

Chapter 8. Understanding the Determinants of Poverty

Summary

A poverty profile describes the pattern of poverty, but is not principally concerned with explaining its causes. Yet a satisfactory explanation of why some people are poor is essential if we are to be able to tackle the roots of poverty.

Among the key causes, or at least correlates, of poverty are:

Regional-level characteristics: these include vulnerability to flooding or typhoons; remoteness; quality of governance; property rights and their enforcement.

Community level characteristics: these include the availability of infrastructure (roads, water, electricity) and services (health, education), proximity to markets, and social relationships.

Household and individual characteristics: Among the most important are:

Demographic: household size, age structure, dependency ratio, gender of head.

Economic: employment status, hours worked, property owned.

Social: health and nutritional status, education, shelter.

Regression analysis is commonly undertaken to identify the effects of each of these characteristics on income (or expenditure) per capita. Care is needed to choose the independent variables carefully, to be sure that they are indeed exogenous. A number of more exotic techniques are now available for this purpose, including CART (classification and regression tree) models and multiple-adaptive regression splines (MARS models).

Regression techniques are good at identifying the immediate (“proximate”) causes of poverty, but are less successful at finding the “deep” causes; they can show that a lack of education causes poverty, but cannot so easily explain why some people lack education.

Learning Objectives

After completing the module on poverty lines, you should be able to:

36. Identify the main immediate (“proximate”) causes of poverty.
37. Classify the main causes of poverty into characteristics related to the country or region, the community, and the household and individual.
38. Explain how regression techniques may be used to identify the proximate causes of poverty and their relative importance.
39. Explain why researchers generally prefer to use regressions to “explain” income (or expenditure) per capita rather than whether an individual is poor or not.
40. Evaluate the assertion that the weakest part of poverty analysis is the understanding of its fundamental causes, and that this represents a “missing middle” that makes it difficult to define a successful anti-poverty strategy.

8.1 What causes poverty?

A poverty profile describes the pattern of poverty, but is not principally concerned with explaining its causes. Yet a satisfactory explanation of why some people are poor is essential if we are to be able to tackle the roots of poverty. This chapter addresses the question of what causes poverty.

Poverty may be due to national, sector-specific, community, household or individual characteristics. This chapter summarizes some of the characteristics of the poor by region, community, household and individual characteristics and then discusses how regression techniques can be used to determine the factors “causing” poverty.

Two cautions are in order. First, it can be difficult to separate causation from correlation. For instance, we know that poor people tend to have low levels of education; but are they poor because they have little education, or do they have little education because they are poor? A statistical association alone is not enough to establish causality, and additional information is likely to be required.

Second, most of the “causes” of poverty that we identify in this chapter are immediate (or “proximate”) causes, but not necessarily “deep” causes. For instance, suppose that we can demonstrate that low levels of education do indeed increase the risk of poverty. This is interesting, but now begs the question of why some people have low levels of education in the first place: Were the school fees too high? Was there no school nearby? Was the quality of the education abysmal? Were their parents unsupportive, or even hostile to education? Was there a concern that an educated woman could not find a husband?

The weakest part of poverty analysis – what Howard White calls the “missing middle” – is developing a clear understanding of the fundamental causes of poverty. Such an understanding is needed if one is to develop an effective strategy to combat poverty. Since there is no reason to believe that the root causes of poverty are the same everywhere, country-specific analysis is essential.

8.1.1 Regional level characteristics

At the regional level, there are numerous characteristics that might be associated with poverty. The relationship of these characteristics with poverty is country-specific. In general, however, poverty is high in areas characterized by geographical isolation, a low resource base, low rainfall, and other inhospitable climatic conditions. For example, many argue that economic development in Bangladesh is severely retarded due to its susceptibility to annual floods; and Nghe An province in north-central

Vietnam is poor in part because it is regularly hit by typhoons, which destroy a significant part of the accumulated stock of capital. In many parts of the world the remoteness of rural areas – which lower the price farmers get for their goods and raise the price they pay for purchases, due to high transport costs – is responsible for generating food insecurity among the poor. Inadequate public services, weak communications and infrastructure, as well as underdeveloped markets are dominant features of life in rural Cambodia, as in many other parts of the world, and clearly contribute to poverty.

