— Collaborative Research Project —

Poverty, Income Distribution and Labour Market Issues in Sub-Saharan Africa

CR-2-11

Empirical Modelling of Investment in Health and Nutrition

by

John Strauss
and Duncan Thomas

July 1998

1. Introduction

The study of poverty not only includes measurement of poverty and the construction of poverty profiles, but it also examines the basic constraints that underlie poverty. Of these, assets have long been considered crucial. While physical assets such as land have often been emphasized, it is human assets that many economists now think are key. T.W. Schultz long ago argued this logic (see his economics of being poor, Nobel Prize lecture, 1980) . He had a broad view of what constituted human assets, or human capital as it became known, a view that included investments in both schooling and health. Economists quickly took up his call to quantify the impact of schooling on economic returns (see, for instance, Psacharopoulos, 1994), but were less quick to explore the role of health in raising productivity (see Strauss and Thomas, 1998, for a recent review).

At present, the state of empirical knowledge of the returns to schooling is now substantial and the literature on health is growing (T.P. Schultz, 1988a; Strauss and Thomas, 1995). It is clear that such investments do indeed have a high return in increasing incomes and therefore reducing poverty, as T.W. Schultz hypothesized. It is also clear that investments in schooling, particularly of women, has a very positive influence on investments in the human capital of children.

Given this conclusion, a critical policy question is why haven't more poor households invested in more schooling and health in order to pull themselves out of poverty. Does their poverty so constrain them that it is extremely difficult to make such investments? This is a common view, but to the extent that this story holds, are there public investments or policies that can increase the probability of the poor so investing? In brief, this set of questions is the focus of this review. Specifically, we address the issues involved in empirically estimating determinants of human capital investments from survey data; it turns out that there are considerable problems involved.

We begin with some economic theory necessary to model such investments, focusing on a basic, static household production model. It is vital to model the behavioural responses of households and individuals to the relaxation of external constraints because it is households and individuals who ultimately make the necessary investments; public policy cannot will outcomes, as much as governments might try. We turn, in Section 3, to measurement issues. In the case of health, in particular, measurement presents some extremely difficult issues; particularly in deciding how to quantify its different dimensions and to do so without bias. This discussion is followed by an actual example of an empirical study, in which we draw from our work on investments in child and adult health in Côte d'Ivoire. That study is prototypical in the literature in that it assumes that public investments are randomly distributed across localities. If they are not, then empirical results can be terribly biased. We discuss the issues involved in systematic programme placement in Section 5, together with a discussion of potential econometric solutions to the problems that are typically encountered. We go on in Section 6 to

discuss another set of possible solutions to the programme placement problem, social experiments, and provide some recent examples to illustrate the associated strengths and weaknesses. We conclude in Section 7.

2. Modelling health (and other human capital) investments

Health production function

We start with a health production function, a biological function that relates health outcomes to health inputs, labour effort, age and sex, family health background, community disease, and public health characteristics, and to factors unobserved to the analyst that affect health (see Grossman, 1972, for the original ideas related to health production functions). If we were interested in modelling schooling investment we would begin with an education production function that likewise related inputs to school outputs (such as test scores).

$$H = H(N, L; A, B', D, \mu, e_h) \qquad (1)$$

H	=	health output
N	=	health inputs
L	=	labour effort
A	=	age and other individual demographics
B'	=	family background, e.g., parental health
D	=	community disease and public health characteristics
μ	=	individual health endowment unobserved to analyst, but known to person
e_h	=	measurement error, unknown to person

We think of health outputs as "final" outputs, although there is ambiguity in this. Examples include a person's height, body mass (weight measured in kilograms/height^2 measured in metres), disease incidence or severity, or ability to function easily at specific physical activities such as walking for 1 kilometre. Mortality, too, may be thought of as a final health outcome.

Examples of health inputs include nutrient intakes such as calories, proteins, iron, vitamins and minerals; type of feeding in the case of infants (e.g., sole breastfeeding, supplemented breastfeeding and so forth); energy expenditure; water and sanitation quality as a measure of exposure to pathogens; exposure to underlying disease conditions, for which one crude proxy may be rainfall; behaviours such as smoking and alcohol consumption; use of health facilities, medications, immunizations and so forth; and underlying healthiness of the person. A good example of a specification of an empirical

health production function is given by the work of the Cebu Team (1992). They specify equations that relate the growth (in weight) of young children to initial weight, diarrheal disease incidence, breastfeeding patterns, solid food intakes, water and sanitation quality, and use of medical facilities, among other factors.

Many potential health inputs are very difficult to measure in surveys and so are not. Some of them are immutable (such as underlying healthiness), while others can be changed through behaviour. If the unobserved immutable factors are known to the individuals and households, knowledge of them may be used to decide input use for those that can be controlled. We can subdivide these unobservables into those that are known by the household and used in input allocation decisions, , and those that are not known to the household and thus cannot be used in decision making, e_h.

Utility function

We assume that the household is the major decision making unit[1] and that the economic problem is the usual one of maximizing a utility function subject to constraints. We represent the utility function as:

$$U = U (C, L; H, S, A, B, \xi) \qquad (2)$$

Where C = goods consumption = $\{C^*,N_1\}$, N_1,cN

ξ = unobserved (to the analyst) taste shifters (that are known by the household)

It is possible to embellish the utility function; for instance, to allow parents to care about the incomes of their children as adults, with adult incomes depending on current health, which in turns depends on past health, including as a child. Then investments in child health (or schooling) are not only valued directly, but also as an instrument to higher adult earnings. Since this representation is static, we do not pursue that avenue.

Budget constraint

The household's budget constraint may be written as:

$$p_cC^* + p_nN = wL + y \qquad (3)$$

where pi = price of good i

y = exogenous non-labour income, if any

In the human capital approach, wages depend on human capital, which includes the factors usually thought of as human capital, such as schooling and its quality, but also health. We model market wages to be a function of human capital:

[1] See Strauss et al. (1998) for a discussion of issues of intrahousehold allocations.

Market wage function

$$w = w (H; A, S, B, I, a, e_w) \quad (4)$$

where S = individual's schooling
B = family background, including parental health and schooling
I = local labour-demand related infrastructure
a = time-invariant, unobserved (to the analyst) market work ability
e_w = measurement error of wage (unknown to the person)

First order conditions

The first order conditions with respect to health inputs are:

$$\frac{\partial U}{\partial H}\frac{\partial H}{\partial N_j} = \lambda(P_n - L[\frac{\partial w}{\partial H}\frac{\partial H}{\partial N_j}]) \qquad (5)$$

where λ = marginal utility of income

Health input demands and output "supplies"

Solving the first order conditions will as usual result in deriving reduced form equations, in this case for health input demands and output "supplies", as a function of exogenous constraint variables.

$$N_J = N_j (p_n, p_c, S, A, B, y, D, I, \alpha, \mu, \xi) \qquad (6)$$

$$H = h (p_n, p_c, S, A, B, y, D, I, \alpha, \mu, \xi) \qquad (7)$$

It is immediately apparent that health inputs, and therefore outputs, vary with individual-specific, unobserved health endowments and market ability, μ, and α, and with tastes, ξ, such as taste for good health. Therefore in the health production function, (1), or in the market wage function, (4), health inputs (and outputs) will be correlated with unobservables, μ, and α!

The import of these correlations is that if one did want to estimate the health biology, or a wage function with health output measures included as covariates, using ordinary least squares (OLS) will result in biased coefficient estimates.[2] For instance, suppose that very unhealthy babies have trouble breastfeeding. If we estimate a production function for infant weight using OLS and omit hard to measure innate child healthiness from our covariates, then the breastfeeding coefficient will be biased upwards; children who are very unhealthy and therefore have low weights (or weight gains) will also be less likely to be breastfed.

We could use instrumental variables estimators, such as 2SLS, provided we can plausibly identify the endogenous variables in the production function. Fortunately the economic theory is helpful for this purpose. From the reduced forms we can see that prices of health inputs, p_N, plus disease and local health infrastructure variables, D, as instrumental variables, affect health inputs, but do not belong in the production function. They are

[2] This is analogous to estimating physical production functions, as pointed out long ago by Mundlak (1963). See Rosenzweig and Schultz (1983) for a good discussion in the context of health production functions.

therefore available, if measured, to use as instruments for health inputs in the production function (or health outputs in a market wage function).

Alternatively, with panel data, it is possible as pointed out by Mundlak (1963) to use fixed effects or first difference estimators, which will sweep out the unobserved μ and α terms, provided that they enter linearly (an assumption that may not always hold).

Conditional demand functions

It is generally the case that poorer households live in areas with fewer social services and so, without good controls for household resources, the estimated effects of community health facility and price characteristics, p_n, may largely reflect household resources. Given that a primary motivation of this study is the estimation of the effects of these community factors, it is key that resource availability be adequately measured. Non-labour income, y, which is notoriously difficult to measure, is not likely to adequately reflect long-run resources. Furthermore, if households smooth consumption over income shocks, then expenditure is likely to be a better indicator of long-run resource availability. We assume that leisure is (weakly) separable from commodity consumption and health in the utility function, in which case it is possible to derive a conditional demand function (Pollak, 1969) that is analogous to (6) or (7), but depends on total expenditure (or income).

If consumption decisions (including leisure) are jointly determined with anthropometric outcomes, then total household expenditure should be treated as endogenous in the conditional demand function. Under the assumption of weak separability between leisure and other elements of the utility function, non-labour income (or the value of assets) is an appropriate instrument. However, even expenditure is a noisy indicator of long-run resources. Taking an instrumental variables approach has the added advantage of ameliorating the effect of measurement error on the estimates.

In models of health of household members, it is inappropriate to treat household composition as exogenous; we control rather crudely for household size by including household per capita expenditure, PCE, as the indicator of resources in the conditional demand function. Then:

$$H = h (p_n, p_c, S, A, B, PCE, D, I, \alpha, \mu, \xi) \qquad (8)$$

It is this model that we will estimate for child and adult anthropometric outcomes. Before turning to the data, however, we need to address some issues of estimation.

3. Measurement of health

Whereas economists have largely focused on understanding the behaviours that underlie associations between health and labour market outcomes, physicians and epidemiologists

have devoted considerable effort to measuring health. This section highlights issues revolving around measurement both because it is important for interpreting empirical relationships with labour outcomes and because of its relative neglect by economists. Specifically, we highlight two aspects of health that set it apart from other indicators of human capital like education: first, it is multi-dimensional and, second, measurement error in health is likely to be related to income and labour market outcomes.

There is a consensus in the literature that years of schooling is a reasonably good indicator of educational attainment. No similar agreement exists for health, in part because health is fundamentally multi-dimensional (Ware et al., 1980).[3] Moreover, different dimensions of health are likely to have different effects on one's productivity or labour supply and these effects may well vary over the life course or wage distribution. We argue, therefore, that it makes good sense to examine the relationship between labour outcomes and *multiple* health indicators simultaneously, wherever possible.

Moreover, it is typically assumed that measurement error in schooling is random (Griliches, 1977). In contrast, many health indicators are measured with error that is systematically related to demand for health and other behaviours that are, in turn, related to wages, labour supply and other socioeconomic characteristics. This complicates interpretation of empirical relationships between health and labour outcomes and seriously compromises the value of standard fix-ups, such as instrumental variables. In addition, the extent and nature of errors are likely to vary from measure to measure. (See Stewart and Ware, 1992, for a comprehensive and thoughtful discussion.)

In this section, a series of different health indicators that have been used in the empirical literature are discussed. We begin with self-reported general health status, which is followed by self-reported morbidities, limitations to normal activities and measures of physical functioning. We then turn to nutrition-based health measures and include both health outputs (anthropometrics) and inputs (nutrient intakes). For each group of indicators, we discuss the nature of the information they are likely to contain along with the kinds of measurement error that are likely to be important. The section ends with a discussion of the implications of random and systematic measurement error for empirical work.

---

[3] In fact, education is more than highest grade completed but may also be characterized by years in school (or grade repetition), test scores, quality of schooling, post-school training and experience. Several studies have included these measures when calculating returns to education (for example, Altonji, 1995).

General health status

Comprehensive clinical evaluations of health status as undertaken in, for example, the United States National Health and Nutrition Examination Surveys are far too expensive to be included in a typical socioeconomic survey, except on a small scale or with very selected samples. This is especially the case in lower income settings where health service infrastructure is relatively weak. Most household surveys have, therefore, relied on interviews with respondents who provide an assessment of their own health.

Within the class of self-evaluations, general health status (GHS) is probably the most widely used indicator in the empirical literature in the United States. Some advocates have argued that it is the best single health index available, citing, for example, the fact that GHS is correlated with subsequent morbidity and mortality (see, for example, Ware et al., 1978).

GHS, however, suffers from two key drawbacks. First, respondents are typically asked to rate their health in one of four or five discrete categories ranging from excellent to poor health. Relying on such a small number of discrete categories cannot possibly do justice to the complexity and diversity of health status of individuals and this is one reason we argue for using multiple indicators simultaneously.

The second problem has to do with measurement. "Good" health may not mean the same thing to all people and respondents are not provided with an established metric against which to compare their own health. Indeed, few surveys provide a clear definition of a reference health status and there is seldom an explicit reference group. Because questions about GHS are typically vague, we have no idea whether respondents are rating their health relative to the national average, relative to the neighbour, or to whom.

Moreover, self-evaluations reflect perceptions of health. While important, perceptions are likely related to values, background, beliefs and information, all of which are systematically related to socioeconomic characteristics, including wages and income. For example, information about own health status is almost surely correlated with the extent of concurrent and prior use of health care because people who have used the health care system are likely to be better informed. Since most people assume they are in good health unless they have information to the contrary, it is plausible to suppose that conditional on a level of health status, those who have little exposure to the health system are likely to report themselves as being in better health. If this is true, given that lower income people are less likely to use health care, particularly in poor societies, measurement error in GHS will be systematically related to income (and wages). There is evidence in the RAND Health Insurance Experiment (HIE) that suggests this is a legitimate concern, at least for those people who, in the baseline, were in poor health and in the bottom quintile of the income distribution. Among these people, those randomly assigned to receive free care used more health care and were, at the end of the experiment, in better health as measured by clinically-evaluated outcomes (such as blood pressure)

and risk of subsequent death. But according to their own evaluation, measured by GHS, their health actually *worsened* (Newhouse et al., 1993). Similar evidence is reported from a health price experiment in Indonesia; where people used more health care in those places where prices were lower, their health improved but their self-reported GHS was worse (Dow *et al*., 1997).

We conclude that GHS does provide information in that it predicts later health problems. But, there are good reasons to be concerned that it is contaminated by measurement error, which is correlated with socioeconomic characteristics, including income.

Self-reported morbidity, illness and "normal" activity

Disease-oriented definitions of health status, favoured by many clinicians and some epidemiologists,[4] have the advantage of a foundation in medical practice. However, even in a clinical setting, it is difficult to unambiguously diagnose all problems, especially in the presence of multiple, inter-related problems. Furthermore, from a social science and public health perspective, it is often the functional consequences of ill-health that are of primary interest and those consequences typically cut across diseases and are exacerbated by interactions among different diseases.

