Health and Poverty in Brazil: Estimation by Structural Equation Model with Latent Variables

Antonio Giuffrida
Roberto F. Iunes
William D. Savedoff

Technical Note on Health No. 1/2005

March 2005

Inter-American Development Bank
Washington, D.C.

Antonio Giuffrida is a Health Specialist at the Social Programs Division, Regional Operations Department 3. Roberto F. Iunes is a Senior Evaluation Officer at the Office of Evaluation and Oversight. William D. Savedoff is a Senior Partner at Social Insight. *We are grateful to André Medici and Marie Moland Gaarder for helpful comments.* *Comments can be sent to Antonio Giuffrida, Inter-American Development Bank, 1300 New York Avenue NW, Mail Stop E-704, Washington, DC 20577, USA. Email: ANTONIOGI@iadb.org.* *To obtain access to this and other related publication at the Inter-American Development Bank, visit our website at http://www.iadb.org/sds/soc.*

Foreword

Brazil is characterized by large inequalities in income, access to health care, and health status. This paper uses data from the 1998 Pesquisa Nacional por Amostra de Domicílios to analyze the complex relationships among health, income, health insurance and health care utilization to gain a better understanding of the various factors determining differences in the health of the population, especially between the poor and non-poor. We use a structural equation model in which health, wealth and access to health care are specified as latent variables. Latent variables are not directly measurable, but are represented by a set of manifest variables, which act as indicators. Separate models are estimated for adult women and men. The estimates suggest that health of adult women and men is improved by water, sanitation and particularly by filtering the water used in the house. Education and wealth improve health status. On the other hand, unemployment, child labor and race discrimination affect negatively health status. The health status of women looks particularly worse compared with men. Women's health status depreciates at a quicker pace and appears to be much more sensitive to negative factors.

.

Contents

Introduction

Health status is important in developing countries, both as a direct indicator of welfare and because of its impact on productivity (see Savedoff and Schultz, 2000). Health care utilization also is of interest, both because of its relation to health status and because it is a particular manifestation of demand and supply of the health system.

However, work on the determinants of health status is problematic because true health status is not directly observable. The indicators of health and nutritional status that have been used in empirical studies (anthropometric measures, days ill, self-reported or clinical disease records, inputs and nutrients, and health care utilization indices) are imperfect indicators of the underlying health condition and generally do not give a congruent representation of that status. This may lead to biases in estimates and incorrect interpretations (Manning et al., 1982).

In this paper, a structural equation model of determinants of health and health care utilization in Brazil is developed using cross-sectional sample survey data. The data set used in the analysis are derived from the 1998 Pesquisa Nacional por Amostra de Domicílios (PNAD-1998). Our approach of dealing explicitly with the complex relationship among health, income, health insurance and health care utilization is in the spirit of van de Ven and van der Gaag (1982). We treat health status, wealth and health care access as unobservable latent variables, which are determined simultaneously. Health care access, which is defined also by health insurance, is considered, among other things as a function of income. Health care utilization is specified as a function of health status, wealth and access to health care and other exogenous variables.

The article follows with some background information on the Brazilian context. In Section 3 we develop the general model, present the estimation procedure and discuss previous relevant researches. Section 4 describes the data and variables used to implement the model. The analysis of the empirical results follows in Section 5. The last section concludes with a discussion of the policy implications.

Background

Brazil is a developing country in the middle of a demographic transition process. The country's population almost quadrupled between 1940 and 2001, rising from 41 to 173 million people in 61 years. In the last 4 decades, the country has undergone considerable economic development and urbanization along with continuing increases in average education levels. In 1940 31% of the population resided in urban areas. This grew to 56% in 1970 and 81% in 2000. A sustained fertility decline, falling infant mortality rates and increasing life expectancy at birth were accompanied by declining but continued rapid population growth (IBGE, 1987; 1997). As a result of these changes, the country does not have the large under-18 population of the past nor the large proportion of older population groups observed in Europe or North America. The current Brazilian demographic structure represents a window of opportunity to implement social policies and reduce inequalities that should not be missed (see IDB, 1998).

Brazil presents great poverty level and enormous inequality in income distribution (see De Barros et al., 2001). Figure 1 shows that in 1999 about 14% of the population live in family in extreme poverty and 34% in family with income below poverty. It means that in Brazil there are about 53 million classified as poor. In the last two decades the percentage of poor decreased from about 40% to little more than 34%. However, the share of poor remained quite stable between 40 and 45% following the macroeconomic dynamic of the country. The maximum level of poverty was at the beginning of the 80s, when more than 50% of the population was below poverty line. However, the origin of poverty in Brazil is not the absolute scarcity of resource, but the enormous inequality in the income distribution. Brazil has one of the most unequal distributions of income in the world. The wealthiest 10% of Brazilians receive 47% of the national income, while the poorest 10 % receive only 0.8% (IDB, 1998).

The Health Care System

The Brazilian national health system was implemented under the principles of public financing, universal access, free care and full coverage. This set of principles also forms its legal essence, including the Brazilian Constitution of 1988. This system is known as SUS (acronym for Sistema Único de Saúde or Unified Health System). Brazil has also an important private sector that provides and finances services. The current health care expenditure in Brazil is estimated at 7.6% of GDP representing about US$ 280 per year (PAHO, 1999).

The Brazilian public-private mix can be better explained with the use of the well-known matrix presented below:

Public/Private Financing and Provision Matrix

		Provision	
		Public	*Private*
Financing	*Public*	A1	A2
	Private	B1	B2

The medical services provided under the SUS fall into cells A1 and A2. Inpatient care is largely provided by (for profit and non-profit) private facilities (cell A2), while ambulatory services are offered by public clinics (cell A1).[1] Hospital admissions are reimbursed by the SUS through a prospective payment system. It is estimated that in 1993 the SUS paid for between 77% and 83% of all hospital admissions in the country (PAHO, 1999).

The pure private market (cell B2) includes the care provided and/or financed by third-party agents like health insurance companies, HMOs, cooperative systems, etc. Because the quality of the services provided within the SUS is poor, companies and individuals that can afford the relatively high premiums charged by the private

[1] Differently from the majority of Latin America, the Brazilian public system provides 70% of ambulatory care, while almost 80% of hospital beds are in private institutions (Campino et al., 2000).

market opt to do so. As a result, almost 28% of the population has some type of private coverage (PNAD 1998), with the SUS clientele constituted mostly by the poor and all those that cannot afford, or whose job does not offer, a private third-party plan.

The SUS has also implemented an aggressive decentralization strategy, transferring the responsibility of ensuring most of the provision of services to the municipalities (see Bossert et al., 2000). Structural inconsistencies that exist within the financing structure of the SUS and the general crisis of the public sector have severely limited the government's ability to provide and finance health care services, further affecting the quality and quantity of the services offered by the system. This is contributing to an expansion of the private market and has resulted in some of the better providers to ending their contracts with the government. As a result, those individuals that depend on the SUS were left not only with fewer services, but also with a generally lower quality of care. The impact of such trend can be better grasped if it is understood that the SUS is responsible for the vast majority of medical services being performed in the country.

The inequalities in income distribution noted before are reflected in the access and utilization of health services as well as in the actual health conditions of individuals across income groups (Campino et al., 2000). Campino and colleagues analyzed a Living Standards Measurement Survey (Pesquisa sobre Padrões de Vida, PPV)[2] and found evidence of pro-wealthy inequalities both in terms of population health and in terms of the utilization of health services.

The study suggested that inequity in the access of health services – particularly preventive health services – has serious implications for the health status of the lower income population.

Data from PNAD-1998 presented in Table 1 confirm the perception that health improves as income increases. The proportion of individuals that describe their health as very good or good increase from 73% for the families with income less than one minimum wage, to 90% for families with income greater than 20 minimum wages. The inverse occurs with those that indicated a health status that is regular, poor or very poor. Also the proportion of people that reported having had a health problem in the last 2 weeks and had their habitual activities limited decreases, though slightly, with family income.

Table 2 shows that coverage by a health insurance plan increases greatly with income. Less than 3% of the poorest reported having a health plan. This percentage increases across income levels reaching 76% among the richest families. Table 2 also shows the proportion of individuals that consulted a doctor or were hospitalized in the last year. The patterns of health care utilization by income groups are not clear. In fact, while the proportion of people reporting consultation increases slightly with family income, the opposite happens in relation to the proportion that reported hospitalization.

In the remaining of the paper we try to disentangle some of the factors that are affecting health care status and health care utilization in Brazil.

