The Dynamics of Health and Its Determinants among Older Adults*

Younoh Kim†

Job Market Paper

October 14, 2012

Abstract

This paper examines the determinants of chronic health conditions and explains their persistence, using a panel data set from the Indonesian Family Life Survey (IFLS). I incorporate dynamics into a health demand function, finding strong correlations between lagged and current health measures when nothing else is controlled. This could represent the influence of lagged health or fixed unobserved factors such as genetic endowments and childhood health. To disentangle these, I estimate the influence of lagged health by using first-difference two-step generalized method of moments (FD-GMM), where the first-differencing removes fixed unobserved factors and keeps only lagged health. I found that it is this fixed effect, representing both genetic endowments and childhood health, that is most important in explaining later life chronic conditions. The impact of past health conditioning on the fixed effect, captured by the coefficients on lagged health measures, is weak, with estimated coefficients relatively close to zero. These results are robust to potential measurement errors in health and to sample attrition. Socio-economic status also has very little influence on current health, again conditioned on the fixed effect and on the influence of lagged health. In order to investigate if past health has different impacts across demographic or economic groups, I disaggregate the sample across age, household per capita expenditure level (PCE), and years of education. The results show that those with less education tend to show more persistency, compared to those with higher education.

Keywords: Elderly Health, Chronic Conditions, Path Dependence, Dynamic Panel Data Model

JEL Classification Numbers: I10, O12

*__Acknowledgments__: I am deeply indebted to my advisor, John Strauss for his excellent guidance and support. I would also like to thank Hyungsik Roger Moon, Jeffrey Nugent, Anant Nyshadham, Firman Witoelar, Vlad Radoias, Fei Wang, and Jay Gwon for valuable comments. All remaining errors are mine.

†**Contact details**: Department of Economics, University of Southern California, younohki@usc.edu

1 Introduction

There has been a large and growing literature on early life origins of health as an adult (eg. Barker et al., 1989; Godfrey and Barker, 2000; Crimmins and Finch, 2004, 2006; Case et al., 2005, and Smith, 2009). These papers find strong correlations between early childhood health and health as older adults. For example, Smith (2009) shows that a measure of childhood health is a good predictor for the onset of diseases during adulthood. However, very little is known about the dynamics of health at older ages.

In Indonesia, for example, hypertension is one of the most common chronic diseases.[1] Using the Indonesian Family Life Survey (IFLS), a panel survey spanning 14 years, Table 1.1 reports the prevalence of hypertension over time among the elderly of age 50 and above. Among those with hypertension in the past, 80% are still hypertensive when measured in the next wave of the survey, between three and seven years later. Elderly women show an even stronger persistence in hypertension than elderly men. The correlations between past and current hypertension are also high, 0.47 and 0.52 for men and women respectively, again implying that it is hard to recover from hypertension, once developed.

The question is why do we observe strong persistence in chronic diseases like hypertension? Is it childhood health that has permanent consequences which is consistent with the findings of early life origins or are there some other factors explaining this persistence? Among possible reasons that might cause persistent chronic health problems are genetic endowments, influence of childhood health and nutrition, health inputs and behaviors as an adult, and the past history of chronic conditions. Health inputs and behaviors will be mediated by socio-economic status. This study makes a start in disentangling the separate influences using longitudinal data from Indonesia.

The main contribution of this paper is to characterize the evolution of health over time, especially during later life. It is reasonable to describe the current health of older adults in a dynamic framework since it is a consequence of past health inputs or behaviors. However, incorporating dy-

[1] Witoelar et al. (2012) find that around 44 percent of elderly men and 53 percent of elderly women have hypertension in Indonesia in 2007.

namics into an empirical model causes potential endogeneity problems because past health is likely to be correlated with time-invariant unobserved characteristics. I address this problem by deriving a health demand function conditioned on lagged health, and estimating it using a first-difference two-step generalized method of moments (FD-GMM), proposed by Arellano-Bond (1991), where the first-differencing removes fixed unobserved factors and keeps only lagged health.

Using a panel data set from Indonesia, the Indonesian Family Life Survey (IFLS), I find that it is genetic endowments and childhood health, captured in the individual fixed effects, that play the most important roles in determining chronic health conditions in later life. The impact of past health conditions, captured in the coefficients on lagged health measures, is weak, conditional on the fixed effects. This result is robust to measurement errors in lagged health and to sample attrition including from selective mortality. Socio-economic status also has very little influence on current health, conditional on both fixed effects and lagged health. These results are consistent with the large and growing literature on early life origins of later life health.

In order to investigate if past health has different impacts across demographic or economic groups, I disaggregate the sample across age, per capita household expenditure level (PCE), and years of education. The results show that those with less education tend to show more persistence of chronic conditions, compared to those with higher education.

In terms of policy, a strong impact of early childhood health implies that investments should be targeted towards the early stages of life. Stronger persistence of chronic conditions of lower educated elderly also suggests that investments should be made not only towards health in early life, but also, towards better education in early life. Among the elderly, the focus should be on those with lower education attainments.

The rest of the paper is organized as follows. Section 2 provides a brief review of the related literature. A theoretical model is presented in section 3. Section 4 describes the data and the variables used in this paper, and section 5 presents the empirical model. The main regression results are discussed in section 6. Concluding remarks follow in section 7.

2 Studies of Health Dynamics

There exists some literature that estimates the degree to which children can recover from health problems caused by nutritional deficiencies at earlier ages (eg. Habicht et. al, 1995; Adair, 1999; Mortorell, 1995, 1999; Hoddinott and Kinsey, 2001; Fedorov and Sahn, 2005; Alderman et.al, 2006; Mani, 2012). In these studies, height is used as an indicator of cumulative health during early childhood. For example, INCAP (Institute of Nutrition of Central America and Panama) conducted randomized experiments in rural Guatemala during the late 1960s and 1970s by randomly giving protein supplementation. They find that for children exposed to supplementation before reaching age of 3 tend to be taller and less stunted. (eg. Habicht et. al, 1995)

Other studies have used panel data sets and examined the magnitudes of coefficients on lagged height for children in dynamic panel models, the "catch-up effect". Hoddinott and Kinsey (2001), for instance, estimate the catch-up effect at 0.56. They adopt both a two-stage least squares (2SLS) method and maternal fixed effects using a sample of children living in rural areas of Zimbabwe. Fedorov and Sahn (2005), on the other hand, estimate a coefficient on lagged height at 0.22 for children in Russia by using an Arellano-Bover (1995) estimation. Alderman et.al (2006) employ a maternal fixed effects instrumental variable estimation (MFE-IV) which yields the catch-up coefficient of 0.43 for children in Zimbabwe. Mani (2012) also finds a partial catch-up effect, 0.23 for malnourished children in Indonesia by using a first-difference GMM method (FD-GMM). While evidence of catch-up effects for children vary some depending on countries and econometric methods, it is worth noting that these results still suggest that children tend to recover from some, but not all, of the height deficits incurred in the past. Examining child health and its consequences is important in the sense that undernourishment at early age may cause permanent growth retardation and in turn, affect future well-being by lowering the human capital such as schooling or earnings. For instance, Hoddinott et al. (2008) find, using the INCAP data, that children who were exposed to protein supplements between age 0 and 2 years have a 46% increase in average wages. Therefore, studies of health and nutrition status during early childhood and whether children could recover from past deficits have received much attention, especially in developing countries where resources

are limited.

However, studies of health dynamics focusing on the elderly are very rare. Adams et al. (2003), for example, use (ordered) probit models and look at the impact of past health and socio-economic status on the future onsets of several diseases among the American elderly. They control for several past health conditions as well as current incidences as covariates and list them by following Wold causal chain such as cancer, heart disease, stroke, degenerative and chronic conditions, and mental condition. The Wold causal chain assumes that the first component causes the second, the first and the second cause the third and so forth. However, this might not be perfectly true since mental conditions, such as depression can increase the risk of having cancer or stroke[2]. Therefore, instead of having several health measures, I analyze one health measure at a time without allowing any contemporaneous influences from other co-morbidities. Adams et al. find a strong impact of past health on current health such as those who had an incontinence problem in the past tend to have the same problem in the next period and the past history of having arthritis, lung problems and the numbers of difficulties of ADLs and IADLs also predict a future incontinence problem.

Smith (2007), examining future onset of disease does not allow for contemporaneous correlation between disease and chronic conditions but controls for several past health measures such as self-assessed general health status and chronic conditions as right-hand side regressors. Using the Panel Study of Income Dynamics (PSID), he finds that the future onsets of disease are strongly correlated with past self-assessed general health status and chronic health conditions. However, neither Adams et al.(2003) nor Smith (2007) addresses the problem arising from the possible correlation between previous health measures and unobserved factors like individual-specific fixed effects. It is most likely that previous health status, measured by lagged health condition, is correlated with individual time-invariant fixed effects and this will create a bias in OLS coefficient estimates. Genetic endowments and childhood health are many times unobserved in data and and so are captured in the individual fixed effect. A person with genetically stronger endowments or who had a healthy childhood is more likely to be healthy at all times, and being healthy leads to higher education and

[2]See Penninx, B. et al.(1998) and Jonas et al.(2000) for details

earnings, which further leads to better health outcomes in the subsequent period. In this case, the impact of previous health can be over-estimated.

Using the National Longitudinal Surveys of Labor Market Experience (NLS), Cooksey et al. (2012) show a strong correlation between past and current depression based on transition matrix using CES-D data. The coefficient on lagged depression is 0.86. However, they found that when the fixed effect is taken into account, this fixed effect explains almost 27% of the persistence, 30% comes from lagged depression, and error terms explain the rest. Moreover, the correlations between past and current depression do not substantially attenuate over time, again suggesting the presence of time-invariant components.

Many studies exist that examine the impact of early childhood health on own later health (eg. Barker et al., 1989; Godfrey and Barker, 2000; Crimmins and Finch, 2004, 2006; Case et al., 2005; Smith, 2009). A leading theory is the early life origins hypothesis. These examples include Barker's argument that organ sizes or function and gene expression may adapt to a new environment in order to raise survival probabilities when exposed to negative shocks during very early childhood or the fetal period (eg. Barker et al., 1989). Crimmins and Finch (2004) argue that inflammation caused by infectious diseases in early life increases the risk of later life morbidity and mortality. Examples of studies that examine the early life origins hypothesis also include Case et al. (2005), who use a 1958 British birth cohort, conditioned on parental income, education and social status and show that respondent's birthweight is strongly associated with poor general health in adulthood, with the correlation becoming larger as people age. As another example, Smith (2009) finds that a measure of childhood health is a good predictor for the onset of disease during adulthood such as cancer, heart problem and hypertension. These studies, however, can be misinterpreted since they do not account for unobserved factors such as individual specific unobserved characteristics or time-varying shocks. The impact of early health might be over-estimated for those with stronger genetic endowments or under-estimated for those exposed to temporary health shocks at the time of the survey of later health status.

3 Model

Health is considered as one of the important components of human capital. Individuals make decisions on health investment or health behaviors such as nutrient intakes, exercise, and health care utilization in order to have better health and achieve both physical and subjective well-being. Similar to agriculture or manufacturing production functions, the health production function represents the technology available to individual or households to convert health inputs to health outcomes. The current health status can be expressed as the outcomes of the health production function. Given that health evolves over the life course, it is reasonable to describe the current health in a dynamic framework, as a function of all current and prior health inputs, individual demographic characteristics, household background characteristics, the disease or public health environment, time-varying health shocks, and time-invariant genetic endowments.

$$H_t = h(N_t, N_{t-1}, .., N_0, I_t, I_{t-1}, .., I_0, D_t, D_{t-1}, .., D_0, \gamma_{et}, \gamma_{et-1}, ..\gamma_{e0}, \gamma_e, \mu_{ht}, \mu_{ht-1}, .., \mu_{h0}, \mu_h) \quad (1)$$

H_t is the current health status measured by the number of difficulties of activities of daily living (ADLs), body mass index (BMI), self-assessed general health status (GHS), hemoglobin, hypertension, and lung capacity. N_t is the health inputs at time t and the health production function contains not only the current health inputs, but also the inputs from all previous periods which may potentially have impacts on the current health status. The model assumes that health inputs should be distinguished from consumption goods, and to the degree that they affect health, consumption goods should be regarded as health inputs. Individuals can not directly obtain utility from consuming health inputs. I_t captures infrastructure availability and the disease or public health environment in the community where the elderly person lives at time t. D_t denotes time-varying demographic characteristics of the older adult such as the respondent's age. γ_{et} reflects time-varying health shocks from birth to present. γ_e is individual time-invariant characteristics including the older adult's gender and genetic endowments. μ_{ht} and μ_h summarize the information about household specific time-varying and time-invariant characteristics respectively. Following

Grossman (1972), I assume that the one-period lagged health is sufficient to capture the information of all previous factors such as health inputs, community environments, and other time-varying demographic characteristics from time 0 until t-1 in the sample.[3] Thus, the current health can be re-expressed as a function of one-period lagged health status, current period health inputs, current period community environment, individual specific demographic characteristics, health shocks, and household specific characteristics.

$$H_t = h(H_{t-1}, N_t, I_t, D_t, \gamma_{et}, \gamma_e, \mu_{ht}, \mu_h) \tag{2}$$

Each household solves an inter-temporal utility maximization problem subject to a lifetime budget constraint and individual's dynamic health production function at time t. The household utility function in each period depends on consumption goods, C_t, leisure, L_t, and the health status of the older adult, H_t. The optimal choice of health inputs can be derived from the following household utility maximization problem:

$$Max : U = E_t \sum_{t=0}^{T} \beta^t u_t[C_t, L_t, H_t] \tag{3}$$

$$s.t. V_T = (\prod_{t=0}^{T}(1 + r_t)V_0) + \sum_{t=0}^{T}(\prod_{t=0}^{\tau}(1 + r_\tau))(w_t(T_t - L_t) + Y_t - P_t^c C_t - P_t^n N_t) \tag{4}$$

$$H_t = h(H_{t-1}, N_t, I_t, D_t, \gamma_{et}, \gamma_e, \mu_{ht}, \mu_h) \tag{5}$$

β is the subjective discount factor reflecting the extent to which household prefer future utility to current one. P_t^c and P_t^n denote the prices of consumption goods and health inputs respectively. T_t is total time endowed to each household and w_t is a wage rate. V_0 is a household assets at time 0. Y_t captures non-labor income at time t.

