Ceaseless Toil?
Health and Labour Supply of the Elderly in Rural China

Dwayne Benjamin
Loren Brandt
Jia-Zhueng Fan

Department of Economics
University of Toronto

This Draft: February 28, 2003[1]

Abstract

Deborah Davis-Friedmann (1991) described the "retirement" pattern of the Chinese elderly as "ceaseless toil": lacking sufficient means of support, the elderly had to work their entire lives. In this paper we re-cast the metaphor of ceaseless toil in a labour supply model, where we highlight the role of age and deteriorating health. The empirical focus of our paper is (1) Documenting the labour supply patterns of elderly Chinese; and (2) Estimating the extent to which failing health drives retirement. We exploit the panel dimension of the 1991-93-97 waves of the China Health and Nutrition Survey, confronting a number of econometric issues, especially the possible contamination of age by cohort effects, and the measurement error of health. In the end, it appears that "ceaseless toil" is an accurate depiction of elderly Chinese work patterns, but failing health plays only a small observable role in explaining declining labour supply over the life-cycle.

[1] This draft has benefited from comments by Mark Stabile, participants at the Canadian Health Economics Study Group, Halifax, NS, May 2002, and seminar participants at McGill, Guelph, Toronto, and UC-Berkeley. Benjamin and Brandt gratefully acknowledge the financial support of the SSHRC.

1.0 Introduction

The elderly in China face considerable economic challenges. In the cities, restructuring of State Owned Enterprises (SOE's) has led to massive layoffs, especially in the form of "early retirement." Compounding difficulties for the retirees, SOE insolvency often implies effective default on their pensions and health insurance coverage. Smaller families from strictly enforced fertility restrictions mean there are fewer children to offer support. Besides, the children are as likely to be unemployed themselves. In the countryside, the situation is no better. Per capita income levels for the elderly are lower to begin with. While not as severe, fertility restrictions in rural areas also reduced family sizes, and even fewer children remain in the villages to take care of their parents. Nor is there is no evidence yet, especially with adverse employment outcomes in the cities, that migrant children's remittances make up for traditional living arrangements-based social security. In rural areas, there are no other sources of support. Not surprisingly, in rural areas, retirement maybe a luxury few can afford. Even under collectivization, however, living standards for the elderly were low, and the elderly worked more than they did before collectivization.[2] Davis-Friedman (1991) characterized this lifetime of work as "ceaseless toil." Our objective is to take her descriptive evidence, and evaluate whether the expression can be given empirical content within a conventional labour supply framework.

Our focus is on quantifying the degree and nature of labor force attachment (working) over the life cycle for men and women. As the image of ceaseless toil suggests, we wish to investigate whether there is evidence that Chinese elderly work until they are no longer physically capable. This entails estimating the role of health in the "retirement" decision. As Davis-Friedman noted, however, the role of health is not independent of economic conditions. It is the underlying lack of resources (wealth or other forms of social security) that necessitates the ceaseless toil. Therefore, we also wish to explore how

[2] Benjamin, Brandt, and Rozelle (2000) discuss the evolution of living standards from the 1930s to the present. They show that the current elderly are relatively worse off than they were in the 1930s. Their results are consistent with the conjectures of other social scientists (primarily in sociology) who have explored this question. One possible reason for the decline in the position of the elderly is the inability for them to have accumulated land, which served the dual role of providing a flow of income, as well as leverage (through inheritance) to enforce inter-generational income transfers.

economic variables – to the limit that we can observe them – interact with health and age in determining labour supply.

While our research is concerned with China, there are parallels between our research agenda and the work of Dora Costa, summarized in Costa (1998).[3] She explores the relative roles that health and income (private pensions and social security) played in the evolution of retirement in the United States over the twentieth century. There is also a large literature on the role of health in labour supply generally, and retirement specifically, in a developed country context.[4] One of the benefits of using Chinese data to estimate linkages between health and labour supply is that health levels are generally lower, and poor health may be a more important factor for physically demanding labour, like farm work. Also, as we shall see, Chinese farmers withdraw from work more gradually, without the complications of social security program parameters, and this affords a better opportunity to observe continuous adjustments of labour supply with respect to health. There are very few studies that look at aging or retirement issues in developing countries, especially in a rural context. Deaton and Paxson (1992) focus on welfare issues pertaining to the elderly, while Mete and Shultz (2002) study urban retirement behaviour in Taiwan. Yet, these issues are very important, especially from a policy perspective. As emphasized in the World Bank (1994) report, "demographic transition" is rapidly increasing the ratio of old to young in developing countries, but few have well-designed old-age security systems in place to meet the possible crunch. At least at the beginning, the elderly will have to fend for themselves, while the near-elderly must prepare for their old age by other means. Understanding the retirement decisions of Chinese elderly thus contributes to the general question of how the elderly support themselves in the absence of government-run social security.

The plan of our paper is as follows. In Section 2 we formalize the notion of "ceaseless toil," casting the work patterns of older Chinese couples in the context of a family labour supply model, and highlighting the ways that health and age may "cause" retirement. In this section we also describe our

[3] See also Costa (1995, 1996, 1997, and 1999), and Costa and Stickel (1997).
[4] See Currie and Madrian (1999), Lumsdaine and Mitchell (1999), and Hurd (1990) for useful summaries of this related literature.

empirical framework. We wish to estimate how much health "explains" observed retirement behavior. In order to do this, we propose to estimate reduced-form labour supply and health age-profiles, and then evaluate the extent to which reductions in health line up with reductions in hours worked. The final part of this argument is the estimate of a "structural parameter" linking health to labour supply. Section 3 describes the CHNS panel sample that we use, and outlines a host of measurement and econometric issues to consider. Section 4 presents the empirical results, beginning with non-parametric explorations of the age profiles. Here, the importance (and potential difficulty) of disentangling age from cohort effects is emphasized. We then report the main results of the paper, including "structural" estimates of the impact of health on labour supply. This requires an instrumental variables procedure designed to address measurement shortcomings of self-reported health. In Section 5, we extend the framework in order to investigate the covariation of the aging and health effects with other economic variables, most notably, household wealth. Section 6 offers some conclusions.

2.0 Modeling Ceaseless Toil

Of course, "ceaseless toil" is nothing more than a metaphor for the tendency of Chinese elderly to work throughout old age, until they are no longer physically capable. The "decision" to choose this pattern of work (like any retirement decision) can readily be incorporated into a labour supply model. As we will see, the metaphor provides no testable implications. However, the labour supply model highlights the economic and other variables that determine the extent of "ceaseless toil." In particular, we focus on the channels by which age and health affect labour supply. We begin with the static model, which yields all of the insight.

2.1 Ceaseless Toil and Labour Supply

A farmer and his wife decide how much to work (and consume). For simplicity, we assume that the separation property holds, that is, that production and consumption decisions are independent. This means that we treat farm profits as exogenous to the labour supply decision, and assume that the farmer's labour productivity can be summarized by market wages.[5] The couple's objective is to maximize household utility:

$$\max_{\ell^M, \ell^F, c} u\left(\ell^M, \ell^F, c; a(h^M, h^F, A^M, A^F, Z)\right) \tag{1}$$

where ℓ^M, ℓ^F are the husband's and wife's non-market time (leisure); c is household goods' consumption; and $a(h^M, h^F, A^M, A^F, Z)$ parameterizes preferences that depend (in general) on the husband's and wife's health (h^M, h^F), their age (A^M, A^F), and other variables, Z.

The family budget constraint is related to health and age in several possible ways:

- Productivity, as reflected in wages, $w^M\left(h^M, A^M, X^M\right), w^F\left(h^F, A^F, X^F\right)$;
- Available time, $T^M(h^M), T^F(h^F)$;
- And "non-labor income," $y\left(A^M, A^F, G\right)$, which includes farm profits, the flow of asset income, and possibly remittances from children;

where X^M, X^F, G are other (exogenous) variables that affect men's and women's productivity, and non-labour income. The budget constraint is therefore:

$$\begin{aligned} &w^M\left(h^M, X^M\right)\ell^M + w^F\left(h^F, X^F\right)\ell^F + pc = \\ &\qquad y\left(A^M, A^F, G\right) + w^M\left(h^M, X^M\right)T^M(h^M) + w^F\left(h^F, X^F\right)T^F(h^F) \end{aligned} \tag{2}$$

and the resulting labour supply functions can be written:

$$L^M = f\begin{bmatrix} w^M\left(h^M, A^M, X^M\right), w^F\left(h^F, A^F, X^F\right), \\ y\left(A^M, A^M, G\right), a\left(h^M, h^F, A^M, A^F, Z\right), T^M\left(h^M\right), T^F(h^F) \end{bmatrix} \tag{3}$$

[5] The separation property unlikely holds in the Chinese context. To begin with, there is no real land rental market. The absence of this market (combined with imperfect labour markets) may artificially tie elderly to their farms, "forcing" them to cultivate when they otherwise would prefer not to. However, the elderly can have their children do the cultivation (implicitly using the land or labour market) and increasingly, markets exist to contract farm labour services to non-family members (i.e., concerns over imperfect farm factor markets are becoming less important).

for men. We can now catalogue the channels by which health affects labour supply. Consider a decrease in a man's health (possibly related to aging). This might affect his (and his wife's) labour supply for a number of reasons:

- *Effect on time endowment*: An adverse health shock may simply reduce the man's available time for work. For example, he might be physically capable of working only four, instead of ten hours per day. In this case, labour supply will be reduced (as in a constrained labour supply model), and there will be a corresponding negative income effect. This adverse income effect will affect optimal consumption of other goods, including the wife's leisure. If the wife's leisure is a normal good, this will lead to an increase in her labour supply.
- *Effect on preferences*: Poor health might increase the "marginal disutility of work," (i.e., affect the marginal rate of substitution between the husband's leisure and other "goods"). This will reduce the husband's labor supply (essentially a substitution effect). Depending on whether the wife's leisure was a substitute or complement for the husband's, her labour supply will increase or decrease. For example, if the wife needs to care for her sick husband, in this simple framework, we can view the husband and wife's non-market time as complementary, and thus her labour supply will decrease with her husband's illness.
- *Effect on own-productivity*: A decrease in productivity – as reflected in a reduction in the husband's wage – will have conventional income and substitution effects, with an ambiguous effect on labour supply. Similarly, the cross-effect on the wife's labour supply is ambiguous, unless the husband and wife's non-market time (leisure) are substitutes, in which case the wife's labour supply will increase in response to her husband's poor health.
- *Health Costs:* The model we sketched does not include the purchase of health care services. However, if the family has to pay for the husband's medical expenses, then we can view this as another adverse income effect, which could (in principle) increase the labour supply of both the husband and wife.
- *Non-labour income:* An adverse health shock may affect non-labour income. For example, a sick farmer may not be able to manage his farm as well, and farm profits will fall. Or, remittances from

relatives may increase in response to illness. In both cases, the health shock will add another income effect.

The main lesson to draw from this theoretical discussion is that adverse health shocks have an ambiguous impact on the labour supply of the husband and wife. Furthermore, there is no obvious way to separate the various possible avenues that health affects labour supply (e.g., separating the effect of health on preferences, productivity, or the time endowment) unless we observe the individual components (like productivity). Nevertheless, the language of income and substitution effects, especially as a consequence of health's effect on productivity (wages), is a useful way to think about ceaseless toil.

Almost all of the above discussion carries over to a discussion of the effect of age on labour supply. For example, we might imagine that labour supply declines in old age because of a systematic decline in productivity: Chinese farmers work on their own farms until their productivity falls below some threshold. But why would Chinese farmers be less likely to retire than the Chinese living in cities, or men in North America? If farm productivity was the main part of the story, then we would have to argue that farm productivity falls more slowly for farmers than university professors or other white collar workers. Alternatively, farm work may be more pleasant than other types of work, so that reservation wages for farm participation are very low. Neither explanation is plausible. More likely, the key variable is "income," or wealth: Chinese farmers have low wealth levels, and thus cannot "afford" to retire. In the context of our model, non-labour income has a different level or trajectory for Chinese farmers than other workers. If they are poor all of their lives, then having a lower level of permanent income means they will have to work more over their entire life-cycle. Or, limited savings mechanisms may prevent farmers from providing for their old-age. Especially if transfers from children are the main returns from "savings", it may take awhile (with imperfect credit markets and low wages for adult children) before elderly workers can "collect" their social security and retire.

Clearly, wealth and productivity may combine to explain the ceaseless nature of work in China as compared to North America. The income effect of permanently lower wages (productivity) may lead to higher lifetime labour supply, while the age-pattern of labour supply tracks the life-cycle trajectory of productivity, including the deterioration in physical strength associated with old age.