As a low cost alternative to clinical evaluations of health status in a household survey, some studies have drawn data from health facility records. But, in low income countries a significant proportion of the population does not use these facilities, and those that do tend to be a select group who are not necessarily those in worst health. In fact, they tend to be higher income people and so it will be very hard to infer anything about the relationship between health and well-being in the whole population without at least some information on the mechanisms underlying the choice to visit a health care facility.[5]

Several surveys have asked questions about illness, or specific symptoms (such as fevers, diarrhea, respiratory problems) during a reference period. As with GHS, these self-evaluations are difficult to interpret if what is deemed an "illness" or a "symptom" is not the same thing for all respondents. In fact, in surveys from low income countries, it is not unusual for the poorest to appear to be the most healthy by this metric! For example, in Ghana and the Côte d'lvoire, the propensity for adults to report being ill in the last four weeks is positively associated with own education (Schultz and Tansel, 1997) and with per capita household expenditures (Over et al., 1992). As with GHS, these indicators

[4] See the discussion, for instance, in Jamison et al. (1993) and World Bank (1993).

[5] For example, Gertler and van der Gaag (1990) report that in Côte d'Ivoire, only 40% of persons who said they were ill in the last month sought medical care at a modern facility and they tended to be from higher income households. A similar pattern is observed in many countries.

are likely to be measured with error that is correlated with use of the health system (and thus income and the price of health services). In addition, other price incentives may influence self-reported morbidity. For instance, an individual may claim to suffer from an illness in order to become eligible for health-related benefits (Parsons, 1982; Bound, 1991).

Another commonly used variant of self-reported illness is to ask whether any days of "normal" activity were lost to ill health. Some argue that this measure is less likely to be contaminated by systematic respondent error. Apart from the fact that "normal" is not well-defined, people whose opportunity cost of time is high (the better educated, say) will have less incentive to miss activities. By this metric, they will appear to be in better health than people with a lower value of time (conditional on a particular "true" health status). This is in contrast with the argument above whereby the better educated report themselves as being in *worse* health because of better knowledge or greater exposure to health services. While the net impact is unclear, there is no reason to expect a closer correspondence between normal days lost with "true" health than with days ill or some other indicator. As it turns out, reported correlations between days lost and education have been both negative (Hill and Mamdani, 1989) and positive (Schultz and Tansel, 1997). Moreover, people with acute health problems are likely to make life-style and occupation choices in response to these problems making it very difficult to interpret "normal".

In general, for all of these measures and for GHS, it is going to be very difficult to separate true health status from measurement error. The key point for our purposes is that this measurement error is likely to be correlated with socioeconomic behaviours and outcomes, such as labour force participation, productivity and wages.

Self-reported physical functioning

Several household surveys have collected information on difficulties with physical functioning that are considered normal activities for people who are in good health; these might include walking a specified distance, lifting a particular weight, bending or climbing stairs. While the notion of "difficulty" is subjective, questions about specific activities of daily living (ADLs) are more precisely defined than "being ill" or "normal activities" and there is some evidence that relative to GHS and morbidity, ADLs are less prone to the type of measurement error discussed above. (See the discussion in Strauss et al., 1993; Schoenbaum, 1995, and Dow et al., 1997, present corroborating evidence.) This is an important advantage of ADLs in a survey setting. However, they have one key drawback: the limitations in physical activities that ADLs typically capture are frequently due to physical health problems such as shortness of breath, joint problems back problems (Stewart et al., 1978). Few prime age adults have difficulty with these activities and so ADLs may not be as useful in studies of health and labour outcomes as they have proved to be in the gerontological literature.

Nutrition-based indicators: Nutrient intakes

Nutrient intakes have been prominent in the empirical literature. It is presumably net energy intake that is related to productivity but since energy expenditures are difficult to measure, studies have tended to focus on nutrient intakes.

Measurement of energy consumption is not nearly as straightforward as might be thought. Several different methods have been adopted and each involves its own strength and weaknesses. Calorie *availability* is computed by converting food quantities (purchases and consumption from own production) into nutrient intakes using standard food composition tables. This has the advantage of being relatively easily calculated using data commonly collected in many household expenditure and farm production surveys. But, the method suffers from several potentially important sources of systematic bias.

First, it assumes no food is wasted: everything that is available is converted into nutrients. It is plausible that very low income (or low wage) households waste less than those that are better off, in which case nutrient intakes will tend to be upward biased and the bias will increase with income. Second, it is very difficult in consumption and production surveys to take into account all meals that are given to guests or employees and all meals that are received in-kind. For example, the members of a household that entertained many guests during the survey period will appear to have consumed more calories than they actually did consume.[6] If the (net) receipt of in-kind food (including transfers and gifts) declines with income and the probability of having guests rises with income, then nutrient intakes will be biased and the bias will, also, be positively correlated with income and wages. As a third example, it is very difficult to measure nutrient intakes for meals eaten away from home; typically, it is assumed that those meals have the same calorie content as meals at home. Clearly, this need not be true. For example, low-income workers are often given food at work; if it is more nutritious than the food eaten at home, measurement error will be negatively correlated with income.

There is *prima facie* evidence that these sorts of leakages (from wastage, food given or received in kind, and meals away from home) are important and they result in systematic errors in estimated nutrient intakes. In most studies that use estimates based on availability, a substantial fraction of the population appears to be consuming at unrealistically low or high levels. Srinivasan (1992), for example, reports that in the 1976 National Sample Survey of India, over 5% of rural households consume fewer than 1,500 calories per capita per day and almost 20% consume over 4,000 daily calories per capita!

An alternative method of collecting nutrient consumption is to obtain information on *intakes* rather than on *availability*. One method weighs ingredients, prior to each meal,

---

[6] While uncommon, it is possible to draw up an entire list of all people at all meals for a week (in a consumption survey) but that is not a realistic option over a longer survey period (such as in a farm household survey).

and wastage, after it, and then converts measured consumption into nutrients. Although meals eaten away from home are not captured, this method is probably the most accurate that has been used to date. The method is prohibitively expensive to field in large-scale household surveys and has, therefore, been used very infrequently.[7]

Thus, the most common strategy used to collect nutrient intakes has been to ask respondents to recall ingredients that went into meals consumed, usually over the previous 24 hours. While this approach has the advantage of potentially excluding leakages (such as meals for guests, transfers of food and wastage of food), it comes at a substantial cost in terms of survey time. Moreover, since there is considerable variation in eating habits, 24-hour recalls are likely to be very noisy; extending the recall period raises concerns about recall bias (which is thought to rise rapidly in this context) and multiple visits further raises the cost of collecting these data.

Systematic evaluation of the accuracy of these different methods of collecting nutrient data is scarce and there are virtually no surveys that contain more than one measure. A survey in Bukidnon, Philippines, conducted by Bouis and Haddad (1992) is an important exception. It has two different measures: intakes from 24-hour recalls and calorie availability from a month's recall of purchases. The survey was conducted over four rounds and the household average of reported daily intakes over the four rounds are very similar: 2,439 per adult equivalent for availability and 2,358 for 24-hour recall although the availability measures have somewhat fatter tails. Key for our purposes is that the difference between availability and intakes for the same household is positively correlated with household characteristics including income, expenditure and education of the head; among low income households, availability tends to be less than intake, but higher income households report higher availability than intakes. This is all consistent with the idea that leakages are missed in availability and tend to be systematically associated with income.

It is possible to get some idea of the relative importance of random noise by regressing intake on availability and vice-versa. If 24-hour recalls are measured without error, then the coefficient when it is the regressand should be 1: it is very close (0.97 with a standard

[7] The method was used in the Brazilian ENDEF survey. To account for the fact that there is a good deal of variation in daily intakes, each of 55,000 households was visited on a daily basis for a week, All food that would be prepared in the following 24 hours and wastage from the previous 24 hours was measured at each visit. The enumerator also listed all household members and guests present at every meal during the previous 24 hours at each visit. It is, therefore, possible to control for a good part of the leakages discussed above. Of interest is the fact that among the poorest households, there is evidence of a decline in intakes after the first day or two; the survey staff attributed this to an initial attempt by the respondents to impress the enumerators (or downplay their own poverty) (Marcello Vasconcellos, personal communication).

error of 0.01). But, as suggested by the fatter-tailed distribution, availability seems to be measured with more noise: when it is the regressand, the slope is significantly smaller (0.86 with a standard error of 0.01). In these data, availability appears to be a worse indicator of calorie intakes, both from the point of view of random and systematic errors.

It is apparent, therefore, that all nutrient intakes are likely to be subject to some random measurement error that is probably greatest for recall-based methods. In addition, availability-based measures suffer from systematic errors that are probably correlated with income, wages and productivity. Since the nature of the errors and the correlations with labour outcomes are likely to differ depending on how intakes are measured, it seems safe to conjecture that interpretation of the impact of nutrient intakes on wages will not be independent of the way the data have been collected.

Nutrition based indicators: Anthropometrics

Having discussed nutrient intakes, which are inputs in a health production function, we turn next to outputs, specifically anthropometrics such as height, weight and body mass index (BMI). It is reasonable to wonder how much information on health can be embodied in anthropometry. Adult stature has been profitably used as an indicator of well-being in both the economic history and development literatures (see the survey by Steckel, 1995, and the discussion in Strauss and Thomas, 1998). Moreover, child height has proven to be an informative indicator of the nutritional status of children and is viewed as a longer-run indicator of nutritional status (see, for example, Waterlow et al, 1977; Falkner and Tanner, 1986; Waterlow, 1988). While height is clearly determined by the time an individual reaches adulthood (apart from shrinking later in life), there is some debate in the literature about the extent to which adult stature is completely determined by the time the child has progressed beyond early childhood. Nevertheless, attained adult height will clearly reflect in part human capital investments made during early childhood.

Whereas height is predetermined by adulthood, weight varies in the short run and so provides a more current indicator of nutritional status. It is hard to interpret weight alone since a light person may also be small, and thus not underweight given height (and, conversely, a heavy, tall person may not be overweight). Nutritionists have found it convenient to analyse weight given height. There are many potential ways of expressing this ratio and one that has been commonly used for adults is body mass index (BMI), the ratio of weight (in kilograms) to height (in metres) squared. On average, a prime age male in the United States has a BMI of about 25; BMIs are considerably lower in poor countries, averaging 22.4 in Cote d'Ivoire, for instance.[8] Extremely low (below 18) and

[8] Moreover, the fraction of the population at either extreme varies dramatically across countries. In the United States, for example, only 1% of men have BMIs below 18, while 6% are below 20 and 13% above 30. In Cote d'lvoire, 3.5% of adult men have BMIs below 18. As a more extreme case, in Viet Nam, where the BMI of the average male is

high (above 30) values have been associated with higher adult mortality (Waaler, 1984; Fogel, 1994) but BMI and mortality are essentially unrelated within that range. The relationship between adult BMI and subsequent mortality, the so-called "Waaler curve", is U-shaped with a long, flat bottom between around 20 and 30. The causal mechanism underlying this association has not been established and the precise definition of extreme values remains debated.

Because they are relatively inexpensive to collect, in many surveys, height, weight and sometimes arm circumference have been measured in the field by an anthropometrist. While the measures may be subject to random error (which can be reduced by good field procedures), a key virtue of these health indicators is the absence of measurement error that is systematically correlated with respondent characteristics (such as income). (But, see Strauss and Thomas, 1996, for a discussion of the National Longitudinal Survey in the United States, which relies on self-reported height and weight, which are measured with error that is related to income.)

Empirical implications of measurement error in health status

Throughout the discussion, we have taken care to distinguish random from systematic errors in measurement of health. This is important since estimated effects of health on labour outcomes will vary with the type of error as will strategies that seek to ameliorate the influence of measurement error. We therefore discuss each in turn although it is important to recognize that a particular health indicator may be affected by both types of error, albeit to different degrees.

Random measurement error in health will bias estimated effects towards zero; this is a case of classic errors-in-variables. It is difficult to think of any health indicator that is not subject to at least some random measurement error. Even weight and height measured by a trained anthropometrist are not immune to error. For example, subjects may not be aligned properly on measuring boards, scales may not be re-calibrated and so forth. Moreover, many indicators vary over the course of a day (like weight) or across days (like nutrient intakes or functional problems). In all of these cases, repeated measures can ameliorate the impact of random error in regression models.

Systematic measurement error poses thornier problems. Consider the example discussed above in which poor health is more likely to be reported if the respondent has more contact with a modern health practitioner. If health care utilization rises with income, then higher wage individuals are more likely to report themselves as ill given a particular level of underlying health status: the impact of income on health will be negatively biased.

---

below 20, 22% are below 18, 66% are below 20 and less than 1% are above 30. The distributions for women are more fat tailed: 3% are below 18 and 18% are above 30 in the United Sates, 5% below 18 in Côte d'lvoire and 24% in Viet Nam.

In contrast with random errors, averaging multiple reports by the same respondent is not likely to help reduce the impact of systematic error. Moreover, it is not easy to think of instruments that are likely to be correlated with "true" health status, but uncorrelated with wages. For example, if the reason people report themselves as being sick is because they are better informed, which is, in turn, a reflection of the availability of health facilities in the community, health infrastructure will be correlated with the measurement error and so is not a valid instrument. Assuming that a person's propensity to report themselves as ill changes slowly over time, repeated observations on the same individual may help since examining differences in health status (transitions) will reduce the influence of systematic measurement error, but at the expense of increasing the relative magnitude of random error.

In general, measurement error in health status or inputs is likely to be partly random and, in many cases, partly systematic. Obviously, the best way to reduce measurement error is to pay more attention to measurement. In some cases this involves taking averages across repeated measures (to reduce random noise), examining differences (to reduce systematic error), taking special care in fieldwork to avoid systematic errors (by carefully measuring "leakages" when collecting nutrient intake data, for example), or by avoiding certain indicators that are especially prone to error. Recent innovations in survey methodology offer some potentially exciting opportunities in this regard. For example, a small number of household surveys have experimented with greater reliance on direct observation such as assessing the incidence of anemia (based on haemoglobin counts), tuberculosis (using sputum) and hypertension (based on blood pressure). In addition, studies have sought to measure glucose levels (with saliva), net energy intake (with labelled water) and lung capacity (using peak flow metres) and also to directly observe functional limitations through timed moves (such as walking a particular distance, standing from a sitting position and so on). Whether these measures will prove to be fruitful in studies of the links between health and productivity remains to be seen.

4. Example of socioeconomic correlates of child growth and adult body mass in Côte d'Ivoire

As an example of how one might go about estimating a reduced form or conditional demand function for health we take up the case of investment in child and adult health in Côte d'Ivoire. We focus on health outcomes and use as our measure of child health, height standardized for age and sex, and weight standardized for height.[9] We use body mass as

[9] We explicitly avoid considering the demand for health inputs. See Gertler and van der Gaag (1990) and Mwabu et al. (1993) for examples of modelling the use of health care facilities, using African data.

our measure of adult health. We relate these health outcomes to household-level factors and also to local investments in health facility infrastructure and to food prices. The discussion and estimates come from Thomas et al. (1996). To motivate the issue and how it is related to more general issues of poverty, we briefly review the economic situation in Côte d'Ivoire.

The Côte d'lvoire enjoyed rapid growth during the two decades after independence (averaging 7% per annum during the 1960s and 1970s); government expenditures grew even faster, with investments in the health sector being concentrated on large hospitals (in urban areas).[10] During the 1970s, there was a shift in focus towards the provision of basic health services to the entire population and so during that decade a large number of health centres were opened throughout the country, especially in rural areas.