In order to solve this optimization problem, following Foster (1995), I assume that the household utility function U is inter-temporally separable and each sub-utility functions, u_t is quasi-concave

[3]This assumption can be considered somewhat strong. However, testing this assumption is beyond the scope of this paper.

and increasing. The household is also assumed to borrow or lend freely against its future in each time period t. Given these assumptions, the household utility maximization problem (3) can be solved subject to a lifetime budget constraint (4) and individual's dynamic health production function (5), and thus, the optimal dynamic conditional health input demand function N_t^* can be expressed as:

$$N_t^* = f(H_{t-1}, I_t, P_t^c, P_t^n, w_t, D_t, Y_t, \gamma_{et}, \gamma_e, \mu_{ht}, \mu_h, E_t(A_{t+j})) \quad , \text{ J} = 1, ..., T-t \tag{6}$$

$$\text{and } A_t = I_t, P_t^c, P_t^n, w_t, D_t, Y_t, \gamma_{et}, \gamma_e, \mu_{ht}, \mu_h$$

By replacing N_t in the equation (5) by N_t^* in equation (6), the dynamic conditional health demand function follows as:

$$H_t^* = h(H_{t-1}, I_t, P_t^c, P_t^n, w_t, D_t, Y_t, \gamma_{et}, \gamma_e, \mu_{ht}, \mu_h, E_t(A_{t+j})) \quad , \text{ J} = 1, ..., T-t \tag{7}$$

$$\text{and } A = I_t, P_t^c, P_t^n, w_t, D_t, Y_t, \gamma_{et}, \gamma_e, \mu_{ht}, \mu_h$$

This dynamic health demand function specifies that the current health status is a function of one-period lagged health, H_{t-1}, resident community environments, I_t, prices of consumption goods P_t^c and health inputs, P_t^n, wage rates, w_t, a set of time-varying and time-invariant characteristics of the elderly and the household as well as household's expectation at time t, $E(A_{t+j})$. Each household forms expectations about the future changes in prices, community environment, demographic characteristics, and other factors. These expectations are used in the optimal choice of current health inputs.

4 Data and Variables

The data comes from the 1993, 1997, 2000 and 2007 waves of the Indonesian Family Life Survey (IFLS). IFLS is a large-scale socio-economic survey conducted in Indonesia which contains extensive information collected at the individual, household and community levels. The survey includes not only indicators of economic, but also of non-economic well-being such as consumption, expenditure, income, education assets, migration, fertility, use of health care, health insurance, marriage, kinship among family members, and labor market outcomes.

The IFLS is a continuing longitudinal survey. The first wave, IFLS1, was conducted in late 1993 and early 1994 and it represented 83% of the Indonesian population residing in 13 of the country's 27 provinces at the time.[4] Within each of the 13 provinces, enumeration areas (EAs) were randomly selected for final survey purpose. In IFLS1, 7,224 households were interviewed and detailed individual level information was collected from over 22,000 respondents. IFLS2 sought to follow up the same sample and re-contacted 94.4% of household interviewed in IFLS1. The re-contact rate of original IFLS1 households was 95.3 % in IFLS3 and 93.6% in IFLS4. For individuals, re-contact rates are also high. In particular, of those individuals who were interviewed in IFLS1, 88% were followed up in IFLS4. Among individuals surveyed in IFLS1, 2,743 had died by IFLS4. Among age groups, the highest re-contact rates (over 90% in 2007, of original 1993 respondents) are observed in those who were older than 40 years in 1993.[5] The attrition level is very low in IFLS because IFLS has tracked individuals who had moved or split off from the origin households. A potential worry associated with attrition is that there might be systematic differences between those who are re-interviewed and those who attrit. As Thomas et al. (2012) show, tracking, particularly over long-distances, in IFLS has helped to mitigate any selection bias.[6]

[4]These are four provinces on Sumatra (North Sumatra, West Sumatra, South Sumatra, and Lampung), all five of the Javanese provinces (Jakarta, West Java, Central Java, Yogyakarta, and East Java), and four provinces covering the remaining major island groups (Bali, West Nusa Tenggara, South Kalimantan, and South Sulawesi).

[5]See Strauss et al.,(2009) and Thomas et al.,(2012), for more details

[6]Thomas et al., (2012) find that attrition is selected on not only age, location and gender but also socio-economic characteristics in IFLS. They show that respondents who were found, but tracked over long distances resemble more attriters who were never found, than they do respondents who were found but never moved.

Sample

In order to focus on the elderly population, I restrict the sample to those respondents who are 50 years or older in 2007 and hence, 36 or older in 1993. The sample sizes are slightly different across health measures because some health measures are available for all four waves, while others for only three, and because of missing data. Some of the information comes from proxy interview, collected from one of the family members, if the main respondents are not available. In this case, self-reported health such as measures of physical functioning and ADLs or general health status can be collected but other bio-markers such as hemoglobin or hypertension cannot. As a final sample, a total of 2351 men and 2768 women are analyzed for the study of general health status dynamics.

Health Measures

Considering that health is a multidimensional composite, I analyze several health measures, both biomarkers and self-reports, known to be important for older adults. These include the number of difficulties with measures of physical functioning and activities of daily living (ADLs), body mass index (BMI), self-assessed general health status (GHS), blood hemoglobin (HB), hypertension, and lung capacity. I use these health measures both as dependent variables and lagged independent variables.

ADLs are defined as routine activities for self-care in the individual's place of residence, as well as outdoor environments. In IFLS, respondents were also asked to choose how easily they could do activities related to physical functioning. Although more activities were added in the 2007 wave, 9 activities that were originally asked in the 1993 and re-asked in subsequent waves are used in the analysis. These are: (a)*carry a heavy load (like a pail of water) for 20 meters*, (b)*sweep the house floor yard*, (c)*walk for 5 kilometers*, (d)*draw a pail of water from a well*, (e)*bow, squat, kneel*, (f)*dress without help*, (g)*stand up from sitting position in a chair without help*, (h)*go to the bathroom without help*, (i)*stand up from sitting on the floor without help*. Respondents can select the degree of difficulty to be either "easily", "with difficulty" or "unable to do it". To construct the variable

for this study, each answer is scored as 1 if respondents report that they can do any of those listed activities only "with difficulty" or "unable to do" it and 0, otherwise. The sum of these scores is then used as the indicator of the number of difficulties with measures of physical functioning and ADLs.

BMI is calculated as weight (in kilograms) divided by height squared (in meters). Following World Health Organization standards (WHO), dummy variables are created for being underweight if BMI is under 18.5, and for overweight if BMI is greater or equal to 25. Many papers have found that BMI has been rising for older people, especially women, in Asia and Indonesia.[7]

In all four waves of the IFLS, respondents were asked to evaluate their own health. The question given to each respondent was *"In general, how is your health?"* with the following options: "very healthy", "somewhat healthy", "somewhat unhealthy", and "unhealthy". Those who answered "somewhat unhealthy", and "unhealthy" are coded as 1, as having *poor health*, while the remaining are coded as 0, as in *good health*. There might be concerns regarding how convincingly GHS works as a reflection of a individual true level of health. As a subjective measure, GHS may be influenced by factors such as the frequency of doctor's visits or mood at the time. However, this measure has been widely used,[8] and many studies have demonstrated that GHS can predict subsequent mortality very well in surveys such as the HRS and ELSA.[9]

Hypertension, a chronic medical condition of having elevated blood pressure, can indicate the risk of coronary heart disease, including stroke, heart attacks, and heart failure. The blood pressure of each respondent was measured three times in the 2007 wave of IFLS and once in the earlier waves, starting in wave 2.[10] For accuracy reasons, the average of the last two measurements is used to construct the hypertension measure in 2007. The dummy variable capturing the prevalence of hypertension is created following the standard definition of the WHO; 1 for those whose *systolic is greater than or equal to 140* or *diastolic is greater than or equal to 90*, and 0 otherwise.

Blood hemoglobin is also one of the important health measures to be examined. It measures

[7]See for example, Popkin, 1994, Monteiro et al., 2004, Strauss et al., 2004, and Witoelar et al., 2012
[8]See Hurd et al. 1995, Adams et al. 2003, Contoyannis and Jones 2004
[9]See Idler et al.(1997), Burstrom and Fredlund 2001, van Doorslaer and Gerdtham 2003
[10]Blood pressure and pulse were measured with an Omron digital measuring device.

anemia, but can also be raised by infections. In IFLS, hemoglobin levels were examined, using the Hemocue meter in waves 2 through 4.[11] A dummy variable is created to capture low hemoglobin level, as equal to 1 for those whose hemoglobin level is below the WHO thresholds (for men: 13g/dL, for women: 12g/dL).

Finally, I examine the lung capacity of older adults. Using the Personal Best Peak Flow Meters, it was measured three times in each wave, starting in wave 2, and the average of three measurements is used in the analysis.[12]

Health Transition Matrices

A very simple picture that shows the health dynamics is presented by the health transition matrices in Tables 1.1 - 1.5. These matrices show the degree to which previous health correlates with current health. Five health measures are studied; low hemoglobin, hypertension, general health status, underweight and overweight. Covariances and correlations between past and current health are computed as follows, taking into account that these health measures are binary variables.[13]

$$\begin{aligned} Cov\left(Y_t, Y_{t-1}\right) &= P\left(Y_t = 1, Y_{t-1} = 1\right) - P\left(Y_t = 1\right) P\left(Y_{t-1} = 1\right) \\ &= \frac{n_{11}}{n_{11} + n_{01} + n_{10} + n_{00}} - \frac{n_{11} + n_{01}}{n_{11} + n_{01} + n_{10} + n_{00}} \frac{n_{11} + n_{10}}{n_{11} + n_{01} + n_{10} + n_{00}}. \end{aligned}$$

where n_{ij} is the number of individuals whose $Y_{t-1} = i$ and $Y_t = j$. Among these, underweight and overweight show the highest correlations between outcomes in one wave and in the prior wave. People who were underweight or overweight in the past mostly tend to remain so. This is not very surprising since height is a stock measure of health and weight has both stock and flow dimensions. General health status has the lowest correlations and especially so for the later waves. Given the nature of subjective measures, this measure can be confounded with many other factors such as

[11] Homocue Method involves pricking a finger and collecting a drop of blood into a cuvette, then inserting the cuvutte into the Hemocue measuring device.

[12] A peak flow meter is a device that measures how fast air comes out of the lungs when one exhales forcefully.

[13] Corr(Y_t, Y_{t-1})=Cov(Y_t, Y_{t-1})/Std.Dev(Y_t)*Std.Dev(Y_{t-1})

subjective mood at the time of survey or their expectation of being healthy. Low hemoglobin has lower correlations across waves. The common causes of low hemoglobin are nutritional deficiency, especially iron, infections from disease and other causes. Depending on changes in diet or whether a person is recovered from certain disease or infection, this measure can vary over time. Hypertension shows a strong correlation over waves.

The main goal of this paper is to explain where this persistence comes from. The strong correlations between lagged and current health measures can represent the impact of lagged health or fixed unobserved factors such as genetic endowments or childhood health. Using first-difference two-step GMM, I attempt to disentangle these factors and discuss the findings in the result section.

Other Control Variables

A set of individual and household characteristics is controlled for in the analysis. The set includes the age and the education level of the respondent, education interacted with age, duration (the gap between two consecutive surveys), duration interacted with respondent's age, and per capita household expenditure (PCE). In order to capture the non-linear impact of age, the dummy variables are created for respondent's age with 5-year intervals: *45-49, 50-54, 55-59, 60-64, 65 and older.*[14] For the same reason, education is also categorized as having *at least some primary school, completed primary school, completed junior high school, and completed high school or more* with no schooling as being the omitted group. The interaction term between respondent's age and the years of education helps to understand whether the education has differential impacts among different age groups. For instance, in the case of GHS, a positive coefficient of this term would indicate that the higher the education, the lower the chance of having poor health in subsequent periods among younger cohorts, possibly because the quality of education has improved over time. A negative coefficient would indicate that schooling makes more difference as people age.

IFLS is not fielded annually and the periods between consecutive surveys are not even. As noted earlier, the first survey was conducted in 1993, and four years later, the second survey was

[14] Age 36-44 is the omitted group.

followed up. IFLS 3 and IFLS 4 were completed in 2000 and 2007 respectively. Following Hoddinot and Kinsey (2001), duration, the period length between the two consecutive surveys, is included to control for uneven gaps between surveys.[15] In the first-difference setting, the information on duration and the age of respondents would be collinear and therefore, one-period lagged age is used instead in the estimation. The regression also includes the interaction term between duration and respondent's lagged age. This term captures age-differential effects over time. In section 3, it is shown that theoretically each household maximizes its utility assuming that they can freely lend or borrow against future periods. However, the assumption of complete markets in developing countries may not be realistic. Thus, I relax this assumption by including a proxy for a household's constraint to access to credit, a one-period lagged log of household per capita expenditure (PCE).[16] [17] All control variables are treated as exogenous. The main statistics are reported in Tables 2.1 and 2.2.

Parental Characteristics

The respondents focused in this study are 50 years and older in 2007 which means their parents would be at least 65 at the time. Several variables are available in IFLS to measure the latest health of each respondent's biological parents and their schooling. Parental schooling is treated as time-invariant since it is reasonable to assume that parents have already completed education given their ages.

The sources of information are different for co-resident parents and for non-coresident ones. For non-coresident parents, it was the adult child respondent who was interviewed about the health status and the level of education of their biological parents. In IFLS4, respondents were asked

[15]The duration is measured in months. For instance, if the first survey was done in June 1993 and the second survey was followed up in July 1997, then the duration would be 49 months.

[16]PCE is converted in real terms across both years and locations. i.e. it is converted into January 2007 Jakarta index.