2.2 A Simple Labour Supply function

Using (3) as a starting point, a linear version of the husband's labour supply function is given by something like:

$$\begin{aligned} L_{it}^{M} &= g_0 + h_M w_{it}^{M} + h_F w_{it}^{F} + h_y y_{it} \\ &+ g_{1M} A_{it}^{M} + g_{1F} A_{it}^{F} + g_{2M} h_{it}^{M} + g_{2F} h_{it}^{F} + g_4 Z_{it} + q_{it} \end{aligned} \tag{4}$$

where *i* indexes an individual, and *t* indexes time. If all variables were observable and perfectly measured, we would be able to estimate (4), and determine the "pure" effect of age and health, controlling for the economic variables. We would also be able to estimate the effect of age and health on the economic variables (wages and non-labour income), in order to distinguish between the various channels discussed previously. For example, the partial own-productivity effect of health on labour supply would be:

$$h_M \frac{dw_{it}^{M}}{dh_{it}^{M}} \tag{5}$$

In this way, we could decompose the total effect of health and aging on the labour supply decision, and completely categorize the dimensions of "ceaseless toil."

Unfortunately, in a rural developing country, measurement of the economic variables is problematic. Wages are unobserved in self-employment, and estimation of "pure" farm profits is difficult. Wages are frequently unobserved in a developed country also, and one could adopt the strategy of Abowd and Card (1989) and treat them as latent variables that shift earnings and hours according to a structural model implicit in (5). For example, with enough structure one can specify a model linking health (and age) to earnings and hours, and thus back-out the implicit impact of age on both productivity and hours. This is the strategy adopted by Laszlo (2002) in estimating the channels by which household education affects household earnings through a labour supply model. Unfortunately, we cannot pursue this strategy because we want to estimate the impact of *individual* health on individual labour supply, but only *household* income is observed. It is virtually impossible to identify the productivity effects in this case, once one admits to the impact of health on other family members' labour supply.

Instead, our objective will be to estimate a "reduced form" version of (4). With this exercise (difficult enough itself), we will be able to estimate the total effect of age and health on labour supply, while being unable to decompose the sub-components of these effects. Substituting out the economic variables yields a reduced form:

$$\begin{aligned} L_{it}^{M} = b_0 + b_{1M} A_{it}^{M} + b_{1F} A_{it}^{F} + b_{2M} h_{it}^{M} + b_{2F} h_{it}^{F} + \\ + b_{3M} X_{it}^{M} + b_{3F} X_{it}^{M} + b_4 Z_{it} + b_5 G_{it} + e_{it} \end{aligned} \quad (6)$$

We estimate variations of this equation, with the objective of estimating b , in order to evaluate the extent that we can link health and labour supply over the life-cycle.

2.3 What if Labour Supply Decisions are Made in a Dynamic Framework?

For simplicity, ignore the family dimension to labor supply, and consider the consequences of the individual making his labour supply decision according to:

$$\max_{L_{it}, c_{it}} \sum_{t=0}^{T} (1+r)^{-t} u_t \left(c_{it}, L_{it}; A_{it}, h_{it}, e_{it}\right) \quad (7)$$

subject to:

$$K_{i0} + \sum_{t=0}^{T} (1+r)^{-t} \left(w_{it}\left(A_{it}, h_{it}\right) L_{it} - p_t c_{it}\right) = 0 \quad (8)$$

With appropriate functional form assumptions, we can concoct a labour supply function like[6]:

$$L_{it} = p_0 + p_1 A_{it} + p_2 h_{it} + p_3 w_{it} + p_4 l_{it} + s_{it} \quad (9)$$

where l_{it} is the marginal utility of relaxing the budget constraint (8).

The main innovation in moving from the static to dynamic model in this case is that (i) we no longer take non-labour asset income as exogenous; and (ii) We recognize that an individual's expected

[6] See Card (1994) for more discussion of intertemporal labour supply models, and in particular, the statistical and modeling issues associated with (7) and (8). He also outlines the possible ways in which the life-cycle model can be used to account for the effect of "age" on labour supply over the life-cycle.

deterioration of productivity due to health and age is summarized in l_{it}. In this way, we can readily compare the life-cycle trajectories of Chinese farmers and U.S. college professors, in terms of their life-time wealth (reflected in l_{it}), and their wage-age productivity profiles. We would employ the language of intertemporal labour supply, where the age- and health-productivity relationship drives wages. Chinese farmers have lower lifetime wealth, and so work more over their entire life-cycle if leisure is a normal good. Furthermore, individuals will time their labour supply to exploit times of relatively high productivity, with farmers taking account of their expected deterioration of productivity associated with old age. Note, it may still be difficult to accommodate the different retirement patterns of farmers and professors within this framework, unless we believe college professors' productivity drops sharply at age sixty-five.

Given the unobservability of wages, we can imagine estimating a reduced form equation like:

$$L_{it} = p_0' + p_1' A_{it} + p_2' h_{it} + p_4' l_{it} + s_{it}' \quad (10)$$

There are subtle differences in interpretation of the impact of health on labour supply in this context. Most importantly, the health coefficient, p_2, captures a pure substitution effect, since the income effect due to anticipated health and productivity decline is controlled for by l_{it}. Similarly, if there is a transitory health shock that does not change long run health prospects, then p_2 will be interpreted as a substitution effect. Even in this framework, however, an unexpected adverse change in health will have p_2 convolute income and substitution effects. Furthermore, there will be a possible statistical complication caused by the correlation of l_{it} and h_{it}, especially as l_{it} is itself unobserved. If those with higher wealth (and lower l_{it}) also have better health, the failure to directly control for l_{it} will generate omitted variables bias. In this case, the negative correlation between l_{it} and h_{it} will impart a negative bias – that is, if p_2 is truly positive, the estimated health effect will be biased towards zero, or the wrong sign. In the dynamic labour supply literature, this is the primary motivation for estimating the model with fixed effects or first differences. This is one reason (among others) that there is a potential gain to using panel data as opposed to cross-section data in the estimation of (6). Note, however, that the FE estimator will not help (in this case) if the changes in health status are permanent and unanticipated, or lead to changes in l_{it}.

3.0 Empirical Implementation

To summarize, our main objective is to estimate the extent of “retirement” in the Chinese countryside, and furthermore, how much deteriorating health “explains” the age-labour supply patterns. The key equation we will estimate is a variant of (6). Our first objective is to estimate the “pure” effects of age on labour supply and health which can be accomplished by estimating:

$$\begin{aligned} L_{it}^{M} &= b_0 + b_1 A_{it}^{M} + v_{it} \\ h_{it}^{M} &= d_0 + d_1 A_{it}^{M} + u_{it} \end{aligned} \tag{11}$$

If we restrict the aging effect to be linear (which we do not), then the effect on hours of of aging ten years (say, from fifty to sixty) would be $10 \times b_1$. If health also declines with age according to (11), and the effect of age on labour supply is entirely attributable to health, then we can then augment the labour supply equation in (11) by controlling for health:

$$L_{it}^{M} = b_0 + b_1 A_{it}^{M} + b_2 h_{it}^{M} + e_{it} \tag{12}$$

And if health is measured perfectly, it will absorb the effect of age on labour supply, yielding an estimate of $b_1 = 0$. But health is definitely not measured perfectly, and other factors besides health link age to retirement.

As a summary of the impact of health on retirement, we estimate (i) the extent to which health declines with age, d_1, and (ii) the impact of health on labour supply, b_2. Our decomposition exercise is thus based on something like:

$$d_1 b_2 \tag{13}$$

More precisely, we estimate the reduced form effect of age on both labour supply and health:

$$L_{it}^{M} = b_0 + \sum_{j=1}^{J} b_{1j} AGEG(j)_{it}^{M} + v_{it}$$
$$h_{it}^{M} = d_0 + \sum_{j=1}^{J} d_{1j} AGEG(j)_{it}^{M} + u_{it} \qquad (14)$$

where $AGEG(j)$ is an age-group indicator for five-year age groups (20-24, 25-29,… 75-79,80 plus). We focus on two age transitions: (i) The implied change in labour supply or health between fifty and sixty, given by $\Delta_{6050}^{L} = b_{1(60-65)} - b_{1(50-55)}$ and $\Delta_{6050}^{h} = d_{1(60-65)} - d_{1(50-55)}$; and (ii) The implied change in labour supply and health between sixty and seventy ($\Delta_{7060}^{L} = b_{1(70-75)} - b_{1(60-65)}$ and $\Delta_{7060}^{h} = d_{1(70-75)} - d_{1(60-65)}$).

We then estimate the "structural" effect of health on labour supply on the basis of:

$$L_{it}^{M} = b_0 + \sum_{j=1}^{J} b_{1j} AGEG(j)_{it}^{M} + b_2 h_{it}^{M} + e_{it} \qquad (15)$$

and define the part of retirement attributed to declining health (with age) as:

$$b_2 \times \Delta_{6050}^{h}, \quad b_2 \times \Delta_{7060}^{h} \qquad (16)$$

3.1 Data

We use the China Health and Nutrition Survey (CHNS) for 1991, 1993, and 1997.[7] Details of the structure of the data set are provided in the data appendix. Of particular note, however, is that we exploit the panel dimension of the CHNS. This means that we restrict our analysis to those individuals observed for the three surveys. We also include people who we could confirm had died between waves of the survey. While the panel affords several key benefits, we must also confront issues of attrition. We discuss some of the general issues below, and describe the nature of the attrition in the data appendix.

We further restrict our sample to those individuals 20 years of age and older, for whom we have a complete set of health and labour supply variables. More restrictive, since we wish to examine the impact of spousal health on labour supply, we only include those individuals for whom we have spousal

[7] The data and complete documentation are available at the website: http://www.cpc.unc.edu/china/home.html.

information. This means that younger single individuals are excluded, as are women who outlive their husbands (widows).

We now discuss an assortment of econometric and measurement issues that need to be confronted before estimates of (14) and (15) can be presented. Along the way, we refer to Table 1, which presents selected summary statistics for our working sample. As Table 1 shows, we have approximately 1200 men and 1200 women that satisfy the sample selection criteria. Of these, there are 375 men, and 296 women who are fifty years or older in 1991. Note that the smaller number of older women reflects the higher mortality of husbands (prior to 1991), and the exclusion of a slightly disproportionate number of older women on the grounds of missing spousal information.

3.2 Measuring Labour Supply

An important issue in the retirement literature is the definition of "retirement." In some studies, retirement is defined to occur when a public or private pension is initiated, irrespective of labour supply. Another possibility would be to define it as complete withdrawal from the labour force. Given the particular possibility of gradual retirement, especially for farmers, we prefer not to restrict ourselves to defining retirement as a distinct state, but instead look at labour supply (hours of work) more broadly. In addition to the quantity of work, it is also worth exploring the type of work done by the elderly: Do they move away from off-farm employment, increasing their share of work on the farm and in other sideline activities?

Table 1 shows average levels of labour market activity. We define "work" as being engaged in income-generating activities. Most notably, this does not include "housework," or working in a garden for the production of home produced vegetables. It does include wage work (off the farm), commercial gardening, farming, raising animals, fishing, and working in a family enterprise. Average "participation rates" are 92 percent and 93 percent for men and women. As an indicator of "ceaseless toil", the employment rate of older men and women is still high past age fifty, at 82 percent. In terms of hours worked, note that women's hours – not even counting housework – exceed those of men, at 2036 versus

1962 per year. Again, labour supply of the elderly is quite high, with annual hours only declining to 1686 and 1565. The drop in labor supply is small by North American standards, and consistent with a metaphor of ceaseless toil. The drop, while small, is more pronounced for women. In terms of the types of work, the majority of time is spent farming, for men and women of all ages. The one age-related pattern is that the share of hours on the farm is higher for older individuals, as they spend less time working in the off-farm labor market. What we cannot tell from the means table, however, is whether this pattern reflects "aging", as older workers quit their off-farm jobs, or whether it reflects cohort effects, where older workers are less likely to have ever worked at a wage job.

3.3 Measuring Health

How can we tell when someone's health has "objectively" declined? This is obviously crucial in identifying the impact of health on the retirement decision. Our main interest is capturing that part of health that is correlated with age, and enters the labour supply decision, possibly according the model described earlier. The CHNS offers several possible health measures, each with its own well-known (potential) problems: (i) Self-reported health status (SRHS); (ii) Body mass index (BMI); (iii) Limitations on activities of daily living (ADL); (iv) Physical function limitations (PF); and (v) Subsequent death. We outline some of the issues associated with each measure in turn, but a further complicating factor is introduced by our focus on panel individuals. The survey did not ask the same sets of health questions in each wave, and in some instances, the details of the questions changed. The need for continuity and comparability further constrains our ultimate choice of health measure.

Self-Reported Overall Health Status (SRHS)

In a survey, interviewers obtain SRHS by asking the question "Right now, how would you describe your health compared to that of other people of your age." Responses are then coded on a scale of one (excellent) to four or five (poor). SRHS is thus a subjective and relative health measure. The CHNS collected SRHS in each wave, and this is the primary health measure we use.

On the positive side, SRHS may contain private health information that no doctor can measure. In fact, previous evidence shows that SRHS has strong predictive power for subsequent mortality, even controlling for a set of more objective health measures (see Deaton and Paxson, 1998, for example). This pattern also holds in the CHNS (as we will see). In our analysis, we distill the SRHS variable into a binary indicator of "good health", which equals one if the person's health is in the top two categories.

