not use all of the data in the econometric analysis, however; some of the data did not cover all provinces, while other data, most notably those from the agricultural survey, did not contain the information needed to incorporate them into the GIS database.

After we collected the data, we standardized them and integrated them within a common data set. This data set contained more than 60 tables with information on the geographic coordinates of all villages, towns, markets, and public facilities; data on the entire road network; socioeconomic and demographic data from a variety of surveys and the population census; and data on the agroclimatic conditions in the country's main provinces. We then organized the data as a GIS, that is, a computer system that allows the analysis and display of geographic and other data.

Table 1. *Data Sources*

Level of aggregation	*Source*	*Coverage*
Household	Priority Survey (1994): data on income and expenditures for 8,642 households	Survey sample (473 villages)
Village	Priority Survey (1994): community component, which covers infrastructure and communal services	Survey sample (473 villages)
Village	National census (1985): demographic data	National
Village	Ministry of Water Management and Infrastructure (1995): data on health and water infrastructure, distances to infrastructure, public administration, and social groupings	25 of 30 provinces
Village	Ministry of Education (1995): data on primary school infrastructure and teacher-pupil ratios	National
Department	Ministry of Agriculture (1993): data on various indicators ranging from average literacy rates to vegetation indexes	National
Department	Directorate of Meteorology (1961–95): data on temperature (31 locations), evapotranspiration (15 locations) and rainfall (160 locations)	National
Province	Ministry of Agriculture (1993): data on cattle per household	National

To illustrate the type of information that we extracted from the GIS for the community study, figure 1 (in the appendix) shows the location of water points and their proximity to the villages in the Department of Karangasso-Vigue. The points on the map that indicate the location of the villages are scaled according to the size of the village population, thus showing the demand pressures on each water point. The map also contains information on road infrastructure, including the quality of roads, and on the hydrographic networks.

We had to limit our study to a smaller data set, because not all data were available for all villages at the time that the data were collected for inclusion in the GIS. In particular, the data obtained from the Ministry of Water Management were limited to 25 provinces, or 5,207 of the country's 6,821 villages. Data for the remaining five provinces were subsequently collected from the Ministry in another survey, but very few variables were comparable between the two surveys. In some villages data on other variables were also missing, and thus we had to reduce the number of villages in the final prediction analysis to 3,871, or 57 percent of the country's total number of villages (descriptive statistics are given in tables 2 and 3).

The lack of sufficient data for all 6,821 villages is, of course, a cause for concern. Some of the missing data are available in the archives of the different ministries and, in principle, could be retrieved. However, if a significant number of villages still do not have all the necessary data, targeting will have to be made at higher levels of geographic aggregation (at the department or province level). To target at these higher levels, we still would have to use the predictions of village-level poverty obtained for all villages outside the sample; we cannot use the household survey directly because the sampling frame and the sample size do not adequately represent all departments and provinces.

Table 2. *Descriptive Statistics on Variables Used in the Estimation*

Aggregation level	*Variable*	*Urban* *Mean*[a]	*Urban* *Standard error*[a]	*Urban* *Number of observations*	*Rural* *Mean*[a]	*Rural* *Standard error*[a]	*Rural* *Number of observations*	*Data source*
Household	Children 0–6 years per adult (15–50 years) in household	0.530	0.495	2,671	0.779	0.598	5,508	Priority Survey
Household	Children 7–14 years per adult in household	0.618	0.590	2,671	0.748	0.640	5,508	Priority Survey
Household	Elderly persons (50+) per adult in household	0.183	0.343	2,671	0.313	0.426	5,508	Priority Survey
Household	Literate head in household	0.477	0.499	2,736	0.134	0.341	5,906	Priority Survey
Household	Percentage male adults literate in household	0.562	0.422	2,736	0.177	0.313	5,906	Priority Survey
Household	Percentage female adults literate in household	0.373	0.397	2,736	0.053	0.174	5,906	Priority Survey
Household	Livestock units per capita	0.123	0.909	2,736	0.442	0.943	5,906	Priority Survey
Village	Distance to nearest rural primary school	n.a.	n.a.	n.a.	2.290	5.640	4,412	Ministry of Water Management and Infrastructure
Village	Teachers per child ages 7–14 years	0.014	0.002	2,736	0.005	0.006	5,760	Ministry of Education
Village	Distance to nearest health facility	n.a.	n.a.	n.a.	4.790	7.770	4,434	Ministry of Water Management and Infrastructure

Village	Nearest facility has safe water	0.820	0.390	2,416	0.034	0.180	4,434	Ministry of Water Management and Infrastructure
Village	Number of pumps per rural community	n.a.	n.a.	n.a.	7.350	10.640	5,241	Ministry of Water Management and Infrastructure
Village	Presence of an all-weather road	n.a.	n.a.	n.a.	0.570	0.500	4,434	Ministry of Water Management and Infrastructure
Department	Cultivated area in department per capita	0.211	0.221	2,736	0.507	0.301	5,760	Famine Early Warning System
Department	Average rainfall, 1980–94	65.800	10.070	2,736	62.500	14.840	5,760	Directorate of Meteorology
Department	Absolute value of deviation of rainfall from average, 1994	19.450	14.490	2,736	22.580	12.960	5,760	Directorate of Meteorology
Department	Average length rainy season, 1982–92	9.520	1.340	2,736	9.530	2.000	5,760	Famine Early Warning System
Department	Average vegetation index, 1982–92	0.114	0.034	2,736	0.136	0.051	5,760	Famine Early Warning System
Department	Homogeneity of rainy season, 1982–92	0.162	0.019	2,736	0.161	0.036	5,760	Famine Early Warning System

n.a. Not applicable.
Note: For village- and department-level variables, the same value is assumed for all households in the community.
a. Weighted using sampling weights.
Source: Authors' calculations.

Table 3. *Descriptive Statistics on Variables Used in the Prediction*

Aggregation level	*Variable*	*Urban*			*Rural*			*Data source*
		Mean[a]	*Standard error*[a]	*Number of observations*	*Mean*[a]	*Standard error*[a]	*Number of observations*	
Village	Children ages 0–6 years per adult (15–50 years) in household	0.656	0.110	300	0.645	0.227	6,818	National Census
Village	Children ages 7–14 years per adult in household	0.593	0.120	300	0.563	0.280	6,818	National Census
Village	Elderly persons (50+) per adult in household	0.320	0.076	300	0.348	0.351	6,818	National Census
Province	Literate head in household	0.450	0.181	191	0.113	0.075	6,711	Priority Survey
Province	Percentage male adults literate in household	0.522	0.147	191	0.141	0.079	6,711	Priority Survey
Province	Percentage female adults literate in household	0.323	0.149	191	0.044	0.034	6,711	Priority Survey
Province	Livestock units per capita	0.147	0.090	191	0.492	0.263	6,711	Ministry of Agriculture
Village	Distance to nearest rural primary school	n.a.	n.a.	n.a.	4.390	5.040	4,556	Ministry of Water Management and Infrastructure
Village	Teachers per child ages 7–14 years	0.023	0.032	295	0.003	0.011	4,753	Ministry of Education
Village	Distance to nearest health facility	n.a.	n.a.	n.a.	6.790	7.460	4,393	Ministry of Water Management and Infrastructure