[17]Current PCE may be exceptionally high or low because of a sudden shock like good rainfall or an accident to a worker preventing them from working. Current PCE and current health are correlated because of the shock. Lagged PCE will then be correlated with the first-differenced shock, $\varepsilon_t - \varepsilon_{t-1}$. To account for this, in some specifications, log PCE_{t-2} and log PCE_{t-3} are used as instruments but with no differences in estimation results.

about the current health of their non-coresident mother and father if their parents were still alive at the time, or the latest health if they died before 2007.[18] For non-coresident parents, therefore, the health information from the 2007 wave is used in the analysis. However, for parents who lived together with respondents at the time of the survey, parents were directly interviewed. Hence, if parents were alive in 2007, the information comes from IFLS4, but if they died between the surveys, a previous wave is used.[19]

More specifically, dummy variables are created for being dead in 2007, having difficulties with ADLs, and poor general health status (GHS). The parents' death dummy variables, one for each biological parent, are equal to 1 if the parent was dead at the time of the survey in 2007. In our sample, only 6 % of fathers and 17% of mothers were still alive at the time, so this dummy indicates a particularly healthy parent if it is 0.[20] The GHS measures are constructed based on the response to the question *"How is the health status of your father/mother now/before his/her death"*. Similar to respondent's own GHS variable, the GHS dummy variable for parents equals 1 if the respondents report the health of their parents as "somewhat unhealthy" or "unhealthy" and 0 otherwise. For difficulties with ADLs, respondents choose "Yes" or "No" for the following question: *"Now/before death does/did your father/mother need help with basic personal needs like dressing, eating, or bathing?"*, and a dummy variable is created as 1 for those who answered "Yes" and 0 for "No". The levels of schooling of each parent are also constructed as dummy variables for level completed: *at least some primary school, completed primary school*, and *junior high school and more* with *no schooling* as being the omitted category.

[18]Some may be concerned that health just prior to death may be worse and do not represent health earlier in life. In our sample, distribution of GHS are actually different, worse for dead parents (a chi-square statistic of differences are 40.7 for fathers and 47.5 for mothers; both are significant at under .01). To address this problem, in the previous paper (Kim et al., (2011)), we compare the estimates from regressing respondent's own health measures on parental characteristics with including the interactions between parents' death and parents' GHS, in the same regression. These interactions are not jointly significant, so we include only the level health variables.

[19]For example, co-resident parents, if they died in 1999 (between IFLS2 (1997) and IFLS3 (2000)), IFLS2 (1997) is used.

[20]In the previous paper on intergenerational transmission of health (Kim et al., (2011)), we tried separate dummies for death before age 60, death after age 60. These turned out not to be significantly different from each other, so we choose the dummies for death, as it is.

Respondent's Birth Cohort

The respondent's birth cohort is constructed by using respondent's age in 2007.[21] The dummy variables are created with 5-year intervals such as 55-59, 60-64, 65-69, 70-74, and 75 and older, to capture different experiences with respect to health, economic and social environments.

5 Empirical Specification

The main goal of this study is to characterize the determinants of chronic diseases and explain their persistence. The empirical framework of the dynamic conditional health demand function can be written as follows:

$$H_{it}^{*} = \beta_H H_{it-1}^{*} + \sum_{j=1}^{S} \alpha_j X_{jit} + \sum_{j=1}^{R} \gamma_j Z_{ji} + \varepsilon_i + \varepsilon_h + \varepsilon_c + \varepsilon_{it} \tag{8}$$

where H_{it}^{*} and H_{it-1}^{*} are the true health status of older adult i as measured by ADLs, BMI, GHS, hemoglobin, hypertension, and lung capacity at time t and t-1, respectively. However, we observe $H_{it} = H_{it}^{*} + \eta_i$ instead, where η_i is measurement error. I assume for now that measurement error is time-invariant.[22] After substituting H_{it} into equation (8), the health demand function becomes:

$$H_{it} = \beta_H H_{it-1} + \sum_{j=1}^{S} \alpha_j X_{jit} + \sum_{j=1}^{R} \gamma_j Z_{ji} + \varepsilon_i + \varepsilon_h + \varepsilon_c + \varepsilon_{it} + (1-\beta)\eta_i \tag{9}$$

where Xs are time-varying regressors such as respondent's age and household incomes. Zs are time-invariant regressors which include respondent's gender and education.[23] Different sources of unobservables are captured by ε_S. First, ε_i captures individual specific unobserved factors that do

[21] Using respondent's age to construct the birth cohort dummy can lead to possible collinearity. However, I use respondent's one-period lagged age as a covariate on RHS. Moreover, lagged age is dropped after first-differencing, since it can be still collinear with duration and duration is measured in months, not years.

[22] This assumption will be relaxed later and time-varying measurement error will be considered as well.

[23] Unlike other papers, the education level of the respondents is treated as time-invariant since the sample includes only elderly who are 50 years and older.

not change over time, such as genetic endowments and childhood health outcomes. Unfortunately in this paper, it is not possible to distinguish genetics from environment in childhood. Second, ε_h, represents time-invariant household specific unobservables, like spousal preference towards respondent's health. ε_c captures time-invariant community specific unobservables such as time-invariant community health factors. Lastly, the household's future expectation, $E_t(A_{t+j})$ in equation (7) enters the empirical specification through ε_{it}.[24] ε_{it} captures the time-varying, individual-specific unobservables, such as individual health or income shocks, or other demographic characteristics that affect current health status but are not observed by econometricians at time period t.

It is difficult to expect a zero correlation between the error term and the explanatory variables when it comes to the inclusion of the lagged dependent variable, H_{t-1} on the right hand side.[25] One-period lagged health status is likely to be correlated with time-invariant unobservables such as $\varepsilon_i, \varepsilon_h$, and ε_c. For example, those born with weaker health endowments are less likely to maintain good health which may lead to having lower education and income, which further leads to poorer health outcomes in subsequent periods. In this case, the estimates without considering the endogeneity issue, would be upward biased. The time-invariant household-specific unobservables and community-specific unobservables also cause correlation with the lagged health measures, creating bias in the estimated coefficients. First-differences can solve this endogeneity problem as it removes all time-invariant unobserved heterogeneity as well as any time-invariant measurement error.

$$\Delta H_{it} = \beta_H \Delta H_{it-1} + \sum_{j=1}^{S} \alpha_j \Delta X_{jit} + \Delta\varepsilon_{it} \tag{10}$$

However, even after first differencing, OLS estimates may still be biased because of the correlation between ΔH_{it-1} and $\Delta\varepsilon_{it}$. ε_t denotes time-varying unobservables such as time-varying health or income shocks that not picked up by PCE. Health at time t-1 is affected by these shocks at t-1 and hence, the endogeniety problem still remains. In this paper, equation (10) is estimated by general method of moments (GMM), where the first-differenced health measure is instrumented

[24] If expectations do not change, they will be part of ε_i and swept away after taking first-difference.
[25] See Deaton(1997), Blundell and Bond(1998) and Wooldridge(2002)

with the two-period (and the three-period) lagged health measure in accordance with Arellano and Bond (1991).[26] In addition, I use external instruments such as parents' health and education, and respondent's own education and birth cohort. These are differenced out of equation (10) but may still be correlated with lagged changes in health when further lags in health are not used as lagged right-hand side variables, as they would not be in a first-stage regression for ΔH_{it-1}.[27]

A valid instrument for the first-differenced model should satisfy the following conditions: (a) it should be correlated with changes in the endogenous regressor and (b) uncorrelated with the differenced error term in the second-stage regression.

The lagged variables of order two or higher, H_{it-2}, H_{it-3} are valid instruments in a first-differenced dynamic panel model since they are correlated with $H_{it-1} - H_{it-2}$ but not with $\varepsilon_{it} - \varepsilon_{it-1}$. This does maintain the assumption that ε_{it} are not serially correlated.[28] Parental characteristics are also strongly associated with the first-differenced lagged health measures. For example, Kim et al. (2011) show that having a mother with poor health status is correlated with changes in the number of ADL problems for elderly men. For women, having a dead mother is positively correlated with moving from no hypertension to having hypertension. If a father has completed primary school, the adult child is less likely to develop hypertension and less likely to have an increased the number of ADL difficulties over time.[29] In addition, changes in one's health status is strongly associated with own socio-economic status including levels of schooling. Respondent's birth cohort is also used as an IV. Indonesia has developed rapidly during the past few decades and therefore, it is more likely that younger respondents had more opportunities for having higher education and had better health infrastructure.

[26] Two-period and three-period lagged health measures are used as internal IVs for ADLs, BMI and GHS, which are available in all four waves. For the other health measures, just two-period lagged health measure is used as an internal IV because they were collected only since IFLS2 in 1997, for three waves.

[27] Place of birth dummies are also used as instruments in a separate specification along with the birth cohort. They are constructed at the district level. The results are similar to those with using only the birth cohort as IV. However, they are not reported in the paper due to the collinearity concern arising from the fact that many people still live in the same place as they were born.

[28] This assumption can be tested. In the next section, the Hansen J statistic, a test of over-identifying restrictions will be reported.

[29] See Kim, Strauss, Witoelar and Sikoki, (2011) for more details.

6 Results

6.1 The Effects of Lagged Health Measures

The results from estimating the dynamic conditional health demand function using GMM estimation are presented in Tables 3 - 8. Results are reported separately for elderly men and women, since the health transitions could be different, depending on gender. In all specifications, an interaction term between place of residence and survey year dummy variables are included to control for time-varying characteristics such as changes in prices and environment at the district level. Some may worry that the place of residence dummy variables after first-differencing might capture migrations, and this might be endogenously determined. For instance, if one decided to move closer to the hospital after getting sick, the previous unobserved health shock could be correlated with the current place of residence. Hence, I use one-period lagged level dummy variables for the place of residence instead.[30]

The first two columns of each table report the estimates from a standard OLS regression, without first-differencing. In the case of low hemoglobin level, the OLS estimates, β_H are 0.230 for men and 0.227 for women; for hypertension, 0.457 for men and 0.440 for women; and 0.576 for men and 0.416 for women for lung capacity. The results from the OLS regression show strong path dependence in several health outcomes. This is consistent with the findings of the transition matrices presented in Tables 1.1 - 1.5. People seem to maintain their health; those who were healthy tend to remain healthy while people with health problems in the past tend to continue to carry them in the subsequent period as well. As discussed earlier, the OLS estimates before taking first difference do not address any endogeneity problems. However, it can still provide a general picture on the extent of the bias that can arise from time-invariant and time-varying omitted variables.

Taking first-differences partially solves the problem by removing the omitted variables which are invariant over time. Running an OLS regression after first-differencing (FD-OLS), the coefficients are estimated at -0.432 for men and -0.492 for women for low hemoglobin level, -0.434 for men and

[30]The equation is written as $\Delta H_{it} = \beta_H \Delta H_{it-1} + \sum_{j=1}^{S} \alpha_j \Delta X_{jit} + dist._{it-1} * Survey\,year + \Delta\varepsilon_{it}$.

-0.462 for women for hypertension, and -0.389 for men and -0.357 for women for lung capacity. The negative coefficients on lagged health represent regression to the mean; those who had more serious health problems tend to recover, while those with less severe problems in the past tend to develop more problems in the subsequent period. Compared to the standard OLS estimates, the magnitude of the coefficients becomes much lower after first-differencing. However, first-differencing only corrects for time-invariant unobservables, while bias could still remain because of time-varying omitted variables. In order to deal with both problems at once, I adopt the first-difference two-step generalized method of moments (FD-GMM).[31]

The first-difference GMM specification (FD-GMM) uses several instruments to identify the first-differenced one-period lagged health status. These instruments are the two-period and three-period lagged health measures, the respondent's education, the parents' health and education, and respondent's birth cohort.[32] The columns 5-6 of each table report the results from using the entire set of instruments. For low hemoglobin, using the two-step GMM, the coefficient on one-period lagged health, β_H, is estimated at 0.0205 for elderly men and 0.0677 for elderly women.[33] This implies that women with low hemoglobin are only about 7 percentage points more likely than healthy women to carry the problem in the subsequent period. Similarly, in the case of hypertension, the coefficient estimates on lagged health are 0.0894 for men and 0.179 for women. This suggests that compared to people who did not have hypertension in the past, men with hypertension are about 9 percentage points more likely to have hypertension in the the subsequent period and for women, it is 18 percentage points more likely. For lung capacity, the estimates are 0.110 for men and 0.0134 for women. The magnitude of the coefficients on lagged health measures indicates the extent to which previous health has an effect on current health, conditioned on fixed effects. A coefficient of one on one-period lagged health measure implies the current health is determined completely by previous health. A coefficient equal to zero would indicate that past health does

[31]Two-step GMM uses the optimal weighting matrix and hence, it is more efficient than 2SLS. Unlike 2SLS, two-step GMM accounts for the case with error terms being heteroskedastic.

[32]For hemoglobin, hypertension and lung capacity, only two-period lagged health measures can be used as internal IVs. Three-period lagged health measures are added as internal IVs for ADLs, BMI and GHS.

[33]See column 5-6, Table 3

not have any effects on subsequent health and hence, suggest that older adults can completely overcome previous health conditions. Given that the coefficients on lagged health status measured by hemoglobin, hypertension, and lung capacity are relatively close to zero, previous health does not matter much in determining current health. Two-step GMM results show similar patterns for other health measures such as ADLs, BMI and GHS. This implies that time-invariant unobservables that were swept away by first-differencing play very important roles in determining current health. The OLS estimates on past health were upward biased because past health conditions are most likely to be strongly correlated with time-invariant fixed effects such as genetic endowments or childhood health. Netting these influences out, little or no correlation is left with lagged health.

In order to be valid, IVs should be not only strongly correlated with endogenous regressors but also uncorrelated with the error term in the second-stage regression. I cannot use the specification test for lack of serial correlation in the first-difference residual proposed by Arellano-Bond (1991) because it needs at least five survey rounds. As an alternative test, the Hansen J statistic, a test for over-identifying restrictions, is used. This is a joint test to check both serial correlation of the error term and the correlations between instruments and error-terms. For low hemoglobin, the p-value of J-statistics are reported as 0.124 for men and 0.784 for women. In the case of hypertension, it is 0.166 for men and 0.307 for women. I found the similar results for other health measures as well. The results suggest that the null hypothesis that all instruments are orthogonal to the error term cannot be rejected and that there is no serial correlations between error-terms.

As another evidence of instrument validity, F-statistics on excluded IVs are reported for the first-stage regressions of changes in health measure in Tables 3-8. A valid instrument (IV) should be correlated with the endogenous regressor. However, if the correlation between IVs and the endogenous variable is too weak, the results from using IVs are likely to be inconsistent and suffer from larger bias than OLS estimates.[34] Following Staiger and Stock (1997), F-statistics from the regressions of endogenous variable on excluded IVs are examined as a test for the relevance of IVs. In the case of low hemoglobin, it is 49.633 for men and 49.555 for women. For hypertension,

[34]See Murray (2006) for more details

F-statistics are 26.581 and 29.580 for men and women respectively. The set of IVs used in this paper are valid and strongly correlated with the endogenous regressor for all health measures since the F-statistics are larger than 10, the rule of thumb values suggested by Staiger and Stock (1997).