There are a number of potentially serious problems with SRHS. McGarry (2002) and Bound (1991) provide excellent reviews of these problems. First, respondents are supposed to net out any effect of age on health. Measured properly, SRHS should be orthogonal to age, and therefore it should be a terrible way to measure the deterioration of health with age. However, respondents do not do a very good job of controlling for age, and SRHS is strongly related to age (see Deaton and Paxson, 1998, for another example). But the effect of age on health may yet be understated, and must be kept in mind as we discuss our empirical results. We could easily underestimate the effect of health on labor supply (because of measurement error, attenuation bias), and compound this by underestimating the deterioration of health with age.

Second, an individual's subjective sense of health may be endogenous to his labour supply decision. If a man is not working, he may justify this by poor health, in which case we exaggerate the relationship between health and labour supply. But this "justification bias" is only one of the possible reasons why health is endogenous to the labour supply equation. Unobserved heterogeneity of the interpretation of self-reported health may be correlated with the economic variables that determine labour supply, in a static or dynamic framework (See Bound, *et al*, 1999). For example, higher income people could have higher "standards" or benchmarks for good health. For two equally healthy people, one rich and one poor, we may find that the poor one reports being in better health, though he works less (or more). Depending on the correlation of the economic variables with labour supply, we could under- or over-estimate the impact of health on labour supply. The reverse (but equally unidentified) bias would occur if the poorer person was in actually poor health, but we fail to control for the economic variables. Third, SRHS may simply be a noisy indicator of underlying latent health, and we will suffer from conventional attenuation bias. Fourth, the timing of observed health may not line up with the "retirement decision," though this problem applies to other health measures.

A number of strategies exist for addressing these problems. For example, other health measures can be used as instruments if they are correlated with SRHS, and otherwise uncorrelated with labour supply. These other health measures can serve as substitutes for SRHS, or can serve as a means of exploring the robustness of conclusions to SRHS. Previous studies, like Baker, Deri, and Stabile (2002), find that the measurement error concerns outweigh the "justification bias", and their work points to the value of using instrumental variables in this setting. Having panel data allows us to address some of the other shortcomings of SRHS. If the subjective benchmark for health is an individual fixed effect, then fixed-effects (FE) estimation will allow us to sweep away this form of heterogeneity. Also, by observing individuals over time, we may better be able to pin down the timing of health shocks and changes in labour supply. As usual, when using FE estimation, however, we may amplify measurement error, in which case instrumental variables techniques are even more important.

Body Mass Index (BMI)

A person's BMI is defined as his weight in kilograms divided by the square of his height measured in metres. It is a measure of "physical robustness", in the sense that a person with especially low BMI may be frail, while a person with especially high BMI is obese. We would thus expect a non-linear effect of BMI on labour supply or other outcomes, as well potential asymmetry of the effect of being either too light or too heavy. Dora Costa (1996, 1998), for example, shows that a "U-shaped" relationship exists between BMI and a variety of other health outcomes, like the number of chronic conditions, bed days, hospitalizations, and doctors' visits. In explorations with the CHNS, we also find that a "U-shaped" relationship exists between BMI and health outcomes like mortality. In order to account for the non-linearity of the effect of BMI, we create indicators of whether an individual's BMI is extremely high or low, defined by whether it is below the twentieth percentile, or above the eightieth.

While objective, BMI is a far from perfect health measure. First, it may (like any measure of health) be endogenous to labour supply. Higher income individuals, or individuals with higher valued economic characteristics may also have "better" BMI's, because of superior nutrition or health care. This could lead to an overstatement of the relationship between health and labour supply. Biasing the other way, BMI may not move much in response to a change in the kind of heath that affects productivity. For example, blindness or a bad back will not be reflected in one's BMI. BMI does not change very much with aging, at least between ages fifty and seventy. When it does change, it usually reflects a significant underlying change in health status. A dramatic loss in weight may be the result of a serious illness. One must then be careful in interpreting the effect of low BMI – it may have nothing to do with low strength, but instead reflect other aspects of poor health. One additional positive aspect of using BMI as a health measure is that it is commonly recorded in surveys, and this permits comparison of our results with other researchers. For example, BMI is the main health measure that Dora Costa uses in her investigation of the role of health in the evolution of retirement in the U.S. BMI is also recorded in all the waves of the CHNS, so we can use it in our panel procedures.

Activities of Daily Living (ADL)

ADL's seem like a promising way to measure physical limitations. The ADL module of the CHNS questionnaire is applied to individuals over fifty years old, and measures an individuals' abilities to carry out a list of daily activities, like taking a bath, being able to eat and drink, use the bathroom, or dress by themselves. In principle, ADLs offer more objective information about health status than SRHS, with an improvement over BMI in terms of capturing functional limitations. Deteriorations of health reflected in ADLS may closely track the sorts of health change related to labour supply.

But ADL's have their own problems, especially in the context of the CHNS. First, we wish to model the effect of health on labour supply more generally, without restricting our analysis to the elderly. While younger people may have no limitations on ADL, it is a shortcoming to not have this information available for individuals under fifty. Second, ADL's were not recorded in all surveys, and are missing from 1991. ADL's also have their own measurement shortcomings. In particular, they are designed to capture extreme disabilities. Thus, for the significant majority of elderly who are not so decrepit, we have no health information to distinguish between their health status (McClellan, 1998). People with diabetes, for example, may have no problem doing all the daily activities, but may decide to work less or retire earlier than otherwise. While we have used ADL's in some explorations, given the survey limitations, we do not use them in our primary analysis.

Physical Function Limitations (PF)

The CHNS asks a series of questions concerning an individual's health that can also be used, like ADLs, to construct an objective "index" of health. The main difference between ADLs and PFs is that PFs do not measure behavioral abilities as ADLs do, but refer to difficulties on specific physical functions such as hearing, eyesight, arms, legs, etc. While the set of questions varies over survey years, a set of five questions (listed in the appendix) provides information on the state of various bodily functions, including some related to the ability to work.[8]

In order to distill the responses to these five questions into one variable, we use principal components analysis create a single index. We also experimented with other indices, yielding virtually identical results. PF's share many of the same pro's and cons as ADLs as health measures to be employed in labour supply functions. In the context of the CHNS, we only have the PF's for 1991 and 1993, and we cannot use them in our complete panel analysis. However, PF's have the advantage over ADL's of being asked of everyone. We will use the PF's as instruments for the SRHS in order to evaluate the possibilities that SRHS suffers from endogeneity of measurement error bias. We defer a more detailed discussion of this procedure until later.

Subsequent Death (Mortality)

[8] The choice of the grouping together of body functions – like heart, lungs, and stomach – into one category seems somewhat mysterious (and slightly amusing), and it is variation in this dimension that restricts comparability over time.

One of the benefits of a longitudinal survey is that we can follow an individual over time. This means that we can observe outcomes like death that occur subsequent to a particular survey year. If death was the result of poor health (as opposed to an accident), these health outcomes may not have been observable to surveyors, or even the respondent, though the underlying poor health may be reflected in labour supply. Previous researchers have found subsequent mortality a useful "objective" health measure.[9] We also use the CHNS to create an indicator of subsequent death, defined from the perspective of 1991 as whether the individual died prior to either the 1993 or 1997 survey.

Because it is a summary of outcomes over the panel, this measure is only available for the 1991 cross-section. We cannot use this as our primary health measure. However, it serves a useful role in cross-validating our other health measures, as well as seeing how much health information may be "missing" from the other measures.

3.4 Preliminary Explorations with the Health Measures

Table 1 provides descriptive statistics concerning some of these health measures. Taken over all age groups, an average of 74 percent of men, and 72 percent of women report being in good health (H12=1). The proportion is dramatically lower for older people, as only 58 percent of men, and an even lower proportion of women (53 percent) report good health. This suggests that SRHS does decline with age, even though individuals are supposed to net out the age effect. The average BMI does not very seem different between the full sample and the older sub-sample. However, this hides some deterioration in health, as a significantly higher proportion of elderly men and women have low BMI than the young. There is not much difference in the indicence of high BMI. The indices of physical function problems (PFs) are also higher in magnitude (more negative) for older individuals. Finally, the probability of subsequent death is much higher for individuals over fifty. Fully twenty percent of men over fifty in 1991 died by 1997. A much smaller fraction of women died by 1997. Note that this does not purely reflect

[9] See Parsons (1980), Hurd and Boskin (1984), and Anderson and Burkhauser (1985), for example.

differences in mortality between men and women, but is a consequence of our sample selection, which is tilted towards younger women, and those with surviving husbands.

Table 2 reports the results of preliminary cross-section regressions to evaluate the information content of the health measures. In the first panel, we report regressions of the indicator of subsequent death on the health measures available in 1991. First note that H12 (self-reported good health) is a statistically significant predictor of mortality across all specifications. Thus, controlling for age, education, province dummies, and health measures like BMI and PFs, we see (like other researchers) that H12 contains important health information. For men (but not women), we also find that having worse physical functions is significantly related to subsequent death.[10]

The second panel shows the results of a similar regression of labour supply (hours of work) on the health measures. This regression also includes controls for age, education, and province. The possible value of subsequent death as a health measure is apparent. By far, it is the strongest variable, and the poor health it captures is statistically significantly negatively related to labour supply. This provides our first concrete evidence that "health matters" in labour supply. While it is not a perfect indicator of current health, it should not suffer from some of the other biases (like justification bias) described earlier. We also see that H12 (good health) is positively correlated with labour supply, statistically significant for older men. The sign patterns of the other health coefficients also make sense, but they are not statistically significant.

3.5 Isolating Age From Cohort Effects

The "pure" effect of age is not so easy to estimate. Consider our labour supply function:

$$L_{it}^{M} = b_0 + b_1 A_{it}^{M} + b_2 h_{it}^{M} + e_{it} \qquad (17)$$

[10] We scale the index of physical functions so that increases in the index reflect improvements in health, so that the signs of the health effects for PF and H12 should be the same.

The age coefficient will be biased if there are any factors in v_{it} that are correlated with age. In particular, birth cohort or "generational" effects may be important, especially for life-cycle behaviour. In that case, we can write:

$$e_{it} = l_c + w_{it} \tag{18}$$

where l_c represents the fixed labour supply pattern of individuals born in cohort c. Goldin (1990), for example, illustrates how such cohort effects can render traditional cross-section age-participation profiles misleading. The key question is to what extent today's sixty year olds are a good predictor of the labour supply attachment of today's fifty years olds when they turn sixty. Especially in a growing economy with decreasing retirement ages, we expect a cross-section age-profile to underestimate the degree of retirement.

The solution is to follow birth-cohorts over time in order to more accurately trace the effects of age. This can be accomplished by including cohort fixed effects in a pooled time-series cross-section specification. With panel data we can go one step further by including individual fixed-effects. The fixed effects will also serve to absorb individual heterogeneity that may be correlated with age or health status. For example, individual "benchmarks" for subjective health can modeled as fixed effects, in which case the fixed effects specification will allow us to adjust for variation across individuals of their perception of (permanent) health. Furthermore, the fixed effects will absorb some of the otherwise unobservable economic variables, like wealth or long-run productivity, that could also be correlated with health.

In the empirical specifications that follow, we will report both fixed-effects (FE) and random-effects (RE) results. The fixed-effects specifications have the advantage of being robust to the problems just described. On the other hand, the FE results may be further biased because of the amplification of measurement error. Furthermore, the random effects estimates admit cross-cohort variation in health and labour supply, which may provide (with appropriate qualifications) a useful source of identification.

3.6 Attrition

While exploiting the panel has its advantages, there are built in problems because of attrition. By restricting our analysis to those individuals who actually survived the 1991-1997 survey cycle, we could actually bias the age and health coefficients. This could arise if:

$$\operatorname{cov}\left(E[v_{it}, u_{it} \mid survive], A_{it}^{M}\right) \neq 0 \tag{19}$$

Specifically, this happens when only healthy or hard-working people live to old age, in which case, we under-state the relationship between age and deterioration of health, or the reduction of labour supply.

There is very little that we can do to address this attrition bias, beyond documenting the extent of attrition, and being aware of situations (which we will see) where it is likely to be a problem. The data appendix provides the first ingredient, with a table documenting the extent of attrition relevant for the construction of our working sample.

4.0 Results

We now provide the results of the exercise outlined above. The first step is a non-parametric exploration of the "pure" relationship between age and our primary variables of interest (labour supply and health). We then estimate the reduced-form relationship between age, health and labour supply. Finally, we estimate the "structural" relationship between health and labour supply, in order to calculate the extent to which retirement can be explained by declines in health.

