ELSEVIER

Journal of Econometrics 77 (1997) 159–185

JOURNAL OF
Econometrics

Health and wages: Evidence on men and women in urban Brazil

Duncan Thomas*[,a,b], John Strauss[c]

[a]*Economics Department, University of California, Los Angeles, CA 90024, USA*
[b]*RAND Corporation, Santa Monica, CA 90047, USA*
[c]*Economics Department, Michigan State University, East Lansing, MI 48824, USA*

Abstract

Survey data indicate that different dimensions of health affect the wages of men and women in urban Brazil. Height has a large and significant effect on wages: taller men and women earn more. Body mass index (BMI) is associated with higher wages of males, especially among the less-educated, suggesting that strength may be rewarded with higher wages. Low levels of *per capita* calorie and protein intakes reduce wages of market-workers, but not the self-employed. After controlling for height, BMI, and calories, the influence of proteins is greater at higher levels, presumably reflecting the impact of higher-quality diets.

Key words: Health; Anthropometrics; Nutrients; Wages; Brazil
JEL classification: J24; I12

1. Introduction

An extensive literature demonstrates there are substantial returns in labor markets to investments in human capital, as measured by education. In contrast, comparatively few studies have examined the returns to other dimensions of

*Corresponding author.

Financial support from National Institutes of Health Grant P50-HD12639-12 is gratefully acknowledged. We have benefited from the comments of the editor, three referees, Angus Deaton, Naihua Duan, Eileen Miech, Mark Rosenzweig, and T. Paul Schultz. Margaret Hwang provided research assistance.

PII S0304-4076(96)01811-8

human capital, such as health, particularly for developing economies. Yet, the link between productivity and health, especially those dimensions related to nutrition, has long played a key role in theories of economic development, through the idea of efficiency wages, and has also taken a central place in the study of economic history.[1] However, until very recently, development economists have typically concluded there is little reliable empirical evidence indicating health has an important impact on labor productivity.[2]

This skeptical view stems, in part, from the paucity of studies on the subject, which reflects the fact that health indicators have seldom been collected in surveys that contain measures of wages or productivity. The skepticism also reflects questions over the proper interpretation of correlations between health and labor outcomes presented in many early studies which paid little or no attention to the direction of causality. Those studies ignored the fact that any component of income, such as wages or labor supply, may affect current behavior which, in turn, affects health, such as consuming a healthier diet, and *vice versa*.

A number of recent studies have begun to grapple with the issue of the potential endogeneity of health status, and a body of evidence is now developing that suggests there are causal relationships between health and labor productivity in low-income countries.[3] These studies have focused on rural populations, mostly male workers, and have seldom examined more than one or two health measures. But, if knowledge is to be advanced in this area and if it is to be potentially relevant for policy, then it is necessary to be more precise and to identify the types of individuals and activities for which the returns to investments in health are the highest. Furthermore, just as education has different dimensions such as its quantity and quality, so too does health. Which dimensions of health have labor market impacts and for whom? Do investments in health as an adult reap returns, or is it only health investments during childhood that matter? And, does the impact of health vary across the income distribution; in particular, does it especially matter among the poor?

In this paper, we contribute to the literature in several ways, using a large and very rich cross-sectional household survey conducted in Brazil. Given income, the country's investment levels in human capital are low in Brazil, and a high fraction of people are in poor health. We investigate the impact of four indicators of health on wages of urban workers: these indicators are height, body mass index (weight divided by height squared), *per capita* calorie intake, and *per capita* protein intake. These indicators do not fully capture 'health', but they do measure different dimensions of it. For example, height is a cumulative measure

[1] See reviews by Rosenzweig (1988), Dasgupta (1993), and Fogel (1994).

[2] See Rosenzweig (1988), for instance.

[3] See Strauss and Thomas (1995) for a review.

reflecting both investments in nutrition during one's life (mostly as a child) and also, possibly, nonhealth human capital investments. Nutrient intakes, in contrast, are inputs into the production function for current health. The effects of the four indicators are examined individually and all together, with body mass index and nutrient intakes being treated as endogenous (or measured with error). Special attention is paid to nonlinearities in the effects of health, guided in our interpretation by the biomedical evidence. Noting that the effects may vary with the nature of the activity, comparisons are drawn between those who work in the market sector and those who are self-employed, treating sectoral choice as endogenous. As a different cut on the same issue, we examine whether returns differ by educational group, which we adopt as a proxy for energy expenditure required for the job. These comparisons also shed light on differences in the returns to health across the income distribution. Finally, contrasts between men and women provide further evidence to support interpretation of the patterns we observe.

Our findings indicate that health, as measured in this paper, yields a substantial return in urban Brazil, at least in the market wage sector. Among men and women who participate in that sector, all four measures of health significantly affect wages even after accounting for endogeneity. Height has a particularly large impact: taller men and women earn more even after controlling for education and other dimensions of health. Body mass index has a positive effect on the wages of men but not of women. This, along with the fact that the effect of height is larger for men, is likely to reflect, in part, a return to strength. Further corroborating evidence along these lines is suggested by the fact that body mass index has a larger (and usually significant) effect on the wages of those with little education (including women). Nutrient intakes are also important among workers in the market sector. More calories are associated with higher wages, but only at very low intake levels. Conditional on calorie intake, mass, and height, additional protein has the greatest return at high levels of intakes, suggesting there is a return to higher quality diets. The evidence among the self-employed is less clear. Height remains a powerful predictor of wages for men, but not for women, and body mass index affects the wages of only those men who have little or no education. Neither of the nutrient intakes appears to significantly affect wages of the self-employed.

The next section outlines a model of wages and health, which guides the empirical analysis. This is followed by a discussion of measurement issues and then a description of the data sources. The regression results follow in Section 5.

2. Model

In order to draw inferences about the effect of health on labor market outcomes, at least two key issues need to be addressed. The first is measurement:

health status is multi-dimensional (Ware, 1987) and difficult to capture in survey data. It is discussed in the next subsection.

The second issue is the direction of causality between measures of health and wages or productivity. While better health may result in a worker being more productive, higher income may be spent on improving one's health. In addition, unobserved factors related to human capital or tastes may affect both current health and productivity. Thus, OLS estimates of the effect of health on wages are likely to be contaminated by both simultaneity and unobserved heterogeneity bias. We first describe a model of health, productivity, and labor force participation, in which all three outcomes are treated as endogenous. The model is then used to identify the effects of health on productivity and wages.

Begin with an unconditional (reduced form) latent labor supply function, h^*, which depends on a vector of observable exogenous individual characteristics, X_i, a vector of community-level characteristics, X_c, such as prices, local demand and infrastructure, and a productivity-related individual-specific unobservable. An individual will choose to work if a wage offer is greater than the shadow wage, in which case, $h^* > 0$.

To examine the influence of health on productivity, consider a standard ln(wage) function, which is conditional on both health, X_h, and labor force participation, $h^* > 0$:

$$\ln w = \omega(X_i, X_h, \bar{X}_c, \mu_i), \qquad h^* > 0, \tag{1}$$

where $\bar{X}_c$ is a subset of community characteristics X_c.

In an attempt to purge correlations between health, X_h, and unobservables, μ_i, we adopt an instrumental variables estimator.[4] Assuming wages do not depend on labor supply, individual and community characteristics which affect health, but not wages, are valid instruments.[5] Examples might include community-level characteristics such as the disease environment, health infrastructure, and prices of health inputs. Since the measures of health used below are all related to food and nutrition, we use as instruments relative food prices, which vary across regions in Brazil.[6] In a static model of wage determination, wealth or nonlabor

[4]Other studies have used fixed effects estimates, taking advantage of repeated observations of the same individual; see, for example, Deolalikar (1988) or Haddad and Bouis (1991). While studies that use only fixed effects side-step the difficult problem of choosing instruments, they do make strong assumptions. In particular, it is assumed that changes in wages are not reflected in contemporaneous changes in health (or nutrition) inputs or outcomes.