6.2 The Effects of Other Co-variates

Duration is included on the right-hand side to control for the uneven gaps between two consecutive surveys in IFLS. In the case of low hemoglobin, the OLS estimates are 0.00475 for elderly men and 0.00677 for elderly women. This implies that for every additional month between surveys, there is a 0.6 percent increase in the chance that elderly women will have lower hemoglobin. However, once the unobservables are taken into account, the duration is no longer significant in FD-GMM specification. On the other hand, duration interacted with the age of the older adult, capturing age-differential effects over time, remains significant in FD-GMM specification. For the study of low hemoglobin, a positive coefficient indicates that the longer the duration is, the higher the chance of having low hemoglobin is in the subsequent period among older cohorts.

Per capita household expenditure (PCE) is controlled for as a socio-economic status (SES) characteristic.[35] The OLS estimate on one-period lagged log(PCE) is 0.0138 for elderly women in the case of being overweight. This captures the positive impact of resources on current health as measured by BMI. The magnitude of the coefficient becomes not significantly different from zero, once fixed effect and the influence of past health are conditioned on, which demonstrates that there is an upward bias in the OLS estimates on log(PCE) due to the possible correlation with the time-invariant unobserved heterogeneity.[36]

[35] The reason that I use PCE instead of household income is the potential worry about systematic measurement errors in income. Bound and Krueger (1991) show that, people in the United States with low-income tend to over-report their income in an effort to hide their situation.

[36] The correlation between household specific unobservables and log(PCE) can be one of many possible examples. For instance, the families who prefer high calorie foods may have higher household expenditure.

6.3 Interactions of Lagged Health with Age and SES

The extent to which the past health has influence on current health can differ across SES or demographic groups. For example, past good health may matter less as people age. Past bad health may matter less for better educated people if they have better knowledge of healthy diets or less risky lifestyles and visit doctors more often when they notice any subtle changes in their health. In addition, people with higher income are generally more likely to utilize the health care system than those with lower incomes. In this case, their past poor health may matter less. To get at these potential differences, the sample is disaggregated by age, PCE, and years of education to see if the impact of previous health differs across groups.

Tables 9.1 - 9.3 report the results from the estimation of disaggregated groups. With the exception of hemoglobin level, the degree of path dependence turns out to be much stronger for those who are less educated (defined as less than completed junior high school). For instance, in the case of hypertension, the path dependence coefficients are 0.088 for men and 0.174 for women with lower years of education, while no significant dependence is observed for the higher educated group. As mentioned earlier, this group might not be as knowledgeable as the higher educated group in terms of healthy behaviors and may not go to doctors as often. I find no significant differences by age group or household per capita expenditure.

6.4 Robustness Checks

Time-varying Measurement Errors

In the previous section, for simplicity, I assume that measurement errors are time-invariant. In that case, external IVs such as parents' health and education, and respondent's education and birth cohort as well as two and three-period lagged health measures are valid IVs.[37] However, if the health measures have time-varying measurement errors, η_{it}, the two-period lagged health

[37] Two and three-period health measures are valid under the assumption that error terms are not serially correlated.

measure is no longer valid. The equation accounting for time-varying measurement errors becomes

$$H_{it} = \beta_H H_{it-1} + \sum_{j=1}^{S} \alpha_j X_{jit} + \sum_{j=1}^{R} \gamma_j Z_{ji} + \varepsilon_i + \varepsilon_h + \varepsilon_c + \varepsilon_{it} - \beta_H \eta_{it-1} + \eta_{it} \tag{11}$$

Then, the first-differenced error term in the second stage regression, $\Delta\varepsilon_{it} - \Delta\beta_H \eta_{it-1} + \Delta\eta_{it}$ will be correlated with two-period lagged health through the measurement error, η_{it-2}. One might think that there would be attenuation bias towards zero in the lagged health coefficient if time-varying measurement error is a problem. Three-period lagged health is still uncorrelated with the first-differenced errors so long as the measurement errors are not serially correlated. Hence, I re-estimate the model using three-period lagged health plus the external IVs as instruments. The results are reported in Tables A.1-A.2. Only GHS, BMI, and the number of difficulties with ADLs are examined, since four waves of data are now required.

For most of the health measures, the coefficients on lagged health again turn out to be close to zero, or change sign and become negative; they do not become more positive. Hence, attenuation bias caused by time-varying measurement error does not seem to be a problem. However, it needs to be noted that the F-tests on identifying instruments do indicate a weak instrument problem, unlike when I also use two-period lagged health.[38]

Attrition and Selective Mortality in Panel Data

Another potential problem that can arise in panel data is attrition, which can result in biased estimates, especially if attrition is correlated with endogenous regressors. In the IFLS sample, the attrition arises from three sources; (a)mortality, (b)refusals[39]and (c)migration.[40]

Given the estimation strategy adopted in this paper, however, the attrition resulting from

[38]I try subsets of the instruments to try to address the weak instrument problem. In some cases, the F-statistics do rise, but not by very much.

[39]Compared to higher-income countries such as the United States, refusal rates in household surveys are low in most developing countries. See Thomas et al. (2012) for details.

[40]Some of the health measures such as self-reported GHS or numbers of difficulties with ADLs or physical functioning also partly come from proxy interviews, if the main respondents are not available or moved out of the origin households and hence, cannot be physically traced down. In this case, biomarkers such as hemoglobin, hypertension or weight cannot be measured.

individual, household, and community specific time-invariant characteristics should no longer be a problem. First-differencing removes all sources of time-invariant unobservables which may correlate with attrition and with any mortality selection. The possible remaining source of attrition is from time-varying shocks. If the shocks in 1993 are correlated with those in 1997 and so forth, those who were exposed to shocks would show much stronger path dependence. If attrition is caused by these shocks and people with stronger path dependence tend to leave the sample, then the estimates will be downward biased. Following Fitzgerald et al.(1998), a linear probability model is estimated to determine whether attrition is related to health shocks. The results are reported in Table B in appendix. The dependent variable is attrition, coded as 1 if the individual was lost in either the third or the fourth wave (regardless of the attrition reasons), and 0 otherwise. The right hand side regressors include changes in respondent's health measured by GHS, BMI, and difficulties with physical functioning or ADLs from 1993 to 1997, age and years of education, log(PCE), parent's health and education, and location variables (rural and the place of residence dummy variables), measured in 1993. The coefficients on changes in health, proxies for initial health shocks, are jointly not significantly different from zero. This indicates that health shocks are not correlated with later attrition including death. The results also show that, in the case of GHS attrition, both men and women with higher PCE are less likely to stay in the survey. In addition, for men, those whose fathers have poor GHS are more likely to leave the sample. Women who live in rural area are more likely to be followed up while having dead fathers has opposite effects on attrition.

7 Conclusion

This paper examines the determinants of chronic diseases and explains their persistence, using a panel data set from the Indonesian Family Life Survey (IFLS). This is done by estimating a dynamic conditional health demand function. To address the endogeneity issue that arises from the potential correlation between lagged health measure and unobservables, I employ a first-difference two-step GMM. Two and three-period lagged health measures, parents' health and education, and

respondent's education and birth cohort are used as IVs to identify the first-differenced one-period lagged health measure.

The correlations of health between waves are high but they fall to small numbers in the FD-GMM specification. The coefficient on one-period lagged health measure in a dynamic health demand function captures the degree to which previous health has an effect on current health, conditional on fixed effects such as through the influences of genetics or childhood health. Given that the coefficients are relatively close to zero, I find that there is little path dependence in health status. Socio-economic status also has very little influence on current health, again conditional on fixed effects and on lagged health. It is the fixed effects that evidently play the most important roles in determining chronic health conditions in later life.

In order to investigate if past health has different impacts across demographic or economic groups, I disaggregate the sample across age, per capita household expenditure level (PCE), and years of education. The results show that those with less education tend to show more persistence, compared to those with higher education, but with no differences by age or PCE.

In terms of policy, a strong impact of early childhood health implies that investments should be targeted towards the early stages of life. Stronger persistence of disease of lower educated elderly also suggests that investments should be made not only towards health in early life, but also towards better education in early life. Among the elderly, the focus should be on those with lower education attainments.

References

[1] Adair, L.S. 1999. "Filipino children exhibit catch-up growth from age2 to 12 years", *Journal of Nutrition*, 129:1140-1148.

[2] Adams, P., Hurd, M.D., McFadden, D., Merril, A. and Ribeiro, T. 2003. "Health, wealthy, and wise? Tests for direct causal paths between health and socioeconomic status", *Journal of Econometics*, 112(2003):3-56.

[3] Alderman, H., Hoddinott, J. and Kinsay, W. 2006. "Long-term consequences of early childhood malnutrition", *Oxford Economic Papers*, 58(3):450-474.

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

[5] Almond, D. and J. Currie. 2011. "Killing me sofely: The fetal origins hypothesis", *Journal of Economic Perspectives*, 25(3):153-172.

[6] Arellano, M. and Bond, S. 1991. "Some tests of specification for panel data: Monte carlo evidence and an application of emplyoment equations", *Review of Economic Studies*, 58:277-297.

[7] Barker, D. 1994. "Mothers, babies and helath in later life", London: BMJ Publishing Group.

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

[9] Blundell, R. and S. Bond. 1998. "Initial conditions and moment restrictions in dynamic panel data models", *Journal of Econometrics*, 87:115-143.

[10] Bound, J. and Krueger, A.B. 1991. "The extend of measurement error in longitudinal earnings data: do two wrongs make a right", *Journal of Labor Economics*, 12: 345-368.

[11] Burstrom, B. and Fredlund, P. 2001. "Self rated health: Is it as good a predictor of subsequent mortality among adults in lower as well as in higher social classes?", *Journal of Epidemiology and Community Health*, 55, 836-840.

[12] Cameron, N., Preece, M.A. and Cole T.J. 2005. "Catch-up growth or regression to the mean? Recovery from stunting revisited", *American Journal of Human Biology*, 17:412-417.

[13] Case, A., Fertig, A., and Paxson, C. 2005 "The lasting impact of childhood health and circumstance",*Journal of Health Economics*, 24, 365-389.

[14] Case, A. and Paxson, C. 2010. "The long reach of childhoood health and circumstance: Evidence from the Whitehall II study", *The Economic Journal*, 121, F183-F204.

[15] Contoyannis, P., Jones, A. 2004. "Socio-economic status, health and lifestyle", *Journal of Health Economics*, 23, 965-995.

[16] Cooksey, E., Eberwein, C. and Olsen, R. 2012. "Depressivity over the Life Course", working paper, Center for Human Resource Research, Ohio State University.

[17] Crimmins, E. and C. Finch. 2004. "Inflammatory exposure and historical changes in human life-spans", *Science* 305, 1736-1739.

[18] Crimmins, E. and C. Finch. 2006. "Infection, inflammation, height and logevity", *Proceedings of the National Academy of Science* 103(2), 498-503.

[19] Currie, J. and E. Moretti. 2007. "Biology as destiny? Short- and long-run determinants of intergenerational transmission of birthweight", *Journal of Labor Economics*, 25(2):231-263.

[20] Deaton, A. 1997. "The analysis of household surveys: A microeconometric approach in deveopment policy", Baltimore, Johns Hopkins University Press.

[21] Deaton, A. and Paxton, C. 1998. "Aging and inequality in health and income", *American Economic Review*, 88(2),248-253.

[22] Dohrenwend, B., et al. 1992. "Socioeconomic status and psychiatric disorders: the causation -selection issue", *Science*, 255,946-952.

[23] Fedorov, L. and Sahn, D. 2005. "Socioeconomics determinants of children's health in Russia: A longitudinal survcey", *Economic Development and Cultural Change*, 479-500.

[24] Fitzgerald, J., Gottschalk, P., and Moffitt, R. 1998. "An analysis of sample attrition in oanel data", *Journal of Human Resources*, 33(2):251-299.

[25] Foster, A. 1995. "Prices, credit markets, and child growth in low income rural areas", *Economic Journal*, 551-570.

[26] Frankenberg, E. and D. Thomas. 2000. "The Indonesia Family Life Survey (IFLS): Study design and results from Waves 1 and 2", DRU-2238/1-NIA/NICHD.

[27] Frankenberg, E. and L.Karoly. 1995. "The 1993 Indonesian Family File Survey: Overview and field report", RAND. DRU-1195/1-NICHD/AID.

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

[29] Goldman, N. 1994. "Social factors and health: the causation -selection issue revisited", *Proceedings of the National Academy of Science*, 91,1251-1255.

[30] Grossman, M. 1972. "On the concept of human capital and demand for health", *Journal of Political Economy*, 80(2):223-255.

[31] Hoddinott, J. and Kinsey, W. 2001. "Child growth in the time of drought", *Oxford Bulletin of Economics and Statistics*, 63(4):409-436.

[32] Hoddinott, J., Maluccio, J., Behrman, J., Flores, R. and Martorell, R. 2008. "Effect of a nutrition intervention during early childhhod on economic productivity in Guatemalan adults", *Lancet*,371, 41116

[33] Hurd, M. and McGarry, K. 1995. "Evaluation of subjective probability distributions in the health and retirement study. ", *Journal of Human Resources*, 30(S):S268-S292.

[34] Idler, E. and Benyamini, Y. 1997. "Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies", *Journal of Health and Social Behavior*, 38(1):21-37.

[35] Indonesian Public Health Association. 1993. "Analysis of the Health Transition in Indonesia:Implications for Health Policy", Jakarta, Indonesia: Indonesian Public Health Association.

[36] Jonas, B. and Mussolino, M. 2000. "Symptioms of depression as a prospective risk factor for stroke", *Phychosomatic Medicine*, 62(4), 463-471

[37] Kim, Y., B. Sikoki, J. Strauss and F. Witoelar. 2011. "Intergenerational correlations of health among older adults: Empirical evidence from Indonesia", RAND Labor and Population Working Paper No. WR-879.

[38] Lleras-Muney, A. 2005. "The Relationship Between Education and Adult Mortality in the United States", *The Review of Economic Studies*, 72 (1): 189-221.

[39] Luo, Y. and Waite, L.J. 2005. "The Impact of Childhood and Adult SES on Physical, Mental, and Cognitive Well-Being in Later Life", *Journal of Gerontology*, 60(2),S93-S101.