Village	Nearest facility has safe water	0.150	0.350	219	0.005	0.073	4,390	Ministry of Water Management and Infrastructure
Village	Number of pumps per rural community	n.a.	n.a.	n.a.	2.350	2.765	5,425	Ministry of Water Management and Infrastructure
Village	Presence of an all-weather road	n.a.	n.a.	n.a.	0.430	0.500	4,618	Ministry of Water Management and Infrastructure
Department	Cultivated area in department per capita	0.669	0.605	300	0.751	0.717	6,821	Famine Early Warning System
Department	Average rainfall, 1980–94	65.520	15.340	300	69.160	16.340	6,821	Directorate of Meteorology
Department	Absolute value of deviation of rainfall from average, 1994	18.900	11.540	300	18.610	13.950	6,520	Directorate of Meteorology
Department	Average length rainy season, 1982–92	10.190	2.310	300	10.770	2.440	6,520	Famine Early Warning System
Department	Average vegetation index, 1982–92	0.126	0.054	300	0.121	0.051	6,821	Famine Early Warning System
Department	Homogeneity of rainy season, 1982–92	0.152	0.038	300	0.153	0.036	6,821	Famine Early Warning System

n.a. Not applicable.
Note: For department-, province-, and village-level variables, the same value is assumed for all households in the community.
a. Weighted using total population relative to village population.
Source: Authors' calculations.

Similar to the LSMS surveys, the sampling for the Priority Survey used to estimate the consumption model was semistratified (INSD 1996). The survey was designed to be representative at both the national and regional levels. The country was divided into seven regions: five rural regions representing five agroclimatic areas and two urban regions, one comprising Ouagadougou and Bobo-Dioulasso, Burkina Faso's two main cities, and the other comprising the remaining cities. From the seven regions 434 enumeration areas were selected on the basis of their socioeconomic characteristics. In each, 20 households were randomly selected. For our econometric estimation, however, we had to reduce the sample size to 5,618 households and 201 enumeration areas because of missing data for five provinces and incomplete data for certain variables in a few other villages.

II. METHODOLOGY

The econometric analysis in this article has two parts. The first estimates a prediction model for household consumption, using the household data of the Priority Survey and the community data from all other sources, in order to determine the variables that best explain household consumption and poverty. The explanatory variables we select for the prediction model include only those for which we have data for all villages outside the Priority Survey sample. In the second part of the analysis we use the prediction model and the village-level data from the GIS database to measure welfare at the village level for all villages outside the Priority Survey sample. In line with similar studies on this subject, we use consumption per standard adult (adult equivalent) as our welfare indicator at the household level and use the headcount index as the measure of poverty.[3]

The Two Models

Let c_{ij} denote the level of consumption per standard adult in household i located in community j. Let z denote the poverty line, and let $y_{ij} = c_{ij}/z$ be the normalized welfare indicator per standard adult. We conduct the analysis in terms of the natural logarithm of y_{ij}. For a poor person, therefore, $y_{ij} < 1$, or $\ln y_{ij} < 0$. The headcount index, H_j, which measures the relative size of the poor population in community j, is equal to the mean value of the individual poverty indicators, H_{ij}, which indicate the probability that the household ij is poor, over all the individuals in that community. The individual poverty indicator is determined by the normalized welfare function as follows:

$$H_{ij} = 1 \text{ if } \ln y_{ij} < 0 \qquad (1)$$
$$H_{ij} = 0 \text{ if } \ln y_{ij} \geq 0.$$

In constructing the prediction model, we represent the individual welfare indicator as a function of a vector of household and community explanatory variables $\mathbf{X}_{ij}$ and a residual term u_{ij}, which is assumed to be normally distributed with

3. We use simple nutritional adult equivalence scales. These are 0.3 for a child less than 5 years old and 0.7 for a child between 5 and 15 years. Each adult counts as 1.0.

$u_{ij} \sim N(0,\sigma_j^2)$, thereby allowing for village-level heteroskedasticity. The prediction model is thus given by:

$$\ln y_{ij} = \boldsymbol{\beta}' \mathbf{X}_{ij} + u_{ij}. \tag{2}$$

As noted earlier, we only select explanatory variables if their mean values are available for all villages in the GIS database. They include community characteristics and mean household characteristics, such as household composition and literacy rates, for all households in the community. We can estimate equation 2 using the maximum likelihood method, with $u_{ij} \sim N[0, \sigma^2 \exp(\gamma \mathbf{X}_j^V)]$, where $\mathbf{X}_j^V$ are the mean values of the explanatory variables in community j. This formulation corrects for heteroskedasticity and generates the estimators $\mathbf{b}$ and s_j of the parameters $\boldsymbol{\beta}$ and σ_j.

We use these estimators and the set of explanatory variables to predict a community's mean consumption for all communities outside the Priority Survey sample. Mean consumption, however, is not necessarily a good predictor of poverty, since the poverty measure is a function of not only mean consumption, but also the distribution of consumption within the community. The term s_j represents one part of that distribution, since the within-community variance is the sum of the variance of the regression and the deviation of predicted household consumption from predicted mean consumption.[4]

Using these estimators and the set of explanatory variables, a consistent estimate of the probability that household ij with characteristics $\mathbf{X}_{ij}$ is poor can then be expressed as:

$$E(H_{ij} \mid \mathbf{X}_{ij}, \mathbf{b}, s_j) = \text{Prob}\,(u_i < -\mathbf{b}'\mathbf{X}_{ij}) = \Phi\,(-\mathbf{b}'\mathbf{X}_{ij} / s_j) \tag{3}$$

where $\Phi(.)$ is the cumulative normal distribution. The predicted incidence of poverty in community j is determined from equation 3 as:

$$E(H_j \mid \mathbf{X}_{ij}, \mathbf{b}, s_j) = E\,[\Phi\,(-\mathbf{b}'\mathbf{X}_{ij} / s_j)]. \tag{4}$$

If complete information on $\mathbf{X}_{ij}$ were available for all households and all villages in the country, this prediction would be fairly straightforward: we would use equation 3 to estimate the probability that each household in the village is poor and equation 4 to predict the incidence of poverty in the community—across all villages outside the sample.[5] However, in Burkina Faso the only data available for all villages outside the Priority Survey sample are the *mean* values of the explanatory variables $\mathbf{X}_j^V$ in each community. Since equation 4 is nonlinear, we cannot use these variables simply to predict village-level poverty. But we can use Taylor expansions to obtain an approximation. Thus we expand

4. The within-village variance of consumption can be written as $E\{[Y_{ij} - E(Y_j)]^2\} = E[(\mathbf{b}'\mathbf{X}_{ij} - \mathbf{b}'\mathbf{X}_j^V)^2] + s_j^2$, in which Y_j is the mean level of consumption in the village. In words, the variance of consumption is the sum of the squared deviation of predicted household consumption from predicted mean consumption in each village and the village-level variance of the prediction model.