[2] This survey took place in urban and rural areas of northeast and southeast Brazil between March 1996 and March 1997.

Overview of structural equation modeling with latent variables

When analyzing the determinant of health and health care utilization, there are two main reasons why structural equation models (SEM) with latent variables have recently been preferred to single equation models with a single health indicator as dependent variable.

SEM enables a researcher to test a set of regression equations simultaneously. Thus, the main advantage of SEM is to construct a model that combines the determinants of health with health care utilization. Using this approach, we can consider simultaneously both the direct and indirect effects of variables such as income, age or education. For example, in the analysis of health care utilization – one of the objectives of this study – it is important to separate the direct effect of some key variables (e.g. wealth) from their indirect effect (e.g. via its effect on health status, or health care access). Conversely, it is not possible to disentangle these factors in a single equation model in which the reduced form parameters include both the direct and indirect effects. Thus the choice of a SEM approach.

In addition, the central concept of health is inherently unobservable and the same can be argued for other concepts as wealth and access to health care. Treating these factors as latent variables allows the inclusion of a variety of indicators that are partial measure of the underlying latent unobservable variables. Manning and colleagues (1982) showed that using multidimensional health status measures rather than a single item measure (e.g. self-assessed health status) increases the precision of the estimation significantly.

SEM language differs from standard econometric jargon. Independent variables, which are assumed to be measured without error are called *exogenous* or *upstream* variables; dependent variables are called *endogenous* or *downstream* variables. *Manifest* or *observed* variables are directly measured by researchers, while *latent* or *unobserved* variables are not directly measured but are inferred by the relationship or correlation among measured variables in the analysis.

The main techniques used to estimate structural equation models with latent variables are the linear structural relationship (LISREL) model developed by Jöreskog and Sörbom (1987) and the partial least squares (PLS) algorithm by Wold (1982).[3] In the health economics literature there are numerous applications of LISREL methodology. On the other hand, even if PLS is widely used in disciplines such as education (see Sellin, 1995) and marketing (see Hulland, 1999), the only applications we are aware of in the health field are Bughin's (1990a; 1990b) studies of Belgium hospitals.

The majority of studies that used LISREL methodology adopted the multiple causes-multiple indicators (MIMIC) model, which is a special case of LISREL with only one latent variable – health status. Robinson and Ferrara (1977) were the first to use this approach. They presented a ranking of the 50 US states for 1969 based on a "health index" estimated by a MIMIC model. Wolfe and van der Gaag' s (1981) study is one of the first published paper that used individual level data. They used data on children from the 1975 Rochester Community Child Health Survey to construct a MIMIC model of health status and health care utilization.[4] Several other studies employed this methodology. For example, Wagstaff (1986; 1993) estimated empirical specifications of the Grossman's (1972) health production model using data from Denmark.

[3] For a comparison of LISREL and PLS methods see Dijkstra (1983) and Krelle (1997).

[4] In a subsequent paper (van der Gaag and Wolfe, 1991) the two authors extended their analysis to consider more health information and data on adult populations.

Häkkinen (1991) estimated a MIMIC model of health and health care utilization using survey data from Finland.

Among studies that allowed for more latent variables, we remember van de Ven and an der Gaag (1982), who estimated a SEM where both health status and income were latent unobservable variables using a health care survey carried out in The Netherlands. They found that the two variables had a mutual, positive impact and that age and education had important impacts on health status.[5] Ersland and colleagues (1995) estimated the impact of environmental pollution on health status and health care utilization treating both variables as latent. The model was estimated using data taken from the German socio-economic panel and the quality of the environment turned out to be an important determinant of health status.

Conversely, there are few health applications of SEM in developing countries. To our knowledge, the only examples are Behrman and Wolfe studies that analyzed health status and care utilization of women and children in Nicaragua using data collected in 1977-78 (Wolfe and Behrman, 1984; Behrman and Wolfe, 1987). The authors expressed doubts on the positive effect of maternal schooling on children health and nutrition. They also concluded that several important characteristics that are usually associated with development had no or had a negative associations with women's health: women's schooling, household resources and women's labor force participation.

[5] Another study that used data from The Netherlands is by van Vliet and van Praag (1987).

Empirical specification

SEM, in its most general form, consists of two parts: the structural equations and the measurement model. The structural equations specify the causal relationships among the latent variables and describe the causal effects and the amount of unexplained variance. The measurement model specifies how the latent variables or hypothetical constructs are measured in terms of the observed variables.

We can write our SEM of health care status and health care utilization as composed by the following four structural equations:

Health status: $$\eta_1 = \beta_1\eta_2 + \gamma_1 z_1 + \nu \quad (1)$$

Wealth: $$\eta_2 = \gamma_2 z_2 + \mu \quad (2)$$

Health care access: $$\eta_3 = \delta_1\eta_1 + \gamma_3 z_3 + \varepsilon \quad (3)$$

Health care utilization: $$y_i = \phi_1\eta_1 + \phi_2\eta_2 + \phi_3\eta_3 + \gamma_4 z_4 + e \quad (4)$$

where, η_1 is the unobservable latent variable health, η_2 is the unobservable latent variable wealth, η_3 is the unobservable latent variable health care access, y_i (i = 1,..,3) represents health care utilization measures, z_1, z_2 and z_3 are vectors of exogenous observable variables, β, γ, δ, and ϕ are parameters to be estimated. Finally ν, μ, ε and e are error terms with zero expectation.

The three latent variables health status, wealth and health care access (η_2, η_2, and η_3) are not observed, but instead vectors x_j (j =1,..,3) of observed variables are available. Thus, the measurement model is defined as:

Measurement model $$x_j = \varphi_j\eta_j + u_j \quad (j=1,..,3) \quad (5)$$

where φ_j are coefficient vectors and u_j represent error terms. Equation (5) represents the multivariate regression of x_j on η_j.

Since the latent variables η_j are unobserved they do not have a definite scale. Both the origin and the unit of measurement in each latent variable are arbitrary. To define the model properly the origin and the unit of measurement of each latent variable must be assigned. The origin is usually assigned by assuming that each variable has zero mean. A convenient way to assign the units of measurement of the latent variables is to fix one of the coefficient in the measurement model to a non-zero value (usually one). This defines the unit of each latent variable in relation to one of the observed variables. Strictly speaking, the unit of measurement of the latent variable will be equal to the unit of measurement of the observed variable minus its error term. An alternative way to define the unit of measurement in the latent variables is to assume that they are standardized (i.e. that they have unit variances).

In summary, the full SEM model is defined by the structural equations (1-4) and the measurement model (5).

The impact of health on wealth

If wealth affects health (as indicated in equation 1) health can also affect wealth. Poor health can reduce both labor supply and wage rates (Luft, 1975; Schultz and Tansel, 1997). A recent volume edited by Savedoff and Schultz (2000) collects numerous empirical studies carried out in Latin America. A study that analyzes data from Mexico suggests that a decline of one year in age at menarche – an indicator that is closely linked with adult health – is associated with an increase in wage of 23-26% (Knaul, 2000). A study in Colombia found that one more day of disability is associated with a decrease in earning by rural women of about 13% and that taller women earn more – about 7 percent more per centimeter (Ribero and Nuñez, 2000). In Peru one less day of reported illness in a month increased the wage rate of urban and rural women by 3.4 and 6.2 percent respectively (Cortez, 2000). A similar effect was estimated among Jamaican women (Neitzert and Handa, 2000). In general the effect of health on wages is found to be stronger among men than among women.

The only study, to our knowledge, that considered the effect of health status on wealth using a SEM framework is the paper by van de Ven and van der Gaag (1982). They found that better health status had a positive impact on wealth, but the relationship was not statistically significant.

In order to analyze this issue, we estimated an extended version our model. The impact of the latent variable health status on wealth is considered by modifying the structural equation that defines the wealth as follows:

Wealth $$\eta_2 = \theta_1 \eta_1 + \gamma_3 z_2 + \mu . \qquad (2b)$$

where θ is a parameter to be estimated. By substituting equation (2b) in place of equation (2) we allow for the relationship between the two latent variables health status and wealth to go either way.

Data and variable definitions

The data set used for the analyses is the 1998 Pesquisa Nacional por Amostra de Domicílios (PNAD-1998). PNAD is one of an annual series of national household surveys produced by the Brazilian Geographical and Statistical Institute (Instituto Brasileiro de Geografia e Estatística, IBGE). It is an annual cross-sectional survey providing information on labor force participation and earnings in conjunction with standard demographic characteristics such as age, schooling and place of residence for all members of the sampled households. The PNAD-1998 includes a special health module with various questions on self reported health status, chronic diseases, health limitation, health utilization and health insurance coverage. The complete sample includes 90,913 households, totaling 344,975 individual observations.