[40] Mani, S. 2012. "Is there complete, partial or no recovery from childhood malnutrition?: Empirical evidence from Indonesia", *Oxford Bulletin of Economics and Statistics*,74(5):691-715.

[41] Marmot, M. 1999. "Multi-level approaches to understanding social determinants", *Social Epidemiology*, in Berkman, Lisa and Ichiro Kawachi, eds. Oxford: Oxford University Press

[42] Monteiro, C., Moura, E., Conde, W., and Popkin, B. 2004. "MSocioeconomic status and obesity in adult populations of developing countries", *Bulletin of the World Health Organization*, 82(12):940-946.

[43] Penninx, B., Guralnik, J. and Havlik, R. 1998., "Chronically depressed mood and cancer risk in older persons", *Journal of The National Cancer Institute*, 90(24):1888-1893.

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

[45] Smith, J.P. and Kington, R. 1997. "Demographic and economic correlates of health in old age", *Demography*, 34(1),159-170.

[46] Smith, J.P. 2007. "The impact of socioeconomic status on health over the life-course", *Journal of Human Resources*, 42(4),739-764.

[47] Smith, J.P. 2009. "The Impact of childhood health on adult labor market outcomes", *The Reivew of Economics and Statistics*, 91(3), 478-489.

[48] Staiger, D. and Stock, J.H. 1997. "Instrumental variables regression with weak instruments", *Econometrica*, 65(May):557-586.

[49] Stock, J., Wright, J. and Yogo, M. 2002. "A survey of weak instruments amd weak identification in generalized method of moments", *Journal of Business and Economics Statistics*, 20(4):518-529.

[50] Strauss, J. and D. Thomas. 1995. "Human resources: empirical modeling of household and family decisions", Handbook of Development Economics, vol.3.

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

[52] Strauss, J., K. Beegle, B. Sikoki, A. Dwiyanto, F. Witoelar and Y. Herawati. 2009. "The Fourth Wave of the Indonesian Family Life Survey (IFLS4): Overview and Field Report", RAND. WR-675/1-NIA/NICHD.

[53] Strauss, J. and D. Thomas. 2008. "Health over the life course", Handbook of Development Economics, vol.4.

[54] Thomas, D., Witoelar, F., Elizabeth, F., Sikoki, B., Strauss, J., Sumantri, C. and Suriastini, W. 2012. "Cutting the costs of attrition: Results from the Indonesia Family Life Survey", *Journal of Development Economics*, 98(1),108-123.

[55] van Doorslaer, E. and Gerdtham, U-G. 2003. "Does inequality in self-assessed health predict inequality in survival by income? Evidence from Swedish data", *Social Science and Medicine*, 57, 1621-1629.

[56] Witoelar F., J. Strauss and B. Sikoki. 2012. "Socioeconomic Success and Health in Later Life: Evidence from the Indonesian Family Life Survey", to appear in M. Majmundar and J.P. Smith (eds.).,*Aging in Asia*, National Academy of Sciences Press.

[57] Wooldridge, J. 2002. Econometric Analysis of Cross Section and Panel Data, Cambridge: MIT Press.

Table 1.1: Transition Matrix - Hypertension

	2007 Men		
	No	Yes	Total
2000 No	970 (70.8%)	400 (29.2%)	1,370
2000 Yes	175 (22.21%)	613 (77.79%)	788
Total	1,145	1,013	2,158
$\text{Corr}(H_{07}, H_{00})$	0.4688		

	2000 Men		
	No	Yes	Total
1997 No	1,139 (78.93%)	304 (21.07%)	1,443
1997 Yes	231 (32.31%)	484 (67.69%)	715
Total	1,370	788	2,158
$\text{Corr}(H_{00}, H_{97})$	0.4558		

	2007 Women		
	No	Yes	Total
2000 No	1000 (63.65%)	571 (36.35%)	1,571
2000 Yes	151 (18.21%)	931 (81.79%)	1,082
Total	1,197	1,456	2,653
$\text{Corr}(H_{07}, H_{00})$	0.5197		

	2000 Women		
	No	Yes	Total
1997 No	1,291 (78.72%)	349 (21.28%)	1,640
1997 Yes	280 (27.64%)	733 (72.36%)	1,013
Total	1,571	1,082	2,653
$\text{Corr}(H_{00}, H_{97})$	0.5049		

Table 1.2: Transition Matrix - Low Hemoglobin

	2007 Men		
	No	Yes	Total
2000 No	1,217 (79.65%)	311 (20.35%)	1,528
2000 Yes	294 (49.41%)	301 (50.59%)	595
Total	1,511	612	2,123
$\text{Corr}(H_{07}, H_{00})$	0.2998		

	2000 Men		
	No	Yes	Total
1997 No	1,143 (79.43%)	296 (20.57%)	1,439
1997 Yes	385 (56.29%)	299 (43.71%)	684
Total	1,528	595	2,123
$\text{Corr}(H_{00}, H_{97})$	0.2408		

	2007 Women		
	No	Yes	Total
2000 No	1,171 (75.79%)	374 (24.21%)	1,545
2000 Yes	569 (53.53%)	494 (46.47%)	1,063
Total	1,740	868	2,608
$\text{Corr}(H_{07}, H_{00})$	0.2321		

	2000 Women		
	No	Yes	Total
1997 No	1,109 (69.57%)	485 (30.43%)	1,594
1997 Yes	436 (43.00%)	578 (57.00%)	1,014
Total	1,545	1,063	2,608
$\text{Corr}(H_{00}, H_{97})$	0.2636		

Table 1.3: Transition Matrix - General Health Status (Poor Health = 1)

	2007		
	Men		
	No	Yes	Total
2000 No	1,607 (79.24%)	421 (20.76%)	2,028
2000 Yes	173 (53.56%)	150 (46.44%)	323
Total	1,780	571	2,351
$\text{Corr}(H_{07}, H_{00})$	0.2061		

	2000		
	Men		
	No	Yes	Total
1997 No	1,842 (89.85%)	208 (10.15%)	2,050
1997 Yes	186 (61.79%)	115 (38.21%)	301
Total	2,028	323	2,351
$\text{Corr}(H_{00}, H_{97})$	0.2723		

	1997		
	Men		
	No	Yes	Total
1993 No	1,904 (89.94%)	213 (10.06%)	2,117
1993 Yes	146 (62.39%)	88 (37.61%)	234
Total	2,050	301	2,351
$\text{Corr}(H_{97}, H_{93})$	0.2468		

	2007		
	Women		
	No	Yes	Total
2000 No	1,697 (75.89%)	539 (24.11%)	2,236
2000 Yes	302 (56.77%)	230 (43.23%)	532
Total	1,999	769	2,768
$\text{Corr}(H_{07}, H_{00})$	0.1683		

	2000		
	Women		
	No	Yes	Total
1997 No	1,929 (83.54%)	380 (16.46%)	2,309
1997 Yes	307 (66.88%)	152 (33.12%)	459
Total	2,236	532	2,768
$\text{Corr}(H_{00}, H_{97})$	0.1572		

	1997		
	Women		
	No	Yes	Total
1993 No	2,080 (86.45%)	326 (13.55%)	2,406
1993 Yes	229 (63.26%)	133 (36.74%)	362
Total	2,309	459	2,768
$\text{Corr}(H_{97}, H_{93})$	0.2102		

Table 1.4: Transition Matrix - Underweight (BMI < 18.5)

2007

	Men		
	No	Yes	Total
2000 No	1,376 (92.16%)	117 (7.84%)	1,493
2000 Yes	104 (31.71%)	224 (68.29%)	328
Total	1,480	341	1,821
$\text{Corr}(H_{07}, H_{00})$	0.5955		

2000

	Men		
	No	Yes	Total
1997 No	1,423 (93.74%)	95 (6.26%)	1,518
1997 Yes	70 (23.1%)	233 (76.9%)	303
Total	1,493	328	1,821
$\text{Corr}(H_{00}, H_{97})$	0.6846		

1997

	Men		
	No	Yes	Total
1993 No	1,444 (94.38%)	86 (5.62%)	1,530
1993 Yes	74 (25.43%)	217 (74.57%)	291
Total	1,518	303	1,821
$\text{Corr}(H_{97}, H_{93})$	0.6784		

2007

	Women		
	No	Yes	Total
2000 No	1,758 (93.46%)	123 (6.54%)	1,881
2000 Yes	120 (29.27%)	290 (70.73%)	410
Total	1,878	413	2,291
$\text{Corr}(H_{07}, H_{00})$	0.6401		

2000

	Women		
	No	Yes	Total
1997 No	1,809 (94.27%)	110 (5.73%)	1,919
1997 Yes	72 (19.35%)	300 (80.65%)	372
Total	1,881	410	2,291
$\text{Corr}(H_{00}, H_{97})$	0.7207		

1997

	Women		
	No	Yes	Total
1993 No	1,805 (95.05%)	94 (4.95%)	1,899
1993 Yes	114 (29.08%)	278 (70.92%)	392
Total	1,919	372	2,291
$\text{Corr}(H_{97}, H_{93})$	0.6736		

Table 1.5: Transition Matrix - Overweight (BMI $\geq$ 25)

2007

Men			
	No	Yes	Total
2000 No	1,470 (92.69%)	116 (7.31%)	1,586
2000 Yes	53 (22.55%)	182 (77.45%)	235
Total	1,523	298	1,821
Corr(H_{07}, H_{00})	0.6355		

2000

Men			
	No	Yes	Total
1997 No	1,545 (96.56%)	55 (3.44%)	1,600
1997 Yes	41 (18.55%)	180 (81.45%)	221
Total	1,586	235	1,821
Corr(H_{00}, H_{97})	0.7598		

1997

Men			
	No	Yes	Total
1993 No	1,565 (96.07%)	64 (3.93%)	1,629
1993 Yes	35 (18.23%)	157 (81.77%)	192
Total	1,600	221	1,821
Corr(H_{97}, H_{93})	0.7321		

2007

Women			
	No	Yes	Total
2000 No	1,477 (88.02%)	201 (11.98%)	1,678
2000 Yes	92 (15.01%)	521 (84.99%)	613
Total	1,569	722	2,291
Corr(H_{07}, H_{00})	0.6957		

2000

Women			
	No	Yes	Total
1997 No	1,605 (93.42%)	113 (6.58%)	1,718
1997 Yes	73 (12.74%)	500 (87.26%)	573
Total	1,678	613	2,291
Corr(H_{00}, H_{97})	0.7893		

1997

Women			
	No	Yes	Total
1993 No	1,656 (90.39%)	176 (9.61%)	1,832
1993 Yes	62 (13.51%)	397 (86.49%)	459
Total	1,718	573	2,291
Corr(H_{97}, H_{93})	0.7106		

Table 2.1: Summary Statistics

	MEN					WOMEN				
	1993	1997	2000	2007	Obs	1993	1997	2000	2007	Obs
Low Hemoglobin		0.32	0.28	0.29	2123		0.39	0.41	0.33	2608
		(0.467)	(0.449)	(0.453)			(0.488)	(0.491)	(0.471)	
Hemoglobin level		13.65	13.73	13.77	2123		12.23	12.12	12.45	2608
		(1.781)	(1.809)	(1.807)			(1.627)	(1.509)	(1.55)	
Hypertension		0.33	0.37	0.47	2158		0.38	0.41	0.55	2653
		(0.471)	(0.482)	(0.499)			(0.486)	(0.492)	(0.498)	
Lung capacity		351.82	354.53	315.86	2092		241.32	244.27	212.60	2478
		(92.464)	(99.798)	(106.382)			(64.496)	(60.807)	(68.674)	
Overweight	0.11	0.12	0.13	0.16	1821	0.20	0.25	0.27	0.32	2291
	(0.307)	(0.327)	(0.335)	(0.37)		(0.4)	(0.433)	(0.443)	(0.465)	
Underweight	0.16	0.17	0.18	0.19	1821	0.17	0.16	0.18	0.18	2291
	(0.367)	(0.373)	(0.384)	(0.39)		(0.377)	(0.369)	(0.383)	(0.384)	
BMI level	21.22	21.28	21.33	23.56	1821	21.93	22.45	22.58	27.41	2291
	(4.898)	(3.094)	(3.284)	(58.844)		(3.753)	(4.14)	(4.415)	(90.935)	
GHS (Poor health)	0.10	0.13	0.14	0.24	2351	0.13	0.17	0.19	0.28	2768
	(0.299)	(0.334)	(0.344)	(0.429)		(0.337)	(0.372)	(0.394)	(0.448)	
# of difficulties w/ ADLs	0.22	0.40	0.43	1.06	2342	0.54	1.02	1.09	1.72	2760
	(0.945)	(1.069)	(1.08)	(1.898)		(1.319)	(1.522)	(1.527)	(2.135)	
Respondent's age	48.40	52.40	55.40	62.40	2351	48.31	52.31	55.31	62.31	2768
	(9.493)	(9.493)	(9.493)	(9.493)		(9.413)	(9.413)	(9.413)	(9.413)	
Log (per capita HH expenditure)	12.72	12.73	12.84		2351	12.70	12.74	12.82		2768
	(0.748)	(0.754)	(0.707)			(0.756)	(0.792)	(0.694)		

Source: IFLS 1993, 1997, 2000 and 2007

Standard deviations are reported in parenthesis.