5. Hentschel and others (1998) use this property to predict regional poverty from census data.

equation 4 around $-\mathbf{b}'\mathbf{X}_j^V/s_j$. Using the property that $E(\mathbf{b}'\mathbf{X}_{ij} - \mathbf{b}'\mathbf{X}_j^V) = 0$, we obtain:

$$(5) \qquad E(H_j) = E[\Phi(-\mathbf{b}'\mathbf{X}_{ij}/s_j)]$$
$$\approx \Phi(-\mathbf{b}'\mathbf{X}_j^V/s_j) + \tfrac{1}{2}(\mathbf{b}'\mathbf{X}_j^V/s_j^3)\phi(-\mathbf{b}'\mathbf{X}_j^V/s_j)\, E(\mathbf{b}'\mathbf{X}_{ij} - \mathbf{b}'\mathbf{X}_j^V)^2$$

where $\phi(.)$ is the normal density function and $E(\mathbf{b}'\mathbf{X}_{ij} - \mathbf{b}'\mathbf{X}_j^V)^2$ is the variance of predicted household consumption around predicted mean consumption within each village (Maddala 1983). In words, the predicted level of poverty for villages outside the sample is a function of the mean level of consumption per adult and the variance around that mean. Equation 5 can therefore predict the incidence of poverty in communities outside the sample, using the estimated parameters of the household consumption function ($\mathbf{b}$ and s_j) and the community-level characteristics $\mathbf{X}_j^V$ of the villages outside the sample.

Empirical Estimation

We use the regression analysis to predict consumption levels for all households rather than to determine whether or not a household is poor. The latter approach would be equivalent to estimating equation 1 directly. That approach is often referred to as a multivariate poverty profile (Ravallion 1996). The individual poverty indicator in equation 1 is binary, so one could use a probit (or an alternative) model to construct a prediction model. As Ravallion points out, a puzzling feature of this approach is that the estimation techniques used typically were developed for situations in which the observed data were dichotomous or truncated at zero, even though consumption was observed.

The standard way of solving this estimation problem is to define a regression model in which a continuous latent (unobserved) variable is regressed on a set of observed explanatory variables (Maddala 1983). A particular error structure (such as the normal distribution for the probit) is then assumed, allowing the parameters of interest to be estimated. These parameters can then be used for inference related to the explanatory variables and the observed limited dependent variable. If this procedure is used on a poverty indicator, such as the headcount index, then the latent variable is an observed variable that was used to calculate the limited dependent variable. Since the latent variable is observed, limited dependent variable estimation of the poverty indicator is not necessary and would be less efficient, since some information actually available (consumption) is not used in constructing the prediction model.

To estimate household consumption, we use a standard reduced-form framework in which income (measured in terms of household consumption) is regressed on household characteristics, including human and physical capital, as well as on community characteristics.[6] Some community characteristics are specified at the village level, whereas others, primarily the agroclimatic conditions, are specified at the department or regional level.

6. Examples of this approach are Glewwe and Kanaan (1989) and Coulombe and McKay (1996). Glewwe (1991) has a useful discussion on the justification for including particular variables in this type of approach. We return to the problems related to this specification below.

The Priority Survey contains a limited but important set of variables that can be used to explain household consumption. We select household-level explanatory variables that allow aggregation at the community level and thus can be used for the prediction model. This limited the choice of household-level variables to those for which the corresponding mean values at the community level are available for all communities in the country. As a consequence, we could not include in the estimation several variables, such as the education of household members (as opposed to the literacy of the household head), that usually are found to be significant in a consumption model. Furthermore, data on household assets and landholdings, which are also significant explanatory variables in most consumption models, are not available in the Burkina Faso Priority Survey. This reduced the explanatory power of the model and very likely created a missing variable problem. Moreover, since the underlying data are cross-sectional, household heterogeneity—a common problem affecting any regression of welfare indicators—is also difficult to address. Despite these reservations, we were able to collect data on most of the important explanatory variables, such as demographic variables and variables representing human and physical capital, and include them in the model (see table 2).

We also include village- and department-level variables. Department-level variables are primarily climatic data and means of certain household variables (such as the average area of cultivated land across all households in the department), which we obtained from the Ministry of Agriculture. We distinguish between the impact of long-term climatic characteristics and the impact of temporary fluctuations by including among the explanatory variables the average level of rainfall in the past 15 years and the absolute value of the deviation of the previous year's rainfall from the long-term average. The village-level explanatory variables also include data on the distance to and quality of schools and health facilities, the quality of access roads, and the quality of the water supply.

III. Estimating Poverty within the Sample

Descriptive statistics on poverty and consumption reveal large differences in the incidence of poverty between urban and rural households and a much higher standard of living in the country's two main cities (table 4). Among rural areas, the West has the lowest rates of poverty. Among the other regions, differences between poverty rates are small.

In the econometric analysis we regress consumption per adult equivalent on the explanatory variables (listed in table 2), according to the linear model given by equation 2. We estimate this model using the maximum likelihood method, in which the regression coefficients and the heteroskedastic errors are estimated jointly (tables 5 and 6).[7] In allowing for heteroskedasticity by community, we can use the community-level information to predict mean consumption per equiva-

7. The regression was weighted with individual sampling weights derived from the original sampling frame used by the World Bank and INSD.

Table 4. *Poverty and Consumption in Burkina Faso*

Region	*Consumption per adult equivalent (Francs CFA per month)*	*Headcount index*[a]
Rural		
West	7,573	0.56
South/South-East	5,699	0.67
Central-North	4,952	0.74
Central-South	5,240	0.75
North	6,122	0.64
Bobo-Dioulasso and Ouagadougou	20,768	0.14
Other urban	12,173	0.39
National	8,766	0.58

a. The poverty line is set at two-thirds of mean consumption.
Source: Priority Survey (1994).

lent adult as well as the variation around this mean. The estimated variance within a community may provide some information on the extent of inequality in the distribution of consumption. We conduct the regression analysis separately for households in rural and urban areas (pooling tests convincingly reject running one national regression). For both urban and rural areas multiplicative heteroskedasticity cannot be rejected at the 1 percent level (table 6).[8]

Because we use maximum likelihood estimation to jointly determine the coefficients in the model and the structure of heteroskedasticity, no simple R^2 can be reported. However, the ordinary least squares (OLS) estimates of the model (derived in the first step) indicate that the adjusted R^2s are low, equal to 0.28 for the urban population and 0.17 for the rural population. These figures are low primarily because of restrictions on the choice of variables included in the model. When we use all the household- and community-level variables that were available in the Priority Survey in the estimation, the R^2 in the regression for rural households rises to 0.50. Because of the low values of the adjusted R^2s, we need to make significant adjustments in applying the results in the prediction model. These adjustments are discussed in section IV.

Interpreting the Results

The results show that the variables included in the model are strongly jointly significant and that a substantial number of household- and community-level variables are highly significant. The following results stand out.

