Chapter 54

HEALTH OVER THE LIFE COURSE*

JOHN STRAUSS

Department of Economics, University of Southern California, 3620 South Vermont Avenue, 306A Kaprielian Hall, Los Angeles, CA 90089-0253, USA

DUNCAN THOMAS

Department of Economics, Duke University, Box 90097, Durham, NC 27708, USA

Contents

Abstract 3376
Keywords 3377
1. Introduction 3378
2. Conceptual framework and methods 3379
2.1. Static models of health investments and health outcomes 3380
2.1.1. Health production functions 3380
2.1.2. Modeling behavioral choices 3381
2.1.3. Wages and health 3382
2.1.4. Reduced form demand for health investments and outcomes 3383
2.1.5. Estimation of the relationship between health and economic prosperity 3384
2.2. Dynamic models of health investments and health outcomes 3385
2.2.1. Dynamic health production function 3385
2.2.2. Allocating resources over the life course 3386
2.2.3. Economic productivity and health over the life course 3387
2.2.4. Dynamic conditional demand for health 3388
2.2.5. Dynamic reduced form demand for health 3389
2.2.6. Estimation of dynamic demand for health functions 3390
2.2.7. Estimation of dynamic health production functions 3392
2.2.8. Estimation of dynamic relationships between health and economic prosperity 3393
2.3. Health, the individual, the household and the family 3395
2.4. Longitudinal study designs 3396
2.5. Studies designed to identify causal effects 3397

* The authors thank Facundo Cuevas for research assistance, Sheila Evans and Eileen Miech for editorial assistance and Jere Behrman, Amar Hamoudi, Chris Ksoll, Doug McKee, Jenna Nobles, Tristan Reed, and T. Paul Schultz for very helpful comments. This research was supported, in part, by NIH grants AG020905, HD047522, HD049766 and HD050452.

Handbook of Development Economics, Volume 4

DOI: 10.1016/S1573-4471(07)04054-5

2.5.1. Experimental designs 3397
2.5.2. Quasi-experimental designs 3399
2.5.3. Other non-experimental methods 3400
2.5.4. Measurement of potential instruments to identify causal effects 3402
3. Meaning of health 3403
3.1. Mortality and life expectancy 3403
3.2. Global burden of disease 3408
3.3. Physical health assessments: Anthropometry 3411
3.4. Physical health assessments and biomarkers 3416
3.5. Self-assessed health 3417
3.6. Measurement error in health 3418
4. Empirical evidence on health over the life course 3419
4.1. Long-run impacts of early life health on later life outcomes 3420
4.1.1. Quasi-experimental studies on fetal health and health in later life 3421
4.1.2. Quasi-experimental studies on fetal health and economic prosperity in later life 3423
4.1.3. Non-experimental evidence on early life health and later life outcomes 3424
4.1.4. Intergenerational transmission of health 3426
4.2. Child and adolescent health, human capital development and economic success 3428
4.2.1. Experimental studies 3429
4.2.2. Quasi-experimental studies 3433
4.2.3. Non-experimental studies 3437
4.3. Adult health and SES 3440
4.4. Health and exogenous shocks 3443
4.5. Health and consumption smoothing 3447
5. HIV/AIDS 3450
5.1. Methodological issues 3451
5.1.1. Sampling 3451
5.1.2. Measurement 3452
5.1.3. Estimation 3453
5.2. Parental death and child schooling 3453
5.3. Parental death and health of other household members 3457
5.4. Parental death and household resources 3458
6. Conclusions 3463
References 3465

Abstract

In recent years, significant advances have been made in better understanding the complex relationships between health and development. This reflects the combined effects of methodological innovations at both the theoretical and empirical level, the integration of insights from the biological and health sciences into economic analyses as well as improvements in the quantity and quality of data on population health and socio-economic

status. To provide a foundation for discussing these advances, we describe static and dynamic models of the evolution of health over the life course in conjunction with the inter-relationships between health, other human capital outcomes and economic prosperity. Facts about health and development at both the aggregate and individual levels are presented along with a discussion of the importance of measurement. We proceed to review the empirical literature with a goal of highlighting emerging lines of scientific inquiry that are likely to have an important impact on the field. We begin with recent work that relates health events in early life, including in utero, to health, human capital and economic success in later life. We then turn to adult health and its relationship with socio-economic success, exploring the impact of health on economic outcomes and *vice versa* as well as the links between health and consumption smoothing. Recent evidence from the empirical literature on the micro-level impacts of HIV/AIDS on development is summarized. We conclude that developments on the horizon suggest a very exciting future for scientific research in this area.

Keywords

health, economic development

JEL classification: O12, I1, D13

1. Introduction

This chapter provides a review of recent advances in microeconomic research on health over the life course, emphasizing empirical findings in developing country contexts. The literature in this field has made substantial progress on many fronts in recent years. This is a reflection of several factors. First, innovative research designs have been used very creatively to shed new light on important and, at least until recently, unresolved questions. The most influential of these designs have a strong foundation in economic theory and, in many cases, also integrate insights from biology. Second, clever ways have been used to isolate causal pathways in studies that use observational or quasi-experimental methods and, increasingly, random assignment treatment-control designs. Third, and closely related, is the fact that studies have exploited new sources of data that contain a rich and improving array of measures of health and socio-economic success, often with repeated measures for the same individual over time.

The relationship between health and development is, and promises to continue to be, an extremely active line of inquiry not only in economics but also in the population and health sciences. By way of examples, recent and on-going research includes studies that highlight the complex dynamics underlying the joint evolution of different domains of health and socio-economic success over the life course. Other work seeks to isolate some of the causal pathways through which different dimensions of health might affect indicators of socio-economic status (SES) and, in turn, how individual, family and community resources affect health. Studies have highlighted the ways in which early life health, including fetal health, potentially influence health and well-being over the entire life course. Those studies underscore the importance of understanding the evolution of health from conception to the end of life. This, and other work, indicates the key role played by genes, and their interaction with the environment, in influencing health outcomes throughout life. This emerging body of research promises to yield important new insights that will contribute to better understanding the relationships between population health and economic development.

Our primary goal is to assess the current state of scientific knowledge regarding the dynamic inter-linkages between health and economic prosperity at the micro level. We do not focus on the policy implications of this research or on the allocation of resources to health or other sectors, health care financing or the production and distribution of health services. These and other issues are discussed in considerable detail in Jamison et al. (2006). The Copenhagen Consensus brought together leading economists to rank the global challenges in terms of their impact on human well-being. They concluded that three of the top four priorities relate to health – controlling HIV/AIDS, providing micronutrients and controlling malaria. See Lomborg (2004, 2006) for a discussion of the costs and benefits associated with such interventions.[1]

[1] Of particular relevance here are Behrman, Alderman and Hoddinott (2004) who discuss the costs and benefits of nutrition-related interventions and Mills and Shillcutt (2004) who discuss communicable diseases.

The next section discusses some key conceptual and methodological issues in the literature, highlighting the dynamics of health over the life course. It is followed by a brief description of some basic facts on health and development in which we pay special attention to concerns about interpretation of different health indicators. Those sections set the stage for a critical assessment of the evidence. We begin with a review of the literature on the long arm of childhood health reaching into the health and economic well-being of adults. The relationships between childhood health and other dimensions of human capital are discussed. We then turn to adult health and its relationship with socio-economic success, exploring the impact of health on economic outcomes and *vice versa*. This is followed by a discussion of health and consumption smoothing. We end with a brief discussion of HIV/AIDS, concentrating on recent studies that examine the impacts of HIV/AIDS on socio-economic outcomes.

2. Conceptual framework and methods

The empirical literature on health is replete with correlations relating measures of resources at the individual or family level, such as income, wealth or education, to many domains of individual health including mortality, other indicators of health status, health behaviors and inputs into the production of health. How these correlations should be interpreted has played a central role in the literature for at least two decades and while substantial progress has been made in understanding the links between health and economic prosperity, many questions remain unresolved.

There are at least three sets of issues that have been stumbling blocks in this literature. First, health varies over the life course and there is limited understanding of the complex dynamics underlying relationships between health, health behaviors and other factors that influence the well-being of individuals; there is even less understanding of how these relationships evolve over the life course. Second, poor health likely affects economic well-being while limited economic resources presumably affect health investments and outcomes, and there are quite likely to be feedbacks between health and resources. It has been extremely difficult to establish causality, which likely operates in both directions. Interpretation of evidence is further complicated by unobserved heterogeneity that affects both health and material well-being. Third, health is multidimensional and hard to measure. Not only do many health markers fluctuate randomly over time but, potentially more pernicious for interpretation, some widely-used indicators of health and health behaviors are likely to be measured with errors that are systematically related to characteristics of interest in many models such as age, education or income.

This section discusses a series of simple economic models of decisions regarding investment in health and health outcomes that provide a starting place to shed light on these issues. To fix ideas, we begin with static models and then proceed to discuss dynamic models of the evolution of health over the life course. The section ends with a discussion of the implications of the models for research design and empirical methods.

2.1. Static models of health investments and health outcomes

2.1.1. Health production functions

Models of the production of health in a static context, and the implications of those models for interpreting relationships between health and development, have been discussed extensively (see for example, Schultz, 1984, 2005; Strauss, 1986, 1993; Behrman and Deolalikar, 1988; Behrman, 1996; Strauss and Thomas, 1995, 1998; Smith, 1999). We briefly review some key models and important conceptual issues to provide a foundation for discussion of some of the dynamic concerns that have received less attention.

We are concerned with health across the entire life course from conception to death. Parents are assumed to make key health decisions for children whereas an adult is assumed to make his or her own decisions. It is important to distinguish health outcomes, such as height, body mass, disease incidence or physical functioning, from health inputs and health behaviors which might include nutrient intakes, exercise, smoking, and utilization of preventive or curative health care.[2]

Assume there is a static health production function for an individual:

$$H = H(N; A, B_H, D, \mu) \tag{1}$$

where H represents an array of measured health outcomes. They depend on a vector of health inputs and behaviors, N. These include, for example, use of health care services, nutrient intake, energy output and time allocated directly or indirectly to the production of health. Examples of behaviors that might contribute to health outcomes include smoking, risky or very stressful lifestyles. The inputs, N, are under the control of the individual. The technology, or shape of the underlying health production function is likely to vary over the life course and so varies with age and, possibly, with other socio-demographic characteristics, A, such as gender. The technology may also vary with dimensions of family background that affect health, B_H, such as parental health and genetic endowment. Technology will likely vary with environmental factors, D, such as the disease environment, public health infrastructure and treatment practices or standards of care.

There are two classes of unobserved characteristics, μ, in the health production function. First, there are unobserved factors that are assumed to be known to the individual but are not known to the econometrician; such as innate healthiness. Second, μ reflects the fact that health and its determinants are potentially measured with error. In some cases, these errors will be random (the "classical" measurement error case) but in some cases the errors may be systematically related to observed or unobserved factors that affect health.

[2] The distinction is not always clean since some health outputs may also be intermediate inputs into other outputs. For example, disease incidence might affect body mass and *vice versa*. We abstract from these complications. Note also that some health outputs may be produced jointly, while others may be produced solely.

The health production function, which is analogous to an agricultural or manufacturing production function, represents the technological and biological constraints that determine how inputs are converted into outputs. In the agricultural economics literature, Nelson (1964) and Hayami and Ruttan (1971) suggested that the available set of technologies is not likely to be the same over time, across countries or even within countries and will depend on the rate of adoption of new technology, the diffusion of technology and the adaption of technology to local settings. Hayami and Ruttan (1971) called the envelope of feasible technological relationships a meta-production function and argued that a farmer at a point in time is likely to face the opportunities on only part of the meta-production function. The same concept probably applies to the production function of health which depends on the development, diffusion, adaption and availability of new technologies (see Cutler, 2004, for an insightful discussion). Moreover, in the case of health, the part of the meta-production function that an individual faces may be a function of his or her own previous choices or the choices of others living in the community. For instance, a person who is HIV positive may be on a different part of the meta-production function than someone who is HIV negative. This substantially complicates identification and estimation of a health production function as it is no longer a technical relationship but incorporates behavioral choices.

2.1.2. *Modeling behavioral choices*

More generally, behavioral choices play a central role in the conceptual framework. Assume that an individual's welfare depends on labor supply, L, and consumption of purchased goods, C. Utility, U, is assumed to depend on health outputs, H, as well as observed characteristics such as socio-demographic characteristics, A, non-health human capital (including schooling) and family background, B_U:

$$U = U(C, L; H, A, B_U, \xi). \tag{2}$$

The unobserved characteristics, ξ, include heterogeneity in tastes which may be related to unobserved characteristics that affect the production of health (1), preferences, for example, may themselves depend on innate healthiness.

Resource allocations are constrained by budget and time constraints in addition to the health production function (1). Suppose that the individual earns wage, w, for each unit of labor supplied and that asset or nonlabor income is V. The budget constraint is:

$$p_c C^* + p_n N^C = wL + V. \tag{3}$$

For expositional simplicity, the vector of consumption, C, has been decomposed into two elements: consumption that is not related to the production of health, C^*, with prices, p_c, and a vector of purchased health inputs, N^C, with prices, p_n. In practice, the division between these groups is not always clear and not all health inputs may be valued directly in the utility function. Time allocated to the production of health, N^T, leisure and labor supply, L, is equal to the individual's total time endowment. Note that

the vector of inputs in the health production, N, includes not only N^C and N^T but also non-purchased inputs that are not valued directly in the utility function.

Solving the system (1) through (3) yields a function for each health input and output which will depend on prices of consumption and health inputs along with the observed and unobserved characteristics that affect health and utility including demography, human capital and background, nonlabor income and the disease environment. In addition, each function depends on wages, which are the outcome of choices made by the individual and so it is necessary to specify a wage function in order to substitute for wages and specify a reduced form demand function for health inputs and outputs. We turn, therefore, to the determination of wages.

2.1.3. Wages and health

Assume that an individual's real wage is equal to his marginal product. A person's wage, w, varies with health outputs, H, socio-demographic characteristics, A, as well as own and family background characteristics, B_w, such as schooling, other non-health human capital, parents' schooling and health:

$$w = w(H; A, B_w, I, \alpha). \tag{4}$$

As discussed below, health outputs (such as height or body mass) may affect wages through better physical or mental health as well as through strength and endurance. Local community infrastructure, I, such as electrification or road density, may be related to labor demand, or to various work characteristics which are valued separately from wages. Wages will also be influenced by unobserved factors, α. These might include ability, effort and school quality (which appear in A if they are measured) along with tastes like ambition and competitiveness that might affect economic productivity as well as tastes for investment in human capital. Note that some of these unobserved characteristics are shared with μ in the health production function (1) and tastes, ξ, that condition welfare in (2). In addition, variation in wages due to measurement error will be reflected in α.

As is the case in the interpretation of schooling in estimated wage functions (Card, 1999), interpretation of health covariates in the wage function (4) is plagued by potential bias due to unobserved heterogeneity that is correlated with both wages and health. For example, unobserved factors that affect wages, α, may well be correlated with health measures, either directly, or through being correlated with unobserved innate healthiness or tastes for investments in human capital, μ. For instance, nutrition as an infant and young child affects height as an adult; moreover the literature indicates that early childhood nutrition is correlated with cognitive development (Pollitt et al., 1993, among numerous others), which may well affect productivity at work. Many empirical wage functions have reported a positive relationship between wages and height. This might be driven by the fact that wages are affected by cognitive achievement which is not observed but is correlated with height because both cognition and height are driven by early childhood investments.

In models of health and wages, there are at least two additional substantively important sources of unobserved heterogeneity. First, the biological and social science literatures suggest there may be contemporaneous feedbacks between health and economic productivity. This will arise, for example, if increases in income caused by improvements in health are spent, at least in part, in ways that further improve health. Specifically, say improvements in the health of a worker cause his or her wages to be higher and that results in higher earnings. There will be a feedback effect if the worker spends some of the additional income on inputs into the health production function (1), such as improved diet (perhaps more food and hence more calories or higher valued foods, that have more protein, iron, vitamin A, or some other micronutrients), improved sanitation and living conditions or improved health care. Feedbacks will arise if greater consumption of these inputs results in improved health. These issues are taken up in more detail below.

As noted above, a second potentially important source of unobserved heterogeneity in models of wages and health revolves around measurement of health. We defer a fuller discussion to the next section.

2.1.4. *Reduced form demand for health investments and outcomes*

Solving the optimization program (1), (2), (3) and (4) for an adult, letting λ be the marginal utility of income and assuming interior solutions,[3] the first-order condition with respect to consumption of the jth health input, N_j, is:

$$\frac{\partial U}{\partial N_j} + \frac{\partial U}{\partial H}\frac{\partial H}{\partial N_j} = \lambda\left(p_{n_j} - L\left[\frac{\partial w}{\partial H}\frac{\partial H}{\partial N_j}\right] - w\left[\frac{\partial L}{\partial H}\frac{\partial H}{\partial N_j}\right]\right) \tag{5}$$

where the first element is zero if N_j is not valued in and of itself (is not an element of C in (2)).

This first-order condition highlights an important substantive point. If health inputs raise wages or labor supply through improving health outcomes, then there will be a decline in the shadow price of inputs that affect health, inducing greater use of those inputs (assuming the usual regularity conditions). In the model, it is possible that the rate of change of health input shadow prices will depend on the level of health, H. Indeed, biomedical evidence suggests that, in some instances, links between health inputs and outputs are not linear. For example, when levels of iron in the blood fall below a threshold (say hemoglobin levels below 13 g/dl), work capacity is reduced; however, additional iron above this threshold has no impact on work capacity. This has been demonstrated in rigorous clinical trials with humans and animals, and the mechanisms underlying the non-linear relationship are well-understood right down to the cell level (Haas and Brownlie, 2001).

[3] Corner solutions complicate the exposition. Under the assumptions of the model, an individual will work if the wage exceeds the shadow value of time and, in that case, the quantity of time allocated to labor supply will equate the wage to the shadow value of time.

There is a reduced form demand function for each element of the vector of health inputs, N, and each of the health outputs, H, which we collect into the vector θ:

$$\theta = \theta(p_n, p_c, A, B, V, D, I, \varepsilon). \tag{6}$$

Each demand function depends on health input prices, p_n, consumption prices, p_c, demographic characteristics, A, non-health human capital and family background, B (which encompasses those dimensions of background that affect the production of health, B_h, productivity in the labor market, B_w, and preferences, B_U), nonlabor income, V, the disease and health environment, D, and non-health determinants of wages, I. All of these characteristics are assumed to be exogenous to the individual's decisions about behaviors and investments related to health, time allocation and resource allocation. In addition, they are all, by definition, observed. There are also factors which influence demand for each of the health inputs and outputs that are not observed, ε. These encompass not only tastes, ξ, but also innate healthiness, μ, and characteristics such as ability and effort, α, which affect wages.

2.1.5. Estimation of the relationship between health and economic prosperity

Under the assumptions we have made, it is clear that health is potentially affected by earnings and that earnings may depend on health. There is a vast literature documenting a strong positive correlation between health and different markers of socio-economic status (SES) across the globe. Identifying the causal pathways underlying this correlation – and the extent to which it is driven by unobserved heterogeneity in a particular empirical model – is not straightforward. We provide a brief discussion of models that examine the impact of health on wages or economic productivity and then turn to models that explore the effect of SES on health.

Inspection of (4) and (6) illustrates one of the difficulties associated with disentangling the causal effect of health on wages and *vice versa*. Specifically, there are unobserved factors that affect both wages and utility as indicated by α appearing in the wage function, (5), and also being an element of ε in the health demand function, (6). Intuitively, this may be because unobserved dimensions of ability which affect wages also impact utility because, for example, more able people transform the same inputs into a different level of utility than less able people. Common unobserved heterogeneity may also reflect dimensions of tastes, such as inter-temporal preferences, that affect both wages and utility. These are examples of contemporaneous feedbacks discussed above. An analogous argument applies to disentangling the causal effect of health on labor supply.

Comparing the functions (4) and (6) also suggests potential instruments to identify the causal effect of health on economic productivity. Conditional on health outcomes, H, and levels of local infrastructure, I, productivity should not be affected by health input prices, p_n, or the disease environment, E. For example, the accessibility, price and quality of health services available to an individual, along with the epidemiological environment in which he or she is living, are potential instruments to the extent they

affect health but have no direct impact on wages in (4). The same characteristics are also potential instruments for inputs into the health production function (1). Similarly, factors, I, that affect wages and hours of work, such as local labor demand conditions and levels of economic infrastructure, are potential instruments for earnings and other indicators of SES in conditional demand functions that relate health to SES. Note that if the health services available to an individual are the outcome of choices – such as decisions about where to live or occupational choice – then they are not appropriate instruments. The same issue applies to prices of local health inputs, the local disease environment and local labor market conditions.

Many studies examine how health covaries with schooling or some other dimension of background, B, or with nonlabor income, V. These estimates can be interpreted in the context of the static reduced form model (6) although, as we shall see below, moving to dynamic framework complicates interpretation. Moreover, there is substantial interest in estimating the demand for health conditional on an individual's earnings or total income. Earnings is the product of wages, w, and labor supply, L, both of which are choices that are likely correlated with unobserved characteristics that affect health in even this simple model. Thus, wages, labor supply and earnings should be treated as endogenous in these conditional demand for health functions. The same principle applies to demand functions that are conditioned on total household income or total expenditure which are potentially more complicated if household living arrangements are associated with unobserved characteristics that affect individual health. The key point is that consistent estimation of these conditional demand for health functions will take the joint determination of health and SES into account.

2.2. *Dynamic models of health investments and health outcomes*

Health evolves over the life course. For at least some health behaviors and health outcomes, health at one point in the life course affects health in later life, and emerging evidence suggests that even health *in utero* is predictive of health as an adult. Incorporating dynamics into models of health not only builds more realism into the models but also yields important theoretical insights as well as useful implications for the implementation and interpretation of empirical applications. Building on the foundation laid in the previous subsection, we turn now to sketching out some of these implications.

2.2.1. *Dynamic health production function*

It is natural to start with a dynamic health production function in which current health status, H_t, depends on all current and prior health inputs, $N_\sigma, \sigma = 0, \ldots, t$, the disease or public health environment, D_σ (Grossman, 1972), current and prior demographic characteristics such as age, A_σ, which are time-varying and other demographic and background characteristics, B_H, which are time invariant.

$$H_t = H(N_t, N_{t-1}, \ldots, N_0, D_t, D_{t-1}, \ldots, D_0, A_t, A_{t-1}, \ldots, A_0, B_H, \mu, \mu_t, \mu_{t-1}, \ldots, \mu_0). \tag{7}$$

Time is treated as discrete and all characteristics are measured at the end of the time period. We have assumed that the unobserved, individual-level health endowment comprises two components: one that is time-invariant, μ, and is known to the individual, and another that varies over time, $\mu_\sigma, \sigma = 0, \ldots, t$. In principle, all past realizations of μ may be known to an individual although this knowledge may come at a cost (such as testing) and the implications of these innovations may not be fully absorbed.

2.2.2. *Allocating resources over the life course*

At each point over the life course, t, an individual will choose consumption, C, and labor supply, L, to maximize the present discounted value of life-time welfare, Ψ, conditional on health outcomes, H, demographic characteristics, A, background, B_U, and tastes, ξ, a component of which is time invariant and a component which may vary over time:

$$\Psi_\tau = \Psi\left\{E_\tau \delta_t^{t-1} U_t(C_t, L_t; H_t, A_t, B_U, \xi, \xi_t)\right\} \quad t = \tau, \tau+1, \ldots, T. \tag{8}$$

Welfare in any period, τ, is updated depending on realizations of unobserved characteristics at that point. The individual's discount rate δ_t is allowed to vary over the life course. We have made no assumptions about separability of preferences over time and so welfare may depend on prior health and demographic characteristics as well as predictions about their trajectory in the future.

Choices are constrained by the period-specific technology of health production, a period-by-period time budget constraint which limits the sum of hours of leisure, work and time spent on investment in health in that period and a lifetime budget constraint:

$$V_T = \prod_{t=1}^{T}(1+r_t)V_0 + \sum_{t=1}^{T}\left(\prod_{\tau=t}^{T}(1+r_\tau)\right)\left\{w_t L_t - \left(p_{ct}C_t^* + p_{nt}N_t^c\right)\right\}. \tag{9}$$

V_T is assets at the end of life which may be positive (bequests), zero or negative (debts). V_0 is assets at the beginning of life (inheritances) and r_t is the interest rate so the first term on the right-hand side of (9) is the present discounted value of initial assets. Asset accumulation in each period $t = 1, \ldots, T$ is the value of the difference between earnings and expenditure that period. Earnings is the product of working L_t hours at wage w_t. Expenditure is given by spending on consumption goods C_t^* and inputs into the health production function N_{ct} valued at prices p_{ct} and p_{nt} respectively. The second term in (9) is the sum of the present discounted value of additions to assets in each period. The extent to which total assets in any period can be negative is limited by whatever constraints exist on borrowing. They are straightforward to include but are not explicitly specified in this general framework.

2.2.3. Economic productivity and health over the life course

To derive dynamic demand functions for health inputs and outcomes, we need to specify the determinants of wages and labor supply. Taking as given that some dimensions of health contain both a stock and a flow component, labor productivity, w, will likely depend not only on current health, but also prior health investments, N, and health outcomes, H. For example, the former will occur if an investment in health in early adulthood (say exercise) affects productivity (because of elevated endurance) over and above the impact of health outcomes on productivity. An example of the latter might arise if being overweight at a point in time affects productivity later in life, conditional on current weight. As above, we collect health inputs and outcomes in the vector θ.

Current labor productivity will also vary with time-invariant background characteristics that affect labor market outcomes, B_w, as well as current and prior demographic characteristics, A, that do vary with time. Wages will also depend on variation in economic infrastructure and labor demand over time, I:

$$w_t = w(\theta_t, \theta_{t-1}, \ldots, \theta_0; B_w, A_t, \ldots, A_0, I_t, \ldots, I_0; \alpha, \alpha_t). \tag{10}$$

There are at least three sources of unobserved heterogeneity, each of which may be correlated with current or prior θ. First, time-invariant unobserved ability, α, may be correlated with health if some people are inherently better able to manage health risks or disease insults. For example, Goldman and Smith (2002) demonstrate that the better educated are better at self-managing health problems by adhering more closely to treatment protocols. Second, time-varying unobserved ability, α_t, may depend, in part, on prior health outcomes. This will arise, for example, if early childhood health affects cognitive development or performance in school (Pollitt et al., 1993; Behrman, 1996). Third, part of α_t reflects measurement error in wages which may be related to current effort that is affected by health status.

Note that the model (10) is specified in a sufficiently general form so that levels of health and changes in health may interact in their influence on productivity and hours of work. For example, the effect of a loss in body mass due to a crippling bout of diarrhea is likely to differ depending on the initial level of body mass of the individual. It is worth emphasizing that the multi-dimensionality of health inputs and outcomes underscores the importance of taking care in specifying non-linearities in relationships between health and economic success. For example, there may be non-linear relationships between wages and body mass as seen in the U-shaped association between BMI and mortality. However, it is also possible that there are important interactions between dimensions of health – and other controls – and those can provide insights into the complementarities and substitutability between different health domains. Crimmins and Finch (2006) provide an example. They argue that part of the dramatic increases in life expectancy in the 19th and 20th century can be attributed to the complementary effects of improved nutrition and reduced inflammation because of lower rates of infection and trauma. Since inflammation and BMI tend to be positively correlated (Crimmins et al.,

2007), a non-linear relationship between BMI and wages might be a reflection of a complementary effect of inflammation.

As in the static case, there is an analogous relationship between labor supply and health, controlling demographics, schooling, background and infrastructure. The same issues that arise with estimation of the wage function (10) also apply in the labor supply case. In addition, labor force participation and hours of work in the current period might be influenced by expectations about future health. One way to mitigate the negative consequences of poor health in the future (or to pay for future health inputs) might be to work more now and take more leisure in the future. This inter-temporal trade will depend on substitutability between leisure and work in good and poor health and, also, on liquidity and credit constraints. If effort expended at work, and hence effective labor supply, is influenced by future health, then the labor supply function will depend not only on current and prior health but also expected future health inputs and outputs. Similar arguments apply for expectations about future demographic characteristics, the disease environment and infrastructure.

2.2.4. Dynamic conditional demand for health

Solving the life cycle optimization program (7)–(10) yields a series of dynamic demand for health functions, θ, one for each health input, N, and outcome, H, which are conditional on health in other periods:

$$\theta_t = \theta\big(\theta_\sigma, m_t(\theta_\tau), p_{nr}, p_{cs}, A_s, B, V_0, D_s, I_s, \varepsilon, \varepsilon_s\big) \tag{11}$$

where

$$\theta_\sigma = \theta_{t-1}, \theta_{t-2}, \ldots, \theta_0; \qquad m_t(\theta_\tau) = m_t(\theta_{t+1}), m_t(\theta_{t+2}), \ldots, m_t(\theta_T),$$
$$Z_s = Z_0, \ldots, Z_{t-1}, Z_t, m_t(Z_{t+1}), \ldots, m_t(Z_T) \quad \text{for } Z = p_c, A, I \text{ and } \varepsilon$$
$$\text{and} \quad Z_r = Z_t, m_t(Z_{t+1}), \ldots, m_t(Z_T) \quad \text{for } Z = p_n \text{ and } D.$$

Each element of the vector of current health inputs and outcomes, θ_{jt}, depends on a vector of lagged health inputs and outcomes which date back to birth (or *in utero*), $\theta_\sigma = \theta_{t-1}, \theta_{t-2}, \ldots, \theta_0$. Since current health depends on prior health inputs, a forward looking person will choose inputs this period in an anticipation of future health inputs and outcomes. In general, therefore, demand for health in the current period, t, will depend on expectations at time t about future health $E_t\theta_\tau$ for each future period from $\tau = t + 1$ until $\tau = T$ which is the (expected) end of life. Since future health is uncertain, it is not only the expected value of future health that will enter (11) but the entire distribution of predicted future health that potentially affects current health and health investments. We denote that distribution evaluated at time t by $m_t(.)$. Obviously expected end of life is endogenous in this model as is the actual end of life since longevity will depend on health investments and resource allocations through the life course.

Current health inputs and outcomes also depend on current health input and consumption prices, p_n and p_c respectively, as well as all expectations about future prices. Prior

prices of inputs into the health production function, p_{nt}, should affect θ_t only through lagged health and so do not belong in the dynamic conditional demand function (11). However, prior prices of all consumption goods (and any inputs that enter directly into the utility function (8) belong in (11). To ease notation, we use the subscript s to denote these past, present and the distribution of predicted future values and subscript r denotes current and expected future values.

The conditional demand for health in any period is dependent on time-varying demographic characteristics, A_s, non-health human capital and background, B, which are time invariant; assets at birth, V_0, as well as local labor demand and infrastructure, I_s. Again, the subscript s denotes all prior, current and predicted future values of time-varying demographic characteristics and infrastructure. As is the case for prices of health inputs, demand for health depends on current and future expectations about the disease and health services environment, D_r. (If there are any elements of D that overlap with infrastructure, I, then their state in prior periods will affect demand.) Without loss of generality, all dimensions of background and human capital that are time-varying are captured in A_s. Time invariant background characteristics, B, include all the factors that affect the production of health, B_H, tastes, B_U, and wages, B_W.

Unobserved heterogeneity, ε, encompasses innate healthiness, μ, in the health production function, innate ability, α, in the wage function, tastes that influence resource allocation in any period, ξ, as well as inter-temporal preferences, δ. Since each of μ, α, and ξ has a time-invariant and time-varying component, unobserved heterogeneity in the conditional demand for health function is also separated into a time invariant, ε, and time-varying component, ε_s. The time-invariant component includes, for example, genetic endowment, ability and those tastes that do not change over the life course. At any point in the life course, the entire history of innovations in unobserved factors $\varepsilon_0, \ldots, \varepsilon_t$ affects demand for health in the current period as do expectations about the future evolution of those factors, $\varepsilon_{t+1}, \ldots, \varepsilon_T$.[4] Future innovations can be ignored if it is only their first moments that affect demand for health, assuming $\varepsilon \sim f(0, \sigma^2)$. In general, higher order moments of innovations will play a role in decisions about health if, for example, the possibility of large negative (or positive) innovations in unobserved dimensions of healthiness affect decisions today perhaps because of risk aversion.