Growth rates were far lower during the 1980s and were negative for the first few years of the decade (1979–1983). The government responded to the economic downturn by instituting a series of economic adjustment programmes that included substantial cuts in public spending. Real government spending rose by about 60% during the 1970s and peaked in 1982 when it accounted for 17% of GDP. Over the following four years, however, the public budget declined in real terms by 14% and expenditures on social services saw even deeper cuts. In 1981, 7.7% of the public budget was spent on health; that share declined to below 7% in 1986, although the share has recovered slightly since then. In real terms, however, public resources devoted to health declined by 12% between 1981 and 1984, which more than offsets the growth in health expenditures during the previous decade. In 1984, therefore, real public spending on health was reduced to its level in 1972. Although private health care is emerging in Côte d'Ivoire, it remains small and concentrated in Abidjan.

Within the health sector, personnel costs remained resilient to cuts in public spending during the 1980s and their share of the public health budget rose during the early years of the economic adjustment programme and so, in 1984, personnel accounted for three-quarters of the total health budget. Thus real public expenditures on medicines and materials were cut by over one-third during the first half of the 1980s. Although health policy has shifted in focus away from large hospitals, trained personnel remain heavily concentrated in large, urban facilities. It is likely, therefore, that less well equipped (rural) facilities bore the brunt of cuts in the health budget during the early 1980s.

Historically, food prices in Côte d'lvoire have been below world prices, partly because of over-valued exchange rates (Krueger et al., 1988). During the 1980s, however, food prices tended to rise (relative to non-food exports) as public policy shifted towards promoting

[10] The following discussion is based on data drawn from République de Côte d'Ivoire (1979, 1981, 1985, 1990); République de Côte d'lvoire/UNICEF (1988); and Economist Intelligence Unit (1986, 1987, 1990, 1991).

food self-sufficiency. Food prices also tended to rise relative to non-foods during the 1980s and particularly after 1985. In 1988 there was a dramatic increase in the relative price of foods; the food price index (for lower income households) rose by over 20% while the general price index rose by less than 4%. Unfortunately, no time series data are available on consumption prices of individual food goods.

How have these changes affected the health of the Ivorian population? The next section sketches a conceptual framework that guides our attempt to answer this question. We also discuss some potentially important econometric issues and introduce the estimator we have adopted. The data are described in the third section, which is followed by our empirical results.

In Côte d'Ivoire, the availability and quality of health infrastructure affects child health and, in particular, height for age. Food prices have a bigger impact on the shorter-run health indicators: weight for height of children and body mass index of adults. In the case of children, price effects are (absolutely) larger among poorer households. Adult health, and in particular women's BMI, is significantly improved as levels of resources in the household rise, although income effects on child health appear to be small.

Data

To study the determinants of anthropometric outcomes, we use individual and household level data from the third wave of the Côte d'Ivoire Living Standards Survey (CILSS) carried out in 1987/88. The survey has a two-stage sampling design; 100 clusters were randomly drawn from a national frame and then 16 households were drawn from each cluster. Half of the clusters are from rural areas, the other half from the urban sector. Simultaneously with the household survey, enumerators collected information on local food prices and, in 1989, a very detailed health facility survey was specially conducted in exactly the same communities covered by the household survey. We match these community level data with the household and individual information to create a very rich multi-level database for our empirical analysis of the impact of health infrastructure and food prices on anthropometric outcomes of Ivorians.[11]

Community characteristics

The health facility data are very detailed, providing extensive information on both the availability and the quality of services in the community. For example, the survey reports the number of doctors, pharmacists, dentists and surgeons who are supposed to

[11] Child anthropometric data from the first two rounds of the CILSS are described in Strauss and Mehra (1989). The household level determinants of child anthropometry are presented in Strauss (1990) for the first round and Sahn (1994) for the second round; Strauss (1990) also examined the impact of community level infrastructure in the rural sector. The demand for health care in Côte d'Ivoire is discussed in Gertler and van der Gaag (1990).

work at the facility as well as the number who were present in the last 24 hours. On average, there were almost two doctors per facility on the books but only 85% of them had been at work during the last day. In the urban sector, there were 3.4 doctors per facility and almost 90% had been in attendance, whereas in the rural sector, not only is the number of doctors on the books smaller (0.53) but only 75% had been at work during the last day.

Summary statistics of the health service measures used in the regressions are reported in the top panel of Table 1, reproduced from Thomas, Lavy and Strauss (1996).[12] On average, there are about seven nurses per doctor and 10 to 15 support personnel (drivers, cleaning staff, administrators) per doctor. While every community has at least one nurse (and one support worker), there is at least one doctor in about 90% of urban communities but one in only 40% of rural communities.

[12] For each of the 100 clusters, the nearest facility was included in the survey. If that facility was not public, then the nearest public facility was also included. In addition, the nearest drugstore or pharmacy was enumerated as was the nearest source of family planning. It turns out that there are very few private health facilities in Côte d'Ivoire: of the 118 facilities surveyed, only 3 are private. In those clusters with more than one facility included in the survey, a service is considered available if it is provided at any of the facilities in the survey; the other measures of health services, such as numbers of doctors, are the sum across all facilities within the community.

Table 1: Sample characteristics-Means and [standard errors]

A: Community

	All		Urban		Rural	
Health personnel						
# doctors, surgeons, dentists present	1.68	[0.24]	2.96	[0.39]	0.40	[0.11]
# nurses, nursing staff present	11.53	[1.53]	19.74	[2.52]	3.32	[0.57]
# support personnel present	19.16	[2.90]	32.12	[5.07]	6.20	[1.22]
Infrastructure						
# usable beds in facility	15.30	[2.07]	20.08	[3.33]	10.52	[2.31]
# working vehicles in facility	1.37		1.87		0.87	
Drug availability						
Proportion with antibiotics in stock	0.77		0.78		0.76	
Aspirin in stock	0.85		0.86		0.84	
Quinine in stock	0.74		0.74		0.74	
Services						
Proportion with immunizations available	0.68		0.78		0.58	
Childbirth facilities	0.60		0.58		0.62	
Growth monitoring programme	0.57		0.70		0.44	
Community prices						
Beef with bones	0.823	[0.013]	0.863	[0.013]	0.784	[0.021]
Fresh fish	0.429	[0.016]	0.427	[0.022]	0.431	[0.022]
Rice (imported)	1.617	[0.008]	1.614	[0.012]	1.620	[0.009]
Palm oil	0.637	[0.022]	0.671	[0.039]	0.603	[0.022]
Eggs (chicken)	0.083	[0.002]	0.065	[0.002]	0.096	[0.003]
Sugar	0.464	[0.070]	0.605	[0.138]	0.324	[0.016]
Plantain	0.118	[0.012]	0.135	[0.015]	0.101	[0.019]
Manioc (unprocessed)	0.111	[0.011]	0.134	[0.019]	0.089	[0.009]

B: Children

	All		Urban		Rural	
Height for age						
z-score	-0.728	[0.03]	-0.586	[0.04]	-0.856	[0.04]
% US median	96.93	[0.11]	97.54	[0.16]	96.39	[0.16]
per cent <90% median	13.6		10.1		16.8	
Weight for height						
z-score	-0.341	[0.02]	-0.308	[0.03]	-0.370	[0.03]
% US median	97.92	[0.22]	98.30	[0.35]	97.57	[0.28]

per cent <80% median	5.1		4.6		5.5			
Child characteristics								
(1) if child of hh head	0.75		0.77		0.74			
(1) male (0) female	0.51		0.49		0.53			
age of child (months)	70.15	[0.64]	70.81	[0.93]	69.55	[0.89]		
Household characteristics								
Per capita expenditure (PCE) (CFAF 000)	163.52	[2.32]	225.88	[4.28]	107.56	[1.23]		
ln(PCE)	11.77	[0.01]	12.11	[0.01]	11.46	[0.01]		
senior female height for age z-score	-0.78	[0.02]	-0.59	[0.02]	-0.95	[0.02]		
(1) can write	0.19		0.33		0.06			
years of education	1.51	[0.05]	2.74	[0.09]	0.41	[0.03]		
senior male height for age z-score	-1.16	[0.02]	-0.91	[0.02]	-1.38	[0.02]		
(1) can write	0.36		0.55		0.18			
years of education	3.08	[0.07]	5.04	[0.12]	1.32	[0.06]		
(1) female head/spouse exist	0.99		0.98		0.99			
(1) male head/spouse exist	0.95		0.91		0.97			
(1) head Ivorian	0.83		0.80		0.87			

C: Adults

	All		**Urban**		**Rural**		**Female**		**Male**
Body mass index (BMI)	22.38	[0.05]	23.26	[0.09]	21.52	[0.06]	22.50	[0.08]	22.23 [0.07]
ln (BMI)	3.10	[0.003]	3.13	[0.004]	3.06	[0.003]	3.10	[0.004]	3.09 [0.004]
% BMI = 18	4.29		2.83		5.76		4.96		3.49
% BMI = 28	5.55		9.46		1.65		7.15		3.61
HH PCE (CFAF 000)	207.80	[3.78]	291.68	[6.95]	125.63	[1.95]	190.43	[4.49]	228.85 [6.32]
ln(PCE)	11.94	[0.01]	12.31	[0.02]	11.57	[0.01]	11.87	[0.02]	12.02 [0.02]
years of education	2.97	[0.07]	4.91	[0.12]	1.08	[0.06]	1.81	[0.08]	4.38 [0.12]
(1) if male	0.45		0.48		0.42				
(1) head of household	0.34		0.37		0.31		0.06		0.68
(1) spouse of head	0.38		0.33		0.42		0.68		
(1) Ivorian	0.84		0.80		0.88		0.85		0.83
(1) married	0.73		0.68		0.80		0.78		0.69

From Thomas, Lavy and Strauss (1996).

About 85% of clusters have facilities with in-patient services. On average there are about 19 beds per cluster, but of these almost 4 were unusable at the time of the survey. In the urban sector, facilities tend to be larger (22 beds per facility) and have a larger proportion of useable beds (85%). The average cluster in the rural sector has 15.5 beds, of which almost a third were unusable, leaving only 10.5 available beds.

Apparently in all these measures of services, there are substantial differences in the resources that are supposed to be available and those that are actually available in each facility. Few facility surveys that can be matched with individual data have collected this sort of detail, although as we shall demonstrate below, the inclusion of poorly-measured indicators of service availability can be very misleading.

Several measures of drug availability are reported in the survey; we focus on three drugs—antibiotics, aspirin and quinine—and whether they were in stock at the time of the survey. Drug stocks are very similar across sectors and about three-quarters of clusters have access to each drug at any time. The facility survey also reports whether particular health services were available. About two-thirds of communities have access to immunization services and about 60% have childbirth and growth monitoring services. Whereas immunizations and child growth monitoring is significantly more common in urban areas, childbirth facilities are slightly more common in the rural sector.

In conjunction with the household survey, a price survey was also conducted. Enumerators were instructed to visit local markets, purchase a range of standardized commodities, and record the weight and cost. For each commodity, up to three prices are recorded. Cluster average prices are reported in the lower part of Table 1, panel A. Critical for this study is the fact that there is substantial cross-sectional variation in prices in Côte d'lvoire. In general, prices tend to be lower in the rural sector, particularly for those commodities that do not require much processing such as sugar (for which the average urban price is nearly double that in the rural sector), plantains and (unprocessed) manioc.[13] Prices of imported products, such as some types of rice, tend to be higher in rural areas. There is also a good deal of heterogeneity in prices within rural and urban areas.

It is these comprehensive and extremely detailed health facility and price data that will form the community covariates in the regressions below. The CILSS also collects information on infrastructure in the community, but only in the rural sector; since we wish to examine the entire country (and stratify the data in several different dimensions), we chose not to use these data.[14]

---

[13] As reported by Kanbur and Grootaert (1994), food prices in Abidjan are about 40% higher than those in the northern Savanna region.

[14] The rural infrastructure data have been analysed by Strauss (1990) using the first round of the CILSS. Moreover, experiments with the data indicate our results are robust to their inclusion (where possible). Our experience has been that with these sorts of data, correlations across groups of characteristics (such as health services and availability of modern sewerage) tend to be quite small although intra-group correlations are typically quite high.

Individual and household characteristics

The characteristics of individuals and households are drawn from the household survey. The CILSS is a broad purpose socioeconomic survey that provides information on household composition and demographics, time use including labour supply, labour income and self-employment enterprises, non-labour income and asset ownership, commodity consumption, and the health of each individual in the household. The breadth of the survey permits the inclusion of comprehensive controls for household resources: in both the child and adult anthropometry regressions, we include (the logarithm of) per capita expenditure[15] together with other indicators of human capital in the household.

All children under 12 years of age are included in our sample.[16] Since child height and weight are systematically related to age and gender it makes sense to control for these characteristics. One strategy would be to include high order polynomials in age and

---

[15] It includes the value of all consumption expenditures and the flow value of durable consumption. Food consumption is the value of purchased foods plus direct estimates of the value of food consumed out of own production. The latter is important among rural households and accounts for about a third of their expenditures. The imputed value of goods in-kind, such as employer-provided transport and housing, are included in expenditure. The flow value of durable consumption includes imputed rents that are based on hedonic housing regressions controlling for selection of owner-occupiers. In these data, the regressions have R^2 varying between 0.64 and 0.84 (see Glewwe, 1987, for a description of the regressions). Other lumpy expenditures (such as durable purchases and large expenditures on ceremonies or health care) are excluded. If market purchases are prone to less measurement error than imputed values, then if we ignore issues of endogeneity of consumption, market expenditure is a good instrument for total expenditure. Some experiments with this specification are reported below.

[16] All households were visited by an anthropometrist twice within a fortnight. Those household members who were absent at the first visit were measured at the second visit. If data checks suggested a measurement from the first visit was unlikely, that person was re-measured, as were a random sub-sample of individuals. Nevertheless, for about 10% of the children included in the survey, we do not have valid height for age measures (either because they were not measured or because the recorded measurements are obviously wrong). Children included in the sample come from slightly higher income and better educated households: comparing the 500 excluded children with those in the sample we analyse, lPCE is 11.68 and 11.77, respectively; the senior male in the household has 2.3 and 3.1 years of education, respectively. Another 10% of children are excluded from the weight for height regressions (for the same reasons) and there are no differences in the household characteristics of those children included and excluded from this sample except that the excluded children tend to be slightly older.

gender in the regressions. We adopt a more parsimonious approach and relate each child's height and weight (conditional on height) to that of a well nourished United States child of the same age and gender (using the NCHS, 1976, tables). Relative to the U.S. median, the average Ivorian child is about 97% as tall and has a mass, given height, of about 98%. Slightly less than 14% of Ivorian children are chronically malnourished (or stunted, height for age less than 90% of the U.S. median) and about 5% are acutely malnourished (or wasted, with weight for height less than 80% of the U.S. median). Urban children in Côte d'Ivoire tend to be taller, heavier given height, and less likely to be stunted or wasted.