The PNAD data define the head of the household as the male spouse, if one is present. If not, a female may be described as the household head. For the purpose of this study, we constructed two different samples: adult women aged 18-60, adult men aged 18-60. We excluded from the analysis individuals living in the house without being part of the family because of uncertainty surrounding income sharing and pooling with other people living in the house. The total number of cases in the two sub-samples, after deletion of cases with missing data on the variables used, were 82,504 adult women and 77,789 adult men.

The SEM used to analyze the two samples follow the same framework described in equation (1-5). The description of all variables used in the analyses together with their means and standard deviations are presented in Table 3.

Health status

The PNAD-98 survey collected information on many indicators of health status. Following the example of other researchers we selected four dimensions of health. First, a measure of the self-rated health status measured into five levels (very good, good, regular, poor, very poor) (HEALTH). Second, the number of chronic diseases (CHRONIC). Third, the number of incapacitates and severe limitation in performing activities such as bathing, taking stairs, etc (LIMIT1). Forth, a dummy variable that indicates whether the responded was unable to perform any habitual activity because of health problem in the previous two weeks (LIMIT2).

Wealth

Individual wealth is considered as a latent variable and four manifest variables are used as indicators in the model related to adult population. The estimated individual WAGE, and the family per capita INCOME can clearly be considered as manifestation of the wealth of the person. We construct an indicator that is equal to the number individuals in the family that are not economically active divided by the number of individuals in the family that are economically active (DEPENDENT), and we expect that this variable is negatively related with the latent wealth. The last observable indicator is the proportion of family income that is earned by the individual in question (SHARE INC). We include this variable because we hypothesize that larger is the share of family income earned by the individual in question, more resources he/she could use for his/her own interest.

Because WAGE is only observed if the individual is working, we estimated them for all men and women using a wage equation that control for sample selection due to labor force employment. The wage equation is specified as $\ln W = \gamma Z + \sigma\lambda + u$, where Z is a vector of individual characteristics thought to affect one's offered wage and γ is the corresponding vector of coefficients, λ is the sample selection correction term σ is λ's coefficient, and u is the error term. The wage equation is estimated for the sub-sample of working men (women) with positive wages and includes and includes as explanatory variables (Z): age, age squared, education, education squared, interaction of education with age.[6] The resulting coefficient vector, γ, is then used to estimate wages as: $WAGE = \exp(\hat{\gamma} Z)$.

Health care access

The concept of health care access is multidimensional and covers both physical (e.g. inadequate healthcare facilities, in-sufficient transportation infrastructure, unequal distribution of healthcare facilities) as well as financial accessibility (e.g. shortage of financial resources, inadequate purchasing power). In relation to the latter, in Brazil health care services provided through the SUS are free of charge. However, use of private health insurance is widespread and about 30% of all individual interviewed had a supplementary health plan. Thus, we include as an indicator of access a variable (INSURANCE) measuring whether the individual in question is covered by health plan.

The second dimension of accessibility is the physical availability of health services. Even if relevant information is scarce in the survey, we account for this factor using a dummy variable measuring whether the individual lives in rural, urban or intermediate areas (CENSUS).[7]

Health care utilization

Using the variables available in the survey, we try to explain for different types of health care utilization. We use three dependent variables in the structural health care utilization equations. The number of visits to a doctor (CONSULT), the number of hospitalization in the last 12 months (HOSPITAL), and a dummy variable indicating if the individual in question had any expenditure during the last three months for health goods or services (EXPEND). The last variable is quite rough as it measures only utilization that determined out-of the pocket expenditures, without accounting for utilization provided by the SUS, which is usually free of charge for the patient.

Independent variables

The survey data include a large number of variables that are hypothesized to affect health status, wealth, health care access and health care utilization.

AGE reflects the health depreciation function emphasized in the Grossman model. This variable is included together with its squared term to explore non-linearity in their relationship with health status, wealth and health care access. EDUCATION is taken as a measure of the efficiency of individuals in combining market goods and time to produce health. EDUCATION is assumed to affect also the other two latent variables (WEALTH and ACCESS to health care) and health care utilization.

We also include three indicators of household conditions that are assumed to affect HEALTH negatively: a dummy variable indicating whether the family has a sewage system (SEWAGE), another indicating whether the house is connected to the water distribution network (WATER), and the presence of water filter (FILTER).

We also include a variable indicating whether the individual is currently UNEMPLOYED and we explore its effect on health status. Some researchers have argued that some forms of non-employment may be unhealthy. In particular, those who were laid off or fired, or who cannot find work, have a low sense of control, low levels of social support, and suffer distress as a consequence. Failure to get or keep a job may result in demoralization and neglect, and thus in poor health (see Pearlin et al., 1981).

We also include a dummy variable equal to one if the respondent started to work at the age of nine or below (CHILD LABOR).

[6] The wage equation is presented in Table A1 in the Appendix. The selection equation includes the number of children in the following age intervals: 0-2, 3-5, 6-10, and 11-14.

[7] We are currently exploring the possibility of complement the survey data with municipal level information on health services availability.

We expect that early entrance into the job market may have a detrimental effect on adult health. Finally we included a dummy variable indicating if the race of the respondent is NON-WHITE.

This variable is hypnotized to affect both health status and health care utilization.

The links among manifest (observed) and latent (unobserved) variables could be represented using path diagrams. Figure 2 shows the model estimated. Ovals represent latent variables, while squares represent measured variables.[8]

[8] In order to improve the fitness of the estimated model we allowed for correlation among the residuals of some of the variables included in the model.

Estimation results

SEM theory has been developed for normally distributed and continuous observed variables. However, these assumptions are never completely met in practice and in the model presented in this paper numerous variables are defined as dichotomous dummy variables. As long as the sample size is sufficiently large, the approach usually adopted is to use the ML estimator and bootstrap simulations to construct confidence intervals and determine the statistical significance of the estimates (Hancock and Nevitt, 1999).[9]

SEM models are usually evaluated in term of how well they fit the data. The goodness of fit of the models measured by the usual chi-square statistics was not satisfactory, showing always very high values. However, the chi-square test of absolute model fit is sensitive to sample size and non-normality in the underlying distributions. Thus, this statistics may not be the best measure of the fitness of the model. We also report the goodness of fit index (GFI) and the adjusted GFI statistics, which are statistics independent of sample size.[10] According to these measures the fit of the estimated models can be viewed as satisfactory since the values of these indices were always greater than 0.9.

Indicator relations

Table A2 in the Appendix presents the estimates of the measurement models. All the coefficients have the *a priori* expected signs and they are all significantly different from zero at the 5 percent probability level.[11]

In order to identify the latent variables, one of the indicators must be constraint to unity. The latent health status in adult women and men is associated most highly with CHRONIC and HEALTH (as measured by the R^2).[12] Turning to wealth, it is very highly associated with the expected WAGE (R^2 is higher than 0.9 in both cases), followed by INCOME (R^2 around 0.5). Finally, considering the indicators that identify health care access, INSURANCE shows the highest association with the latent variable. Thus, our indicators generally seem to be associated fairly strongly with our latent variables.

Structural Model

The maximum-likelihood estimation results for the structural equations determining health status are displayed in Table 4. The first column presents the estimated unstandardized regression coefficients, which represent the amount of change in the dependent variable per single unit change in the predictor variable. We also report the standardized estimates, which represent the amount of change in the dependent variable that is attributable to a single standard deviation unit's worth of change in the predictor variable. The squared multiple correlation (R^2) of the latent health status regression in the women sample is 0.25, while in the men sample is 0.30.
Most of the estimated parameters are statistically significant and have the expected sign. Health

[9] An alternative option is to use Browne's (1984) *Asymptotically Distribution Free* (ADF) estimator. However, this estimator can be adopted only in fairly simple models due to the computational requirements of the estimation procedure (Muthén, 1993) which failed to converge in our estimations.

[10] The GFI statistic was devised by Jöreskog and Sörbom (1984). It is always between zero and unity, where the latter indicates a perfect fit. The adjusted GFI takes into account the degrees of freedom available for testing the model. It is bounded above by 1, which indicates a perfect fit, but it is not bounded below by zero.

[11] The analysis was performed using the AMOS (Analysis of Moment Structure) software (Arbuckle, 1999).