Table 2.2: Summary Statistics

Variables	Description	MEN		WOMEN	
		Mean	Std. Dev	Mean	Std. Dev
Respondent's Edu	At least some primary	0.30	(0.458)	0.30	(0.458)
	Completed primary	0.29	(0.455)	0.18	(0.388)
	Completed Jr. H	0.09	(0.291)	0.07	(0.247)
	Completed High and more	0.16	(0.367)	0.08	(0.277)
Resp's yrs of Edu		5.58	(4.459)	3.56	(4.133)
Father's Education	At least some primary school	0.09	(0.284)	0.07	(0.252)
	Completed primary school	0.25	(0.43)	0.24	(0.428)
	Completed junior high school and more	0.06	(0.229)	0.05	(0.22)
Mother's Education	At least some primary school	0.07	(0.252)	0.05	(0.227)
	Completed primary school	0.17	(0.374)	0.16	(0.365)
	Completed junior high school and more	0.02	(0.14)	0.02	(0.143)
Father's Health	Death =1 if dead in 2007	0.95	(0.221)	0.94	(0.233)
	ADL problem = 1 if need help with basic personal needs now or before death	0.23	(0.421)	0.23	(0.422)
	GHS = 1 if Somewhat unhealthy or very unhealthy now or before death	0.50	(0.5)	0.47	(0.499)
Mother's Health	Death =1 if dead in 2007	0.83	(0.377)	0.83	(0.376)
	ADL problem = 1 if need help with basic personal needs now or before death	0.25	(0.432)	0.25	(0.434)
	GHS = 1 if Somewhat unhealthy or very unhealthy now or before death	0.47	(0.499)	0.44	(0.496)
Birth Cohorts	born before 1932	0.12	(0.323)	0.11	(0.318)
	born between 1932 and 1936	0.11	(0.314)	0.10	(0.302)
	born between 1937 and 1941	0.15	(0.359)	0.16	(0.37)
	born between 1942 and 1946	0.15	(0.355)	0.16	(0.366)
	born between 1947 and 1951	0.22	(0.415)	0.20	(0.396)
Observations		2351		2768	

Source: IFLS 1993, 1997, 2000 and 2007

Table 3.1: Dynamic Health Demand Function Estimation - Low Hemoglobin

	OLS		FD-OLS		FD-GMM	
	Men	Women	Men	Women	Men	Women
Lagged Low HB	0.230***	0.227***	-0.432***	-0.492***	0.0205	0.0677**
	(0.0164)	(0.0139)	(0.0196)	(0.0163)	(0.0335)	(0.0338)
45-50 (lagged age)	0.00787	-0.0246				
	(0.0197)	(0.0200)				
50-55	0.0588**	-0.00203				
	(0.0253)	(0.0246)				
55-60	0.0834***	0.0102				
	(0.0314)	(0.0294)				
60-65	0.119***	0.00868				
	(0.0389)	(0.0374)				
65-	0.170***	0.00894				
	(0.0517)	(0.0494)				
At least some primary	-0.0174	0.0221				
	(0.0274)	(0.0243)				
Completed primary	-0.00489	0.00108				
	(0.0383)	(0.0432)				
Completed Jr. High	0.00303	0.0112				
	(0.0525)	(0.0616)				
Completed High & more	-0.0277	0.0157				
	(0.0678)	(0.0820)				
Yrs of edu * lagged age	-0.0000899	-0.0000713	0.000200	0.000579	0.000165	0.00113
	(0.0000905)	(0.000119)	(0.000858)	(0.000967)	(0.000933)	(0.00110)
Duration	0.00475	0.00677**	-0.00356	0.00505	-0.00410	0.00847
	(0.00310)	(0.00339)	(0.00430)	(0.00423)	(0.00450)	(0.00515)
Duration * lagged age	0.0000240	0.0000718***	0.0000430*	0.000106***	0.0000362	0.0000783***
	(0.0000252)	(0.0000246)	(0.0000221)	(0.0000217)	(0.0000238)	(0.0000257)
Lagged log (PCE)	-0.000947	-0.0166*	0.000101	-0.000718	-0.000295	-0.0264*
	(0.00984)	(0.00927)	(0.0154)	(0.0136)	(0.0167)	(0.0160)
Lagged dist. * year dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	2123	2608	2123	2608	2123	2608
F stat on excluded IVs					49.633	49.555
Hansen J statistics					36.755	22.829
(p-value)					(0.124)	(0.784)

Source: IFLS 1997, 2000 and 2007 *** significant at 1%, ** significant at 5%, * significant at 10%

Robust standard errors adjusted for clustering at the community level are reported in parenthesis.

low HB:1 for those whose hemoglobin level is below the threshold (for men: 13g/dL, for women: 12g/dL)

Dummy variables are included to capture missing values for parent's education, GHS, ADL and household's PCE.

In FD-GMM, $low\ HB_{t-2}$, respondent's edu, parents' health and edu, and birth cohort are used as IVs.

Table 3.2: Dynamic Health Demand Function Estimation - Hemoglobin Level

	OLS		FD-OLS		FD-GMM	
	Men	Women	Men	Women	Men	Women
Lagged HB Level	0.380***	0.337***	-0.452***	-0.448***	0.134**	0.151***
	(0.0180)	(0.0163)	(0.0255)	(0.0227)	(0.0525)	(0.0434)
45-50 (lagged age)	-0.00675	0.0833				
	(0.0801)	(0.0638)				
50-55	-0.197**	0.0241				
	(0.0906)	(0.0740)				
55-60	-0.266**	-0.0104				
	(0.116)	(0.0901)				
60-65	-0.417***	0.145				
	(0.139)	(0.114)				
65-	-0.542***	0.0976				
	(0.187)	(0.153)				
At least some primary	-0.00203	-0.0186				
	(0.100)	(0.0747)				
Completed primary	-0.0410	0.0486				
	(0.147)	(0.134)				
Completed Jr. High	-0.0823	-0.0711				
	(0.211)	(0.187)				
Completed High & more	0.0376	-0.0809				
	(0.260)	(0.249)				
Yrs of edu * lagged age	0.000257	0.000108	0.000804	-0.00415	0.00643*	-0.00615*
	(0.000352)	(0.000362)	(0.00305)	(0.00306)	(0.00355)	(0.00333)
Duration	0.0606***	0.0575***	0.00378	0.00908	-0.00325	0.00460
	(0.0187)	(0.0187)	(0.0158)	(0.0148)	(0.0174)	(0.0142)
Duration * lagged age	-0.000102	-0.000268***	-0.0000922	-0.000375***	0.0000490	-0.000302***
	(0.0000910)	(0.0000788)	(0.0000813)	(0.0000688)	(0.0000892)	(0.0000784)
Lagged log (PCE)	0.0793*	0.0973***	-0.0484	0.0169	-0.0173	0.0286
	(0.0420)	(0.0357)	(0.0569)	(0.0419)	(0.0613)	(0.0470)
Lagged dist. * year dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	2123	2608	2123	2608	2123	2608
F stat on excluded IVs					27.334	43.875
Hansen J statistics					47.706	35.71
(p-value)					(0.012)	(0.182)

Source: IFLS 1997, 2000 and 2007 *** significant at 1%, ** significant at 5%, * significant at 10%

Robust standard errors adjusted for clustering at the community level are reported in parenthesis.

Dummy variables are included to capture missing values for parent's education, GHS, ADL and household's PCE.

In FD-GMM, HB_{t-2}, respondent's edu, parents' health and edu, and birth cohort are used as IVs.

Table 4: Dynamic Health Demand Function Estimation - Hypertension

	OLS		FD-OLS		FD-GMM	
	Men	Women	Men	Women	Men	Women
Lagged Hypertension	0.457***	0.440***	-0.434***	-0.462***	0.0894**	0.179***
	(0.0146)	(0.0128)	(0.0212)	(0.0193)	(0.0448)	(0.0449)
45-50 (lagged age)	0.0395*	0.0763***				
	(0.0206)	(0.0186)				
50-55	0.0700***	0.106***				
	(0.0252)	(0.0227)				
55-60	0.0897***	0.129***				
	(0.0307)	(0.0260)				
60-65	0.131***	0.192***				
	(0.0382)	(0.0339)				
65-	0.114**	0.251***				
	(0.0494)	(0.0438)				
At least some primary	-0.0101	0.0265				
	(0.0274)	(0.0220)				
Completed primary	0.0139	0.0709*				
	(0.0405)	(0.0385)				
Completed Jr. High	0.0494	0.0751				
	(0.0569)	(0.0563)				
Completed High & more	0.0943	0.118				
	(0.0740)	(0.0736)				
Yrs of edu * lagged age	-0.000101	-0.000219**	-0.000335	-0.00137	-0.00124	-0.00153
	(0.0000996)	(0.000105)	(0.000862)	(0.000886)	(0.000957)	(0.00101)
Duration	0.00589*	0.00161	0.00837*	-0.00514	0.00646	-0.00395
	(0.00315)	(0.00519)	(0.00489)	(0.00434)	(0.00564)	(0.00486)
Duration * lagged age	0.0000214	-0.0000297	0.0000323	-0.00000338	0.00000413	-0.0000146
	(0.0000248)	(0.0000219)	(0.0000215)	(0.0000194)	(0.0000234)	(0.0000222)
Lagged log (PCE)	0.00662	0.0146	0.0195	-0.00241	0.0103	-0.00261
	(0.00974)	(0.00901)	(0.0148)	(0.0128)	(0.0164)	(0.0160)
Lagged dist. * year dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	2158	2653	2158	2653	2158	2653
F stat on excluded IVs					26.581	29.580
Hansen J statistics					35.148	32.303
(p-value)					(0.166)	(0.307)

Source: IFLS 1997, 2000 and 2007 *** significant at 1%, ** significant at 5%, * significant at 10%

Robust standard errors adjusted for clustering at the community level are reported in parenthesis.

Hypertension: 1 for those whose systolic $\geq$ 140 or diastolic $\geq$ 90

Dummy variables are included to capture missing values for parent's education, GHS, ADL and household's PCE.

In FD-GMM, $hypertension_{t-2}$, respondent's edu, parents' health and edu, and birth cohort are used as IVs.

Table 5: Dynamic Health Demand Function Estimation - Lung Capacity

	OLS		FD-OLS		FD-GMM A	
	Men	Women	Men	Women	Men	Women
Lagged Lung Capacity	0.576***	0.416***	-0.389***	-0.357***	0.110**	0.0134
	(0.0157)	(0.0148)	(0.0217)	(0.0203)	(0.0557)	(0.0307)
45-50 (lagged age)	-3.152	-4.967**				
	(3.731)	(2.312)				
50-55	-9.439**	-6.551**				
	(4.582)	(2.762)				
55-60	-22.53***	-15.77***				
	(5.591)	(3.262)				
60-65	-32.17***	-10.62***				
	(6.709)	(3.983)				
65-	-31.92***	-13.86***				
	(8.930)	(5.298)				
At least some primary	12.76***	6.454**				
	(4.780)	(2.647)				
Completed primary	20.20***	10.72**				
	(7.199)	(4.555)				
Completed Jr. High	29.26***	15.86**				
	(10.15)	(6.578)				
Completed High & more	49.72***	22.10***				
	(13.24)	(8.521)				
Yrs of edu * lagged age	-0.0296*	0.000797	0.0480	0.160	0.0294	0.0551
	(0.0178)	(0.0124)	(0.149)	(0.102)	(0.164)	(0.100)
Duration	3.377***	1.878***	3.996***	0.991**	3.785***	0.941**
	(0.497)	(0.403)	(0.787)	(0.418)	(0.831)	(0.427)
Duration * lagged age	-0.0154***	-0.0203***	-0.0236***	-0.0185***	-0.0153***	-0.0174***
	(0.00430)	(0.00268)	(0.00359)	(0.00234)	(0.00392)	(0.00237)
Lagged log (PCE)	4.069**	2.256**	1.683	-0.818	3.877	-0.523
	(1.620)	(1.078)	(2.273)	(1.435)	(2.456)	(1.487)
Lagged dist. * year dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	2092	2478	2092	2478	2092	2478
F stat on excluded IVs					16.172	43.346
Hansen J statistics					18.757	20.542
(p-value)					(0.906)	(0.875)

Source: IFLS 1997, 2000 and 2007 *** significant at 1%, ** significant at 5%, * significant at 10%

Robust standard errors adjusted for clustering at the community level are reported in parenthesis.

Dummy variables are included to capture missing values for parent's education, GHS, ADL and household's PCE.

In FD-GMM, *lung capacity*$_{t-2}$, resp's edu, parents' health and edu, and birth cohort are used as IVs.

Table 6.1: Dynamic Health Demand Function Estimation - Overweight

	OLS		FD-OLS		FD-GMM	
	Men	Women	Men	Women	Men	Women
Lagged Overweight	0.723***	0.731***	-0.446***	-0.353***	0.0453	0.249***
	(0.0167)	(0.0107)	(0.0347)	(0.0228)	(0.0877)	(0.0653)
45-50 (lagged age)	0.00757	-0.0172				
	(0.0107)	(0.0122)				
50-55	0.0127	-0.0204				
	(0.0128)	(0.0138)				
55-60	-0.00143	-0.0383**				
	(0.0155)	(0.0170)				
60-65	0.0116	-0.00970				
	(0.0190)	(0.0207)				
65-	0.0396*	-0.0276				
	(0.0239)	(0.0288)				
At least some primary	0.00108	0.0111				
	(0.0119)	(0.0130)				
Completed primary	0.0165	0.0250				
	(0.0206)	(0.0237)				
Completed Jr. High	0.0239	-0.00560				
	(0.0318)	(0.0356)				
Completed High & more	0.0511	0.0172				
	(0.0400)	(0.0449)				
Yrs of edu * lagged age	0.0000185	0.0000633	0.000504	0.000172	0.000236	-0.000154
	(0.0000583)	(0.0000659)	(0.000329)	(0.000381)	(0.000267)	(0.000370)
Duration	0.000752	-0.00569	0.00292	-0.00442	0.00290	-0.00179
	(0.00406)	(0.00413)	(0.00321)	(0.00321)	(0.00331)	(0.00495)
Duration * lagged age	-0.0000375***	-0.0000352**	-0.0000298**	-0.0000390***	-0.0000210	-0.0000313**
	(0.0000141)	(0.0000158)	(0.0000125)	(0.0000139)	(0.0000144)	(0.0000154)
Lagged log (PCE)	0.0100**	0.0138**	0.00348	0.0108*	0.00402	0.00222
	(0.00497)	(0.00561)	(0.00573)	(0.00602)	(0.00732)	(0.00714)
Lagged dist. * year dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	1821	2291	1821	2291	1821	2291
F stat on excluded IVs					18.134	15.576
Hansen J statistics					87.74	96.33
(p-value)					(0.235)	(0.150)

Source: IFLS 1993, 1997, 2000 and 2007 *** significant at 1%, ** significant at 5%, * significant at 10%

Robust standard errors adjusted for clustering at the community level are reported in parenthesis.

over BMI = 1 for those whose BMI $\geq$ 25

Dummy variables are included to capture missing values for parent's education, GHS, ADL and household's PCE.

In FD-GMM, $OverBMI_{t-2,\ t-3}$, respondent's edu, parents' health and edu, and birth cohort are used as IVs.