The household variables that are correlated most closely with the level of consumption in both rural and urban areas are the adult literacy rates. (The variables describing adult literacy also include literacy of the household head.) The

8. The Breusch-Pagan Lagrange-multiplier test convincingly rejects homoskedasticity (see table 6). The Glesjer (1965) test indicates that in both urban and rural areas the null hypothesis of multiplicative heteroskedasticity cannot be rejected at the 1 percent level.

Table 5. *Regression Results: Poverty and Consumption*

	Rural		*Urban*	
Variable	*Coefficient*	*t-value*	*Coefficient*	*t-value*
Constant	7.71	52.48*	10.82	21.99*
Children 0–6 years per adult (15–50 years) in household	0.02	1.55	0.01	0.40
Children 7–14 years per adult in household	–0.03	–1.67***	–0.04	–1.60
Elderly persons (50+ years) per adult in household	0.03	1.24	–0.13	–2.91*
Literate head in household	0.18	3.66*	0.33	7.63*
Percentage male adults literate in household	0.13	2.48**	0.16	3.18*
Percentage female adults literate in household	0.55	8.41*	0.42	10.11*
Livestock units per capita(/10)	0.93	11.06*	0.31	1.49
Distance to nearest rural primary school (/100)	–0.48	–2.80*		
Teachers per child ages 7–14 years (*10)	0.21	1.07	–0.39	–0.22
Distance to nearest health facility (*100)	0.18	1.58		
Nearest facility has safe water	0.14	2.92*	0.92	5.24*
Number of pumps per rural community (/100)	0.34	3.27*		
Presence of all-weather road	0.10	4.83*		
Cultivated area in department per capita	0.01	0.32	1.66	5.29*
Average rainfall, 1980–94 (/100)	0.53	3.39*	–1.64	–2.28**
Absolute value of deviation of rainfall from average, 1994 (/100)	–0.22	–2.78*	–0.44	–2.03**
Average length of rainy season, 1982–92	–0.01	–0.69	–0.06	–0.98
Average variable vegetation index, 1982–92	–0.54	–1.81***	8.17	3.24*
Homogeneity of rainy season, 1982–92	2.50	8.38*	–12.26	–3.63*
F-joint significance regression	34.58* *F*[19, 4107]		53.70* *F*[15, 2346]	
Number of valid observations	4,119		2,362	

***Significant at the 10 percent level.
**Significant at the 5 percent level.
*Significant at the 1 percent level.
Note: Dependent variable is log of consumption per standard adult.
Source: Authors' calculation.

dependency rates, namely the number of children and elderly persons per adult in the household, do not have a clear effect on consumption in rural households. But in urban households the number of elderly persons per adult in the household has a significant impact on the level of consumption per adult equivalent.

The number of livestock units per capita—the only proxy for the household's physical assets available in the Priority Survey—is significantly and positively correlated with consumption in rural areas. The community-level variables that characterize agroclimatic conditions also have a strong impact on consumption in rural areas. Consumption per capita typically is higher in rural areas that have relatively high levels of long-run average rainfall, relatively normal rain in the survey year, and low rainfall variability over the rainy season.

However, the agroclimatic variables representing the average level of rainfall and the homogeneity of the rainy season seem to have a negative effect on consumption in urban areas. A possible explanation is that the consumption basket

Table 6. *Regression Results: Estimated Variance with Multiplicative Heteroskedasticity*

	Rural		Urban	
Variable	*Coefficient*	*t-value*	*Coefficient*	*t-value*
Constant	0.12	5.03*	0.03	1.75***
Children 0–6 years per adult (15–50 years) in household (community mean)	0.40	2.85*	–0.86	–3.68*
Children 7–14 years per adult in household (community mean)	0.79	6.27*	1.12	5.08*
Elderly persons (50+) per adult in household (community mean)	–0.29	–1.64***	0.47	1.53
Literate head in household (community mean)	0.49	3.86*	0.09	0.91
Percentage male adults literate in household (community mean)	–0.26	–1.96**	0.12	1.12
Percentage female adults literate in household (community mean)	0.11	0.90	0.05	0.70
Livestock units per capita (community mean)	–0.05	–0.94	0.00	0.04
Distance to nearest rural primary school	0.00	0.07		
Teachers per child ages 7–14 years (*100)	0.20	4.70*	2.00	4.95*
Distance to nearest health facility	0.00	–0.97		
Nearest facility has safe water	–0.38	–3.28*	–1.47	–3.63*
Number of pumps per rural community	0.01	2.16**		
Presence of all-weather road	0.20	4.03*		
Cultivated area in department per capita	0.33	3.05*	–1.03	–1.33
Average rainfall, 1980–94	0.02	4.20*	0.02	1.38
Absolute value of deviation from average, 1994	0.00	1.14	0.00	–0.38
Average length rainy season, 1982–92	0.09	2.78*	0.00	–0.01
Average variable vegetation index, 1982–92 (*10)	0.31	4.36*	–3.19	–5.56*
Homogeneity of rainy season, 1982–92 (*10)	–0.16	–2.25**	4.35	5.65*
Breusch-Pagan LM heteroskedasticity	603.35** (19 degrees of freedom)		158.33 (15 degrees of freedom)	
Glesjer test multiplicative heteroskedasticity	3.66* $F[19, 4107]$		3.66* $F[15, 2328]$	

***Significant at 10 percent level.
**Significant at 5 percent level.
*Significant at 1 percent level.
Note: A positive coefficient on the explanatory variable indicates that this variable has the effect of raising the variance; a negative coefficient indicates that this variable has the effect of lowering the variance.
Source: Authors' calculations.

of urban households typically includes commodities that were not recorded in the Priority Survey (which includes only a small number of consumption items); the reduction in consumption during the normal years recorded in the survey is therefore spurious. Another possible explanation is that these variables are correlated with significant missing variables that have a negative impact on the consumption of urban households. The data we had at our disposal did not allow us to analyze these effects further.

In rural areas consumption in villages that are farther away from schools is generally lower. No similar effect is revealed with respect to the distance to health facilities. One explanation may be that in some regions villages located farther from a health facility receive services from mobile health clinics. In both urban and rural areas the quality of services in the health facility—approximated by the availability of safe drinking water in the facility—is significantly correlated with the level of per capita consumption in the surrounding villages and urban neighborhoods. Only about one-third of the health facilities in Burkina Faso have safe drinking water.

In rural areas the quality of infrastructure, indicated by the availability of safe drinking water in the village (measured by the number of functioning pumps), and the quality of access roads to the village have a significant and positive impact on consumption. Mean consumption in villages with an all-weather access road is nearly 10 percent higher than in villages without one. The greater opportunity to trade, rather than produce for own-consumption, and the better alternatives for nonagricultural work that access to an all-weather road provides are the main reasons for this effect.

The coefficients that determine the pattern of the village-level error terms indicate that in villages where a relatively high proportion of household heads are literate, the distribution of per capita consumption is less equal than in villages where a low proportion of household heads are literate. It may be that in a village with more literate adults the income differences between households with less educated heads and households with more educated heads are relatively larger.