[5]This abstracts from endogenous program placement or selective migration (Rosenzweig and Wolpin, 1986, 1988).

[6]Since there are substantial cost-of-living differences between regions in Brazil, it is important to control for overall interregional price differentials; an aggregate price index and region control are included in the wage functions.

income and, possibly, characteristics of other household members may be potential instruments. We use measures of nonlabor income (of the individual and other household members). These nonlabor income and price characteristics also serve to identify selection into the labor force.

Modeling sectoral choice decisions is, in principle, analogous to the participation decision with polychotomous choices, and so sector-specific wage functions might be estimated with the same identification strategy. However, identification of health effects is slightly more difficult for the self-employed than for those working for market wages. Productive, quasi-fixed household assets belong in the net self-employment wages function, as do characteristics of all household members who work in the enterprise. Provided that, conditional on health, the household self-employment production decisions and its consumption decisions can be modeled recursively (Singh, Squire, and Strauss, 1986), health infrastructure and price variables (unrelated to the self-employment enterprise) are still legitimate instruments for the self-employment wage function when it is conditioned on health (Pitt and Rosenzweig, 1986).

3. Dimensions of health, productivity, and labor supply

The next key issue is measurement, which has spawned a large literature of its own. At a conceptual level, it is not clear how to measure health and, more practically, different dimensions of health are likely to have different effects on wages. These effects may also differ depending on the nature of work: a laborer, for example, may suffer a larger decline in income because of physical injury than a more sedentary worker.

Measures of morbidity incidence among adults appear to have fairly small effects on incomes and wages (Pitt and Rosenzweig, 1986; Schultz and Tansel, 1992). Clearly, incidence of ill health is a short-term measure (unless the illness is chronic), and for many bouts of illness, incomes are likely to be little affected in the longer run (say, over a year). Moreover, illness is rather hard to measure in surveys, and most studies have relied on self-reported morbidities, which are prone to both random and systematic reporting error.[7]

Anthropometric measurements, in particular height and weight, have been suggested as less subjective indicators of health status, although they measure different dimensions of health. Height may be directly related to productivity, but it also reflects previous health investments, primarily early in life, and it is possibly correlated with nonhealth human capital investments made during

[7] If higher-income individuals are more likely to report themselves ill, *ceteris paribus*, it will be hard to separately identify the effect of reporting error and illness incidence on wages. See, for example, Hill and Mamdani (1989).

childhood. Long-term increases in heights in the United States and Europe have been related to increases in national incomes and declines in mortality rates (Fogel, 1994). Weight is also potentially related to productivity, at least among those who are very light or very heavy, through such mechanisms as metabolic efficiency and maximum physical capacity. But a light person may also be small, and so not underweight given height (and, conversely, heavy, tall people may not be overweight). Thus nutritionists have found it convenient to analyze weight given height. While different ways of expressing this ratio are possible, one that is often used is the body mass index or BMI.[8] BMI has been shown to be related to maximum physical capacity independent of energy intake (Spurr, 1983; Martorell and Arroyave, 1988). Furthermore, energy can be stored in the body and expended when needed. Thus, BMI is likely to vary in the short run and may be affected by contemporaneous movements in income or prices. Nevertheless, it is important to note that BMI, as well as height, may partly reflect previous health investments, and these may be correlated with other past human capital investments that directly affect productivity and labor supply.

A third set of health-related factors are nutrient intakes, which are likely to vary over both the short and long run. We will focus on calories and protein. Current intakes of energy may enhance productivity in some jobs, for example by increasing maximum oxygen uptake, which is also related to maximum work capacity. On the other hand, many jobs do not require maximum physical effort, so it is not obvious that energy, or other nutrient intakes, should be correlated with either productivity or labor supply. If, as some claim, the body adapts to changes in energy intakes so that functioning is kept intact, then it is only at extremely low levels of nutrient intakes that productivity or labor supply should suffer;[9] this suggests the relationship between intakes and wages may be very nonlinear.

An important contribution of this study is the simultaneous examination of four different dimensions of health. There are biomedical reasons to expect that these measures will not all affect wages in the same way. For example, current weight or BMI is likely to have a different effect than current calorie intake, even if the body is being drawn upon as a source of energy expenditure. This is because the conversion of body fat into calories may not be an equally efficient source of energy as current caloric intake. Furthermore, conditional on calorie intake, BMI may serve as a source of strength which may be important for some

[8] BMI is defined as weight (in kilograms) divided by height (in meters) squared. Extreme values (below 20 and above 30 or so) are associated with risk of adult mortality (Waaler, 1984; Fogel, 1994).

[9] See Payne (1992) for a recent exposition of this view and Gopalan (1992) or Dasgupta (1993) for the counter-view. The empirical evidence in support of the adaption hypothesis is far from overwhelming; see, for example, Bhargava (1992), Schultink et al. (1993), and Spurr et al. (1994) who provide evidence against the hypothesis.

jobs. It is also likely that current calorie intake will operate differently from protein intakes, and these effects may differ across the intake distribution. Furthermore, Foster and Rosenzweig (1992, 1994) point out there may be economic reasons for labor market rewards to vary with health measures in the event that productivity is not costlessly observable, because some health measures are more readily observed (such as BMI) by an employer than others (such as nutrient intakes). To speak to this issue, in part, separate analyses are presented for workers in the market sector and the self-employed.

Whether the four health measures do have different effects on wages is fundamentally an empirical issue; to the extent they do, then examining them in combination permits testing several more subtle hypotheses. For example, holding calories constant, higher protein diets will be more expensive and thus of higher quality (in terms of not only proteins but also other nutrients); it will thus be possible to assess whether higher-quality diets are rewarded in the labor market. Similarly, holding nutrient intakes constant, the effects of weight and height on wages can be interpreted as reflecting the effect of factors such as strength. Further evidence regarding these issues will be provided by the comparisons below of men and women, as well as comparisons of the effects of health across the education distribution.

4. Data

The data are drawn from a broad-purpose household budget survey, *Estudo Nacional da Despesa Familiar* (ENDEF), collected from August 1974 through August 1975 in Brazil. The survey gathered detailed information from about 53,000 households on incomes, expenditures, and sociodemographic characteristics of household members.

Every respondent (over 14 years of age) reported labor supply and income for the previous year. In addition to annual earnings in each job, the respondent reported the number of hours worked during the previous month and whether or not the person was working in that job for each month of the previous year. An implied wage for each respondent aged 25 to 50 is calculated, assuming the hours reported for the previous month reflect the average for the year. We distinguish wages earned in the market and self-employed sectors because returns to health may differ between the sectors, because wages may not be fully comparable across the sectors, and because labor markets may be segmented.

Self-employment income is notoriously difficult to measure; in ENDEF, the enumerators sought to obtain income net of business expenses. Our measure of self-employment earnings may, however, reflect returns on capital used in that enterprise. It is not entirely clear how to attribute income from family enterprises in which more than one household member is working. Since this is a significant problem in the rural sector, where many households operate farms, we focus only on the urban sector in this paper and use data on 16,169 men and 17,925 women from the South and Northeast of Brazil.[10]

In addition to expenditures on food and nonfood items, the survey gathered very detailed information on quantities of foods purchased and consumed. Each household was visited on seven consecutive days, and foods consumed were weighed, with careful account being taken of wastage. Over 300 foods were distinguished, and these intake quantities were then converted into nutrients using standard composition tables. The number of people eating at every meal during the week was recorded and used to convert nutrients into a *per capita* basis. For this paper, we use *per capita* calorie and protein intake of household members as measures of nutrient intake. Ideally, we would like to observe individual nutrient intakes but, as it is, this was a very intrusive survey and it is hard to imagine being able to measure nutrient intakes in this way at an individual level for an entire week. Furthermore, only foods consumed at home were measured; we assume that foods consumed inside and outside the home have the same nutrient content.[11] Clearly, these measures of nutrient intakes are error-ridden, and we attempt to take account of this fact in the empirical work below.