[12] The coefficient of determination (R^2) represents the proportion of the total variance of the indicator explained by the latent variable.

status depreciates with age at increasing pace.[13] EDUCATION is directly related with better health status in both women and men, thought the effect is slightly stronger in the former group.

The variables measuring housing standards, WATER, SEWAGE and FILTER, are all significantly associated with better health status, with the latter indicator showing the largest positive impact on health especially among women. The two workforce variables (UNEMPLOYED and CHILD LABOR) are strongly associated with poorer health. The estimated coefficient of CHILD LABOR support the result of a recently published study that, using data from another Brazilian household survey, found that early entrance into the labor market is associated with worse self-reported health status (Kassouf et al., 2001).

Comparing the coefficients between gender we note that the coefficients of the variables UNEMPLOYED and CHILD LABOR estimated among women are respectively twice and three times larger than among men. Finally, the structural estimate of the impact of NON WHITE on health status is positive, statistically significant and more than twice larger among women.[14]

The latent variable health status is a measure with an arbitrary unit of measurement. To understand better the magnitude of the impact of the independent variable we calculate that between the age of 40 and 45 woman's health status index increases, on average, by 0.107 and man's health status index by 0.065. Thus, if a woman started working at the age of 9 or below, at the age of 40 her health status would be as bad as that of a 45 year old who had not started working at such an early age.

[13] Note that the health status indicators are defined in such manner that higher values indicate worse health status. Thus, a positive coefficient indicates that the independent variable produces a worsening of the latent health status.

[14] See Buvinic et al. (2002) for a survey of gender inequality in health related to the Latin American region.

Health care utilization

Table 5 shows the estimates of the health care utilization structural equations. Health status is positively associated with the three dependent variables: CONSULT, HOSPITAL and EXPEND. The coefficients are positive and highly significant in both women and men samples, indicating that individuals with worse health status utilize more health care resources, as we would expect intuitively.[15]

Better access to health care (i.e. having a private health insurance and living in urban areas) is positively associated with health care utilization. This result is in line with our expectations and with previous results (van de Ven and van der Gaag, 1982; Erbsland et al., 1995).

We estimate a negative effect of wealth on CONSULT and HOSPITAL, which suggests pro-poor inequality in the health care utilization. One plausible explanation of this finding is that the quality of care received may be different. It is likely that, *ceteris paribus*, wealthier people have access to better quality care, and thus they require less units of care. A second explanation is that wealthier people may attach higher value to their time, thus they may prefer using less medical care that is time consuming (i.e. consultation or hospitalization) (see Mullahy, 1999). Similar results were found in other studies. For example van de Ven and van der Gaag (1982) reported a negative effect of family income on OTC pharmaceutical expenditure and Erbsland et al. (1995) estimated that higher income earners had fewer general practitioners visits and spent fewer days in hospital. On the other hand, we find a positive effect of wealth on EXPEND, which indicates as expected that wealthier individuals are more likely to incur in out-of-the-pocket expenditure for health services or goods.

[15] Surprising, most of the studies that used SEM found the opposite (e.g. van de Ven and van der Gaag, 1982; Wolfe and Behrman, 1984; Wagstaff, 1986; Häkkinen, 1991; Erbsland et al., 1995). However, Wagstaff (1986) showed that this result (worse health related with less health care utilization) is compatible with the Grossman' model if health capital adjustments are not instantaneous.

EDUCATION is positively related with CONSULT and HOSPITAL, but negatively with EXPEND.[16] According to the Grossman's theoretical model the coefficient of EDUCATION in the structural medical care utilization equations is expected to be negative because better-educated individuals are hypothesized to be more efficient producers of health. As suggested by Wagstaff (1986), this result may reflects a more favorable relationship (e.g. cultural accessibility) between better-educated individuals and health service providers. On the other hand, when utilization is measured by EXPEND (i.e. expenditure on health services and goods) the effect of fact that better-educated individuals are more efficient producers of health as hypothesized by Grossman, might be dominant.

Finally, NON WHITE individuals appear to use health care less than white people in Brazil. In the model we are controlling for many important factors that may be related with race such as wealth and access and health status. Thus, these differences may reflect either differences in tastes or some degree of racial discrimination.

Latent access and latent wealth

Turning to the estimates of the structural equation health care access, we found that its relationship with AGE differs among sexes (see Table 6). In the men sample, we observe a "U" shaped relationship. Access, on average, gets worse with age up to the age of 46, but after improves (i.e. be more likely to be covered by private health insurance and live in more accessible urban areas). On the other hand women show an almost linear positive relationship between AGE and health care access.

The effect of wealth on health care access suggests, as we were expecting, that wealthier persons are more likely to buy health insurance and to live in urban areas. The relationship between health status and access does not appear to be significant. This result implies that unhealthy people are not more likely to seek health insurance or to live in urban areas, thus it suggests absence of adverse selection in the population studied. Finally, there are also gender differences in the impact of EDUCATION on access. More educated men present better access to health care, the opposite is found in the women's sample.

In the model we also allow for AGE and EDUCATION to influence wealth. In the men sample, AGE appears to be positively related with wealth until the age of 54, but after this age men, on average, are less wealthy. On the other hand, the relationship between AGE and wealth in the women sample appears monotonically increasing. Finally, EDUCATION appears to have an expected positive impact, without large gender differences.

Total effects of wealth, education and age

The structural equation estimate presented so far, show the direct effect of one variable on another. However, in the model we have specified one variable may affect another indirectly via a third variable. Table 7 presents direct, indirect and total effects of wealth, EDUCATION and AGE on health care utilization, latent health and access.

The effect of EDUCATION on health status is particularly interesting. Beside the positive direct effect we have already mentioned, more education leads to higher wealth, which in turn leads to better health status. Table 7 shows a strong total positive effect of EDUCATION on the health status of both women and men, but a smaller effect is estimated in the latter. The indirect effect of EDUCATION (via wealth) on health exceeds the direct effect in women sample, as found in the studies of van de Ven and van der Gaag (1982) and Grossman (1975). On the other hand, in the men sample the direct effect is the stronger than the indirect effect.[17]

[16] Positive effect of education on consultations and hospitalizations was also found in Wolfe and Behrman (1984) and Wagstaff (1986).

[17] For example, the indirect effect of EDUCATION on health status (via latent wealth) is the product of the estimated effect of EDUCATION on wealth times the estimated effect of wealth on health status (in the

Turning to the effect of AGE on health status, the strong negative direct effect is partially mitigated by the indirect positive effect via wealth. Figure 3 plots the total effects of age on health depreciation for the two sexes, which clearly shows that the depreciation rate is faster in the women sample.

AGE also affect health care utilization indirectly via health status, wealth and access. In the men samples, CONSULT are monotonically increasing with AGE, while women present a "U" shaped relationship. Similar patterns are observed in the case of HOSPITAL. On the other hand quite different is the relationship with EXPEND, while women present a monotonically increasing relationship between AGE and EXPEND, in the men sample we observe an inverse "U" relation.

A priori the total effect of wealth on health care utilization is undetermined. The direct effect of wealth on CONSULT and HOSPITAL is negative. On the other hand, there are two indirect effects: the first via health (a negative effect) the second via access (a positive effect). Our estimation results indicate the indirect effect of wealth on utilization is positive in all cases (i.e. the indirect effect via access dominates the indirect effect via health) but in the case of HOSPITAL utilization in the woman sample. However, the positive indirect effect of wealth on CONSULT and HOSPITAL is smaller than the negative direct effect, thus the resulting total effect remains negative. Conversely, since the indirect effect of wealth on EXPEND is positive, the resulting total effect both in women and in men is larger and highly significant.

Similarly complex is the total effect of EDUCATION on health care utilization. The indirect effect of EDUCATION on utilization works via wealth, health status and access. Altogether, the indirect effects of EDUCATION on CONSULT and HOSPITAL are negative, while the indirect effect on EXPEND is positive. The estimated total effects of EDUCATION on CONSULT and EXPEND are positive, while the total effect of EDUCATION on HOSPITAL is negative in both women and men.

Effects of health on wealth

We also estimate an extension of the model that, including equation (2b), allows for a direct effect of health status on wealth. In this model, the two latent variables HEALTH and WEALTH constitute a non-recursive subset. That is, in the path diagram of the model, it is possible to start at any one of the variables in the subset, and, by following a path of single-headed arrows, return to the original variable while never leaving the subset.

The estimated relationship is presented in Table 8. We can see that in both adult men and women healthier individuals are also wealthier. Similarly to the results found in the literature, the direct effect is stronger among men. On the other hand the estimates of the total effect of health on wealth are much smaller, though statistically significant.