Other important regional and national characteristics that affect poverty include good governance, a sound environmental policy, economic, political and market stability, mass participation, global and regional security, intellectual expression and a fair, functional, and effective judiciary. Regional-level market reforms can boost growth and help poor people, but it is important to note that they can also be a source of dislocation. The effects of market reforms are complex, deeply linked to institutions and to political and social structures. The experience of transition, especially in countries of the former Soviet Union, vividly illustrates that market reforms in the absence of effective domestic institutions can fail to deliver growth and poverty reduction. There is also a case for bringing vulnerability and its management to center stage. Participatory poverty work underlines the importance of vulnerability to economic, health, and personal shocks.

Inequality is also back on the agenda. New work shows the importance of gender, ethnic, and racial inequality as a dimension – and a cause – of poverty. Social, economic, and ethnic divisions in regions are often sources of weak or failed development. In the extreme, vicious cycles of social division and failed development erupt into internal conflict (within or across regions), as in Bosnia and Sierra Leone, with devastating consequences for people.

8.1.2 Community level characteristics

As with regional characteristics, there are a variety of community-level characteristics that may be associated with poverty for households in that community. At the community level, infrastructure is a major determinant of poverty. Indicators of infrastructure development that have often been used in econometric exercises include proximity to paved roads, whether or not the community has electricity, proximity to large markets, availability of schools and medical clinics in the area, and distance to local administrative centers. Other indicators of community level characteristics include average human resource development, access to employment, social mobility and representation, and land distribution.

Recent research has also stressed the importance of social networks and institutions, and “social capital” (which includes, for instance, the level of mutual trust in the community) xxx: Robert Putnam. In addition to removing social barriers, effective efforts to reduce poverty require complementary initiatives to build up and extend the social institutions of the poor. Social institutions refer to the kinship systems, local organizations, and networks of the poor and can be thought of as different dimensions of social capital. Research on the roles of different types of social networks in poor communities confirms their importance. An analysis of poor villages in North India, for example, shows that social groups play an important role in protecting the basic needs of poor people and in mediating risk. Studies of agricultural traders in Madagascar show that social relationships are more important to traders than input prices. Close relationships with other traders are used to lower the transactions costs of exchange, while ties to creditors and others who can help out during times of financial hardship are vital sources of security and insurance.

How does social capital affect development? The narrowest view holds social capital to be the social skills of an individual – one’s propensity for cooperative behavior, conflict resolution, tolerance and the like. A more expansive meso view associates social capital with families and local community associations and the underlying norms (trust, reciprocity) that facilitate coordination and cooperation for mutual benefit. A macro view of social capital focuses on the social and political environment that shapes social structures and enables norms to develop. This environment includes formalized institutional relationships and structures, such as government, the political regime, the rule of law, the court system, and civil and political liberties. Institutions have an important effect on the rate and pattern of economic development.

An integrating view of social capital recognizes that micro, meso, and macro institutions coexist and have the potential to complement one another. Macro institutions can provide an enabling environment in which micro institutions develop and flourish. In turn, local associations help sustain regional and national institutions by giving them a measure of stability and legitimacy – and by holding them accountable for their actions. Social capital is clearly a complicated characteristic and often researchers find it difficult to identify appropriate variables that measure social capital quantitatively.

8.2 Household and individual level characteristics

Some of the important characteristics in this category would include the age structure of household members, education, gender of the household head, and the extent of participation in the labor force. In

recent times, other components that fall under this category have included domestic violence prevention, and gender-based, anti-discrimination policies. The following discussion organizes these characteristics into groups and discusses them in greater detail. These groups are demographic, economic and social characteristics.