Table 6.2: Dynamic Health Demand Function Estimation - Underweight

	OLS		FD-OLS		FD-GMM	
	Men	Women	Men	Women	Men	Women
Lagged Underweight	0.643***	0.651***	-0.418***	-0.420***	0.181***	0.144**
	(0.0157)	(0.0139)	(0.0271)	(0.0247)	(0.0692)	(0.0594)
45-50 (lagged age)	0.0175	0.0210**				
	(0.0109)	(0.00978)				
50-55	0.0245	0.0382***				
	(0.0154)	(0.0134)				
55-60	0.0379*	0.0489***				
	(0.0200)	(0.0181)				
60-65	0.0351	0.0867***				
	(0.0258)	(0.0237)				
65-	0.0514	0.0715**				
	(0.0337)	(0.0304)				
At least some primary	0.00424	-0.00631				
	(0.0172)	(0.0127)				
Completed primary	-0.00163	-0.0154				
	(0.0226)	(0.0208)				
Completed Jr. High	0.00587	-0.00341				
	(0.0295)	(0.0279)				
Completed High & more	-0.0104	-0.0100				
	(0.0383)	(0.0378)				
Yrs of edu * lagged age	-0.0000596	-0.0000525	-0.000906***	-0.000624**	-0.000256	-0.000494*
	(0.0000534)	(0.0000578)	(0.000326)	(0.000286)	(0.000276)	(0.000253)
Duration	0.00422	-0.00212	0.000284	0.00267	0.00308	0.000756
	(0.00410)	(0.00348)	(0.00329)	(0.00260)	(0.00768)	(0.00111)
Duration * lagged age	0.0000202	-0.00000310	-0.00000916	-0.0000113	-0.00000372	-0.00000844
	(0.0000186)	(0.0000171)	(0.0000160)	(0.0000144)	(0.0000162)	(0.0000121)
Lagged log (PCE)	-0.0135**	-0.00540	-0.00314	0.0121**	-0.00174	0.00758
	(0.00580)	(0.00534)	(0.00727)	(0.00567)	(0.00814)	(0.00877)
Lagged dist. * year dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	1821	2291	1821	2291	1821	2291
F stat on excluded IVs					14.705	19.29
Hansen J statistics					59.12	72.96
(p-value)					(0.954)	(0.777)

Source: IFLS 1993, 1997, 2000 and 2007 *** significant at 1%, ** significant at 5%, * significant at 10%

Robust standard errors adjusted for clustering at the community level are reported in parenthesis.

under BMI = 1 for those whose BMI < 18.5

Dummy variables are included to capture missing values for parent's education, GHS, ADL and household's PCE.

In FD-GMM, $UnderBMI_{t-2,\ t-3}$, respondent's edu, parents' health and edu, and birth cohort are used as IVs.

Table 6.3: Dynamic Health Demand Function Estimation - BMI Level

	OLS		FD-OLS		FD-GMM	
	Men	Women	Men	Women	Men	Women
Lagged BMI level	0.744***	0.553***	-0.0147	-0.436**	-0.0574	0.00389
	(0.284)	(0.191)	(0.0574)	(0.222)	(0.117)	(0.250)
45-50 (lagged age)	-2.351	-0.0900				
	(2.206)	(0.964)				
50-55	-6.015	0.615				
	(5.704)	(2.346)				
55-60	-10.62	-0.0732				
	(9.458)	(2.163)				
60-65	-14.17	-3.536				
	(13.07)	(2.398)				
65-	-17.24	-0.224				
	(15.72)	(2.224)				
At least some primary	-2.083	0.436				
	(1.305)	(0.667)				
Completed primary	-2.438	0.703				
	(1.695)	(3.294)				
Completed Jr. High	-5.110	-3.194				
	(4.058)	(3.739)				
Completed High & more	-6.788	-3.668				
	(6.170)	(4.776)				
Yrs of edu * lagged age	0.0118	0.00180	0.0277	-0.0679	0.00650**	-0.00406
	(0.00997)	(0.00699)	(0.0245)	(0.0446)	(0.00315)	(0.00895)
Duration	-0.440	-0.428	-0.0490	-0.372	0.0442	-0.241
	(0.294)	(0.281)	(0.109)	(0.531)	(0.163)	(0.562)
Duration * lagged age	0.0120	0.000498	0.00774	0.00118	0.000464	0.000522
	(0.0116)	(0.00169)	(0.00738)	(0.00257)	(0.000436)	(0.00148)
Lagged log (PCE)	0.266	0.171	0.0684	-0.287	0.00413	0.209
	(0.237)	(0.495)	(0.287)	(1.087)	(0.154)	(0.448)
Lagged dist. * year dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	1821	2291	1821	2291	1821	2291
F stat on excluded IVs					3.978	4.404
Hansen J statistics					39.81	41.04
(p-value)					(0.999)	(0.999)

Source: IFLS 1993, 1997, 2000 and 2007 *** significant at 1%, ** significant at 5%, * significant at 10%
Robust standard errors adjusted for clustering at the community level are reported in parenthesis.
Dummy variables are included to capture missing values for parent's education, GHS, ADL and household's PCE.
In FD-GMM, $BMI_{t-2,\ t-3}$, resp's edu, parents' health and edu, and birth cohort are used as IVs.

Table 7: Dynamic Health Demand Function Estimation - GHS (Poor Health)

	OLS		FD-OLS		FD-GMM	
	Men	Women	Men	Women	Men	Women
Lagged GHS	0.232***	0.155***	-0.502***	-0.533***	0.0249	-0.0564
	(0.0173)	(0.0140)	(0.0191)	(0.0152)	(0.0731)	(0.0530)
45-50 (lagged age)	-0.0141	-0.0221*				
	(0.0126)	(0.0134)				
50-55	-0.00771	0.0162				
	(0.0165)	(0.0167)				
55-60	-0.000671	0.0264				
	(0.0213)	(0.0212)				
60-65	0.0182	0.0525*				
	(0.0270)	(0.0271)				
65-	-0.00682	0.0799**				
	(0.0359)	(0.0362)				
At least some primary	-0.000229	0.0431***				
	(0.0172)	(0.0160)				
Completed primary	0.0266	0.0546**				
	(0.0249)	(0.0276)				
Completed Jr. High	0.0193	0.0502				
	(0.0336)	(0.0402)				
Completed High & more	0.0547	0.0470				
	(0.0443)	(0.0534)				
Yrs of edu * lagged age	-0.000155**	-0.000162**	0.0000398	-0.000469	-0.000955	-0.00232*
	(0.0000609)	(0.0000791)	(0.000387)	(0.000441)	(0.00123)	(0.00126)
Duration	-0.0110**	-0.00499	-0.00547	-0.00528	-0.0171	-0.0175
	(0.00445)	(0.00465)	(0.00379)	(0.00409)	(0.0217)	(0.0156)
Duration * lagged age	0.0000955***	0.0000255	0.0000844***	0.0000187	0.0000598	-0.0000393
	(0.0000198)	(0.0000195)	(0.0000185)	(0.0000189)	(0.0000606)	(0.0000362)
Lagged log (PCE)	0.000418	0.0108	0.00283	0.0119	0.00818	0.0132
	(0.00651)	(0.00664)	(0.00851)	(0.00859)	(0.0271)	(0.0226)
Lagged dist. * year dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	2351	2768	2351	2768	2351	2768
F stat on excluded IVs					41.929	60.294
Hansen J statistics					28.740	27.840
(p-value)					(0.941)	(0.979)

Source: IFLS 1993, 1997, 2000 and 2007 *** significant at 1%, ** significant at 5%, * significant at 10%

Robust standard errors adjusted for clustering at the community level are reported in parenthesis.

GHS (Poor Health) = 1 for those who evaluated own health as *somewhat unhealthy, unhealthy*

Dummy variables are included to capture missing values for parent's education, GHS, ADL and household's PCE.

In FD-GMM, $GHS_{t-2,\ t-3}$, respondent's edu, parents' health and edu, and birth cohort are used as IVs.

Table 8: Dynamic Health Demand Function Estimation - Number of Difficulties with ADLs

	OLS		FD-OLS		FD-GMM	
	Men	Women	Men	Women	Men	Women
Lagged ADLs	0.352***	0.276***	-0.382***	-0.457***	0.0664	-0.0303
	(0.0312)	(0.0184)	(0.0336)	(0.0214)	(0.0408)	(0.0287)
45-50 (lagged age)	-0.147***	-0.0872*				
	(0.0373)	(0.0450)				
50-55	-0.261***	-0.00526				
	(0.0602)	(0.0645)				
55-60	-0.225***	0.118				
	(0.0855)	(0.0917)				
60-65	-0.141	0.341***				
	(0.117)	(0.121)				
65-	-0.158	0.688***				
	(0.146)	(0.163)				
At least some primary	-0.0535	-0.0108				
	(0.0668)	(0.0653)				
Completed primary	0.106	0.000175				
	(0.0983)	(0.114)				
Completed Jr. High	0.117	0.0612				
	(0.126)	(0.162)				
Completed High & more	0.272	-0.128				
	(0.170)	(0.216)				
Yrs of edu * lagged age	-0.000665***	0.0000521	-0.00335**	-0.00252	0.000433	-0.0000955
	(0.000249)	(0.000341)	(0.00134)	(0.00157)	(0.00106)	(0.00154)
Duration	-0.0561***	-0.120***	-0.0564***	-0.0729***	-0.0489***	-0.0912***
	(0.0181)	(0.0187)	(0.0151)	(0.0143)	(0.0148)	(0.0207)
Duration * lagged age	0.000839***	0.000686***	0.000755***	0.000547***	0.000701***	0.000626***
	(0.0000914)	(0.0000907)	(0.0000764)	(0.0000769)	(0.0000788)	(0.0000879)
Lagged log (PCE)	0.0141	0.0337	0.0125	0.00305	0.0143	0.0116
	(0.0227)	(0.0247)	(0.0290)	(0.0310)	(0.0330)	(0.0358)
Lagged dist. * year dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	2342	2760	2342	2760	2342	2760
F stat on excluded IVs					36.917	50.197
Hansen J statistics					98.97	125.66
(p-value)					(0.179)	(0.008)

Source: IFLS 1993, 1997, 2000 and 2007 *** significant at 1%, ** significant at 5%, * significant at 10%

Robust standard errors adjusted for clustering at the community level are reported in parenthesis.

Dummy variables are included to capture missing values for parent's education, GHS, ADL and household's PCE.

In FD-GMM, $ADL_{t-2,\ t-3}$, respondent's edu, parents' health and edu, and birth cohort are used as IVs.

Table 9.1: Disaggregated Groups 1

	Hemoglobin Level				Low Hemoglobin				Hypertension			
	Men		Women		Men		Women		Men		Women	
	Coeff.	N	Coeff.	N	Coeff.	N	Coeff.	N	Coeff.	N	Coeff.	N
Age in 2007 < 65	0.196***	1371	0 .235***	1721	0.014	1371	0.088**	1721	0.070	1401	0.182***	1752
	(0.075)		(0.055)		(0.04)		(0.041)		(0.054)		(0.057)	
Age in 2007 ≥ 65	0.083	752	-0.139*	887	0.016	752	-0.036	887	0.171*	757	0.072	901
	(0.082)		(0.082)		(0.055)		(0.057)		(0.091)		(0.075)	
Years of education < 9	0.093	1566	0.125***	2224	0.053	1566	0.042	2224	0.088**	1597	0.174***	2264
	(0.064)		(0.046)		(0.039)		(0.036)		(0.052)		(0.048)	
Years of education ≥ 9	0.335**	557	-0.515	384	-0.073	557	-0.394	384	-0.238	561	0.332	389
	(0.142)		(0.821)		(0.061)		(1.187)		(0.151)		(0.669)	
Lagged log(PCE) < median	0.138*	994	0.101*	1232	0.062	994	0.032	1232	0.058	1012	0.079	1237
	(0.078)		(0.055)		(0.046)		(0.046)		(0.065)		(0.059)	
Lagged log(PCE) ≥ median	0.129	1129	0.018	1376	-0.014	1129	0.039	1376	0.060	1146	0.181***	1416
	(0.091)		(0.068)		(0.046)		(0.045)		(0.064)		(0.062)	

Table 9.2: Disaggregated Groups 2

	BMI Level				Overweight				Underweight			
	Men		Women		Men		Women		Men		Women	
	Coeff.	N	Coeff.	N	Coeff.	N	Coeff.	N	Coeff.	N	Coeff.	N
Age in 2007 < 65	-0.042	1166	0.136	1522	-0.094	1166	0 .232***	1522	0.151**	1166	0.124	1522
	(0.035)		(0.314)		(0.068)		(0.08)		(0.075)		(0.08)	
Age in 2007 ≥ 65	-0.091	655	-0.318	769	0.019	655	0.018	769	0.132	655	0.098	769
	(0.284)		(0.317)		(0.133)		(0.137)		(0.118)		(0.112)	
Years of education < 9	-0.074	1357	-0.013	1951	0.149	1357	0.173**	1951	0.149*	1357	0.145**	1951
	(0.13)		(0.219)		(0.138)		(0.07)		(0.077)		(0.063)	
Years of education ≥ 9	0.082	464	0.167	340	-0.026	464	-0.205	340	-0.091	464	-1.007	340
	(0.256)		(0.826)		(0.108)		(0.319)		(0.191)		(0.642)	
Lagged log(PCE) < median	-0.162*	827	-0.331	1047	-0.064	827	0.157*	1047	0.099	827	0.074	1047
	(0.093)		(0.655)		(0.1)		(0.083)		(0.091)		(0.075)	
Lagged log(PCE) ≥ median	-0.220	994	0.244	1244	-0.009	994	0.170*	1244	0.143*	994	0.096	1244
	(0.337)		(0.323)		(0.069)		(0.092)		(0.078)		(0.081)	