2.2.5. *Dynamic reduced form demand for health*

Before discussing estimation of the life cycle models of health and productivity, it is useful to derive the reduced form demand for health inputs and outcomes at time t by substituting all endogenous variables in (11) with their determinants. These potentially endogenous variables include all prior levels of health $\theta_{t-1}, \theta_{t-2}, \ldots, \theta_0$, expectations about future health $m_t(\theta_{t+1}), m_t(\theta_{t+2}), \ldots, m_t(\theta_T)$, and predictions about the evolution of socio-demographics, A, the disease environment, D, and infrastructure, I. The

[4] Note that prior innovations in innate healthiness, μ, play no role in ε.

determinants of these potentially endogenous variables are prices and the information set that governs predictions about the future:

$$\theta_t = \theta(p_{n\sigma}, p_{c\sigma}, A_\sigma, B, V_0, D_\sigma, I_\sigma, \upsilon, \upsilon_\sigma)$$
$$\text{where} \quad Z_\sigma = Z_0, \ldots, Z_{t-1}, Z_t \quad \text{for } Z = p_n, p_c, A, D, I \text{ and } \upsilon. \tag{12}$$

It is important to note that all dimensions of A, D and I that are potentially correlated with unobserved characteristics that affect health do not belong in the reduced form demand function. These will include, for example, the local disease or infrastructure environment if the individual chooses to live in a particular location because of their impact on health either directly (through reduced infection) or indirectly (through improved earnings opportunities).

2.2.6. Estimation of dynamic demand for health functions

In principle, estimation and interpretation of dynamic reduced form demand for health inputs and outcomes, (12), is straightforward. However, the data demands are beyond the reach of current population survey data resources since the model calls for knowledge of all prices of health inputs and consumption goods in the current period and in every period back to conception along with a complete demographic and environmental history for each respondent.

Estimation of the dynamic conditional health function is even more demanding of data. Not only does current health depend on the complete health history, price, demographic and environmental history but also on expectations about all of these processes including the evolution of future health.

In the absence of such data, it is necessary to place restrictions on the model and reduce the dimensionality of the estimation problem. We describe three separability assumptions that are commonly made, often only implicitly, in this literature.

First, few empirical studies of health demand functions examine more than one or a small number of domains of health at any time. These studies assume that the specific markers capture all domains of health or that other health inputs and outcomes have no direct impact on the health markers of interest.

Second, if health in the previous period is a sufficient statistic for all health prior to that period, then the conditional demand function simplifies to:

$$\theta_t = \theta\big(\theta_{t-1}, m_t(\theta_{t+1}), p_{nr}, p_{cs}, A_s, B, V_0, D_r, I_s, \varepsilon, \varepsilon_s\big)$$
$$\text{where} \quad Z_r = Z_t, m_t(Z_{t+1}) \quad \text{for } Z = p_n \text{ and } D$$
$$\text{and} \quad Z_s = Z_0, \ldots, Z_{t-1}, Z_t, m_t(Z_{t+1}), \ldots, m_t(Z_T) \quad \text{for } Z = p_c, A, I \text{ and } \varepsilon. \tag{13}$$

The conditional demand model can be estimated with knowledge of health and health input prices in only three periods: the prior period, current period and expectations about the next period. However, all other control variables (p_c, A and I) continue to enter the demand function in all past, current and future periods.

The third assumption, which is often invoked in models of life cycle consumption and labor supply, requires that preferences are additively (or strongly) separable over time:

$$\Psi = E\sum_{t=1}^{T}\delta_t^{t-1}U_t(C_t, L_t; H_t, A_t, B_U, \xi, \xi_t) \tag{14}$$

which places restrictions on the way that out-of-period characteristics affect current period decisions. Specifically, the marginal utility of income (or its inverse, the price of utility) is fixed over time and so, given that price, decisions in any one period can be made without regard for any prices or other characteristics in any other period, conditional on the price of utility.

Combining these assumptions the conditional dynamic demand function simplifies dramatically to:

$$\begin{aligned}\theta_t &= \theta\big(\theta_{t-1}, m_t(\theta_{t+1}), p_{nr}, p_{cs}, A_s, B, V_0, D_r, I_s, \varepsilon, \varepsilon_s\big)\\ &\quad \text{where} \quad Z_s = Z_{t-1}, Z_t, m_t(Z_{t+1}) \quad \text{for } Z = p_n, p_c, A, I \text{ and } \varepsilon\\ &\quad \text{and} \quad Z_r = Z_t, m_t(Z_{t+1}) \quad \text{for } Z = p_n \text{ and } D.\end{aligned} \tag{15}$$

Taken together, these are very powerful assumptions. The conditional demand model can be estimated with knowledge of not only health but also all control variables in only three periods: the prior period, current period and expectations about the future. This contrasts with (13) which depends on all past, current and future values of the control variables. Note also that these assumptions have no effect on the reduced form which continues to depend on all prior values of all covariates.

The assumption of inter-temporal separability of preferences in combination with assuming that θ_{t-1} is a sufficient statistic for health prior to that date is not only powerful. These are also very strong assumptions. The credibility of additive separability will depend on the completeness of financial markets. The credibility of assuming previous period health is a sufficient statistic for all prior health will vary depending on the specific measure of health and whether prior health problems are likely to have long-lived effects on later health.

An increasing amount of non-experimental evidence has accumulated that establishes an association between health *in utero* or in very early childhood and later life health and non-health outcomes (see Barker, 1994, and Godfrey and Barker, 2000 for instance). It seems that it would be profitable to draw on insights from the biological and biomedical literatures on the persistence of changes in specific domains of health to aid the selection of plausible exclusion restrictions for the model. Depending on the data available, it may be possible to allow health in a limited number of other periods to affect current health. Alternatively, the exclusion restrictions can be subjected to testing by determining whether health and the control variables measured in other periods have a direct impact on current health (appropriately taking account of unobserved heterogeneity).

Some surveys collect information about expectations regarding the future, including future health. In the absence of that information, it will be necessary to specify a

model of expectation formation in order to estimate the dynamic conditional demand functions (11) or (15).

There are at least two additional challenges to estimation of the functions. First, it is likely that unobserved heterogeneity, ε, is correlated with one or more of the observed characteristics in the model. Consider, first, the component of unobserved heterogeneity that is time-invariant. This includes genetic endowments and innate healthiness, ability and tastes that do not change over the life course. These are likely to be correlated with health in any period as well as family background such as parental health and schooling. This suggests including an individual-specific fixed effect in the model. This raises two concerns. First, the fixed effect will also absorb all observed characteristics that are fixed, such as height, that enter the model in a linear and additive way. This is unfortunate since the relationship between these observed characteristics and health are often of substantive interest. Second, while convenient, the assumption that all time-invariant observed and unobserved characteristics enter the model in a linear and additive way is typically *ad hoc* and may not be innocuous. For example, as noted above, it is possible that height or innate healthiness, ability or completed education might interact with other time-varying covariates that are included in the model; in that case, the fixed effects estimates will be biased.

The second challenge to the estimation of the dynamic conditional demand function is that prior health and expectations about future health will also be correlated with unobserved heterogeneity that is time-varying. For example, unobserved factors, μ_{t-1}, that affect the production of health outcomes, H_{t-1}, and possibly inputs into that production, N_{t-1}, will be correlated with θ_{t-1}. This suggests an instrumental variables approach to estimation, possibly in combination with fixed effects. In the general form of (11), there are no obvious instruments. However, given some exclusion restrictions on the lags and leads of θ included in the model, all covariates from all other periods are potential instruments. For example, imposing the assumptions underlying (15), all prices, demographics, disease environment and infrastructure from $t = 0$ through $t - 2$ and all expectations of those characteristics from $t = t + 2$ through T are candidate instruments. Health inputs and outcomes from the same periods are also potential instruments.

2.2.7. *Estimation of dynamic health production functions*

As with the other models, estimation of the dynamic health production functions (7) is very demanding of data. Invoking the same separability assumptions used above, rewrite the production function with health in the previous period serving as a sufficient statistic for all health inputs and outcomes for all other prior periods:

$$H_t = H(H_{t-1}, N_t, L_t; D_t, A_t, B, \mu, \mu_{t-1}). \tag{16}$$

Subtracting H_{t-1} from both sides, the specification can be interpreted as relating changes in health (such as growth in child height or changes in weight) to health in

prior periods and other contemporaneous inputs into the production function. The assumption that H_{t-1} is a sufficient statistic for all prior health reduces the empirical problem to explaining flows in health and not the evolution of the stock of health over the entire life course. Estimation needs to take into account the fact that prior health, H_{t-1} and contemporaneous inputs, N_t, are likely correlated with time invariant innate healthiness, μ, and time-varying unobserved variation in healthiness, μ_{t-1}. An individual fixed effect will absorb the impact of μ if its influence is approximately linear and additive. Prices of inputs and earlier period characteristics in the production function are potential instruments for prior health and contemporaneous inputs. Again, it may be profitable to combine a fixed effects and instrumental variables approach and it is straightforward to assess the validity of the assumptions underlying the specification using over-identification tests.

The instrumental variables approach was adopted in an important paper by the Cebu Study Team (1992) which used extremely rich data to estimate a complicated dynamic weight function for children every two months from birth to 24 months of age in Cebu, Philippines. The study assumes that weight in any period depends on inputs that period and weight in the prior period; weight and inputs in other periods do not influence current weight except through previous-period weight.

2.2.8. *Estimation of dynamic relationships between health and economic prosperity*

Examination of the dynamics underlying the relationship between health, labor market prosperity and socio-economic success provides several additional substantive insights and empirical challenges. We begin with models of the effect of health on economic prosperity and then examine dynamic demand for health conditional on indicators of SES.

As above, it is necessary to impose exclusion restrictions to yield an empirically tractable model. The simplest model treats current health as a sufficient statistic for all past and future health inputs and outcomes

$$w_t = w(\theta_t; B_w, A_t, I_t, \alpha, \alpha_t). \tag{17}$$

Dynamics are embedded in time-invariant unobserved heterogeneity, α, which, in a pioneering paper, Deolalikar (1988) treated as a fixed effect. This amounts to examining the effect of changes in health on changes in wages.

Consistent estimation of (17) requires that differences in the time-varying errors, $(\alpha_t - \alpha_{t-1})$, be uncorrelated with differences in time-varying health characteristics, $(\mu_t - \mu_{t-1})$. This is plausible if the time-varying errors only contain random measurement error. The assumption is less appealing if the errors also comprise omitted characteristics which affect wages since part of these wage shocks may be invested in health. This will arise, for example, if a worker experiences a surprisingly good year and spends some of the unanticipated income on health-augmenting inputs (perhaps by eating more nutritious foods). In that case, it makes sense to combine the fixed effect with an instrumental variables approach. The prices of health inputs in prior periods

and expected prices in future periods are potential instruments. Prices of consumption goods in periods other than t are also potential instruments as long as they have no direct effect on wages. This precludes, for example, the prices of crops produced by a farmer. In this specification, prior realizations and future expectations about demographics and infrastructure are also instruments if they do not have a direct impact on wages.

Returning to the assumption that previous period health is a sufficient statistic for all prior health, the conditional wage function becomes:

$$w_t = w(\theta_t, \theta_{t-1}, E_t\theta_{t+1}; B_w, A_t, A_{t-1}, E_tA_{t+1}, I_t, I_{t-1}; \\ E_tI_{t+1}, \alpha, \alpha_t, \alpha_{t-1}, E_t\alpha_{t+1}) \tag{18}$$

where the moments, $m(.)$, are replaced with expectations, $E_t(.)$, to facilitate exposition. As in our earlier notation for moments, $m(.)$, expectations, $E_t(.)$, are taken over all future periods, from $t+1$ onwards. The model in (18) can be estimated using a fixed effects instrumental variables approach (with instruments drawn from observed characteristics in periods not included in the model). Separating θ_t into a part that is anticipated at time $t-1$, $E_{t-1}\theta_t$, and a part that is not, $\theta_t - E_{t-1}\theta_t$, highlights the differences between the effect of anticipated changes in health and health shocks. Prior health, demographics and infrastructure can also be separated into a component that is anticipated and a component that is not.

This extended model provides opportunities to address several important questions about the relationship between prior health, changes in health, current health and labor outcomes. For example, to what extent do current episodes of ill-health, especially unforeseen health shocks, affect labor market outcomes now and in the future? Does it matter whether ill-health episodes are expected or unexpected, as in principle it should? How important is previous health, conditional on current health, in determining labor market outcomes and health itself, over the life course? What types of health shocks have permanent effects on health, what types of shocks can be overcome by later periods of good health, and what types of shocks can be ameliorated by positive health interventions? By allowing for separate effects of current and past health on wages in (18), it is possible to examine the impact of *changes* in health status, or of the joint effects of levels and changes in health, on labor market, health and schooling outcomes. To date, there is only limited evidence on these types of questions.

Note also that the same conceptual framework described for wages in (10), (17) and (18) can be applied to explore the relationship between health and other labor outcomes such as labor force participation, sectoral choice and hours of work although identification in those models is more complicated. First, it is likely that demographics and infrastructure will have a direct impact on labor supply choices. For example, if labor demand is expected to fall in the future, an individual may choose to work more hours in this period. Second, prices of any goods that affect the real shadow price of time are not valid instruments for health.

As in the static model, there is substantial interest in the reverse relationship: the demand for health conditional on economic prosperity. The reduced form, (12), provides a basis for interpreting the impact on the evolution of health over the life course

of schooling, any other background indicator included in B, and assets at birth, V_0. Models of health conditional on earnings or total income are more difficult to interpret. Substituting for prior and future health in the conditional demand function (11), health in the current period depends on wages, w, and labor supply, L, as well as wealth, V, input and consumption prices, p, demographics, A, background, B, and the disease environment, E. All current and prior realizations of these control variables along with predictions about how they will evolve in the future belong in the model:

$$\begin{aligned}\theta_t &= \theta(w_s, L_s, V_s, p_{ns}, p_{cs}, A_s, B, D_s, \varsigma, \varsigma_s)\\ &\text{where} \quad Z_s = Z_0, \ldots, Z_{t-1}, Z_t, m_t(Z_{t+1}), \ldots, m_t(Z_T)\\ &\text{for } Z = w, L, V, p_n, p_c, A, D \text{ and } \varsigma. \end{aligned} \tag{19}$$

As above, it is useful to separate prior realizations of each covariate into a part that is predicted and a part that is not anticipated. Estimation of (19) is a challenge. If a fixed effect is included in the model to absorb time-invariant unobserved heterogeneity, ς, then, as long as individuals are able to transfer resources across periods, it is only unanticipated components of realized wages, hours of work, wealth and all other covariates that will affect current health status. The longevity of unanticipated fluctuations in these covariates is an empirical question and will depend on both the extent to which the measure of health reflects the role of a stock or flow and the specific covariate.

2.3. *Health, the individual, the household and the family*

The discussion thus far has focused on decisions about health investments, resource and time allocation made by an individual. Many of these decisions are made within the context of decisions made by other household or family members and take into account their health and well-being. For example, if one member of a family falls ill, other members might respond by reallocating time and goods to provide care to that person, to avoid other members also falling sick and to offset transitory fluctuations in income or expenditures associated with the illness. Reactions by family members will likely be different if the illness is expected to be persistent, or if the person dies, and there is a permanent reduction in family resources.

The extent to which household and family members are able to smooth fluctuations due to anticipated and unanticipated variation in health of a family member is of substantial interest. In addition, there are important questions about how the burden of poor health is distributed within the family. There is a substantial literature on the allocation of resources among children with different endowments. (See, for example, Behrman, Pollak and Taubman, 1982.) On the one hand, from the perspective of maximizing family resources, it may be optimal for a family to invest most in the child with the greatest endowment of innate healthiness. On the other hand, families may allocate resources to offset these innate differences (Pitt, Rosenzweig and Hassan, 1990).

More generally, how family members with different economic opportunities, endowments and preferences co-ordinate decisions and allocate resources is an active area of research. Incorporating insights from those models into the conceptual framework described above is beyond the scope of this chapter. Browning, Chiappori and Weiss (2007) provide an excellent review of the literature; see also Behrman (1997), Pitt (1997) and Mwabu (2008). We will explore some of the empirical evidence on these issues below. For example, we will discuss the impact of illness and death in areas which HIV/AIDS is prevalent on adults and children; we will also discuss evidence on how the impact of income shocks on health is distributed within families.

The discussion of conceptual models thus far has highlighted at least two issues that are important for the design of research on health and development: the likely contributions of longitudinal studies and the potential value of studies designed to pin down causal mechanisms. We take up those issues next.

2.4. Longitudinal study designs

As discussed in the previous section, studies of the relationships between health and other life outcomes that rely exclusively on contemporaneous associations substantially limit the scope for understanding the complexities underlying these relationships. Dynamic models of behavior call for measurement of health and life outcomes over the life course. Investments in high-quality longitudinal data sets have had substantial payoffs in terms of their contributions to the field. The recent literature on the long arm of early childhood health and its impact on later life outcomes suggests that it may be profitable to develop long-term longitudinal studies which follow individuals from the fetal period through to death and, arguably, studies that follow the children of those respondents. While this is an expensive endeavor, it is not impossible as evidenced by several high-quality cohort studies which have followed individuals from conception through adolescence into adulthood. These include, for example, five British cohorts, the Cebu study conducted by Popkin, Adair and associates, and the Birth to Twenty cohort in South Africa (Barbarin and Richter, 2001).

Attrition is the Achilles heel of longitudinal surveys. In telephone and web-based surveys, participant refusal drives much of the attrition. Refusal rates are typically very low in developing country studies. Most of those surveys involve face-to-face interviews and attrition is primarily due to migration away from the study sites since movers are the hardest and most expensive to follow-up. Until recently, many panel surveys were designed to exclude migrants from follow-up.

Migration is typically positively correlated with human capital and so is likely to be associated with some dimensions of health. The extent to which attrition is correlated with observed characteristics that affect health inputs and outcomes will complicate interpretation of the models discussed in the previous section. Attrition that is correlated with unobserved characteristics in the models poses substantially greater empirical challenges. At the very least, it will be prudent to assess the extent to which attrition is correlated with observed characteristics that are included in the models of interest and

take those results into account in the interpretation. To the extent possible, it would also be worth relating attrition to observed characteristics, which are not included in the model, but are likely to be correlated with unobserved factors in the model of interest. The choice of characteristics will be model-dependent but might include the example of migration above, the value of time, performance on cognitive tests and so on.

It is possible to model attrition drawing, say, on characteristics of the survey itself and the enumerators as predictors of survey quality that are not related to the health behaviors of respondents. (See, for example, Thomas, Frankenberg and Smith, 2001.) It is also possible that judiciously combining longitudinal surveys with data collected on health status and health inputs retrospectively will yield useful resources for the research community. (See, for example, Schultz and Tansel, 1997; Smith, 2004.) Of course, none of these is a substitute for spending resources on surveys that are well-designed and maintain high retention rates across waves.

2.5. *Studies designed to identify causal effects*

The second issue highlighted by the discussion of the conceptual models is that identifying causal effects is far from straightforward. This is not a new insight. However, in recent years, there has been greater emphasis on designing studies that have the potential to disentangle causal pathways and those studies promise to provide important new evidence. It is useful to identify three classes of studies: randomized experiments, quasi-experiments and other non-experimental methods. See Duflo, Glennerster and Kremer (2008) for an extensive discussion of the theory and practice of randomization in development economics. Angrist and Krueger (1999) provide an excellent overview of non-experimental methods.

2.5.1. *Experimental designs*

Experimental studies are designed to randomly assign respondents into treatment and control groups. In principle, a comparison of changes in behavior between the groups after the treatment provides evidence on the causal effect of the treatment. Examples of treatment-control designs include providing nutritional supplements (as in the INCAP study or the Mexican Oportunidades program); providing drugs to treat a health problem (such as drugs to treat HIV/AIDS, schistomasias or worms, Miguel and Kremer, 2004), changing the epidemiological environment (by draining rivers and canals or providing bed nets in areas where malaria is endemic, Watson, 1953) or changing the prices of health services (as in the RAND Health Insurance Experiment, Newhouse, 1996) or the availability, quality and thus effective price of health services (by providing incentives for health service providers to attend clinics, Banerjee, Deaton and Duflo, 2004).

It is estimated that nearly 2 billion people suffer from iron deficiency anemia (IDA) with prevalence rates being highest in Asia. While rigorous clinical studies on animals

and humans have demonstrated that IDA reduces work capacity,[5] there is little evidence on the impact of iron deficiency on wages, time allocation and other economic choices. The Indonesian Work and Iron Status Evaluation (WISE) was specifically designed to provide evidence on the extent to which there is a causal effect of health on economic prosperity. About half the 17,000 subjects were assigned to receive a weekly iron supplement. The other half received an identical-looking placebo. Take up of the supplement was monitored very closely. Subjects were interviewed every four months – for a year before supplementation, the year of supplementation and for two years after supplementation. The impact of supplementation on iron in the blood was measured at each interview along with a rich array of economic and social indicators. It is thus possible to trace out the impact over time of improved nutritional status by contrasting changes in the lives of the treatments relative to the controls (Thomas et al., 2006).

In practice, treatment-control design experiments can be difficult to implement successfully. For example, if the treatment is effective then respondents in control groups are likely to observe others who are benefiting from the treatment and respond by seeking out the treatment. In order to minimize this effect, several experiments randomize at the level of the village or school so that all peers receive the same treatment. This complicates interpretation of evidence at any other level (such as the individual) and thus dissipates some of the key benefits of the experimental design. More generally, subjects are likely to change their behavior in response to the experiment and these behavioral responses further contaminate interpretation of experimental evidence.

Moreover, experimental designs are often expensive to implement and there are many key questions about health which are either difficult or impossible to address using experimental designs. For example, it is difficult to prospectively randomize *in utero* development to measure the causal impact of fetal deprivation on outcomes in later life. Because most experiments are conducted on specific sub-populations for a limited period of time or, if the experiment is long-lived, there is often substantial attrition from the sample and the degree to which attrition is different between treatments and controls is likely to be related to the perceived net benefits of the treatment. This substantially contaminates interpretation of results from experiments. More broadly, it is often difficult in social experiments to establish the generalizability of results, assess the longevity of impacts or provide insights into likely general equilibrium effects.

With all of these concerns in mind, and as discussed in more detail below, randomization has been, and promises to continue to be, an extremely valuable tool in this literature. It seems likely that combining the exogenous variation provided by randomization with non-experimental methods that take into account behavioral responses is likely to be a very productive approach to addressing some of the most pressing unresolved questions in the literature. We turn, therefore, to a discussion of non-experimental methods to pin down causal effects.

[5] The biochemical literature has established the key role that iron plays in transporting oxygen through the blood and the mechanisms underlying the causal effect of iron on VO_2max, say, are well-understood right down to the cell level.

2.5.2. Quasi-experimental designs

Many important studies have used a quasi-experimental design exploiting variation in the social, economic or health environment surrounding individuals. To the extent that this variation is outside of the control of the individual (or is uncorrelated with observed and unobserved characteristics that affect the outcome of interest in the empirical model), it provides plausibly exogenous variation which is exploited to pin down causal effects. For example, studies have exploited variation in weather, variation due to natural disasters, unanticipated changes in the epidemiological environment, the introduction of new drugs and changes in health services, as well as changes in public policies associated with the provision of health, education, income or some other type of support.

Studies of the 1944 Dutch Winter Famine provide an example of a quasi-experimental design that has proved to be very powerful (Stein et al., 1975; Ravelli et al., 1998 and Roseboom et al., 2001b, for example). In the winter of 1944/1945, western Holland was under Nazi occupation and, as punishment for resistance activities, the Germans gave each person a very small daily ration of between 400 and 800 calories. Children who were *in utero* during this period have been weighed and measured from birth throughout their lives. Their children have also been followed to investigate the inter-generational consequences of deprivation during the fetal period.

In practice, it is often difficult to identify a "treatment effect" by comparing health inputs, outcomes or behaviors of the same group of people before and after some "treatment" because at least part of the observed differences may be due to secular change. To control for this, many studies compare differences over time among the treated group with differences over the same period among a control group and attribute the difference-in-difference to the causal effect of the 'treatment.' A key issue is selection of control groups that mimic the treatment group.

Spatial or temporal variation in the introduction of new programs, policies or health services is often exploited to measure the effect of the change. In some cases, this variation can be treated as randomly assigned, say, because the new program is introduced in a phased way and the phasing is unrelated to observed or unobserved characteristics in the model of interest. However, in many cases, programs are targeted to place where the need is greatest or where implementation is easiest because of the existence of other, complementary infrastructure. In these cases, the difference-in-difference estimates will be misleading unless the differences are fully controlled. That said, arguably, developing countries are an especially good context for studies that exploit a quasi-experimental design precisely because of the many rich sources of variation they provide both across time and space which can be treated as randomly assigned.

Several studies have exploited changes in the environment in combination with insights from the biological literature to develop quasi-experimental designs. As an example, under the assumption that child length, conditional on age, is little affected by health and nutrition inputs after age 3 or 4 years (Habicht, Martorell and Rivera, 1995), it is possible to measure the impact of health interventions, school feeding programs and public support programs, *inter alia*, by comparing the length of children exposed

with those not exposed to the program. Such studies include examination of the impact of the old age pension program in South Africa (Case and Deaton, 1998; Duflo, 2000; Hamoudi and Thomas, 2005) and the introduction of midwives to villages in Indonesia (Frankenberg, Suriastini and Thomas, 2005).

To be sure, not all observed variation is "natural" and it is imperative that the first step in any quasi-experimental design is to demonstrate that the sources of variation that are being exploited can be treated as exogenous. See, for example, Rosenzweig and Wolpin (2000) for an insightful discussion.

2.5.3. Other non-experimental methods

Non-experimental data are the workhorse of empirical research in development. In reduced form models, estimates of causal effects are provided by ordinary least squares (OLS) or matching methods. Matching, be it on covariates or through estimation of propensity scores (Rosenbaum and Rubin, 1983), seeks to replicate the advantages of a treatment-control design. Importantly, matching places fewer parametric restrictions on the relationships of interest relative to OLS which assumes the model is linear and additive in unobserved characteristics. It is also straightforward to estimate the distribution of the effects of characteristics in these models using non-parametric methods (Hardle, 1992). All of these methods assume selection is only on observed characteristics.

Estimation of causal effects in production functions and conditional demand functions is more complicated because selection is likely to be on unobserved characteristics that are correlated with covariates in the model. The biases caused by these unobserved characteristics will affect OLS and matching method estimators. To the extent that these unobserved characteristics are fixed for an individual and affect the outcome in a linear and additive way, they can be absorbed by including a fixed effect in the model. As noted above, a drawback of a fixed effect is that it absorbs all characteristics that are fixed, such as height or education, and enter the model in a linear and additive form. In addition, it is well known that the inclusion of a fixed effect exacerbates problems associated with classical measurement error in covariates – an issue that is of considerable import when estimating relationships between health and indicators of SES (Griliches and Hausman, 1985).

In general, it will be necessary to identify instruments that are not correlated with the unobserved characteristics in the model but are correlated with the "endogenous" covariates or "treatments." Instrumental variable (IV) methods are well established in the literature but involve strong parametric assumptions. If the unobserved characteristics that affect the "treatments" are additively separable from the observed and unobserved characteristics affecting the outcome of main interest, then IV will identify the local average treatment effect (LATE). Notice this rules out the possibility of heterogeneity of treatment effects across the distribution of unobserved characteristics. This is a potentially important restriction in the health and development literature. Not only is the distribution of "effects" of paramount interest in some cases, but the biological literature suggests that at least for some health indicators such as glucose, cholesterol,

inflammation and stress markers, there may be heterogeneous "treatment effects" on, say, economic productivity across characteristics that are not usually observed. The latter might include genetic endowment, prior and concurrent health and related behaviors.

Greater flexibility can be achieved with control function methods, pioneered by Heckman (Heckman, 1978; Heckman and Robb, 1985), or local instrumental variable methods (Heckman and Vytlacil, 2006). For example, control function methods include controls for unobserved heterogeneity that is correlated with covariates in the model of interest (the so-called "λ-method") and these controls can be estimated nonparametrically. (See Blundell and Powell, 2003, and Heckman and Navarro-Lozano, 2004, for excellent discussions of these methods.)

Studies have attempted to exploit the temporal ordering of variation in health and economic status to isolate "Granger" causality (Granger, 1969). Using longitudinal data on older Americans from Asset and Health Dynamics among the Oldest Old (AHEAD), Adams et al. (2003) relate the incidence of new health conditions to prior health conditions and SES in the previous period (which is two years earlier). Finding that SES does not predict future health problems (except for mental health), they conclude that SES does not have a causal impact on physical health in this sample. Conversely, controlling prior wealth, prior health does predict current wealth and so causality from health to wealth cannot be ruled out.

Inspection of the dynamic model (7)–(12) above highlights one of the difficulties with this approach. In general, the evolution of both health and SES are the outcomes of choices made over the entire life course and the relationship between health and SES in a short window may not be informative about the causal mechanisms that underlie their association over the life span. This concern is likely to be particularly germane in early life (when SES amounts to parental resources) and later life (when health and SES reflect the cumulation of decisions made earlier in life).

In addition, all the issues discussed above regarding identification of the causal effect of health on SES and *vice versa* are relevant in tests of Granger causality and they are not easy to address. These include concerns regarding correlations between covariates included in the empirical models and unobserved characteristics. For example, wealth in any period is a function of prior spending and savings decisions as well as portfolio allocations; those choices are likely to have affected prior health and possibly will affect health in the future over and above the value of wealth in the prior period. More generally, innate healthiness and preferences, which are typically not observed, are likely to be correlated with prior health and SES, and estimation strategies that fail to take into account correlations between covariates and unobserved heterogeneity will not be consistent. These and other issues are discussed in detail in Adams et al. (2003) and related commentaries (Adda, Chandola and Marmot, 2003; Hausman, 2003; and Heckman, 2003, *inter alia*).

An important feature of the estimation approaches discussed thus far is that it is not necessary to fully specify the health production function (7) or preferences (8). An alternative approach specifies functional forms for these relationships and makes stronger assumptions about the stochastic components of the models and then, using dynamic

programming methods, solves the lifetime optimization problem from end of life back to birth, at least in principle. There are at least two advantages to this approach. First, spelling out the full lifetime optimization program clarifies what assumptions are necessary to pin down relationships of interest and imposes greater discipline on the interplay between relationships of interest. Second, it is possible to recover parameters of interest which can then be used for out-of-sample predictions. What would happen if there was a significant increase in health investments or if an epidemic swept through a population? The cost, of course, is that the estimates are only as valid as the assumptions and data that underlie them. See, for example, Eckstein and Wolpin (1989) for a discussion of the approach and McKee (2006) for a recent application to labor supply and health in a developing country context.

2.5.4. Measurement of potential instruments to identify causal effects

The discussion about models, in the previous section, and methods, in this section, highlights the importance of designing studies that include data on characteristics that can serve as instruments that address concerns about unobserved heterogeneity in the models of interest. Prices of health inputs are one set of potential instruments that were noted several times in the previous subsection. Yet, relatively few surveys allocate significant resources to the design and collection of high-quality data on prices of health inputs. There are a large number of potential health inputs that might be important: they include health services, drugs, nutrients, and the epidemiological environment. Moreover, prices should be interpreted broadly to include the availability of inputs (or the time costs associated with obtaining the input) and the quality of the inputs. On the one hand, this further complicates the study design but, on the other hand, it also provides potentially exogenous and possibly important variation that can be exploited in empirical models.

There are potentially substantial payoffs to integrating individual-level survey data with community-level surveys or administrative records. For example, virtually every country collects detailed data on prices in multiple markets; and many countries track the availability and quality of public health services. With the increasing prevalence of data on precise locations of markets, service providers and respondents in surveys (through, for example, GPS measurements), it is in principle straightforward to significantly enrich individual-level survey data by integrating them with community-level information on the local environment (including prices and historical rainfall). This becomes especially important when studies use retrospective measures of health (or health indicators that do not change during prime ages, such as height) as long as the survey data also provides a residential location history (or at least birth place) for each respondent and the residential histories can be linked to administrative or survey records.

3. Meaning of health

The discussion of health thus far has been largely abstract. Defining and measuring health is far from easy. This section takes up these issues from the perspective of empirical research in economics and provides a context for our exploration of the literature in the next sections.

It is widely recognized that health is multi-dimensional and reflects the combination of an array of factors that include physical, mental and social well-being, genotype and phenotype influences as well as expectations and information. A multitude of health indicators have been used in scientific studies of health and economic well-being drawing on data from both the developed and developing world. These have included mortality, reports of morbidities, health-related behaviors, self-assessments of overall health, assessments of physical functioning or activities of daily living, information on specific morbidities and a battery of physical assessments including biomarkers. In order to interpret the evidence on the relationships between health and development, it is critically important to understand what each of the indicators measures.