The height, weight, and arm circumference of every individual in the survey, including adults, was measured by an anthropometrist. We use height and BMI as two additional indicators of health (and nutritional) status, which are not as prone to measurement error as nutrient intakes.

We create unit prices for foods from the expenditure and quantity data. Some of the variation in household-level prices may result from measurement error and some from quality variation (Deaton, 1988). Thus, we take median prices for 135 commodities for market areas defined by state and by whether the urban area is metropolitan. We then create Tornquist indices for 15 commodity groups, including 11 food groups. Ten of the food price indices are used as instruments for BMI and nutrient intakes.[12]

The sample is described in detail in Thomas and Strauss (1995). Of particular interest is the fact that a large fraction of men and women in Brazil are in poor

[10] Less than 5% of individuals reporting some self-employment income live in households with unpaid family workers, who are presumably working in the same enterprise. We ignore this problem and treat reported income as income earned by that individual. Women who were pregnant or lactating at the time of the survey are also excluded from the analysis.

[11] This is a strong assumption and will be violated if employers provide meals to their workers which are more (or less) nutritious than those they eat at home.

[12] These are rice, wheat, tubers, beans, sugar, fruits and vegetables, meat, dairy, fish, and oils and fats. The aggregate price index is formed from these subindices plus other foods, clothing, fuel, household goods, and housing. Because there was 32% annual inflation, all indices and earnings are deflated to August 1974 using a national monthly deflator. There is substantial price heterogeneity even between the Northeast and South. For example, the log price of wheat is 0.012 in the South but -0.121 in the Northeast (standard errors are 0.001). The log price differential on meat is smaller (0.051 in the South and 0.055 in the Northeast), although even this difference is significant ($t = 5.4$). There is also heterogeneity within regions since prices are defined at the level of the state-strata.

health, at least according to these measures. For example, a common reference point is two standard deviations below the median height of an adult in the United States. In the United States, about 2.5% of men and women fall below this line. In Brazil, however, nearly 25% of urban adults are below the line. Furthermore, a substantial fraction of adults have very low weights given their stature, as indicated by the fact that one-fifth of women and one-sixth of men have very low BMIs (below 20). Corroborating evidence for significant malnutrition problems are reflected in the nutrient intake data. Thus, it would not be surprising to observe health effects on wages in ways that might not be likely in North America or Europe.

Nonparametric estimates of the relationships between wages and health are displayed graphically in Thomas and Strauss (1995). Wages and height are positively correlated, with the relationship being linear in logs for both men and women. Wages and BMI are positively correlated among men whose BMI lies between about 20 and 26, but there is no correlation among those who are very light or heavy, given height; the positive relationship is only evident among women with low BMIs. The relationship between wages and nutrient intakes are concave, flattening out when calorie intakes are over 2,400 and protein intakes are over 90 grams. These patterns are consistent with those suggested by nutritionists. However, no account is taken of the potentially confounding effect of observed and unobserved heterogeneity and measurement error. Thus, we turn next to regression estimates which attempt to control for these factors.

5. Empirical results

Of the four health indicators discussed above, only one, height, might be thought of as predetermined by the time an individual reaches adulthood. We treat the other three, BMI, *per capita* calories, and *per capita* protein intakes, as jointly determined with wages and also allow nutrient intakes to be measured with error. Instrumental variables estimates of (ln) wage functions are reported separately for men and women. Following the model outlined above, relative food prices (for ten commodity groups) along with (polynomials in) nonlabor income of the worker and nonlabor income of all other household members are assumed to be the identifying instruments. To control for selection into the labor force and the choice between the market and self-employment sectors, the regressions include estimated hazard rates based on multinomial logits (Heckman, 1974; Lee, 1983) using the same set of identifying instruments.[13]

[13]About 60% of the women in the sample do not participate in the labor market, and among those who do work, two-thirds are employed in the market sector and one-third are self-employed. This trichotomous choice is estimated as a multinomial logit and the appropriate hazard rates included in each of the sector-specific wage functions. Since all but 95% of men are working, we model only the choice between self-employment and market work for them, again using a logit.

In this model, calculation of the variance–covariance matrix is not straightforward and must take into account the endogeneity of BMI and nutrient intake variables as well as the stochastic nature of the estimated hazard rate. In addition, it is prudent to allow for general forms of heteroskedasticity in both the first and second estimation stages (and, as discussed below, possible cluster effects). All of these goals can be accomplished simultaneously by applying the method of the bootstrap (Efron, 1982). The bootstrap is fundamentally data-driven and has the advantage of imposing no *a priori* structure on the distribution of the errors in the model.[14] Cluster effects do not appear to be important in these models, and so the main results do not take them into account; they are discussed further below.

5.1. First-stage results

We begin with a brief description of the first-stage results and focus, in particular, on the predictive power of the identifying instruments.[15] Not only do better predictions in the first stage improve efficiency, but very poor predictions can lead to wildly incorrect inferences.[16] Useful summary statistics for assessing the potential biases in the second stage are the first-stage F-statistics for the significance of the identifying instruments.[17] These are reported in Table 1.

In all cases, the F-statistics on prices (with and without nonlabor income) are high. Nonlabor income is a significant predictor of health in all but two cases: BMI of females and household calorie intakes in the male sample. However, in these two cases relative prices do a good job of prediction. In spite of these fairly large F-statistics, a good deal of unexplained heterogeneity remains, as indicated by the low R^2s, particularly for BMI. The regression results should be interpreted with this fact in mind.

[14] In each replication, a new random sample is drawn from the data and the first-stage and second-stage estimates are computed and stored; the variance in the latter estimates is the basis for our estimated variance–covariance matrix. The appropriate choice of number of bootstrap replications is an empirical issue. It appears that in these models, the estimated variance–covariance matrix stabilizes after 100–200 replications; all estimates are based on 400 replications. Increasing the number of replications to 1,000 (for men) resulted in estimated standard errors changing only in the third or fourth decimal place.

[15] Since the first-stage health regressions are linear in parameters, variance–covariance matrices are estimated using the infinitesimal jackknife (Huber, 1967; White, 1982). For a description of the results and all the estimates, see Thomas and Strauss (1995). In view of the large sample sizes, the reader may prefer to adopt a Bayesian criterion, such as suggested by Schwarz (1978), which picks the *a posteriori* most likely model. According to that criterion, the critical value of the χ^2 is $r\ln N$ where r is the number of restrictions and N is the sample size.

[16] The issue has received considerable attention recently; see Bhargava and Sargan (1983), Nelson and Startz (1990), Bhargava (1991), Staiger and Stock (1994), and Bound, Jaeger, and Baker (1995).

[17] The inverse of the F-statistic is proportional to the bias in the second stage.