8.2.1 Demographic characteristics

Indicators of household size and structure are important in that they show a possible correlation between the level of poverty and household composition. Household composition, in terms of the size of the household and characteristics of its members (such as age), is often quite different for poor and non-poor households. The Cambodian CSES of 1993/94 shows that the poor tend to live in larger households, with an average family size of 6.6 persons in the poorest quintile compared to 4.9 in the richest quintile; similar patterns are found in most countries. The poor also tend to live in younger households – with the bottom quintile having twice as many children under 15 per family as the top quintile – and slightly fewer elderly people over age 60. Better-off households also tend to have heads that are somewhat older.

The dependency ratio is calculated as the ratio of the number of family members not in the labor force (whether young or old) to those in the labor force in the household. This ratio allows one to measure the burden weighing on members of the labor force within the household. One might expect that a high dependency ratio will be associated with greater poverty.

It is widely believed that the *gender of the household head* significantly influences household poverty, and more specifically that households headed by women are poorer than those headed by men. This is of particular importance to Cambodia. Due to male casualties in past wars, women are often the heads of households. Women play an important role in the labor force, both in the financial management of the household and in the labor market, but appear to face large degree of discrimination. They are severely affected by both monetary and non-monetary poverty; for example, they have low levels of literacy, are paid lower wages, and have less access to land or equal employment. According to a report based on a joint conference between the Cambodian Institute for Cooperation and Peace and the World Bank Institute, 43 percent of women are illiterate and 90 percent of these women are poor. According to the Cambodian Ministry of Women's and Veterans' Affairs, the following contribute to poverty—lack of opportunities for employment and education, lack of access to finance, mass illiteracy, food insecurity, malnutrition, human trafficking, powerlessness, no resources, overwork in wage employment and in households, discrimination in the labor market and in work places, and domestic violence. So many observers are surprised to learn that poverty rates are *not* higher among female-headed than male-headed

households in Cambodia. Likewise, female-headed households in neighboring Vietnam are no more likely to be in poverty than their male-headed counterparts.

8.2.2 Economic characteristics

Apart from income or consumption – which are typically used to define whether a household is poor – there are a number of other economic characteristics that correlate with poverty, most notably household employment and the property and other assets owned by the household.

There are several indicators for determining *household employment*. Within this array of indicators, economists focus on whether individuals are employed; how many hours they work; whether they hold multiple jobs; and how often they change employment.

The *property of a household* includes its tangible goods (land, cultivated areas, livestock, agricultural equipment, machinery, buildings, household appliances and other durable goods) and its financial assets (liquid assets, savings and other financial assets). These indicators are of interest as they represent the household's inventory of wealth and therefore affect its income flow. Furthermore, certain households, especially in rural areas, can be poor in terms of income, but wealthy when their property is taken into consideration. Despite its importance, property is difficult to value in practice in any reliable way. First, one encounters the same problem of under-declaration. Second, it is very difficult to measure certain elements of property such as livestock. Finally, the depreciation of assets may be difficult to determine for at least two reasons: 1) the life span of any given asset is variable; 2) the acquisition of these assets occurs at different moments in each household. Therefore, property is more difficult to use than certain other elements in the characterization of poverty.

8.2.3 Social characteristics

Aside from the demographic and economic indicators, several social indicators are correlated with poverty and household living standards. The most widely used are measures of health, education and shelter.

Four types of indicators are normally used to characterize *health* in analyzing a household's living standards. These indicators include

- Nutritional status, for example, anthropometric indicators such as weight for age, height for age, and weight for height;

- Disease status, for example, infant and juvenile mortality and morbidity rates as related to certain diseases such as malaria, respiratory infections, diarrhea and sometimes poliomyelitis;
- The availability of health care services such as primary health-care centers, maternity facilities, hospitals and pharmacies, basic health care workers, nurses, midwives, doctors and traditional healers, and medical service such as vaccinations, access to medicines and medical information; and;
- The use of these services by poor and non-poor households.

Three types of indicators are normally used to characterize **education** in an analysis of household living standards. These include the level of education achieved by household members (basic literacy; years of education completed); the availability of educational services, such as proximity to primary and secondary schools; and the use of these services by the members of poor and non-poor households. For this last item, commonly used measures include children's registration in school, the dropout rate of children by age and gender and reasons for dropping out, the percentage of children who are older than the normal age for their level of education, and average spending on education per child registered.