Households in villages with relatively high average levels of rainfall differ more widely in per capita consumption, possibly because in these villages some households are better equipped and more able to take advantage of superior agricultural conditions. Villages with higher average landholdings per household show larger variability in per capita consumption.

Accounting for Endogeneity

Several of the interpretations of the results suggest possible causal relationships between the explanatory variables and the dependent variable. However, these interpretations are intended primarily as background for a more thorough evaluation of the possible policy implications. We make them with the usual caveat of potential endogeneity of the community-level variables, which means that correlation need not imply causality. Thus, for example, the government's policy of locating more public education facilities in poorer villages as part of its antipoverty program will lead to a high negative correlation between average per capita consumption and the proximity of the village to a school (Rosenzweig and Wolpin 1986). To take another example, the availability of a large number of water pumps in a village need not be the cause of a relatively high standard of living, but rather may be the result of a higher demand for safe drinking water among more affluent villagers.

The purpose of our regression estimates is, however, to construct a prediction model that can identify the poorest and wealthiest villages. The quality of such predictions depends only on the degree of correlation between the explanatory and the dependent variables, irrespective of whether or not the correlation indicates causality. If, for example, health facilities are intentionally placed in poorer villages, then the distance from a village to a health facility can be useful for predicting the standard of living.[9] Nevertheless, the possibility of endogeneity forces us to use special care in interpreting the results for policy purposes.[10] Although the significance and size of the coefficients are suggestive, additional work is needed to design appropriate policies for reducing poverty.

IV. Predicting the Geographic Distribution of Poverty

In the next step we apply the regression results (obtained for a sample of communities) to the data available in the GIS database in order to predict the distribution of poverty across all communities in Burkina Faso. We focus on the 3,871 villages, out of a total of more than 6,000, for which all the necessary information was available. We calculate the headcount index of a community from equation 5, using the parameter estimates that were obtained in the regression analysis. To use this equation, we also need estimates of mean consumption per adult in the community and the variance of consumption. For the term $E(\mathbf{b}'\mathbf{X}_{ij} - \mathbf{b}'\mathbf{X}_j^V)^2$ in equation 5 we use the average value per region (rather than per community) in the survey data, and we obtain s_j using the coefficients given in table 6. We predict mean consumption per standard adult in the communities outside the sample from the mean values of the explanatory variables for each of these communities, using the coefficients given in table 5.

Before applying these predictions for all communities outside the sample, we assess their quality by comparing them with direct estimates of poverty from the sample of 201 communities included in the Priority Survey. We derive the correlation coefficients of the predicted and calculated level of poverty for these villages. The value of the Pearson-correlation coefficient is 0.51, and it is strongly statistically significant. For policy decisions, however, the more relevant criterion is the order of villages along the poverty scale. To test this aspect of the prediction model, we calculate the Spearman rank correlation coefficient between the order established by the direct estimates of poverty and the order established by the model's predictions. That coefficient is also strongly significant at 0.43.

Another way to test the quality of the predictions is to compare the estimated and the predicted values of the headcount index. We do this for selected communities in three provinces (table 7). Although the predicted values of the headcount index in each community often fall outside the confidence interval of the calcu-

9. This assumption also requires that the same program placement rule be used outside the sample as inside the sample. Since the sample is nationally representative, this may be an appropriate assumption.

10. There are other possible sources of endogeneity. For example, we assume that location is not a choice variable. We therefore do not consider migration explicitly.

Table 7. *Comparison of Predictions and Direct Estimates of the Headcount Measure of Poverty for Sample Villages in Three Provinces*

Province	*Village identification number*	*Within-sample estimates*[a]	*Outside-sample predictions*
Kossi	4426	0.28 (0.01)	0.24
	3786	0.33 (0.01)	0.67
	512	0.54 (0.01)	0.64
	2936	0.54 (0.04)	0.65
	5266	0.57 (0.02)	0.56
	5117	0.64 (0.02)	0.69
	1626	0.68 (0.02)	0.70
	1556	0.69 (0.01)	0.61
	1290	0.78 (0.01)	0.70
	250	0.80 (0.01)	0.57
Kouritenga	744	0.64 (0.03)	0.66
	6233	0.72 (0.03)	0.66
	657	0.75 (0.01)	0.57
	1627	0.80 (0.03)	0.74
	2943	0.83 (0.01)	0.65
	3213	1.00 (0.00)	0.64
	3828	1.00 (0.00)	0.76
Mouhoun	1278	0.23 (0.01)	0.32
	790	0.36 (0.02)	0.52
	6753	0.48 (0.03)	0.57
	4982	0.48 (0.02)	0.56
	740	0.50 (0.02)	0.52
	5149	0.57 (0.02)	0.52
	5635	0.72 (0.01)	0.60
	6674	0.72 (0.01)	0.65

a. Standard errors are in parentheses (from Deaton 1997:47). The figures are weighted by household size.
Source: Authors' calculations.

lated measure, the rank order of communities from richest to poorest in each province is similar.

Despite these results, the low R^2 values in the regression analysis and the poor quality of the data prevent us from using these predictions directly. Moreover, these predictions rely on the assumption of normality of the error term. One common test for normality is the Jarque-Bera test. Our estimate of the Jarque-Bera statistic is 11.8, and we therefore have to reject the normality hypothesis.[11] As a result, we do not use the prediction to establish a complete order of communities on the poverty scale. Instead, we divide the 3,871 villages and urban communities into four categories of poverty, ranging from the poorest to the wealthiest, according to predicted levels of poverty. Despite the errors in these predictions, our results suggest that most of the villages categorized as "poorest"

11. This statistic has a chi-square distribution with two degrees of freedom. The normality hypothesis has to be rejected at a probability of 0.997. However, the Jarque-Bera test is not robust to the presence of heteroskedasticity, which could not be rejected by the Breusch-Pegan LM test and the Glesjer test. We are not aware of a test of normality in the presence of heteroskedasticity, but the high value of the Jarque-Bera statistic suggests that it is highly probable that the residuals are not distributed normally.

are likely to have a higher incidence of poverty than most of the villages categorized as "least poor." The villages in the poorest category are therefore candidates for targeted poverty alleviation programs, and the villages in the least poor category are candidates for cost-recovery programs.

Given the data limitations in Burkina Faso, effective targeting would have to focus only on these two extreme categories in order to reduce leakage as much as possible and keep within the budget constraint. We can further improve targeting in the present circumstances—given the limited availability and poor quality of the data—by dividing the villages into a larger number of categories and concentrating on villages in the two extreme categories. Future research intended to improve targeting will have to focus, however, on efforts to improve the quality of the data as well as generate additional series of geo-referenced data.