The household roster in the CILSS records the relation of each individual to the household head. About one-quarter of the children in the sample are not children of the head or spouse; most of these children are other relatives of the head. Recent studies of child anthropometry have restricted the sample to those children for whom the parents' characteristics are reported (see Strauss, 1990, and Sahn, 1994, on the Côte d'Ivoire) in order to control for household background (Thomas et al., 1990). If we restrict the sample to those children for whom maternal characteristics are reported, we lose 17% of the children; requiring observations on paternal characteristics as well cuts the sample another 13%. Yet, it is these 30% of all children who may be the most vulnerable to price fluctuations and reduced social spending: we therefore prefer not to exclude them from our analysis. We choose, instead, to include the characteristics of the head and spouse as measures of household resources in the child anthropometry regressions.

If the head is male, we assign his characteristics to the *senior male*, grid the characteristics of his spouse to the *senior female*. If the head is a woman, then her characteristics are assigned to the *senior female*. The education of the head and spouse are likely to affect resource availability, through their impact on the value of time and also, possibly, through resource allocation decisions. About a fifth of senior women and almost twice that proportion of men are literate and men have spent about twice as long at school as women; these differences are even wider in the rural sector. The (standardized) height of the senior male and female are included as controls for genetics, endowments and other background factors. Controls for the existence of a senior male and female in the household as well as whether the head is Ivorian are included.

Characteristics of adults aged 20 through 60 are included in the final panel of Table 1. Body mass index (measured in kilograms per metre squared) is, on average, 22 and is slightly higher in the urban sector; men have slightly lower BMIs. Biomedical evidence suggests that a BMI below 18 or above 28 is associated with higher mortality risk; around 5% of Ivorian adults are in this range with relatively more in the upper range in the urban sector and more in the lower range in the rural sector. The distribution of BMI among women is fatter tailed, and so they are more likely to have very low or very high BMIs than men.

Households in the child sample tend to be early in the life cycle and thus have relatively young heads whereas those in the adult sample are later in the life-cycle: per capita expenditure is thus higher in the adult sample. Slightly less than half the sample is male, a third are heads of their own households, another third are spouses and almost three-quarters are married. Urban residents are much better educated than their rural counterparts and, on average, men have received more than twice the education of women.

Results

Estimates of the conditional anthropometric output functions (8) are reproduced from Thomas et al. (1996) as Table 2, child height for age and weight for height, and as Table 3 for adult body mass index. All these models include (the logarithm of) per capita expenditure, which is treated as endogenous: instruments used in the tables include (polynomials of) the value of land owned and livestock, financial and business assets.[17] Test statistics in the tables are based on heteroskedasticity consistent estimates of the variance-covariance matrix, which permit intra-cluster correlations and take account of the endogeneity of per capita expenditure.

Child anthropometry

The community and household determinants of child height for age and weight for height are reported in Table 2 for all children in the sample and separately for those living in the rural and urban sector. In Table 3, the data have been stratified by age of the child (into two groups, less than three years and older than three) and also by education of the head or spouse (again into two groups, households in which neither the head nor spouse has

[17] Over 50% of the variance of lnPCE is explained in the first stage regression for the whole country. In the child height for age regressions, reported in Table 2, for example, the R^2 is 0.54 for the entire country with an $F_{58,4180}$ of 94.7. Taken together, the identifying instruments are jointly significant ($F_{15,4180} = 17.8$), with the value of land and livestock ($F_{6,4180} = 14.8$) and financial assets ($F_{2,4180} = 35.4$) being particularly important determinants of lnPCE. In the urban sector, the R^2 is also 0.54 ($F_{57,1977} = 44.2$) and the value of financial assets is a key instrument ($F_{2,1977} = 50.9$). The regressions are not as good in the rural sector where only a third of the variation in lnPCE is explained by the covariates ($F_{56,2203} = 20.6$) and, not surprisingly, the value of land and livestock are important determinants of household consumption ($F_{6,2203} = 11.0$). GMM overidentification tests indicate the instruments are valid; for example, p-values of χ^2 tests for the significance of the instruments in explaining the residual from the height for age regressions are over 30% in the urban and rural sectors. Furthermore, Durbin–Wu–Hausman tests for the erogeneity of income are rejected; for example, in the urban sector, the t statistic on unexplained lnPCE from the first stage regression is over 4 in a subsidiary regression.

any education and households in which either has some education). The dependent variables have been standardized on age and gender and are expressed as z-scores.[18]

[18] The descriptions in the last section were based on percentages of U.S. medians to be comparable with other studies (Waterlow et al., 1977). In order to account for the fact that measures of both location and dispersion of child anthropometric outcomes vary with age and gender, we use z-scores, $(h - h_{med})/\sigma_h$, in the regressions.

Table 2*:* Community determinants of child height for age and weight for height

(z score)	**Height for age**				**Weight for height**			
(Children under 12)	*All*	*Urban*	*Rural*		*All*	*Urban*	*Rural*	
Health services and facilities								
Health personnel (*10)								
#doctors, surgeons,	0.757	0.592	-1.902		-0.339	-0.406	0.994	
and dentists	[2.96]	[2.11]	[1.32]		[1.27]	[1.30]	[0.75]	
# nurses, nursing staff	-0.121	-0.150	0.315		0.091	0.118	0.394	
	[2.25]	[2.73]	[0.83]		[1.64]	[1.88]	[1.27]	
# support staff	-0.029	-0.021	0.411		0.020	0.028	-0.322	
	[1.07]	[0.75]	[1.82]		[0.72]	[0.88]	[1.90]	
Infrastructure								
# usable beds (*100)	-0.286	-0.209	-1.984		-0.479	-0.151	0.117	
	[1.77]	[0.99]	[3.06]		[3.34]	[0.73]	[0.24]	
# working vehicles	0.010	0.042	-0.087		0.001	-0.009	0.062	
	[0.72]	[2.72]	[1.41]		[0.04]	[0.60]	[1.31]	
Drug availability: (1) if								
Antibiotics in stock	0.236	0.268	0.236		0.111	0.221	0.052	
	[3.23]	[2.85]	[1.28]		[1.72]	[2.52]	[0.38]	
Aspirin in stock	-0.032	0.057	-0.175		0.033	0.156	-0.188	
	[0.39]	[0.49]	[1.16]		[0.44]	[1.48]	[1.45]	
Quinine in stock		0.093	-0.194	0.291		-0.005	-0.018	-0.013
	[1.31]	[1.82]	[2.17]		[0.07]	[0.18]	[0.13]	
Services: (1) if available								
Immunizations	0.209	-0.124	0.258		-0.010	0.037	0.010	
	[2.94]	[1.05]	[2.48]		[0.16]	[0.31]	[0.12]	
Childbirth services	0.039	0.278	-0.023		-0.119	-0.242	-0.272	
	[0.54]	[2.34]	[0.19]		[1.821]	[2.09]	[2.75]	
Growth monitoring pgm	-0.110	-0.164	-0.000		-0.009	0.023	0.090	
	[1.67]	[1.73]	[0.00]		[0.14]	[0.25]	[0.98]	
Community prices								
Beef with bones	0.074	-0.967	0.518		-0.132	-0.107	-0.570	
	[0.25]	[2.02]	[1.20]		[0.52]	[0.23]	[1.74]	
Fresh fish	-0.411	0.076	-0.406		-0.495	0.566	-0.981	
	[1.68]	[0.18]	[1.03]		[2.68]	[1.42]	[3.57]	
Rice (imported)	0.108	-0.279	-0.037		-0.841	-1.813	-0.643	
	[0.29]	[0.49]	[0.06]		[2.75]	[3.72]	[1.47]	
Palm oil	0.127	0.054	-0.140		-0.524	-0.252	-1.954	
	[0.99]	[0.35]	[0.38]		[4.42]	[1.71]	[6.09]	
Eggs (chicken)	0.373	1.222	-0.014		-0.281	-0.276	-0.612	
	[1.85]	[2.81]	[0.05]		[2.03]	[0.79]	[3.45]	
Sugar	-0.045	-0.204	-0.516		-0.104	-0.251	0.518	
	[0.76]	[2.83]	[1.33]		[1.74]	[3.86]	[1.62]	
Plantains	0.116	-2.052	0.517		-0.425	-1.305	0.175	
	[0.45]	[4.12]	[1.62]		[1.87]	[2.31]	[0.67]	
Manioc (unprocessed)	0.806	2.018	0.244		-1.005	0.124	-4.659	
	[1.93]	[3.59]	[0.23]		[2.54]	[0.24]	[5.92]	
Wald test statistics for joint significance								
Health facilities: personnel	14.28	14.06	12.51		6.93	9.65	4.81	
	[0.00]	[0.00]	[0.01]		[0.07]	[0.02]	[0.19]	
Drug availability	18.87	10.65	18.60		4.17	9.73	2.28	
	[0.00]	[0.01]	[0.00]		[0.24]	[0.02]	[0.52]	
Services	11.05	10.32	7.12		4.29	4.47	7.82	
	[0.01]	[0.02]	[0.07]		[0.23]	[0.21]	[0.05]	
All health characs	54.59	42.48	42.16		30.43	21.38	22.41	
	[0.00]	[0.00]	[0.00]		[0.00]	[0.03]	[0.02]	
Prices	14.71	36.55	11.15		90.48	71.52	111.70	
	[0.06]	[0.00]	[0.19]		[0.00]	[0.00]	[0.00]	

Table 2 (continued)

Household determinants of child height for age and weight for height

(z score)	Height for age			Weight for height		
(children under 12)	*All*	*Urban*	*Rural*	*All*	*Urban*	*Rural*
ln(per capita expenditure)	0.294	0.613	0.127	0.207	0.075	0.317
	[1.17]	[2.53]	[0.44]	[1.12]	[0.29]	[1.71]
Standardized height						
Senior female	0.122	0.103	0.153	-0.037	-0.047	-0.050
	[4.19]	[2.27]	[4.00]	[1.57]	[1.26]	[1.75]
Senior male	0.133	0.087	0.165	-0.013	0.006	0.003
	[4.82]	[1.94]	[4.53]	[0.60]	[0.16]	[0.11]
Education						
Senior female:	0.217	0.316	0.282	0.013	-0.150	0.354
(1) can write sentence	[1.14]	[1.33]	[0.86]	[0.07]	[0.60]	[1.42]
years of education	-0.032	-0.061	0.081	0.010	0.052	-0.141
	[0.74]	[1.15]	[0.88]	[0.23]	[0.91]	[1.99]
years education^2	-0.001	-0.001	-0.018	0.001	-0.001	0.011
	[0.30]	[0.21]	[1.98]	[0.40]	[0.25]	[1.38]
Senior male:	0.316	0.460	0.330	-0.237	-0.121	-0.139
(1) can write sentence	[1.98]	[2.44]	[1.07]	[1.86]	[0.80]	[0.55]
years of education	-0.108	-0.119	-0.158	0.032	0.034	-0.013
	[3.01]	[2.90]	[1.99]	[1.07]	[0.98]	[0.21]
years education^2	0.007	0.007	0.012	-0.002	-0.002	0.001
	[3.51]	[3.19]	[2.05]	[1.13]	[1.14]	[0.13]
Household composition						
(1) senior female exists	0.131	-0.052	0.755	0.036	0.069	-0.069
	[0.62]	[0.19]	[2.51]	[0.22]	[0.30]	[0.32]
(1) senior male exists	0.280	0.066	0.634	-0.014	-0.024	-0.183
	[2.26]	[0.43]	[2.87]	[0.13]	[0.17]	[1.05]
(1) head of hh Ivorian	0.112	-0.256	0.402	0.206	0.318	-0.045
	[1.27]	[2.08]	[2.96]	[2.79]	[2.93]	[0.41]
Child characteristics						
(1) child of head	-0.078	-0.157	0.007	0.242	0.187	0.209
	[1.22]	[1.69]	[0.08]	[4.90]	[2.34]	[3.32]
(1) male	-0.080	-0.121	-0.032	-0.089	-0.115	-0.087
	[1.56]	[1.64]	[0.44]	[2.03]	[1.61]	[1.57]
Wald test statistics for joint significance						
Education of senior female	2.96	6.17	6.31	3.08	2.48	4.14
(include can write)	[0.40]	[0.10]	[0.10]	[0.38]	[0.48]	[0.25]
(years of education only)	2.86	6.10	5.57	1.22	1.63	4.13
	[0.24]	[0.05]	[0.06]	[0.54]	[0.44]	[0.13]
Education of senior male	12.68	10.48	5.49	4.73	1.41	3.63
(include can write)	[0.01]	[0.01]	[0.14]	[0.19]	[0.70]	[0.30]
(years of education only)	12.39	10.17	4.36	1.30	1.30	0.06
	[0.00]	[0.01]	[0.11]	[0.52]	[0.52]	[0.97]
F(all covariates)	8.22	5.17	5.00	7.32	3.83	7.78
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
LM test statistics						
Heteroskedasticity	117.03	67.76	70.82	69.22	34.64	52.53
	[0.00]	[0.01]	[0.00]	[0.01]	[0.84]	[0.03]
Block effects	34.88	4.42	0.36	327.41	56.75	101.15
	[0.00]	[0.04]	[0.55]	[0.00]	[0.00]	[0.00]
Number of observations	4180	1977	2203	3726	1752	1974

Notes: Regressions include nine age dummies and dummies for location of household in Abidjan and in rural sector. InPCE treated as endogenous; identifying instruments are measures of household assets and non-labour income. Health services and community prices measured at cluster level. Variance covariance matrixes estimated by infinitesimal jackknife accounting for block structure of community data. Wald and LM test statistics are χ^2. Degrees of freedom for Wald tests equal number of covariates included in test; LM test for heteroskedasticity has dof equal number of covariates in regression (=45 in first and fourth columns); block effects LM test has 1 dof. [t statics] below coefficient estimates; [p values] below test statistics.

Source: From Thomson, Lavy and Strauss (1996)s

Health services and facilities. We discuss first the health facility characteristics, which, taken together, explain a significant proportion of the variance in both child anthropometric outcomes in rural areas, in urban areas and in the country as a whole. Health personnel characteristics are significantly associated with child height but not with weight for height. Children tend to be taller in communities with more doctors (in the urban sector and in the country as a whole), but facilities with more nurses tend to be associated with shorter children. This could reflect quality of services in dimensions we cannot measure: those facilities with more doctors may provide better services, perhaps because of better equipment and infrastructure within the facility. We also find that the number of working vehicles is positively associated with child health (at least height in the urban sector).

On the other hand, communities with larger facilities tend to be associated with less healthy children: weight for height is significantly negatively correlated with the number of useable beds as is height for age, especially in the rural sector. This correlation may be due to public policy decisions to locate larger facilities in areas that need them most (see Section 5).

Recall that because of the detail of the health facility survey, our measures of health services probably better reflect what is actually available in the community than what is recorded as supposed to be available. Does this distinction make any difference? In theory, if the services that are actually available are measured with random error by the services that are supposed to be available, then invoking a standard errors-in-variables argument their estimated impacts should be smaller.