4.1 Non-parametric Explorations

Figures 1 through 4 provide non-parametric estimates of the following relationship[11]:

$$y_i = g(Age_i) + f_i \tag{20}$$

Where y_i refers to (i) Hours of work; (ii) Participation (positive hours worked); (iii) Good Health (H12); and (iv) The fraction of hours spent working off the farm. In each figure we estimate the cross-section

[11] We use the Fan (1992) estimator, described in detail (with examples) by Deaton (1997).

age-profile for men and women for survey year 1991. In order to evaluate the extent to which cross-section profiles are an accurate predictor of dynamic (life-cycle) behaviour, we also show estimates of:

$$\Delta y_i = f(Age_i) + j_i \quad (21)$$

In this case, we look at the *ex post* change in hours and health from 1991 to 1997 for each person arrayed by their age in 1991. With this function, we can compare the predictions of the cross-section with the realized change in y_i. If the predictions line up, then cohort-effects are not likely to be a problem, and the "pure" effect of age will be easier to identify.

Figure 1 illustrates "ceaseless toil" more clearly than any other result in this paper. Here, we can see in the top panel that the age-hours profile for men is much flatter than in North America, or other developed countries. The average seventy year old Chinese man works almost 1000 hours per year, which is about half the peak of 2000 hours per year. Average hours begin to decline after age forty, so the only evidence of retirement behaviour is this gradual decline of hours worked. The same pattern holds for women, though "retirement" is more pronounced: seventy-year old women work an average of 500 hours per year, which is approximately one-quarter of their peak labour supply of 2000 hours per year.

The bottom two panels allow us to explore the possible contamination of the cross-section age profile by cohort effects. Here we compare the predictions implied by the cross-section to what actually happened to labour supply between 1991 and 1997. Take fifty year olds as an example. The 1991 cross-section estimate of aging from fifty to sixty suggests that annual hours will drop from about 2000 to 1500 hours, a 500 hour decline. Inaccuracy of this prediction may arise if there are permanent differences in life-time hours between fifty and sixty year olds in 1991. For example, if fifty year olds are richer (from a richer generation), then they may reduce their labour supply more than predicted. These unobserved differences between cohorts may result in different "intercepts" for labour supply, or more likely, different age profiles.

Using the 1991 cross section, we can predict the change in hours associated with six years of aging (from 1991 to 1997). The predicted change in hours is given by the dashed line in the middle panel.

For fifty year olds, we predict a decline of approximately 200 hours per year. As the solid line shows, however, their actual hours dropped by 700! One possibility is that this reflects a significant change in retirement propensities, or a shift to “early retirement.” But a quick glance at the change in hours for other ages suggests this is not the story. In fact, it appears that the actual reduction in hours is a parallel shift downward of the predicted change in hours. The difference of approximately 500 hours is common to all age groups. This can more accurately be described as a “year effect.”

Why did hours decline so much for everyone? We explored a number of possible explanations. Almost all of the decline in total hours is due to reductions in time spent farming. Possibly the survey question is different in 1997 than 1991? However, the question is identical in the surveys. This does not preclude the possibility of different instructions being given to the enumerators. However, it is striking that the decline is so uniform across provinces and age groups. We were also unable to line up the change in hours with observable economic variables, like wages or crop prices. One possibility is that reported hours in agriculture now more accurately conform to “hours worked”, rather than time spent “idle” on the farm. As Benjamin and Brandt (2002) show using a different Chinese survey, there appears to be a great deal of inefficient time spent in “farming” which appears to decline as economic opportunities improve. Whatever the explanation, we have no reason to believe that the drop in hours represents something that substantively affects our interpretation of the impact of age on labour supply. From this point, we will agnostically label this a “year effect”. Note also that we are not interested in disentangling (identifying) the possible cohort and year effects. What this exercise also underlines is the importance of looking at the labour supply of all ages, not just the elderly in isolation, in identifying “retirement.”

Our main concern in this exercise is that the cross-section provides a poor estimate of the impact of age on labour supply. This will be reflected in non-parallel differences between the predicted and actual change in hours. The bottom panel (for men) suggests in fact, that the difference between predicted and actual changes is unrelated to age, in which case cohort effects may not be an issue. In the parametric analysis that follows, we will formally test whether the cross-section (random effects) and fixed-effects estimates are the same, which is essentially a test for the contamination of the age coefficients by cohort effects.

Turning to the similar analysis for women, we expect cohort effects to be a more serious problem. In addition to a possible shift towards early retirement, the changing economic role of women, including that caused by fertility restrictions, might render the cross-section misleading (as was the case in Goldin (1990)). However, similar to the men, the cross-section overpredicts hours worked in 1997, consistent with a year effect of about 500 hours. For women, there is more correlation of the gap with age, i.e., the actual change in hours is not a simple parallel shift of predicted hours. The drop in hours is slightly smaller for younger women, which is consistent with increasing labor supply by young women (in their early 20s). But the gap is actually smallest for older women. This is consistent with older women working relatively *more* than predicted (once we account for a common year effect). Whatever this reflects, it does not appear that there is a movement towards early retirement for women in China. If the strong attachment to work is “ceaseless toil,” there is no sign of it abating.

Figure 2 shows the corresponding results for participation (positive hours worked in the previous year). This allows us to disentangle the zero hours from total, and evaluate the to which there is discrete withdrawal from the labour market, as is the case in developing countries. Furthermore, these figures should be relatively robust to the measurement issues concerning hours that we just discussed. The basic picture that emerges for participation is similar to the one for total hours. Men and women both have a high rate of participation, which only gradually declines at age fifty. Even by age seventy, over half of men and women are working.

Concerning possible cohort or year effects, the middle panel for men shows that the drop in participation is greater than predicted between 1991 and 1997. The gap is just under 10 percentage points, which is non-trivial, given that actual withdrawal was about twice as high as predicted. For the youngest workers, the increase in their participation was about 10 percent less than predicted. But for prime age workers (between 35 and fifty), the gap was much smaller. This pattern by age is not consistent with a simple common year effect, though the correlation of the gap with age may not be statistically significant.

The bottom two panels for women tell a story quite similar to Figure 1. Actual participation dropped more than predicted (between 5 and 10 points), but the gap appears neutral with respect to age. If anything (as with hours), older women's participation decreased relatively less than other ages.

Figure 3 addresses the type of work done over the life-cycle, particularly whether people shift towards farm work from non-agricultural jobs. The top panels show that older men and women are definitely less likely to work off the farm, as they work a lower fraction of their hours in non-agricultural activities. This is an age profile one would expect to be especially contaminated by cohort effects, if there are trends across cohorts towards off-farm work. In the middle panel, we see that these cohort effects are quite pronounced for middle-aged men. The cross-section predicts that forty year-old men would drop their share of hours off the farm by more than 5 percent, but instead they actually increased their relative time off the farm. For older men, the actual drop exceeded that predicted by the cross-section.

The figures for women underline the growing importance of non-farm work. Women of all ages increased their share of work off the farm, contrary to the prediction of the cross-section. Apparently, the cross-section age profile is mostly a "cohort," not "age" profile. Concerning retirement, there is no evidence that older women shift to farming from non-agricultural pursuits.

Figure 4 shows the age profiles for our key health variable, H12. These graphs illustrate the difficulty of disentangling age from cohort effects, and also the potential biases introduced by attrition. What we might expect to see is a steady deterioration of health with age that lines up with the decline in labour supply seen in Figures 1 and 2. The top panels for both men and women suggest this is the case. About 90 percent of twenty year old men, and 80 percent of twenty-year old women report being in good health, compared to sixty percent of sixty year old men, and 40 percent of sixty year old women. Note the slight "uptick" in average health for seventy year old men compared to sixty year olds. It appears that health actually increases with old age. An alternative explanation is that the otherwise unhealthy men are dead by age seventy, so that only the healthy men are left to answer the survey. This is prototypic selection bias that results from panel attrition.

In an economy with rapidly rising incomes, health is expected to improve with time. It might be the case that younger cohorts are permanently healthier than older ones, in which case, the cross-section age profile will be a misleading predictor of the evolution of health with age. In fact, the cross-section and longitudinal data line up for most ages, except the oldest age groups, where the selection bias induced by attrition is worst. The self-reported health status of sixty-five year old men deteriorated much more than predicted by the cross-section. But note the scale of health deterioration for both men and women: Only 2.5 percent of fifty year old men, and about 7.5 percent of sixty year old men saw their health status decline from "good health." If retirement is driven by declines in H12, then it will have to be the case that health has a large effect on desired labour supply, given how few elderly report declines in health.

4.2 Reduced Form Effects of Age

We now provide more precise estimates of the impact of aging on health and labour supply. The regressions are slight variations on (14), where we have added controls for years of schooling (EDU), and province and year dummies:

$$\begin{aligned} L_{it}^{M} &= b_0 + \sum_{j=1}^{J} b_{1j} AGEG(j)_{it}^{M} + b_3 EDU_{it} + b_4' PROV_{it} + b_5' YEAR_{it} + v_{it}^{M} \\ h_{it}^{M} &= d_0 + \sum_{j=1}^{J} d_{1j} AGEG(j)_{it}^{M} + d_3 EDU_{it} + d_4' PROV_{it} + d_5' YEAR_{it} + u_{it}^{M} \end{aligned} \tag{22}$$

Rather than show all of the coefficients, we report the implied effect of aging from fifty to sixty years old ($\Delta_{7060}^{L} = b_{1(70-75)} - b_{1(60-65)}$ and $\Delta_{7060}^{L} = b_{1(70-75)} - b_{1(60-65)}$, and the analogously defined $\Delta_{5060}^{h}, \Delta_{6070}^{h}$). We estimate (22) using both fixed and random effects, and also report the Hausman-Wu test for the equality of these estimates. Note that the test does not distinguish which age coefficients are different, and provides an overall test-statistic for the possible contamination of cohort effects.

The top panel of Table 3 presents the results for men, while Table 4 shows the estimates for women. Looking at the first row, the random effects (RE) estimated decline in labour supply from fifty to sixty years old is 356 hours per year. The fixed effects estimate, on the other hand, is an *increase* of 79 hours, though this is statistically insignificant. The RE and FE estimates are more similar for the change between sixty and seventy, with an approximate decline of 500 hours for ten years of aging. Not surprisingly, the Hausman test rejects the equality of the RE and FE estimates more generally. So, looking across age cohorts, there is a difference in labour supply that looks like "retirement." However, for fifty to sixty year olds, the within-cohort (within-person) estimate is *no* change in labour supply. In identifying the possible role of health in explaining retirement, our benchmark will be how much health can explain (i) the difference in labour supply across cohorts (RE), and (ii) the actual decline in labour supply associated with aging (FE). While there is a significant difference across age cohorts, it does not appear that there is much retirement behaviour to explain from ages fifty to sixty, while there is some between sixty and seventy. The second row shows the results for participation. The patterns mirror those for hours, except that the estimated aging effects are more statistically significant for age sixty to seventy, and there is virtually no difference between FE and RE for this transition.

The key health results are in the next three rows. While there are apparent declines in health reflected in the RE estimates, there does not appear to be any significant decline in health – whether we measure it by H12 or extreme BMI – as people age, when we control for individual (and cohort) fixed effects. Regardless of the eventual estimated impact of health on individual labour supply, it appears unlikely that we will be able to line up declines in labour supply with declines in health.

The results for women are quite similar. Basically, we find no evidence of reduced hours using the FE estimator, although there are differences in hours across age groups using the RE estimator. From ages sixty to seventy, we do see a significant decline in participation, and as with men, this is the only robust evidence of retirement behavour. Similar to men, we do not find systematic declines in health with age, controlling for fixed effects. Furthermore, the Hausman tests suggest we prefer the FE to the RE estimates. It would appear, then, that health can only have a limited role to play in explaining retirement. This comes from the dual finding of (i) very little actual "retirement" (i.e., ceaseless toil), and (ii) even less observed decline in health.

4.3 Structural Estimates and Decompositions

We now explore the possible connections between health and labour supply, by estimating the "structural model," and providing the "bottom line" decompositions, $b_2 \times \Delta^h_{6050}$ and $b_2 \times \Delta^h_{7060}$. The base equation is a variation of (15):

$$L^M_{it} = b_0 + \sum_{j=1}^{J} b_{1j} AGEG(j)^M_{it} + b_{2M} h^M_{it} + b_{2F} h^F_{it} + b_3 EDU_{it} + b_4' PROV_{it} + b_5' YEAR_{it} + e^M_{it} \qquad (23)$$

where we wish to identify the effect of own and spouse's health on labour supply, b_{2M}, b_{2F}. For men, we report the structural estimates in Table 5, and the decompositions back in Table 3 (alongside the reduced forms). For women, the results are reported in Tables 6 and 4.

For men, Table 5 shows the RE (non-instrumented, "OLS") estimate of good health on labour supply is a small, but statistically significant 161.8 hours per year. If this is used to explain retirement, given declines in good health of 0.097 and 0.073 for fifty to sixty and sixty to seventy year olds, we can account for 15.7/356 hours of the decline in hours from ages fifty to sixty, and 11.9/542 of the decline for sixty to seventy years old. These estimates suggest that declining health has nothing to do with retirement. The fixed-effects results tell the same story, with none of the decline between fifty and sixty explained by health, and only 11.3/453.8 explained of the change in hours between sixty and seventy. Wife's health has a statistically significant impact on hours as well, consistent with an "added worker" effect: if men's wives experience a decline in health, they increase their hours of work. Table 5 also shows the results for BMI. Only in the FE specification (which nets out "normal BMI") do we see that adverse changes in BMI have a non-trivial impact on hours worked. However, the conclusions from the decompositions do not change, even with the additional impact of these health measures. Declining health goes no way in explaining retirement.