As noted above, health at a point in time combines the cumulative effects of phenotype factors including an individual's behavior through the life course as well as the health and socio-economic environments to which the individual has been exposed starting *in utero*. Health at any point in the life course also depends on genotype influences and, potentially, interactions between genotype and phenotype factors. Some measures of health can be thought of as stocks; attained height as an adult is an example which provides information about early life experiences. Other measures are flows and vary with high frequency; these include blood pressure or cortisol,[6] which respond to challenges; or the level of glucose in the blood which varies with food intake.

This section begins with a discussion of the relationship between life expectancy and development and then highlights key differences in cause of death between poor and better-off countries. We turn next to individual-level information and discuss the measurement of health of respondents in a socio-economic survey.

3.1. Mortality and life expectancy

Mortality, perhaps the ultimate measure of health, is the most widely-used indicator of the health of a population. At the population level, age-specific mortality rates are often converted to life expectancy. Figure 1 provides a visual summary of the relationship across countries between life expectancy at birth and GDP per capita for three years:

[6] Cortisol is a hormone produced by the adrenal cortex and is released in response to stress. The amount of cortisol present in serum follows a diurnal pattern being highest in the early morning and lowest in the evening after the onset of sleep. In addition, cortisol varies rapidly during the course of the day in response to challenges in the environment. In general, this variation is healthy, but consistently high levels of cortisol reactivity have been shown to be associated with elevated morbidity, reduced cognitive functioning and high risk of mortality. (See Seeman et al., 1997, 2004 for more information.)

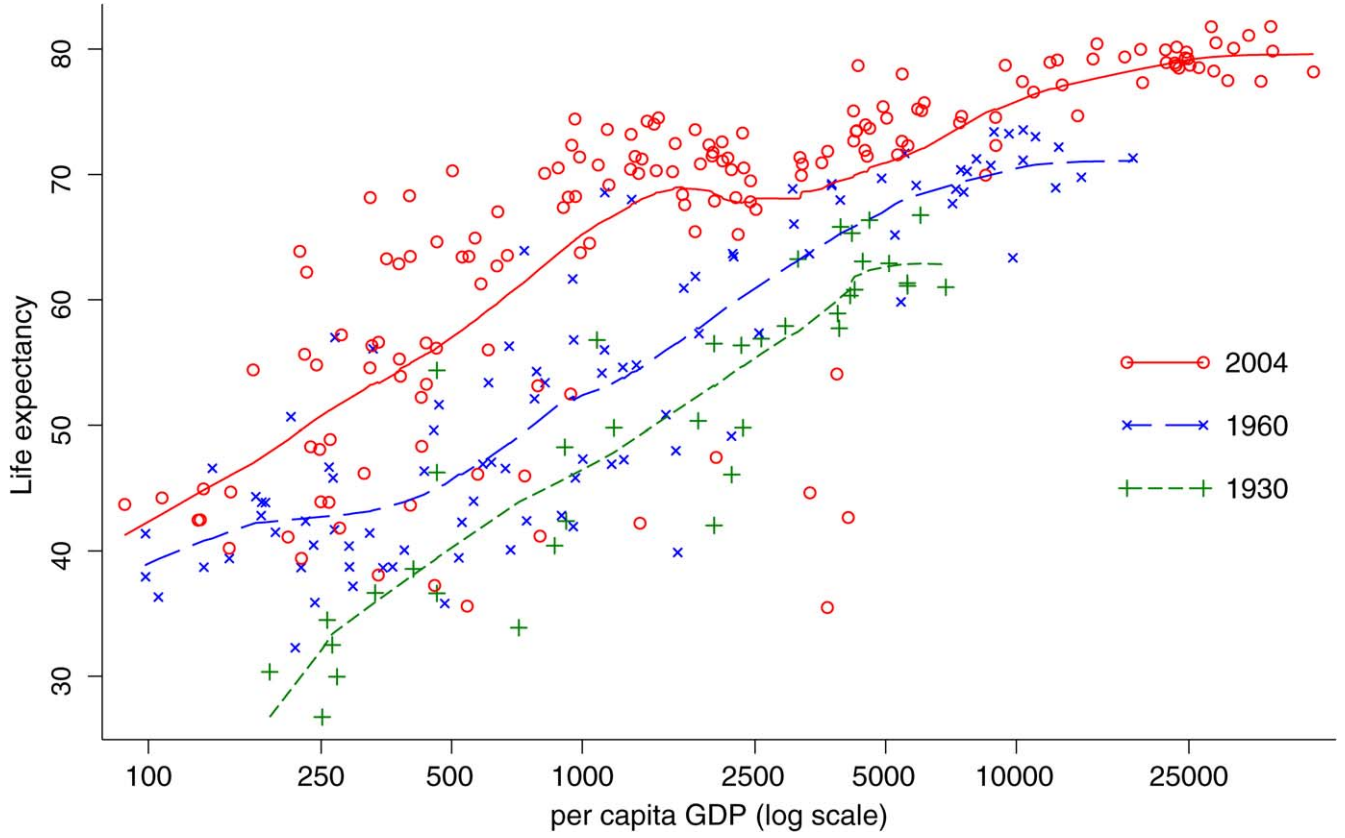

Figure 1. Life expectancy and GDP per capita: 1930, 1960 and 2004.

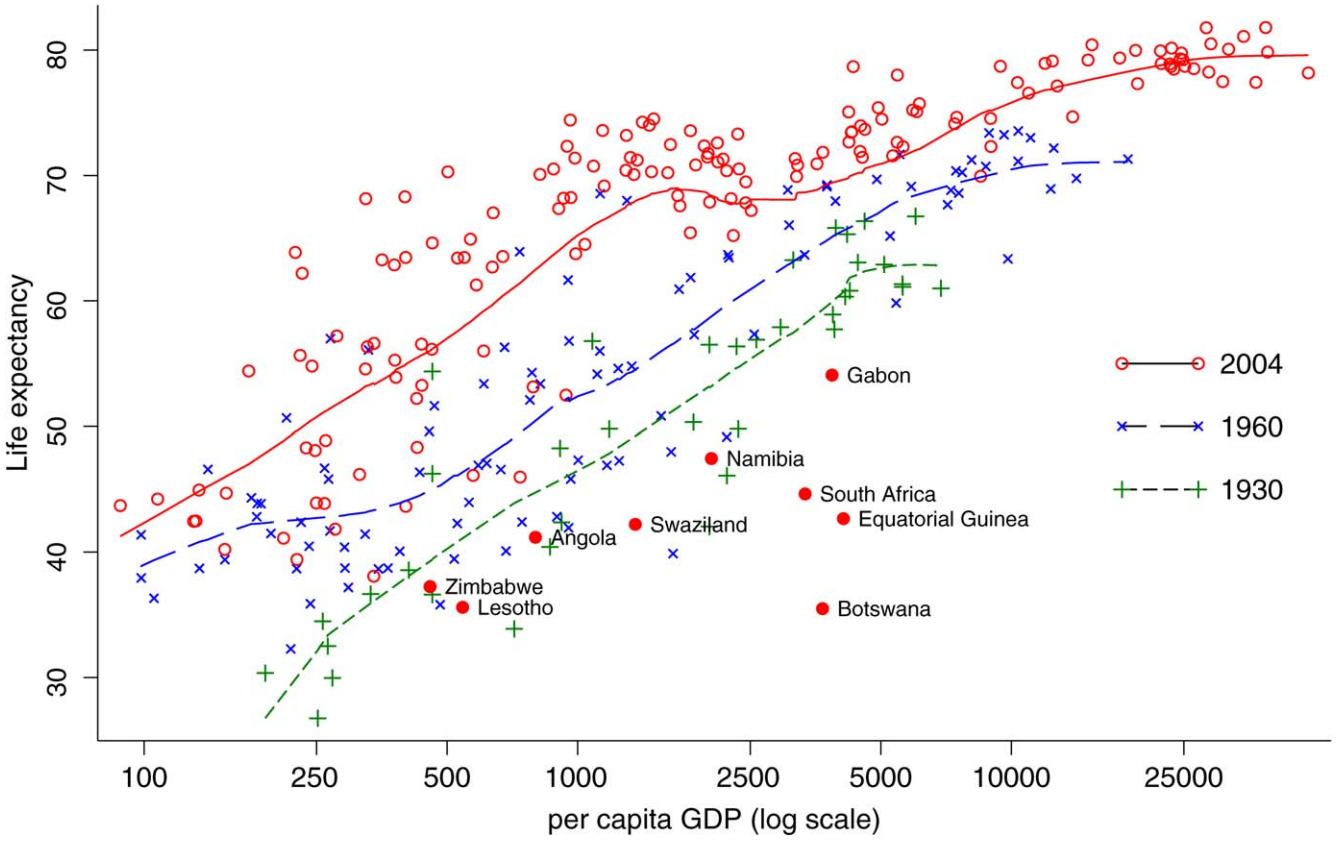

Figure 2. Life expectancy and GDP per capita: Highlighting outliers in 2004.

1930, 1960 and 2004.[7] It is important to recognize that mortality data are often incomplete in lower-income countries and so the estimates of life expectancy are likely to have large standard errors. Nonetheless, the "Preston Curves" (Preston, 1975) make three points very effectively.

First, people in richer countries can expect to live longer than those in poor countries. In 2004, a person born in a low-income country like Mali, where GDP per capita is $250, can expect to live to age 50, whereas someone born in the United States, where GDP per capita is over $35,000, can expect to live until 80. In the last half-century, diminishing marginal returns to income have set in. The Preston Curve remains steepest among the poorest countries, but the income level at which the curve flattens has declined over time.

Second, as indicated by the upward shift of the Preston Curve, there have been dramatic improvements in the health of populations over time.[8] These improvements have been greatest among the poorest. For example, in 2004, the average person born in a country where GDP per capita is about $250 (say, Mali) can expect to live about as long as someone born in a country with over twice that level of GDP in 1960 (say, Thailand) and five times the level of GDP in 1930 (say, Spain).

As Preston points out, if all the improvements in health were driven by aggregate income growth, countries would have merely moved along the Preston Curve over time. The upward shift of the curve, however, suggests that factors other than income are important, and Preston highlights the role of education, improved technology including the availability of vaccinations and oral rehydration therapy, expansion of public health services and nutrition. (See, also, Fogel, 2004 and Cutler, 2004.) Of course, as emphasized in the previous section, the positive association between income and health cannot be assigned a causal interpretation. Indeed part of the growth in income, itself, likely reflects technological progress.

Moreover, evidence suggests that all the benefits of improved technology and health services have not been fully realized by the poorest. For example, effective immunizations against measles have been available for over 40 years and childhood mortality associated with measles has essentially disappeared in richer countries. A WHO/UNICEF initiative, which began in 2000, targeted 45 low-income countries where measles was a leading cause of childhood mortality and provided immunizations. Wolfson et al. (2007)

[7] The data for 1930 are from Preston (1975); the data for 1960 and 2004 are from the World Bank's World Development Indicators (including the ex-Soviet bloc for 1960 and 2000). GDP per capita is measured in 2000 $US using official exchange rates, since purchasing power parity exchange rates are not available for 1930. It is important to underscore that the countries are not the same in each year with a larger sample of countries in the later years. This reflects an increase in the number of countries and improvements in the quality of data over time as more countries have estimates of life expectancy and GDP today than in 1930. The figure includes a non-parametric estimate of the association between life expectancy and per capita GDP for each of the three years. We use a locally weighted smoothed scatterplot estimator (Cleveland, 1979) with a 35% bandwidth for 1960 and 2000 and 70% bandwidth for 1930.

[8] Of course, this effect stretches back well before 1930. See, for example, Fogel (2004).

estimate the program resulted in an over 50% reduction of measles-related child mortality. Differences in the availability of treatments for HIV/AIDS in richer and poor countries is another of many examples

The third point that emerges clearly from Fig. 1 is that in 2004 there are several countries where life expectancy is substantially lower than its predicted level. For nine countries, highlighted in Fig. 2, life expectancy in 2004 is below the level that would have been predicted in 1930 given GDP per capita. Have the technological improvements of the last 70 years passed over these countries?

Six of the nine countries are in Southern Africa and it is HIV/AIDS that has resulted in substantially reduced life spans since the 1990s. Between 20 and 25% of 15–49 year-olds in those countries are thought to be infected with HIV/AIDS. The economies of Angola, Equatorial Guinea and Gabon are dominated by oil, and the benefits of rapid growth in recent years have not reached into population health.

Figure 3 presents time series of real GDP per capita and life expectancy for four countries from 1960 through 2004. Indonesia has seen a four-fold increase in per capita GDP and a 75% increase in life expectancy from around 40 to 70 years. In the United States, GDP has doubled and life expectancy has increased by about 10% to 78. The positive correlation between life expectancy tends to be larger in lower-income settings mimicking the cross-country evidence in Fig. 1.

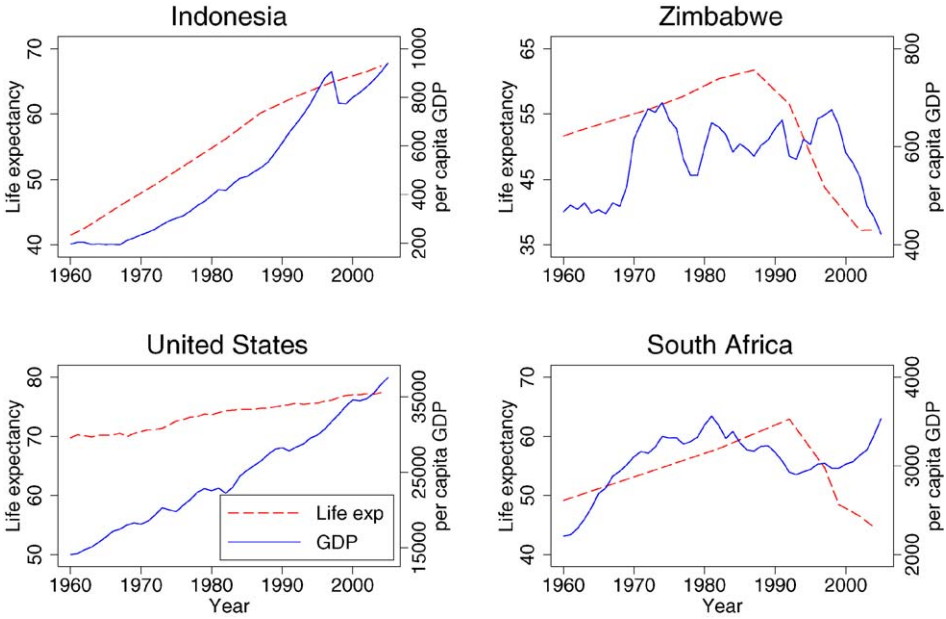

Figure 3. Life expectancy and GDP per capita.

Zimbabwe and South Africa are different: the effect of HIV/AIDS is clear. In Zimbabwe, life expectancy has plummeted from 60 in the late 1980s to 36 in 2004, among the lowest in the world. As in Indonesia and the United States, there is a positive correlation between life expectancy and aggregate income. Moreover, since the decline in life expectancy precedes the collapse of the economy, one might be tempted to conclude that health causes income. That conclusion would be premature.[9] In all the other southern African countries, there is little evidence that HIV/AIDS and GDP per capita are correlated. In South Africa, for example, income has grown substantially since the early 1990s but life expectancy has declined from 65 to 45 during the same period.[10]

Correlations between aggregate income and mortality within countries cannot be interpreted as uncovering causal relationships. Over and above the issues discussed above, it is far from clear that longevity and income should move in lock step even if they are related. Moreover, the literature has emphasized the potentially important role played by early life experiences on later life outcomes. Mortality reflects the cumulation of these experiences which further complicates the interpretation of cross-country and time series associations between life expectancy and current levels of income (or income growth).

Cutler, Deaton and Lleras-Muney (2006) provide an insightful discussion of the evidence on global mortality in considerably more depth. They conjecture that the preponderance of evidence suggests that it is science and technological innovation, including improved public health services, that have been the driving forces behind increased longevity of the world's population over the last century. Given the current state of knowledge, as they note themselves, their conclusions are tentative since key causal relationships underlying their argument have not been established.

3.2. Global burden of disease

Mortality is a very useful indicator of overall population health. However, health is far more than life or death and, as the southern African countries demonstrate, it is important also to understand the causes of death. That turns out to be complicated because the quality of data on cause of death varies dramatically across time and across countries and because in many cases only one cause of death is recorded, masking a far more complex reality. With these caveats in mind, Table 1, which is from Mathers, Lopez and Murray (2006), provides an overview of the global burden of disease.

Whereas, in the early twentieth century, infectious diseases accounted for a large fraction of deaths across the world, today they have essentially been eradicated in high-income countries. Nevertheless, they still account for almost one quarter of deaths in

[9] The collapse of the Zimbabwean economy can be traced to factors that are unrelated to the incidence of HIV/AIDS although arguably the severity of the health crisis is related to declining income in a complex way.
[10] Relying on the sort of cross-country evidence in Figs. 1 and 3, Pritchett and Summers (1996) conclude that it is income (or wealth) that drives health. That conclusion also seems premature.

Table 1
Causes of death in low/middle and high income countries (2001)

	1. Low and middle income countries		2. High income countries		3. All countries	
	# deaths (thousands)	% deaths	# deaths (thousands)	% deaths	# deaths (thousands)	% deaths
I. Communicable diseases, pregnancy outcomes, nutritional deficiencies	**17,622**	**36.4**	**552**	**7.0**	**18,174**	**32.3**
A Infectious and parasitic diseases	10,692	22.1	152	1.9	10,844	19.3
HIV/AIDS	*2554*	*5.3*	*22*	*0.3*	*2576*	*4.6*
Diarrheal diseases	*1778*	*3.7*	*6*	*0.1*	*1784*	*3.2*
Tuberculosis	*1590*	*3.3*	*16*	*0.2*	*1606*	*2.9*
Childhood diseases	*1363*	*2.8*	*2*	*0.0*	*1365*	*2.4*
Malaria	*1208*	*2.5*	*0*	*0.0*	*1208*	*2.1*
B Respiratory infections	3483	7.2	349	4.4	3832	6.8
C Perinatal conditions	2951	6.1	32	0.4	2983	5.4
D Nutritional deficiencies	451	0.9	18	0.2	469	0.8
II. Noncommunicable conditions	**26,037**	**53.8**	**6,868**	**86.5**	**32,905**	**58.5**
E Cardiovascular disease	13,362	27.6	3, 039	38.3	16,401	29.1
F Malignant neoplasms	4957	10.2	2, 066	26.0	7023	12.5
G Respiratory diseases	3127	6.5	477	6.0	3604	6.4
H Digestive diseases	1602	3.3	335	4.2	1937	3.4
I Diabetes mellitus	758	1.6	202	2.6	960	1.7
J Neuropsychiatric disorders	701	1.4	378	4.8	1079	1.9
K Congenital anomalies	477	1.0	30	0.4	507	0.9
III. Injuries	**4717**	**9.8**	**471**	**5.9**	**5188**	**9.2**
L Unintentional	3216	6.6	321	4.0	3537	6.3
M Intentional	1501	3.1	151	1.9	1652	2.9

Source: Mathers, Lopez and Murray (2006, Table 6.4).

low- and middle-income countries. In 2001, HIV/AIDS and tuberculosis, which often go hand in hand, accounted for over 8% of deaths (nearly 5 million people); diarrheal diseases and childhood diseases account for 6% of deaths in these countries. Many of these deaths would be prevented if the people benefited from the health services and technologies that are available in richer countries. That they have not reflects the combined effect of limited resources at the individual and family level, information, behavioral choices, and the health environment including lack of clean water, poor sanitation, nutrition and medical care as well as limited access to drugs.

Perinatal conditions account for another 5% of deaths in low- and middle-income countries but less than $\frac{1}{2}$% of deaths in high-income countries. The latter reflects spectacular technological innovations in the care of newborns and infants in richer countries.

The literature suggests that in many low-income countries, improved prenatal care, other health inputs and behavioral change during pregnancy are likely to significantly reduce infant mortality.

The vast majority of deaths in high-income countries are attributed to non-communicable diseases with cardiovascular disease and malignant neoplasms (cancers) accounting for nearly two thirds of all deaths. Many of the low- and middle-income countries are now passing through the epidemiological transition as infectious diseases diminish and non-communicable diseases emerge as major causes of death.[11] It is estimated that about 18 million people in low- and middle-income countries died from heart disease or cancers (over one third of all deaths). This number is over four times larger than the deaths from HIV/AIDS and tuberculosis combined. To be sure, heart disease and cancers are not only prevalent throughout the globe but will increase in prevalence as infectious diseases become better controlled.

Just under 5 million people (almost 1 in 10 deaths) died in 2001 in developing countries because of injuries, which is more people than died from HIV/AIDS or tuberculosis. The causes include traffic accidents, violence and abuse as well as injuries from natural events.

Interventions in high-income countries have demonstrated that deaths from traffic accidents can be avoided through enforcement of vehicle safety checks, driver education, seat-belt usage and speed control. In middle- and low-income countries, it is prime-age adults who are most likely to die in a traffic accident (World Health Organization, 2004). Bertrand et al. (2006) estimate that in a study of Indian driver license applicants (average age = 25), about two-thirds of those who obtained a license could not pass a basic driving test.

The December 2004 Indian Ocean tsunami killed approximately 250,000 people and the 2005 Pakistan earthquake killed nearly 75,000 people. Natural events tend to take a much larger toll in lower-income settings where the infrastructure is poorly suited to withstand the force of nature. In high-income countries, building codes, restrictions on development and early warning systems have proved effective mechanisms to minimize death and injury in these events.

Cause-of-death information provides insights only into the burden of disease at the end of life. In an effort to provide a richer summary measure of the health of a population at the aggregate level, estimates of healthy life expectancy (HALE) have been developed which include an adjustment for time in poor health. See Lopez et al. (2006) for a comprehensive discussion of these methods and description of changes in the global burden of disease in the last decade. We follow a complementary approach and focus on measurement of health of individuals in population surveys.

Before turning to that discussion, two points are worth emphasizing. First, the technology exists to avoid a substantial fraction of deaths in low- and middle-income countries. Identifying the impediments to bringing technologies and health services to these

[11] See Omran (1971) who was the first to describe this process.

populations is likely to have a high pay-off in terms of improving their well-being and contributing to scientific understanding of the determinants of population health.

Second, underlying the aggregate mortality estimates in Fig. 1 is considerable heterogeneity in the nature of the disease burden. As childhood diseases decline, life expectancy will increase and the burden of disease will shift towards more non-communicable diseases, many of which involve expensive treatments that can extend life for many years. The demographic transition with declining fertility and prolonged life spans will result in older people constituting a larger fraction of the population. The combination of aging populations and shifts in the burden of disease will have profound effects on the entire economy including the public and private health sectors, old age support and the distribution of resources between generations.

3.3. Physical health assessments: Anthropometry

Anthropometric measures, particularly height and weight, have been the workhorse in research on health in both the history and development literatures. Studies have also used such measurements as the circumference of the head, arm, waist and hip, among others, and skin-fold thickness in various body locations. They are all easy and inexpensive to measure in a field setting, can be measured accurately by a trained person and provide a wealth of information that can be compared across populations and over time.

Height (or length of babies and infants) reflects the combination of both genotype influences and phenotype influences *in utero* and during the first few years of life (Martorell and Habicht, 1986). Because attained height is fixed for adults (until shrinkage in older ages), it has provided important insights into differences in the early life health across cohorts. Moreover, height is positively correlated with cognitive achievement and schooling outcomes and predicts economic productivity and mortality, among others. These associations tend to be especially strong in lower-income settings.

Weight, a more contemporaneous indicator of general health and nutrition, reflects the combined effects of energy intake (food and diet) and energy output (physical activity). Weight is most easily interpreted when combined with height. Among young children, weight conditional on height along with growth in height are key markers for nutritional status and are typically compared with standard growth tables (see, for example, de Onis et al., 2006; CDC, 2000).

Body mass index (BMI), which is weight (in kg) divided by height (in m) squared, has proved to be a very convenient summary, particularly for adults. Extreme values of BMI (underweight BMI < 18.5 and overweight BMI > 25) have been shown to be associated with elevated morbidity and mortality (Waaler, 1984; Fogel, 2004). An overweight adult, and particularly an obese adult (BMI > 30), is at elevated risk of *inter alia* heart disease, dislipidemia, type II diabetes, stroke and some types of cancers. Many of these are the non-communicable diseases that are prevalent among older adults in high-income countries. This suggests that the epidemiological transition is likely to be accompanied by a nutrition transition (Popkin, 1994, 2003).

Figure 4 and Table 2 display the distribution of BMI for adult males and females (age 22 through 75) using survey data from six countries.[12] There are stark differences between the three poorer countries in the left panel of the figure and the three richer countries in the right panel: in general, the distribution of BMI tends to shift to the right as development proceeds.

In Bangladesh, the poorest of the six countries, over half the adult population is underweight and less than 5% are overweight. Moving up the GDP distribution to China, only about 10% of the population is underweight whereas about 15% is overweight. The distribution in Indonesia is similar although a smaller fraction of the population is in the healthy ranges of BMI. Obese people make up a very small fraction of the population in the three poorest countries.

In all six countries, women are more likely to be overweight and obese than men. These differences are nowhere greater than in South Africa. Whereas about 10% of South African males are underweight and 10% are obese; very few women are underweight and over one third of the women are obese. Continuing up the GDP ladder, in Mexico there is very little undernutrition but almost three quarters of the population is overweight and nearly a third are obese. The Mexican and United States distributions are very similar although GDP is about 5 times higher in the United States. The figures suggest that obesity rates rise with economic development which is troubling given the relationship between obesity and cardiovascular disease, stroke, diabetes and possibly cancer. The figures also suggest that growth in obesity is driven by more than income growth alone. Information and technology likely play a key role with physical activity at work and in leisure, food intake and the composition of the diet are all candidate proximate contributors to increased rates of obesity.

Figure 4 and Table 2 indicate that BMI rises with aggregate income. Does BMI increase with SES at the individual level? To answer this question, Fig. 5 uses individual-level data from the same six surveys and displays the relationship between BMI and education for males and females controlling age, in a regression framework.[13] The countries are ordered according to GDP.

In the five developing countries, BMI of males is positively correlated with education at all levels of education but in the United States BMI and education are not correlated among male high school dropouts and negatively correlated for better-educated males. Among females, there tends to be a positive correlation between BMI and education at

[12] Data are from Bangladesh, Indonesia, China, South Africa, Mexico and the United States.The South African, Mexican and United States surveys are nationally representative; the Indonesian survey is representative of about 80% of the Indonesian population; the Chinese survey is representative of 9 provinces and the Bangladesh survey is representative of one district. Figure 4 presents Epanechnikov kernel density estimates for males and females separately. The bandwidth is 0.7.

[13] The shapes in the figures are essentially identical if we relate the probability of being overweight to education. Education is specified as a spline with knots at 5 and 10 years of education; each regression includes splines for age (with knots at 10-year intervals). Separate regressions are estimated for males and females for each survey. Adults age 22 through 70 are included in the regressions.

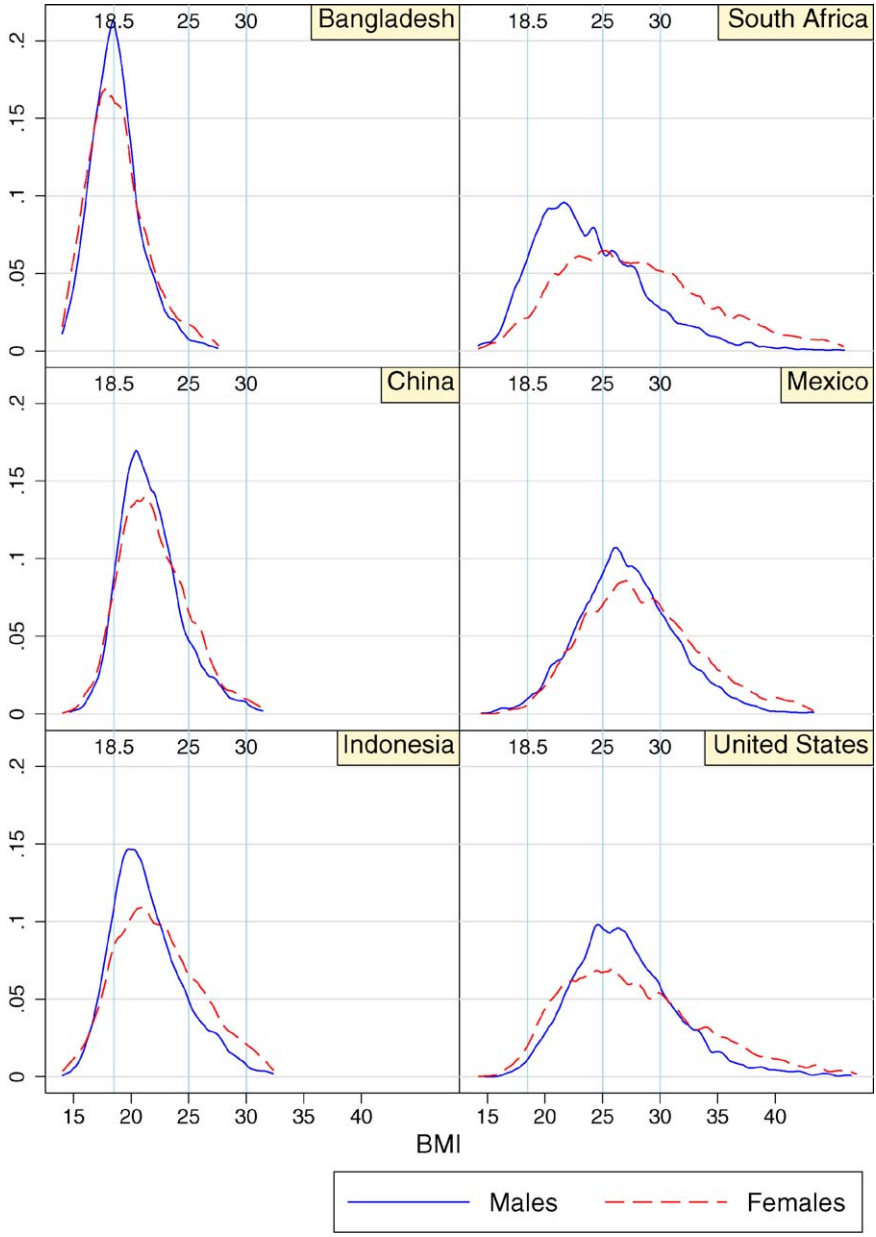

Figure 4. Distribution of BMI of males and females.

Table 2
Distribution of BMI in 6 countries

Indicator	Gender	Bangladesh	China	Indonesia	South Africa	Mexico	United States
All ages (25–70 yrs)							
BMI (Mean)	Male	18.7	21.8	21.5	24.0	27.2	27.1
	Female	18.7	22.2	22.5	28.2	28.7	28.1
% Underweight	Male	50.4	8.7	15.9	10.1	1.4	1.1
[BMI $\leqslant$ 18.5]	Female	52.7	10.3	14.7	4.7	0.9	2.4
% Overweight	Male	1.7	11.9	13.9	36.0	69.5	63.3
[BMI $>$ 25]	Female	3.0	17.5	25.1	63.6	74.1	62.7
% Obese	Male	0.1	1.2	1.3	11.4	23.8	22.5
[BMI $>$ 30]	Female	0.3	2.5	4.8	35.1	36.0	32.9
Prime age (25–44 yrs)							
% Underweight	Male	43.5	6.4	13.6	10.0	1.3	1.1
[BMI $\leqslant$ 18.5]	Female	45.2	8.3	10.5	5.0	0.8	2.7
% Overweight	Male	1.8	9.5	13.5	31.6	68.7	59.3
[BMI $>$ 25]	Female	3.3	14.3	25.2	59.9	70.9	57.0
Older age (45–70 yrs)							
% Underweight	Male	57.3	12.0	19.7	10.2	1.6	1.2
[BMI $\leqslant$ 18.5]	Female	62.5	13.3	21.2	4.2	0.9	2.1
% Overweight	Male	1.6	15.2	14.8	42.1	70.6	67.8
[BMI $>$ 25]	Female	2.6	22.3	25.0	68.1	79.4	69.9
Sample size	Male	3449	3257	7815	3547	4863	5636
	Female	4222	3457	8423	5242	6662	6431
Survey		MHSS	CHNS	IFLS	SADHS	MxFLS	NHANES III
Survey year		1996	1991	2000	1998	2002	1988–1994
GDP per capita ($US 2000)		310	421	800	2974	5852	27,833

Notes. MHSS is the 1996 Matlab Health and Socioeconomic Survey (Rahman et al., 1999). CHNS is the 1991 wave of the China Health and Nutrition Survey (Popkin, 1993). IFLS is the 2000 wave of the Indonesia Family Life Survey (Frankenberg and Karoly, 1995; Frankenberg and Thomas, 2000; Strauss et al., 2004). SADHS is the 1998 South African Demographic Health Survey (Demographic and Health Surveys, 2002). MxFLS is the 2002 wave of the Mexican Family Life Survey (Rubalcava and Teruel, 2004). NHANES III is the National Health and Nutrition Examination Survey Wave III (National Center for Health Statistics, 1994).

the bottom of the education distribution which turns negative at higher levels of schooling.