We construct the geographic distribution of rural and urban communities across these categories of well-being within each province (table 8). We divide the villages into the four categories, ranging from poorest to least poor, using the predicted values of poverty incidence and allocating the entire population in each village to the corresponding category. We set the categories so that the popula-

Table 8. *Distribution of Communities by Poverty Category*
(percent)

	Poverty category				
Province	*Poorest*	*Lower-middle*	*Upper-middle*	*Least poor*	*Percentage of total population*
Bam	0	10	36	54	0.90
Bazega	13	36	35	16	4.85
Boulgou	37	22	32	9	7.04
Boulkiemde	41	36	15	7	5.99
Ganzourgou	47	33	19	2	2.74
Gnagna	24	13	21	41	3.73
Gourma	34	17	23	25	5.55
Kossi	26	30	29	14	6.22
Kouritenga	7	25	43	25	3.11
Mouhoun	41	39	18	2	6.06
Nahouri	23	29	40	8	1.71
Namentenga	4	25	18	52	2.16
Oubritenga	9	24	40	27	4.73
Oudalan	24	36	19	21	1.69
Passore	11	2	18	68	3.77
Sanguie	31	23	25	21	4.64
Sanmatenga	17	11	25	46	7.64
Seno	26	29	19	26	4.70
Sissili	41	43	16	0	3.69
Soum	3	14	21	61	2.54
Sourou	12	12	12	64	5.31
Tapoa	60	17	22	1	2.12
Yatenga	24	24	31	2	6.72
Zoundweogo	3	59	30	8	2.40

Note: The poverty line is set at two-thirds of mean consumption.
Source: Authors' calculations.

tion in each represents 25 percent of the country's total population. The distribution of the population within provinces is significantly different, however. For example, 41 percent of the population in the province of Boulkiemde lives in villages classified as poorest, and only 7 percent of the population lives in villages classified as least poor.

Consider, as an illustration, an antipoverty program targeted to the five provinces in which at least 40 percent of the population resides in villages classified as poorest. Under this criterion 21 percent of the country's population will be included in the target provinces. Only 3 percent of the population in the five target provinces (which account for only 0.6 percent of the country's total population), however, lives in villages classified as least poor, suggesting that leakage is likely to be small. Nearly 43 percent of the population in the five target provinces lives in villages classified as poorest, accounting for 36 percent of the country's population that lives in the poorest villages. At the other extreme, a cost-recovery program that is targeted to the seven provinces in which more than 40 percent of the population lives in villages classified as least-poor will cover 26 percent of the total population but only 13 percent of the population that lives in the poorest villages.

Targeting antipoverty programs at the province level, however, is likely to be less effective than targeting at the village level. The reason is that targeting provinces is bound to include villages of the higher categories in which the incidence of poverty is likely to be less than that in villages of the lowest category. Therefore, under a given budget constraint, a targeted program at the village level that focuses only on villages categorized as poorest is likely to cover a larger share of the country's poor population and entail less leakage, despite the prediction errors in classifying villages.

Most of the urban communities are classified as least poor, largely because of the much higher standard of living and much lower incidence of poverty in urban areas. There are also several other, more technical, explanations. In urban areas the distinction between poor and nonpoor communities is less clear than in rural areas. In many developing countries it is not uncommon for poor households to reside in relatively affluent urban communities and vice versa. Further, urban communities, as defined in the household surveys, are, in fact, enumeration areas that have been demarcated by local authorities for administrative purposes, and their borders are often quite arbitrary. Whereas in rural areas enumeration areas are generally limited to one or two neighboring villages that have similar living standards, in urban areas, where the distance between neighborhoods is small, enumeration areas often include communities with widely different living standards. In this study we had access to community-level data in urban areas only in the household survey.

Another reason is that in all other data sources the towns, including Ouagadougou and Bobo-Dioulasso, were considered as single points in the GIS data set. In the econometric analysis all the enumeration areas from each of the large towns have the same community characteristics and thus have to be considered as a single entity.

Of the villages in the Province of Sanguie, most of those considered poorest are located farthest from urban centers and are not connected to an all-weather road (figure 2, in the appendix). Targeting an antipoverty program to the entire population of the province is bound to include many nonpoor villages, whereas excluding this province from the program will leave out a considerable number of poor villages.

V. Simulating the Impact of a Village-Level Targeting Scheme

To evaluate the effectiveness of community targeting, we conduct a simple simulation experiment. We use the consumption data in the 201 communities for which we have complete information from the Priority Survey. The simulation design follows closely the framework of Baker and Grosh (1994). We assume that the government has a given budget for transferring income to the target population. The effect of these transfers on poverty is evaluated using the actual household consumption data. We select which villages to target, however, using the predicted levels of poverty estimated from our model. The simulations thus can evaluate how effectively these predictions identified the poor by estimating leakage and undercoverage. We make the estimates for the households included in the survey, but we use the individual sampling weights to measure the impact on the total population at the regional and national levels. The reliability of the regional and national results is therefore affected by the sampling errors that are due to the sampling frame of the Priority Survey.

We compare the outcome of this simulation with an untargeted uniform transfer scheme in which all individuals in the country receive a transfer. We also consider two other targeted programs: a village-level targeted program that uses actual poverty levels to identify the poor villages included in the program and a "perfect" targeting program that uses actual household consumption data to identify the poor villages included in the program.

The three targeted programs are designed to include 30 percent of the population.[12] To achieve this, we first set the poverty line so that 30 percent of the country's population is identified as poor. We select villages for the three targeted programs as follows. For the program that targets at the village level based on the predicted level of poverty, we rank all villages in the sample. Starting from the poorest village, we select villages for targeting until (at least) 30 percent of the population is included in the program. We follow the same procedure for the program that targets at the village level based on the actual level of poverty. For the program targeted at the household level based on actual consumption per

12. In our simulation we use a lower poverty line than in the previous section to focus on attempts to target a relatively small part of the population. Using two-thirds of mean consumption as the poverty line, poverty is estimated at 58 percent, suggesting a transfer program that attempts to include nearly two-thirds of the population. The issues of undercoverage and leakage thus become less interesting to study. The population considered in this simulation includes the households for which we have the complete set of variables in the prediction model.

adult equivalent, we again rank households, this time according to consumption level. We include households starting from the poorest until 30 percent of the population is covered.

We first consider a targeted program using actual village poverty levels.[13] The simulation results suggest that 44 percent of the poor would not be covered. With the targeted program that uses predicted village poverty levels, undercoverage would rise to 56 percent as a result of prediction errors. As an indicator of the accuracy of the predictions, this implies that about 79 percent of the poor who could be reached through targeting using actual poverty data could also be reached using predicted poverty data.

By design, untargeted transfers result in no undercoverage, but leakage is high. We define leakage as the number of nonpoor covered by a program divided by the total number of people covered. For the untargeted program leakage is (by design) 70 percent, since we consider 30 percent of the population to be poor. When using the village-level poverty estimates based on the Priority Survey data to establish criteria for targeting, leakage falls to 44 percent. The distribution of leakage across (rural) regions is similar. When using the predictions on village-level poverty to select the communities included in the scheme, leakage increases to 56 percent. Still, this amount is far less than the leakage implied by undifferentiated transfers.[14]

VI. Conclusion

Geographical targeting of antipoverty programs can be an effective way to reach the poor and contain program costs in countries where information on individual households is incomplete or unavailable, making individual or household targeting impossible. By identifying the geographic areas where the poor are concentrated, these programs can reduce leakage so that a larger share of the poor population can be reached with a given budget or a larger share of this budget can reach the poor. However, in most countries the target areas are regions, states, or entire rural areas. Although targeting at these levels can offer considerable savings compared with nontargeted programs, it invariably results in substantial leakage. Narrow targeting at the level of the community or administrative department can be an effective way to reach the poor for two main

13. Further details, including those on the interregional distribution of leakage and undercoverage, are given in Bigman and others (1999).