By replacing the health personnel and infrastructure characteristics with what is supposed to be available in the height regressions in Table 2, we find that the impact of the number of nurses hardly changes (in the urban sector and whole country). In the urban sector and whole country regressions, however, the estimated effects of doctors on child height are reduced to half their value in Table 2 and also become insignificant; in the rural sector, the impact of doctors switches from negative (and insignificant) to positive (and insignificant). The effect of vehicles is also cut in half (in urban Côte d'Ivoire) and switches sign in the rural sector. The (absolute value of the) impact on child height of the number of beds collapses to about a quarter of its value in Table 2 and turns insignificant in the rural sector; the estimated effect switches sign in the urban sector (and for the whole country).[19]

---

19 Effect of different measures of health services on child height for age:

In Côte d'lvoire, there is a divergence between health services that are supposed to be available and those that are actually available. Failure to take account of this fact when evaluating the returns to public investments is likely to result in severely misleading conclusions.

The availability of drugs in the community has a positive effect on child height (in both the rural and urban sector) and also on weight for height (in the urban sector). The availability of antibiotics (in the urban sector) and of quinine (in the rural sector) is significant individually although it seems likely that these are only indicative of an entire package of drugs being available in the facility. The effect of having all three drugs in stock is both statistically significant (χ^2_3 = 18.9) and quite substantial. For the country as a whole, the z-score of a child's height in a community where all three drugs are available would be about 0.3 higher than, *ceteris paribus*, the height of a child where none of the drugs is in stock: this would reduce the deficit in height of the average Ivorian child, relative to a U.S. child, by over a third. Ensuring facilities have a stock of basic drugs is likely to realize a substantial return in terms of improved child health.

The availability of immunizations, childbirth services and child growth monitoring programmes (which typically incorporate a nutrition education component) are also associated with children. Immunizations are especially important in the rural sector (where just over half the communities have these services), whereas childbirth services are more important in the urban sector.

The availability of immunizations and drugs (particularly antibiotics and possibly quinine) and the number of doctors in the community have a significantly positive impact on the height of only older children (three years or more). The presence of growth

(1)	(2)	[se]	(1)	(2)	[se]	(1)	(2)	[se]			
#doctors			0.34	0.76*	[0.26]	0.31	0.59*	[0.28]	1.98	-1.90	[1.44]
#nurses			-0.12*	-0.12*	[0.05]	-0.14*	-0.15*	[0.05]	-0.16	0.32	[0.38]
#support			-0.02	-0.03	[0.03]	-0.03	-0.02	[0.03]	-0.15	0.41	[0.23]
#beds (*100)			0.02	-0.29	[0.16]	0.11	-0.21	[0.21]	-0.49	-1.98*	[0.65]
#vehs (10)			0.22	0.10	[0.14]	0.23	0.42*	[0.15]	0.45	-0.87	[0.62]

Specification (1) includes those services that are *supposed* to be available in the community; specification (2) are based on services *actually* available (and reported in Table 2). Both specifications include all other covariates in Table 2. Standard errors in both specifications are similar: only the standard errors associated with specification (2) are reported. Asterisks denote significance (at 5% size of test). See Table 2 for more detail.

monitoring programmes appears to be negatively correlated with their height and positively, but not significantly, correlated with the length of infants and babies. Larger facilities are associated with shorter infants and babies (but older children near large facilities and those with more doctors tend to be lighter, given height). Taking all the health facility characteristics together, they affect the height only of older children, but the weight for height of both infants and older children seems to respond to changes in these services.

It is also of interest to determine whether a reduction in social service expenditures is likely to have a bigger effect on children from poorer households. Simply stratifying on household expenditure, which is endogenous, complicates interpretation;[20] instead we stratify on the education of household members (which is, itself, a good predictor of long-run resource availability). The sample is split into two groups: those in which either the head or spouse has at least some education (about 40% of the sample) and those with neither having any schooling. Household per capita expenditure is around CFAF 184,000 and 100,000 for the two groups, respectively. For the sake of brevity, we will refer to the latter as poorer households although we recognize that education not only affects household income but may also be related to the gathering and processing of information related to child-rearing and health care practices (Thomas, Strauss and Henriques, 1991).

In the results, reported in Thomas, Lavy and Strauss (1996), it is apparent that on the one hand, more operational vehicles and the availability of childbirth services have a positive impact on the height only of children in less poor households. On the other hand, more doctors and fewer nurses affect the heights only of children in poorer households. These results are quite difficult to interpret and they may reflect (unobserved) heterogeneity in both the availability and quality of services.

Less ambiguous, however, is the fact that immunizations have a significant impact only on the heights of poorer children and that these effects are significantly larger than they are on less poor children. Similarly, the availability of drugs (in particular antibiotics and quinine) has a significantly positive impact only on children from poorer households. The provision of these very basic services will not only have a positive effect on the health of children in Côte d'Ivoire but will have the biggest effect on those who are worst off. For example, the z-score of the height of a child in a household with neither head nor spouse educated will, on average, be 0.8 higher if immunizations are accessible and there is an adequate supply of drugs than if none of these services is provided; this is a big difference and more than offsets the difference between the average child in Côte d'lvoire and the United States. If distributional issues are of concern in the formulation of public policy, then these results suggest that investing in basic health services will realize a substantial payoff.

---

[20] Including interactions among all covariates and household expenditures results in problems with multicollinearity.

It was argued above that failure to control for household resources complicates interpretation of the impact of community services. If the poorest live in poor neighbourhoods that are ill-served by public facilities, then estimated facility effects may simply be picking up the role of household resources. If public services are targeted towards poor neighbourhoods and resources are not controlled, then facility effects will tend to be downward biased. Whether this is important is an empirical issue and can be tested by excluding household expenditure from the regressions. The evidence is mixed, going in both directions depending on the covariate. For example, in the child height for age regressions in the urban sector, failure to control for household resources would substantially overestimate the impact of the number of doctors (the coefficient increases by 28%). This presumably reflects the fact that doctors are clustered in better-off areas. Childbirth services also tend to be located in these areas (their effect increases by 10% when expenditure is excluded). On the other hand, the impact of the availability of antibiotics appears to be greatest among the poorest as its effect declines by 10% when expenditure is not controlled; similarly the impact of working vehicles falls by 25% in this case. Controlling for household resources does seem to be important in these sorts of studies.

Local market prices. In general, the availability and quality of health infrastructure has a larger effect on height for age, a longer-run measure of health status. A shorter-run measure of child health, weight for height on the other hand, is very responsive to price variation, whereas height, for age is little affected by prices especially in the rural sector. There are also striking differences in the impact of prices across sectors.

Weight for height is negatively affected by the prices of fresh fish, eggs, palm oil and manioc in the rural sector and the whole country, but in the urban sector it is the prices of rice, sugar and plantains that are negatively associated with weight for height. Furthermore, price elasticities are larger in the rural sector (greater than unity for fish, palm oil and manioc) than in the urban sector (where the elasticity for sugar and plantains is close to 0.5). Prices (individually or jointly) have no significant impact on the height for age of rural children; but in the urban sector, prices of plantains and beef have a negative effect on height for age, whereas higher manioc and egg prices are associated with *taller* children.

The weight for height of babies and infants is somewhat protected from price rises since these effects tend to be more negative on older children (except for fish and palm oil, which are the only commodity prices that significantly reduce the mass, given length, of young children). In contrast, all but beef and fish prices have a significantly negative impact on the weight for height of children aged three and older. Price rises also tend to have an (absolutely) larger impact on the weight for height of children in poorer households (those with neither the head nor spouse educated): this is true for palm oil, eggs, manioc and, especially, fish.

These price effects hold constant per capita consumption expenditures. However, nominal incomes of food producers will rise as food prices rise and so price effects may

become positive if PCE is excluded from the regressions. Among rural households, most of whom are food producers, unconditional price effects are bigger but the magnitudes of the differences are not large enough to overturn any of the signs in the tables. In sum, our results suggest that stabilization programmes that induce increases in food prices are likely to have a deleterious impact on child weight for height, with older children and possibly those from poorer households being the most vulnerable.

Household characteristics. The household determinants of child height for age and weight for height using the first two waves of the CILSS have been described elsewhere (see Strauss, 1990; Sahn, 1994); we thus discuss these results only briefly. Household resources (lnPCE) have a positive impact on child health, but this is significant only in the case of height for age among urban children.[21] The heights of the senior male and female in the household also have a positive impact on child height (but not weight) reflecting both genetic and family background influences. Few women in Côte d'Ivoire have much education, and child health is unrelated to the level of education of the senior female in the household. The education of the senior male does affect child height (but not weight for height) and, again, these effects are significant only in the urban sector. If the senior male is literate (can write a simple sentence), then child height is significantly greater; but additional years of education are associated with *lower* height (until eight years of schooling, which accounts for about half the sample). A literate male with five years of schooling (the average) would, nevertheless, tend to have taller children in his household relative to an illiterate head with no education.

About 75% of the sample are children of the head: there is very little evidence suggesting that they receive preferential treatment in resource allocation. On average, their height for age is no different from other children, but sons of the head are significantly taller than other boys in the household. While children of' the head tend to be heavier, given height, than other children, this pattern is reversed among older children; this might be, perhaps, because these older children provide labour services to the household, which works to maintain them in good health (see Ainsworth, 1996, for a discussion of child fostering practices in Côte d'lvoire). The height of the senior male and female in the household has a slightly bigger effect on the height of their own children, but the impact remains significant on all other children. Either those children who are not direct descendants of the head and spouse are quite closely related, or the height of senior household members reflects more than just genotype influences. Among the prices and health facility services, only the availability of drugs (in particular antibiotics) has a significantly

[21] Ignoring endogeneity of expenditure and placing the spotlight on measurement error, market expenditure is a valid instrument for total expenditure. We have experimented with this strategy. Market expenditure is a good predictor of total expenditure (with R^2s around 0.60). The estimated effects of InPCE are remarkably similar to those reported in the tables that use non-labour income and assets as instruments: in the height for age regression, for example, the effects are 0.30, 0.54 and -0.05 in the all, urban and rural regressions, respectively.

different effect on children of the head, relative to others, with the impact being bigger on the height for age of those who are not sons or daughters of the head.

Adult anthropometry

The individual, household and community determinants of (the logarithm of) adult body mass index are reported in the first column of Table 3 for all (measured) adults aged 20 through 60, and stratified on both gender and sector in the last four columns.[22]

The measures of availability and quality of health services in the community used in the child regressions have no impact (individually or jointly) on the body mass index of adults in Côte d'Ivoire; local market prices, on the other hand, do significantly affect adult health outcomes. We report regressions, therefore, excluding health facility characteristics.[23]

In general, higher food prices are associated with lower BMI of adults and these effects are strongest among rural dwellers: for example, doubling beef and fish prices in rural areas is likely to reduce rural adult BMI by 10 to 20% . The impact of increasing the price of eggs and manioc will be even larger (with elasticities around 0.80) and falls mainly on the shoulders of urban men and women. The most dramatic effects, however, are in the price of plantains, which reduce the BMI of urban women and rural men (with elasticities of 1.0 and 0.7, respectively). Stabilization programmes that involve large increases in prices are likely to have a substantial impact on the health of both adults and children in Côte d'Ivoire, and these effects are not uniformly distributed across the population.

We turn next to the role of individual and household characteristics. The impact of household per capita expenditure and an individual's years of education are not, in general, linear: both are permitted to be quadratic. Overall, adult BMI is positively associated with PCE, which, recall, is treated as endogenous. It should not, therefore, reflect reverse causality of health status on current earnings and income, at least to the extent that the instruments are uncorrelated with current productivity. The impact of PCE on adult BMI is different for urban and rural men and women, however. The relationship is positive for all women (although it is convex in urban areas and concave in rural areas) and the elasticity (at the mean) is around 0.09 for rural women and 0.11 for urban dwellers:

---

[22] Of all adults in this age group who were enumerated in the CILSS, 14% are excluded from the sample because either height or weight is missing. There are no systematic differences by gender or household per capita expenditure between those included in our sample and those excluded. Almost all heads (96%) and 90% of their spouses are included in our sample; BMI is missing, however, for about a quarter of their children. Those excluded are thus younger, they are more likely to live in urban areas and are, therefore, slightly better educated.

[23] Some of the health service indicators are clearly child specific (availability of immunizations) and so should not directly affect adult health; we have experimented with other indicators of health infrastructure and found little of interest.

increasing household resources is likely to be associated with substantial improvements in the health of women. For men, higher PCE is associated with higher BMI for all (but 5%) of those in the urban sector and also for those in the bottom eight deciles of PCE in the rural sector. The elasticity for men, however, is much smaller than it is for women (being 0.01 in the rural sector and 0.04 in the urban sector), suggesting that men's health would be less affected by declines in household income.

Over and above household resources, education has little impact on body mass index, except in the case of urban females.[24] For the approximately 25% of women with more than seven years of education, BMI declines with schooling and increases for those with less schooling. When household resources are dropped from the regressions, then education has a positive impact on the BMI of all urban males, all but about 15% of urban women and all rural women, but it still has no significant impact on the body mass index of rural men.

[24] Literacy is excluded from the regressions as it has no independent impact on BMI.

Table 3*:* Determinants of adult (log) body mass index by sector and gender

In(BMI)*100	Urban			Rural	
Community prices					
Beef with bones	-6.016	0.990	-6.109	-5.599	-12.181
	[1.90]	[0.13]	[0.89]	[1.63]	[2.88]
Fresh fish	-5.361	-3.664	3.716	-8.651	-5.817
	[3.08]	[0.58]	[0.78]	[3.77]	[2.21]
Rice (imported)	5.976	27.194	0.853	-3.342	-11.285
	[1.14]	[1.69]	[0.10]	[0.70]	[1.70]
Palm oil	-0.941	-1.251	2.312	-3.482	-5.538
	[0.69]	[0.60]	[1.13]	[1.01	[1.52]
Eggs (chicken)	-6.402	-11.886	-20.238	-4.003	-2.195
	[2.18]	[1.04]	[1.39]	[2.28]	[0.93]
Sugar	-0.428	-0.116	-1.583	0.375	-2.241
	[0.76]	[0.14]	[2.15]	[0.08]	[0.52]
Plantains	-4.317	-4.131	-13.470	-7.434	1.131
	[2.05]	[0.55]	[2.12]	[2.42]	[0.44]
Manioc (unprocessed)	-9.446	-26.746	2.296	-10.627	-12.996
	[1.60]	[1.58]	[0.32]	[1.48]	[1.63]
Household characteristics					
ln (per capita expenditure)	-27.149	-93.479	88.330	132.180	-22.149
	[0.74]	[1.34]	[2.04]	[2.53]	[0.31]
ln (per capita expenditure)^2	1.451	4.089	-3.257	-5.477	1.302
	[0.96]	[1.45]	[1.86]	[2.47]	[0.43]
Education (years)	0.670	0.359	1.352	0.343	-0.083
	[3.05]	[0.97]	[3.44]	[1.02]	[0.10]
Years^2	-0.061	-0.029	-0.094	-0.027	0.035
	[2.97]	[0.98]	[3.15]	[0.79]	[0.33]
(1) male	-2.638				
	[3.08]				
(1) head of household	-0.496	-3.650	5.133	-1.637	-1.721
	[0.48]	[1.57]	[2.25]	[1.09]	[0.55]
(1) spouse of head	-1.674		1.861		-1.477
	[1.73]		[0.99]		[1.09]
(1) Ivorian	-1.108	0.282	-3.831	-0.032	-1.278
	[1.44]	[0.20]	[2.42]	[0.03]	[0.85]
(1) Married	3.552	5.513	2.954	2.668	0.616
	[4.03]	[2.37]	[1.44]	[2.18]	[0.41]
(1) if Abidjan	-0.853	0.934	-3.012		
	[0.64]	[0.40]	[1.32]		
(1) if rural	-3.002				
	[3.27]				
Test statistics:					
F (all covariates)	16.48	4.32	5.92	2.89	3.62
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
χ^2 Wald tests: Prices	34.25	9.37	12.07	38.42	21.55
	[0.00]	[0.31]	[0.15]	[0.00]	[0.01]
lnPCE	26.94	12.36	20.60	9.81	9.93
	[0.00]	[0.00]	[0.00]	[0.01]	[0.01]
Education	9.70	0.99	11.97	1.21	0.53
	[0.01]	[0.61]	[0.00]	[0.55]	[0.77]
χ^2 LM (block effects)	29.62	3.78	4.14	6.16	14.02
	[0.00]	[0.00]	[0.04]	[0.01]	[0.00]
χ^2 LM (heteroskedasticity)	34.44	16.44	26.45	17.07	16.76

	[0.00]	[0.69]	[0.19]	[0.58]	[0.67]
Number of observations	3710	884	952	794	1080

Notes: See Table 2. Dependent variable is log(BMI)* 100 of all measured adults, 20 to 60 years of age. Covariates include six age dummies.