Table 9.3: Disaggregated Groups 3

	GHS (Poor Health)				ADLs				Lung Capacity			
	Men		Women		Men		Women		Men		Women	
	Coeff.	N	Coeff.	N	Coeff.	N	Coeff.	N	Coeff.	N	Coeff.	N
Age in 2007 < 65	-0.017	1455	-0.104*	1719	-0.019	1451	-0.051*	1716	0.073	1387	-0.00057	1715
	(0.07)		(0.063)		(0.029)		(0.03)		(0.07)		(0.038)	
Age in 2007 ≥ 65	0.023	896	-0.011	1049	0.051	891	-0.065	1044	0.098	705	0.063	763
	(0.245)		(0.091)		(0.102)		-(0.075)		(0.094)		(0.064)	
Years of education < 9	0.104	1754	0.016	2355	0.062	1746	-0.022	2347	0.093	1536	0.011	2097
	(0.067)		(0.696)		(0.044)		(0.031)		(0.064)		(0.035)	
Years of education ≥ 9	0.003	597	-0.797	413	-0.237**	596	-0.323	413	-0.039	556	0.264	381
	(0.303)		(28.856)		(0.102)		(0.579)		(0.121)		(0.621)	
Lagged log(PCE) < median	0.007	1069	-0.014	1269	0.048	1064	-0.096***	1264	0.105	978	0.010	1160
	(0.083)		(0.048)		(0.045)		(0.034)		(0.083)		(0.042)	
Lagged log(PCE) ≥ median	0 .030	1244	-0.111	1499	-0.032	1278	-0.024	1496	-0.019	1114	-0.005	1318
	(0.078)		(0.075)		(0.066)		(0.047)		(0.096)		(0.051)	

Appendix

Table A.1: FD-GMM: In the Presence of Time-Varying Measurement Errors

	Overweight		Underweight		BMI level	
	Men	Women	Men	Women	Men	Women
Lagged health measure	-0.258	-0.0654	-0.208	-0.196	-0.525*	-0.504
	(0.187)	(0.128)	(0.253)	(0.335)	(0.296)	(3.713)
Yrs of edu * lagged age	0.000827	0.000400	-0.000310	-0.000344	-0.00000606	-0.0242
	(0.000625)	(0.000827)	(0.000457)	(0.000489)	(0.0226)	(0.0925)
Duration	-0.000216	0.00125	-0.000159	0.00284	0.00188	-0.0791
	(0.00573)	(0.00413)	(0.00399)	(0.00298)	(0.327)	(0.569)
Duration * lagged age	-0.0000188	-0.0000311*	0.000000645	-0.00000852	-0.000325	-0.000308
	(0.0000134)	(0.0000161)	(0.0000154)	(0.0000163)	(0.000734)	(0.00220)
Lagged log (PCE)	0.00534	0.0169*	-0.0135	0.0114	0.0561	0.437
	(0.00967)	(0.00913)	(0.0116)	(0.0104)	(0.435)	(1.221)
Lagged dist. * year dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	1821	2291	1821	2291	2351	2768
F stat on excluded IVs	1.443	1.52	1.27	0.77	1.231	0.746
Hansen J statistics	50.07	66.55	48.14	57.54	15.21	6.83
(p-value)	(0.431)	(0.084)	(0.508)	(0.278)	(0.999)	(0.999)

Table A.2: FD-GMM: In the Presence of Time-Varying Measurement Errors

	General Health Status (Poor Health)		Number of difficulties with ADLs	
	Men	Women	Men	Women
Lagged health measure	-0.390	-0.471**	0.0989	-0.458*
	(0.315)	(0.211)	(0.376)	(0.278)
Yrs of edu * lagged age	-0.000735	0.000144	-0.00293	0.00297
	(0.00163)	(0.00150)	(0.00308)	(0.00364)
Duration	-0.0157	-0.00101	-0.0645**	-0.0786***
	(0.0146)	(0.0132)	(0.0263)	(0.0256)
Duration * lagged age	0.000112**	0.0000385	0.000706***	0.000639***
	(0.0000439)	(0.0000338)	(0.0000812)	(0.0000853)
Lagged log (PCE)	0.0147	0.0128	-0.0225	-0.0416
	(0.0187)	(0.0235)	(0.0564)	(0.0510)
Lagged dist. * year dummies	Yes	Yes	Yes	Yes
N	2351	2768	2342	2760
F stat on excluded IVs	1.028	1.578	0.813	0.756
Hansen J statistics	31.45	32.28	50.77	97.15
(p-value)	(0.831)	(0.833)	(0.707)	(0.001)

Table B: Determinants of Sample Attrition

Attrition = 1	Men	Women	Men	Women	Men	Women	Men	Women
Δ in BMI	-0.000301	0.00108					-0.000314	0.00108
	(0.000286)	(0.000948)					(0.000294)	(0.000948)
Δ in ADL			-0.000131	-0.000147			0.000369	-0.000103
			(0.000417)	(0.000739)			(0.000377)	(0.000670)
Δ in GHS					-0.00126	-0.000388	-0.00650*	0.00119
					(0.00375)	(0.00431)	(0.00350)	(0.00350)
Age in 1993	0.0000572	-0.000272*	0.0000420	-0.000340**	-0.0000536	-0.000198	0.0000455	-0.000267*
	(0.000171)	(0.000151)	(0.000183)	(0.000159)	(0.000178)	(0.000164)	(0.000168)	(0.000156)
Yrs of education	0.000151	-0.0000756	-0.000818	-0.00140	-0.000846	-0.00131	0.000149	-0.0000714
	(0.000717)	(0.000635)	(0.000888)	(0.00102)	(0.000865)	(0.00102)	(0.000717)	(0.000634)
Log (PCE)	0.00418	0.00285	0.00719**	0.00544**	0.00650**	0.00613**	0.00410	0.00287
	(0.00304)	(0.00178)	(0.00324)	(0.00244)	(0.00301)	(0.00245)	(0.00305)	(0.00180)
Rural	0.00270	-0.0127*	-0.00458	-0.0152**	-0.00666	-0.0157**	0.00279	-0.0126*
	(0.00732)	(0.00648)	(0.00747)	(0.00669)	(0.00750)	(0.00652)	(0.00734)	(0.00645)
Father's education								
At least some primary	0.00341	-0.00890	0.000199	-0.0120	-0.000298	-0.0114	0.00350	-0.00897
	(0.00451)	(0.0102)	(0.00480)	(0.0106)	(0.00464)	(0.0105)	(0.00453)	(0.0102)
Completed primary	0.0101	0.00452	0.00473	0.00410	0.00566	0.00424	0.00982	0.00451
	(0.00622)	(0.00624)	(0.00639)	(0.00679)	(0.00624)	(0.00731)	(0.00613)	(0.00625)
Jr. high & more	-0.00265	0.0178	0.00900	0.0283	0.0207	0.0187	-0.00275	0.0178
	(0.00539)	(0.0152)	(0.0112)	(0.0179)	(0.0152)	(0.0204)	(0.00536)	(0.0152)
Mother's education								
At least some primary	-0.00259	-0.00565	-0.00449	0.00599	-0.00605	0.00671	-0.00289	-0.00561
	(0.00746)	(0.0102)	(0.00687)	(0.0132)	(0.00661)	(0.0128)	(0.00750)	(0.0103)
Completed primary	-0.00773	-0.00853	-0.00323	-0.000782	-0.00663	-0.000863	-0.00796	-0.00856
	(0.00712)	(0.00716)	(0.00707)	(0.00762)	(0.00727)	(0.00780)	(0.00717)	(0.00717)
Jr. high & more	0.0227	-0.0267**	0.0535	-0.0274**	0.0424	-0.00332	0.0224	-0.0268**
	(0.0243)	(0.0123)	(0.0350)	(0.0139)	(0.0353)	(0.0270)		
Parent's health								
Father's GHS	0.00816*	0.000672	0.00920*	0.00204	0.00874*	0.00305	0.00794*	0.000658
	(0.00483)	(0.00484)	(0.00532)	(0.00502)	(0.00496)	(0.00502)	(0.00477)	(0.00478)
Father's ADL	-0.00217	-0.00422	-0.00845	-0.00533	-0.00699	-0.00544	-0.00205	-0.00425
	(0.00750)	(0.00624)	(0.00818)	(0.00614)	(0.00778)	(0.00596)	(0.00749)	(0.00625)
Father's death	-0.00498	0.00802**	-0.00318	0.00880**	-0.00301	0.00731**	-0.00489	0.00797**
	(0.00971)	(0.00324)	(0.00957)	(0.00362)	(0.00915)	(0.00359)	(0.00972)	(0.00331)
Mother's GHS	-0.00128	-0.00537	0.000257	-0.00720	0.00149	-0.00619	-0.00128	-0.00542
	(0.00433)	(0.00463)	(0.00449)	(0.00470)	(0.00432)	(0.00466)	(0.00432)	(0.00462)
Mother's ADL	0.00503	0.00682	0.0109	0.00647	0.0104	0.00497	0.00518	0.00690
	(0.00648)	(0.00593)	(0.00751)	(0.00580)	(0.00735)	(0.00577)	(0.00650)	(0.00598)
Mother's death	0.00696	0.00821*	0.00579	0.00430	0.00724	0.00430	0.00713	0.00820*
	(0.00548)	(0.00426)	(0.00596)	(0.00469)	(0.00574)	(0.00458)	(0.00551)	(0.00426)
Residence place in 1993 (dist.)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	2705	3089	3039	3321	3223	3460	2705	3089
F-test for health shocks (p-value)							(0.2950)	(0.6709)

Source: IFLS 1993,1997

Δ in BMI = $BMI_{1997} - BMI_{1993}$

Δ in ADLs = $ADLs_{1997} - ADLs_{1993}$

Δ in GHS = $GHS_{1997} - GHS_{1993}$

*** significant at 1%, ** significant at 5%, * significant at 10%

Robust standard errors are reported in parenthesis.

Attrition =1 if the individual are dropped out of the sample, and 0 otherwise.

Table C.1: First-stage Regression for Low Hemoglobin Reported in Column 5-6, Table 3.1

		Men		Women	
		Coef.	Std. Err.	Coef.	Std. Err.
		Excluded IVs			
	$LowHB_{t-2}$	-0.8029***	(0.023)	-0.7517***	(0.0203)
Birth Cohorts	55 - 60	-0.0042	(0.035)	-0.0370	(0.0359)
	60 - 65	0.0606	(0.0592)	-0.0368	(0.0572)
	65 - 70	0.0527	(0.0815)	-0.0238	(0.0771)
	70 - 75	0.0393	(0.107)	-0.0214	(0.1018)
	75 -	0.0890	(0.1457)	-0.0754	(0.1381)
Resp's Edu	At least some primary	0.0210	(0.0384)	-0.0203	(0.0344)
	Completed primary	0.0047	(0.0562)	-0.1137*	(0.0647)
	Completed Jr. High	-0.0212	(0.0792)	-0.1555	(0.0955)
	Completed High & more	-0.0965	(0.1063)	-0.2277*	(0.1295)
Mother's Edu	At least some primary	0.0289	(0.0442)	0.0492	(0.0435)
	Completed primary	-0.0042	(0.0329)	-0.0055	(0.0327)
	Jr. High & more	0.0793	(0.0704)	-0.0586	(0.0823)
Father's Edu	At least some primary	0.0119	(0.0391)	-0.0820*	(0.0421)
	Completed primary	-0.0215	(0.0294)	-0.0298	(0.0295)
	Jr. High & more	-0.0481	(0.0537)	-0.0192	(0.0604)
Parent's health	Mother Death	0.0023	(0.0268)	-0.0184	(0.0267)
	Father Death	0.0527	(0.0384)	0.0003	(0.0406)
	Mother GHS	0.0192	(0.0252)	-0.0247	(0.0252)
	Father GHS	-0.0303	(0.0241)	0.0209	(0.0258)
	Mother ADL	0.0140	(0.0283)	-0.0086	(0.0277)
	Father ADL	0.0053	(0.0296)	0.0121	(0.0289)
		Included IVs			
	Δ Education * lagged age	0.0008	(0.0025)	0.0047	(0.0032)
	Δ Duration	0.00001	(0.0065)	-0.00495	(0.0058)
	Δ Duration * lagged age	0.00007	(0.0001)	0.00010	(0.0001)
	Δ Lagged log (PCE)	0.0046	(0.014)	0.0273**	(0.0132)
Lagged dist. * year dummies		Yes		Yes	
F statistics on excluded IVs		49.633		49.555	
(p-value)		(0.0000)		(0.0000)	

Table C.2: First-stage Regression for Hypertension Reported in Column 5-6, Table 4

		Men		Women	
		Coef.	Std. Err.	Coef.	Std. Err.
		Excluded IVs			
	$hypertension_{t-2}$	-0.5495***	(0.0216)	-0.5190***	(0.0186)
Birth Cohorts	55 - 60	0.0576	(0.0355)	0.0260	(0.0312)
	60 - 65	0.0297	(0.0574)	0.0393	(0.0505)
	65 - 70	0.0864	(0.0794)	0.0068	(0.0675)
	70 - 75	0.0780	(0.105)	0.0350	(0.0889)
	75 -	0.0360	(0.1401)	0.0761	(0.1226)
Resp's Edu	At least some primary	-0.0225	(0.0381)	0.0605**	(0.0306)
	Completed primary	0.0217	(0.0581)	0.1203**	(0.0555)
	Completed Jr. High	0.0485	(0.084)	0.1920**	(0.0833)
	Completed High & more	0.1824	(0.1134)	0.2276**	(0.1134)
Mother's Edu	At least some primary	0.0111	(0.0391)	0.0039	(0.039)
	Completed primary	0.0249	(0.0321)	0.0091	(0.0306)
	Jr. High & more	-0.1649**	(0.0779)	0.0204	(0.0813)
Father's Edu	At least some primary	-0.0914**	(0.0363)	-0.0401	(0.0382)
	Completed primary	-0.0417	(0.0294)	-0.0339	(0.0266)
	Jr. High & more	-0.0127	(0.054)	-0.0568	(0.0543)
Parent's health	Mother Death	-0.0230	(0.0278)	0.0255	(0.0228)
	Father Death	-0.0028	(0.0396)	0.0844**	(0.0372)
	Mother GHS	0.0468*	(0.0256)	0.0111	(0.0218)
	Father GHS	-0.0051	(0.025)	-0.0451**	(0.0221)
	Mother ADL	-0.0216	(0.0293)	0.0177	(0.025)
	Father ADL	-0.0002	(0.0299)	-0.0075	(0.025)
		Included IVs			
	Δ Education * lagged age	-0.0023	(0.0027)	-0.0053*	(0.0027)
	Δ Duration	-0.00101	(0.0058)	-0.00386	(0.0053)
	Δ Duration * lagged age	0.00006	(0.0001)	0.00007	(0.0001)
	Δ Lagged log(PCE)	0.0052	(0.0141)	-0.0043	(0.0126)
Lagged dist. * year dummies		Yes		Yes	
F statistics on excluded IVs		26.581		29.580	
(p-value)		(0.0000)		(0.0000)	