Recall that, on average, females have higher BMI than males. This is true at lower levels of education, but at the top of the education distribution in China, Mexico and the United States, males have higher BMI than females. Similarly, in South Africa the gap in BMI between males and females is very small among the best educated relative to those with no education.

Behind Figs. 3 and 4 there is substantial heterogeneity across age groups, some of which can be attributed to cohort differences. The lower panels in Table 2 display the fraction of prime-age adults and older adults who are underweight and overweight. In

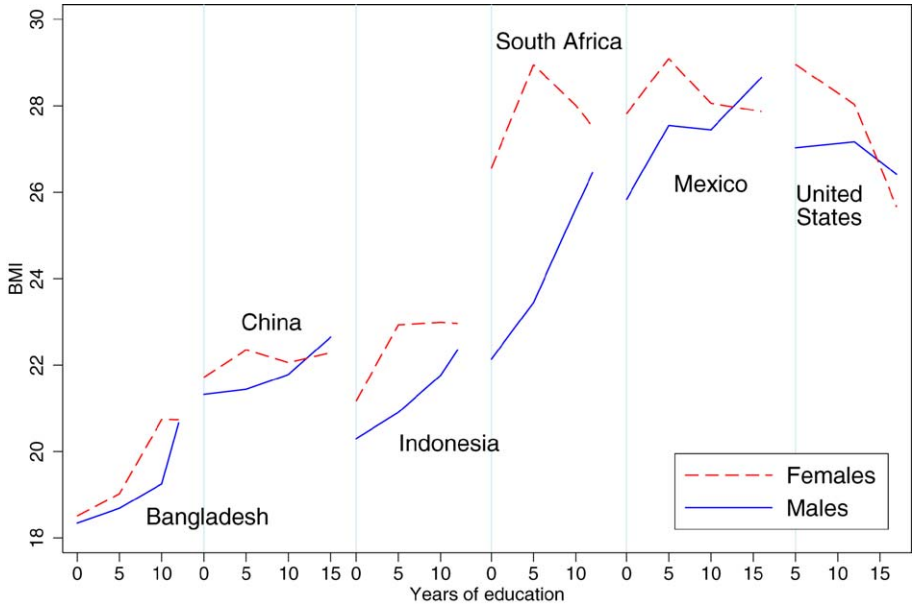

Figure 5. Relationship between BMI and education.

the three poorest countries, older adults are more likely to be underweight than prime-age adults indicating that in terms of improved nutrition, younger adults benefit more from economic growth. In contrast, in the richer countries, older adults are more likely to be overweight than prime-age adults suggesting that as populations move through the nutrition transition, it is the prime-age adults who are more responsive to the negative effects of being overweight. Notice also that in China (and among prime-age Indonesians) twice as many females are overweight as underweight suggesting rapid change in the nutrition profile is underway in these countries.

In fact, the evidence on differences across cohorts indicates that the rate of increase in obesity in many developing countries is far faster than it was in the more-developed countries. See, for example, Popkin (1993) and Monteiro et al. (2004) who discuss differences in BMI by SES, their causes and implications. The inverted U shape in Fig. 5 is consistent with evidence in the literature indicating that as a population moves through the nutrition transition, it is the most educated (and highest income) who are the first to exit under-nutrition. They are also the first to adjust their diet and physical activity to avoid the deleterious effects of being overweight, suggesting that behavioral changes have important impacts on health outcomes. Figure 5 also indicates that it is women who tend to lead this transition. The relative importance of information, resources, technology and other factors in these processes has not been established.

BMI cannot distinguish between lean tissue weight and weight due to body fat. Moreover, the location of fat on the body is predictive of elevated morbidity. For example,

elevated levels of stress tend to result in body fat being deposited around the abdomen, elevated waist to hip circumference is a marker for elevated stress. However, higher waist to hip ratios can be caused by several other factors. More precise measures of fat and its distribution on the body are not as easy to assess although new technologies promise simple, non-invasive and rapid measurement of body composition in the future. At this point in time, a potential disadvantage of relying exclusively on anthropometric measures is that they provide information about general health and nutrition but little insight into the pathways through which health affects (or is affected by) life outcomes.

3.4. Physical health assessments and biomarkers

Several recent studies have attempted to measure more specific domains of health status through a wide array of physical assessments. These include simple physical activities that can be completed in a short period of time such as walking a specific distance, balance (standing on one leg for a specific time) or timed repeated stands from a sitting position. These assessments are easy to administer and can provide especially useful information when combined with self-assessed activities of daily living (ADLs).

Recent innovations in the development of non-invasive biomarker measures that can be implemented on a large scale in a field setting have had a major impact on the extent and quality of health measurement in socio-economic surveys. In the last decade, a large number of socio-economic surveys have included the collection of several biomarkers such as blood pressure, lung capacity, measures of strength, as well as measures of sight, hearing and oral health.

Some important indicators of health status can only be measured in biological samples. Studies have collected blood, cheek swabs, hair samples and urine which provide a host of genetic and non-genetic information about each respondent. In some cases, the biomarkers are measured in the home and are very straightforward to implement. For example, blood from a pin prick to the finger is sufficient to measure hemoglobin, glucose, a lipid profile, albumin and HIV status in a few minutes in the home. (Measurement of glucose and triglycerides requires fasting which complicates administration of the assessment.)

Technological innovation in health measurement promises to have a significant impact on empirical research in the area of health and development with less expensive, simpler and more robust methods for the collection, storage and assaying of biological materials being introduced at a rapid rate. One especially promising set of innovations involves the development of technology for the collection and analysis of dried blood spots (DBS). Blood from a finger prick is spotted on filter paper which is dried and stored for later analysis. It is typically feasible to collect several spots from a single finger prick and each spot can be used to measure a different marker. Protocols for over one hundred analytes have been validated for DBS (McDade, Williams and Snodgrass, 2006). These include markers for stress (such as antibodies to Epstein–Barr virus), inflammation (such as C-Reactive Protein, CRP), micronutrients (such as Transferrin Receptors, TfR, an indicator of iron stores), metabolic functioning (such as glycosylated

hemoglobin, Hb_{A1c}, which has been implicated in type II diabetes) and reproductive health. Innovations in measurement of biomarkers using saliva, cheek swabs, hair and other minimally invasive methods that are easy to administer in a field setting are also likely to have a major impact on research in this area. (See Lindau and McDade, 2007, for an excellent summary of the state of the art.)

The inclusion of a broad array of biomarkers in population-based studies will not only provide a stronger biological foundation for research on health but also explore new hypotheses. For example, the metabolic syndrome highlights the cumulative effects of obesity, high blood pressure, high cholesterol and hyperglycemia on the risk of diabetes and heart disease (Haller, 1997). McEwen (1998) and his collaborators have highlighted the cumulative effects on health of multiple stress events, each of which triggers a biological response. Allostatic load is the wear and tear on the body due to the over-stimulation of these responses which either fail to shut down or fail to respond. Developing countries are ideal contexts for more fully exploring these constructs.

While innovations in health measurement are very exciting and promise to revolutionize the nature and quality of data on health that is collected in large-scale, population-based socio-economic surveys, to a large extent the benefits remain a promise. To date, only a relatively small number of studies have made effective use of biomarkers to provide new insights into the relationships between health over the life course and other indicators of well-being.

3.5. Self-assessed health

Many of the early studies of health in the social sciences relied on self-assessments of health and health behaviors. While these assessments provide important insights into the distribution of health and characteristics associated with better health, self-assessments do not only reflect intrinsic health. In many cases, these indicators also reflect perceptions of health and information about health, both of which are likely correlated with education, income and use of health services. This substantially complicates interpretation of the indicators.

The RAND Health Insurance Experiment (HIE) provides an example for General Health Status (GHS), a widely-used overall indicator of health which is very easy to administer in a survey and which has been shown to be predictive of future mortality (Ware and Sherbourne, 1992). In the RAND HIE, subjects were randomly assigned to receive free health care or to one of several programs that varied the price of care. Health care use rose the most among those who received free care and, among those who were in poor physical health at the beginning of the study, their physical health improved (Newhouse, 1996). However, the subjects who received free care reported their general health to be worse at the end of the study relative to other subjects, suggesting that GHS is influenced by exposure to health service providers who provide information about health problems which, in turn, affects the subject's own evaluation of his or her health. Similar results are reported for a health price experiment in Indonesia (Dow et al., 2004). This provides one explanation for why GHS tends to be lower for

higher income respondents in many surveys from developing countries although mortality and physical health measures suggest that they are in better health than those with less income.

Similar issues arise with all self-reported health indicators to a greater or lesser extent. Many surveys include questions about whether the respondent suffered from one of several morbidities such as diarrhea, nausea, dizziness, fatigue, runny nose, flu, respiratory problems and so on. Not only is it likely that the meaning of "suffer from" varies across SES but the definition of a "runny nose" may well vary depending on what is "expected" health. Some studies ask about diagnosed conditions such as elevated cholesterol or blood pressure, and those reports will provide a distribution of population health that is selected on use of health services.

It has been suggested that ADLs, such as walking a specific distance or carrying a specific load for a specific distance, are less prone to these sorts of biases. It is unlikely, however, that these indicators are not also influenced by a respondent's perception of "good health" or "difficulty" conducting an activity. Moreover, in some cases, the activity might be outside the domain of the person's experience. A standard ADL is difficulty walking up stairs which is unlikely to be relevant in many rural settings. Since ADLs were originally developed to assess health of older adults, many of the standard questions such as difficulty getting up from the floor are not likely to provide much information about the health of prime-age adults.

Recent studies have attempted to develop vignettes which provide a mechanism to standardize each respondent's perception and information about "good health" (King et al., 2004). The value of the vignette approach remains an open question.

Some domains of health can only be assessed with self-reports. Psycho-social health is a good example. It has proved to be difficult to develop and validate instruments that can be fielded across cultural and social contexts. The WHO World Mental Health Surveys, which fielded an adaption of the Composite International Diagnostic Interview (CIDI) in 14 countries, provide an important baseline (WHO Mental Health Survey Consortium, 2004). There would seem to be substantial scope for important contributions to methods and measurement in this area in developing countries where poverty, uncertainty, death and disease likely take their toll.

3.6. Measurement error in health

The fact that health is multidimensional complicates interpretation of specific health markers as covariates in models if those markers are correlated with other health indicators that are not included in the model. In these cases, the health indicator included in the model will be correlated with unobserved indicators not in the model and the estimated "effect" will be biased. Since the included health measure proxies for other dimensions of health, it is not straightforward to interpret estimated "effects."

Health is also likely to be subject to random measurement error. For some physical assessments, such as height, error is attributable to poor measurement by enumerators and can be minimized with good training. Other sources of error affect physical assess-

ments such as weight and blood pressure which vary over the course of the day. More of a concern are errors that depend on respondent performance such as timed chair stands; if more energetic and enthusiastic respondents complete the task more quickly, it is important to recognize these physical assessments capture more than, say, mobility and lower body strength.

Self-assessed indicators of health raise special concerns. In both developing and developed country contexts, interpretation of evidence that relies on self-reports will be complicated to the extent that those indicators reflect not only true health status but other influences that are systematically related to outcomes of interest in specific models, or to characteristics that are included in those models. For example, if individuals who earn higher wages are more likely to use health care services and, because of that use of care, tend to report their health as being worse than it would be if they did not use care, then it will be very difficult to interpret the relationship between wages and self-assessed health status. One approach to addressing this issue is to use instruments for self-reported health status. Given the non-classical nature of this type of measurement error, in general, it will be difficult to identify plausible instruments. This concern applies to all the self-assessed health indicators discussed above and needs to be born in mind when interpreting evidence in the literature.

4. Empirical evidence on health over the life course

We turn to a review of some of the key contributions in the literature, beginning with a body of new evidence that suggests that health in early life may have substantial effects on health and well-being throughout the entire life course. The origins of this work can be traced back to seminal work by Fogel and others in economic history and development as well as innovative lines of research in nutrition, epidemiology and the biomedical sciences. This is a rapidly developing line of inquiry and a very promising avenue for new research that brings together the best practices of these disciplines. It is also important from the perspective of the dynamic models. We begin with evidence on the impact of health *in utero* on mortality and later life health and then turn to economic outcomes in adulthood. This is followed by a brief review of the links between early life health and outcomes during adolescence. This is an area in which there is a substantial literature with contributions from many disciplines and is the subject of one of the chapters in this volume (Glewwe and Miguel, 2008).

We next examine health and socio-economic success in adulthood both from the perspective of the impact of health and health investments on economic productivity and the impact of income and resources on health. The dynamics underlying these relationships naturally leads to a discussion of the role that health may play in smoothing consumption which relates to a much broader literature in micro- and macro-economics. We end with a discussion of HIV/AIDS, arguably the most pressing public health crisis of our time.

4.1. Long-run impacts of early life health on later life outcomes

In recent years, a significant body of evidence has accumulated that establishes an association between health *in utero* or in very early childhood and later life outcomes including health status. Godfrey and Barker (2000) review the industrial country literature. Alderman and Behrman (2006) review studies of low birthweight impacts on later life outcomes in developing countries and provide estimates of the likely welfare costs of low-birthweight babies.

A leading explanation for the relationship between early and later life health has been that poor nutrition of the mother during pregnancy, or of the child during very early childhood, may induce adaptations in organ function or organ size, metabolism or cause gene expression to adapt to a new environment in order to raise survival probabilities through the early years but cause problems later in life (Barker et al., 1989; Fogel, 2004; Komlos, 1994).

Several biological mechanisms have been suggested for this relationship. Thinness at birth may result in slower than normal child growth as an adaptation to reduced nutrients. In the long run, it may lead to negative consequences such as a higher probability of elevated blood pressure, glucose intolerance, a poor lipid profile and elevated risk of type 2 diabetes and coronary heart disease. A corollary to the Barker hypothesis is the notion that a "thrifty phenotype" will develop as an adaption to adverse nutrition *in utero* which results in permanent metabolic and endocrine changes. These changes would be beneficial if food availability continued to be scarce after birth. However, if food is plentiful, the same metabolic changes are disadvantageous and they are associated with elevated risks of obesity, dyslipidemia and glucose intolerance in later life (Hales and Barker, 1992). Thus, in this model, it is those who are born small and thin but become overweight in adulthood that are at greatest risk of heart disease and type 2 diabetes in later life. Quantifying the extent of these effects seems crucial given the rapid rates of growth in obesity, heart disease and type 2 diabetes across the globe (Robinson, 2001).

The biomedical literature suggests that the timing of *in utero* nutrition insults matters. As an example, poor maternal nutrition during the period when arteries form in the fetus have been associated with elevated risk of hardened arteries when the individual reaches mid-life. Thus, it is not just those babies who are very small at birth that are at risk of suffering longer-term consequences of early life deprivation. Even normal birthweight babies are at risk and the extent of these effects depend on subsequent nutrition and other health inputs over the life course (Godfrey and Barker, 2000).

A second, possibly complementary, hypothesis is that inflammation caused by infections, even in very young children, can have deleterious long-run health consequences. The argument is that exposure to infectious disease results in inflammations that ultimately lead to the development of atherosclerosis, which is a major source of cardiovascular disease (Finch and Crimmins, 2004). Studies indicate that high levels of inflammations can promote atherogenesis, the process of forming plaques in the inner

linings of arteries, even in the absence of high-fat diets. This process begins *in utero* and continues after birth, throughout life.

Experimental studies with rat models have provided support for both the nutrition and inflammation hypotheses (Godfrey and Barker, 2000; Finch and Crimmins, 2004). Although the nutrition literature is replete with randomized interventions that provide an array of supplements to pregnant women, the impact of those interventions beyond birth outcomes has been largely ignored. Long-term follow-up of those children have the potential to provide important new evidence. To date, therefore, most research on longer run effects of fetal health in humans has been limited to non-experimental designs. Elo and Preston (1992) provide a comprehensive and insightful review of the early literature.

4.1.1. Quasi-experimental studies on fetal health and health in later life

Perhaps some of the best empirical evidence comes from quasi-experimental studies. An important set of research exploits the fact that there was food rationing for 5 to 6 months during the winter of 1944–1945 in some Dutch cities and not in others (Stein et al., 1975). By using a difference-in-difference procedure, comparing birth cohorts exposed to the famine with cohorts not exposed and comparing those born in cities with rationing relative to those born in cities with no rationing, Stein et al. establish that exposure to the famine raised infant mortality, particularly in the first few months after birth, although there was little impact on mortality through age 18, conditional on surviving infancy.[14] Longer-term follow-up studies have mostly examined health at age 50, comparing cohorts exposed to the famine with cohorts not exposed (see Roseboom et al., 2001a, 2001b, for summaries). Mortality at age 50 is apparently unaffected by the famine, conditional on surviving infancy. However, exposure to the famine is associated with elevated risk of glucose intolerance among men and women, worse lipid profiles and elevated risk of coronary heart disease. Women exposed during the first trimester as a fetus tended to have higher BMI and waist circumference but not if the fetus was only exposed later in the pregnancy; there was no evidence of an impact on male body mass. Nor is there evidence that blood pressure is related to exposure to the famine (Ravelli et al., 1998, 1999).[15]

Luo, Mu and Zhang (2006) follow a similar approach and examine the long-run impact of the Great Chinese famine of 1959–1962 on body mass of men and women. The Chinese famine killed between 20 and 30 million people – far more than the Dutch Winter Famine – with deaths being concentrated in rural areas and distributed very unevenly across provinces. Using data from the China Health and Nutrition Survey (CHNS) Luo

[14] They also report no differences in an IQ-type test at age 18.

[15] An advantage of the study is that it is possible to draw on administrative data to identify those who have died and those who have emigrated. Although there are no detailed analyses of attrition in the literature, of 2400 individuals in the famine birth cohorts from Amsterdam, 90% were followed up when they were 50. One hundred and ninety-nine had emigrated and most of the rest had died.

et al. examine the BMI in adulthood comparing those born during the famine or soon afterwards and also drawing contrasts across regions of birth in a difference-in-difference framework. They find women born during the famine are more likely to be overweight 30–40 years later; they find no differences for men. Of course it is possible that other changes across time and across regions explain these differences. Perhaps more problematic in this case might be mortality selection. If more frail babies died during the famine, or were not brought to term, they might have been in worse health later in life, which would bias even a difference-in-difference estimator against finding significant effects.

Other studies are limited to comparing birth cohorts. While this has the disadvantage that all time-varying differences attributed to the cohort might in fact reflect confounding time or period effects, the long time series exploited in several important studies mitigates these concerns.

Crimmins and Finch (2006) estimate the impact of infant mortality on older-age mortality using aggregate data by birth cohort on age-specific mortality from different European countries in the 19th century. Specifically they regress mortality from age 70 to 74 years, $q_{70–74}$, on infant and child mortality rates for that same birth cohort.[16] They argue that infant and childhood mortality in 19th century Europe was a result largely of infectious disease, for which they have supporting evidence from death by cause data. Thus high infant and child mortality rates would be a proxy for high levels of infectious disease and thus inflammation. On the other hand, older-age mortality even in the 19th century was largely from chronic disease, of which heart disease was a major part. Again, they show some evidence in support of this claim. Of course, they have to assume that the only cohort-specific effect derives from different levels of childhood infection, and that period or time effects, possibly correlated with birth cohort effects, are not contaminating the results. Both are strong assumptions. Since they have no direct measures of infection or inflammation, this evidence is indirect.

In a similar study, Bengtsson and Lindstrom (2000) use vital records maintained by churches in four parishes in southern Sweden for the period 1760 to 1894. Using a Cox proportional hazard model they estimate how the probability a person dies, conditional on surviving to the prior period, varies with infant mortality in the year and area of birth as covariates, food prices and socio-economic background of the individual, controlling for cohort effects with a time trend. They find that both birth cohort and region-specific infant mortality are positively and significantly related to older-age mortality, similar to the findings of Crimmins and Finch (2006). However food prices at birth and child (0–5 years) mortality rates are not significantly related to subsequent mortality.

[16] They use $q_{0–1}$, $q_{1–4}$, $q_{5–9}$ and $q_{10–14}$ simultaneously. The pre-adolescent mortality rates tend to have the largest coefficients, but even controlling for cohort-specific later childhood mortality, infant mortality is significant for Sweden and Switzerland. On the other hand, results for France and England show a lack of significance or a negative coefficient for infant mortality, although later child mortality rates are significantly positive predictors of old-age mortality.

Using almost two decades of data from the Netherlands, van den Berg, Lindeboom and Portrait (2006) explore the relationship between being exposed to an economic recession during the first few years of life and later life mortality. They find that being born during a recession results in an 8 percent increase in the mortality rate after the first year of life and that exposure to reduced resources at the macro level during years 1 through 6 of a child's life have no longer-run implications. This can be interpreted as further evidence that conditions during the fetal period have long-lasting impacts on later life health.

In another historical cohort study, Kannisto, Christensen and Vaupel (1997) analyzed the impact of a large famine in Finland in 1866, 67 and 68, a period of three successive years of serious drought. (At the time, 92% of the Finnish population was rural.) Infant mortality rose to 40% (from 20%) in the last year of the famine, dropping back below 20% right after, so the famine's effect on infant mortality was large. Kannisto et al. compare cohorts born in the three famine years with cohorts born just before and just after. They find a large drop in survival to age 17 in the affected cohorts, consistent with the large period rise in infant mortality, but no impact on subsequent mortality from age 17 to older ages (40, 60 and 80). While later emigration out of Finland could cause biases, if the emigration were different for the different birth cohorts, data on emigration after age 17 shows roughly the same rates across birth cohorts. Of course, to the extent that the more frail died at young ages for cohorts exposed to the famine, this would reduce any observed mortality differences later in life. Unfortunately the data do not provide information on health conditions of the living when older. However, Kannisto et al. demonstrate that the impacts of these major health shocks *in utero* on adult health outcomes is not uniformly negative, as Barker and others hypothesize.

4.1.2. Quasi-experimental studies on fetal health and economic prosperity in later life

An innovative study by Almond (2006) explores the impact of impaired fetal health in the US on adult health, human capital accumulation and labor market outcomes. He compares people who were conceived during the Spanish flu pandemic in late 1918 (and born in 1919) with those born before the pandemic and those born at least nine months after the pandemic ended. By allowing for secular trend in the outcomes, he identifies the effect of being *in utero* during the flu pandemic by deviations from the trend. He also exploits variation in timing of the onset of the flu pandemic across the country. He finds large and significant effects of poor fetal health on disability in mid-life, educational attainment and income.

As he notes, Almond is not able to control for other factors that change abruptly because of the flu pandemic. A plausible candidate is food and other commodity prices since there was a positive spike in prices in 1918 associated with demands from World War I which may have affected women's health during pregnancy. Whatever the underlying reason, the evidence in the study suggests that insults to fetal health affect both health and economic prosperity well into adulthood.

Twin studies provide another type of quasi-experimental study. For example, Behrman and Rosenzweig (2004) uses data from a survey they conducted on a sample of same sex monozygotic twins from the Minnesota Twins Registry to estimate the impact of birthweight on adult height, body mass, schooling and wages. Since monozygotes share an identical genetic endowment, differences in their birthweights must reflect differences in their fetal health and nutrition. The average age of the sample is 45, so the impact of birthweight can be examined on outcomes in early and midlife adulthood.

Because the distribution of birthweight differences between twins is much smaller than that between two siblings or between two otherwise similar people in the general population, there may not be sufficient power to detect the impact of birthweight on these outcomes. In fact, examining differences in each outcome as a function of birthweight differences, Berhman and Rosenzweig find that the higher birthweight twin is taller, completed more education and has higher wages. In these data, BMI as an adult does not depend on birthweight. The literature suggests that birthweight may have a non-linear association with health and labor market outcomes. Behrman and Rosenzweig are not able to detect non-linear effects on any of the outcomes other than wages which may be a reflection of insufficient power.

4.1.3. Non-experimental evidence on early life health and later life outcomes

There is a very large literature that uses observational data to draw out the longer-term impacts of early life health. The historical and epidemiological literatures have documented the relationships between height and weight as adults, and subsequent adult mortality (Waaler, 1984; Fogel, 2004). To the extent that height as adults is positively conditioned by height attainment as a child, this evidence demonstrates links between childhood nutrition and health and health as an adult. Of course, if it could be that time persistent unobservables related to innate healthiness or lifetime resource constraints or preferences (perhaps correlated between parents and children) that are the underlying causes of these correlations.

One of the early epidemiological studies linking health as a young child and health when older is Barker and Osmond (1986) who show that in communities in England and Wales that had high neonatal and post-neonatal mortality in the early 1900s people had high adult death rates from coronary disease later in the century (Barker and Osmond, 1986). However, as Elo and Preston (1992) note, half of the population had migrated from its region of origin (and are therefore out of the sample) by the later survey dates, making the results less convincing. At the more micro level, death rates from coronary heart disease among British men and women born between 1911 and 1930 were higher among those persons with low birthweight (Godfrey and Barker, 2000). There is a real issue of how much of these simple correlations represent causal relationships. Many of these studies are bivariate only. Even multivariate studies in general can only control for a subset of potentially important influences. Hence, unobserved factors could easily result in these correlations, confounding the analyses. Still, these studies raise issues that cry out for further scientific investigation.

McDade et al. (2001) use multivariate methods with the Cebu Longitudinal Child Health and Nutrition Survey to examine the correlations between *in utero* health (that is intrauterine growth retardation as indicated by low birthweight for full-term babies) and antibody response to vaccination against typhoid when the same infants were adolescents. They chose a small, random sub-sample of 103 full-term babies of the Cebu study to participate in the vaccination study. Each child was vaccinated against typhoid, and antibody production was measured after vaccination. In their empirical analysis, McDade et al. estimate a health production function, including covariates for birthweight for gestational age, current (adolescent) body mass index, interactions between those two, whether the timing of maturation was early or late, weight gain in the first year of life, whether breastfeeding was long and interactions between those two covariates. They find that low birthweight is associated with a lower likelihood of exhibiting normal antibody responses to the typhoid vaccination, making the vaccination less effective. This negative effect of low birthweight is magnified if the adolescent also has low current body mass. The strength of this analysis is the rich set of health input variables spanning the entire lifetime of these adolescents available in the data. The weakness of the analysis is that these inputs might be determined jointly with the health outcomes. As discussed in Section 2, estimated effects will be biased if there are characteristics omitted from the antibody function that are correlated with the health inputs.

The McDade et al. (2001) study does not directly bear on very long-run health impacts of early childhood insults, as their Cebu sample uses data on adolescents. This issue is addressed by a recent study by Case, Fertig and Paxson (2005), who examine data from the National Child Development Study (NCDS) which collected information on a cohort of all children born in the UK during one week of March in 1958 and has reassessed those respondents periodically through 2004. Case et al. use data from birth through middle-age (42 years old). From the time of birth, they use birthweight and whether the mother smoked heavily during pregnancy. Heights at age 16 are available, as are certain household characteristics from that time period, such as family income. At age 42 and earlier adult ages, the health variable that is available is self-reported health: excellent, good, fair, poor. This is not ideal for the reasons explained above.

Using multivariate regression models, Case et al. examine the extent to which fetal and child health predict self-reported general health status at ages 23, 33 and 42 controlling socio-demographic characteristics. Low birthweight is associated with worse health, and taller children at age 16 report better health in later life although these effects are not always significant.[17] Conditional on these characteristics, whether the mother smoked during pregnancy is negatively associated with later life health which is interpreted as suggesting that *in utero* health affects health in middle age even after controlling birthweight. As Case et al. note, it is possible that these results are driven

[17] The mechanisms through which height at age 16 affects later life outcomes in not obvious given that height at this age is substantially affected by the timing of the puberty spurt and has not been shown to be associated with health and well-being in later life. A correlation between height at 16 and earnings at age 32 was reported by Persico, Postlewaite and Silverman (2004) using the same data.

by unobserved heterogeneity. In the context of the conceptual model discussed in Section 2, the estimated models are conditional demand functions with the conditioning variables being potentially correlated with factors that are not observed in the models. For example, a mother who smokes may also invest less in other dimensions of the health and human capital development of her children and so, controlling birthweight and genetic endowment, her children will perform less well in school, have more health problems and perform less well in the labor market in later life than similar children whose mother did not smoke.

More generally, studies that report correlations between early life health and later life outcomes are both suggestive and intriguing but do not establish causality. Finding that birthweight is predictive of health or human capital outcomes in later life may indicate that low birthweight is a marker for poverty or other sources of deprivation, broadly defined, while *in utero*. If that deprivation persists through life, it is not clear whether it is these other (unobserved and persistent) factors or low birthweight that causes poorer outcomes in later life. Distinguishing unobserved heterogeneity from state dependence is a problem that plagues many studies of behavioral choices and plays a central role in the literature on health.

The temporal nature of the measures rules out the possibility that later life outcomes cause early child health. There may well be unobserved heterogeneity that is correlated with both early life health and later life outcomes. Examples include unobserved factors that are associated with innate healthiness, tastes for investments in health, preferences towards good health and health behaviors.[18] Moreover, in the forward-looking model described above, it is possible that early life investments in health are made because of their implications for health in later life. Establishing that early life health has a causal impact on later life outcomes is not straightforward. In the absence of experimental or quasi-experimental variation in early life experiences, isolating the causal effect of, say, birthweight on later life outcomes, calls for measures of health input prices during the fetal period, that can be used as instruments in IV procedures. These might include the price, availability and quality of prenatal care, relative prices of foods that are rich in iron or other micronutrients that are critical during pregnancy.

4.1.4. Intergenerational transmission of health

Whereas there is a large literature on the inter-generational transmission of education and SES, there is substantially more limited evidence for health outside the biomedical and genetics literature. This is surprising given the fact that genetic endowments are transmitted across generations and, if fetal or infant health affects health and well-being in later life, it seems natural to assess the extent to which health is transmitted across generations.

[18] As discussed above, the same empirical issue arises in models of health and productivity. (See Behrman and Deolalikar, 1988 or Strauss, 1993, for early critiques.)

Animal studies have shown that fetal growth is retarded in offspring of mother rats that are fed protein-restricted diets over several generations. The offspring are born small relative to controls on a normal diet. However, when a later generation of those rats on the restricted diet were fed a normal diet, starting from conception, the rat is larger in size than controls. If these rats continue with the normal diet then within two generations, there are no differences in the size of the treatments and controls (Stewart et al., 1980).

Emanuel et al. (1992) show that in the NCDS, the 1958 British cohort study, there is a positive correlation between birthweights of mothers and children. This result has been replicated for the state of Washington by Emanuel et al. (1999). Historical data from a hospital in China shows that a higher BMI mother (greater than 26 compared to less than 23 in the 38th week of pregnancy) was more likely to give birth to a child who would have lower blood glucose and higher insulin concentrations when he/she reached age 60 (Barker, 2006). The same study demonstrates that maternal pelvic size is correlated with the odds of the child having a stroke or diagnosed hypertension as an adult.

An extensive literature in developmental psychology indicates that there are significant positive inter-generational correlations in health-related behaviors including smoking, drinking, substance abuse, diet and eating disorders as well as in psycho-social and emotional problems. (See Serbin and Karp, 2004 and Chapman and Scott, 2001, for reviews.) For example, the Concordia study (Serbin et al., 1991) indicates that children with mothers that had a history of aggression tend to have reduced cortisol reactivity when challenged by a verbal conflict task with the mother. Moreover, the same study reports a positive association between cortisol reactivity in the mothers and their children.

Using unique data from Denmark, Eriksson, Bratsberg and Raaum (2006) examine the inter-generational correlation in health of adults in their late forties with that of their parents focusing on a series of diagnosed conditions that include cancer, hypertension and heart disease, respiratory problems, allergies, migraines and psycho-social problems. Both own and parental reports are provided by the adult child. The adult child is at substantially higher risk of suffering from any of the ailments if either parent suffered from the same ailment. For example, if a father is reported to have had psycho-social problems, the child is eight times more likely to also report psycho-social problems; the risk is five times higher if the mother had psycho-social problems. Heart disease in either parent is associated with an elevated risk that is four times higher in the adult child. While some of these correlations may be driven by common measurement error (given the way the data were collected), they are large and consistent with evidence in the biomedical literature. Moreover, controlling own health status absorbs around a quarter of the inter-generational correlation in earnings.