As discussed previously, however, the estimated health effects may be biased because of the endogeneity of H12. Especially, given the possibility of measurement error (attenuation bias) suggested by Baker, Deri, and Stabile (2002) (and others), it is worth investigating whether we have understated the possible role of health. We use "objective" measures of physical function limitations as instruments, and report the 2SLS results alongside OLS.[12] Essentially, by instrumenting H12, we are emphasizing the correlation with labour supply of that part of self-reported health that is driven by changes in "objective" physical functions. This exercise yields strikingly larger health effects. For RE, the estimated H12 coefficient increases in magnitude by a factor of six, while for FE, the coefficient increases fifteen times. This has a significant effect on the results of the decompositions. For the younger men (aging from fifty to sixty), the RE estimate is 104/356, while the FE estimate is basically zero. Even with a higher estimated impact of health, the FE estimate shows no decline in health with age. So, it is only in comparing labour supply across different cohorts that we see any association of health with hours, in explaining hours differences between fifty and sixty year olds. For the movement from sixty to seventy, though, we explain a much higher fraction of the decline in labour supply. The RE decomposition yields 79/542 hours, while the FE result is 200/453. If we take the instrumented FE coefficients as our preferred results, we see that about 45 percent of the decline in hours between sixty and seventy can be accounted for by declining health.

The bottom panel of Table 5 shows the results for participation. The sign and statistical significance patterns are similar to those for hours, except that only the RE estimates are statistically significant (at the 5% level). If we again take the FE-IV results as the preferred estimates, we cannot explain "retirement" (the discrete withdrawal from working) by changes in health. Deteriorating health has a larger impact on hours worked than on participation.

Tables 6 and 4 show the estimates for women. As was the case for men, the OLS estimates of own-H12 are quite small, though not statistically significant. The only significant own-H12 effects are for the instrumented H12 in the RE specification. In this case, 121/431 (for fifty to sixty) and 149/382 (for sixty to seventy) of the decline in hours can be explained by failing health. Unlike the case for their

[12] Recall that PFs are only available for 1991, 1993. For purposes of comparison, we could add a third set of estimates: OLS for the 1991, 1993 sample. However, the results are essentially the same as for 1991-1997 (with smaller standard errors). We are thus confident that the differences between the reported OLS and IV coefficients are due to bias in OLS, as opposed to the smaller 1991-93 sample.

husbands, the linkage between own-health and labour supply is very weak, especially in the FE equation. Possibly, this suggests that other factors – economic variables?—are more important for women. The spousal effects, however, are stronger for women than men. Indeed, the estimated coefficients on spouse's health are about twice as high. This suggests that a wife's labour supply is significantly more elastic with respect to her husband's health, and that women work more if their husband's cannot. As suggested by Berger (1983) and Berger and Fleisher (1984), spousal health is a significant determinant of labour supply. But, the most notable result from these tables is the small role that health plays in explaining changing work patterns with age.

One possibility is that we have diluted the impact of health on labour supply by restricting young and old to be affected the same by health shocks. In the appendix, we replicated Tables 3-6, but for the sub-sample 40 years and older. The estimated health effects are slightly larger (as expected), though not enough to change the story based on the entire sample. The only notable difference is that the share of the decline in labour supply for men between sixty and seventy using the FE-IV estimates is greater than one half. In summary, to the extent that we regard "ceaseless toil" as working until it is no longer physically possible, it is only for men as they age from sixty to seventy that we can attribute much of the change in labour supply to observable declines in health. We do find statistically and economically significant estimates of the effect of individual health on labour supply for the other age groups: it is just that we do not observe declines in health with age that line up with hours. One important caveat: attrition may be flattening the age profile for health (as discussed previously).

5.0 Extensions

Our results so far suggest a relatively small role for health in explaining retirement, especially at younger ages (between fifty and sixty). This raises the possibility that other variables – possibly economic variables – are more important. Of greatest interest is the role of "income" or "wealth," as this is presumably one of the greatest differences between Chinese elderly and their counterparts in rich countries. Possibly, the Chinese elderly are so poor that they cannot retire. Only death offers respite from work. Indeed, this is one of the important elements in Davis-Friedman's original portrait of the Chinese elderly. Income was also the key factor highlighted by Costa (1998) in explaining trends towards earlier retirement in the United States over the twentieth century.

While we would like to assess the role of economic variables in the retirement decision, the inability to measure the marginal returns to work makes disentangling income from substitution effects impossible. Nevertheless, it is still interesting -- even if the evidence is only suggestive – to contrast the retirement behaviour of "rich" and "poor." In order to do this, we construct a measure of household wealth using various parts of the CHNS. Household wealth is comprised of productive assets (5%) like farm machinery and draft animals, and non-productive wealth, like housing (78%), transportation equipment, livestock, and consumer durables.[13] In the regressions that follow, we use the log of per capita household wealth as the wealth measure, dropping those households in the top and bottom wealth percentiles, in order to avoid using households with extreme levels of wealth. We interact wealth with the age and health variables, and estimate the equations by "OLS" (i.e. no instrumental variables) for both RE and FE specifications.

What might we expect? As Costa (1998) notes when discussing a similar exercise (she finds no significant interactions), the *a priori* effect of wealth (or income) on the response of labour supply to health is ambiguous. On one hand, wealth may facilitate the reduction of labour supply in response to an adverse health shock. If a household is rich and the husband doesn't feel very well, he may more comfortably reduce his labour supply (i.e., his "marginal utility" of leisure may be higher than a poor person, who more highly values consumption). On the other hand, higher income may reduce the linkages between health and labour supply, as Costa suggests became the case in North America. In this case, the

[13] By far the most important source of wealth is housing. All households were surveyed in 1991 about the value of their house. This is the value we use in calculating wealth for all years, except for those households who reported they moved, in which case an updated house value was recorded. For households with missing housing values, we imputed the value of the house using a hedonic regression including housing characteristics (e.g. square footage, building materials, year of construction).

income effect of higher wealth dominates, and health is a minor determinant of the retirement decision: only the poor work until they no longer cannot.

Table 7 shows results for the interactions with wealth. Our focus is on the sign and statistical significance patterns of the interactions with age and health. Looking first at the columns for men, a few suggestive conclusions emerge. First, the coefficient on wealth itself is positive, suggesting that those who are richer work more (thus illustrating the typical problem of identifying income effects in labour supply regressions). However, the age profile for men after age 55 is significantly steeper with greater wealth. If we define retirement as a *reduction* of labour supply, it seems that wealth facilitates retirement. However, the wealth effect is not so strong as to imply that richer men enjoy a greater amount of leisure in old age – merely that they “slow down” more. For own health, the interaction effect is negative, which is consistent with income weakening the link between health and labour supply (similar to Dora Costa’s story).

The results for women provide an interesting contrast. The age profile is similar to their husband’s: richer women work more overall, and have steeper declines in hours after age fifty. But the health-wealth interaction effects are the opposite of the men’s. There is a significant positive interaction term, so that as income increases, the sensitivity of labour supply to health also rises. Own-health matters more for richer women, which means that if they suffer an adverse health shock, they can more readily reduce their labour supply. The cross-health effect (for husband’s health) is also highly dependent on wealth level. The negative interaction effect implies that richer women with healthy husbands are less likely to work (or work fewer hours). This is consistent with the link between health and labour supply weakening with wealth. Certainly, the results of own and spousal health-wealth interactions suggest that women have more elastic labour supply to these income earning activities.

In our final table, we report an even more speculative exercise, and explore the possible links between health and household welfare, as summarized by log per capita household income. These results need to be interpreted with caution: if health is endogenous to labour supply, it is even more so in an income regression. Most obviously, higher income households may be healthier, so the positive association between health and income may have nothing to do with causality running from health to income.[14]

Nevertheless, given the apparent shifts in labour supply in response to health shocks, it is worth exploring whether there is any net effect on income, and if so, whether it is proportional to reductions in labour supply. Furthermore, we can document the links between age and household income. Note that the links between age or health and income may be weakened by a variety of off-setting factors besides family labour supply. Elderly or sick households may receive remittances from other family members, either in the form of cash, or the in-kind provision of farm labour. Households in principle can also hire labour in response to a health shock, in which case farm profits may be immune to the impact of a health shock, as Pitt and Rosenzweig (1986) found in Indonesia. Finally, households are comprised of more than the husband and wife, and the income or labour supply of other family members may more than offset any lost work from aging or sick parents.

Table 8 shows the results of estimating the reduced form and structural equations, with log per capita income replacing hours worked as the dependent variable. We report the main age profiles and the resulting "decompositions", for the RE and FE specifications, with and without instrumenting self-reported health with PF's. We also estimate the equations separately for men and women, as if we are "attributing" household income to the individual.

In the top panel, we report the reduced form age profiles. For aging from fifty to sixty, we estimate no decline in income for men, in either the FE or RE specification. Note that the similarity of the RE and FE results is a little surprising, as we would expect there to be significant cohort effects for household income in a growing economy. In such a case, older individuals would appear poorer, not because of "aging", but because their "generation" (cohort) was always poorer. If that is true, then the FE estimator should sweep out the spurious aging effect, and yield a smaller estimate of the "pure" aging effect. This does not happen for men. For women, we find that the FE estimates are significantly more

[14] See Case, Lubotsky, and Paxson (2002) for a recent discussion and survey of interpretation issues concerning the income-health relationship. While their paper addresses child health, the identification problems in disentangling the direction of causality between income and health are discussed, and a rich set of references is provided.

negative than RE. This suggests that the "pure" aging effect is worse than the apparent RE estimate which would be the case if cohort living standards are actually *declining* (relative to the young) for the elderly. Given we are using household income, how can older women be in poorer households than older men? This is driven by the age difference between husbands and wives: older women live with even older men, and they are thus poorer. They may also be widows. Clearly, older women seem to have worse retirement-income profiles than men. For aging between sixty and seventy, we also find significant declines in income, with the FE estimates more negative than the RE. Again, this is consistent with declining living standards, and may be part of the reason (or another reflection) that retirement ages have not fallen over the 1990s. We also see a steeper decline for women than men, for aging between sixty and seventy.

Turning to the health coefficients, note first that the potential for bias is clearly reflected in the BMI coefficients: high BMI is associated with higher income, while low BMI is associated with lower income. Self-reported health is also positively related to income, but the relationship is surprisingly weak. Even with the IV specifications, none of the H12 coefficients are statistically significant. The results of the decomposition exercise, not surprisingly, reflect the insignificant role of observable health in explaining declining per capita incomes with age.

From this exercise, the most interesting results concern the steep drops in income associated with old age. In fact, the drops in income are greater in percentage terms than the declines in hours, especially for aging between sixty and seventy. This possibly suggests that productivity as measured by income per hour worked, declines with age. Perhaps this can explain the drop in labour supply? Maybe the life-cycle model has more content than previously believed, and retirement behaviour is simply the movement along an intertemporal labour supply function in response to falling (shadow) wages? Such an interpretation would be a stretch in this case (and others, as documented by Card (1994)), and without more formal modeling (with highly restrictive assumptions), it is impossible to identify a causal link between productivity and hours worked.[15] The non-significance of the health variables in the income equation, especially given their apparent role in labour supply, and the large effects of aging, underline the inability of observed health to "explain" the joint deterioration of income and hours with ageing. Possibly, physical strength declines with age, reducing productivity, leading to a reduction in labour supply. Or possibly, hours simply decline, and in concert with reduced productivity, income falls proportionately more.

6.0 Conclusions

To the extent that "ceaseless toil" conjures a bleak image of working until you die, the labour supply pattern of Chinese elderly is consistent with this image. Using longitudinal data from rural China, we document a number of empirical patterns:

- Men and women both work long hours (approaching 1000 hours per year) into their seventies;
- Even with the high rates of economic growth experienced in rural China over the 1990's, we do not see any movement towards "earlier retirement" as one would expect based on the experiences of other countries. This suggests that the elderly may not be enjoying the benefits of growth;
- Using observed health as a means by which to quantify the importance of physical limitations in driving retirement, we find that (i) for men and women under sixty, health plays almost no role. This is primarily because health is not observed to deteriorate at this age; and (ii) For men over sixty (and women to a lesser extent), declining health explains about half of the decline in hours worked by age seventy.
- For both men and women, but especially women, we find that labour supply is sensitive to the health of the spouse, in a pattern consistent with cross-wage elasticity, where women's labour supply is more elastic;

[15] In an ideal world, we would be able to pursue a latent variable approach, as employed by Abowd and Card (1989) and Laszlo (2002). However, the inability to attribute earnings to an individual makes this approach essentially infeasible.