However, in order to produce meaningful results, a non-recursive model must be stable. In our estimation, the statistics of the "stability index" appears to be greater than one, indicating that the system of equations is unstable. In other words we may obtain different results depending on the order followed in solving the system of equations. Thus, the results present problems of interpretation and further research in needed to improve the specification of this part of the model.

---

sample of women aged 18-60: -0.0003 × 559167. = -0.0147). The total effect is the sum of the direct and indirect effect (-0.0092 -0.0147 = -0.0239).

Conclusion

In this paper we have presented some preliminary results of a SEM of health care status and health care utilization in adult women and men in Brazil.

The estimates suggest that health of adult women and men is improved by water, sanitation and particularly by filtering the water used in the house. Education and wealth improve health status. On the other hand, unemployment, child labor and race discrimination affect negatively health status. The health status of women looks particularly worse compared with men. Women's health status depreciates at a quicker pace and appears to be much more sensitive to negative factors.

We acknowledge some caveats in the analysis. SEM assumes continuos and normally distributed observed variables and these assumptions are violated in our analysis. However, this situation is typical in SEM application in the health field and we have tried to overcome this limitation by constructing distributionaly-free standard errors by bootstrapping. Second, the indicators of health care access used in the analysis are only partial and we are trying to enhance the analysis using municipal level data on medical input access. We also recognize that part of the model is constructed in an *ad hoc* manner. Although, the formulation fits in the demand for health framework developed by Grossman's (Wagstaff, 1986; 1993).

What do these estimates suggest concerning the health impact of the development process of the region?
At aggregated level, Latin America and the Caribbean show important development in health improving factors. For the region as whole it is estimated that between 1980 and 1997 access to drinkable water grew from 60.3 to 73.5%; access to sewage services expanded from 43.3 to 69.7%; the percentage of literate population increased from 80.1 to 87.2% (PAHO, 1998). Regional gain in education were more marked for female. Gender differences in the average years of education historically favoring men disappeared with the cohorts born in 1970 and nowadays favor women (Duryea and Székely, 2000). Since these factors have a significant positive effect on health the trends are encouraging as they suggest that they may translate also in health improvement particularly among women.

However, what it is of concern is the distribution of opportunities in the region. In Brazil during the last decades we witnessed an increase in the inequality in the distribution of education and other health improving factors (see see De Barros et al., 2001 and IDB 1998). Thus, the underlying inequality factors presents serious challenges to the promotion of health and well-being in Brazil as well as in the Latin American region more in general.

It is also important to recognize the high probability that inequalities in the distribution of disadvantages would cumulate. For example, the poor are disproportionately more likely to be uneducated, non-white, to start working at early age, and not being covered by health insurance. To deal with individuals in particular disadvantaged position, action should be taken in the form of positive measures and targeted programs to improve women and men's education and training, as well as to provide support services, improve employment opportunities and access to health services. It is also important that interventions are introduced at an early age to avoid the process of cumulating disadvantages.

References

Arbuckle JL. 1999. *Amos User's Guide Version 4.0*. SmallWaters Corporation, Chicago.

Behrman JR, Wolfe BL. 1987. How does mother's schooling affect family health, nutrition, medical usage, and household sanitation? *Journal of Econometrics* 50: 183-203.

Bossert T, Larrañaga O, Ruiz Meir F. 2000. Decentralization of health systems in Latin America. *Pan American Journal of Public Health* 8(1/2): 84-92.

Browne MW. 1984. Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistics Psychology* 37:62-83.

Bughin J. 1990a. *Supplier-Induced Demand in the Hospital Sector: Partial Least Squares Estimates for Belgium*. Serie Recherche No. 106, Faculte Universitaire Notre-Dame de la Paix, Faculte des Sciences Economiques et Sociales, Namur Belgium.

______. 1990b. *Using Partial Least Squares Methodology for Disequilibrium Economics: Experiments with a Price Function*. Serie Recherche No. 107, Faculte Universitaire Notre-Dame de la Paix, Faculte des Sciences Economiques et Sociales, Namur Belgium.

Buvinic M, Giuffrida A, Glassman A. 2003. *Gender Inequality in Health and Work: the Case of Latin America and the Caribbean*. In Jody Heymann (Eds.) Global Inequalities at Work. Oxford: Oxford University Press.

Campino A, Díaz MDM, de Oliveira RG, Nuñes A. 2000. *Health Sector Inequalities and Poverty in Brazil. In Health Sector Inequalities and Poverty in Latin America and the Caribbean*. Results of PAHO-UNDP-World Bank EquiLAC-IHEP Project.

Cortez R. 2000. Health and Productivity in Peru: Estimates by Gender and Region. In Savedoff WD, Schultz TP. (eds.) *Wealth from Health. Linking Social Investments to Earnings in Latin America*. Inter-American Development Bank, Washington DC.

De Barros RP, Henriques R, Mendonça R. 2001. *A Estabilidade Inaceitável: Disigualdade e Pobreza no Brasil*. Texto Para Discussão No. 800. Instituto de Pesquisa Econômica Aplicada, Rio De Janeiro.

Dijkstra T. 1983. Some comments on maximum likelihood and partial least squares methods. *Journal of Econometrics* 22: 67-90.

Duryea S, Székely M. 2000. Labor markets in Latin America: a look at the supply-side. *Emerging Markets Review* 1(3): 199-228.

Dwyer D, Bruce J. 1988. *A Home Divided, Women and Income in the Third World*. Stanford University Press, Stanford CA.

Ersland M, Ried W, Ulrich V. 1995. Health, health care, and the environment. Econometric evidence from German micro data. *Health Economics* 15: 169-182.

Grossman M. 1972. On the concept of health capital and the demand for health. *Journal of Political Economy* 80: 223-255.

Grossman M. 1975. *The Correlation Between Health and Schooling*. In Terleckyj N. (eds.) *Household Production and Consumption*. Columbia University Press, New York.

Hancock GR, Nevitt J. 1999. Bootstrapping and the identification of exogenous latent variables within structural equation models. *Structural Equation Modeling* 6(4): 394-399.

Hulland J. 1999. Use of partial least squared (PLS) in strategic management research: a review of four recent studies. *Strategic Management Journal* 20: 195-204.

IBGE (Instituto Brasileiro de Geografia e Estatística). 1987. *Estatísticas Históricas do Brasil. Volume* 3. IBGE, Rio de Janeiro.

______. 1997. *Contagem da População 1996. Volume* 1. IBGE, Rio de Janeiro.

IDB (Inter-American Development Bank). 1998. *Facing up Inequality in Latin America. Economic and Social Progress in Latin America 1998-1999 Report*. Johns Hopkins University Press, Baltimore.

Jöreskog KG, Sörbom D. 1984. *LISREL-IV User's Guide*. Scientific Software, Mooresville, Indiana.

Kassouf AL, McKee M, Mossialos E. 2001. Early entrance to the job market and its effect on adult health: evidence from Brazil. *Health Policy Planning* 16: 21-28.

Knaul FM. 2000. Health, Nutrition and Wages: Age at Menarche and Earnings in Mexico. In Savedoff WD, Schultz TP. (eds.) *Wealth from Health. Linking Social Investments to Earnings in Latin America*. Inter-American Development Bank, Washington DC.

Krelle W. 1997. *How to Deal with Unobservable Variables in Economics*. Discussion Paper No. B-414. University of Boon.

Luft H. 1975. The impact of poor health on earnings. *Review of Economics and Statistics* 57: 43-57.

Manning WG, Newhouse JP, Ware JE Jr. 1982. The Status of Health in Demand Estimation: Beyond Excellent, Good, Fair and Poor. In Fuchs V. (eds) *Economic Aspects of Health*. University of Chicago Press, Chicago.

Mullahy J. 1999. It'll only hurt a second? Microeconomic determinants of who gets flu shots. *Health Economics* 8(1): 9-24.

Muthén B. 1993. Goodness of Fit with Categorical and other Nonormal Variables. In Bollen K, Long JS. (eds.) *Testing Structural Equation Models*. SAGE Publications, Newbury Park, CA.

Neitzert M, Handa S. 2000. *Gender and the Impact of Health on Labour Force Participation in Jamaica*. Mimeo.

PAHO. 1998. *Health Situation in the Americas: Basic Indicators 1995-1998*. Pan American Health Organization, Washington ,DC.

PAHO. 1999. *Brazil: Basic Country Health Profiles, Summaries 1999*. Pan American Health Organization, Washington ,DC.