Table 1
First-stage F-statistics for significance of identifying instruments

	Males			Females		
	BMI	Calorie intakes	Protein intakes	BMI	Calorie intakes	Protein intakes
Prices & nonlabor income	11.7	10.7	20.6	10.3	15.1	34.6
p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Prices	13.5	14.1	23.7	14.4	18.0	29.7
p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Nonlabor income	6.5	1.9	12.2	0.2	8.6	43.9
p-value	(0.00)	(0.12)	(0.00)	(0.92)	(0.00)	(0.00)
R^2	0.07	0.16	0.16	0.09	0.16	0.17

5.2. *Instrumental variables estimates for male market wages*

Instrumental variables estimates of the impact of health indicators on the wages of men and women in urban Brazil are reported in Tables 2 through 5. Market wage (earnings per hour) functions for men are presented in Table 2. The robustness of these results is examined in Table 3. Table 4 reports results for market wage functions for women and self-employment wage functions for both men and women. In Table 5 the sample is stratified by level of education.

Standard errors are presented below each coefficient estimate. They are followed by Durbin–Wu–Hausman χ^2 statistics for endogeneity of the health measures, GMM tests for overidentification, and tests for the joint significance of sets of covariates; p values are reported below these tests statistics. Unless stated otherwise, all test statistics are based on the bootstrapped estimates of the variance–covariance matrix.

In addition to the covariates listed, the regressions include age, its square, an aggregate commodity price index to account for regional heterogeneity in price levels, and levels of infrastructure.[18] Seasonal variation (and also the effect of

[18] If the overall price index captures only price-level variation, then its estimated coefficient should be one. Deviations from one should reflect the impact of demand-side factors. It is close to unity in all regressions. It is also very well determined (with t-statistics around 20 in most cases) and, in some cases, it is significantly different from unity. For example, in the first regression in Table 2, the coefficient on the price index is 1.10 (with a standard error of 0.05); when all the health indicators are included in the regression, the coefficient on the price index has declined to 0.95 (with a standard error of 0.11).

inflation) is picked up by a set of month dummies. A dummy identifying those who live in the poorer Northeast of Brazil is included to control for ethnic differences (which may be related to height and BMI but are not measured in the survey) and to also account for variation in prices, labor demand, and infrastructure not captured in the overall price index. All covariates reported in the tables (apart from education and the hazard rate) are in logarithms and so the coefficient estimates can be interpreted as elasticities.

In the survey, education is only reported in groups: three education dummies are included (with illiterates being the excluded category). Perhaps one of the most striking empirical facts about earnings and wage functions in Brazil is that education explains a very large fraction of the variation in wages (Lam and Levison, 1991; Strauss and Thomas, 1991). The first column of Table 2 includes only education (along with the covariates listed above): almost half the variation in ln (market wages) is accounted for in these regressions. Relative to an illiterate worker in the market sector, a literate man earns about 50% more, a man with elementary schooling earns about 130% more, and having at least secondary schooling results in nearly 550% higher wages. The hazard rate for sectoral choice is significantly positive, indicating that those with unobserved characteristics that raise wages are more likely to enter market work.

Height, which is treated as predetermined, is added to the model in the second column. Like schooling, height has a positive and significant effect on wages. A 1% increase in the height of a man is associated with a 2.4% increase in his wage, an estimate which is precisely determined. It is useful to place the magnitude of this elasticity in perspective by comparing it with the effect of education on wages. For example, to have the same effect on log wages as becoming literate (without completing primary schooling), log male height would have to increase by 0.16, which is equivalent to nearly 16 standard deviations in the data. To earn the same wage, an illiterate would thus have to be about 30 cm taller than a literate male, *ceteris paribus*.

The addition of height in the male market log-wage equations changes the education coefficients only slightly. Thus, although height is presumably capturing the effect of both cumulative health status as well as past nonhealth human capital investments, it does not seem to be highly correlated with the schooling dummy variables. On the other hand, the hazard rate coefficient halves and, since the standard error is unchanged, it is no longer significantly different from zero, indicating that height is capturing part of the unobservable factors that affect sectoral choice. Given that the education groups over which the dummy variables are defined are quite broad (and dictated by the nature of the data), it is also possible that height may be capturing some component of either years or quality of schooling that varies within these groups. With these data, it is not possible to disentangle these sources.

In the third column, the logarithm of BMI is added to the regression. BMI is also positively associated with market wages for men. Not only do taller men earn higher wages, but so do heavier men. This may reflect greater physical

Table 2
Males in market sector: Impact of health characteristics on ln(wages)

Covariates	No health (1)	Height only (2)	Add BMI (3)	Add calories (4)	Add protein (5)	All health (6)
ln(height)	.	2.431 (0.17)	2.407 (0.17)	2.832 (0.44)	1.437 (0.29)	3.921 (0.98)
ln(body mass index)	.	.	2.223 (1.08)	.	.	4.740 (2.29)
ln(per capita calories)	.	.	.	88.763 (35.94)	.	163.759 (74.75)
– squared	.	.	.	– 5.860 (2.37)	.	– 10.964 (4.96)
ln(per capita protein)	.	.	.	.	27.537 (13.67)	– 28.848 (29.73)
– squared	.	.	.	.	– 2.049 (1.06)	2.301 (2.29)
Education						
(1) literate	0.398 (0.02)	0.391 (0.02)	0.338 (0.03)	0.262 (0.07)	0.201 (0.06)	0.223 (0.08)
(1) elementary	0.830 (0.03)	0.803 (0.02)	0.709 (0.05)	0.636 (0.09)	0.484 (0.08)	0.515 (0.10)
(1) secondary +	1.867 (0.03)	1.791 (0.03)	1.642 (0.09)	1.606 (0.12)	1.372 (0.10)	1.338 (0.13)
Hazard rate	0.337 (0.12)	0.140 (0.12)	0.041 (0.13)	0.222 (0.13)	0.215 (0.14)	0.104 (0.20)
Tests for						
Endogeneity	.	.	329.34 (0.00)	123.60 (0.00)	297.74 (0.00)	882.07 (0.00)
Overidentification	.	.	27.98 (0.00)	19.61 (0.00)	17.85 (0.00)	6.17 (0.00)
Joint significance						
Education *p*-value	6019.83 (0.00)	5643.11 (0.00)	675.28 (0.00)	692.87 (0.00)	483.49 (0.00)	146.75 (0.00)
Calories *p*-value	.	.	.	6.10 (0.05)	.	7.78 (0.02)
Protein *p*-value	.	.	.	.	21.59 (0.00)	9.68 (0.01)
Nutrients *p*-value	.	.	.	6.10 (0.05)	21.59 (0.00)	25.33 (0.00)
BMI & Intakes *p*-value	.	.	.	.	.	27.02 (0.00)
All health *p*-value	.	202.47 (0.00)	196.64 (0.00)	140.57 (0.00)	183.71 (0.00)	108.54 (0.00)
All covs *p*-value	64166.54 (0.00)	68825.11 (0.00)	65027.27 (0.00)	47622.54 (0.00)	61065.45 (0.00)	26345.47 (0.00)

BMI, calories, and protein and sectoral choice are treated as endogenous; identifying instruments are ten relative food prices, own nonlabor income, its square, nonlabor income of other members and its square. All regressions include controls for age, aggregate price index, region, and month of survey. Standard errors below coefficient estimates; *p*-values below test statistics. Variance–covariance matrix estimated by method of the bootstrap. Endogeneity is Durbin–Hausman–Wu χ^2; overidentification is GMM χ^2. Sample selection hazard rate based on logit for all men; there are 10,675 males working in the market sector.

strength and greater maximum physical capacity; it probably also reflects better nutrition and well-being.

The elasticity is 2.2. What does this mean in practical terms? The effect of going from the bottom decile of predicted BMI, 20, to the top decile, 24, is associated with an increase in log wages of 0.4. This is slightly larger than the impact of becoming literate, but far less than the benefit associated with completing primary school.