Literacy and schooling are important indicators of the quality of life in their own right, as well as being key determinants of poor people's ability to take advantage of income-earning opportunities. Based on CSES data, Cambodia by 1993-94 had achieved a self-reported basic literacy rate of 67 percent among adults (older than age 15), implying a high degree of literacy among the poor. However, the literacy gap remained quite large, with literacy ranging from just over half of adults (58%) among the poorest 20 percent of the population to 77% among the richest 20% quintile. Much larger differentials appear in the distribution of schooling attainment: adults in the poorest quintile averaged 3.1 years of schooling, compared with 5.3 years among the richest quintile. Men averaged 5.1 years of education, compared with 3.2 years for women.

Shelter refers to the overall framework of personal life of the household. It is evaluated, by poor and non-poor household groups, according to three components (some of which overlap with the indicators mentioned above): housing, services, and the environment. The housing indicators include the type of building (size and type of materials), the means through which one has access to housing (renting or ownership), and household equipment. The service indicators focus on the availability and the use of drinking water, communications services, electricity, and other energy sources. Finally, the environmental indicators concern the level of sanitation, the degree of isolation (availability of roads and

paths which are usable at all times, length of time and availability of transportation to get to work) and the degree of personal safety.

Example. It is generally established that poor households live in more precarious, less sanitary environments, which contribute to the poorer health and lower productivity of household members. To illustrate: the data from the Cambodian Socio-Economic Survey of 1993/94 show that water and sanitation are especially important influences on health and nutritional status. The CSES showed that only 4 percent of the poorest quintile had access to piped water, while more than 17% of the richest quintile had the same. Similar differences are apparent in access to sanitation. Just 9% of the poor had access to a toilet in the home, while around half of the richest quintile did.

Another indicator of housing standards is access to electricity. Here again the access of the poor lagged far behind. Access to electricity from a generator or line connection rose sharply with income, from a mere 1 percent among people in the bottom quintile to 37 percent of Cambodians in the richest quintile. Other indicators of household wealth include ownership of transportation. Access to bicycles is quite evenly distributed, with at least one half of households owning a bicycle in every quintile, even the poorest. However, access to cars, jeeps or motorbikes is very rare among the poor and rises sharply with income.

A summary of the main influences on poverty is provided in Table 8.1.

8.3 Analyzing the determinants of poverty: Regression techniques

Tabulated or graphical information on the characteristics of the poor is immensely helpful in painting a profile of poverty. However, it is not always enough when one wants to tease out the relative contributions of different influences on poverty. For example, data from the Vietnam Living Standards Survey of 1998 showed per capita expenditure to be significantly higher in female-headed households than in households headed by a man. However, after controlling for other influences – where the household lived, the size of the household, and so on – the effect proved to be statistically insignificant.

Table 8.1 Main determinants of poverty	
Regional characteristics	Isolation/remoteness, including less infrastructure and poorer access to markets and services Resource base, including land availability and quality. Weather (e.g. are typhoons or droughts common) and environmental conditions (e.g. frequency of earthquakes) Regional governance and management Inequality
Community characteristics	Infrastructure (e.g. is there piped water, access to a tarred road) Land distribution Access to public goods and services (e.g. proximity of schools, clinics) Social structure and social capital
Household characteristics	Size of household Dependency ratio (i.e. unemployed old and young relative to working age adults) Gender of head; or of household adults on average Assets (typically including land, tools and other means of production, housing, jewelry) Employment and income structure (i.e. proportion of adults employed; type of work – wage labor or self employment; remittance inflows) Health and education of household members on average
Individual characteristics	Age Education Employment status Health status Ethnicity

By far the most widespread technique used to identify the contributions of different variables to poverty is regression analysis. There are two main types of analysis:

- (i) Attempts to explain the level of expenditure (or income) per capita – the dependent variable – as a function of a variety of variables (the “independent” or “explanatory” variables). The independent variables are typically of the type discussed above in section 8.2.
- (ii) Attempts to explain whether a household is poor or not, using a logit or probit regression. In this case the independent variables are as in (i), but the dependent variable is binary, usually taking on a value of 1 if the family is poor and zero otherwise.