Source: From Thomas, Lavy and Strauss (1996).

Conditional on household resources and human capital, men tend to be significantly lighter, given height as are rural dwellers. Married people, on the other hand, tend to be heavier. Stratifying individuals into those who have no education and those who have been to school, we find no evidence that there are significant differences in the impact of prices and household resources on the BMI of these two groups.

Conclusions

Examination of the determinants of individual anthropometric outcomes using the CILSS suggests an important role for public policy in affecting the health of Ivorians, especially children. Our results also indicate some of the dilemmas governments are likely to face as they design stabilization and price reform policies.

Increasing food prices to be in line with world prices may lead to more efficient resource allocation within the economy. Higher food prices should also be associated with higher incomes, at least among those rural households who are net food producers, although these
benefits may only be realized in the longer run. In the shorter run, then, higher food prices are likely to have a detrimental impact on the health of Ivorians as measured by weight for height among children (especially those aged three and older) and body mass index among adults. Higher income, on the other hand, is associated with better health of both adults and children, although the magnitude of the income effects are small except for adult women. Thus, very large increases in income will be needed to offset the negative effects of higher food prices at least in the case of child health.

If reductions in social spending result in lower availability of health care services, then child health (particularly height for age) is likely to be adversely affected. The provision of basic services (such as immunizations) and ensuring health facilities are equipped with simple materials (such as having basic drugs in stock) will probably yield high social returns in terms of improved child health.

Our analysis has relied heavily on the fact that we have had access to very detailed information on the characteristics of the local community, such as the level of health services *actually available*. Without this detail, some of our inferences would have been severely misleading. It would appear that amassing this kind of detailed information will substantially aid the design of good public policy.

5. Endogenous programme placement and selective migration

The estimates of programme availability, quality and pricing impacts just discussed assume that no correlations exist between programme availability or quality variables and unobserved components in the outcome being analysed. Two sources of such bias have been discussed in the literature: purposive programme placement and selective migration.

Some programmes may be placed using criteria that are related to the outcomes being studied (Rosenzweig and Wolpin, 1986). For instance, health clinics may be placed first in less healthy areas. In the absence of perfect measurement of the health environment, clinics will appear to be less effective than they are in reality. Indeed, observing a negative correlation between public health investments (numbers of nurses, say) and health outcomes might be construed as indicating that nurses make people less healthy, but the correlation may, in fact, reflect effective targeting of public investments. The implications of these two interpretations for public health policy are completely different and so it becomes important to understand the mechanisms underlying programme placement. This is area that has been underexplored in the socioeconomic literature. Along the same lines, secondary schools may be located in areas with high demand for schooling; without taking account of the reasons for the location of those schools, evaluations will overstate the impact of secondary schools on school enrolments or completion.

Selective migration will contaminate estimates of programme effectiveness if different levels of infrastructure or subsidies across regions attract people with selectively different tastes for the particular service or endowment level (Rosenzweig and Wolpin, 1988; Schultz, 1988b). For instance, if residential location is related to quality of local schools for families with unobserved higher tastes for education, or an ability to better complement learning in schools, then school effectiveness will be over-rated.

In order to fix ideas we consider a simple example of a programme that has a simple additive impact, identical for all households, with no time lags.[25]

Let ** denote an outcome with an intervention and * an outcome without an intervention, so that if Y is the outcome variable, Y** is the outcome after the programme. Let

$$Y^{**} = Y^{*} + \alpha d \qquad (9)$$

where α = treatment effect and d is a dummy variable equal to 1 if the household receives treatment. Y* may depend on covariates X and unobservables; for instance, we could have:

$$Y^{*} = X\beta + \varepsilon \qquad (10)$$

---

[25] This discussion borrows heavily from that in Moffitt (1991).

We would like to estimate: $\hat{\alpha} = E(Y^{**} | d = 1) - E(Y^{*} | d = 1)$, however, we don't observe E(Y*|d=l). That is we don't observe the counterfactual experiment of what would have resulted *without* the programme for the households who received programme treatment.

One estimate, possible to use with cross section data, is to take the differences between those who received treatment and those who didn't:

$$\tilde{\alpha} = E(Y^{**} | d = 1) - E(Y^{*} | d = 0) \tag{11}$$

The two estimates, $\hat{\alpha}$ and $\tilde{\alpha}$, are equal if: *E(Y*|d=*l) = E(Y*|d=0).* That is, conditional on receiving the treatment or not, the outcomes with no programme treatment should not have been systematically different between the two groups. This condition would be met if one conducted a randomized control trial in which assignment to the treatment or control groups was random, a major advantage of social experiments to which we will return later.

However, in many cases, especially in nonexperimental settings, this condition will not be met. For example, suppose data on some key variable that is related both to the outcome and to being in the treatment group goes unmeasured. Introduce time now and let

$$Y_t^{*} = X_t + \mu + \varepsilon_t \tag{12}$$

where μ is a time-invariant variable unobserved to us but known by the household. Allowing for other covariates, taking differences between treatment and controls amounts to running the following pooled regression:

$$Y_t = X_t\beta + \alpha d_t + \mu + \varepsilon_t \tag{13}$$

with α being the measure of the mean difference in outcomes between treatments and controls.

However, suppose that we can also derive a reduced form model of whether the household is in the treatment or control group as:

$$d_t = Z_t\gamma + \theta\mu + \zeta_t \tag{14}$$

where Z is a vector of observed variables that likely will include most or all of X. In this case the time-invariant unobservable, μ, affects both the outcome and being in the programme. Since μ is correlated with d, the OLS (or simple difference) estimates of α will be biased.

This is a selection problem. The example given above, that nurses may be targeted to less healthy areas, conforms to this context; d would indicate being exposed to nurses in the

local area and μ to the underlying healthiness of the area. In this case the unobservable is at the region or community level. This will often be the case when we are using availability measures of community infrastructure. However, if we are using a measure of whether the individual household uses the programme, then the μ is more likely to correspond to individual household characteristics.

There are a number of potential solutions to this problem. (See Heckman and Robb, 1985, for a classic discussion; Moffitt, 1991, is a much more readable presentation without nearly the detail, however.)

Solution 1: Control in the analysis for unobservables, μ, that cause Y* and d. Since μ is an omitted variable(s), measuring and controlling for it is a good way to proceed if possible. This of course requires that we have better and more complete data. Again, in the setting of using community availability or quality variables, it is missing information on other community infrastructure that is likely to be causing the underlying problem.

Solution 2: Use instrumental variables procedures. This requires having data on identifying variables, Z, that affect d, but not Y*. In general such variables will be difficult to find.[26] A counter-example, provided by Pitt and Khandker's (1996) analysis of household impacts of the Grameen Bank, uses the existence of a programme eligibility criterion that either is not related to the programme outcome or else is related to eligibility in a very nonlinear way, different from how it affects the outcome variable.

To be eligible for membership in Grameen, landholdings have to be less than 0.5 hectare (at least theoretically). A dummy variable of owning under 0.5 hectares would then be a good explanatory variable for whether the household is a Grameen member. If the impact of land ownership on outcomes, say child nutritional status, is reasonably linear, then this dummy variable can be used to predict membership, while a linear land owned variable is used to predict child nutritional status. Pitt and Khandker did even better; they had data from villages that weren't covered by Grameen as well as those that were. Thus they could use an interaction between dummies for the village being covered by the programme and whether the household met the eligibility criterion. This effectively allowed comparisons between landless and smallholders in villages where Grameen does and doesn't operate. They also included village level dummies to capture any unobserved village variation that might be responsible for villages being chosen purposively by Grameen Bank to work in.

Solution 3: Use selection correction procedures such as developed by Heckman (1979). This is similar to solution 2, except that instead of using instrumental variables to predict d, the actual d is used in an OLS regression together with a Mill's ratio term related to the

[26] While empirical modelling of the processes of government investment is possible in principle, see for instance Besley and Case (1995), the variables measured and used tend to be extremely crude, making it unclear how helpful such an approach would be in practice.

probability that d=l (at least if a two-step estimator is used). In this case it is preferred, though not necessary, to have some identifying variables that help predict the probability of participating in the programme, but not the outcome variable. The identification issues are identical to those just discussed for solution 2.

Solution 4: Instead of comparing levels, compare changes between treatments and controls. This corresponds to the classical quasi-experimental setup and in recent literature has been named the difference-in-difference estimator. It requires longitudinal data, ideally including a baseline survey taken before the programme begins and one taken after the programme has ended. The estimator is thus:

$$\tilde{\alpha} = E(Y_t^{**} - Y^*_{t-1} \,|d= 1) - E(Y^*_t - Y^*_{t-1} \,|d = 0) \quad (15)$$

The estimator we would ideally like is:

$$\hat{\alpha} = E(Y_t^{**} - Y^*_{t-1} \,|d= 1) - E(Y^*_t - Y^*_{t-1} \,|d = 1) \quad (16)$$

which would come out of a randomized controlled experiment. The two will be the same so long as:

$$E(Y_t^{*} - Y^*_{t-1} \,|d= 0) - E(Y^*_t - Y^*_{t-1} \,|d = 1) \quad (17)$$

This assumption allows the initial levels of Y* to differ between treatments and controls, for instance by μ, which is differenced away in this estimator. The estimator does assume, however, that any time trends are identical for the two groups.[27]

The regression that corresponds to this method is the first-differenced regression (fixed effects would do just as well and may be preferred on efficiency grounds):

$$Y_t - Y_{t-1} = (X_t - X_{t-1})ß + \alpha d_t + (e_t - e_{t-1}) \quad (18)$$

That is, one compares *changes* in outcomes across areas with and without a programme (policy), d=l versus d=0. We can see that differencing (or taking fixed effects) eliminates the linear unobservable, μ. In an ideal situation, the baseline survey has been fielded before the programme begins so that d=0 at time t-1, so that only d at time t (whether the observation is a treatment or control) appears in the regression. Assuming that d is uncorrelated with ε, using OLS to estimate this regression gives unbiased parameter estimates.

This class of problems has been faced in much of the programme evaluation literature (including training, welfare reform, technological innovation and health interventions).

---

[27] If not we would want to take another difference:

$$\alpha = E[(Y_t^{**} - Y^*_{t-1}) - E(Y^*_{t-1} - Y^*_{t-2})|d = 1] - E(Y_t^{**} - Y^*_{t-1}) - E(Y^*_{t-1} - Y^*_{t-2})|d = 0)$$

This is the so-called difference-in-difference-in-difference estimator.

The most commonly used solution is number 4: use of statistical procedures that mimic the quasi-experimental design of pre- and post-programme comparisons for experimental and control groups. A few studies have used individual-level longitudinal data with individual fixed effects (Rosenzweig and Wolpin, 1986), but many by aggregate individual-level observations into time-series cross-section data and use regional fixed effects (Schultz, 1973; Montgomery and Casterline, 1993; Pitt, Rosenzweig and Gibbons, 1993; Gertler and Molyneaux, 1994; Frankenberg, 1995).[28]

The assumption that all unobserved heterogeneity is swept out by these methods may not be innocuous. There may be interactions between programme effectiveness and unobserved community characteristics, such as between a health programme and the underlying healthiness of an area. Alternatively, if new programmes are located in areas where there is an increase in demand for the service (that is if the source of unobserved heterogeneity is not fixed), then fixed effects estimates will obviously be biased. Similarly, they will be biased if there are changes in other community-level factors that are, in turn, correlated with programme changes.[29] Intuitively, examining the impact of changes in programmes on outcomes naturally raises questions about what determines those changes and one may suppose that they are even *more* endogenous than levels of services.

In general, survey data do not follow a strict control-experimental group pattern. Instead, spatial variation in the intensity of a programme (Schultz, Pitt, Rosenzweig and Gibbons, Frankenberg, Gertler and Molyneaux) or the length of exposure to it (Rosenzweig and Wolpin) is used to identify the programme effect. An exception is provided by Foster and Roy (1997), who analyse data from the quasi-experimental family planning programme run by the International Center for Diarrheal Disease Research—Bangladesh (ICDDR,B) in the Matlab region. ICDDR,B began a family planning programme in 1978 in one part of Matlab, whereas in a contiguous area no such programme was initiated.[30]

---

[28] Frankenberg uses a difference-in-difference type estimator as she randomly pairs two children per local area, born in different periods and thus with different programme exposure, and examines subsequent survival.

[29] Pitt et al. (1993) allow for a constant term in their fixed effects estimates, which implies a time trend in the level estimates. They find that inclusion of the constant reduces the significance of programme variables. Similarly, Montgomery and Casterline (1993) allow for time effects. Clearly, if there is a strong trend component to the programme, controlling for time will result in an underestimate of programme effectiveness, just as omitting a time variable will result in an overestimate if other factors are trending together. A somewhat different approach would be to examine whether two or more series of programme changes are cointegrated and with how many underlying factors. However, this would require a longer time series than is typically available in survey data.

[30]

While one might expect programme leakages due to diffusion of information, apparently this did not occur on a large-scale. As documented by Foster and Roy (1997), contraceptive rates rose and age-specific fertility rates fell in the experimental area as compared with the control area with both starting from similar levels. What Foster and Roy show is that the family planning programme had a major impact on child schooling as well. In the year prior to the programme's initiation, schooling completion rates were similar for both areas, whereas 12 years after programme initiation, schooling levels for both boys and girls were significantly higher in the treatment area. Moreover, the gender differential in schooling attainment had narrowed considerably in the treatment area, but not in the control area. This estimated programme effect survives controls for parental education, owned land and distance to local schools. It is possible that unmeasured factors differ between the two areas and may contribute to increasing differences in fertility and schooling levels. To check this possibility, Foster and Roy also use mother-level fixed effects and, if anything, find the results to be strengthened.