Adding BMI to the model does not perceptibly change the estimated impact of height on wages, although the schooling coefficients do decline, especially among men at the bottom of the education distribution. The coefficient on the hazard rate declines dramatically and is now zero. Apparently height and BMI capture much of the unobserved heterogeneity in worker characteristics that are related to sectoral choice.

Recall that BMI is weight divided by the square of height and so, as pointed out by Bhargava (1994), including both the log of height and BMI in a linear model is equivalent to including the log of height and weight separately: one is simply a linear transformation of the other. Moving to a nonlinear model, however, does impose restrictions on the data, as also pointed out by Bhargava. We have experimented with a number of different specifications; estimates (with standard errors) based on a quadratic form in log height ($\ln ht$) and weight ($\ln wt$) are:

$$\begin{aligned} \ln(wage) = \beta_0 &+ \underset{(856.9)}{117.12}\ln ht + \underset{(211.2)}{21.45}\ln ht^2 - \underset{(515.1)}{37.82}\ln wt + \underset{(22.8)}{6.32}\ln wt^2 \\ &- \underset{(128.0)}{25.29}(\ln ht^*\ln wt) + \beta_6' X + u. \end{aligned} \tag{2}$$

It is easy to show that a quadratic in log height and log BMI imposes the restriction that $\tau = 4\beta_4 + \beta_5 = 0$. Our estimate of τ is only 0.01 and not different from zero. In view of these results, rather than specify the model in terms of height and weight, we prefer to use BMI because of the literature demonstrating links between BMI and maximal oxygen uptake and to subsequent adult mortality.[19]

Having accepted the log BMI specification, the fourth column of Table 2 presents the effect of the first of two measures of nutrient intake: *per capita*

[19] Furthermore, a model which is linear in log(height) and quadratic in log(BMI) would imply that $\beta_3 = 4\beta_4$; that restriction is not rejected, which is not surprising in view of the linearity between log wages and log height discussed above. In fact, we find no evidence for the kinds of nonlinearities described in the biological literature. In part, this is because there is simply not enough variation in our data since the bottom decile of (predicted) BMI is 20 and the top decile is 24, which lies within the range over which BMI is not thought to be extremely low (around 18 or 19) or too high (around 26 or 27).

calorie intake. Calories are only just significant at the 5% size of test. At the bottom of the calorie (and income) distribution, the estimated impact is large relative to the top of the distribution. At the bottom quartile of *per capita* calories (around 1,700 per day), the elasticity is 1.6, but it diminishes rapidly and turns negative around 1,950 calories per day.

Additional protein intake is also associated with higher wages (column 5). Unlike calories, the relationship is upward-sloping through most of the distribution of protein (albeit at a declining rate). At the bottom quartile of the protein distribution, a 1% increase in protein is associated with a 1.9% rise in wages and the elasticity falls to about 0.2 at the top quartile.

These estimates have taken the nutrient intakes separately, but individuals with high protein intakes will generally have high calorie intake as well. Furthermore, calorie and protein needs will vary with body mass. If our four health measures do represent distinct components of health, then it is of interest to examine them in combination, and so the final column of the table includes all the health measures.

Conditional on health, education remains a significant determinant of wages (although standard errors are larger). Relative to returns without conditioning on health status (in column 1), the estimated returns to education are about 45% smaller for literate men and 30% less for men with secondary education or more. Hence, it appears that an important part of the effect of education on wages operates through health, particularly for men with little education. Similarly, the hazard rate coefficient is less than one-third its size in the regression without any health covariates and it is not significant.

Taking all the health measures together, including height, it is clear they are significant predictors of market wages of men. Moreover, each of the health measures is individually significant at the 0.05 level or better. The calorie elasticity is now more nonlinear, further reinforcing the idea that it is only workers with extremely low intakes who benefit from greater calorie intakes given protein intake, height, and BMI. Of particular interest is the fact that the impact of protein on log wages changes shape and becomes positive and *convex*. Conditional on calories, additional protein may be interpreted as indicative of quality of diet and probably reflects, in part, a higher share of animal proteins.[20] Not only is a better diet associated with higher wages, but the effect is largest among those men with the best diets.[21]

Exogeneity of BMI and the nutrient intake variables is unambiguously rejected by Durbin–Wu–Hausman tests in all specifications. The combination

[20]The influence of other nutrients, such as iron or micronutrients, are likely to be subsumed in the protein variable. Basta et al. (1979) show an effect of iron supplementation on labor productivity of male rubber tree tappers.

[21]For example, at the top quartile of the protein distribution, the elasticity is about 1.9.

Table 3
Males in market sector: Specification testing and robustness

Covariates	Baseline (1)	Cluster effects: Excluded (2)	Cluster effects: Included (3)	Instruments exclude nonlabor income (4)
ln(height)	3.921	3.651	3.651	3.355
	(0.98)	(0.58)	(0.58)	(0.33)
ln(body mass index)	4.740	5.596	5.596	1.052
	(2.29)	(1.12)	(1.12)	(1.68)
ln(per capita calories)	163.759	140.778	140.778	121.742
	(74.75)	(38.66)	(38.81)	(94.71)
– squared	– 10.964	– 9.416	– 9.416	– 8.080
	(4.96)	(2.56)	(2.57)	(6.28)
ln(per capita protein)	– 28.848	– 40.328	– 40.328	– 28.914
	(29.73)	(15.90)	(15.97)	(27.83)
– squared	2.301	3.198	3.198	2.277
	(2.29)	(1.23)	(1.23)	(2.15)
Hazard rate	0.104	.	.	– 0.617
	(0.20)			(0.24)
Tests for				
Overidentification	6.17	.	.	3.87
	(0.00)			(0.00)
Joint significance				
χ^2 (Education)	146.75	210.79	210.79	186.67
p-value	(0.00)	(0.00)	(0.00)	(0.00)
χ^2 (Calories)	7.78	19.69	19.59	1.69
p-value	(0.02)	(0.00)	(0.00)	(0.43)
χ^2 (Protein)	9.68	19.72	19.62	3.64
p-value	(0.01)	(0.00)	(0.00)	(0.16)
χ^2 (Nutrients)	25.33	30.82	30.82	3.73
p-value	(0.00)	(0.00)	(0.00)	(0.44)
χ^2(BMI & Intakes)	27.02	54.61	54.47	4.17
p-value	(0.00)	(0.00)	(0.00)	(0.52)
χ^2 (All health)	108.54	123.25	123.19	147.66
p-value	(0.00)	(0.00)	(0.00)	(0.00)
χ^2 (All covs)	26345.47	24010.92	23690.11	34872.47
p-value	(0.00)	(0.00)	(0.00)	(0.00)

See Table 2. Variance–covariance matrix estimated by method of bootstrap in (1) and (4), infinitesimal jackknife in (2), and infinitesimal jackknife with cluster effects in (3).

of unobserved heterogeneity and contemporaneous feedbacks between wages (income) and health are important to take into account in the estimation of these wage–health relationships. Tests of overidentification, however, also fail, although the F-statistic in the final column with all health measures is relatively small. Since the R^2 on the auxiliary regression of residuals on excluded instruments is only 0.007, the large sample size bears a good deal of the responsibility for failure of these tests. As pointed out by Staiger and Stock (1994), when faced with a choice between bias that arises from using instruments which are weakly correlated with the health measures and bias suggested by failure of overidentification tests, it is prudent to not place blind faith in these test statistics. Failure of the tests does, however, raise issues of robustness, a subject to which we now turn.