We now consider each of these in somewhat more detail.

A regression estimate shows how closely each independent variable is related to the dependent variable (e.g. consumption per capita), *holding all other influences constant*. There is scope for a wide variety of regressions; for instance the dependent variable could measure child nutrition, or morbidity, or schooling, or other measures of capabilities; the regressions could be used to examine the determinants of employment or labor income; or regressions could be used to estimate agricultural production functions (which relate production with information on type of crops grown per area, harvest, inputs into agricultural production, and input and output prices). For an accessible discussion and lots of examples, in the context of Vietnam, see Haughton et al., *Health and Wealth in Vietnam* (1999).

A typical multiple regression equation, as applied to poverty analysis, would look something like this:

$$\log(y_i / z) = \alpha_0 + \alpha_1 X_i^1 + \alpha_2 X_i^2 + \dots + \alpha_n X_i^n$$

where z is the poverty line, y_i is (per capita) income or consumption, the X_j^n are the “explanatory” variables and the α_j are the coefficients that are to be estimated. Note that y_i/z is in log form, which is a common way of allowing for the log normality of the variable. Since we are interested in the determinants of individual poverty, but typically have information at the level of the household, it is standard (but in this context, not universal) to estimate the regression using weights that reflect the size of the household. The `regress` command in STATA is flexible and allows one to use weights.

The independent (i.e. right hand side) variables may be continuous variables, such as the age of the individual. But often we want to represent a categorical variable – the gender of the person, or the region in which he or she lives. In this case we need to create a “dummy” variable – for instance, the variable might be set to 1 if the person is a man and 0 for a woman. If there are, say, 10 regions in a country, each region would need to have its own dummy variable; but one of the regions needs to be left out of the regression, to serve as the point of reference.

Often we believe that the determinants of poverty differ from one area to the next, which would mean that there are differences in “structure.” In this case we could estimate separate regressions for each area – for instance, for each region in a country. Sometimes it is sufficient to specify the regression equation in a way that is flexible enough to allow for such differences, for instance by allowing interactive effects: for example, one could create a variable that multiplies educational level by age, instead of estimating separate regressions for individuals in different age groups.

The fit of the equation is typically measured using \bar{R}^2 (“adjusted R squared”), which will vary between 0 (no fit) and 1 (perfect fit). There is no hard and fast rule for determining whether an equation fits well, although with household survey data one is often pleased to get an \bar{R}^2 of 0.5 or more.

We also need to know how much confidence to place in the accuracy of the coefficients as guides to the truth; this is commonly done by reporting t-statistics, which are obtained by dividing a coefficient by its standard error. The rule of thumb is that if the t-statistic is, in absolute terms, smaller than 2, then the coefficient is not statistically significantly different from zero (at about the 95% confidence level); in other words, we cannot be sure that we have picked up an effect, and it is possible that the coefficient just

reflects noise in the data. Many researchers prefer to report p-values, which give the confidence level directly; a p-value of, say, 0.03 indicates that we are 97% confident that the coefficient is not zero. So we hope to find low p-values (and we usually do when working with large data sets). Arbitrarily, it is standard to consider a coefficient to be statistically significant if the p-value is less than 0.05, but this rule is not graven in stone.

Table 8.2 shows typical regression output, from an example based on data from the Côte d'Ivoire. Here the dependent variable is the log of per capita household expenditure. The figures in parentheses are t-statistics. Separate regressions were estimated for households in urban, and in rural, areas, on the thinking that the determinants of poverty might be quite different in these two areas.

The results of the urban equation show that education is an important determinant of expenditure per capita. The coefficients for most of the educational variables are statistically significant and quite large; having an elementary education boosts income by approximately 38% relative to someone with no education; this comes from the coefficient of 0.38, and the fact that the dependent variable is in log form.