In most of the studies to date, estimates of programme impacts are quite different when individual-level fixed effects are used. However, the pattern is not uniform, which is to be expected since any bias should depend on how programmes are distributed, and this is apt to differ across programmes and between countries. For instance, Rosenzweig and Wolpin (1986) examine changes in standardized child heights and weights over a four-year period using data from Laguna Province in the Philippines. They calculate measures of the time children were exposed to health and family planning programmes, using changes in the fraction of life children were exposed as regressors along with child-level fixed effects.[31] The fixed effects results show that exposure to family planning increases growth in both standardized height and weight, in contrast to results based on cross-sectional data, which show insignificant effects.[32] Evidently health programmes were

---

A maternal and child health programme, including oral rehydration therapy and vaccinations, was later gradually initiated in the experimental area.

[31] Rosenzweig and Wolpin aggregate all ages under 18. This combines adolescents with infants, who have distinct growth trajectories. Standardizing on levels, rather than growth rates, is unlikely to correct for this adequately. Indeed, there exist distinct age profiles in standardized heights and weights (see Martorell and Habicht, 1986; or Thomas et al. 1990). Thus, the estimated programme exposure effects are also likely to be picking up age effects. Moreover, it is also likely that it matters *when* during childhood the programme is experienced, just as the effect of maternal education differs by age of the child (Barrera, 1990; Thomas et al. 1990). Most of the children who have differences in programme exposure between the two time periods, measured as a fraction of lifetime, are likely to be older, with a spell of non-exposure early in life.

[32] It is arguable that even fixed effects estimates are inappropriate in this case because of the stock nature of height and weight. A dynamic production function might relate growth (or changes) to levels of current and past inputs and to a child-level unobservable. In that case, a reduced form in differences will not sweep out child heterogeneity.

placed in areas with less healthy children.[33] Similarly, Montgomery and Casterline (1993) find the estimated impact of family planning on fertility in Taiwan to increase threefold when fixed effects are used; Schultz (1973), using a different and shorter time series, also finds the effectiveness of family planning programmes on birth outcomes in Taiwan increases with fixed effects, as do Duraisamy and Malathy (1991) for fertility in India.

On the other hand Pitt et al. (1993) demonstrate that negative and significant impacts of family planning clinics on school attendance by young children in Indonesia disappear when a fixed effects procedure is used.[34] However, the impact of availability of local schools remains significantly, positively related to schooling. In child fertility equations, the impacts of schools and family planning clinics also disappear with fixed effects. Gertler and Molyneaux (1994), using somewhat different Indonesian data, also find that school and health clinic availability become insignificant in birth probability models once community-level fixed effects are controlled. They find that measures of family planning intensity do not affect birth probabilities on balance, but do increase the odds of using contraceptives and also lower the age at which women marry; these two effects cancel each other. Frankenberg (1995) examines child mortality in Indonesia. She finds that a large effect of having a private toilet, estimated with cross-sectional data, turns insignificant when fixed effects are used, but the availability of maternity clinics and doctors becomes significantly and negatively correlated with child mortality.

As noted by Foster and Roy, the fixed effects estimates may suffer from inattention to lag structure, a problem driven by data availability. For instance, Pitt et al. (1993) relate changes in schooling, fertility and mortality from 1980 to 1985 to programme changes from 1980 to 1986. Yet if family planning programmes provide information, not just subsidies, and this information diffuses gradually by area, longer lags might be expected. The estimates of Frankenberg (1995) and Gertler and Molyneaux (1994) may be similarly affected. Schultz accounts for this by cumulating person-months of programme time over a period of time and relating changes in these stock measures of family planning inputs to changes in children born over a five-year period. This amounts to relating the fertility changes to a cumulative measure of inputs over the same five-year period. Montgomery and Casterline (1993) use a dynamic model with lagged fertility levels in the same and nearby districts to examine the possibility of diffusion.[35] Rosenzweig and Wolpin's (1986) results are less susceptible to this difficulty because they are using data on time exposure, although they do not allow effects to differ depending on when the programme

---

[33] This is corroborated by a regression of local health facilities on locally averaged predicted child heights.

[34] The data they use are aggregated to a regional level.

[35] This requires good data on other time- and district-varying covariates. Montgomery and Casterline use information on child mortality rates, local population density, percent of population employed in agriculture and a schooling variable.

was experienced. Foster and Roy, however, are able to allow the impact of the family planning programme to differ for different lags of exposure.

Less work has focused on biases from selective migration. Schultz (1988b) shows that urban female migrants in Colombia have fewer children than women of equivalent age and education levels in origin areas, but have more children than like women in the urban destination area. If this behaviour is related to possible underlying heterogeneity of the women and not to programme effects, then migration selection will bias estimates of programme effects. One way to circumvent this problem is to compare programme effectiveness among those who are not migrants. Rosenzweig and Wolpin (1988) pursue this line of reasoning in analysing weight gains of children in a city in Colombia who were exposed to a programme of nurse volunteers. Rosenzweig and Wolpin use child-level fixed effects to purge unobserved heterogeneity that might cause bias in such a comparison.[36] They compare their estimates from the sample of non-migrants to estimates from an inclusive sample and find that programme impacts are overestimated for children in high income households, but underestimated for children of low income households.[37]

The general problem in all these models may be cast as one of omitted variables: some studies have shown that the inclusion of additional characteristics can mitigate apparent selectivity bias. For example, Behrman and Birdsall (1983) find the effect of schooling on wages in Brazil is different for migrants relative to non-migrants, indicating there may be migration selection. However, when the authors take school quality into account, this difference disappears. Few studies have attempted to include controls for the process underlying programme placement although this seems like a natural next step.

On balance it appears that evidence is emerging that systematic programme placement can severely bias programme evaluations. The evidence on selective migration is more limited. It may be a less severe source of bias than purposive programme placement to the extent that the prime motivating forces behind migration are relative wages, spreading income risk or marriage and not programmes (see, for instance, Stark, 1991) and to the extent that sources of underlying individual heterogeneity can be measured.

6. Social experiments

The fix-ups for biases caused by endogenous programme placement or participation discussed above are mostly econometric. As discussed, even the most advanced of these depend on assumptions that may not always hold. That potential weakness has led some social scientists to explore use of social experiments as an alternative. Social experiments involve randomization between treatment and control groups, either over individuals or households, or sometimes communities. The random assignment can avoid problems of

[36] The issues of dynamics and age standardization raised above also apply in this case.

[37] Interactions are included between income and programme exposure.

endogenous programme placement or participation as then E(Y'|d=l) = E(Y*|d=O), so that if no trends exist, any differences after the programme must be caused by the programme. If trends do exist they should be the same for randomly assigned treatments and controls, so that one can simply compare changes in their outcomes; that is take the difference in changes, or difference in differences. It is the beauty of being able to make such simple comparisons that makes experiments so enticing.

Nor are experiments as impractical as many times thought. For example, a set of programme investments may be planned, but because of budget constraints they cannot be implemented all at once. In such a case, it may well be practical to begin with randomly selected regions instead of purposively chosen ones. Then, there would be randomization over communities, although not necessarily over households within. The Indonesian Resource Mobilization Study, to be discussed below, is one such example.

Social experiments have been used in the United States for such purposes as to evaluate workforce training programmes as well as numerous other social programmes; Grossman (1994) provides a good recent summary. A few, but increasing number of experiments have been initiated in developing countries as well (see Newman et al., 1994). We will discuss three examples below.

While they do have their advantages, experiments are not without pitfalls. Experiments can be expensive to run on more than a small sample and for more than a limited time. Having only a limited time means that the results are likely to inform only on short-run impacts and maybe only very short term. For instance, it may take time for an improvement in health facility quality to have an impact on health outcomes, while the impact on utilization may occur sooner.

Small samples make it harder to look for differential effects by socioeconomic groups, which may be of utmost importance, particularly from the point of view of reducing poverty, as the Thomas et al. (1996) results demonstrate.

The longitudinal nature of experiments raises the potential of sample attrition. This may have several undesired consequences. If attrition is random, sample size will be reduced with attendant problems. However, if attrition is non-random and related to the outcome being investigated, it can confound the results just like selective migration. In this case one has to resort to econometric fixes, such as using Heckman selection correction procedures (see Heckman and Hotz, 1989, for an example).

As a practical matter, attrition is often a major issue. Furthermore, truly randomizing assignment can be difficult in some situations. There may be ethical or political considerations that make randomization difficult. Sometimes it may not be possible to randomize individuals, but it may be necessary to settle for randomization over localities. Examples of this abound, for instance in randomizing price increases or facility quality improvements.

Randomizing only at the community level makes it more difficult to interpret results, because it is possible that the communities may accidentally differ in certain characteristics or trends that can be confused with the treatment. When only a small number of communities are part of the study it is more likely that such small sample problems might arise. In the extreme case of only two communities, it will be very hard to distinguish programme impacts from other differential changes.

Finally, potential spillover effects across regions may make it more difficult to find a programme impact. For instance, controls may learn from treatments, or if randomization is by area and areas are too close, diffusion of ideas (or migration of people) may occur. As an example, in the Matlab experiment described above, it may be that contraception knowledge may have diffused into the control region, leading to an underestimate of the treatment impact.

Example: Investing in school quality in Kenya

Kremer et al. (1997) describe an experiment involving investments in school quality (supplying textbooks) and providing a subsidy (paying for uniforms) in rural primary schools in Kenya. Randomization was at school level, with 7 out of 14 schools run by the same NGO randomly chosen for "treatment".

Kremer et al. find that dropout rates declined in treatment relative to control schools and that attendance rates increased. The programme caused a large inflow of students to take advantage of the uniforms subsidy (US$10), which for many households is large relative to their budget.

This selective in-migration could cause major problems in analysis if migrants were used as part of the sample; however, in the main analysis Kremer et al. use only those students who were in the treatment and control schools in the baseline survey (i.e., before the programme began). On the positive side, the in-migration of students allowed another quasi-experiment: a rise in student–teacher ratios in the treatment schools.

Moreover, relative test scores, taken both at baseline and after, show no differences between treatment and control schools. Hence while the provision of textbooks may not have aided achievement, the increase in class sizes did not worsen test results. Hence, Kremer et al. argue, it is possible in these schools to increase student–teacher ratio and use the savings to help pay for school uniforms (and perhaps make some quality adjustments), while obtaining an increase in attendance and reduced dropout rate, without a decline in learning as measured by test scores. This is an important finding and one that would have been very difficult to discover from nonexperimental, cross-section data.

There are, however, some potential confounding factors. Of particular concern is the fact that some 120% of students in control schools dropped out, compared with only 4% in treatment schools. Kremer et al. use as their sample for test score analysis, students who began and remained in a particular school. For control schools this may be a selected sample if those who dropped out had characteristics related to low test scores, which

seems likely. If weaker students are more likely to remain in the treatment schools and if they are less likely to learn, then test score improvements would be understated in the comparison of treatment with control schools. In such a case it might be that the changes in treatment schools did help to raise test scores. Whether this occurred is not clear from the studies to date.

Example: Evaluating impacts of raising user fees at health clinics in Cameroon

In their highly cited study conducted during the late 1980s, Litvack and Bodart (1993) concluded that introducing fees at health centres in Cameroon for clinic use and for drug prescriptions actually resulted in an increase in utilization, unlike what doomsday preachers were arguing. The fees were reinvested into drug supplies, assuring a more stable supply than before. It is certainly possible for such a policy to have a positive impact on utilization and the study has been influential in the design of user fee policy.

The study had a quasi-experimental design to it, but upon closer inspection, it was really not implemented as an experiment. Still, it is useful to review in order to explore some of the pitfalls that can occur in designing social experiments.

Some communities were to receive the "treatment" earlier than other, providing the scope for random assignment of who received early treatment. However, communities were not allocated to groups randomly, rather they were assigned by the Health Ministry based upon its list of first or second wave communities. It is not clear from the original article on what basis these decisions might have been made. This raises the potential for purposive programme placement confounding results, the very reason we might have been interested in a social experiment in the first place. Fortunately it turned out that in the baseline, treatments and controls appeared to be reasonably similar with respect to measured characteristics.

Not only were treatment and control communities not randomly allocated, but the investigators did not follow up on the same individuals and households from the baseline to the post-intervention survey. Rather, investigators drew independent samples of households for the baseline and post-intervention. This meant that a simple difference-in-difference (or difference in changes) estimate could not be used to measure the programme impact. Since the villages were the same, however, village averages could be computed for treatments and controls separately from the baseline and post-intervention surveys and differences examined at the village level. The trouble with this strategy is that since only a few villages were included in the design, there is a problem of a rather small sample.

Furthermore, the households were selected for inclusion in the sample based on whether a person was reported as ill during prior month. In Section 3 we discussed issues of measurement of health and potential biases brought in through systematic mismeasurement. Here the issue is that persons in households that use modern health care facilities are more likely to report an illness. Consistent with this hypothesis, no

differences were found in illness rates across income levels, a possible combination of a negative true relationship and positive reporting error.

Indeed, the systematic reporting error can explain the result of apparent increases in health care utilization in treatment areas that experienced the price increases. Suppose that in fact, relative utilization by treatments did fall among those who were in the baseline survey. Post-intervention, those individuals would be less likely on average to report being ill and therefore would be less likely to be included in the post-intervention sample. Those persons in the treatment areas who reported themselves as being ill and who are in the post-intervention survey will have become a more select group and are *more* likely to have sought treatment than those in the baseline sample. Hence the degree of selection is likely to have changed differentially between treatment and control areas in response to relative price changes. It is thus not so surprising that observed utilization would increase in treatment areas, which is exactly what is reported!

Example: Indonesian Resource Mobilization Study

This is similar in some respects to the Cameroon study, in that some areas received the "treatment" of increased user fees for clinics with some quality improvements, while other areas had the onset of fee increases delayed. However, areas were randomized over treatment and control groups and the same individuals and households were interviewed in both the baseline and post treatment surveys!

Even then, it turned out that what treatment meant was different in differing areas, because the private health sector responded by increasing their fees, but differently in different areas. Nevertheless it is possible to examine differences in differences, (changes) between treatment and control households to get the impact of the *joint* treatments. In doing so, Gertler and Molyneaux (1996) find that utilization did go down relatively in areas that were subject to price increases. Dow et al. (1997) find that people do report greater health problems for health measures that are more objective, such as whether one can walk 1 kilometre easily.

7. Conclusions

We have discussed in this paper a basic household production model of investments in human capital, using health investments as our prime example. We have explored problems and potential solutions when empirically estimating such models, emphasizing issues related to measurement error and endogeneity. This literature is still growing, and comparatively few health and schooling investment functions have been estimated for sub-Saharan African countries. Given the high and possibly growing amount of poverty on this sub-continent, and given strong donor pressure related to social sector pricing policies, it would seem imperative for such studies to be carried out.

References

Ainsworth, M. 1996. "The demand for children in Côte d'Ivoire: Economic aspects of fertility and child fostering". In T.P. Schultz, ed., *Research in Population Economics, Volume 8*. Greenwich, Connecticut: JAI Press.

Ainsworth, M. and J. Munoz. 1986. "The Côte d'lvoire living standards survey: Design and implementation". Living Standards Measurement Study Working Paper 53. Washington, D.C.: The World Bank.

Aldeman, H. 1990. "Nutritional status in Ghana and its determinants". Social Dimensions of Adjustment in sub-Saharan Africa Working Paper No. 3. Washington, D.C.: The World Bank.

Altonji, J. 1995. "The effects of high school curriculum on education and labour market outcomes". *Journal of Human Resources*, 30(3): 409─38.