5.3. Robustness issues

Table 3 examines the robustness of these results for male market workers and focuses on three issues: controlling for sectoral choice, the impact of sampling design (or cluster) effects on the size of standard errors, and variation in the instrument set. Since we are primarily interested in the robustness of the estimated effects of health, all other covariates are suppressed from the table.

The final column of Table 2, which includes all the health variables, is repeated as the baseline in the first column of Table 3. Conditional on these characteristics, the hazard rate for selection into the market sector is not significant, suggesting that ignoring selection bias will not substantially change the estimates. To test this hypothesis, the second column reports estimates of a wage function conditional on participating in the market sector, but without a selectivity control. The estimated health effects in this model are very close to those in the first column, indicating our conclusions are robust to ignoring sectoral choice.[22] Since the conditional model is linear, the variance–covariance matrix is estimated by the instrumental variables version of the infinitesimal jackknife. The standard errors are considerably smaller in the conditional function and lend further credence to our conclusions in the previous section regarding the effect of health on wages.

Survey data are seldom random but typically involve a multi-stage design; if there are common unobservables within clusters, then, as has been pointed out by Moulton (1990), among others, standard errors will be downward-biased. This potential bias is examined in column 3.

Taking account of cluster effects is not straightforward. First, defining the appropriate cluster is ambiguous. For this analysis, it is possible to think about

[22] Within the range of the sample data, the nutrient intake functions are very close. For example, the estimated elasticity at mean protein intake is 0.90 in column 1 and 1.02 in column 2; the standard error on the latter is 0.3.

at least two levels: the state–strata level, at which prices are defined, and the primary sampling unit of about 20 households. The state–strata definition is used in column 3, although calculations using the primary sampling unit yield essentially identical results. Another issue is whether to constrain the correlations within clusters to be equal; while this is common practice, it is clearly restrictive and has little *a priori* justification. We thus allow correlations to vary by cluster. Finally, correcting for within-cluster correlation does not take into account heteroskedasticity, which is often an important issue in analyses using household survey data. Results in column 3 are based on an estimated variance–covariance matrix that takes into account both cluster effects and general forms of heteroskedasticity by incorporating cluster effects in the infinitesimal jackknife (or Huber–White) 2SLS calculations. The differences in standard errors with and without cluster effects are trivial in this instance. Put another way, taking account of heteroskedasticity of a general form appears to perform well in the model even in the presence of a cluster design.[23] Furthermore, in this case, the average intracluster correlation is tiny (0.007) and has a short interquartile range (– 0.0004 to 0.0033).

A third issue is identification of the endogenous health measures. Thus far, ten relative food prices, own nonlabor income (and its square), and nonlabor income of other members (and its square) have served as identifying instruments. It may be argued that nonlabor income reflects previous labor supply choices and so is not a valid instrument. However, recall that food prices are measured at the community level, and so relying on them alone as instruments would leave us with no individual-specific instruments. One might anticipate that predictions in the first stage will be poor and this does, indeed, turn out to be the case. Estimates using only prices as instruments are reported in column 4 of Table 3. There are three main results. First, the estimated effect of BMI on wages collapses to zero while the nutrient intake estimates are remarkably stable; this is probably a reflection of the fact that relative food prices are powerful predictors of intakes but less successful at predicting BMI. Second, the (endogenous) health effects are estimated less precisely and none is significant. Third, the hazard rate is now negative (and significant) indicating negative selection of men into the market sector! Taking all these results together, we judge that it is preferable to include nonlabor income in the instruments but recognize the need for caution in interpreting the results. Our judgement is bolstered by the fact that a Wu–Hausman test for the inclusion of nonlabor income in the instrument set, maintaining prices belong in the set, is not rejected (with a p-value of 0.08). We now turn to a discussion of estimates for other subpopulations.

[23]In these models, Lagrange Multiplier tests for homoskedasticity are rejected; for example, in column 2 the test statistic is 1,063 which is distributed as a χ^2_{25}.

5.4. Estimates for self-employed men and working women

The relationship between wages and health of self-employed men and working women (distinguishing market workers from the self-employed) are reported in Table 4. For the sake of brevity, two specifications are reported in each case: height and BMI are included in the first column of each panel, and all six health variables are included in the second column.

There are sound economic reasons to examine differences across sectors. First, health may be used as a signal by employers in the market sector if pay is not piece-rate, whereas this is less likely to be an issue among the self-employed, for whom wages are probably a closer reflection of productivity. Differences in estimated effects of health may, therefore, be indicative of differences due to signalling and observability (Foster and Rosenzweig, 1994). Second, different dimensions of human capital are likely to be rewarded differently depending on the nature of the work: distinguishing sectors is a first step. Further analyses in this direction are discussed below.

The restriction that there are no differences across the self-employed and market sectors is rejected for men (with a χ^2 for equality of coefficients of 235 in the first column and 167 in the second). For women, equality is rejected in the specification that includes height and BMI.

Height is associated with higher wages for self-employed men, and the effect is bigger than it is for men in the market sector (the difference in elasticity is 0.6 with a standard error of 0.3). The magnitude of the elasticity for women is essentially the same in both sectors but significant only in the market sector, which probably reflects lack of precision in the self-employed sector. The fact that the height elasticity is considerably smaller for women and largest for self-employed men (many of whom are manual laborers) suggests that it reflects, in part, a return to strength. Corroborating evidence is provided by the fact that BMI has a small and insignificant effect on wages of women but a large, significant effect on self-employed males (and the latter is also bigger than the effect on males in the market sector).

The importance of distinguishing sectors is even clearer when comparing the effects of nutrient intakes. As described above, among men in the market sector, there is clear evidence that nutrient intakes affect wages. The evidence for the self-employed is weak. When they are included separately, there is some evidence that calories and protein are associated with higher wages among self-employed men, but these effects decline in magnitude and become insignificant when controls for BMI are included in the regression. For the self-employed, it seems that strength is far more important than quality of diet. The evidence for women is broadly similar. Calorie and protein intakes have a significant impact on wages in the market sector, and with similar nonlinear shapes found for men, but they have no impact on the wages of the self-employed. These results contrast with Foster and Rosenzweig (1994) who report that in the rural

Table 4
Males and females: Self-employed and market sector workers

Covariates	Males, self-employed: Hgt & BMI (1)	Males, self-employed: All health (2)	Females, market sector: Hgt & BMI (1)	Females, market sector: All health (2)	Females, self-employed: Hgt & BMI (1)	Females, self-employed: All health (2)
ln(height)	3.085	3.580	2.089	2.458	2.003	−1.002
	(0.41)	(1.50)	(0.32)	(0.67)	(1.49)	(3.40)
ln(body mass index)	4.943	5.177	1.292	−0.412	0.516	−3.918
	(1.52)	(2.78)	(0.78)	(1.44)	(3.43)	(6.37)
ln(per capita calories)	.	113.431	.	186.68	.	68.686
		(101.85)		(75.11)		(244.57)
– squared	.	−7.547	.	−12.415	.	−4.405
		(6.76)		(4.94)		(16.06)
ln(per capita protein)	.	−10.910	.	−54.237	.	−51.579
		(40.51)		(30.62)		(103.81)
– squared	.	0.901	.	4.303	.	4.124
		(3.12)		(2.37)		(103.81)
Education						
(1) literate	0.520	0.400	0.447	0.368	0.409	0.344
	(0.07)	(0.12)	(0.05)	(0.10)	(0.20)	(0.50)
(1) elementary	0.958	0.759	0.976	0.801	0.774	0.461
	(0.09)	(0.17)	(0.05)	(0.15)	(0.17)	(0.74)
(1) secondary +	1.762	1.469	2.081	1.716	1.720	1.039
	(0.13)	(0.21)	(0.08)	(0.26)	(0.31)	(0.90)
Hazard rate	−0.496	−0.532	0.155	−0.177	1.109	1.023
	(0.18)	(0.22)	(0.11)	(0.19)	(0.80)	(1.17)