14. In Bigman and others (1999) we also introduce actual income transfers and evaluate the poverty impact of these transfers for a given budget. This budget is equally divided among all individuals included in each program. Using a budget that is just more than one-half the actual poverty gap, undifferentiated transfers reduce the headcount measure of poverty by a relatively high 22 percent, while village-level targeting using actual poverty levels reduces that measure of poverty by 33 percent. Using predicted poverty levels reduces the measure of poverty by about 27 percent. These are only modest gains, but giving all households living in villages included in a program the same monetary transfer is by no means optimal when minimizing poverty. Our scheme is aimed at identifying poor communities, and it attempts to minimize leakage and undercoverage.

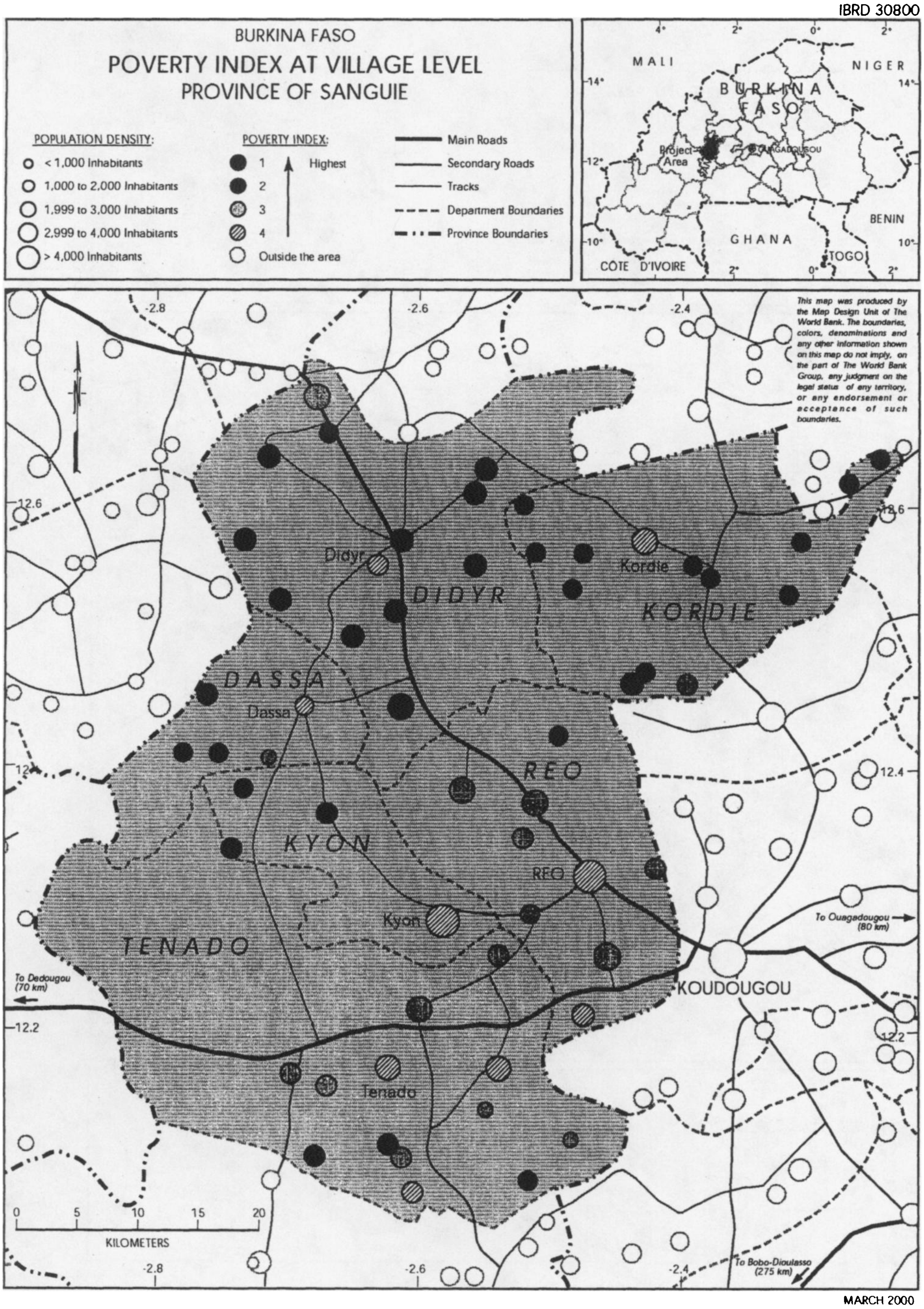
reasons. First, in most developing countries, particularly countries in Sub-Saharan Africa, poverty tends to be concentrated in villages and certain parts of towns. Second, geographical targeting entails relatively low administrative costs and, by relying primarily on local authorities, may ensure that a large portion of the benefits will reach the target population.

This article presented a methodology for using data from many different sources in order to establish criteria for targeting poverty-reduction programs at the level of the village, urban community, or local administrative department. This methodology consists of collecting data from several sources, aggregating them at the village level, and arranging them as a geographic information system. We conducted an econometric analysis using data from a household survey to identify the variables that best explain household consumption. The explanatory variables included important characteristics of the community and of households in that community. We selected the household-level explanatory variables whose mean values in each community were available for most of the communities in Burkina Faso not included in the household survey. This made it possible to use the model estimated in the regression analysis with the data of the Priority Survey to predict the incidence of poverty in all the villages outside the Priority Survey sample and thereby identify the spatial distribution of poverty at the community level.