Barrera, A. 1990. "The role of maternal schooling and its interaction with public health programmes in child health production". *Journal of Development Economics*, 32: 69─91.

Behrman, J. and N. Birdsall. 1983. "The quality of schooling: Quantity alone is misleading". *American Economic Review*, 73(5): 928─46.

Behrman, J. and A. Deolalikar. 1988. "Health and Nutrition". In H. Chenery and T.N. Srinivasan, eds., *Handbook of Development Economics*, Vol. 1, Amsterdam: North Holland Press, pp. 631-711.

Besley, T. and A. Case. 1995. "Incumbent behaviour: Vote seeking, tax setting and yardstick competition". *American Economic Review*, 85(l): 25-45.

Bouis, H. and L. Haddad. 1992. "Are estimates of calorie–income elasticities too high? A recalibration of the plausible range". *Journal of Development Economics*, 39(2): 333–364.

Bound, J. 1991. "Self-reported objective measure of health in retirement models". *Journal of Human Resources*, 26(l): 106–38.

Cebu Study Team. 1992. "A child health production function estimated from longitudinal data". *Journal of Development Economics*, 38(2): 323–351.

Dow, W., P. Gertler, R.F. Schoeni, J. Strauss and D. Thomas. 1997. "Health care prices, health and labour outcomes: Experimental evidence, Labour and Population Programme". Working Paper Series 97-01, RAND. Santa Monica, California.

Duraisamy, P. and R. Malathy. 1991 "Impact of public programs on fertility and gender-specific investment in human capital of children in rural India: Cross-sectional and time series analyses". In T.P. Schultz, ed., *Research in Population Economics, Volume 7*. Greenwich, Connecticut: JAI Press, pp. 157–90.

Economist Intelligence Unit. 1986, 1987, 1990, 1991. *Country Report: Côte d'Ivoire*. London.

Falkner, F. and J. Tanner. 1986. *Human Growth: A Comprehensive Treatise*, Vol 3, 2nd Edition. New York: Plenum Press.

Feachem, R., W. Graham and I. Timaeus. 1989. "Identifying health problems and health research priorities in developing countries". *Journal of Tropical Medicine and Hygiene*, 92: 133–91.

Fogel, R. 1994. "Economic growth, population theory and physiology: The bearing of long-term processes on the making of economic policy". *American Economic Review*, 84: 369–395.

Foster, A. and N. Roy. 1997. "The dynamics of education and fertility: Evidence from a family planning experiment". Mimeo. Department of Economics, Brown University, Providence, Rhode Island.

Frankenberg, E. 1995. "The effects of access to health care on infant mortality in Indonesia". *Health Transitions Review*, 5(2): 143–162.

Gertler, P. and J. Molyneaux. 1994. "How economic development and family planning combined to reduce Indonesian fertility". *Demography*, 31(l): 33-64.

Gertler, P. and J. Molyneaux. 1996. "Pricing public health services: Results From a social experiment in Indonesia". Mimeo, RAND, Santa Monica, California.

Gertler, P. and J. van der Gaag. 1990. *Measuring the Willingness to Pay for Medical Care in Developing Countries: Evidence from Two Developing Countries.* Baltimore: Johns Hopkins University Press.

Glewwe, P. 1987. "The distribution of welfare in the Republic of Côte d'lvoire in 1985". Living Standards Measurement Study Working Paper No. 29. Washington,. D.C.: The World Bank.

Griliches, Z. 1977. "Estimating the returns to schooling: Some econometric problems". *Econometrica*, 45(l): 1─22.

Grossman, J.B. 1994. "Evaluating social policies: Principles and U.S. experience". *World Bank Research Observer*, 9(2): 159─180.

Grossman, M. 1972. "On the concept of health capital and the demand for health". *Journal of Political Economy*, 80(2): 223─55.

Habicht J.-P., R. Martorell, C. Yarbrough, R. Malina and R.E. Klein. 1974. "Height and weight standards for pre-school children: How relevant are ethnic differences in growth potential?" *Lancet*, 1: 611─615.

Heckman, J. 1979. "Sample selection bias as specification error". *Econometrica*, 47(l): 153─61.

Heckman, J. and V.J. Hotz. 1989. "Choosing among alternative nonexperimental methods for estimating the impact of social programmes: The case of manpower training". *Journal of the American Statistical Association*, 84: 862─74.

Heckman, J. and R. Robb. 1985. "Alternative methods for estimating the impact of interventions". In J. Heckman and B. Singer, eds., *Longitudinal Analysis of Labour Market Data*. Cambridge: Cambridge University Press, pp. 156─246.

Hill, A. and M. Maindani. 1989. "Operational guidelines for measuring health through household surveys". Mimeo. Centre for Population Studies, School of Hygiene and Tropical Medicine, University of London.

Jamison, D. et al., eds. 1993. *Disease Control Priorities in Developing Countries.* Oxford: Oxford University Press.

Jimenez, E. 1987. *Pricing Policy in the Social Sectors: Cost Recovery for Education and Health in Developing Countries.* Baltimore: Johns Hopkins University Press.

Kanbur, R. and C. Grootaert. 1994. "A new regional price index for Côte d'Ivoire using data from the International Comparisons Project". *Journal of African Economies*, 3: 114–141.

Kremer, M., S. Moulin, D. Myatt and R. Namunyu. 1997. "The quality–quantity tradeoff in education: Evidence from a prospective evaluation in Kenya". Mimeo. Department of Economics, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Krueger, A., M. Schiff and A. Valdes. 1988. "Agricultural incentives in developing countries: Measuring the effect of sectoral and economy-wide policies". *World Bank Economic Review*, 2: 255–272.

Litvack, J. and C. Bodart. 1993. "User fee plus quality equals improved access to health care: Results of a field experiment in Cameroon". *Social Science and Medicine*, 37(3): 369–383.

Lopez, A. 1993. "Causes of death in industrial and developing countries: Estimates for 1985–1990". In D. Jamison et al., eds., pp. 35–50.

Martorell, R. and J. Habicht. 1986. "Growth in early childhood in developing countries". In F. Falkner and J. Tanner, eds., *Human Growth: A Comprehensive Treatise*, Vol 3, 2nd edition. New York: Plenum Press.

Moffitt, R. 1991." Programme evaluation with nonexperimental data". *Evaluation Review*, 15(3): 291-314.

Montgomery, M. and J. Casterline. 1993. "The diffusion of fertility control in Taiwan: Evidence from pooled cross-section, time-series models". *Population Studies*, 47(3): 457–79.

Mundlak, Y. 1963. "Estimation of production and behavioral functions from a combination of time series and cross section data". In C. Christ et al., eds., *Measurement in Economics: Studies in Mathematical Economics and Econometrics.* Stanford: Stanford University Press, pp. 138–66.

Mwabu, G., M. Ainsworth and A. Nyamete. 1993. "Quality of medical care and choice of medical treatment in Kenya". *Journal of Human Resources*, 28(4): 838–862.

National Center for Health Statistics (NCHS). 1976. *Growth Charts*. U.S. Department of Health and Human Services, Public Health Service, Health Resources Administration.

Newhouse, J. and the Insurance Experiment Group. 1993. *Free for All? Lessons From the RAND Health Insurance Experiment.* Cambridge: Harvard University Press.

Newman, J., L. Rawlings and P. Gertler. 1994. "Using randomized control designs in evaluating social sector programs in developing countries". *World Bank Research Observer*, 9(2): 181–202.

Over, M, R. Ellis, J. Huber and O. Solon. 1992. "The consequences of adult ill-health". In R. Feachem et al., eds., *The Health of Adults in the Developing World.* Oxford: Oxford University Press, pp. 161–208.

Parsons, D. 1982. "The male labour force participation decision: Health, reported health and economic incentives". *Economica*, 49(193): 81–91.

Pitt, M. and S. Khandker. 1996. "Household and intrahousehold impact of the Grameen Bank and similar targeted credit programmes in Bangladesh". World Bank Discussion Paper No. 320. Washington, D.C.: The World Bank.

Pitt, M. and M. Rosenzweig. 1985. "Agricultural prices, food consumption and the health and productivity of Indonesian farmers". In I. Singh, L. Squire and J. Strauss, eds., *Agricultural Household Models: Extensions, Applications and Policy.* Baltimore: Johns Hopkins University Press.

Pitt, M.M., M.R. Rosentweig and D.M. Gibbons. 1993. "The determinants and consequences of the placement of government programmes in Indonesia". *World Bank Economic Review*, 7(3):319–348.

Pitt, M., M. Rosenzweig and M. N. Hassan. 1990. "Productivity, health and inequality in the intrahousehold distribution of food in low-Income countries". *American Economic Review*, 80(5): 1139–56.

Pollak, R. 1969. "Conditional demand functions and consumption theory". *Quarterly Journal of Economics,* 83: 70–78.

Psacharopoulos, G. 1994. "Returns to investment in education: A global update". *World Development*, 22(9): 1325–43.

République de Côte d'lvoire. 1979. *1975–1977 En Chiffres*. Abidjan.

République de Côte d'Ivoire. 1981. *1978–1980 En Chiffres*. Abidjan.

République de Côte d'lvoire. 1985. *Memento Chiffre de la Côte d'Ivoire, 1983–4*. Abidjan.

République de Côte d'lvoire/UNICEF. 1988. *Analyse de la Situation des Enfants et des Femmes en Côte d'lvoire*. Abidjan.

République de Côte d'Ivoire. 1990. *Projet de Plan de Developpement Economique, Social et Culturel 1986-90*. Abidjan.

Rosenzweig, M. and T.P. Schultz. 1983. "Estimating a household production function: Heterogeneity, the demand for health inputs and their effects on birthweight". *Journal of Political Economy*, 91(5): 723–46.

Rosenzweig, M. and K. Wolpin. 1986. "Evaluating the effects of optimally distributed public programmes". *American Economic Review*, 76: 470–482.

Rosenzweig, M. and K. Wolpin. 1988. "Migration selectivity and the effects of public programmes". *Journal of Public Economics*, 37(3): 265–89.

Sahn, D. 1994. "The contribution of income to improved nutrition in Côte d'lvoire". *Journal of African Economies*, 3: 29–61.

Schoenbaum, M. 1995. "The health status and labour force behaviour of the elderly in Taiwan". Comparative Study of the Elderly in Asia, Research Report No. 95-32. Population Studies Center, University of Michigan.

Schultz, T.P. 1973. "Birth rate changes over time and space: A study of Taiwan". *Journal of Political Economy*, 31 (Supplement): S238–74.

Schultz, T.P. 1988a. "Education investment and returns". In H. Chenery and T.N. Srinivasan, eds., *Handbook of Development Economics*, Volume 1. Amsterdam: North Holland Press.

Schultz, T.P. 1988b. "Heterogeneous preferences and migration: Self-selection, regional prices and programs and the behaviour of migrants in Colombia". In T.P. Schultz, ed., *Research in Population Economics*, Volume 6. Greenwich, Connecticut: JAI Press, pp. 163–82.

Schultz, T.P. and A. Tansel. 1997. "Wage and labour supply effects of illness in Côte d'lvoire and Ghana". *Journal of Development Economics*, 53(2): 251–86.

Schultz, T.W. 1980. "Nobel Lecture: The economics of being poor". *Journal of Political Economy*, 88(4): 639–51.

Srinivasan, T.N. 1992. Undernutrition: Concepts, measurements and policy implications". In S.R. Osmani, ed., *Nutrition and Poverty*. Oxford: Clarendon Press, pp. 97–120.

Stark, O. 1991. *The Migration of Labour*. Cambridge: Basil Blackwell.

Steckel, R. 1995. "Stature and the standard of living". *Journal of Economic Literature,* 33(4): 1903–1940.

Stewart, A. and J. Ware. 1992. *Measuring Functioning and Well-Being: The Medical Outcomes Study.* Durham: Duke University Press.

Stewart, A. et al. 1978. *Conceptualization and Measurement of Health Status for Adults in the Health Insurance Study: Volume II, ploical health in terms of functioning.* R-1987/2-HEW, RAND, Santa Monica, California.

Strauss, J. 1990. "Households, communities and preschool children's nutrition outcomes". Evidence from rural Côte d'lvoire". *Economic Development and Cultural Change,* 36(2): 231–262.

Strauss, J. and K. Mehra. 1989. "Child anthropometry in Côte d'lvoire. Estimates from two surveys, 1985 and 1986". Living Standards Measurement Study Working Paper 51.

Strauss, J. and D. Thomas. 1995. "Household resources: Empirical modelling of household and family decisions". In J. Behrman and T.N. Srinivasan, eds., *Handbook of Development Economics,* Volume 3. Amsterdam: North Holland Press.

Strauss, J. and D. Thomas. 1996. "Measurement and mismeasurement of social indicators". *American Economic Review Papers and Proceedings,* 86(2): 30–34.

Strauss, J. and D. Thomas. 1998. "Health, nutrition and economic development". *Journal of Economic Literature,* 36(2): 766–817.

Strauss, J., G. Mwabu and K. Beegle. 1998. Intrahousehold allocations: A review of theories, empirical evidence and policy issues". African Economic Research Consortium Plenary Sessions, 1997–8, *Journal of African Economies,* Supplement, forthcoming.

Strauss, J., P. Gertler, O. Rahman and K. Fox. 1993. "Gender and life-cycle differentials in the patterns and determinants of adult health". *Journal of Human Resources,* 28(4): 791–837.

Thomas. D. and J. Strauss. 1992. "Prices, infrastructure, household characteristics and child height". *Journal of Development Economics,* 39: 301–331.

Thomas, D. and J. Strauss. 1997. "Health and wages: Evidence on men and women in urban Brazil". *Journal of Econometrics,* 77(1): 159–186.

Thomas, D., V. Lavy and J. Strauss. 1996. "Public policy and anthropometric outcomes in the Côte d'lvoire". *Journal of Public Economics,* 61: 155-192.

Thomas, D., J. Strauss and M. Henriques. 1990. "Child survival, height for age and household characteristics in Brazil". *Journal of Development Economics,* 33(2): 197–234.

Thomas, D., J. Strauss and M. Henriques. 1991. "How does mother's education affect child height". *Journal of Human Resources,* 26(2): 183–211.

Waaler, H. 1984. Height, weight and mortality: The Norwegian experience". *Acta Medica Scandinavia*, Supplement No. 679, Stockholm, Sweden.

Ware, J. 1989. Standards for validating health measures: Definition and content". *Journal of Chronic Diseases*, 40(6): 473–80.

Ware, J., A. Davies-Avery, and R. Brook. 1980. *Conceptualization and Measurement of health Status for Adults in the Health Insurance Study: Volume VI, Analysis of Relationships among Health Status Measures.* R-1987/6-HEW, RAND, Santa Monica, California.

Ware, J., A. Davies-Avery, and C. Donald. 1978. *General Health Perceptions.* R-1987/5-HEW, RAND, Santa Monica, California.

Waterlow, J., ed. 1988. *Linear Growth Retardation in Less Developed Countries.* Nestle Nutrition Workshop Series Volume 14. New York: Raven Press.

Waterlow, J., R. Buzina, W. Keller, J. Lane, M. Nichaman, J. Tanner. 1977. "The presentation and use of height and weight data for comparing the nutritional status of groups of children under the age of ten years". *Bulletin of the World Health Organization,* 55: 489–98.