Pearlin LI, Lieberman MA, Menaghan EG, Mullan JT. 1981. The stress process. *Journal of Health and Social Behavior* 22:337-356.

Ribero R, Nuñez J. 2000. Adult Morbidity, Height and Earnings in Colombia. In Savedoff WD, Schultz TP. (eds.) *Wealth from Health. Linking Social Investments to Earnings in Latin America*. Inter-American Development Bank, Washington DC.

Robinson PM, Ferrara MC. 1977. The Estimation of a Model for an Unobservable Variable with Endogenous Causes. In Aigner DJ, Goldberger AS. (eds.) *Latent Variables in Socio-Economic Model*. North-Holland, Amsterdam.

Savedoff WD, Schultz TP. 2000. Wealth from Health. *Linking Social Investments to Earnings in Latin America*. Inter-American Development Bank, Washington DC.

Schultz TP, 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.

Sellin N. 1995. Partial Least Square Modeling in Research on Educational Achievement. In Bos W, Lehmann RH. (eds.) *Reflections on Educational Achievement; Papers in Honour of T. Neville Postlethwaite.* Waxmann Munster, New York.

Wagstaff A. 1986. The demand for health. Some new empirical evidence. *Journal of Health Economics* 5: 195-233.

______. 1993. The demand for health: an empirical reformulation of the Grossman model. *Health Economics* 2: 189-198.

Wold H. 1982. Soft Modeling. The Basic Design and Some Extensions. In H. Wold and K. Jöreskog (Eds.), *Systems Under Indirect Observation: Causality, Structure, Prediction (Vol. I)*, Amsterdam: North-Holland, 263-270.

Wolfe BL, Behrman JR. 1984. Determinants of women's health-status and health-care utilization in a developing-country: a latent variable approach. *Review of Economics and Statistics* 66(4): 696-703.

Wolfe BL, van der Gaag J. 1981. A New Health Status Index for Children. In van der Gaag J, Perlman M. (eds.) *Health, Economics, and Health Economics*. North-Holland, Amsterdam.

Van de Ven WPMM. van der Gaag J. 1982. Health as an unobservable: a MIMIC model for health care demand. *Journal of Health Economics* 1: 157-183.
Van der Gaag J, Wolfe BL. 1991. *Estimating Demand for Medical Care: Health as a Critical Factor for Adults and Children*. In Duru G, Paelinck JHP. (eds.) *Econometrics of Health Care*. Kluwer, Amsterdam.
Van Vliet RCJA, van Praag BMS. 1987. Health status estimation on the basis of MIMIC-health care models. *Journal of Health Economics* 6: 27-42.

Appendix

Table 1 Health status indicators by family income

Family income [1] (number of minimum wage)	Self-rated health status (%) Very good or good	Regular	Poor or very poor	Population having health problem in the last 2 weeks that limited habitual activities (%) No	Yes	Population suffering from chronic diseases (%) No	Yes
Less than 1	72.54	21.10	6.30	91.78	8.20	66.83	33.08
From 1 to 2	74.59	20.04	5.31	92.68	7.29	67.75	32.11
From 2 to 3	75.21	20.06	4.67	92.81	7.14	67.02	32.86
From 3 to 5	78.26	18.27	3.43	93.92	6.06	68.36	31.50
From 5 to 10	81.66	15.89	2.39	94.52	5.46	68.25	31.63
From 10 to 20	84.98	13.04	1.93	94.96	5.00	68.71	31.12
More than 20	90.13	8.65	1.20	94.99	5.01	70.05	29.81
Without income	81.58	14.91	3.39	93.48	6.49	76.58	23.32
Not declared	80.34	16.05	3.42	94.04	5.79	65.81	33.92

1: Excluding income of non-family member living in the house

Table 2 Health insurance and health care utilization by family income

Family income [1] (number of minimum wage)	Population covered by private health insurance (%)	Population having medical consultations in the last 12 months (%)	Population hospitalized in the last 12 months (%)
Less than 1 minimum wage	2.56	49.67	8.71
From 1 to 2	4.83	50.01	8.02
From 2 to 3	9.36	51.67	7.17
From 3 to 5	18.58	53.44	6.45
From 5 to 10	34.72	56.51	5.90
From 10 to 20	54.03	60.78	5.85
More than 20	76.18	67.17	6.13
Without income (2)	6.39	55.74	11.52
Not declared	35.19	55.18	6.19

1: Excluding income of non-family member living in the house

Table 3 Description of variables used in the analyses

		Women 18-60 (observations 82,504)		Men 18-60 (observations 77,789)	
Variables	*Description*	*Mean*	*Std. Dev.*	*Mean*	*Std. Dev.*
HEALTH	Self-rated health status, 5 categories: very good (=1), good, regular, poor, very poor (=5)	2.083	0.777	1.951	0.733
CHRONIC	Number of chronic diseases suffered (12 in total): affecting the spine or vertebras, arthritis or rheumatism, cancer, diabetes, bronchitis or asthma, hypertension, heart problem, kidney problem, depression, tuberculosis, tendons, cirrhosis	0.782	1.147	0.513	0.890
LIMIT1	Number of incapacitates or great difficulties in (7 in total): eating, washing, toiletry; running, lifting heavy objects, doing sport; domestic work; take stairs; standing-up, bending; walk more than 1 Km; cover a distance of 100 meters.	0.273	0.871	0.139	0.621
LIMIT2	Dummy variable, 1 if reported a health problem in the last 2 weeks that limited habitual activities, 0 otherwise	0.070	0.255	0.047	0.211
CONSULT	Number of medical consultation in the last 12 months	2.958	4.648	1.384	3.202
HOSPITAL	Number of time hospitalized in last 12 months	0.130	0.473	0.056	0.340
EXPEND	Dummy variable, 1 if incurred in any cost in last 3 months for health services or goods, 0 otherwise	0.175	0.380	0.258	0.438
INCOME	Monthly income per family member	307.00	511.90	322.42	616.83
DEPENDENT	Number of non economically active members for each economically active member of the family	1.336	1.284	1.244	1.247
SHARE INC	Share of family income earned by the individual	23.914	30.096	56.412	35.766
WAGE	Predicted monthly wage in Real	425.13	357.65	196.35	188.35
INSURANCE	Dummy variable, 1 if covered by a private health insurance, 0 otherwise	0.306	0.461	0.264	0.441
CENSUS	Dummy variable, 1 if area of residency was rural in the last census, 0 otherwise	0.117	0.321	0.134	0.341
AGE	Age in years	34.951	11.392	34.749	11.468
EDUCATION	Number of years in education	7.997	4.259	7.698	4.193
WATER	Dummy variable, 1 if the water used in the house comes from the public distribution network, 0 otherwise	0.822	0.383	0.802	0.398
SEWAGE	Dummy variable, 1 if the house is connected to the urban sewage system or has a septic tank, 0 otherwise	0.690	0.463	0.674	0.469
FILTER	Dummy variable, 1 if in the house there is a filter for the water, 0 otherwise	0.598	0.490	0.580	0.494
UNEMPLOYED	Dummy variable, 1 if unemployed, 0 otherwise	0.073	0.260	0.061	0.240
CHILD LABOR	Dummy variable, 1 if started working at the age of 9 or below, 0 otherwise	0.078	0.268	0.155	0.362
NON WHITE	Dummy variable, 1 if the race is non-white, 0 if the race is white	0.444	0.497	0.462	0.499

Table 4 Structural Equation. Dependent variable: Latent Health Status

	Women aged 18-60			Men aged 18-60		
Variables	*Estimate Non-stand.*	*Standard Error*	*Estimate Stand.*	*Estimate Non-stand.*	*Standard Error*	*Estimate Stand.*
AGE	0.0066 **	0.0010	0.1487 **	-0.0011 ***	0.0005	-0.0361 ***
AGE^2	0.0002 ***	9.4E-06	0.3163 ***	1.7E-04 ***	4.6E-06	0.4286 ***
EDUCATION	-0.0092 ***	0.0018	-0.0773 ***	-0.0081 ***	0.0008	-0.0999 **
WEALTH (latent)	-0.0003 **	3.2E-05	-0.2539 **	-8.8E-05 ***	1.3E-05	-0.1426 **
WATER	-0.0151 **	0.0049	-0.0113 ***	-0.008 ***	0.0032	-0.0093 ***
SEWAGE	-0.0156 ***	0.0040	-0.0142 ***	-0.0148 **	0.0027	-0.0204 **
FILTER	-0.0359 ***	0.0038	-0.0347 ***	-0.0195 **	0.0026	-0.0282 **
UNEMPLOYED	0.0490 **	0.0071	0.0250 **	0.0229 ***	0.0053	0.0161 ***
CHILD LABOR	0.1168 ***	0.0070	0.0616 **	0.036 ***	0.0035	0.0383 ***
NON WHITE	0.0504 **	0.0038	0.0493 **	0.0207 **	0.0026	0.0303 **
R^2	0.25			0.30		
System estimate statistics						
χ^2	4,335 ***			4,085 ***		
GFI	0.946			0.972		
Adjusted GFI	0.985			0.996		