Martorell et al. (1981) report a positive correlation between maternal height and the survival probabilities of their children in Guatemala. Thomas, Strauss and Henriques (1990) use multivariate models to demonstrate a positive association between a mother's height and the mortality of her children, for children from Brazil; controlling for factors

such as mother's and father's schooling and household nonlabor income. These results may be driven by the impact of maternal pelvic size. However, Thomas, Strauss and Henriques (1990) also show a strong partial correlation between mother's and father's height and the age/sex standardized heights of children, again controlling for parental schooling and household income.[19]

In a clever and insightful study, Almond and Chay (2006) note that the health of blacks improved dramatically during the 1960s while the health of whites changed relatively little. This is indicated by the infant mortality rate of blacks declining from around 40 in the mid-1960s to below 30 in the mid-1970s; the absolute reduction in infant mortality rates was smaller for whites albeit about the same proportionate change. They argue that black women born in the late 1960s are substantially healthier than those born a decade before them. To examine the inter-generational transmission of health, they use data from the annual Natality Detail files and compare the health of babies born to black women born in these two cohorts and find that babies born to the later cohort are less likely to be low birthweight and have better APGAR scores. This is not driven by secular change in the broader society since there is no corresponding improvement in the health of babies born to white women from these two cohorts. As a further check on this, Almond and Chay compare the improvement of child outcomes across the cohorts of black mothers born in Mississippi, a state that had very large declines in black infant mortality, to black mothers born in Alabama, where mortality declines were smaller. These results reinforce their other results. Almond and Chay conclude that there is substantial and significant inter-generation transmission of health.

There are surely several mechanisms underlying observed correlations in health across generations. Part is likely to be attributable to inter-generational transmission of genetic endowment. Part of the correlation may reflect non-genetic dimensions of parental health being transmitted across generations. It is possible that parental health is a marker for unobserved ability in managing the inputs into the production of one's child's health (μ in the health production function) as well as tastes for health and other goods (ξ in the utility function) and that these are transmitted across generations. The work of Eriksson, Bratsberg and Raaum (2006) suggests that part of the unobserved factors that affect labor market success may be correlated with health. While the evidence at hand does not establish which of these mechanisms is important across contexts, it seems that this line of research has the potential to open new windows into the ways in which health investments and health outcomes affect the well-being of individuals and their progeny.

4.2. Child and adolescent health, human capital development and economic success

The previous subsection focused on the relationship between health *in utero* or during infancy and well-being, including health, in later life. Moving to the next stage in the life

[19] Also see Horton (1986) for an earlier study which uses mother's height to explain child's height.

course, in this section we explore the relationships among health and other indicators of human capital during childhood and adolescence. We also examine the longer-term effects on health in childhood on health and well-being during adulthood. Health in adulthood is discussed in the following subsection.

Studies of the links between child health, cognitive achievement and schooling outcomes are summarized by Behrman (1996) and discussed in detail in Glewwe and Miguel (2008). Therefore, we focus our discussion on relationships between different dimensions of health and human capital as a child including child growth. We then highlight the implications for well-being in later life.

Many of the early studies in this literature relied on non-experimental cross-sectional data and, by necessity, made strong assumptions. However, in recent years, a substantial body of research has developed which uses innovative experimental or quasi-experimental designs, in combination with longitudinal data, to provide new insights into the causal effect of child and adolescent health on well-being during childhood and as an adult.

4.2.1. Experimental studies

One of the most innovative studies in the field of health and development, and probably the best known large-scale treatment-control experimental study linking child health to cognitive development, is the INCAP longitudinal study (1969–1977) in rural Guatemala. Subsequent follow-ups in 1988–1989, 1997–1998 and 2002–2004 provide uniquely rich data spanning over three decades to examine the links between child health and well-being in adulthood. The motivation, study design and some of the key results are described in Martorell, Habicht and Rivera (1995), Grajeda et al. (2005a) and Martorell et al. (2005).

Four villages were chosen for the study. Pairs of villages were matched using fifteen village characteristics such as population, age distribution, nutritional status and health status. For each pair, one village was randomly assigned to treatment and the other to control. Using only four villages and randomizing at the village level limits the power of the experiment and also limits the extent to which matching yields a truly random design.

The study focused on young children, from birth to seven years in 1969–1977, and pregnant and lactating women. The treatment involved providing a daily nutritional supplement in the form of a fortified drink, Atole, which contained about 39 kilojoules of energy and 11.5 g of protein along with some micronutrients including iron and niacin. (This amounts to about 10% of recommended daily intake of energy and 20% of protein in a well-nourished population.) The placebo drink, Fresco, was provided in the control villages and contained around 14 kilojoules of energy from sugar.

Two years after the start of the intervention, both drinks were fortified with additional micronutrients so that Atole and Fresco contained the same amounts of iron, niacin, vitamin A, thiamine, fluorine and ascorbic acid. Health care services, including free curative care and immunizations, were provided in all four villages. To the extent that

these health services, and the change in micro-nutrient content of the drinks, do not have the same impact on treatments and controls, then inferences about the impact of the protein/calorie supplement will be confounded.

The drinks were provided at centrally located venues in each village during the day. Participation was voluntary and about three quarters of all age-eligible children and pregnant or lactating women participated. Modeling participation is not straightforward and so most analyses have examined the intent to treat effect on all eligible villagers. There are two particularly troubling aspects of this dimension of the design. While participation rates are high relative to other, similar studies, in the treatment villages, lower-SES children were more likely to consume the drink; however no such difference is observed in the control villages. Second, since children older than 7 were not eligible for the drink, it is possible that families substituted food in favor of older children because younger children were receiving the supplement. In fact, children in larger families were more likely to participate in the program. These are additional factors that potentially confound interpretation of differences in outcomes between treatments and controls and between exposed cohorts and those cohorts not intended to be exposed.

The amount of the drink consumed by each subject in the study also differed. In the control villages, young children (age 0–3 years) drank far less of the drink than children in the treatment villages. Thus, the latter ingested more calories, protein and micronutrients. However, mothers in the control villages drank more than mothers in the treatment villages and so intakes of calories were roughly the same for all women and micro-nutrient intake was higher among the women in control villages. It is thus very difficult to interpret differences between treatment and control women. For example, a comparison of birthweight of children born to women in treatment relative to control villages yields no difference. However, birthweight is positively correlated with calorie intake from the drinks which, of course, reflects at least in part choices by the women. (See Martorell et al., 2005, for an excellent discussion of these issues.)

Baseline measurements of many health and nutrition factors were taken in this study and the children and mothers were followed up with measurements on breastfeeding, illness symptoms, solid food intakes, and anthropometry at intervals which depended on the child's age. Many other measurements were taken, particularly on cognitive development, starting during infancy, up to age 7 years during the 1969–1977 period. Taking the daily supplementation was closely monitored and records kept. While relatively little economic information was collected in the early phase of the study, the economic content of the follow-up surveys was substantially expanded in later waves.

The data from this study have been very extensively used, although only a subset of the studies exploit the experimental design. As a good example of the latter, Habicht, Martorell and Rivera (1995) use a difference-in-difference approach to measure the impact of supplementation on child growth. Comparing subjects from treatment villages with those from control villages and also comparing birth cohorts that were included in the study with those not included, they find that supplementation affects linear growth to around 36 months of age but not thereafter. This has been a very influential study and indicates that nutrition interventions in very early life are likely to have significant

benefits in terms of improved health and well-being throughout life (see Martorell et al., 2005, for a summary).[20]

The experimental design has also been exploited to measure the impacts of nutrition on cognitive learning and schooling (Stein et al., 2005, has a summary). The first studies focused on children up to age three and reported very small, but significant, effects of the supplementation on a series of development markers. Subsequent studies on children when they were older found larger impacts on test scores, particularly among treated children from low-SES families (Pollitt et al., 1993). While the study exploits the experimental design, it also interacts the treatment group with an SES index at the time of the measurement and with years of schooling of the child, both of which are arguably related to unobserved characteristics that affect cognition and, therefore, are endogenous in the model. Moreover, recall that in treatment villages children of low-SES families were more likely to participate in the study which is, itself, another choice.

Maluccio et al. (2006) use the 2002–2004 follow-up data together with the original 1969–1977 data. The 2002–2004 follow-up data collected a lot of economic data, in addition to adding to the health data collected in prior waves. This round also administered achievement tests and Ravens' colored progression matrix tests to the survey respondents.

Maluccio et al. examine whether schooling attainment and test scores were affected by the supplementation given to treatments. Most respondents were in their 30s by 2002–2004 and thus had completed their schooling. Maluccio et al. estimate a reduced-form demand function for schooling attainment, test scores and other outcomes. To measure the intent to treat effect of the supplementation they use a difference-in-difference approach with all children in the study, irrespective of their participation status. They compare children exposed to the Atole treatment relative to those exposed to the Fresco control, drawing comparisons between children in birth cohorts that were exposed to the study from birth through 36 months of age with cohorts not fully exposed during those ages. Village fixed effects absorb the main effects of differences between villages. They find significant and positive effects of exposure to supplementation on a test for abstract reasoning, the Ravens' colored progression matrix test. Performance in school, however, was only enhanced among female subjects: they completed more grades of school by age 13, completed an extra year of schooling by early adulthood and they performed better on Spanish reading tests. It is not clear what drives these differences between males and females.

It is critical in any experiment that all subjects be followed. This is harder the longer-term the study. In the case of the INCAP surveys, the follow-up rates of children in the 1988–1989 survey were 70–75%, depending on the village and there is no evidence of differential attrition between the villages (Martorell, Habicht and Rivera, 1995). Whereas among those who did not move from the village, re-contact

[20] Several analyses which use these data but do not exploit the experimental design draw similar conclusions. These studies are plagued with potential confounds associated with unobserved heterogeneity.

was 86–90%; among migrants, 40–45% were re-interviewed (in part because tracking of movers was limited to the local district capital or Guatemala City). By the 2002–2004 wave, more of the study subjects had moved from the original village and attempts were made to re-contact all respondents no matter where they lived. They found about 60% of the original sample of children. Since 11% of the respondents had died, this amounts to a recontact rate of 70% among all eligible respondents (Maluccio et al., 2006). Their own analyses and those presented by Grajeda et al. (2005b) indicate that men, younger people and people who were not living with their parents in prior waves were less likely to be re-contacted. On the positive side, Maluccio et al. (2006) find insignificant effects of the exogenous intent to treat variables on individual attrition, which suggests attrition is unrelated to the treatment. They also use the method suggested by Fitzgerald, Gottschalk and Moffitt (1998) to reweight the sample and thereby control for attrition bias that can be attributed to selection on observed characteristics. The results are almost identical to those that ignore attrition, suggesting selective attrition that is correlated with treatment status is not important in this model.

A second important experimental study provides deworming treatments to school-age children in rural Kenya and examines the impact of reducing worm loads on school attendance and school performance (Miguel and Kremer, 2004). Intestinal helminth (worms) infections are prevalent in many parts of the developing world, and when worm burdens are high, they can result in iron deficient anemia, protein-energy deficiency and abdominal pain. Over one third of the children in the study suffered from moderate to heavy worm loads at the initiation of the study.

Most types of worms are readily treated with single-dose oral therapies. Seventy-five primary schools were randomly assigned to one of three groups. Children in schools in the first group received deworming treatments for two years starting at the initiation of the study. Children in schools in the second group received treatments in the second year only. Children in schools in the third group did not receive treatments until three years after the start of the study.

The treatments were administered in the schools and so children in treatment schools who were absent on the day of administration failed to get the treatment. About one-quarter of the children in the target schools did not receive the treatment with the compliance rate being lower in the second year.

After treatment, children in the treatment schools tended to be in better health than children in control schools. Deworming significantly increased school attendance in treatment schools by seven percentage points, which is a 25% reduction in school absenteeism. These gains are largest among the youngest children. Test scores are not affected by deworming.

The study demonstrates that health has a causal impact on school attendance although the evidence for benefits in terms of improved learning and cognitive development are less clear. The study also highlights the importance of externality benefits of deworming since reduced incidence of worms will lower infection rates for others in the school or

community. For more discussion, see the chapter by Glewwe and Miguel (2008) in this volume.

Another set of important experimental studies exploits data from an evaluation sample of villages in Mexico that were randomly assigned to receive PROGRESA benefits immediately or with a delay. Those studies are discussed by Parker, Rubalcava and Teruel (2008) in this volume.

4.2.2. *Quasi-experimental studies*

Several quasi-experimental studies in low-income settings have used weather shocks to identify the effect of health shocks on subsequent health (weight gain) (Foster, 1995) and schooling outcomes (Jacoby and Skoufias, 1997). Strauss and Thomas (1998) survey that literature. Recent studies have examined the longer-term consequences of health shocks on health and schooling using the dynamic conditional demand function framework (15) discussed in Section 2.

A very useful set of studies uses data collected by Bill Kinsey from households living in rural land resettlement schemes in Zimbabwe. The project began in the early 1980s in rural Zimbabwe in an effort to assess the impact of a resettlement program initiated by the newly-elected majority-rule government. Households were surveyed in mid-1983 through early 1984, in 1987, 1992 and then annually through 2001.

Hoddinott and Kinsey (2001) use anthropometric data collected annually between 1993 and 1997 to examine growth in height among children age 12–24 months in 1993. They regress the change in child height (in cms) between adjacent years on lagged height (in cms), gender, age in 1993, time between measurements and time interacted with initial age, a series of mother's characteristics, including schooling and height, a series of household characteristics, including livestock and land holdings, and time dummies, including one for the drought year.

The OLS estimate of the coefficient on lagged height is -0.33 suggesting significant but incomplete catch-up in linear growth.[21] This is also consistent with reversion to the mean driven by measurement error.[22] As Hoddinott and Kinsey note, and as discussed in Section 2, it is not clear what to make of OLS estimates of this model. They adopt two different empirical strategies.

First, they apply an IV approach to account for the endogeneity of lagged height, using the child's birthweight and whether birthweight is known as instruments. There are two concerns. First, it is assumed that birthweight affects current height through height in the previous year. The limited evidence that exists suggests this is unlikely

[21] There was a severe drought in 1993–1994 which is associated with reduced linear growth, particularly among the youngest children (age 12–24 months).

[22] The authors were careful to take out some gross outliers based on height and that resulted in the -0.33 coefficient on lagged height. When they use all of the observations, the coefficient on lagged height falls to -0.59, which again is consistent with random measurement error imparting a negative bias to the growth equation estimate being a serious concern.

to be true at least for very young children. (See the discussion above of the INCAP results, in which birthweight is found to be negatively correlated with child growth up to 36 months, or the hypotheses on the impacts of early infections proposed by Crimmins and Finch, 2006.) The second and probably more important issue with the instruments is that birthweight is reported by the mother and known for only a fraction of the children. Putting aside recall issues, a mother will only know the birthweight if it was measured which, in Zimbabwe, typically means the birth took place in a clinic or hospital. Low-SES mothers are likely to give birth at home and so will not know the birthweight; if their children tend to grow less well than children of higher-SES mothers, the instrument will likely be correlated with unobserved characteristics in the model of linear growth.

As a second approach, Hoddinott and Kinsey estimate the growth in height model including mother fixed effects. The coefficient on lagged height is −0.81 suggesting almost complete catch-up growth. However, it is well known that in a model specified in terms of growth, random measurement error in the lagged dependent variable (lagged height in this case) will impart a negative bias in the estimates, apart from the usual bias towards zero. Comparing siblings (by including mother fixed effects) typically raises the noise to signal ratio and exacerbates the negative bias due to random measurement error. This is likely to at least partly explain the decline in the coefficient on lagged height from −0.33 to −0.81. Random measurement error in lagged height is likely to be exacerbated with the differencing involved with a mother's fixed effects estimator, then this negative bias may be enhanced, which would explain the drop in the lagged height coefficient from −0.33 to −0.81. In any event, mother fixed effects estimation will only control for correlations between birthweight and omitted variables if parents do not compensate for differences between children that are correlated with birthweight. (See Rosenzweig and Wolpin, 1988, for a classic treatment.)

Using data from Russia, Federov and Sahn (2005) estimate the impact of child health on later child health with the GMM method proposed by Arellano and Bond (1991). This amounts to taking first differences of individual child height to eliminate child-specific time invariant factors and then using instrumental variables to take into account correlations between lagged health and time-varying unobserved factors including measurement error. The assumption necessary for these estimates to be consistent is that there is no serial correlation in time-varying unobserved characteristics such as innate healthiness or ability. This is a strong assumption.

Recognizing this, Alderman, Hoddinott and Kinsey (2006) combine maternal fixed effects with instrumental variables to analyze the impact of pre-school height on later health and schooling using the Zimbabwean data. They examine three outcomes: adolescent height, age at which the child started school and the level of schooling completed by 2000. Pre-school height, which was measured in 1983, 1984 or 1987, was assumed to have been affected by two exogenous shocks: the 1982–1984 drought, the worst drought in living memory, and the 14-year civil war which ended with one-man one-vote elec-

tions in 1980. It is assumed that, controlling pre-school height, these shocks had no independent effect on subsequent height and schooling outcomes.[23]

In the first stage regressions, the drought and civil war are significant predictors of pre-school height. In the second stage, pre-school height has a positive and significant impact on adolescent height and years of attained schooling. Taller children also tend to start school earlier although this effect is not significant. The estimated effects are large. The drought resulted in adolescents being 2.3 cm shorter and completing 0.7 years less schooling. (By comparison, adult height increased by about 1 cm per decade in many developing countries in the last half of the twentieth century. See Strauss and Thomas, 1995.)

There was a substantial amount of attrition with around 60% of age-eligible respondents for this study being assessed in the 2000 wave (which is the wave when adolescent height and school attainment are measured). Attriters tend to be female, earlier birth cohorts and were taller prior to starting school. Conditional on age, sex and village controls, pre-school height is not a significant predictor of attrition suggesting that attrition bias is not related to unobserved characteristics correlated with pre-school height. Whether attrition is related to later height and schooling outcomes is, of course, an open question.

While the identifying assumptions in this paper are arguably weaker than those in most of the non-experimental literature on this topic, it is not clear that the 1982/1984 drought and civil war were "shocks" in the sense they were unanticipated and exogenous (to child human capital). The severity of the 1982–1984 drought might have been unexpected, droughts are common in southern Africa and the impact of the drought likely lasted several years through its impact on grain storage and prices and on livestock holdings and prices. This would likely affect human capital accumulation after the drought ended and so identification based on the years of the drought alone may not be appropriate beyond the drought. The civil war directly affected virtually every family in Zimbabwe and carried with it a tremendous amount of dislocation and destruction. Its end was not unanticipated and the impact of the subsequent rebuilding of infrastructure and civil society possibly affected adults and children for many years beyond the end of the war.

The impact of child health, as indicated by height, on later schooling outcomes is also examined by Alderman et al. (2001). Using 5 years of panel data from Pakistan, they exploit variation in prices, rather than weather variation and civil war, to identify pre-school height in a child schooling equation. Specifically, they model school enrollment at age 7 as a function of standardized height at age 5, controlling parental schooling, household composition and a three-year average of household per capita expenditure over three years (to represent long-run resources), and current prices of wheat, rice

[23] Specifically the 1982–1984 drought shock is an indicator variable for children who were born in 1981 or 1982 and so were age 12 to 36 months during the drought. The child's exposure to the civil war is measured by (the logarithm) of the age of the child (in days) when the war ended on August 18, 1980. These shock variables amount to cohort effects.

and milk. Village-level fixed effects are included in the models, so that the prices can be interpreted as deviations from village means which they interpret as current price shocks. The variation of price shocks over time is the source of the identification. Current prices enter the school enrollment probit, while lagged prices enter the pre-school height regression as identifying instruments. Village-level dummies, or fixed effects, capture expected prices. To obtain extra power in the first stage, the community-level price variables are interacted with household-level variables: mother's being schooled and gender of the child, as additional instruments.

The IV results indicate a strong, positive impact of height at age 5 on girls' enrollment at age 7, but not on boys' enrollment, an effect which is much larger than the OLS estimates. Paralleling the approach used in the Zimbabwe case (Alderman, Hoddinott and Kinsey, 2006), a key assumption is that price variation at age 5 affects height at that age but has no longer-lasting impacts on budget allocations, height and schooling when the child is older. If that assumption is not correct then other health variables belong in the second stage enrollment equations and if one does not include them, then lagged prices would affect the school enrollment decisions. In this case the price shocks would not be valid instruments. Given the discussion in the previous subsection, one might be concerned that since the initial measurement is at age 5, this may be too late for current prices to influence current height in a way that is likely to have a long-lasting impact on human capital. Alternatively, if we interpret these prices as being proxies for prices at age 2 (assuming that the prices are strongly serially correlated) then one can interpret the results more easily.

Rosenzweig and Zhang (2006) use an innovative sample of twins and non-twins, aged 7–18, taken in Kunming, China, to examine the impact of birthweight (and twins) on schooling and health of children. The researchers worked with one of the government's statistics bureaus and examined the 2000 Census data to identify likely twins in each household. These households were then surveyed, if found, in 2002, and a sample of non-twins were taken in the same areas. The twins were not necessarily monozygotic. Rosenzweig and Zhang (2006) use a within-twins fixed effects estimator to measure the impact of birthweight on non-required schooling expenditures per child (in China some expenditures, such as fees, would be required, while others, say for tutors, are not). The sample used in this exercise is from urban areas affected by the one-child policy, so that parents are constrained from having more than one child unless they had a twin. They are able to look at twins at first birth, which can be taken as exogenous in the China setting. They find a significant, positive impact on child educational expenditures, but one that is much lower than in a simple OLS estimate.

In a different set of regressions, they take the sample of urban children and regress measures of expected college completion (reported by the parents), scores on math and literature achievement tests, self-reported health being excellent or good, weight, height and BMI on whether the child is in a household with a first-born twin's birth and birthweight. Of interest here are the birthweight coefficients, which Rosenzweig and Zhang (2006) find to be significantly and positively related to all of the health variables, but not

to the achievement test scores and only weakly to the parents' expectations of college enrollment.

One unknown in this study is how select the analysis sample is because of requiring data on birthweight. As explained above, birthweight data coming from self-reports, which these data presumably do, may suffer from selection issues related to who knows about birthweight.

4.2.3. Non-experimental studies

In a closely related paper to Maluccio et al. (2006), Behrman et al. (2006) estimate test score production functions for reading and cognitive skills, separately, for the INCAP 2002–2004 sample. The tests were given to the respondents in the latter survey period, when they were adults. Behrman et al. use variables measuring whether the respondent was stunted at age 7, years of schooling, age and age squared at the test, and tenure for having a skilled job, as covariates. The production functions do not use the experimental nature of the data, so non-experimental techniques are required for estimation. Behrman et al. treat all the covariates except age as endogenous and use instrumental variables to estimate the production function. For the instruments, however, the experimental nature of the data is used. As instruments for the variable stunting at age 7, being in a birth cohort fully exposed to either the supplementation or placebo, alone and interacted with a dummy for being in a treatment village, are used. These are the variables used by Maluccio et al. (2006) to measure intent to treat effects. Other instruments include household wealth when young, mother's and father's education, and variables measuring village teacher–student ratio at age 7 and whether a secondary school was available in the village at age 7. Behrman et al. find positive impacts of not being stunted at age 7 on these later adult test score outcomes.

There are two related issues with the interpretation of these estimates. First is whether the instruments are appropriately excluded from the achievement test production function. Take the intent to treat variables for instance. Excluding them assumes that the only nutrition variable that affects test outcomes is being stunted at age 7. But it is possible that the impact of height is not discrete, that weight-for-height at age 7 might matter or perhaps height at earlier ages might affect test outcomes. Indeed, the fetal programming hypothesis would suggest that birthweight may matter, after conditioning on height at age 7. Likewise, one can reasonably ask why school quality measures such as teacher–student ratios do not directly affect test score achievement. The second issue, which comes up in most attempts to estimate a human capital production function, is how completely and how well inputs are measured. The inputs used in this paper are rather sparse. As a result, it is likely that many of the instruments will proxy for those inputs and so should not be excluded from the second stage.

Glewwe, Jacoby and King (2001) use longitudinal data to also examine the impacts of early child health on later schooling outcomes. Using the Cebu Longitudinal Child and Nutrition Survey, they estimate an achievement production function in which lagged child height is one of the inputs. These data are, in principle, perfect for this purpose,

since children are followed from birth to 11 years. Unfortunately Glewwe et al. do not use anthropometric data from the first two years of life, except as instruments. Rather they examine the impact of child height at the time of school enrollment (about age 7 in the Philippines) on school achievement test scores at age 11 or so. Other input variables are variables measuring gender, age at first enrollment, months in school in 1st and 2nd grades and in 3rd to 6th grades (as separate variables), months repeating these grades, months not in school and mother's years of schooling. The reason that heights before 2 years are not used in the test score equations is that family fixed effects are used in the estimation, together with IV. The family fixed effects involves taking the difference of covariates between children. To use height at age 2 years as a covariate would then require that both children have these data available. However, while the Cebu data has heights at birth and up to 24 months for index children, for younger siblings, height data are only available at sporadic ages, up to age 7, and in general not at age 2 years, so there is a missing data problem.

However, heights at age 7 are available for both index children and younger siblings. The family fixed effects controls for unobserved characteristics that are the same for both siblings. This may include parental preferences for schooling and health investments, for example. However, as long as parents observe innate ability of the individual children and act on that while making their investment decisions, there will still be omitted variables correlated with the sibling height difference, thus causing bias. To control for this, Glewwe et al. use instruments separately for the school inputs and the lagged height. The instruments for the schooling input variables are month of birth dummy variables for both siblings, age difference of testing of the siblings, and various interactions. For the lagged standardized height-for-age, the instrument is the height-for-age of the older sibling at birth, 12 and 24 months. The idea (from Rosenzweig and Wolpin, 1988) is that early childhood height of the older child is known when schooling investments are made in both children and should be correlated with the difference in sibling heights, but that this early height should be uncorrelated with later schooling outcomes conditional on later heights. A critical assumption being made is that time-invariant unobserved factors in a child's health production function are uncorrelated with unobserved factors that affect scores in a test score production function. The assumption is not tested. In addition, it may be that shocks to child health at birth or ages 12 and 24 months affect child cognitive development, even controlling child height at age 7. In this case, again, these variables belong in the achievement production function in addition to height at age 7.

Glewwe and King (2001) take advantage of the time-varying anthropometric measurements on the index children in the Cebu sample and estimate conditional demand functions for cognitive achievement, as indicated by a score on the Raven's colored progressive matrix assessment. Their covariates are birthweight, growth in height in 6 month intervals from 0–6 months, up to 24 months, and then growth from 24 months to 8 years; plus parental schooling and household wealth (which is time-varying). They use lagged rainfall and prices to instrument the lagged height growth variables, adding mother's height and arm circumference in some specifications. In the first stage esti-

mates, few variables are individually significant, but Glewwe and King do not present F-statistics to test the identifying variables as a group. They find that while OLS results are positive and highly significant, IV estimates tend not to be significant, and when they are, are not consistent across different instrument sets.

While a number of studies have now appeared that examine impacts of exogenous shocks, rainfall or price, on health, few studies examine what can be done to counter these impacts. One exception is the Guatemala supplementation experiments. Another is a study by Yamano, Alderman and Christiaensen (2005) that examines whether food aid, given to households, results in enhanced child growth, using panel data from Ethiopia from 1995/1996. Yamano et al. have data on heights and weights of children six months apart. They know whether the household received food aid during this period, and how much. They estimate a dynamic conditional demand child growth equation, similar to the studies discussed previously, in which the change in child height is regressed on initial child height, age and a series of detailed household characteristics having to do with education and asset holdings within the household. Among the household variables is the amount of food aid received over the period. Community fixed effects are used to capture price, rainfall, and other community shocks. Yamano, Alderman and Christiaensen (2005) find that food aid is positively and significantly related to child growth in height, after controlling for a rich set of factors.

In their analysis, lagged height is treated as endogenous, but due to random measurement error. Lagged weight is used as the instrument, which requires the assumptions that measurement errors in height and weight are uncorrelated (which seems reasonable) and that lagged weight does not belong in the height growth equation conditional on lagged height (a stronger assumption, but one regularly made in this literature). This method does not correct for potential bias due to omitted variables; the authors argue that the bias due to measurement error may be the more important of these, but there is no way to know for certain.

Household (and community) receipt of food aid is also treated as endogenous, since it is the case that food aid was not distributed randomly across localities. Following a paper by Jayne et al. (2002) that used the same data, Yamano, Alderman and Christiaensen (2005) use the central government's assessment of each community's food aid needs during the famine period of 1984/1985 as an identifying instrument for household receipt of food aid, after conditioning in the second stage on measures of current year crop shocks at the household level. The idea is that it appears that the regional distribution of food aid in Ethiopia is quite similar to the distribution in the famine period, even ten years later. This is so controlling for many factors such as the current rainfall shock in each area, the coefficient of variation of long-run rainfall in each area and even larger area fixed effects. Hence, past central government assessments of food aid needs still strongly predicts whether the village (and households within the village) gets food aid and how much.

4.3. Adult health and SES

A large and growing literature documents a positive correlation between many dimensions of health as an adult and several indicators of socio-economic status.[24] As noted above, studies have demonstrated that early life health has important consequences for later life outcomes. This is manifest in one of the most robust findings in the literature relating health to socio-economic status in developing countries: in a wide variety of contexts, adult height is a powerful predictor of economic productivity as indicated, for example, by hourly earnings. In general, the association between height and economic success is substantially stronger in low-income countries than more developed economies. This likely reflects differentials in the returns to strength in the labor market, the incidence of poor nutrition in early life and differences in the extent of inter-generational transmission of human capital (Strauss and Thomas, 1998).

Pinning down the causal pathways that underlie the associations between health and socio-economic success has proved to be difficult. To date, experimental and quasi-experimental methods have provided the most compelling evidence on the impact of health on economic success.

Joshi and Schultz (2006) use "randomly-assigned" variation in access to contraceptives in rural Bangladesh to examine the impact on later female health, productivity, earnings and wealth. Using data from the 1996 Matlab Health and Socioeconomic Survey, Joshi and Schultz exploit the fact that starting in 1977, half the villages in Matlab were assigned to receive services from an intensive family planning and maternal-child health outreach program involving home visits every other week and encouragement to adopt modern contraceptives. While the selection of treatment and control areas was not strictly random, it is close for this analysis; an analysis by Joshi and Schultz (2006) shows that baseline fertility and schooling are uncorrelated with being in a treatment or control area. A comparison of completed fertility of women in treatment relative to control villages demonstrates the outreach program resulted in a large and significant decline in fertility. Joshi and Schultz investigate whether lowering the effective price of reproductive health services also affects other dimensions of the lives of women twenty years after the program was initiated. They report that female health, as indicated by BMI, is significantly improved and estimate that child mortality risks are reduced by about one-fifth. Better-educated women who lived in treatment villages reported higher earnings and total income and greater wealth.

Variation in the price of health care is also provided by the RAND Health Insurance Experiment (HIE) which randomly assigned subjects to different combinations of deductibles and co-payments. As noted above, those who received free care used more health care and were in better physical health at the end of the intervention. Dow et al. (2004) exploit the variation in price of health care to examine the short-run impact on labor supply. They report that females who received free care were more likely to

[24] Harris, Gruenewald and Seeman (2007) provide a recent review.

work than other females. A similar result is reported for males who had not completed high school. An experiment in Indonesia involved changes in the prices of health services. User fees at public health centers were raised in randomly selected "treatment" districts while prices were held constant (in real terms) in neighboring "control" districts. Two years after the intervention, relative to control areas, health care utilization and labor force participation had declined in treatment areas (where prices increased). Reductions in employment were particularly large (and significant) for men and women at the bottom of the education distribution, those whom we would expect to be the most vulnerable. The most plausible interpretation of both the HIE and Indonesian results is that the average treatment effects on labor supply indicate a causal role of improved health on the allocation of time to the labor market (Dow et al., 2004).

Experimental designs are well-suited to isolate the impact of specific nutrients on labor outcomes. Iron deficiency is prevalent throughout the developing world, particularly in south and southeast Asia. In addition, several studies have demonstrated there is a causal effect of iron deficiency on reduced work capacity suggesting there may be a direct link between iron deficiency and earnings (Haas and Brownlie, 2001). Iron plays an essential role in oxidative energy production. Iron deficient anemia affects physical activity through two main pathways. As hemoglobin levels decline, the maximum amount of oxygen the body can use (aerobic capacity) declines. As iron stores are depleted, the amount of oxygen available to muscles declines, reducing endurance and causing the heart to work harder for the same activity. Iron deficiency is also associated with, *inter alia*, greater susceptibility to disease, fatigue and reduced cognitive development.