Table 8.2: Determinants of household spending levels in Côte d'Ivoire, c.1993				
	Urban		Rural	
	Coefficient	t-statistic	Coefficient	t-statistic
Dependent variable: ln(expenditure/capita)				
Educational level of most-educated male				
Elementary	0.38	(5.3)	0.04	(0.6)
Junior secondary	0.62	(8.6)	0.08	(0.9)
Senior secondary	0.80	(9.6)	0.05	(0.4)
University	0.93	(9.4)	-	
Elementary	0.11	(1.7)	0.07	(1.0)
Junior secondary	0.24	(3.1)	0.27	(2.2)
Senior secondary	0.34	(4.1)	-	
University	0.52	(4.1)	-	
Value of selected household assets				
Home	0.06	(5.3)	-	
Business assets	0.04	(3.3)	0.16	(4.9)
Savings	0.08	(4.7)	-	
Hectares of agriculture land				
Cocoa trees			0.17	(4.3)
Coffee trees			0.04	(1.3)
Distance to nearest paved road			-0.04	(-2.9)
Distance to nearest market			-0.09	(-3.3)
Unskilled wage			0.37	(6.4)
<i>Source: Adapted from Grosh and Munoz 1996, p.169, based on Glewwe 1990.</i>				

However, in rural areas education does not appear to explain expenditure per capita levels very much – a not uncommon finding. On the other hand, the infrastructure variables have substantial predictive power:

households located in villages that are nearer to both paved roads and public markets are better off, as are households living in areas with higher wage levels. The results raise further questions about the quality of education in rural areas and the importance of rural infrastructure in helping families grow out of poverty, which could be addressed in putting together a poverty reduction strategy.

It is vital to choose the independent variables carefully, and to be sure that they are truly exogenous. For instance, in the above example, one could have included income as an independent variable, along with education, assets, and the like. But that does not advance us much, because income is in turn determined by such variables as educational levels and household assets. In our drive to find the underlying causes of poverty, we need to dig deep to find variables that are indeed pre-determined. A good start is to work with the variables identified in Table 8.1.

When multiple cross-sectional surveys are available, the same regression can be repeated for different years to see how the association of certain correlates with income or consumption varies over time. Variations over time will be reflected in changes in coefficients or parameters. The results of repeated cross-section regressions can also be used to decompose variations in poverty in terms of changes in household characteristics, and changes in the returns to (or impact of) these characteristics (e.g., Baulch et al. 2001; van de Walle 2000; Wodon 2000). xxx

Some researchers prefer to use, on the left hand side, a binary variable that is set equal to 1 if the household is poor, and to 0 otherwise. Some of the information is lost by doing this, and the resulting logit or probit regression is relatively sensitive to specification errors, which is why this is rarely the preferred approach to take. However an analysis of this kind is likely to be useful when designing targeted interventions (e.g. educational vouchers for poor households), as it allows one to assess the predictive power of various explanatory variables used for means testing. It is also possible to undertake a multiple logit analysis, where the dependent variable could be in one of several categories – for instance, expenditure quintiles.

Recent research has explored more exotic forms of analysis, including non-parametric regression, classification and regression trees (CART models), and multiple-adaptive regression splines (MARS models). The goal of all such efforts is to unearth a parsimonious number of determinants of poverty, and quantify their effects, even when those effects are highly nonlinear.

References

Baulch et al. 2004 In Paul Glewwe, Nisha Agrawal and David Dollar (eds.), *Economic Growth, Poverty and Household Welfare in Vietnam*, World Bank 2004:

1. Ethnic Minority Development in Vietnam: A Socio-Economic Perspective [With Bob Baulch, Truong Thi Kim Chuyen and Dominique Haughton.]

Gibson. Cambodia. ??

VLSS

Haughton et al., 1999, *Health and Wealth in Vietnam*.

Grosh and Munoz. 1996.

van de Walle. 2000. Van de Walle, Dominique, and Dileni Gunewardena. 2001. "Sources of Ethnic Inequality in Vietnam." *Journal of Development Economics*, 65(1):177–207.

Wodon. 2000.