*** indicates $p \leq 0.01$; ** indicates $0.01 < p \leq 0.05$; * indicates $0.05 < p \leq 0.1$

Significance levels estimated by bootstrapping with 2,000 iterations

Table 5 Structural Equations. Dependent variables: Health Care Utilization

Women aged 18-60

	Dependent variables								
	CONSULT			HOSPITAL			EXPEND		
Variables	*Estimate Non-stand.*	*Standard Error*	*Estimate Stand.*	*Estimate Non-stand.*	*Standard Error*	*Estimate Stand.*	*Estimate Non-stand.*	*Standard Error*	*Estimate Stand.*
HEALTH (latent)	4.1759 **	0.0502	0.4035 **	0.2294 **	0.0046	0.2123 ***	0.0478 ***	0.0036	0.0561 ***
ACCESS (latent)	8.3545 ***	0.1888	0.5679 ***	0.5871 ***	0.0165	0.3823 ***	0.7678 ***	0.0177	0.6336 ***
WEALTH (latent)	-0.0084 ***	1.1E-04	-0.7866 ***	-0.0009 ***	1.0E-05	-0.7806 ***	1.4E-04 ***	9.8E-06	0.1583 **
EDUCATION	0.2182 ***	0.0062	0.1766 ***	0.0228 **	5.8E-04	0.1766 **	-0.0147 ***	5.4E-04	-0.1447 **
NON WHITE	-0.1587 **	0.0306	-0.0150 **	-0.0015 ***	0.0033	-0.0014 ***	-0.0051 ***	0.0024	-0.0058 ***
R^2	0.54			0.18			0.30		

Men aged 18-60

	CONSULT			HOSPITAL			EXPEND		
Variables	*Estimate Non-stand.*	*Standard Error*	*Estimate Stand.*	*Estimate Non-stand.*	*Standard Error*	*Estimate Stand.*	*Estimate Non-stand.*	*Standard Error*	*Estimate Stand.*
HEALTH (latent)	4.6471 **	0.0528	0.4769 **	0.3534 **	0.0052	0.3401 **	0.0808 ***	0.006	0.0539 **
ACCESS (latent)	2.0459 **	0.0592	0.2354 ***	0.0771 **	0.006	0.0832 ***	1.0124 **	0.0218	0.7569 ***
WEALTH (latent)	-0.0012 **	1.0E-05	-0.2029 **	-0.0001 **	1.2E-06	-0.1953 **	0.0002 ***	2.0E-04	0.2035 ***
EDUCATION	0.0905 **	0.0033	0.1145 **	0.0073 **	0.0004	0.0867 **	-0.024 ***	0.0007	-0.1977 ***
NON WHITE	-0.0826 **	0.0218	-0.0124 **	-0.0057 **	0.0024	-0.0081 **	-0.0315 ***	0.0026	-0.0308 ***
R^2	0.28			0.17			0.65		

*** indicates $p \le 0.01$; ** indicates $0.01 < p \le 0.05$; * indicates $0.05 < p \le 0.1$
Significance levels estimated by bootstrapping with 2,000 iterations

Table 6 Structural Equations. Dependent variables: Latent Access and Wealth

ACCESS (latent)

	Women aged 18-60			*Men aged 18-60*		
Variables	*Estimate Non-stand.*	*Standard Error*	*Estimate Stand.*	*Estimate Non-stand.*	*Standard Error*	*Estimate Stand.*
AGE	-0.0141 **	6.2E-04	-0.4483 **	0.0042 ***	0.0005	0.1256 ***
AGE^2	1.5E-04 **	6.1E-06	0.3629 **	1.5E-04 ***	6.1E-06	-0.0669 ***
WEALTH (latent)	0.0010 **	2.1E-05	1.3558 **	0.0003 **	2.1E-05	0.4095 **
HEALTH (latent)	0.0039	0.0037	0.0055	-0.0062	0.006	-0.0055
EDUCATION	-0.0060 **	0.0012	-0.0714 **	0.0321 ***	0.0008	0.3536 ***
R^2	0.75			0.54		

WEALTH (latent)

	Women aged 18-60			*Men aged 18-60*		
Variables	*Estimate Non-stand.*	*Standard Error*	*Estimate Stand.*	*Estimate Non-stand.*	*Standard Error*	*Estimate Stand.*
AGE	-0.2888 ***	0.0011	-0.5018 **	34.731 **	0.1327	0.7236 **
AGE^2	29.6739 **	0.1015	0.6896 **	-0.3196 **	0.0013	-0.4958 **
EDUCATION	55.9167 ***	0.2094	0.4858 **	56.282 ***	0.2368	0.4288 **
R^2	0.96			0.95		

*** indicates $p \leq 0.01$; ** indicates $0.01 < p \leq 0.05$; * indicates $0.05 < p \leq 0.1$. Significance levels estimated by bootstrapping with 2,000 iterations

Table 7 Total, Direct and Indirect Effects of Selected Variables

		Women aged 18-60			
	Effect	EDUCATION	AGE	AGE^2	WEALTH (latent)
CONSULT	Direct	0.2182 ***	-	-	-0.0084 ***
	Indirect	-0.1608 ***	-0.1279 ***	0.0024 ***	0.0072 ***
	Total	0.0573 ***	-0.1279 ***	0.0024 ***	-0.0013 **
HOSPITAL	Direct	0.0228 **	-	-	-0.0009 ***
	Indirect	-0.0255 **	-0.0173 **	0.0002 **	0.0005 ***
	Total	-0.0027 ***	-0.0173 **	0.0002 **	-0.0004 **
EXPEND	Direct	-0.0147 ***	-	-	0.0001 ***
	Indirect	0.0445 **	0.0158 ***	-0.0001 ***	0.0007 ***
	Total	0.0297 **	0.0158 ***	-0.0001 ***	0.0009 **
HEALTH (latent)	Direct	-0.0092 ***	0.0066 **	0.0002 ***	-0.0003 **
	Indirect	-0.0147 **	-0.0078 **	0.0001 **	-
	Total	-0.0239 ***	-0.0012 ***	0.0003 **	-0.0003 **
ACCESS (latent)	Direct	-0.0060 **	-0.0141 **	0.0002 **	0.0010 **
	Indirect	0.0552 **	0.0293 **	-0.0003 **	-1.0E-06 **
	Total	0.0492 ***	0.0153 ***	-0.0001 **	0.0010 **

		Men aged 18-60			
		EDUCATION	AGE	AGE^2	WEALTH (latent)
CONSULT	Direct	0.0905 **	-	-	-0.0012 **
	Indirect	-0.0309 **	-0.0329 **	0.0011 **	0.0002 **
	Total	0.0596 **	-0.0329 **	0.0011 **	-0.0010 **
HOSPITAL	Direct	0.0073 **	-	-	-1.3E-04 **
	Indirect	-0.008 **	-0.0047 **	0.0001 **	-9.2E-06 ***
	Total	-0.0007 ***	-0.0047 **	0.0001 **	-1.3E-04 ***
EXPEND	Direct	-0.024 ***	-	-	0.0002 ***
	Indirect	0.0583 **	0.0204 **	-2E-04 **	0.0003 **
	Total	0.0343 **	0.0204 **	-2E-04 **	0.0005 **
HEALTH (latent)	Direct	-0.0081 **	-0.0011 ***	0.0002 ***	-8.8E-05 ***
	Indirect	-0.005 ***	-0.0031 ***	0 ***	-
	Total	-0.0131 ***	-0.0041 ***	0.0002 **	-8.8E-05 ***
ACCESS (latent)	Direct	0.0321 ***	0.0042 ***	-3E-05 ***	0.0003 **
	Indirect	0.016 **	0.0099 **	-9.2E-05 **	5.4E-07 **
	Total	0.0482 ***	0.0141 ***	-1.2E-04 **	0.0003 **