Rigorous clinical trials with animals and humans demonstrate a causal relationship between iron deficiency and reduced maximum aerobic capacity (VO_2max) which changes by about 25–30% as subjects are made anemic or receive adequate iron supplementation. Iron deficiency is also associated with reduced endurance at below maximal work rates.

Demonstrating iron deficiency impedes maximal capacity and endurance does not reveal the economic consequences of iron deficiency in daily life. Those consequences may be more closely aligned with energy efficiency (the amount of physiological energy required to perform a given task). In fact, evidence from randomized treatment-control studies of Chinese female cotton mill workers and Sri Lankan female tea plantation workers suggest that elevated productivity resulted in changes in time allocation. Specifically, both studies found no evidence of greater productivity per hour among iron-deficient subjects who received iron supplements relative to subjects who were not supplemented. However, in both studies, treated women re-allocate time away from work and towards non-work activities (Li et al., 1994; Edgerton et al., 1979). Results from a random assignment treatment-control iron supplementation study of male rubber workers, however, indicates that treatments who were anemic at the initiation of the study were able to tap around 20% more rubber after supplementation, relative to anemic controls (Basta, Soekirman and Scrimshaw, 1979). Interpretation of this result is, however, complicated by the fact that attrition from the study exceeded 50%; if those

subjects who did not benefit from the study were more likely to attrition, the estimated benefits will be biased upwards.

The Work and Iron Status Evaluation (WISE) is designed to examine the immediate and longer-term impact of providing iron supplements to older adults in Central Java, Indonesia (Thomas et al., 2006). In the population, about one-quarter of older men and a third of older women are iron deficient (as indicated by low hemoglobin). Older males are randomly assigned to receive a weekly iron supplement for slightly over a year or an identical appearing placebo. Everyone in the man's household was similarly assigned to treatment or control (in the small fraction of households with more than one older male, a random male was selected to determine assignment to treatment or control for the entire household).

Respondents who were iron deficient prior to the intervention and who were assigned to the treatment have higher levels of iron in the blood and are able to cycle for longer on an ergocycle at the end of the intervention, relative to comparable control subjects. Iron deficient treatments are better off in terms of physical health, psycho-social health and economic success. Relative to similar controls, treated subjects are more likely to be working, lose less work time to illness, are more energetic, more able to conduct physically arduous activities and their psycho-social health is better.

About half the male workers in the study are self-employed (primarily rice farmers) and the other half are paid a time wage. There is no evidence that hours of work responded to the treatment for time-wage workers or the self-employed. Among males who earned a time wage, there is no evidence of changes in productivity as indicated by their hourly earnings. Of course, if their wages are set by an employer, it is not obvious the worker will reap the benefits of greater productivity. This is not true for the self-employed. Males who were iron deficient and self-employed at baseline reported around 20% higher hourly earnings after six months of supplementation relative to similar controls. Since there was no change in their hours of work, this translates into 20% higher income from labor.

Wage workers who received the treatment reduced the amount of time spent sleeping by around 40 minutes and reallocated all of this time towards leisure. Self-employed workers made no such adjustments. A picture emerges of iron deficiency having a causal impact on work capacity and energy needed to complete tasks. The self-employed who benefited from the treatment allocate the additional energy to their fields and work harder, produce more and earn more per hour of work. In contrast, wage workers were able to channel the greater energy to reduced sleep and more time allocated to leisure. After twelve months of supplementation, there is evidence that some of the treated iron-deficient males who were working for a time wage at baseline had shifted to self-employment or taken up an additional job. While the study demonstrates iron deficiency has a causal impact on time allocation and economic productivity, it also highlights the importance of taking behavioral responses to the experiment itself in assessing the impact of the treatment.

4.4. Health and exogenous shocks

A literature has developed that explores the impacts of exogenous shocks, such as from weather, on health and other outcomes of interest. Most of this literature explores only short-run effects of shocks. Estimating longer-run effects is much more difficult, as explained in Section 4.1, because in part, other changes will occur after the shock, some because of the shock, and it will often be difficult to parse out the causality between the shock and other events.

Much of the earlier literature on health impacts of shocks focused on economic shocks such as unanticipated changes in the prices of staple foods. Alternative to using measured changes in prices or incomes, a number of papers compare health outcomes of birth cohorts (usually of young children) that were affected by the particular shock to adjacent birth cohorts, not so affected (for instance, Razzaque et al., 1990, short-run analysis of the 1974 famine in Bangladesh; Stein et al., 1975, long-run analysis of the Dutch winter famine; or Almond, 2006, long-run analysis of the 1919 flu pandemic). See the extended discussion of these historical and long-run studies of cohort "shocks" in Section 4.1. The one difficulty in these types of cohort studies is to ensure that there were not other factors changing over the same period that could confound the effects of the shock being measured. If regional variation is also present in exposure to the shock, this can help alleviate such potential biases (see Duflo, 2001, for such a modeling example using cohorts and regional differences in a different setting), as can appropriate interactions of the shocks with household-level variables. Interacting a community-level shock with a household-level variable can also greatly increase the power of test for the shock impact, but may suffer from endogeneity problems if the household variable is not exogenous. In practice, many studies use household assets because households with low assets may have more difficulty in obtaining loans, therefore the shocks may hit them harder. However, in a dynamic model many types of household assets will be endogenous. One asset that may be exogenous in some parts of the world is land owned. In south Asia, it is argued for example, that land sales markets are very thin and thus land rarely changes hands except for bequests (Rosenzweig and Wolpin, 1985).

A good example of a paper that uses rainfall shocks to investigate impacts on health is Rose (1999). She uses the NCAER ARIS national Indian household panel data from 1969–1971 and merges rainfall data at the district level. She examines the impact of rainfall shocks on the *relative* survival probability of girls to boys, by birth mother, finding that higher rainfall increases the odds that a surviving child is a girl. If households were perfectly able to smooth consumption in the face of shocks, then rainfall should have no impact on relative mortality, except through possible biological mechanisms. The fact that it does is consistent with the inability of households to perfectly smooth their consumption, since a positive rainfall shock should result in higher income. If there exists boy preference for any reason, then when incomes go down unexpectedly, it may be girls that suffer more than boys in terms of their allocations of health inputs. Then, when incomes improve, girls should benefit (see Alderman and Gertler, 1997, for a simple model in which this result emerges). Further, Rose finds that the rainfall

shock effects on the relative survival of girls is stronger in landless households than households with land, consistent with a credit market constraint story affecting landless households more.

Jensen (2000) also uses historical rainfall data to form a measure of shock for areas in Cote d'Ivoire between 1986 and 1987, and using the World Bank's Living Standards Surveys, and relates the shock measures to income declines and then to investment in children, using a difference-in-difference approach. He categorizes regions into those affected by and not affected by rainfall shocks between 1986 and 1987 and then looks at differences in the fraction of standardized weight-for-height z-scores for children 0–10 years that are under -2 (wasted), and takes the difference-in-difference as his estimate of the shock. He finds significant, positive impacts of being in a rainfall shock region for boys' being stunted and positive, significant at 10%, for girls.

More recent papers that use the exposed and unexposed birth cohort comparison include Paxson and Schady (2005) and Rukumnuaykit (2003). Both papers examine the effects of economic crises on infant mortality: Paxson and Schady looking at the collapse of the Peruvian economy in the late 1980s and Rukumnuaykit at the impact of the financial crisis in Indonesia in the late 1990s. Paxson and Schady examine year-to-year changes in infant mortality rates using Demographic and Health Survey (DHS) data. Rukumnuaykit uses different waves of the Indonesia Family Life Survey (IFLS) and also constructs year-by-year changes in both neonatal and post-neonatal mortality. She also uses hazard analysis, allowing for trends in mortality, to estimate the change in mortality for exposed birth cohorts, relative to trend. Both find substantial impacts, increases of 1.4 percentage points in Indonesia and 2.5 percentage points in Peru.

In a slightly different, though related, study, Cutler et al. (2002) use a triple difference estimator to estimate the impacts of economic crisis in Mexico on age-specific mortality of children and the elderly. Cutler et al. look at changes in mortality rates for their "treatment" groups before and after each crisis, relative to changes in mortality for the same group in the years just prior to each crisis. Subtracting the latter change controls for trend changes in mortality before the crisis. They then use prime-aged male adults as a control group, arguing that mortality for this group should not have been much affected by an economic shock. They take difference-in-differences for this control group and subtract that from the difference-in-difference for each treatment group to arrive at their triple differences. They find that mortality during the 1995/1996 crisis did rise for both children under 5 and the elderly, relative to mortality changes for prime-aged men.

What is not clear from the Cutler et al. (2002) study is whether those young children and older persons who died were fragile in the first place and died a little earlier than they would have anyway. Results of McKenzie (2006) show that less-educated Mexican households during the 1995 Peso crisis cut back their expenditures on durables and semi-durables, and cut back by much less their expenditures on basic foods, possibly to better maintain their health.

Jensen and Richter (2003) investigate the impacts of the Russian economic crisis on health of adults. They note that in 1996 there was a major crisis in the pension system, that accompanied the overall economic collapse, the consequence of which was that

many pensioners had their payments withheld (this was separate from the withholding of wage payments, which was even more common). They use a difference-in-difference regression strategy, defining dummy variables for whether the individual had pension arrears and another for the year 1996, when the pension arrears were binding, plus an interaction variable to identify the pension effect. Jensen and Richter (2003) find that calorie and protein intakes of men and women are lower if the person has pension arrears. Also they find that the probability of having a checkup in the last year is lowered, as is the likelihood of taking medication for a chronic health condition in the past week (conditional on the person having had the condition before the crisis). In terms of more chronic conditions, for men, the ADL index was worsened and men were more likely to report experiencing chest pains. Finally, Jensen and Richter look at the impact on subsequent mortality, but their regression is more like a health production function, including many health inputs, such as whether the person is obese, had experienced chest pains, drinks and smokes. Even holding constant all of these factors, having pension arrears is found to have a positive impact on subsequent mortality for men. However, one factor that reduces the impact on health is the coping mechanism reaction of these people. Jensen and Richter find that labor supply goes up and that assets are drawn down for those individuals whose pensions are in arrears.

Shocks other than economic have been examined for their impacts on health. For instance, in late 1997 there were a series of serious forest fires in Indonesia that greatly affected air pollutants in parts of Indonesia and neighboring countries. Frankenberg, McKee and Thomas (2005) use detailed data from the National Space and Aeronautical Administration's (NASA) Total Ozone Mapping Spectrometer (TOMS) data, which are daily, to measure haze from the fires, and matched those to the Indonesia Family Life Survey data from late 1997 and early 1998 using latitude and longitude data. They examine the short-run impacts of the haze on three health indicators of adults: difficulty in carrying a heavy load, whether coughing was reported in the past one month and whether general health was reported as poor; looking separately at prime-aged women, prime-aged men and the elderly (defined as over 55). Their empirical specification uses data from the 1993 and 1997 waves of IFLS. Many communities in the 1997 wave were not exposed to the haze, but many were. Frankenberg, McKee and Thomas (2005) look at the difference-in-difference between 1993 and 1997 health outcomes, between individuals in heavily exposed areas and individuals in not heavily exposed areas.

In their regression analysis they use individual fixed effects, together with exposure dummies that are timed as to whether exposure ended or started at the time of the survey or at least one month before. They find that exposure to the haze indeed did positively affect having difficulty in carrying a heavy load, and also affected coughing and being in poor health. From the time-dated exposure variables they discovered that people's health rebounded quite quickly to the end of haze exposure. The timing of exposure also helps to reassure one that one possible confounding variable, being exposed to cigarette smoke in the house, probably is not an issue. In studies of impacts of air pollution on health in the United States, it has been found (e.g., Mullahy and Portney, 1990) that areas that have greater air pollution are poorer areas, where cigarette smoking is

more prevalent. In Indonesia, where 70% of men currently smoke (Witoelar, Strauss and Rukumnuaykit, 2006), the socio-economic gradients with smoking are weaker than in the US, but more to the point, do not vary over time much (Witoelar, Strauss and Rukumnuaykit, 2006). This implies that one should not expect to see rapid recovery in health impacts of haze, if they were really due to being exposed to cigarette smoke.

Another study, Sastry (2002), finds short-run impacts of haze from the Indonesian fires on mortality of infants and older persons 65–74 in Kuala Lumpur, and on older persons in Kuching. Sastry uses aggregate city mortality data by day from vital registration systems, merged with data on haze, that is more crude than the TOMS data that Frankenberg, McKee and Thomas (2005) use and estimates count models of mortality. Like the Cutler et al. study, it is not clear whether it was the frail old and very young who died, perhaps just a little before they would have anyway.

Jayachandran (2006) also uses smoke data for Indonesia and examines, by birth cohort, whether smoke during the prenatal or postnatal periods (separately) affected the size of the cohort in the Census year 2000. Size of cohort in 2000 is used as a proxy for infant mortality, which she does not observe in her data. Unfortunately there are many very strong assumptions that are required for birth cohort size in 2000 to be an unbiased proxy for infant mortality for past birth cohorts, so that the data are not the most suitable for testing of this hypothesis. Nevertheless, she finds a strong negative effect of smoke exposure on cohort size.

Pitt, Rosenzweig and Hassan (2005) examine the impacts of indoor smoke from cooking stoves on health of adult women, especially those who cook, and their children. In many poor countries cooking with wood stoves, or using fuels such as charcoal or kerosene, is common. Use of these fuels causes a lot of smoke, which may be quite harmful to health, especially respiratory health. They use survey data that contain detailed time allocation information, from which they can determine who in the household is cooking, and therefore most exposed to the smoke, and for how much time. They also are able to measure the permeability of the walls and roof, which will affect how well air circulation will spread. Unfortunately, they do not have any direct measures of smoke or particular matter in the air of the house (or kitchen of the house), nor do they have direct measures of how clean are the lungs of individuals living in the household.

They measure self-reported coughing, having difficulty breathing and having a fever, which they use as symptoms of respiratory problems. The discussion in Section 3.6 is relevant here. While there is likely to be real information in these self-reports, they may be biased in that higher-SES women may be more likely to have gone to the doctor and to have been properly diagnosed. Indeed, in Tables 2 and 4 of Pitt, Rosenzweig and Hassan (2005) the partial correlation of schooling with having respiratory or intestinal symptoms is positive and significant for women over 16 years.

Pitt, Rosenzweig and Hassan (2005) estimate a conditional health "demand" function for having one or more respiratory illness symptoms, as a function of time spent cooking, age, sex, schooling, per capita expenditure, having a permeable roof, walls, and whether the kitchen is outdoors. They use three different estimation techniques: household random effects, household fixed effects, and household fixed effects plus in-

strumental variables. As instruments they use variables that measure whether the woman is the spouse of the household head, or his daughter-in-law, the number of daughters-in-law in the household, and interactions between them. They argue that separate spheres of activities exist within the household, with the senior spouse being primarily responsible for the cooking. A key assumption is that this set of relationships with the household head affects the allocation of time to cooking but has no direct effect on health. While this assumption is plausible, it is not subjected to testing. There are also plausible conditions under which it will be violated. For example, if one's relationship with the household head affects spending on individual-specific health inputs, then there will be direct effects on individual health outcomes.

Pitt, Rosenzweig and Hassan (2005) do find that cooking time does significantly raise the likelihood of having respiratory symptoms, especially for women over 16 years. On the other hand, it does not affect intestinal symptoms, which is good for their argument, because if it had, one might have argued that the correlation with respiratory symptoms was spurious, based on omitted variables not controlled for by taking household fixed effects, or by the instruments. They also find that the permeability of walls and roofs has no effect in ameliorating the negative health impacts of smoke. Finally they find that among the determinants of time spent cooking, it is the inherent healthiness of the woman that matters, with less healthy women being more likely to work in the kitchen. If poor health has a bigger negative impact on agricultural productivity than productivity in cooking, then it would be sensible for the woman in the household who has the poorest health to spend the most time cooking in the kitchen.

4.5. Health and consumption smoothing

Since the work of Rosenzweig (1988) and Townsend (1994), empirical tests of consumption smoothing and the methods that households use to smooth their consumption have blossomed (see Dercon, 2004, for a recent compilation of papers). As an example, Foster (1995) uses data on flooding in an area in rural Bangladesh having a very good population surveillance system run by the International Centre for Diarrhoeal Disease Research, Bangladesh, to examine how young child standardized weights were affected. Foster has two periods of data available, one before the flood and one after, and looks at differences over time. The geographic area is small, so there is not regional variation in how hard the floods hit. However, households differ in ex ante assets, particularly land-owned, with some households being landless. Asset ownership arguably may condition how badly the flood affected children's weights. Foster builds a dynamic structural model of weight change under expected utility maximization and derives Euler equations that he estimates. The key point of the model is that changes in child weight may respond to household borrowing, and average (and variance of) village borrowing if interest rates vary within and across villages, perhaps because of segmented markets.[25] Indeed Foster finds that child weight gain after the flood is lower

[25] Foster uses borrowing, and not interest rates, because the latter are not observable in his data.

if the individual household borrows, conditional on mean village borrowing, but that this is true only for landless households. This result is consistent with many studies that find different interest rates for different households within a village, based on cost of lending to each household.

Foster's test is not specifically a test for income pooling within the village, certainly not like Townsend. However, of interest here is another covariate, the change in the proportion of days on which diarrheal illness was reported. While diarrheal disease incidence would normally be considered as measured with error, in these data it is likely measured much better. This variable may be a better measure of a health shock on child weight. Foster finds that when the proportion of diarrheal disease days is reduced, weight gain increases, as one would expect. The impact is greater for children from landless households, perhaps because they have a more frequent incidence of diarrheal disease days (and thus a higher marginal product of disease day reduction), or perhaps because they face more stringent credit market constraints.

Foster is cognizant of the potential endogeneity of changes in diarrheal disease days. Shocks which increase diarrheal disease days may also lower weight gains. Thus even though weight is specified in first differences, the diarrheal disease variable may be correlated with unobserved characteristics in the model. To correct for this, Foster uses initial-period diarrheal disease days as an instrumental variable for changes in diarrheal disease days. So long as diarrheal disease is not serially correlated, this IV is consistent. However, it seems likely that diarrheal disease might be serially correlated, in which case this strategy is not sufficient to obtain unbiased estimates.

Gertler and Gruber (2002) present a Townsend-type test of income pooling in response to a health shock, using Indonesian data. The nature of their test is simple; they regress change in variables that might be affected by a health shock (non-medical consumption, household head's labor supply, head's imputed earnings, imputed earnings of others in the household) on changes in household head's health, plus a series of other characteristics in the initial period (age, education, marital status). Thus the specification is a mix of a change and a growth equation.

Gertler and Gruber use three different measures of health shocks for the household head: illness symptoms, chronic symptoms, and an index of activities of daily living (ADLs). They find that positive changes in the ADL index are associated with positive, large changes in non-medical consumption and hours of work of the household head. Illness and chronic symptom changes, on the other hand, are not associated with changes in consumption or head's labor supply.

Of concern here is that both omitted variables may be causing changes in both variables, that there may be direct reverse causality, and measurement error. To explore this in more detail, it is useful to think about what we want to measure and then what we may be measuring in the first difference specification. We would like to measure an exogenous health shock when taking first difference of our health measure. There are likely to be differences in stocks and flows in this regard. For instance, reported health symptoms, one of the types of health measures used by Gertler and Gruber, are likely to be flow variables, whereas chronic disease symptoms and measures of activities of daily

living (ADLs) are stocks. For a stock, first differencing will result in a health investment equation. The health investments will reflect choices of endogenous health inputs and behaviors, as well as exogenous unobserved healthiness. The endogenous health inputs and behaviors, in turn, will reflect income and price shocks, in addition to exogenous health shocks. Thus if the estimating equation does not have good proxies for price and income shocks, those are likely to be important omitted variables that would cause bias to the resulting estimates. Gertler and Gruber do include community fixed effects in their estimates. Since there are only two periods of data that they use, this will account for price and other community-level shocks. However, feedback effects from income shocks to health are not taken into account.

Gertler and Gruber are aware of these issues and test for omitted variables bias and feedback effects, but the tests are not very satisfactory. For instance they look at whether change in the head's ADL index matters differently for non-medical consumption if the head is a worker or not. The idea is that if the causation goes from health to consumption, then worker's ADLs should have a larger effect than non-workers. However, if, as Mirrlees (1975) pointed out long ago, workers get allocated nutrients on a favorable basis in order to make them more productive, then one might see a positive correlation between improvement in consumption and in health of workers that would bias upwards a health-consumption relationship even in changes.

Measurement error is acknowledged, but nothing is done about it. As we discuss, systematic measurement error may be differenced out by first differencing, but at the expense of random measurement error. The symptom data are likely to be more prone to both systematic and random measurement error.

Dercon and Krishnan (2000) write down a much more structural model than do Gertler and Gruber (2002) and estimate it using panel rural Ethiopian data. They relate changes in body mass index of adults to shocks of various kinds: regional rainfall shocks; household-level exogenous shocks, such as shocks to crops and livestock; recent household deaths (as a shock to labor supply), and the number of days a person was unable to work because of illness. The days a person could not work, Dercon and Krishnan view as a health shock variable; they are interested in examining its impact on own health as measured by BMI. Dercon and Krishnan treat own days lost to illness as endogenous, but not so whether an adult in the household died since the last interview, breastfeeding or pregnancy status. Thus they are not consistent in their treatment of health outcome and input variables, all of which can appropriately be considered as endogenous. This is a weakness in their approach. Change in lagged BMI is also used in the BMI change equation, and is treated as endogenous, identified by further lags in BMI, and the number of sick days taken in the last 5 years. This identification strategy, which is like the Arellano–Bond estimator in dynamic panel models, can work provided that the assumption is made that the time-varying error term in the BMI change equation is not serially correlated. If the change in adult BMI is related to early life events as suggested by the discussion in Section 4.1, this assumption may fail.

Days of lost work due to illness is recognized as being potentially correlated with omitted time-varying variables in the BMI change equation. Dercon and Krishnan in-

clude only non-gastrointestinal illnesses, because the latter might affect BMI directly. This is problematic, however, in that the severity of even non-gastrointestinal illnesses, such as measles for instance, may be affected by levels of nutrition.[26] Instead of using days of lost work due to illness, Dercon and Krishnan use the residual from an individual fixed effect regression of days of lost work on lagged BMI, lagged lost work days, individual and household characteristics and village/time dummies. The problem with this specification is that there may be other unobserved time-varying shock variables at the individual and household levels that are not controlled for that cause both nutrition and health shocks. While Dercon and Krishnan discount this, they have no evidence in support of their view, which does not seem convincing.

One other issue with the Dercon–Krishnan work is the health variable they use for shock: days of work lost to illness. This variable, as noted in Section 3, is just as much a labor supply variable as it is a health variable. Substitution effects would suggest that it is likely that persons with low values of time will be more likely to take off work from a given illness than someone with a higher value of time, while income effects would be reversed. The dominance of income effects may explain some of Dercon and Krishnan's results, specifically that it is for lower-income women in the south that the illness shock is most related to lowered BMI.

5. HIV/AIDS

In this section, the spotlight is placed on HIV/AIDS which has arguably posed the greatest global public health challenges over the last quarter century. It is estimated that in 2006 almost 40 million people were infected with HIV/AIDS and nearly 3 million died that year (UNAIDS/WHO, 2006; UNAIDS, 2006). The vast majority of the burden has fallen on developing countries, particularly sub-Saharan Africa, where it is estimated that 25 million people are infected with HIV/AIDS and over 2 million died in 2006.

We focus here on empirical evidence at the micro level which speaks to the impact of this major epidemic on the well-being of populations in developing countries. We have chosen to highlight HIV/AIDS because it is important, because there is a substantial literature on the economic impact of the disease and because the literature provides additional insights into understanding how poor health and reduced life expectancy affects the well-being of individuals and their family members.

The next subsection discusses measurement and methodological issues. It is important to note that much of the evidence on the impact of HIV/AIDS on economic outcomes has relied on information about the death of a prime-age adult to identify this effect. While many of these deaths in southern and eastern African populations are

[26] Scrimshaw, Taylor and Gordon (1968) showed that child death rates were much higher for children who had both diseases such as measles and were undernourished, compared to those who had measles only. The effects were found to be interactive.

likely to be associated with HIV/AIDS, the evidence should be interpreted as measuring the effect of the death of an adult. As more population-based socio-economic surveys include tests for HIV status, this literature is likely to yield substantially new insights into the links between HIV/AIDS and population well-being.

There is little experimental evidence on the impact of HIV/AIDS, which is unfortunate in view of the large number of treatment trials that have been conducted. We discuss the non-experimental evidence, focusing first on the impact of parental death on child schooling which has received much of the attention in the literature. The final subsection discusses the impact on other indicators of well-being. Our exploration of the micro evidence provides some suggestions for why forecasts in the macroeconomic and general equilibrium literatures of economic collapse because of HIV/AIDS have not been borne out. (See, for example, Bell, Devarajan and Gersbach, 2006.)

We do not attempt to review the literature on the epidemiology of HIV/AIDS or the development and availability of treatments and vaccines, but note several themes in common to that and the economics literatures. (See, for example, Bertozzi et al., 2006.) For example, the high prevalence of HIV/AIDS in sub-Saharan Africa reflects the influence of prior health, particularly untreated sexually transmitted diseases (STDs) which were identified as a key mechanism through which infections were transmitted as early as the late 1980s (Cates and Bowen, 1989). The combination of weak public health infrastructure and inadequate understanding of the risk factors, as well as potential consequences of HIV infections, has compounded the impact of the disease in sub-Saharan Africa. This contrasts sharply with the evidence of substantial behavioral change as well as broad availability of therapies that has occurred in the United States and developed countries (Ahitov, Hotz and Philipson, 1996).[27]

5.1. Methodological issues

Three methodological issues have limited progress in this field. First, most samples used in analyses are not random and span a short time horizon. Second, until recently, HIV/AIDS status was seldom recorded in survey data. Third, selectivity of those infected and those who die has often been ignored. We discuss each in turn.

5.1.1. Sampling

Many of the early studies of HIV/AIDS used samples drawn from "sentinel sites," which are often public ante-natal clinics, and samples were comprised of pregnant women.

[27] We could highlight other pernicious diseases such as malaria or an influenza pandemic which some predict will sweep through the world in the coming years. Many of the issues that arise in the examination of HIV/AIDS are relevant for these diseases including weak health infrastructure, inadequate adoption of relatively simple and well-understood technologies, co-morbidities and compromised immune systems associated with prior health insults.

Putting aside the difficulties of using pregnant women to infer population-level infection rates for all women, let alone men, this approach is likely to yield a selected sample which will potentially overstate the incidence of HIV. Specifically, in poor countries, many women give birth at home and are thus not counted in sentinel sites. These women tend to live in rural areas, are older, less well educated and poorer than the average woman. Early in the pandemic, these women tended to have lower rates of infection and so estimates based on sentinel site data overstated the prevalence of HIV/AIDS. Zambia provides an example. In 2001, using sentinel site data, it was estimated that 21.5% of the adult population was infected with HIV/AIDS (WHO and UNAIDS, 2003). The Zambia Demographic Health Survey (DHS) conducted in 2001 tested HIV status on a random sub-sample of the population and estimated prevalence to be 15.6% (Central Statistics Office et al., 2003). These estimates are also potentially biased because 20% of women and 25% of men refused to be tested. High refusal rates have been the norm in most household surveys that measures HIV status. Whether those who refused are more or less likely to be HIV positive is unknown, although the highest rate of refusal was among urban men suggesting there may be selection on HIV status. Similar evidence in several other sub-Saharan African countries has led the Population Division of the United Nations' Department of Economic and Social Affairs to lower prevalence estimates (WHO and UNAIDS, 2003).

In order to estimate behavioral relationships using sentinel data, it is necessary to assume that the probability a pregnant woman attended an antenatal clinic in the sentinel site is independent of her own HIV/AIDS status and that of her family members. This is a strong assumption. If higher-income women are more likely to attend antenatal clinics and if the incidence of HIV/AIDS is higher among those with more income, the assumption will be violated. Estimation of behavioral relationships using survey data suffers from similar problems if those who refuse to be tested are not random.

5.1.2. Measurement

Few socio-economic surveys have actually measured HIV status of respondents although this is changing rapidly with the availability of simple tests using mouth swabs. Many of the early studies followed the lead of the first Kagera Health and Development Survey (KHDS) and treat deaths of prime-aged adults as a proxy for HIV/AIDS (Ainsworth et al., 1992). During the first wave of the KHDS between 1991 and 1994, it was thought that the Kagera area in rural, northwest Tanzania had much higher rates of infection than the rest of Tanzania at that time, and it seemed plausible that the majority of prime-aged adult deaths were AIDS related. Proxy-reports on symptoms of those who died were consistent with AIDS being the major cause of these deaths (Ainsworth et al., 1992). A recent survey from South Africa that does collect HIV status suggests that relying only on adult deaths misses part of the reaction to HIV/AIDS because part of the behavioral reaction begins when the household recognizes that a member has HIV/AIDS, which can be well before death (Linnemayr, 2005).

The KHDS has one other feature that is very relevant to analysis. Because adult deaths are such a rare event, the sampling in KHDS was choice-based, with different probabilities of being in the sample depending on the 1988 Census estimates of adult deaths and depending on whether a village census showed there was a recent death in a particular household. The benefit of this sampling strategy is the power that comes with sufficient deaths to test hypotheses. The cost is that, with cross-sectional analyses, special estimation procedures are required to obtain unbiased estimates (Manski and Lerman, 1977).

5.1.3. Estimation

Many empirical models have related some indicator of well-being to whether a prime-age adult in a household has died or whether a child in the household is an orphan. These estimates cannot be given a causal interpretation since adult (or parental) death is not randomly assigned. Including a household fixed effect will yield unbiased estimates if there are time invariant household-level characteristics that underlie the death of the adult. However, if there are time-varying factors, such as variation in household structure, then these estimates will be biased.

As an example, an orphan may move into a new household. Including a household fixed effect in the destination household does not solve the problem of unobserved heterogeneity in parental death. Including a household fixed effect in the origin household raises a different source of selection bias – the selection of household to which an orphan moves, if the orphan does move. An alternative approach is to estimate models with child fixed effects in which case identification is based on time-varying exposure to the adult's death. These selection issues are complicated and the biases associated with them have not been fully explored in the literature. There is little guidance on the possible direction and magnitude of biases due to selectivity which are likely to vary depending on the context and period covered by a particular study.[28]

5.2. Parental death and child schooling

The issue that has received the most attention in the literature on the economic consequences of HIV/AIDS has been the impact of the death of a prime-age adult on child schooling. We start with an examination of that literature focusing on evidence based on

[28] Some studies report the association between SES and adult death is positive (Ainsworth and Semali, 1998); others indicate the association is zero (Chapoto and Jayne, 2005). De Walque (2006) reports no correlation between education and testing positively for HIV using DHS data from 5 sub-Saharan African countries (Burkina Faso, Cameroon, Ghana, Kenya and Tanzania) but notes that members of households with more assets are more likely to refuse to be tested. He also documents that the association between education and HIV status among young adults in southern rural Uganda switched from zero (or slightly positive) in 1990 to significantly negative in 2000 which he attributes to the impact of an anti-AIDS information campaign that had its biggest impact on young, better educated women (de Walque, 2004).

longitudinal data which are better suited to addressing concerns regarding unobserved heterogeneity and selection.

Ainsworth, Beegle and Kodi (2005) use the KHDS to examine the relationship between death of a parent and both attendance at school by 7 to 14-year-old children and the hours they spent at school in the previous week. They distinguish children who are paternal orphans (10%), maternal orphans (5%) and dual orphans (both parents are dead, 2.5%). They find a significant negative association between being a maternal orphan and school attendance for younger children (7–10 year-olds) but not for older children, suggesting the association works through delayed enrollment. Being a paternal orphan is not significantly associated with school attendance and dual orphanhood is no different from being a maternal orphan. Having any adult death in the last six months also has a negative association with attendance, again only for younger children.

The hours in school models are estimated conditional on being in school and are specified in terms of changes in hours over time. They find a negative association with maternal death but only for girls. They explore whether these changes precede, coincide with or follow the adult death and find that for both sons and daughters, hours in school decline during the period 4 to 6 months prior to a death and then return to prior levels after the mother's death. This indicates that the time of children is shifted away from school when the mother is very ill but this shift is only temporary.

Ainsworth, Beegle and Kodi (2005) measure the short-run relationships between school and parental death. Using a follow-up survey of the KHDS conducted in 2004, 13 years after the first wave, Beegle, De Weerdt and Dercon (2007) explore the longer-run relationship between orphanhood and attained schooling by age 19. Of the individuals who were children (aged 0–15 years) in the first wave, and alive in 2004, 80% were recontacted. Older children in the first wave, those from higher-income households, and those with better-educated parents were less likely to be recontacted if they were alive.