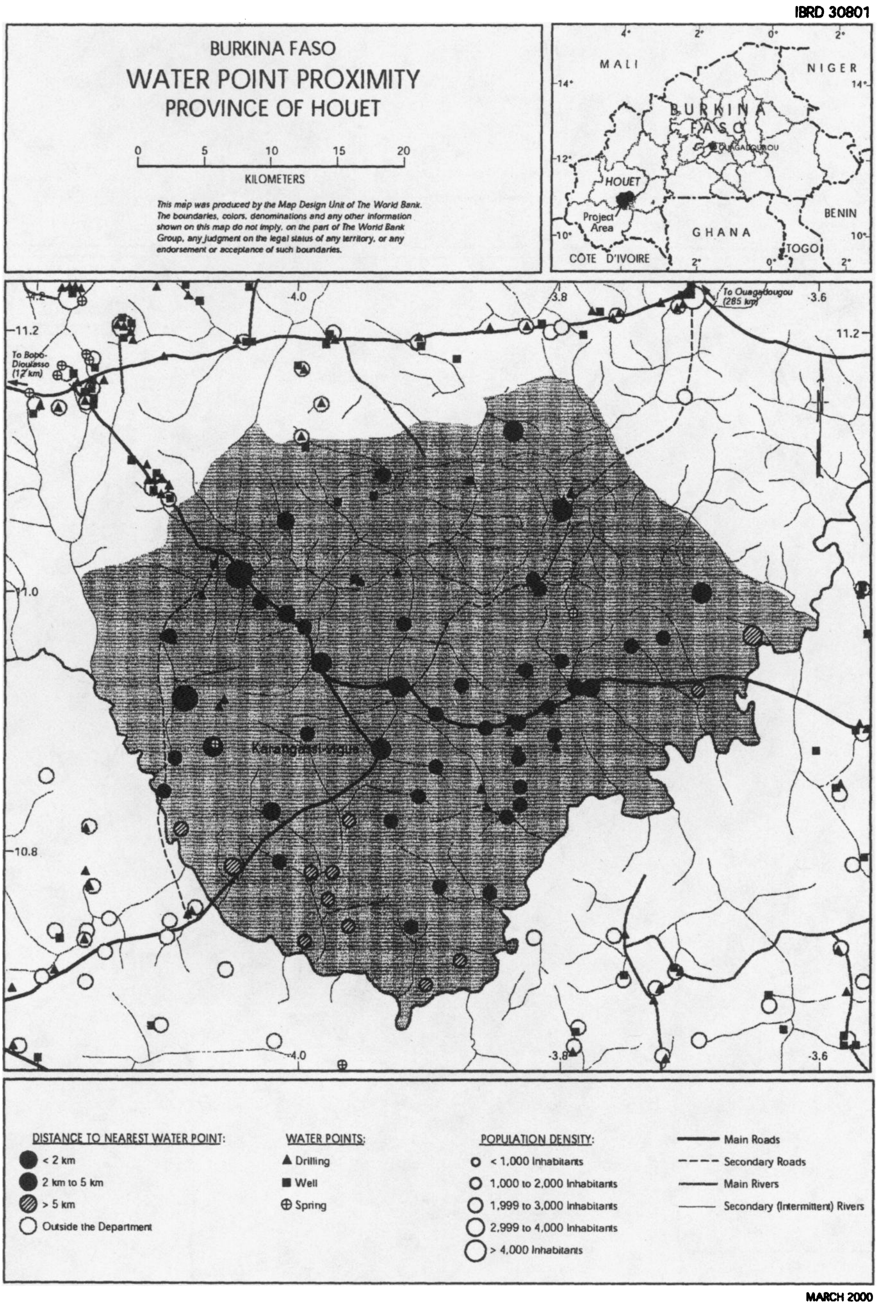
Constraints on the availability and quality of data for Burkina Faso led to considerable prediction errors and prevented us from using the complete ordering of the villages as predicted by the econometric analysis. To reduce the impact of these errors, we divided the villages into several categories and focused only on the villages categorized as poorest and least poor. Indeed, practical considerations in applying antipoverty programs and tight budget constraints are likely to reduce the need for a complete ordering. Poverty alleviation programs are more likely to focus on villages at the lower end of the distribution, and cost-recovery programs are more likely to focus on villages at the higher end. Nevertheless, the limited availability of geo-referenced data and the low quality of the data currently available reduced the predictive power of our econometric analysis. Further work is needed to augment and improve the stock of relevant data.

Targeting poverty alleviation or cost-recovery programs at the level of the village or department has other advantages as well. First, budget constraints are likely to restrict programs that are targeted to larger geographic areas, such as regions or states, and, as a result, the errors of inclusion and exclusion are likely to be high. Targeting smaller geographic areas can reach many more of the country's poor, given the same budget constraints. Second, lower-level targeting is likely to include villages and districts in all regions or states and thus be less divisive and contentious on ethnic, social, or political grounds. Third, whereas differences in the incidence of poverty among regions are primarily due to differences in agroclimatic conditions, differences in the incidence of poverty among villages within the same region often reflect past policy biases that led to differences in the quality of access roads or public services. Targeting future policies in light of these criteria can remedy past biases.

Appendix figure 1.



Appendix figure 2.



References

The word "processed" describes informally reproduced works that may not be commonly available through library systems.

Baker, Judy L., and Margaret Grosh. 1994. "Measuring the Effects of Geographic Targeting on Poverty Reduction." LSMS Working Paper 99. Development Research Group, World Bank, Washington, D.C.

Besley, Tim, and Ravi Kanbur. 1991. "The Principles of Targeting." In V. N. Balasubramanyam and Sanjaya Lall, eds., *Current Issues in Development Economics*. London: Macmillan.

Bigman, David, Stefan Dercon, Dominique Guillaume, and Michel Lambotte. 1999. "Community Targeting for Poverty Reduction in Burkina Faso." CES Discussion Paper Series 99.10. Center for Economic Studies, Catholic University of Leuven. Processed.

Coulombe, Harold, and Andrew McKay. 1996. "Modeling Determinants of Poverty in Mauritania." *World Development* 24(6):1015–31.

Deaton, Angus. 1997. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. Washington D.C.: World Bank and Johns Hopkins University Press.

Glesjer, Henri. 1965. "A New Test for Heteroscedasticity." *Journal of the American Statistical Association* 60(6):539–47.

Glewwe, Paul. 1991. "Investigating the Determinants of Household Welfare in Côte d'Ivoire." *Journal of Development Economics* 35(2):307–37.

Glewwe, Paul, and Oussama Kanaan. 1989. "Targeting Assistance to the Poor Using Household Survey Data." Policy, Planning, and Research Working Paper 225. Development Research Group, World Bank, Washington, D.C. Processed.

Hentschel, Jesko, Jean Olson Lanjouw, Peter Lanjouw, and Javier Poggi. 1998. "Combining Survey Data with Census Data to Construct Spatially Disaggregated Poverty Maps: A Case Study of Ecuador." Development Research Group, World Bank, Washington, D.C. Processed.

INSD (Institut National de Statistique et de Démographie). 1996. "Le profil de pauvreté au Burkina Faso." Ouagadougou. Processed.

Maddala, G. S. 1983. *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge, U.K.: Cambridge University Press.

Ravallion, Martin. 1993. "Poverty Alleviation through Regional Targeting: A Case Study for Indonesia." In Karla Hoff, Avishay Braverman, and Joseph E. Stiglitz, eds., *The Economics of Rural Organization: Theory, Practice, and Policy*. New York: Oxford University Press.

———. 1996. "Issues in Measuring and Modelling Poverty." *Economic Journal* 106(September):1328–43.

Rosenzweig, Mark, and Kenneth Wolpin. 1986. "Evaluating the Effects of Optimally Distributed Public Programs." *American Economic Review* 76(3):470–82.

van de Walle, Dominique. 1995. "Public Spending and the Poor: Theory and Evidence: Introduction." In Dominique van de Walle and Kimberly Nead, eds., *Public Spending and the Poor: Theory and Evidence*. Baltimore, Md.: Johns Hopkins University Press.