Tests for

Endogeneity	205.68	404.09	12.30	148.88	36.00	149.77
p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Overidentification	4.60	2.66	4.74	1.23	3.87	1.21
p-value	(0.00)	(0.00)	(0.00)	(0.24)	(0.00)	(0.26)
Joint significance						
Education	226.56	71.28	875.82	44.01	42.02	2.64
p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.45)
Calories	.	1.33	.	14.50	.	0.29
p-value		(0.51)		(0.00)		(0.87)
Protein	.	2.54	.	17.16	.	1.35
p-value		(0.28)		(0.00)		(0.51)
Nutrients	.	6.57	.	21.26	.	1.97
p-value		(0.16)		(0.00)		(0.74)
BMI & Intakes	.	16.54	.	22.71	.	1.98
p-value		(0.01)		(0.00)		(0.85)
All health	70.57	62.83	48.16	60.57	2.23	3.98
p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.33)	(0.68)
All covs	16553.74	12889.18	16730.04	12599.46	399.23	309.67
p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Sample size	4403		4948		2653	

See Table 2.

Philippines, nutrients are not rewarded in daily-wage-remunerated market work but are in piece-rate work and argue that this is because employers cannot observe what workers (and their families) eat. However, employers can observe the outputs, such as better general health, improved pallor and higher levels of energy and effort, and those indicators may be used in setting wages of their workers. Moreover, as we will see below, the fact that wages of the self-employed do not respond to diet may reflect differences in health effects for different types of work.

Education is a powerful predictor of the wages of self-employed men, having a larger impact than among men in the market sector, especially those with little education. For women the reverse is true: schooling returns are higher in the market sector. In fact, in the market sector, the return to female schooling is higher than for men. The hazard rate coefficients are negative and significant for men, indicating an absolute advantage for men who enter the market sector. Unlike the market sector results, the hazard term remains significant for self-employment wages even when health variables are included in the model. Women in both the market and self-employment sectors evidently command a higher wage than a randomly chosen woman. For women in the market, as for men, the inclusion of health characteristics reduces the magnitude of the selection term and it becomes insignificant. The selection term remains large (albeit imprecisely estimated) among self-employed women. In all cases, exogeneity of BMI and nutrient intakes is rejected. The evidence on overidentification is mixed. Among self-employed men, the restrictions are rejected, although the F-statistic is fairly low, especially in the specification including all health measures. For women, the overidentification restrictions cannot be rejected in those models.

In sum, taking all health indicators together, the wages of men and women who work in the market sector are significantly associated with better health. Among self-employed women, there is little evidence that health *per se* matters and, in fact, there is no evidence that even height is associated with higher productivity of women working in this sector. Among self-employed men, however, the evidence is less clear-cut. Certainly height and possibly BMI are positively associated with productivity. There is some evidence that nutrient intakes affect wages, but these independent effects seem to disappear when we control for mass. In part, this reflects the fact that the three measures of current health status are correlated with one another and their predicted values are even more closely correlated:[24] taking BMI, calorie, and protein intakes together, then there is evidence that better health is associated with productivity gains. The addition of height to the set of health indicators serves to strengthen this argument.

---

[24]For example, among men, the correlation between BMI and either protein or calorie intake is about 0.65; the correlation between the two nutrient intakes is 0.80.

5.5. Estimates by level of education

The effects of (these indicators of) health status on labor market outcomes are likely to vary, depending on the kind of work in which an individual is engaged. A laborer, for example, would presumably reap returns from strength and stature, whereas those characteristics are unlikely to be rewarded, in and of themselves, in a more sedentary occupation. Health status and choice of occupation are likely to be jointly determined, thus adding another layer of simultaneously determined covariates. Unfortunately, the ENDEF data do not provide large enough samples to estimate meaningful relationships between health and wages for a particular occupation. The nature of one's work, however, is likely to be related to education: among male workers, the work of 80% of those who are illiterate involves moderate or heavy activity, whereas only 20% of those with secondary schooling or more are in this sort of work. Among female workers, 90% of those with secondary schooling are involved in light activities, but only 10% of women who are illiterate are engaged in light tasks. We have, therefore, stratified on four education categories in order to determine whether returns to health vary over the education (and thus, presumably, the occupation) distribution.

Since samples are quite small, we focus on the anthropometric measures and report results for men and women in Table 5. For the same reason, all women having completed primary schooling who are working in the self-employed sector are grouped together.

Among men in the market sector, the return to stature tends to rise with education, although among self-employed men the pattern is, if anything, reversed. If height is capturing human capital investments over and above (gross measures of) education, then the market wage results suggests that those investments may be larger among the better educated. Among women working in the market sector, there is no clear pattern in the return to stature across the education distribution although the evidence among self-employed women parallels that for men.

If BMI represents strength alone, we would expect its effect to decline with education, since occupations of those with little education are likely to be more physically demanding than the work of the better educated. There is some evidence to support this view. In the self-employed sector, the impact of BMI is large and significant for illiterate men but declines with education and becomes insignificant for men with at least elementary schooling. In the market sector, it is only illiterate men who are rewarded for greater mass, given height. A similar pattern emerges for women, although the coefficients for the self-employed are imprecisely estimated. Hence once we stratify by schooling level, we do find a significant, positive impact of BMI, at least for women working in the market sector in contrast with the pooled results discussed above. This is consistent with the notion that women who have little education tend to work in strenuous occupations such as domestic service.

Table 5
Males and females: Effect of health on ln(wages) by level of education

	Males: Education level				Females: Education level			
	Illiterate (1)	Literate (2)	Completed elementary (3)	Completed secondary (4)	Illiterate (1)	Literate (2)	Completed elementary (3)	Completed secondary (4)
ln(market wage)								
ln(height)	1.265	1.940	2.096	4.164	1.842	2.605	1.081	1.580
	(2.48)	(6.86)	(5.17)	(7.42)	(2.49)	(4.11)	(1.36)	(2.95)
ln(body mass index)	3.100	1.148	−0.460	1.302	2.611	2.925	−2.676	−0.693
	(2.51)	(1.22)	(0.31)	(0.41)	(2.17)	(2.53)	(1.03)	(0.53)
χ^2(Hgt & BMI)	10.32	47.35	26.79	75.95	8.53	16.96	8.73	12.93
p-value	(0.01)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)
χ^2(All covs)	2070.16	19003.02	15767.92	22143.89	223.33	1033.62	1878.42	12758.38
p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Sample size	1159	4161	2977	2378	658	1488	1088	1714
ln(self-employment wage)								
ln(height)	4.023	2.627	3.080	2.376	0.770	2.302		3.263
	(3.45)	(4.70)	(3.93)	(2.52)	(0.68)	(2.74)		(2.68)
ln(body mass index)	5.590	3.126	2.221	3.633	2.916	1.487		−3.477
	(2.30)	(2.11)	(0.97)	(1.03)	(1.49)	(0.77)		(0.03)
χ^2(Hgt & BMI)	16.72	24.14	17.31	9.15	2.55	7.80		11.72
p-value	(0.00)	(0.00)	(0.00)	(0.01)	(0.28)	(0.02)		(0.00)
χ^2(All covs)	358.42	4051.00	4638.68	5910.62	97.60	252.34		669.84
p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)		(0.00)
Sample size	550	1943	1121	789	568	1225		860

See Table 2.