Beegle et al. restrict analysis to children who were not orphans in the 1992–1994 baseline and compare the educational attainment of those who become orphans with those who did not. They find children who were orphaned have significantly less schooling (0.9 years of schooling) than other children. The estimate can be interpreted as causal only if there are no unobserved characteristics that affect the probability a child becomes an orphan and the child's schooling. This is a strong assumption. Beegle et al. probe the assumption and report the height of 15-year-olds who became maternal orphans is 1.3% less than the height of other children not orphaned. Unless these differences are driven by the impact of maternal death that occurred during the first few years of life, the evidence suggests that children who were orphaned during the hiatus between the survey waves are not randomly selected from the population, but tend to have lower human capital than other children. It is not clear how much of the difference in schooling at age 19 should be attributed to this selection. (Beegle et al. also adopt a propensity score approach which, recall from above, assumes mortality selection is fully captured by observed characteristics.)

Two additional results are worth highlighting. First, the negative association between orphanhood and schooling is observed only among children who were not enrolled in

school when the mother died. Second, there is no relationship between orphanhood and schooling among those children who moved away from the natal home. The first result does not appear to be driven by wealth although the second might be. Both results suggest that unobserved characteristics including parental preferences, short-term resource constraints and child ability might play a central role in explaining the estimated negative association between orphanhood and schooling.

Two related, important studies by Evans and Miguel (2007) and Case and Ardington (2006) use panel data from western Kenya and Kwa-Zulu Natal, South Africa, respectively. Both studies use a sample of children who are initially not orphans, with extensive information on the households prior to orphanhood. Both studies fully exploit their data to minimize the impact of potential biases due to selectivity of, for example, adult mortality and child mobility. Evans and Miguel use child-level fixed effects, and Case and Ardington use several different strategies, including both family and child fixed effects.

Evans and Miguel follow up children who were in the Miguel and Kremer (2004) worms study in western Kenya four years after the 1998 baseline. The sample is children age 5 to 18 years old, enrolled in primary school and not orphaned at baseline. Some of the analyses use sub-samples that have more baseline information, which requires that the child attended school on the day the worms study collected household-level information (which may impart some choice-based sampling bias). Around one quarter of the sample has missing or unreliable information on parental death or schooling.

Evans and Miguel find a significant, negative association between being an orphan and school attendance, with maternal deaths being associated with bigger reductions in schooling than paternal deaths. They report that declines in education begin one to two years before the parent's death and cumulate until three years after the death (which is the longest time period they can examine). There are no differences between sons and daughters, between younger and older children or between children from households with more or less assets. They do find that orphans who scored higher on school achievement tests in the 1998 baseline were more likely to be enrolled in school in the 2002 follow-up.

Building on prior work by Case, Paxson and Ableidinger (2004), Case and Ardington use very rich demographic surveillance data from the Africa Center in Kwa-Zulu Natal, South Africa, to which they added two rounds of a socio-economic survey they organized, in 2001 and 2004. They examine impacts of orphanhood on current school attendance, years of completed schooling and log of monthly educational expenditures. Orphanhood rates are high in these data, 9% of children are maternal and 15% are paternal orphans. Case and Ardington find strong, negative associations between mother's death and years of completed schooling, current enrollment, and education expenditures. They do not find significant associations with a father's death and the impact of being a double orphan are approximately equal to being a maternal orphan. Like Evans and Miguel, Case and Ardington find a cumulative impact of maternal death on child years of schooling, that is increasing in time since the mother died. They also find a significant interaction of the mother's death with whether a female pensioner lives in the household.

Case and Ardington conduct several analyses in order to determine whether the results are driven by unobserved heterogeneity. In some models, they include destination household fixed effects; in others, they include extensive household characteristics to control for potential omitted variables bias. They report that completed schooling at baseline, in 2001, is negatively associated with whether the mother was dead in 2001, but also knowing whether the mother was dead in 2004 does not help predict schooling in 2001. This is reassuring since it suggests that the estimates are not contaminated by time-invariant omitted variables. They proceed to use first differences (which is equivalent to including a child fixed effect) and report that maternal death between 2001 and 2004 is negatively related to the change in years of completed schooling as well as changes in enrollment between the waves. There is no association between paternal death between the waves and changes in the school outcomes of the child.

In a different study, Yamano and Jayne (2005) use three waves of a nationwide panel of rural households in Kenya. The first wave, fielded in 1997, was a farm household survey that did not collect any child schooling information, but they went back in 2000 and 2002 and did add child schooling information, as well as information on adult deaths and timing of those deaths. Importantly, they have information on schooling of children even if they left the household, so that potential biases from migration of children are avoided. The disadvantage of these data is that parental identification is not available except for children of the household head, so they use as their adult death variable a marker for any prime-aged adult in the household who died, hence the children in households with adult deaths are not necessarily orphans. Although households that moved are not followed, household attrition is low suggesting that selection biases associated with migration are not likely to be large. Yamano and Jayne (2005) use both household and child fixed effects estimators in their analyses, as well as models with household and individual controls, to control for mortality selection.

They find that without using any fixed effects, children from households with low initial wealth in 1997 and a subsequent household death experience lower enrollment. The impact starts before the death for girls, while for boys the largest impact comes right after the death. When adding household fixed effects, the impact for girls of a future death remains, but the impact on boys disappears. They also find that lagged regional HIV prevalence rates of pregnant women (from sentinel sites) are negatively associated with enrollment for boys, although the estimates become imprecise after taking household fixed effects, and no other community covariates are included in the model making omitted community variables a potential issue.

Several common threads emerge from these studies. First, there appears to be a negative impact of being orphaned on child schooling, particularly if the child's mother died. These effects tend to be larger for daughters, especially during the period prior to the death. This may be because the daughters are caring for the ill household members, or substituting for the mother in home production, rather than going to school. The evi-

dence on whether these effects are larger or smaller for children in poorer households is ambiguous.[29]

5.3. Parental death and health of other household members

This subsection discusses the relationship between death of an adult and health of other family members. The next subsection explores the relationship with indicators of economic well-being at the individual and household level.

The elderly in a household may bear much of the brunt of an adult child having HIV and dying of AIDS. Knodel (2008) shows that in Thailand and Cambodia over 60% of adult children who died of AIDS lived with an elderly parent before death and that for over 90% of these adult children, a parent supplied at least some care during the last days of illness, and for over 70% the parent was the main care giver. Such care by the elderly for their adult children may well have health consequences on the elderly through creating greater stress for the older parents, not having enough younger members to care for the parents, or other causes.

Ainsworth and Dayton (2003) use the KHDS to examine the impact of the death of a prime-age adult on health, as measured by BMI, of other elderly household members (age 50 years or more). They relate changes in BMI to whether a death occurred before the survey, around the time of the survey or after the survey, controlling change in demographic characteristics, household composition and assets as well as community health infrastructure. They find very little evidence of an association between BMI of the elderly and an adult death although there is a suggestion that, among better-off households, BMI of the elderly tends to decline slightly prior to the adult's death and rise after the death. Overall, the health of the elderly does not appear to be related to the death of a prime-age household member.

Ainsworth and Semali (2000) examine the short-run impacts of adult deaths in the KHDS on child health measured by height-for-age, weight-for-height and proxy reported illness of young children (under 60 months). They report negative, significant associations between both mother's and father's death and child height, but the associations are not significant when child-level fixed effects are included in the models. There is some evidence that a recent adult death (not necessarily of a parent) in the

[29] The longer-run implications of being an orphan are investigated by Yamano (2006). He uses data on adults living in rural Kenya to examine whether completed schooling is associated with being an orphan by age 15. Orphanhood is measured retrospectively. While the models include respondent characteristics and community fixed effects, the estimates need to be interpreted cautiously since they are likely to be contaminated by unobserved characteristics including time-invariant parental characteristics and mortality selection. He reports that cohorts who went to school in the pre- and immediate post-colonial periods lost an average of one year of schooling if they were maternal orphans by age 15. These were cohorts whose maternal deaths came *before* HIV/AIDS was prevalent. Among younger cohorts, whose mothers were exposed to HIV/AIDS, there are much smaller schooling gaps between orphans and other adults. These differences across cohorts might indicate differences in selection, differences in period effects, or they may indicate that some families and communities are more resilient than others in the face of parental death.

last six months is associated with reduced height-for-age among children from poor households. As Ainsworth and Semali point out, in the fixed effects specification, the adult death indicates whether a child became an orphan between the survey rounds and very few children in the KHDS were orphaned during the first wave of fieldwork.

5.4. Parental death and household resources

A central issue regarding HIV/AIDS is the effect on both the household economy and the economy in aggregate. Although much has been written about the topic, there have been rather few reliable studies at the micro level.

In an important study, World Bank (1997) documents that households in the KHDS reporting a prime-aged adult death in the last year had 7.5% lower expenditures per adult equivalent (including durables) than households without an adult death. Households with a death had much higher funeral and health care expenditures just before and at the time of the death. This, together with the lower overall expenditures meant that other expenditures had to be lower. Of course this does not control for mortality selection, which in this case could result in an underestimate since the Kagera households having adult deaths tended to be higher-income households.

Other studies have examined the impacts of adult deaths on per capita consumption (Beegle, De Weerdt and Dercon, 2008, and Case and Ardington, 2006); on incomes (Yamano and Jayne, 2004; Chapoto and Jayne, 2008); on assets (Yamano and Jayne, 2004; Case and Ardington, 2006; Chapoto and Jayne, 2008) and on labor productivity (Fox et al., 2004). As for the child schooling-orphan studies, there are differences in the controls that are available for mortality selection and in the magnitudes of the results, but most studies report a negative impact of an adult death on these outcomes.

Beegle, De Weerdt and Dercon (2008) use the first and second waves of the Kagera data to estimate the impact of a prime-aged adult death in a household between the 1991 and 2004 waves on the growth of household per capita expenditures, from the initial interview (usually in 1991) to 2004. The adult deaths are categorized into dummy variables showing male or female deaths in one of three five-year intervals: 1991–1995, 1996–2000 and 2000–2004. They are set to one if the individual was living with that person at the time of death. This specification amounts to a difference-in-difference estimator. While unobserved factors related to the *level* of consumption are removed in this way, there might be omitted variables that are correlated with *changes*, particularly with the death of an adult member. To control for some of these omitted factors, Beegle et al. control for initial period household fixed effects. If subsequent adult deaths are correlated with characteristics of the initial household, then this procedure will control for the selectivity of households that had an adult death. This procedure induces a different problem since comparisons will be drawn between members of the original household who stay and those who move. To address the possibility that other shocks may be corre-

lated with consumption changes, they also control for self-reported agricultural shocks that affected each household.

Beegle, De Weerdt and Dercon (2008) find that recent deaths are associated with 7% lower consumption growth during the first five years after the death, but that consumption seems to recover after that, albeit not to the pre-death level. A 7% decline in consumption is not trivial, but is small in comparison with changes in consumption associated with large economic shocks such as the 1998 financial crisis in Indonesia (Frankenberg, Thomas and Beegle, 1999) or the large depression in Peru in the late 1980s (Glewwe and Hall, 1994). The consumption response to adult death does not appear to be related to whether the household is initially poor or how long the adult was ill prior to death. The evidence is suggestive that families respond to minimize the impact of an adult death on consumption, although what mechanisms are used is not explored.

Case and Ardington (2006) look at the impact of adult deaths in their South African sample, on log per capita expenditures and on assets owned, using the same specifications as they used to analyze orphan schooling. They find a significant negative impact of a father's, but not mother's death on both outcomes before they first difference. For assets, they are able to take first differences (regressing the change in assets on whether there was a death between the survey rounds), thereby eliminating time-invariant omitted variables, which might confound the estimates. The coefficient on assets is negative but halved in magnitude and not significant, suggesting the omitted variables bias is substantial in the models specified in levels.

The relationship between adult death and income is explored in Yamano and Jayne (2004) and Chapoto and Jayne (2008). Yamano and Jayne (2004) measure the impacts of adult deaths in a rural panel in Kenya on a number of farm income and input variables. The data were collected in 1997 and again in 2000 on a sample of 1500 farms throughout Kenya. Because the sample is random, deaths are very uncommon. Only 46 of the households reported a male adult death. Of those, only 27 of the deaths occurred between 1997 and 2000 – the time frame examined by Yamano and Jayne. (There were 37 female deaths reported in the survey.) The small number of "treated" households raises questions about the power of the tests.

Yamano and Jayne (2004) use a difference-in-difference estimator to estimate the impacts of adult deaths from 1997 to 2000. They have "baseline" information in 1997 and can split the sample into treatment and control groups based on whether an adult death occurred, or not. They regress the change in a dependent variable on whether a death occurred in the 3-year period, plus a set of village dummy variables, to absorb variation in prices, weather and other village-level characteristics. Fixed household characteristics are differenced out.

They find no effect of an adult death on the area cultivated for cereals or root crops. Evidently food crop areas are considered inviolate by these Kenyan farmers. However there is a significant decline of almost an acre in crops with high potential returns, such as coffee, tea, sugarcane, fruits and vegetables, when a male head of household dies (though not for deaths of other family members). When they examine the impact on gross and net farm output, they find that a death of a male head is associated with a

decline in both gross and net farm output, but not on output per acre. Apparently households are choosing to reduce the land that is cropped, particularly for non-cereal crops, but there is no change in the intensity of input use and hence no change in yields. Neither, apparently are farmers choosing to reduce lower-productivity land, which would have resulted in a rise in yields. Again, however, the very small sample size of households with an adult death is likely to be affecting the possibility to observe such subtle impacts.

Yamano and Jayne also look at what happens to assets of different types and to off-farm income. They find that off-farm income does respond negatively to the death of a male head, while the value of small ruminants declines with the death of women who are not spouses of the head. This might be because such woman may take responsibility for the oversight of such animals. Finally, they find that the impacts of death of male heads of households are not uniform, but occur primarily among households in the bottom half of the initial 1997 asset distribution.

In a related study, Chapoto and Jayne (2008) look at a panel of nationally representative, rural Zambian households in 2001 and 2004. They report a high correlation of 0.84 between a household indicator of adult death and a community indicator of HIV prevalence, reported from sentinel sites, thus verifying the usefulness of reported adult mortality as an indicator of HIV exposure. Chapoto and Jayne also have a relatively small number of households that had adult deaths, but a substantially larger number than in the Yamano and Jayne (2004) study. In this case there were 90 households that had death of a male household head, and substantially more with any adult death (over 170 with a male death and over 220 with a female death).

Chapoto and Jayne (2008) estimate their models in household first differences, which will take out household time-invariant unobserved characteristics that may cause mortality selection. However, they go further. They have data on area rainfall, which they convert to rainfall deviations. In 1994/1995 there was a major drought that affected Zambia. Chapoto and Jayne argue that that shock might have helped induce migration which could affect mortality by 2004. They interact these rainfall shocks with age cohort, since the 1994 drought would have affected different age cohorts differently. These are the instruments that they use to predict adult deaths from 2001 to 2004. They find that the rainfall deviations interacted with age cohort are highly significant predictors of adult mortality in the household. Having instruments, they compute Wu–Hausman test statistics and find that adult mortality is not endogenous when the models are estimated in first differences. Hence using household first differences seems to "solve" the mortality selection issue in this study. They also use Wooldridge's (2002) inverse probability weighting estimator to try to correct for attrition (using interviewing team dummies for identification), but like Yamano and Jayne (2005) find the results to be very robust to this.

Chapoto and Jayne (2008) find that adult deaths do result in declines in household size, but less than one for one. They find that death of a male head is associated with a decline in total area cultivated, including a decline of area in cereals and in crops with high potential returns (in Zambia, these are cotton, coffee, tobacco, sunflower,

vegetables and fruits). Death of a woman who is not a spouse is associated with less land in root crops. This makes sense because the main root crop is cassava, which is intensive in female labor for processing and is highly perishable once harvested; hence less available female labor will make it more difficult to grow. Unlike in Kenya, the gross value of production, either total or per hectare, is invariant to adult deaths, even of the male head. The value of livestock, however, especially small animals, but also cattle, are sensitive to the death of the male head and females for small animals. Off-farm income, however, is not sensitive to adult deaths. When they interact adult deaths with a measure of poverty status (being in the bottom half of the distribution of assets at baseline in 2000), Chapoto and Jayne (2008) find that the poor are more likely to reduce land area upon death of the male head, but gross output per cultivated hectare rises, particularly for the poor. This suggests that the poor are intensifying their agriculture with the death of a male head. A different measure of poverty in Zambia is land size. Chapoto and Jayne (2008) find that households with more land reduce their land area by more upon death of a household head, while these households do not intensify their input use, which results in lower output per hectare.

In sum, the results in both the Yamano and Jayne (2004) and Chapoto and Jayne (2008) studies suggest there is a large decline in income or output when an adult dies. This is in contrast with the early claims of HIV/AIDS resulting in economic collapse. However, family responses to the death of an adult differs across contexts. Yamano and Jayne indicate there is a decline in the area allocated to non-cereals but no change in yields in Kenya. Chapoto and Jayne (2008) demonstrate a small decline in area planted, but no decline in total output, in Zambia.

Fox et al. (2004) use unusually detailed data on daily tea picking by workers on a tea estate in Kenya to examine the impact of being near the end of life with HIV/AIDS on productivity. The tea workers on this estate are paid piece rate for plucking tea leaves and the company provided detailed individual records on worker output over several years. The company also owns a local hospital which cares for sick workers and their family members. From the hospital records, Fox et al. obtained names of workers who had died from HIV/AIDS as diagnosed by a medical practitioner or who were hospitalized and close to dying, again from diagnosed AIDS. The sample size is small: only 51 workers who had died. If workers went home to die of AIDS, as documented by studies such as Ainsworth, Ghosh and Semali (1995) and Chapoto and Jayne (2008), they would not be included in the sample as the hospital would not have their records. This could result in some (unknown) biases in terms of who is in the sample. Fox et al. obtained daily work records, both days worked and daily kilograms plucked, for these workers for the four years prior to death. They then matched workers who were working in the same field as the worker who died and obtained historical productivity records for the matched workers. The matching had to be done manually because of limited access to the company records. The matched sample is not perfect: there are more males, older workers and more experienced workers.

Fox et al. (2004) find that around the time of death, those who died picked about 20% less than the matched sample who were not infected with HIV/AIDS. The productivity

differences emerged as being significant about 1.5 years prior to the death when there was a 10% productivity gap. There was no productivity gap 3 years prior to the death. In addition to these productivity differences, there were important differences in days worked. In the two years before death, AIDS-infected workers had double the missed days of work as the matched comparison workers, reaching 70 days missed in the year before death.

Fox et al.'s study of piece-rate workers on plantations provides evidence on the effect of HIV/AIDS on the productivity of individuals rather than family responses on small-holder plots. This is important since, as noted by Pitt and Rosenzweig (1986), in a recursive farm household production model, illness of a family member will have no impact on farm output or productivity. Specifically, if family and hired labor are perfect substitutes and if farmers are price takers for piece rate wages (or wages per efficiency unit of labor), then when a household member falls ill, the household will hire in labor to replace the ill person and there will be no change in demand for farm inputs, supply of output or farm profits (since those choices are driven by prices and technology which does not change when a household member is ill). Full income will decline because the hired worker is paid. Clearly, the combination of studies that examine the impact of HIV/AIDS on piece rate work and on family farm (and non-farm) businesses provides important insights into the likely behavioral responses of family members to the death of an adult.

Another set of studies has examined the link between HIV/AIDS and labor supply of household members who are not sick. Using the first KHDS, Beegle (2005) distinguishes hours worked on the farm, on household chores, in wage employment and on non-farm self-employment, for both men and women. She uses dummy variables measuring a death in the 12 months prior to the death and 12 months after the death. Estimation uses individual fixed effects to try to capture mortality selection, plus control for other unobserved factors such as preferences for work, time and risk preferences. Time-varying individual characteristics are also included. Beegle finds essentially little or no effects of an adult death on labor supply to the farm, to wage work, non-farm self-employment or to chores. She finds limited evidence that participation in coffee farming is reduced after a male death. She suggests that new members join a household after an adult death and thereby replace the labor. Yamano and Jayne (2004) provide evidence in support of this result and report that household size falls by less than one after an adult death.

Thirumurthy, Zivin and Goldstein (2006) conducted a survey in western Kenya in early 2004 and 2005 which comprised two samples: a random sample of households and a sample of households from which individuals had used a local HIV clinic. From the clinic, Thiramurthy et al. obtained records of how long each patient had undergone ARV therapy, plus medical records on two key biomarkers: CD4 protein count and BMI. Thirumurthy et al. begin by analyzing whether the ARV treatment had effects on CD4 counts and BMI, which it should have. They estimate a patient fixed effects model using longitudinal data on each patient in the HIV clinic and find that being on ARV therapy for at least 3 months had a large impact, but the impact dropped being on ARV for

longer. For BMI, they found an immediate impact of being on ARV for 1 month, and a larger jump if the patient had been on ARVs for 6–9 months.

To assess the impacts of improved health on labor supply, Thirumurthy et al. estimate "reduced form" models. In doing so, they pool the random sample with the HIV infected sample. These reduced form models take being HIV positive as exogenous but treat "health" as measured by CD4 count or BMI as endogenous. Individual fixed effects models are estimated with the two periods of data, 2004 and 2005, which may account for the potential selection effect of HIV status, provided the true model is additively linear, with no separate trends for the two groups. Since individuals from the random sample are included, this amounts to a difference-in-difference procedure, with the time differences among the random sample, not infected by HIV, picking up any macro time trends. In addition, models are estimated that allow differential impacts of being on ARV at different times before the survey.

Thirumurthy et al. find large inter-temporal impacts of ARV treatment on labor supply, participation rising 20 percent, relative to those not HIV infected, and hours rising by 35 percent. Increases in family businesses comprise the major form of labor force increase. They also find that labor supply of women and young boys declines after ARV treatment begins, implying that these groups were substituting for the labor of the HIV infected persons.

In sum, studies of the relationship between of HIV/AIDS and household socio-economic outcomes are mixed. While several studies report negative associations and some indicate no association, none reports a large negative effect of HIV/AIDS on household economic status. Some of the differences in estimated effects are due to differences in outcomes, differences in methodologies and differences in contexts. Studies that examine individual-specific differences (include individual fixed effects) tend to yield robust results, although for many outcomes that approach is not feasible. Above all, it is apparent that scientific evidence in this area remains limited and claims that HIV/AIDS will have a devastating impact on economic prosperity are at best premature.

6. Conclusions

Overall there have been dramatic improvements in population health across the globe in the last century as the poorest countries have emerged from high levels of infant and child mortality, infectious diseases have come under better control and the incidence of under-nutrition among young children has been substantially reduced. The advent of HIV/AIDS has been a significant setback, primarily in southern and eastern Africa. As longevity increases, the fraction of the population that is older will increase and the burden of disease will shift to non-communicable diseases. This will bring new challenges to the health systems and economies of developing countries. Indeed, already several low-income populations that are emerging from the nutrition transition are experiencing rapid rates of increase in obesity, diabetes, cardio-vascular disease and cancers.

In recent years, there have been significant advances in understanding of the complex relationship between the many different domains of health and economic progress. Studies have highlighted the importance of distinguishing stocks from flows of health. This has led to greater emphasis on the development of dynamic models of health and prosperity that incorporate the complex feedbacks between health and productivity, and *vice versa*, both contemporaneously and over time. Research has also established the important role that households and families play in providing insurance to mitigate the adverse consequences on an individual of his or her own poor health.

Several important studies in economics have been substantially enriched by reaching into the biological and biomedical literatures. Those literatures are a source of extremely useful insights into likely causal pathways as well as potential mechanisms through which specific dimensions of health might impact other life outcomes. New evidence on the biological foundations underlying associations between early-life health, including fetal health, and later life outcomes has powerful implications for the design of policy and has suggested substantially new lines of research. Understanding these links is all the more pressing if fetal and childhood nutrition contribute to the rapid increase in rates of adult obesity – and the associated health problems – in low-income settings. More generally, further integration of theory and evidence between the biological and social sciences has the potential to yield major breakthroughs in understanding the links between health and development.

The Copenhagen Consensus highlighted poor health and nutrition as a key constraint to economic development. The evidence reviewed here identifies several pathways through which improved health contributes to development. Good health in early life is correlated with better health in later life. Health is positively associated with other dimensions of human capital, including schooling. Better health is associated with economic success. Isolating causality in these relationships continues to be a challenge although significant progress has been made on this front in recent years.

There are developments on the horizon that suggest a very exciting future for scientific research in this area. The development of high quality, long-term longitudinal socio-economic surveys of individuals, households, families and communities that are guided by insights from theory will provide the foundation for the empirical implementation of improved models of behavioral choices over the life course. These data, coupled with remarkable innovations in health measurement that enable population-based surveys to collect substantial biomarker and genetic information at low cost, promise to revolutionize the field of population health. As researchers integrate these data with theoretical insights from both the biological and social sciences and draw on innovative experimental and non-experimental methods to pin down causal mechanisms, the opportunities for significant progress in the health and development fields seem unparalleled.

References

Adams, P., Hurd, M., McFadden, D., Merrill, A., Ribeiro, T. (2003). "Healthy, wealthy and wise: Tests for direct causal paths between health and socioeconomic status". Journal of Econometrics 112 (1), 3–56.

Adda, J., Chandola, T., Marmot, M. (2003). "Socio-economic status and health: causality and pathways". Journal of Econometrics 112 (1), 57–63.

Ahitov, A., Hotz, V.J., Philipson, T. (1996). "The responsiveness of the demand for condoms to the local prevalence of AIDS". Journal of Human Resources 21 (4), 869–897.

Ainsworth, M., Beegle, K., Kodi, G. (2005). "The impact of adult mortality on primary school enrollment in northwestern Tanzania". Journal of Development Studies 41 (3), 412–439.

Ainsworth, M., Dayton, J. (2003). "The impact of the AIDS epidemic on the health of older persons in northwestern Tanzania". World Development 31 (1), 131–148.

Ainsworth, M., Ghosh, S., Semali, I. (1995). "The impact of adult deaths on household composition in Kagera Region, Tanzania". Development Research Group, World Bank.

Ainsworth, M., Semali, I. (1998). "Who is most likely to die of AIDS? Socio-economic correlates of adult deaths in Kagera region, Tanzania". In: Ainsworth, M., Fransen, L., Over, M. (Eds.), Confronting AIDS: Evidence from the Developing World. European Union, Brussels.

Ainsworth, M., Semali, I. (2000). "The impact of adult deaths on children's health in Northwestern Tanzania". Working paper 2266. Policy Research, The World Bank, Washington, DC.

Ainsworth, M., Koda, G., Lwihula, G., Mujinja, P., Over, M., Semali, I. (1992). "Measuring the impact of fatal adult illness in sub-Saharan Africa". Working paper 90. Living Studies Measurement Study, The World Bank, Washington, DC.

Alderman, H., Behrman, J.R. (2006). "Reducing the incidence of low birth weight in low-income countries has substantial economic benefits". World Bank Research Observer 21 (1), 25–48.

Alderman, H., Gertler, P. (1997). "Family resources and gender differences in human capital investments: The demand for children's medical care in Pakistan". In: Haddad, L., Hoddinott, J., Alderman, H. (Eds.), Intrahousehold Resource Allocation in Developing Countries: Models, Methods and Policy. Johns Hopkins Univ. Press, Baltimore, pp. 231–248.

Alderman, H., Hoddinott, J., Kinsey, W. (2006). "Long-term consequences of early childhood malnutrition". Oxford Economic Papers 58 (3), 450–474.

Alderman, H., Behrman, J.R., Lavy, V., Menon, R. (2001). "Child health and school enrollment: A longitudinal analysis". Journal of Human Resources 31 (1), 185–205.

Almond, D. (2006). "Is the 1918 Influenza pandemic over? Long-term effects of *in utero* exposure in the post-1940 US population". Journal of Political Economy 114 (4), 672–712.

Almond, D., Chay, K. (2006). "The long-run intergenerational impact of poor infant health: Evidence from cohorts born during the Civil Rights era". Mimeo. Department of Economics, University of California, Berkeley.

Angrist, J., Krueger, A.B. (1999). "Empirical strategies in labor economics". In: Ashenfelter, O., Card, D. (Eds.), Handbook in Labor Economics. North-Holland, Amsterdam, pp. 1277–1366.

Arellano, M., Bond, S. (1991). "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations". Review of Economic Studies 58, 277–297.

Banerjee, A., Deaton, A., Duflo, E. (2004). "Wealth, health, and health services in rural Rajasthan". American Economic Review, Papers and Proceedings 94 (2), 326–333.

Barbarin, O., Richter, L. (2001). Mandela's Children: Growing up in Post-apartheid South Africa. Routledge, New York.

Barker, D.J.P. (1994). Mothers, Babies and Health in Later Life. BMJ Publishing Group, London.

Barker, D.J.P. (2006). "Esther and Isadore Kesten memorial lecture. Leonard Davis School of Gerontology." University of Southern California, November 2.

Barker, D.J.P., Osmond, C. (1986). "Infant mortality, childhood nutrition and ischaemic heart disease in England and Wales". Lancet 1, 1077–1081.

Barker, D.J.P., Winter, P.D., Osmond, C., Margetts, B., Simmonds, S.J. (1989). "Weight in infancy and death from ischaemic heart disease". Lancet 2, 577–580.

Basta, S., Soekirman, K., Scrimshaw, N. (1979). "Iron deficiency anemia and productivity of adult males in Indonesia". American Journal of Clinical Nutrition 32, 916–925.
Beegle, K. (2005). "Labor effects of adult mortality in Tanzanian households". Economic Development and Cultural Change 53 (3), 655–683.
Beegle, K., De Weerdt, J., Dercon, S. (2007). "Orphanhood and the long-run impact on children". Mimeo. World Bank, Washington, DC.
Beegle, K., De Weerdt, J., Dercon, S. (2008). "Adult mortality and consumption growth in the age of HIV/AIDS". Economic Development and Cultural Change, in press.
Behrman, J.R. (1996). "The impact of health and nutrition on education". World Bank Research Observer 11 (1), 23–38.
Behrman, J.R. (1997). "Intrahousehold distribution and the family". In: Rosenzweig, M.R., Stark, O. (Eds.), Handbook of Population and Family Economics, vol. 1A. North-Holland, Amsterdam, pp. 125–187.
Behrman, J.R., Alderman, H., Hoddinott, J. (2004). "Hunger and malnutrition". In: Lomborg, B. (Ed.), Global Crises, Global Solutions. Cambridge Univ. Press, pp. 363–420.
Behrman, J.R., Deolalikar, A. (1988). "Health and nutrition". In: Chenery, H., Srinivasan, T.N. (Eds.), Handbook of Development Economics, vol. 1. North-Holland, pp. 632–711.
Behrman, J.R., Pollak, R., Taubman, P. (1982). "Parental preferences and provision for progeny". Journal of Political Economy 90 (1), 52–73.
Behrman, J.R., Rosenzweig, M.R. (2004). "Returns to birthweight". Review of Economics and Statistics 86 (2), 586–601.
Behrman, J.R., Hoddinott, J., Maluccio, J., Soler-Hampejsek, E., Behrman, E.L., Martorell, R., Ramirez-Zea, M., Stein, A. (2006). "What determines adult cognitive skills? Impacts of pre-schooling, schooling and post-schooling experiences in Guatemala". Mimeo. Department of Economics, University of Pennsylvania.
Bell, C., Devarajan, S., Gersbach, H. (2006). "The long-run economics costs of AIDS: A model with an application to South Africa". World Bank Economic Review 20 (1), 55–90.
Bengtsson, T., Lindstrom, M. (2000). "Childhood misery and disease in later life: The effects on mortality in old age of hazards experienced in early life, southern Sweden, 1760–1894". Population Studies 54, 263–277.
Bertozzi, S., Padian, N.S., Wegbreit, J., DeMaria, L.M., Feldman, B., Gayle, H., Gold, J., Grant, R., Isbell, M.T. (2006). "HIV/AIDS prevention and treatment". In: Jamison, D.T., Breman, J.G., Measham, A.R., Alleyne, G., Claeson, M., Evans, D.B., Jha, P., Mills, A., Musgrove, P. (Eds.), Disease Control Priorities in Developing Countries, Chapter 18. World Health Organization, Geneva, pp. 331–369.
Bertrand, M., Djankov, S., Hanna, R., Mullainathan, S. (2006). "Obtaining a driving license in India: An experimental approach to studying corruption". Mimeo. Harvard University.
Blundell, R., Powell, J.L. (2003). "Endogeneity in nonparametric and semiparametric regression models". In: Dewatripont, M., Hansen, L.P., Turnovsky, S.J. (Eds.), Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress. Cambridge Univ. Press, Cambridge.
Browning, M., Chiappori, P.-A., Weiss, Y. (2007). Economics of the Family. Cambridge Univ. Press, Cambridge.
Card, D. (1999). "The causal effect of education on earnings". In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics, Chapter 30. North-Holland, Amsterdam, pp. 1801–1863.
Case, A., Ardington, C. (2006). "The impact of parental death on school enrollment and achievement: Longitudinal evidence from South Africa". Demography 43 (3), 401–420.
Case, A., Deaton, A. (1998). "Large cash transfers to the elderly in South Africa". Economic Journal 108 (450), 1330–1361.
Case, A., Fertig, A., Paxson, C. (2005). "The lasting impact of childhood health and circumstance". Journal of Health Economics 24, 365–389.
Case, A., Paxson, C., Ableidinger, J. (2004). "Orphans in Africa: Parental death, poverty and school enrollment". Demography 41 (3), 483–508.
Cates, W., Bowen, G.S. (1989). "Education for AIDS prevention: Not our only voluntary weapon". American Journal of Public Health 79 (7), 871–874.