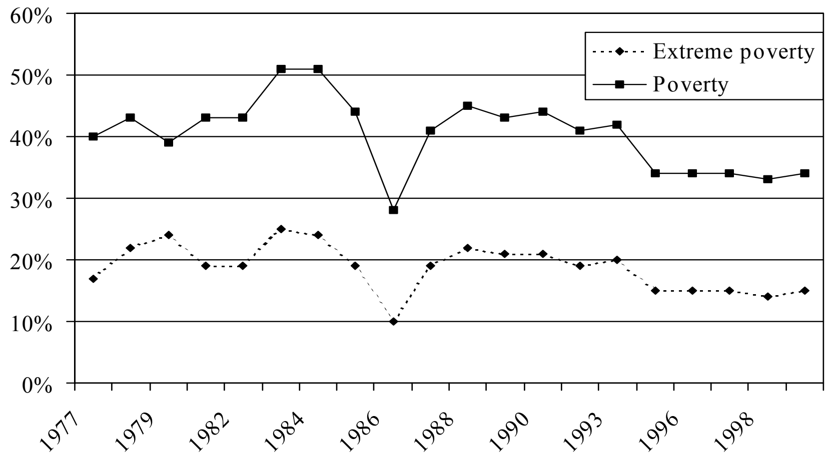
*** indicates $p \leq 0.01$; ** indicates $0.01 < p \leq 0.05$; * indicates $0.05 < p \leq 0.1$. Significance levels estimated by bootstrapping with 2,000 iterations

Table 8 Total, Direct and Indirect Effects of Health on Wealth

	Women aged 18-60		Men aged 18-60	
	Non-standardized	Standardized	Non-standardized	Standardized
Direct Effects	-2,499 **	-2.7622 **	-11,948 **	-7.2697 ***
Indirect Effects	2,491 **	2.7527 **	11,948 **	7.2696 ***
Total Effects	-8.5515 ***	-0.0095 ***	-0.1919 ***	-0.0001 ***
Stability index	291		623	

*** indicates $p \leq 0.01$; ** indicates $0.01 < p \leq 0.05$; * indicates $0.05 < p \leq 0.1$. Significance levels estimated by bootstrapping with 2,000 iterations

Figure 1 Poverty in Brazil 1977-1999



Source: PNADs 1977-1999

Figure 2 Path Diagram of the Model

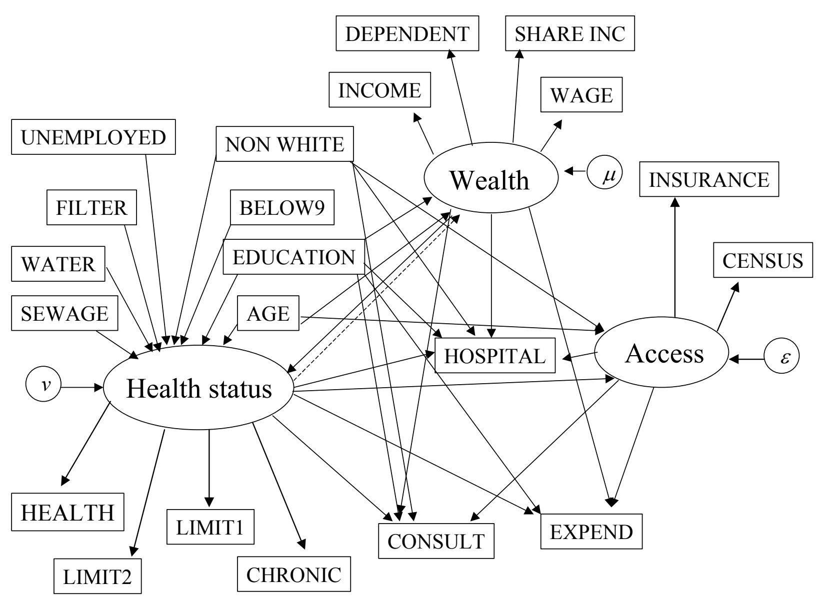


Figure 3 Total effect of age on health status

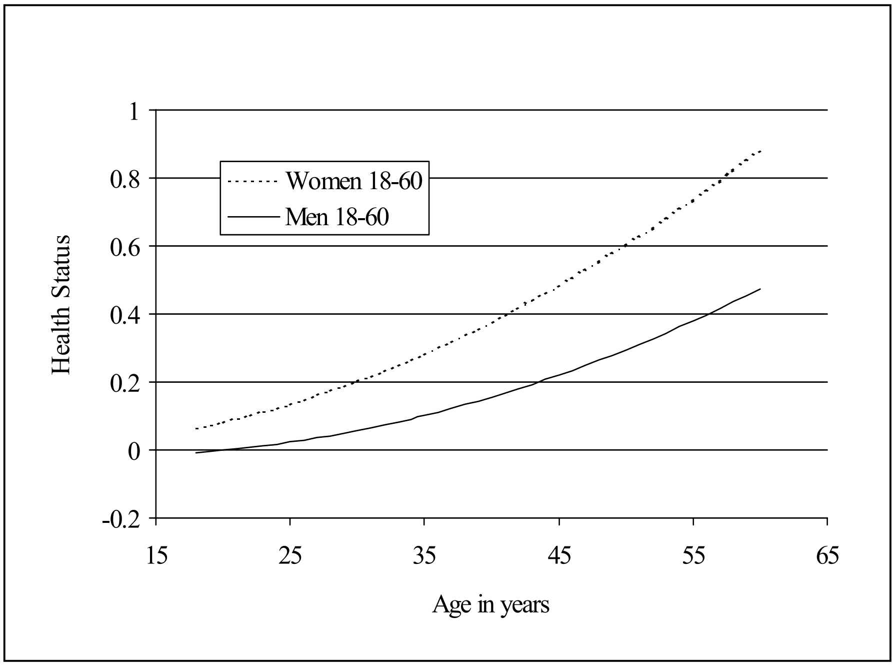


Table A1 The Wage Equations. Dependent Variable: Natural Log of Weekly Wage

Variables	Women aged 18-60		Men aged 18-60	
	Coefficient	Standard Error	Coefficient	Standard Error
AGE	0.0942 ***	0.0029	0.1198 ***	0.0022
AGE^2	-1.1E-03 ***	3.6E-05	-0.0014 ***	2.7E-05
EDUCATION	0.0074	0.0051	0.0296 ***	0.0034
EDUCATION^2	0.0071 ***	2.0E-04	0.0027 ***	1.5E-04
AGE × EDUCATION	0.0009 ***	8.6E-05	0.0018 ***	6.3E-05
CONSTANT	2.3675 ***	0.0688	2.6327 ***	0.0448
Test of independence $\rho = 0$	478.73 ***		289.63 ***	
Number of observation	94214		88191	
Censored observation	51271		17366	
Uncensored observation	42943		70825	

*** indicates $p \leq 0.01$; ** indicates $0.01 < p \leq 0.05$; * indicates $0.05 < p \leq 0.1$

Dummy variables for each of the 27 States were included in the estimation

Table A2 Measurement Equations. Indicator Relations

	Women aged 18-60			*Men aged 18-60*	
			HEALTH (latent)		
Variables	*Estimate Non-stand.*	*Standard Error*	R^2	*Estimate Non-stand.*	*Standard Error*
CHRONIC	1.6340 **	0.0120	0.53	1.8195 ***	0.0154
HEALTH	0.9925 ***	0.0076	0.43	1.4529 ***	0.0124
LIMIT1 ^{a}	1	-	0.34	1	-
LIMIT2	0.1771 ***	0.0021	0.13	0.2267 **	0.0029
			WEALTH (latent)		
Variables	*Estimate Non-stand.*	*Standard Error*	R^2	*Estimate Non-stand.*	*Standard Error*
DEPENDENT	-0.0008 ***	9.4E-06	0.08	-0.0001 **	1.3E-05
INCOME ^{a}	1	-	0.55	1	-
SHARE INC	0.0311 **	2.3E-04	0.22	0.025 ***	0.0002
WAGE	0.6665 **	0.0022	0.95	1.2485 ***	0.0046
WAGE FATHER					
WAGE MOTHER					
			ACCESS (latent)		
Variables	*Estimate Non-stand.*	*Standard Error*	R^2	*Estimate Non-stand.*	*Standard Error*
INSURANCE ^{a}	1	-	0.50	1	-
CENSUS	-0.2619 ***	0.0035	0.08	-0.2332 **	0.0036

*** indicates $p \leq 0.01$; ** indicates $0.01 < p \leq 0.05$; * indicates $0.05 < p \leq 0.1$. Significance levels estimated by bootstrapping with 2,000 iterations

a: Normalized to be one to identify the model