6. Conclusions

The influence of health on wages of men and women in urban Brazil is not unidimensional. We have used four distinct measures: height, body mass index, *per capita* calorie intake and *per capita* protein intake, For adults, height is predetermined and reflects cumulative health as well as other human capital investments made by the individual's parents. Height may also have a direct effect on wages through strength. The other covariates all depend, at least in part, on current-period allocations of the individual. They are treated as being determined simultaneously with wages along with sectoral choice. Identification is based on the assumption that relative food prices and nonlabor income have no direct effect on wages.

In urban Brazil, height is clearly an important determinant of wages of men and also of those women who work in the market sector. Body mass is a significant positive predictor of male wages, even after controlling for height and nutrient intakes, and these impacts are generally largest among the less-educated. This suggests that mass is of value in particular types of work. For women, BMI is only related to wages for the least educated. *Per capita* calorie and protein intakes are also significantly related to wages of both men and women who work in the market sector. This is true even after holding constant height and BMI. The positive effect of calories disappears rather rapidly, consistent with calories being important only for the very malnourished. On the other hand, holding calories constant, protein intakes do positively influence wages, and more so at higher levels of intake. This probably represents an effect of diet quality, of which protein (and its quality) is an important dimension.

If the assumptions of the empirical model are correct, then taking these results together, the evidence suggests that health, as measured here, yields a substantial return in the formal sector of Brazilian labor markets. Different dimensions of health affect wages differentially, and a nontrivial part of the impact of education on wages appears to operate through these measures of health.

References

Basta, S., Soekirman, D. Karyadi, and N. Scrimshaw, 1979, Iron deficiency anemia and productivity of adult males in Indonesia, American Journal of Clinical Nutrition 35, 916–925.

Bhargava, A., 1991, Identification and panel data models with endogenous regressors, Review of Economic Studies 58, 129–140.

Bhargava, A., 1992, Malnutrition and the role of individual variation with evidence from India and the Philippines, Journal of the Royal Statistical Society A 155, 221–231.

Bhargava, A., 1994, Modeling the health of Filipino children, Journal of the Royal Statistical Society A 157, 417–452.

Bhargava, A. and J.D. Sargan, 1983, Estimating dynamic random effects models from panel data covering short time periods, Econometrica 51, 1635–1660.

Bound, J., D. Jaeger, and R. Baker, 1995, Problems with instrumental variable estimation when the correlation between the instruments and the endogenous explanatory variable is weak, Journal of the American Statistical Association 90, 443–450.

Dasgupta, P., 1993, An inquiry into well-being and destitution (Oxford University Press, Oxford).

Deaton, A., 1988, Quality, quantity and spatial variation of price, American Economic Review 78, 418–430.

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

Efron, B., 1982, The jackknife, the bootstrap and other resampling plans, SIAM CBMS-NSF monograph 38.

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

Foster, A. and M. Rosenzweig, 1992, Information flows and discrimination in labor markets in rural areas in developing countries, Proceedings of the annual conference on development economics: Supplement to the World Bank Economic Review 6, 173–204.

Foster, A. and M. Rosenzweig, 1994, A test for moral hazard in the labor market: Effort, health and calorie consumption, Review of Economics and Statistics 76, 213–227.

Gopalan, C., 1992, Undernutrition: Measurement and implications, in: S.R. Osmani, ed., Nutrition and poverty (Clarendon Press, Oxford) 17–48.

Haddad, L. and H. Bouis, 1991, The impact of nutritional status on agricultural productivity: Wage evidence from the Philippines, Oxford Bulletin of Economics and Statistics 53, 45–68.

Heckman, J., 1974, Shadow prices, market wages and labor supply, Econometrica 42, 679–694.

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

Huber, P.J., 1967, The behavior of maximum likelihood estimates under nonstandard conditions, Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability 1, 221–233.

Lam, D. and D. Levison, 1991, Declining inequality in schooling in Brazil and its effects of inequality in earnings, Journal of Development Economics 37, 199–226.

Lee, L.-F., 1983, Generalized econometric models with selectivity, Econometrica 51, 507–512.

Martorell, R. and G. Arroyave, 1988, Malnutrition, work output and energy needs, in: K. Collins and F. Roberts, eds., Capacity for work in the tropics (Cambridge University Press, Cambridge) 57–76.

Moulton, B., 1990, An illustration of a pitfall in estimating the effects of aggregate variables on micro units, Review of Economics and Statistics 72, 334–338.

Nelson, R. and R. Startz, 1990, The distribution of the instrumental variables estimator and its t-ratio when the instrument is a poor one, Journal of Business 63, S125–S140.

Payne, P., 1992, Assessing undernutrition: The need for a reconceptualization, in: S.R. Osmani, ed., Nutrition and poverty (Clarendon Press, Oxford) 49–91.

Pitt, M. and M. Rosenzweig, 1986, Agricultural prices, food consumption and the health and productivity of Indonesian farmers, in: I. Singh, L. Squire, and J. Strauss, eds., Agricultural household models: Extensions, applications and policy (Johns Hopkins University Press, Baltimore, MD) 153–182.

Rosenzweig, M., 1988, Labor markets in low income countries, in: H. Chenery and T.N. Srinivasan, eds., Handbook of development economics, Vol. 1 (North-Holland Press, Amsterdam) 714–762.

Rosenzweig, M. and K. Wolpin, 1986, Evaluating the effects of optimally distributed programs, American Economic Review 76, 470–482.

Rosenzweig, M. and K. Wolpin, 1988, Migration selectivity and the effects of public programs, Journal of Public Economics 37, 265–290.

Schultink, J., J. Vanraaij, and J. Hautvast, 1993, Seasonal weight loss and metabolic adaption in rural Beninese women: The relationship with body mass index, British Journal of Nutrition 70, 689–700.

Schultz, T.P. and A. Tansel, 1992, Measurement of returns to adult health: Morbidity effects on wage rates in Côte d'Ivoire and Ghana, Economic Growth Center discussion paper no. 663 (Yale University, New Haven, CT).

Schwarz, G., 1978, Estimating the dimension of a model, Annals of Statistics 6, 461–464.

Singh, I., L. Squire, and J. Strauss, 1986, Agricultural household models: Extensions, applications and policy (Johns Hopkins University Press, Baltimore, MD).

Spurr, G., 1983, Nutritional status and physical work capacity, Yearbook of Physical Anthropology 26, 1–35.

Spurr, G., D. Dufour, J. Reina, R. Hoffman, C. Waslien, and L. Staten, 1994, Variation of the basal metabolic rate and dietary energy intake of Colombian women during 1 y, American Journal of Clinical Nutrition 59, 20–27.

Staiger, D. and J. Stock, 1994, Instrumental variables regression with weak instruments, Mimeo. (Kennedy School of Government, Harvard University, Cambridge, MA).

Strauss, J. and D. Thomas, 1991, Wages, schooling and background: Investments in men and women in urban Brazil, Economic Growth Center discussion paper no. 649 (Yale University, New Haven, CT).

Strauss, J. and D. Thomas, 1995, Human resources: Empirical modeling of household and family decisions, in: J. Behrman and T.N. Srinivasan, eds., Handbook of development economics, Vol. 3 (North-Holland Press, Amsterdam) 1883–2023.

Thomas, D. and J. Strauss, 1995, Health, wealth and wages of men and women in urban Brazil, RAND labor and population working paper 95-11 (RAND, Santa Monica, CA).

Waaler, H., 1984, Height, weight and mortality, Acta Medica Scandinavia, Supplement no. 679 (Stockholm).

Ware, J., 1987, Standards for validating health measures: Definition and content, Journal of Chronic Diseases 40, 473–480.

White, H., 1982, Instrumental variables regression with independent observations, Econometrica 50, 483–500.