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

Center for Disease Control (CDC) (2000). CDC 200 Growth Charts, United States. CDC, Atlanta.

Central Statistics Office (2003). Central Board of Health, ORC Macro. Zambia Demographic and Health Survey, 2001–2002. Calverton, MD.

Chapman, D.A., Scott, K.G. (2001). "Intergenerational Risk Factors and Child Development". Developmental Review 21, 305–325.

Chapoto, A., Jayne, T. (2005). "Characteristics of individuals affected by AIDS related mortality in Zambia." Working paper 14. Food Security Research Project Department of Agricultural Economics, Michigan State University.

Chapoto, A., Jayne, T. (2008). "Impacts of AIDS-related mortality on farm household welfare in Zambia". Economic Development and Cultural Change, in press.

Cleveland, W.S. (1979). "Robust locally weighted regression and smoothing scatterplots". Journal of the American Statistical Association 74, 829–836.

Crimmins, E., Finch, C. (2006). "Infection, inflammation, height and longevity". Proceedings of the National Academy of Science 103 (2), 498–503.

Crimmins, E., Frankenberg, E., McDade, T., Seeman, T., Thomas, D. (2007). "Inflammation and socio-economic status in a low income population". Mimeo. UCLA.

Cutler, D. (2004). Your Money or Your Life. Oxford Univ. Press, Oxford.

Cutler, D., Deaton, A., Lleras-Muney, A. (2006). "The determinants of mortality." Working paper 11963. NBER.

Cutler, D., Knaul, F., Lozano, R., Mendez, O., Zurita, B. (2002). "Financial crisis, health outcomes and aging: Mexico in the 1980s and 1990s". Journal of Public Economics 84 (2), 279–303.

de Onis, M., Garza, C., Onyango, A.W., Martorell, R. (Eds.) (2006). WHO Child Growth Standards. Acta Pædiatrica, vol. 95. Supplement 450.

de Walque, D. (2004). "How does the impact of an HIV/AIDS information campaign vary with educational attainment: Evidence from rural Uganda". Working paper 3289. World Bank.

de Walque, D. (2006). "Who gets AIDS and how? The determinants of HIV infection and sexual behaviors in Burkina Faso, Cameroon, Ghana, Kenya and Tanzania". Working paper 3844. World Bank.

Demographic and Health Surveys (2002). "South African DHS: 1998 – Final Report". Measure/DHS, Columbia, MD.

Deolalikar, A.B. (1988). "Nutrition and labor productivity in agriculture: Estimates for rural south India". Review of Economics and Statistics 70 (3), 406–413.

Dercon, S. (Ed.) (2004). Insurance Against Poverty. Oxford Univ. Press, Oxford.

Dercon, S., Krishnan, P. (2000). "In sickness and health: Risk sharing with households in rural Ethiopia". Journal of Political Economy 108 (4), 688–727.

Dow, W., Thomas, D., Gertler, P., Schoeni, R., Strauss, J. (2004). "Health prices, health outcomes and labor outcomes". Mimeo, UCLA.

Duflo, E. (2000). "Child health and household resources in South Africa: Evidence from the old age pension program". American Economic Review 90 (2), 393–398.

Duflo, E. (2001). "Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment". American Economic Review 91 (4), 795–813.

Duflo, E., Glennerster, R., Kremer, M. (2008). "Using randomization in development economics research: A toolkit". In: Schultz, T.P., Strauss, J. (Eds.), Handbook of Development Economics, vol. 4. North-Holland, Amsterdam.

Eckstein, Z., Wolpin, K.I. (1989). "The specification and estimation of dynamic stochastic discrete choice models: A survey". Journal of Human Resources 24 (4), 562–598.

Edgerton, V., Gardner, G.W., Ohira, Y., Gunawardena, K.A., Senewiratne, B. (1979). "Iron-deficiency anemia and its effect on worker productivity and activity patterns". British Medical Journal 2, 1546–1549.

Elo, I., Preston, S. (1992). "Effects of early-life conditions on adult mortality". Population Index 58 (2), 186–211.

Emanuel, I., Filakti, H., Alberman, E., Evans, S. (1992). "Ingenerational studies of human birthweight from the 1958 birth cohort: Evidence for a multigenerational effect". British Journal of Obstetrics and Gynaecology 99 (1), 67–74.

Emanuel, I., Leisenring, W., Williams, W., Kimpo, C., Estee, S., O'Brien, W., Hale, C. (1999). "The Washington State intergenerational study of birth outcomes: Methodology and some comparisons of maternal birthweight and infant birthweight and gestation in four ethnic groups". Paediatric and Perinatal Epidemiolgy 13 (3), 352–371.

Eriksson, T., Bratsberg, B., Raaum, O. (2006). "Earnings persistence across generations: Transmission through health?". Mimeo, University of Oslo.

Evans, D., Miguel, E. (2007). "Orphans and schooling in Africa: A longitudinal analysis". Demography 44 (1), 35–58.

Federov, L., Sahn, D. (2005). "Socioeconomic determinants of children's health in Russia: A longitudinal study". Economic Development and Cultural Change 53 (2), 479–500.

Finch, C., Crimmins, E. (2004). "Inflammatory exposure and historical changes in human life spans". Science 305, 1736–1739.

Fitzgerald, J., Gottschalk, P., Moffitt, R. (1998). "An analysis of sample attrition in panel data". Journal of Human Resources 33 (2), 251–299.

Fogel, R.W. (2004). The Escape From Hunger and Premature Death, 1700–2100. Cambridge Univ. Press, Cambridge.

Foster, A. (1995). "Prices, credit markets and child growth in low-income rural areas". Economic Journal 105, 551–570.

Fox, M.P., Rosen, S., MacLeod, W.B., Wasunna, M., Bii, M., Foglia, G., Simon, J.L. (2004). "The impact of HIV/AIDS on labour productivity in Kenya". Tropical Medicine and International Health 9 (3), 318–324.

Frankenberg, E., Karoly, L. (1995). "The 1993 Indonesian Family Life Survey: Overview and Field Report". RAND DRU-1195/1-NICHD/AID.

Frankenberg, E., McKee, D., Thomas, D. (2005). "Health consequences of forest fires in Indonesia". Demography 42 (1), 109–129.

Frankenberg, E., Suriastini, W., Thomas, D. (2005). "Can expanding access to basic healthcare improve children's health status? Lessons from Indonesia's 'midwife in the village' program". Population Studies 59 (1), 5–19.

Frankenberg, E., Thomas, D. (2000). "The Indonesia Family Life Survey (IFLS): Study Design and Results from Waves 1 and 2". RAND DRU-2238/1-NIA/NICHD.

Frankenberg, E., Thomas, D., Beegle, K. (1999). "The real costs of Indonesia's economic crisis". Working paper 99-04. Labor and Population, RAND, Santa Monica, CA.

Gertler, P., Gruber, J. (2002). "Insuring consumption against illness". American Economic Review 92 (1), 51–70.

Glewwe, P., Hall, G. (1994). "Poverty, inequality and living standards during unorthodox adjustment". Economic Development and Cultural Change 42 (4), 689–718.

Glewwe, P., Jacoby, H., King, E. (2001). "Early childhood nutrition and academic achievement: A longitudinal analysis". Journal of Public Economics 81 (3), 345–368.

Glewwe, P., King, E. (2001). "The impact of early childhood nutritional status on cognitive development: Does the timing of malnutrition matter?" World Bank Economic Review 15 (1), 81–113.

Glewwe, P., Miguel, E. (2008). "The impact of child health and nutrition on education in less developed countries". In: Schultz, T.P., Strauss, J. (Eds.), Handbook of Development Economics, vol. 4. North-Holland Press, Amsterdam.

Godfrey, K., Barker, D. (2000). "Fetal nutrition and adult disease". American Journal of Clinical Nutrition 71 (suppl), 1344S–1352S.

Goldman, D., Smith, J.P. (2002). "Can patient self-management help explain the SES health gradient?". Proceedings of the National Academy of Sciences 99 (16), 10929–10934.

Grajeda, R., Behrman, J.R., Flores, R., Stein, A.D., Maluccio, J.A., Martorell, R. (2005a). "Design and implementation of the INCAP early nutrition, human capital and economic productivity follow-up study, 2002–2004". Food and Nutrition Bulletin 26 (2), S15–S24.

Grajeda, R., Behrman, J.R., Flores, R., Maluccio, J., Martorell, R., Stein, A. (2005b). "The human capital study 2002–2004: Tracking, data collection, coverage and attrition". Food and Nutrition Bulletin 26 (2, Supplement 1), S15–S24.

Granger, C. (1969). "Investigating causal relations by econometric models and cross-spectral methods". Econometrica 37, 424–438.

Griliches, Z., Hausman, J.A. (1985). "Errors in variables in panel data: A note with an example". Journal of Econometrics 31, 93–118.

Grossman, M. (1972). "On the concept of health capital and the demand for health". Journal of Political Economy 80 (2), 223–255.

Haas, J., Brownlie, T. (2001). "Iron deficiency and reduced work capacity: A critical review of the research to determine a causal relationship". Journal of Nutrition 131 (supplement), 676S–688S.

Habicht, J.-P., Martorell, R., Rivera, J. (1995). "Nutritional impact of supplementation in the INCAP longitudinal study: Analytic strategies and inferences". Journal of Nutrition 125 (4S), 1042S–1050S.

Hales, C.N., Barker, D.J. (1992). "Type 2 (non-insulin-dependent) diabetes mellitus: the thrifty phenotype hypothesis". Diabetologia 35, 595–601.

Haller, H. (1997). "Epidemiology and associated risk factors of hyperlipoproteinemia". Z. Gesamte Inn. Med. 32 (8), 124–128.

Hamoudi, A., Thomas, D. (2005). "Pension income and the well-being of children and grandchildren: New evidence from South Africa". Working paper 043-05, CCPR.

Hardle, W. (1992). Applied Nonparametric Regression. Cambridge Univ. Press, Cambridge.

Harris, J.R., Gruenewald, T.L., Seeman, T. (2007). "An overview of biomarker research from community and population-based studies on aging". Chapter 5 in: Weinstein, M., Vaupel, J.W., Wachter, K.W. (Eds.), Biosocial Surveys. National Academies Press, Washington, DC.

Hausman, J.A. (2003). "Triangular structural model specification and estimation with application to causality". Journal of Econometrics 112 (1), 107–113.

Hayami, Y., Ruttan, V.W. (1971). Agricultural Development: An International Perspective. Johns Hopkins Univ. Press, Baltimore.

Heckman, J. (1978). "Dummy endogenous variables in a simultaneous equation system". Econometrica 46 (4), 931–959.

Heckman, J. (2003). "Conditioning, causality and policy analysis". Journal of Econometrics 112 (1), 73–78.

Heckman, J., Navarro-Lozano, S. (2004). "Using matching, instrumental variables, and control functions to estimate economic choice models". Review of Economics and Statistics 86 (1), 30–57.

Heckman, J., Robb, R. (1985). "Alternative methods for evaluating the impact of interventions". In: Heckman, J., Singer, B. (Eds.), Longitudinal Analysis of Labor Market Data. Wiley, New York, pp. 156–245.

Heckman, J., Vytlacil, E. (2006). "Econometric evaluation of social programs". In: Heckman, J., Leamer, E. (Eds.), Handbook of Econometrics, vol. 6. North-Holland, Amsterdam.

Hoddinott, J., Kinsey, B. (2001). "Child growth in the time of drought". Oxford Bulletin of Economics and Statistics 63 (4), 409–436.

Horton, S. (1986). "Child nutrition and family size in the Philippines". Journal of Development Economics 23 (1), 161–176.

Jacoby, H., Skoufias, E. (1997). "Risk, financial markets and human capital in a developing country". Review of Economic Studies 64, 311–335.

Jamison, D.T., Breman, J.G., Measham, A.R., Alleyne, G., Claeson, M., Evans, D.B., Jha, P., Mills, A., Musgrove, P. (2006). Disease Control Priorities in Developing Countries. Oxford Univ. Press and World Health Organization.

Jayachandran, S. (2006). "Air quality and early life mortality: Evidence from Indonesia's wildfires". Mimeo. UCLA.

Jayne, T.S., Strauss, J., Yamano, T., Molla, D. (2002). "Targeting of food aid in rural Ethiopia: Chronic need or inertia?" Journal of Development Economics 68 (2), 247–288.

Jensen, R. (2000). "Agricultural volatility and investment in children". American Economic Review Papers and Proceedings 90 (2), 399–404.

Jensen, R., Richter, K. (2003). "The health implications of social security failure: Evidence from the Russian pension crisis". Journal of Public Economics 88 (1–2), 209–236.
Joshi, S., Schultz, T.P. (2006). "Family planning as an investment in development and female human capital: Evaluating the long-run consequences in Matlab, Bangladesh". Mimeo. Yale University.
Kannisto, V., Christensen, K., Vaupel, J. (1997). "No increased mortality in later life for cohorts born during famine". American Journal of Epidemiology 145, 987–994.
King, G., Murray, C., Salomon, J., Tandon, A. (2004). "Enhancing the validity and cross-cultural validity of measurement in survey research". American Political Science Review 98 (1), 191–207.
Knodel, J. (2008). "Poverty and the impact of AIDS on older persons: Evidence from Cambodia and Thailand". Economic Development and Cultural Change, in press.
Komlos, J. (Ed.) (1994). Stature, Living Standards and Economic Development: Essays in Anthropometric History. Univ. of Chicago Press, Chicago.
Li, R., et al. (1994). "Functional consequences of iron supplementation in iron-deficient female cotton workers in Beijing, China". American Journal of Clinical Nutrition 59, 908–913.
Lindau, S.T., McDade, T.W. (2007). "Minimally invasive and innovative methods for biomarker collection in population-based research". Chapter 13 in: Weinstein, M., Vaupel, J.W., Wachter, K.W. (Eds.), Biosocial Surveys. National Academies Press, Washington, DC.
Linnemayr, S. (2005). "Consumption smoothing and HIV/AIDS: The case of two communities in South Africa". Mimeo. Paris-Jourdan Sciences Economiques, Paris.
Lomborg, B. (Ed.) (2004). Global Crises, Global Solutions. Cambridge Univ. Press, Cambridge.
Lomborg, B. (Ed.) (2006). How to Spend $50 Billion to Make the World a Better Place. Cambridge Univ. Press, Cambridge.
Lopez, A.D., Mathers, C.D., Ezzati, M., Jamison, D., Murray, C. (2006). Global Burden of Disease and Risk Factors. Oxford Univ. Press, Oxford.
Luo, Z., Mu, R., Zhang, X. (2006). "Famine and overweight in China". Review of Agricultural Economics 28 (3), 296–304.
Maluccio, J., Hoddinott, J., Behrman, J.R., Quisumbing, A., Martorell, R., Stein, A.D. (2006). "The impact of nutrition during early childhood on education among Guatemalan adults". Mimeo. IFPRI, Washington, DC.
Manski, C., Lerman, S. (1977). "The estimation of choice probabilities from choice-based samples". Econometrica 45 (7), 1977–1988.
Martorell, R., Habicht, J.-P. (1986). "Growth in early childhood in developing countries". In: Falkner, F., Tanner, J.M. (Eds.), Human Growth, vol. 3, Methodology. Ecological, Genetic, and Nutritional Effects on Growth. second ed. Plenum, New York, pp. 241–262.
Martorell, R., Habicht, J.-P., Rivera, J. (1995). "History and design of the INCAP Longitudinal Study (1969–77) and its follow-up (1988–89)". Journal of Nutrition 125 (4S), 1027S–1041S.
Martorell, R., Delgado, H., Valverde, V., Klein, R. (1981). "Maternal stature, fertility and infant mortality". Human Biology 53, 303–312.
Martorell, R., Behrman, J.R., Flores, R., Stein, A. (2005). "Rationale for a follow-up study focusing on economic productivity". Food and Nutrition Bulletin 26 (2, supplement 1), S5–S14.
Mathers, C., Lopez, A., Murray, C. (2006). "The burden of disease and mortality by condition: Data, methods and results for 2001". In: Lopez, A., Mathews, C., Ezzati, M., Jamison, D., Murray, C. (Eds.), Global Burden of Disease and Risk Factors. Oxford Univ. Press, Oxford, pp. 45–235.
McDade, T.W., Williams, S.R., Snodgrass, J.J. (2006). "What a drop can do: Expanding options for the analysis of blood-based biomarkers in population-level health research". Mimeo. Northwestern University.
McDade, T.W., Beck, M.A., Kuzawa, C., Adair, L.S. (2001). "Prenatal undernutrition, postnatal environments, and antibody response to vaccination in adolescence". American Journal of Clinical Nutrition 74 (4), 543–548.
McEwen, B.S. (1998). "Protective and damaging effects of stress mediators". New England Journal of Medicine 338, 171–179.
McKee, D. (2006). "Forward thinking and family support: Explaining retirement and old age labor supply in Indonesia". Working paper 005-06. CCPR, UCLA.

McKenzie, D. (2006). "The consumer response to the Mexican Peso crisis". Economic Development and Cultural Change 55 (1), 139–172.

Miguel, E., Kremer, M. (2004). "Worms: Identifying impacts on education and health in the presence of treatment externalities". Econometrica 72 (1), 159–217.

Mills, A., Shillcutt, S. (2004). "Communicable diseases". In: Lomborg, B. (Ed.), Global Crises, Global Solutions. Cambridge Univ. Press, Cambridge, pp. 62–114.

Mirrlees, J. (1975). "A pure theory of underdevelopment". In: Reynolds, L.G. (Ed.), Agriculture in Development Theory. Yale Univ. Press, New Haven.

Monteiro, C.A., Moura, E.C., Conde, W.L., Popkin, B.M. (2004). "Socioeconomic status and obesity in adult populations of developing countries: A review". Bulletin of the World Health Organization 82 (12), 940–946.

Mullahy, J., Portney, P.R. (1990). "Air pollution, cigarette smoking and the production of respiratory health". Journal of Health Economics 9 (2), 193–205.

Mwabu, G. (2008). "Health economics for low-income countries". In: Schultz, T.P., Strauss, J. (Eds.), Handbook of Development Economics, vol. 4. North-Holland, Amsterdam.

National Center for Health Statistics (NCHS) (1994). National Health and Nutrition Examination Survey Data Wave III. US Department of Health and Human Services, Hyattsville, MD. Centers for Disease Control and Prevention, http://www.cdc.gov/nchs/about/major/nhanes/nh3data.htm.

Nelson, R. (1964). "Aggregate production function and medium-range growth projections". American Economic Review 53, 974–1004.

Newhouse, J. (1996). Free for All?: Lessons from the RAND Health Insurance Experiment. Harvard Univ. Press.

Omran, A.R. (1971). "The epidemiologic transition: A theory of the epidemiology of population change". Millbank Memorial Fund Quarterly 49 (4.1), 509–538.

Parker, S., Rubalcava, L., Teruel, G. (2008). "Evaluating conditional schooling and health programs". In: Schultz, T.P., Strauss, J. (Eds.), Handbook of Development Economics, vol. 4. North-Holland, Amsterdam.

Paxson, C., Schady, N. (2005). "Child health and economic crisis in Peru". World Bank Economic Review 19 (2), 203–224.

Persico, N., Postlewaite, A., Silverman, D. (2004). "The effect of adolescent experience on labor market outcomes: The case of height". Journal of Political Economy 112 (5), 1019–1053.

Pitt, M.M. (1997). "Specification and estimation of the demand for goods within the household". In: Haddad, L., Hoddinott, J., Alderman, H. (Eds.), Intrahousehold Resource Allocation in Developing Countries: Models, Methods and Policy. Johns Hopkins Univ. Press, Baltimore.

Pitt, M.M., Rosenzweig, M.R. (1986). "Agricultural prices, food consumption and the health and productivity of Indonesian farmers". In: Singh, I., Squire, L., Strauss, J. (Eds.), Agricultural Household Models: Extension, Application and Policy. Johns Hopkins Univ. Press, Baltimore.

Pitt, M.M., Rosenzweig, M.R., Hassan, M.N. (1990). "Productivity, health and inequality in the intrahousehold distribution of food in low-income countries". American Economic Review 80 (5), 1139–1156.

Pitt, M.M., Rosenzweig, M.R., Hassan, M.N. (2005). "Sharing the burden of disease: Gender, the household division of labor and the health effects of indoor pollution". Working paper No. 93, BREAD.

Pollitt, E., Gorman, K.S., Engel, P., Martorell, R., Rivera, J. (1993). Early supplementary feeding and cognition. Monographs of the Society for Research in Child Development, Serial, vol. 235, 58(7).

Popkin, B.M. (1993). China Economic, Population, Nutrition and Health Survey, 1991. Questionnaires and Interview Manual. Online at http://www.cpc.unc.edu/projects/china/data.

Popkin, B.M. (1994). "The nutrition transition in low-income countries: An emerging crisis". Nutrition Reviews 52 (9), 285–298.

Popkin, B.M. (2003). "The nutrition transition in the developing world". Development Policy Review 21, 581–597.

Preston, S.H. (1975). "The changing relation between mortality and level of economic development". Population Studies 29 (2), 231–248.

Pritchett, L., Summers, L.H. (1996). "Wealthier is healthier". Journal of Human Resources 31 (4), 841–868.

Rahman, O., Menken, J., Foster, A., Peterson, C.E., Khan, M.N., Kuhn, R., Gertler, P. (1999). The 1996 Matlab Health and Socio-economic Survey: Overview and User's Guide. RAND, DRU-2018/1.

Ravelli, A.C.J., van der Meulen, J.H.P., Michels, R.P.J., Osmond, C., Barker, D., Hales, C.N., Bleker, O.P. (1998). "Glucose tolerance in adults after prenatal exposure to famine". Lancet 351, 173–177.

Ravelli, A.C.J., van der Meulen, J.H.P., Osmond, C., Barker, D., Bleker, O.P. (1999). "Obesity at the age of 50 years in men and women exposed to famine prenatally". American Journal of Clinical Nutrition 70, 811–816.

Razzaque, A., Alam, N., Wai, L., Foster, A. (1990). "Sustained effects of the 1974–75 famine on infant and child mortality in a rural area of Bangladesh". Population Studies 44 (1), 145–154.

Robinson, R. (2001). "The fetal origins of adult disease. No longer just a hypothesis and may be critically important in south Asia". British Medical Journal 322, 375–376.

Rose, E. (1999). "Consumption smoothing and excess female mortality in rural India". Review of Economics and Statistics 81 (1), 41–49.

Roseboom, T.J., van der Meulen, J.H.P., Osmond, C., Barker, D.J., Blecker, O.P. (2001a). "Adult survival after prenatal exposure to the Dutch famine 1944–1945". Paedeatric and Perinatal Epidemiology 15, 220–225.

Roseboom, T.J., van der Meulen, J.H.P., Ravelli, A., Osmond, C., Barker, D.J., Blecker, O.P. (2001b). "Effects of prenatal exposure to the Dutch famine on adult disease in later life: An overview". Twin Research 4 (5), 293–298.

Rosenbaum, P.R., Rubin, D.B. (1983). "The central role of the propensity score in observational studies for causal effects". Biometrika 70, 41–55.

Rosenzweig, M.R. (1988). "Risk, implicit contracts and the family". Economic Journal 98, 1148–1170.

Rosenzweig, M.R., Wolpin, K.I. (1985). "Specific experience, household structure and intergenerational transfers: Farm family land and labor arrangements in developing countries". Quarterly Journal of Economics 100 (Supplement), 961–987.

Rosenzweig, M.R., Wolpin, K.I. (1988). "Heterogeneity, intrafamily distribution and child health". Journal of Human Resources 23 (4), 437–461.

Rosenzweig, M.R., Wolpin, K.I. (2000). "Natural 'natural experiments' in economics". Journal of Economic Literature 38 (4), 827–874.

Rosenzweig, M.R., Zhang, J. (2006). "Do population control policies induce more human capital investment? Twins, birthweight and China's 'One Child' policy". Working paper No. 933. Economic Growth Center, Yale University.

Rubalcava, L., Teruel, G. (2004). "The Mexican Family Life Survey: Overview and Study Description". CIDE/UIA.

Rukumnuaykit, P. (2003). "Economic crises and demographic outcomes: Evidence from Indonesia". PhD dissertation. Department of Economics, Michigan State University, East Lansing, MI.

Sastry, N. (2002). "Forest fires, air pollution and mortality in southeast Asia". Demography 39 (1), 1–23.

Schultz, T.P. (1984). "Studying the impact of household economic and community variables on child mortality". In: Moseley, W.H., Chen, L. (Eds.), Child Survival: Strategies for Research. Population and Development Review 10 (Supplement) 215–236.

Schultz, T.P. (2005). "Productive benefits of health: Evidence from low-income countries". In: Lopez-Casasnovas, G., Riveras, B., Currais, L. (Eds.), Health and Economic Growth: Findings and Policy Implications. MIT Press, Cambridge MA, pp. 257–286.

Schultz, T.P., Tansel, A. (1997). "Wage and labor supply effects of illness in Cote d'Ivoire and Ghana: Instrumental variable estimates for days disabled". Journal of Development Economics 53 (2), 251–286.

Scrimshaw, N.S., Taylor, C.E., Gordon, J.E. (1968). Interactions of Nutrition and Infection. World Health Organization, Geneva.

Seeman, T.E., McEwen, B., Singer, B., Albert, M., Rowe, J. (1997). "Increase in urinary cortisol excretion and declines in memory: MacArthur Studies of Successful Aging". Journal of Clinical Endocrinology and Metabolism 82, 2458–2465.

Seeman, T.E., Crimmins, E., Huang, M.H., Singer, B., Bucur, A., Gruenewald, T., et al. (2004). "Cumulative biological risk and socio-economic differences in mortality: MacArthur studies of successful aging". Social Science and Medicine 58 (10), 1985–1997.

Serbin, L.A., Karp, J. (2004). "The intergenerational transfer of psychosocial risk: Mediators of vulnerability and resilience". Annual Review of Psychology 55, 333–363.

Serbin, L.A., Schwartzman, A.E., Moskowitz, D.S., Ledingham, J.E. (1991). "Aggressive, withdrawn, and aggressive/withdrawn children in adolescence: Into the next generation". In: Pepler, D.J., Rubin, K.H. (Eds.), The Development and Treatment of Childhood Aggression. Erlbaum, Hillsdale, NJ, pp. 55–70.

Smith, J.P. (1999). "Healthy bodies and thick wallets: The dual relation between health and economic status". Journal of Economic Perspectives 13 (2), 145–167.

Smith, J.P. (2004). "The impact of health on SES". Mimeo, The RAND Corporation.

Stein, Z., Susser, M., Saenger, G., Marolla, F. (1975). Famine and human development: The Dutch hunger winter of 1944–45. Oxford Univ. Press, Oxford.

Stein, A., Behrman, J.R., DiGirolamo, A., Grajeda, R., Martorell, R., Quisumbing, A.R., Ramakrishnan, U. (2005). "Schooling, educational achievement and cognitive functioning among young Guatemalan adults". Food and Nutrition Bulletin 26 (2), S46–S54.

Stewart, R.J.C., Sheppard, H., Preece, R., Waterlow, J.C. (1980). "The effect of rehabilitation at different stages of development of rats marginally malnourished for ten to twelve generations". British Journal of Nutrition 43 (3), 403–412.

Strauss, J. (1986). "Does better nutrition raise farm productivity?" Journal of Political Economy 94 (2), 297–320.

Strauss, J. (1993). "The impact of improved nutrition on labor productivity and human resource development". In: Pinstrup-Andersen, P. (Ed.), The Political Economy of Food and Nutrition Policies. Johns Hopkins Univ. Press, Baltimore.

Strauss, J., Thomas, D. (1995). "Human resources: Empirical modeling of household and family decisions". In: Behrman, J.R., Srinivasan, T.N. (Eds.), Handbook of Development Economics, vol. 3A. North-Holland, Amsterdam, pp. 1185–2023.

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

Strauss, J., Beegle, K., Sikoki, B., Dwiyanto, A., Witoelar, F., Herawati, Y. (2004). The Third Wave of the Indonesia Family Life Survey (IFLS3): Overview and Field Report. RAND, WR-144/1-NIA/NICHD.

Thirumurthy, H., Zivin, J.G., Goldstein, M. (2006). "The economic impact of AIDS treatment: Labor supply in western Kenya". Discussion paper 947. Economic Growth Center, Yale University.

Thomas, D., Frankenberg, E., Smith, J.P. (2001). "Lost but not forgotten: Attrition in the Indonesia Family life Survey". Journal of Human Resources 36 (2), 556–592.

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

Thomas, D., Frankenberg, E., Friedman, J., Habicht, J.-P., Ingwersen, N., McKelvey, C., Hakimi, M., Jaswadi, P., Jones, N., Sikoki, B., Pelto, G., Seeman, T., Smith, J.P., Sumantri, C., Suriastini, W., Wilopo, S. (2006). "Causal effect of health on labor market outcomes: Experimental evidence". Mimeo, UCLA.

Townsend, R. (1994). "Risk and insurance in village India". Econometrica 62 (3), 539–592.

UNAIDS (2006). "2006 Report on the Global AIDS Epidemic." New York.

UNAIDS/WHO (2006). AIDS epidemic update: December 2006. United Nations and World Health Organization, New York.

van den Berg, G., Lindeboom, M., Portrait, F. (2006). "Economic conditions early in life and individual mortality". American Economic Review 96 (1), 290–293.

Waaler, H. (1984). "Height, health and mortality: The Norwegian experience". Acta Medica Scandinavica Supplement 679, 1–51.

Ware, J.E., Sherbourne, C.D. (1992). "The MOS 36-item short-form health survey (SF-36): I. Conceptual Framework and item selection". Medical Care 30 (6), 473–483.

Watson, M. (1953). African Highway, The Battle for Health in Central Africa. Murray, London.

World Health Organization Mental Health Survey Consortium (2004). "Prevalence, severity, and unmet need for treatment of mental disorders in the World Health Organization World Mental Health Surveys". JAMA 291, 2581–2590.

Witoelar, F., Strauss, J., Rukumnuaykit, P. (2006). "Smoking behavior among youth in a developing country: The case of Indonesia". Mimeo, University of Southern California.

Wolfson, L.J., Strebel, P.M., Gacic-Dobo, M., Hoekstra, E.J., McFarland, J.W., Hersh, B.S. (2007). "Has the 2005 measles mortality reduction goal been achieved? A natural history modelling study". The Lancet 369 (9557), 191–200.

Wooldridge, J. (2002). Econometric Analysis of Cross Section and Panel Data. MIT Press, Cambridge.

World Bank (1997). Confronting AIDS: Public Priorities in a Global Epidemic, Oxford Univ. Press, Oxford.

World Health Organization (WHO) (2004). World Report on Road Traffic Injury Prevention.

World Health Organization (WHO) and UNAIDS (2003). "Reconciling antenatal clinic-based surveillance and population-based survey estimates of HIV prevalence in sub-Saharan Africa". Geneva.

Yamano, T. (2006). "Long-term impacts of orphanhood on education attainments and land inheritance and adults in rural Kenya". Mimeo. Foundation for the Advanced Study of International Development, Tokyo. Paper presented at the Population Association of America annual meetings, April 2006, Los Angeles.

Yamano, T., Alderman, H., Christiaensen, L. (2005). "Child growth, shocks and food aid in rural Ethiopia". American Journal of Agricultural Economics 87 (2), 273–288.

Yamano, T., Jayne, T.S. (2004). "Measuring the impacts of working-age adult mortality on small-scale farm households in Kenya". World Development 32 (1), 91–119.

Yamano, T., Jayne, T.S. (2005). "Working-age adult mortality and primary school enrollment in rural Kenya". Economic Development and Cultural Change 53 (3), 619–653.