- Furthermore, when we explore the covariation of age and health effects with household wealth, we find generally that wealth mitigates the impact of health on the retirement decision, consistent with Costa's suggestion in an American context. Wealth has the greatest impact on women's own and spousal health responses, and generally increases their elasticity of labour supply (to income generating activities) with respect to health;
- Finally, the per capita household income associated with older individuals is significantly lower than the young. If, as seems to be the case, the elderly are ceaselessly toiling in order to support themselves because they have "no choice", further research on social security arrangements, especially the role played by the family, is an important avenue for future work.[16]

Along the way, we also noted the importance of considering several econometric issues, especially, measurement error in the health variables, and the difficulty of disentangling age from cohort effects.

[16] While we provide no discussion in the paper, we investigated whether there were any clear patterns of the outcomes described here by the living arrangements of the elderly (i.e., whether or not they lived with adult children). No pattern emerged. See Benjamin, Brandt, and Rozelle (2000) for a more detailed discussion of living arrangements of the elderly in rural China, and the conceptual difficulties in estimating the "impact" of living arrangements on elderly outcomes.

7.0 References

Abowd, John and David Card. 1989. "On the Covariance Structure of Earnings and Hours Changes." *Econometrica*, 57:2, pp. 411-45.

Anderson, Kathryn H. and Richard V. Burkhauser. 1984. "The Importance of the Measure of Health in Empirical Estimates of the Labor Supply of Older Men." *Economic Letters*, 16, pp. 375-80.

Baker, Michael, Catherine Deri, and Mark Stabile. 2001. "What do self-reported, objective, measures of health measure?" *NBER Working Paper #8419*.

Benjamin, Dwayne and Loren Brandt. 2002. "Property Rights, Labour Markets, and Efficiency in a Transition Economy: The Case of Rural China." *Canadian Journal of Economics*, 35:4.

Benjamin, Dwayne, Loren Brandt, and Scott Rozelle. 2000. "Aging, Well-being, and Social Security in Rural North China." *Population and Development Review*, 26:0, pp. 89-116.

Berger, Mark. 1983. "Labor Supply and Spouse's Health: The Effects of Illness, Disability, and Mortality." *Social Science Quarterly*, 64:3, pp. 494-509.

Berger, Mark and Belton Fleisher. 1984. "Husband's Health and Wife's Labor Supply." *Journal of Health Economics*, 3:1, pp. 63-75.

Bound, John. 1991. "Self-Reported versus Objective Measures of Health in Retirement Models." *Journal of Human Resources*, 26:1, pp. 106-38.

Bound, John, Michael Schoenbaum, Todd R. Stinebrickner, and Timothy Waidmann. 1999. "The Dynamic Effects of Health on the Labor Force Transitions of Older Workers." *Labour Economics*, 6:2, pp. 179-202.

Card, David. 1994. "Intertemporal labour supply: an assessment," in *Advances in Econometrics: The Sixth World Congress*. C.A. Sims ed. New York: Cambridge University Press, pp. 49-78.

Case, Anne, Darren Lubotsky, and Christina Paxson. 2002. "Economic Status and Health in Childhood: The Origins of the Gradient." *American Economic Review*, 92:5, pp. 1308-34.

Costa, Dora. 1995. "Pensions and Retirement: Evidence from the Union Army Veterans." *Quarterly Journal of Economics*, 110:2, pp. 297-320.

Costa, Dora. 1996. "Health and labor force participation of older men, 1900-1991." *Journal of Economic History*, 56:1, pp. 62-89.

Costa, Dora. 1997. "Displacing the Family: Union Army Pensions and Elderly Living Arrangements." *Journal of Political Economy*, 105:6, pp. 1269-92.

Costa, Dora. 1998. *The Evolution of Retirement: An American Economic History, 1880-1990.* Chicago: University of Chicago Press.

Costa, Dora. 1999. "A House of Her Own: Old Age Assistance and the Living Arrangements of Older Nonmarried Women." *Journal of Public Economics*, 72:1, pp. 39-59.

Costa, Dora and Richard H. Steckel. 1997. "Long-Term Trends in Health, Welfare, and Economic Growth in the United States," in *Health and welfare during industrialization*. Richard H. Steckel and Roderick Floud eds. Chicago: University of Chicago Press, pp. 47-89.

Currie, Janet and Brigitte C. Madrian. 1999. "Health, Health Insurance and the Labor Market," in *Handbook of Labor Economics*. Orley Ashenfelter and David Card eds. New York and Oxford: Elsevier Science, North Holland, pp. 3309-416.

Davis-Friedmann, Deborah. 1991. *Long lives: Chinese elderly and the communist revolution*. Stanford: Stanford University Press.

Deaton, Angus. 1997. *The analysis of household surveys: A microeconometric approach to development policy*. Baltimore and London: Johns Hopkins University Press for the World Bank.

Deaton, Angus and Christina Paxson. 1992. "Patterns of Aging in Thailand and Cote d'Ivoire," in *Topics in the economics of aging*. David A. Wise ed. Chicago: University of Chicago Press, pp. 163-202.

Deaton, Angus and Christina Paxson. 1998. "Aging and Inequality in Income and Health." *American Economic Review*, 88:2, pp. 248-53.

Fan, Jianqing. 1992. "Design-adaptive Nonparametric Regression." *Journal of the American Statistical Association*, 87, pp. 998-1004.

Goldin, Claudia. 1990. *Understanding the Gender Gap: An Economic History of American Women*. New York: Oxford University Press.

Hurd, Michael. 1990. "Research on the Elderly: Economic Status, Retirement, and Consumption and Saving." *Journal of Economic Literature*, 28:2, pp. 565-637.

Hurd, Michael and Michael Boskin. 1984. "The Effect of Social Security on Retirement in the Early 1970s." *Quarterly Journal of Economics*, 99:4, pp. 767-90.

Laszlo, Sonia. 2002. "Education, Labour Supply and Market Development in Rural Peru." *Ryerson University Working Paper, #34*.

Lumsdaine, Robin and Olivia S. Mitchell. 1999. "New Developments in the Economic Analysis of Retirement," in *Handbook of Labor Economics*. Orley Ashenfelter and David Card eds. New York and Oxford: Elsevier Science, North Holland, pp. 3261-307.

McGarry, Kathleen. 2002. "Health and Retirement: Do Changes in Health Affect Retirement Expectations?" *NBER Working Paper, #9317*.

McLellan, Mark. 1998. "Health Events, Health Insurance, and Labor Supply: Evidence from the Health and Retirement Survey," in *Frontiers in the economics of aging*. David A. Wise ed. Chicago: University of Chicago Press, pp. 301-46.

Mete, Cem and T. Paul Schultz. 2002. "Health and Labor Force Participation of the Elderly in Taiwan." *Manuscript, Economic Growth Center, Yale University*.

Parsons, Donald O. 1980. "The Decline in Male Labor Force Participation." *Journal of Political Economy*, 88:1, pp. 117-34.

Pitt, Mark and Mark Rosenzweig. 1986. "Agricultural Prices, Food Consumption, and the Health and Productivity of Indonesian Farmers," in *Agricultural household models: Extensions, applications, and policy*. Inderjit Singh, Lyn Squire and John Strauss eds. Baltimore and London: Johns Hopkins University Press for the World Bank, pp. 153-82.

World Bank. 1994. *Averting the old age crisis: policies to protect the old and promote growth*. Oxford: Published for the World Bank by Oxford University Press.

8.0 Appendix

8.1 The CHNS Sample and Patterns of Attrition

The first wave of the CHNS was carried out in 1989, with subsequent waves in 1991, 1993, 1997, and 2000. The wage income and health data are problematic for 1989, so we use th1 1991-1997 data (the 2000 data are not yet available). The sampling pattern of individuals, households, and communities for the CHNS from 1991 to 1997 is summarized in Table A1. Of 143 communities surveyed in 1991, only 108 survived through all subsequent waves, primarily because Liaoning province was dropped, and "replaced" Heilongjiang in 1997.

The attrition rate increases with the unit of observation, being smallest for communities, and highest for individuals. Just under two-thirds (61%) of households, but only 46 percent of individuals were surveyed in all three years.

As a simple summary of the characteristics of individuals who attrited from the 1991 sample, we estimated a linear probability model of whether an individual is in our panel sample (which also includes people who died and are accounted for). The results are as follows:

Linear Probability of Remaining in Sample for 1991-1997

Gender	edu	h12	lnpcy	a1019	a2029	a3039	a5059	a6069	a7079	a80
-0.11 (11.1)	-0.01 (6.3)	0.00 (0.0)	-0.02 (2.8)	-0.11 (7.5)	-0.22 (13.6)	-0.01 (0.6)	-0.02 (1.1)	-0.04 (1.9)	-0.04 (1.3)	-0.12 (2.3)

Notes:

1. Sample Size is 7830.
2. The dependent variable is an indicator of whether an individual remains in the sample all three years (or dies and is accounted for);
3. The sample includes all men and women in 1991 aged 10 years and older;
4. The absolute t-statistics are reported in parentheses.
5. The regression also includes dummies for the 125 communities, the F-statistic for which is 24.7, which is highly significant.

These results suggest that younger individuals (especially age 20 to 29 in 1991) were least likely to remain in the panel, as were those over 80 years old (who presumably died, but their deaths were not recorded in subsequent surveys). More educated individuals, and to some extent, those from higher income households were also most likely to leave the sample, probably though migration. Women were also more likely to leave the sample, as it is the custom for women to marry outside their village. As far as health is concerned, it does not appear that there is any correlation between health and attrition (conditional on our including subsequent deaths in our sample).

8.2 Definitions of Physical Functions

The following table summarizes the five CHNS questions pertaining to physical function limitations that we can use (comparably) for 1991 and 1993:

Question	Answers
How is the present condition of your heart, your lungs, and your stomach?	1. normal. 2. occasionally affect daily activities and work 3. frequently affect daily activities and work. 4. unable to carry out daily activities and work.
How is the present condition of your upper extremities, shoulders, upper back, and neck?	1. functioning normally. 2. having some problems, but not affecting daily activities and work. 3. slightly affecting daily activities and work, some degree of help is needed. 4. affecting daily activities and work, help is required.
How is the present condition of your lower extremities and spinal cord, and does this affect your walking?	1. functioning normally. 2. having some problems, but can still walk alone. 3. needing some help walking. 4. cannot walk, confined to bed, using wheelchair, or carried by others.
How is the present condition of your hearing, eyesight, and speaking?	1. functioning normally. 2. wearing glasses, hearing aid, or having some loss of ability. 3. deaf in ear, blind in one eye, some loss of speech, or serious loss of vision, hearing or speech. 4. completely deaf, blind, or unable to speak.
How is the present condition of your urine control and bowel control?	1. normal. 2. nighttime or occasional loss of urine or bowel control. 3. frequent loss of urine or bowel control. 4. total loss of urine or bowel control.

Table 1
Descriptive Statistics for Sample Individuals, 1991

	All Men		Men 50+		All Women		Women 50+	
Labour Supply								
Positive Weeks Worked	0.92		0.82		0.93		0.82	
Annual Hours Worked	1962		1686		2036		1565	
Wage Work (and %)	495	(25.2%)	256	(15.2%)	233	(11.4%)	50	(3.2%)
Gardening (and %)	129	(6.6%)	160	(9.5%)	171	(8.4%)	171	(10.9%)
Farming (and %)	1037	(52.9%)	1031	(61.2%)	1118	(54.9%)	863	(55.1%)
Animal Husbandry (and %)	121	(6.2%)	142	(8.4%)	460	(22.6%)	444	(28.4%)
Fishing (and %)	10	(0.5%)	4	(0.2%)	2	(0.1%)	1	(0.1%)
Family Business (and %)	168	(8.6%)	94	(5.6%)	53	(2.6%)	36	(2.3%)
Health								
Good Health (H12)	0.74		0.58		0.72		0.53	
BMI	21.15		20.91		21.50		20.96	
Low BMI	0.20		0.26		0.20		0.30	
High BMI	0.11		0.13		0.20		0.17	
Index of Physical Functions	-2.27		-2.40		-2.29		-2.43	
Subsequent Death	0.07		0.17		0.03		0.06	
Demographic Characteristics								
Age	43.6		59.3		41.6		58.6	
Education	5.9		3.5		3.5		0.8	
Household Size	4.6		4.4		4.6		4.3	
Sample Size	1253		375		1266		296	

Notes:

1. The full sample is defined as all men or women over the age of 20 who were in the 1991-1997 panel, unless they "left the sample" because of death. Furthermore, the sample is restricted to those individuals with observations for all variables, including spouse's health status. This effectively restricts the sample to married individuals with a living spouse. Men and Women 50+ is the subsample of individuals 50 years and older.
2. The Labour supply variables are: (1) an indicator of whether the person worked positive weeks in the previous year; (2) hours worked in the previous year, including zeroes; (3) Hours spent in various types of work, (with the percentage of total hours in parentheses).
3. The health measures are: (1) H12 (self-reported health in the top two categories); (2) BMI; (3) Low BMI (BMI less than the 20th percentile); (4) High BMI (BMI greater than the 80th percentile); (5) An index of physical infirmities (described in the appendix); and (6) An indicator of whether the person died between 1991 and 1997.

Table 2
The Information Content of Health Measures
(absolute *t*-values in parentheses, * indicates statistical significance at the 5% level)

	Men		**Women**	
	Full Sample	**Older Sample**	**Full Sample**	**Older Sample**
Subsequent Death:				
H12	-0.035*	-0.062*	-0.031*	-0.045*
	(2.3)	(2.6)	(3.0)	(2.6)
Low BMI	-0.024	-0.042	0.017	0.026
	(1.4)	(1.5)	(1.4)	(1.3)
High BMI	-0.019	-0.024	0.000	-0.014
	(0.9)	(0.7)	(0.0)	(0.7)
Index of Physical Functions	-0.560*	-0.588*	-0.071	-0.041
	(3.9)	(3.0)	(0.8)	(0.3)
Hours Worked:				
Subsequent Death	-320.5*	-363.5*	-729.1*	-839.9*
	(2.2)	(2.2)	(3.1)	(3.3)
H12	133.8	223.9*	-122.5	-65.7
	(1.7)	(2.2)	(1.4)	(0.6)
Low BMI	-15.9	-133.5	-90.5	42.2
	(0.2)	(1.1)	(0.9)	(0.3)
High BMI	-118.2	-98.1	-67.9	16.6
	(1.1)	(0.7)	(0.7)	(0.1)
Index of Physical Functions	490.7	570.6	1009.5	1539.2
	(0.7)	(0.7)	(1.4)	(1.8)
Sample Size	1226	716	1237	637

Notes:

1. In the first panel, the reported coefficients are from regressions of an indicator of whether an individual died between 1991 and 1997, as a function of his/her health in 1991. The regression includes other covariates: age-group dummies (defined in the text, and in Table 3), years of education, and province dummies.
2. The second panel shows cross-section estimates of individual labour supply (hours worked) in 1991 as a function of health status in 1991, including a variable indicating whether the person subsequently died (by the 1997 survey). The regression includes the standard covariates (age dummies, education, and province dummies).
3. The regressions are estimated separately for men and women, for the full sample, and for the sample 40 and older.
4. The health indicators are described in the text, and in Table 3 (H12 is the indicator of self-reported good health; Low and High BMI; and an index of physical functions (where higher numbers mean better health).

Table 3
Health and Labour Supply: Men
(p-values in parentheses, * indicates significance at the 5% level)

	Age 50 to 60		Age 60 to 70		Hausman Test
Reduced Form Age Profiles:	**RE**	**FE**	**RE**	**FE**	**RE vs. FE**
Hours	-356.2*	79.0	-542.3*	-453.8	30.2*
	(0.000)	(0.681)	(0.000)	(0.057)	(0.011)
Participation	-0.144*	-0.035	-0.301*	-0.261*	38.2*
	(0.000)	(0.430)	(0.000)	(0.000)	(0.001)
H12	-0.097*	0.003	-0.073	-0.082	27.9*
	(0.006)	(0.963)	(0.186)	(0.377)	(0.023)
Low BMI	0.093*	0.026	0.036	0.045	28.1*
	(0.001)	(0.601)	(0.424)	(0.468)	(0.021)
High BMI	-0.008	0.004	0.080*	0.070	8.1
	(0.754)	(0.926)	(0.041)	(0.201)	(0.921)
The Estimated Effect of Health on Labour Supply:					
Change in hours:	-356.2	79.0	-542.3	-453.8	
H12 (OLS)	-15.7	0.5	-11.9	-11.3	
H12, BMI (OLS)	-20.5	-4.4	-19.1	-20.4	
H12 (IV)	-104.3	8.9	-79.0	-208.2	
H12, BMI (IV)	-99.0	4.5	-89.2	-217.7	
Change in participation:	-0.144	-0.035	-0.301	-0.261	
H12 (OLS)	-0.004	0.000	-0.003	-0.004	
H12, BMI (OLS)	-0.004	-0.001	-0.007	-0.007	
H12 (IV)	-0.011	0.000	-0.008	-0.011	
H12, BMI (IV)	-0.009	0.000	-0.012	-0.014	

Notes:

4. The estimated age profile is based on the difference in age coefficients from a regression of the dependent variable on covariates: age-group dummies (20-24, 25-29,…, 80+); year dummies for 1993, 1997; years of education; and province dummies. The Age 50 to 60 "profile" is the difference between the age 60-64 and age 50-54 coefficients, while the Age 60 to 70 "profile" is the difference between the age 70-74 and age 60-64 coefficients. P-values are reported for the hypothesis that the difference is zero.
5. The regressions are estimated over the pooled sample (1991, 1993, and 1997) for men 20 years and older. Sample size is 3600.
6. "RE" refers to the random effects specification, while "FE" refers to the (individual) fixed effects specification. The Hausman Test column reports the results of the Hausman test of random versus fixed effects (p-values in parentheses).
7. The dependent variables for the labour supply measures are: Hours (hours worked in the previous year, including zero); Participation (whether the individual worked at all the previous year); H12 (self-reported health in the top two categories); Low BMI (BMI less than the 20^{th} percentile); and High BMI (BMI greater than the 80^{th} percentile).
8. The estimated effect of health on labour supply is the predicted impact of the change in health (H12, or H12 and BMI) multiplied by the estimated "structural" effect of health on labour supply (from Table 3). OLS refers to predictions based on OLS coefficients, while IV refers to the 2SLS-based coefficients.

Table 4
Health and Labour Supply: Women
(p-values in parentheses, * indicates statistical significance at the 5% level)

	Age 50 to 60		Age 60 to 70		Hausman Test
Reduced Form Age Profiles:	**RE**	**FE**	**RE**	**FE**	**RE vs. FE**
Hours	-431.0*	-341.7	-382.9*	97.6	11.3
	(0.000)	(0.098)	(0.039)	(0.749)	(0.734)
Participation	-0.123*	-0.055	-0.246*	-0.154*	17.6
	(0.000)	(0.234)	(0.000)	(0.024)	(0.283)
H12	-0.083*	0.045	-0.101	-0.146	14.6
	(0.035)	(0.583)	(0.150)	(0.225)	(0.483)
Low BMI	0.012	-0.046	0.050	0.030	27.4*
	(0.705)	(0.387)	(0.370)	(0.706)	(0.025)
High BMI	0.015	0.083	-0.092	-0.044	11.5
	(0.621)	(0.109)	(0.092)	(0.567)	(0.716)
The Estimated Effect of Health on Labour Supply:					
Change in hours:	-431.0	-341.7	-382.9	97.6	
H12 (OLS)	-1.6	-0.8	-2.0	2.6	
H12, BMI (OLS)	-2.0	-7.5	1.9	6.9	
H12 (IV)	-121.9	47.0	-148.6	-153.3	
H12, BMI (IV)	-121.2	52.1	-147.3	-154.8	
Change in participation:	-0.123	-0.055	-0.246	-0.154	
H12 (OLS)	-0.001	0.000	-0.001	-0.001	
H12, BMI (OLS)	-0.001	0.002	0.001	-0.002	
H12 (IV)	-0.008	-0.004	-0.009	0.013	
H12, BMI (IV)	-0.008	-0.001	-0.007	0.012	

Notes:

9. The estimated age profile is based on the difference in age coefficients from a regression of the dependent variable on covariates: age-group dummies (20-24, 25-29,…, 80+); year dummies for 1993, 1997; years of education; and province dummies. The Age 50 to 60 "profile" is the difference between the age 60-64 and age 50-54 coefficients, while the Age 60 to 70 "profile" is the difference between the age 70-74 and age 60-64 coefficients. P-values are reported for the hypothesis that the difference is zero.
10. The regressions are estimated over the pooled sample (1991, 1993, and 1997) for women 20 years and older. Sample size is 3602.
11. "RE" refers to the random effects specification, while "FE" refers to the (individual) fixed effects specification. The Hausman Test column reports the results of the Hausman test of random versus fixed effects (p-values in parentheses).
12. The dependent variables for the labour supply measures are: Hours (hours worked in the previous year, including zero); Participation (whether the individual worked at all the previous year); H12 (self-reported health in the top two categories); Low BMI (BMI less than the 20th percentile); and High BMI (BMI greater than the 80th percentile).
13. The estimated effect of health on labour supply is the predicted impact of the change in health (H12, or H12 and BMI) multiplied by the estimated "structural" effect of health on labour supply (from Table 5). OLS refers to predictions based on OLS coefficients, while IV refers to the 2SLS-based coefficients.

Table 5
Structural Models: The Effect of Health on Labour Supply
Men 20 and Older
(absolute *t*-values in parentheses, * indicates statistical significance at the 5% level)

	Random Effects		Fixed Effects	
	OLS	**IV**	**OLS**	**IV**
Hours:				
H12	161.8*	1076.1*	137.9	2539.8*
	(2.7)	(3.0)	(1.8)	(2.6)
Wife's H12	-145.3*	-458.2	-104.2	-1301.6
	(2.6)	(1.2)	(1.4)	(1.5)
Low BMI	-57.6	44.5	-182.7*	-162.9
	(1.1)	(0.6)	(2.2)	(0.9)
High BMI	-64.7	-147.1	-11.3	-29.5
	(1.1)	(1.8)	(0.1)	(0.2)
Participation:				
H12	0.043*	0.112	0.046*	0.133
	(3.1)	(1.4)	(2.6)	(0.7)
Wife's H12	-0.049*	-0.169	-0.040*	-0.299
	(3.6)	(1.9)	(2.3)	(1.7)
Low BMI	-0.007	0.017	-0.049*	-0.024
	(0.5)	(1.0)	(2.5)	(0.7)
High BMI	-0.048*	-0.052*	-0.018	-0.027
	(3.3)	(2.8)	(0.8)	(0.7)

Notes:

1. The estimated coefficients are the coefficients from a regression of labour supply (hours or participation) on health variables (H12 = good health; wife's H12; and own Low and High BMI indicators), plus covariates (age group dummies (the same as table 2), years of education, year and province dummies. Absolute *t*-values for the coefficients are shown in parentheses.
2. "OLS" is estimated on the full pooled sample of men 20 and older (1991,1993,1997), and is estimated by Random and Fixed Effects. The sample size is 3600.
3. "IV" is 2SLS estimation for the two available years (1991 and 1993) with an index of physical "infirmities" for the individual and spouse as instruments for individual and spouse H12. The sample size is 2257. The IV specification is also estimated by Random and Fixed Effects.
4. The *t*-value for the difference between the OLS and IV own H12 coefficients (i.e., the Hausman Tests) are (1) For RE Hours, 2.59; (2) For FE Hours, 2.47; (3) For RE Participation, 0.88; (4) For FE Participation, 0.46.

Table 6
Structural Models: The Effect of Health on Labour Supply
Women 20 and Older
(absolute *t*-values in parentheses)

	Random Effects		Fixed Effects	
	OLS	**IV**	**OLS**	**IV**
Hours:				
H12	19.6	1476.4*	-18.0	1049.7
	(0.3)	(3.2)	(0.3)	(1.1)
Husband's H12	-87.3	-996.4*	23.4	-2317.4
	(1.5)	(2.8)	(0.3)	(1.9)
Low BMI	12.7	48.1	142.8	226.1
	(0.2)	(0.6)	(1.7)	(0.9)
High BMI	-35.2	11.6	-0.7	187.9
	(0.7)	(0.2)	(0.0)	(0.8)
Participation:				
H12	0.007	0.091	0.008	-0.092
	(0.6)	(1.0)	(0.5)	(0.5)
Husband's H12	-0.026	-0.056	-0.018	-0.345
	(1.9)	(0.8)	(1.1)	(1.4)
Low BMI	0.018	0.022	0.011	0.081
	(1.5)	(1.4)	(0.6)	(1.7)
High BMI	-0.013	-0.017	0.028	0.081
	(1.0)	(1.2)	(1.5)	(1.7)

Notes:

5. The estimated coefficients are the coefficients from a regression of labour supply (hours or participation) on health variables (H12 = good health; husband's H12; and own Low and High BMI indicators), plus covariates (age group dummies (the same as table 4), years of education, year and province dummies. Absolute *t*-values for the coefficients are shown in parentheses.
6. "OLS" is estimated on the full pooled sample of women 20 and older (1991,1993,1997), and is estimated by Random and Fixed Effects. The sample size is 3602.
7. "IV" is 2SLS estimation for the two available years (1991 and 1993) with an index of physical "infirmities" for the individual and spouse as instruments for individual and spouse H12. The sample size is 2137. The IV specification is also estimated by Random and Fixed Effects.
8. The *t*-value for the difference between the OLS and IV own H12 coefficients (i.e., the Hausman Tests) are (1) For RE Hours, 3.19; (2) For FE Hours, 0.93; (3) For RE Participation, 0.88; (4) For FE Participation, 0.54. For the effect of husband's H12 (on hours), the RE Hausman test is 2.59, and for FE it is 1.92.

Table 7
The Interaction of Wealth and Labour Supply
OLS Estimates
(absolute *t*-values in parentheses, * indicates statistical significance at the 5% level)

	Men				Women			
	Hours		Participation		Hours		Participation	
	RE	FE	RE	FE	RE	FE	RE	FE
H12	921.3	1767.2*	-0.110	0.201	-1078.6	-1713.7*	-0.044	-0.196
	(1.6)	(2.2)	(0.8)	(1.2)	(1.9)	(2.4)	(0.4)	(1.3)
Spouse's H12	-599.3	-610.1	0.184	0.015	1561.4*	2807.9*	0.164	0.352*
	(1.0)	(0.8)	(1.4)	(0.1)	(2.6)	(3.7)	(1.3)	(2.2)
Wealth×H12	-88.4	-200.3	0.022	-0.019	157.1*	242.2*	0.006	0.026
	(1.1)	(1.9)	(1.3)	(0.8)	(2.1)	(2.6)	(0.4)	(1.3)
Wealth× SH12	44.8	51.5	-0.032	-0.009	-221.5*	-387.8*	-0.024	-0.049*
	(0.6)	(0.5)	(1.9)	(0.4)	(2.8)	(3.9)	(1.4)	(2.3)
Wealth	147.9	257.2*	0.008	0.038	169.5*	318.3*	0.019	0.029
	(1.8)	(2.0)	(0.4)	(1.4)	(2.1)	(2.9)	(1.1)	(1.3)
Wealth×A2024	267.3	506.2	-0.029	-0.017	310.7	607.9	0.076	0.077
	(1.2)	(1.3)	(0.6)	(0.2)	(1.5)	(1.9)	(1.7)	(1.2)
Wealth×A2529	-19.0	111.3	0.022	0.054	-110.5	-9.0	0.004	0.025
	(0.2)	(0.6)	(0.9)	(1.3)	(1.0)	(0.1)	(0.2)	(0.8)
Wealth×A3034	-70.6	-127.0	-0.014	-0.034	-144.1	-180.6	-0.018	-0.024
	(0.7)	(0.9)	(0.6)	(1.1)	(1.6)	(1.5)	(1.0)	(1.0)
Wealth×A3539	-44.8	-157.6	0.002	-0.015	-81.2	-119.5	-0.013	-0.002
	(0.5)	(1.3)	(0.1)	(0.6)	(1.0)	(1.2)	(0.7)	(0.1)
Wealth×A4549	35.6	-11.2	-0.002	-0.007	-72.7	2.8	-0.011	-0.013
	(0.4)	(0.1)	(0.1)	(0.3)	(0.9)	(0.0)	(0.6)	(0.6)
Wealth×A5054	7.5	-146.1	-0.017	-0.041	-188.3	-256.7	-0.051*	-0.060*
	(0.1)	(1.1)	(0.8)	(1.4)	(1.9)	(1.9)	(2.4)	(2.2)
Wealth×A5559	-197.0	-201.4	-0.017	0.002	-198.1	-283.3	-0.024	-0.015
	(1.8)	(1.3)	(0.7)	(0.1)	(1.7)	(1.8)	(1.0)	(0.5)
Wealth×A6064	-239.2	-321.6	-0.026	-0.020	-245.4	-308.5	-0.039	-0.038
	(1.9)	(1.7)	(0.9)	(0.5)	(1.8)	(1.6)	(1.4)	(1.0)
Wealth×A6569	-319.3*	-449.4	-0.108*	0.015	-300.7*	-471.7*	-0.055	-0.098*
	(2.3)	(1.9)	(3.4)	(0.3)	(2.0)	(2.0)	(1.7)	(2.0)
Wealth×A7074	-352.2	-419.8	-0.117*	0.033	-159.0	-415.7	-0.103*	-0.204*
	(1.8)	(1.4)	(2.6)	(0.5)	(0.8)	(1.3)	(2.5)	(3.1)
Wealth×A7579	-268.9	-576.5	-0.172*	-0.447	-258.1	-307.7	-0.192*	-0.241*
	(1.1)	(0.5)	(3.1)	(1.8)	(0.9)	(0.7)	(3.2)	(2.5)
Wealth×A80	-333.2	567.4	-0.212*	-0.856*	-265.3	-495.3	-0.300	-0.732*
	(0.9)	(0.5)	(2.4)	(3.7)	(0.2)	(0.4)	(1.3)	(2.6)

Notes:

1. The coefficients are from a regression of labour supply (hours and participation) on the base covariates (the same as Tables 3 and 5), with log per capita wealth included, and interacted with the health and age-group dummies. The omitted age category is A4044.
2. The regressions are estimated for men 20 years and older (sample size = 2183) and women 20 years and older (sample size = 2141)
3. Models are estimated by "OLS" (i.e., not instrumental variables) by Random Effects (RE) and Fixed Effects (FE).

Table 8
Health, Aging, and Log Per Capita Income
(* indicates statistical significance at the 5% level)

	Men				Women			
Reduced Form Aging Profiles (*p*-values in parentheses):								
	RE		**FE**		**RE**		**FE**	
Age 50 to 60	-0.012		0.123		-0.126		-0.319*	
	(0.858)		(0.354)		(0.076)		(0.022)	
Age 60 to 70	-0.158		-0.365*		-0.113		-0.485*	
	(0.131)		(0.028)		(0.373)		(0.018)	
Hausman Test	34.31*				25.44*			
(RE vs. FE)	(0.003)				(0.044)			
"Structural" Models of the Impact of Health on Log Per Capita Income (absolute *t*-values in parentheses):								
	RE		**FE**		**RE**		**FE**	
	OLS	**IV**	**OLS**	**IV**	**OLS**	**IV**	**OLS**	**IV**
H12	0.095*	0.252	0.070	0.161	0.013	-0.475	0.033	-0.149
	(2.3)	(1.0)	(1.3)	(0.3)	(0.3)	(1.7)	(0.7)	(0.3)
SH12	0.032	-0.347	0.033	-0.154	0.108*	0.289	0.048	0.217
	(0.8)	(1.3)	(0.7)	(0.3)	(2.6)	(1.3)	(1.0)	(0.3)
LBMI	-0.064	-0.056	-0.003	0.041	-0.031	-0.104*	0.035	-0.054
	(1.7)	(1.1)	(0.1)	(0.4)	(0.8)	(2.1)	(0.6)	(0.4)
HBMI	0.128*	0.105	-0.003	0.054	0.124*	0.080	0.069	-0.021
	(3.0)	(1.9)	(0.0)	(0.5)	(3.3)	(1.6)	(1.2)	(0.2)
Sample Size	3537	2225	3537	2225	3541	2108	3541	2108
Predicted Effect of Own Health on Log Per Capita Income:								
Age 50 to 60	-0.012		0.123		-0.126		-0.319	
H12 only	-0.009	-0.024	0.000	0.001	-0.001	0.039	0.001	-0.007
H12 and BMI	-0.016	-0.031	0.000	0.002	0.001	0.039	0.006	-0.006
Age 60 to 70	-0.158		-0.365		-0.113		-0.485	
H12 only	-0.007	-0.019	-0.006	-0.013	-0.001	0.048	-0.005	0.022
H12 and BMI	0.001	-0.012	-0.006	-0.007	-0.014	0.035	-0.007	0.021

Notes:
1. The first panel reports the reduced form aging profiles (analogous to Tables 2 and 4) for log per capita household income. See the notes to Tables 2 and 4 for details. The Hausman Test concerns the hypothesis that the random and fixed effects estimators are the same.
2. The second panel reports the structural estimates of the effect of health on log per capita household income, analogous to Tables 3 and 5. See the notes to tables 3 and 5 for details.
3. The third (and bottom) panel shows the estimated effect of health on the "reduced form" change in log per capita income, analogous to Tables 2 and 4.
4. The estimates are shown for men and women (20 years and older), and models are estimated by "OLS" (RE and FE), and Instrumental Variables (IV), where H12 and spouse's health (SH12) are treated as endogenous, and the instruments are the same as described in Tables 3 and 5.

Appendix Table A1
Sampling Patterns (Attrition) for CHNS (Rural)

	Communities		Households		Individuals	
	Number	**Percentage**	**Number**	**Percentage**	**Number**	**Percentage**
In 1991 Only			91	3.0	1,201	9.4
In 1993 Only			12	0.4	309	2.4
In 1997 Only	17	11.9	577	19.0	2,071	16.2
In 1991 & 1993	18	12.6	411	13.5	2,006	15.7
In 1993 & 1997			36	1.2	504	3.9
In 1991 & 1997			55	1.8	768	6.0
In 1991, 1992, & 1993	108	75.5	1,863	61.2	5,926	46.4
Total	143	100.0	3,045	100.0	12,785	100.0

Notes:

1. The table provides in formation on how many individuals, households, and communities appear in the 1991, 1993, and 1997 CHNS for each year and combination of years.
2. The percentages are column percentages for each category

Appendix Table A2
First Stage Regressions
(absolute t-values in parentheses, * indicates statistical significance at the 5% level)

Men

	Full Sample				Older Sample (40 and Older)			
	Random Effects		Fixed Effects		Random Effects		Fixed Effects	
Instrumented Variable:	**H12**	**SH12**	**H12**	**SH12**	**H12**	**SH12**	**H12**	**SH12**
Physical Functions Index	1.560* (8.3)	0.152 (0.8)	0.916* (2.8)	-0.471 (1.3)	1.634* (7.8)	0.220 (1.0)	0.876* (2.3)	-0.557 (1.4)
Wife's Physical Functions Index	0.223 (1.2)	1.423* (7.4)	0.391 (1.2)	1.219* (3.6)	0.234 (1.1)	1.556* (7.1)	0.674 (1.8)	1.569* (3.9)
Low BMI	-0.009 (0.4)	0.060* (2.4)	0.049 (1.0)	0.103* (2.0)	-0.013 (0.4)	0.070* (2.1)	0.061 (0.9)	0.116 (1.6)
High BMI	0.027 (1.0)	0.004 (0.1)	0.036 (0.6)	-0.022 (0.3)	0.044 (1.1)	0.013 (0.3)	-0.025 (0.3)	-0.095 (0.9)
Chi-squared Test (p-value)	80.9* (0.000)	61.3* (0.000)	5.8* (0.003)	6.4* (0.002)	70.6* (0.000)	58.6* (0.000)	5.9* (0.003)	7.5* (0.001)
Sample Size	2257	2257	2257	2257	1395	1395	1395	1395

Women

	Full Sample				Older Sample (40 and Older)			
	Random Effects		Fixed Effects		Random Effects		Fixed Effects	
Instrumented Variable:	**H12**	**SH12**	**H12**	**SH12**	**H12**	**SH12**	**H12**	**SH12**
Physical Functions Index	1.370* (6.9)	0.233 (1.2)	1.220* (3.4)	0.293 (0.9)	1.426* (6.3)	0.241 (1.1)	1.511* (3.6)	0.474 (1.2)
Husband's Physical Functions Index	0.113 (0.6)	1.683* (9.2)	-0.527 (1.5)	0.791* (2.4)	0.170 (0.8)	1.736* (8.4)	-0.605 (1.5)	0.793* (2.1)
Low BMI	-0.051* (2.0)	-0.010 (0.4)	0.149* (2.8)	0.096 (1.9)	-0.051 (1.4)	-0.024 (0.7)	0.204* (2.6)	0.094 (1.3)
High BMI	-0.008 (0.3)	0.005 (0.2)	0.119* (2.0)	0.105 (1.9)	0.014 (0.4)	0.029 (0.8)	0.079 (0.9)	0.096 (1.2)
Chi-squared Test (p-value)	52.8* (0.000)	98.5* (0.000)	6.0* (0.003)	4.0* (0.018)	45.5* (0.000)	80.7* (0.000)	6.6* (0.002)	3.7* (0.026)
Sample Size	2137	2137	2137	2137	1190	1190	1190	1190

Notes:

1. The reported coefficients are from the first stage equations corresponding to the "structural" labour supply equations summarized in Tables 3, 5, 3A and 5A.
2. The dependent variables are the self reported health measures: own good health (H12) and spouse's good health (SH12). The explanatory variables are the full set of covariates (age group dummies, years of education, year dummy, and province dummies), plus the reported health coefficients.
3. The excluded (identifying) instruments are the own and spouse's index of physical functions (problems). The Chi-square test (and p-value) is of the null hypothesis that these two variables can be excluded from the first-stage equation. The hypothesis is rejected in each case.

Appendix Table A3
Health and Labour Supply: Men over 40
(p-values in parentheses, * indicates statistical significance at the 5% level)

	Age 50 to 60		**Age 60 to 70**		**Hausman Test**
Reduced Form Age Profiles:	**RE**	**FE**	**RE**	**FE**	**RE vs. FE**
Hours	-351.1*	261.7	-532.0*	-285.0	32.6*
	(0.000)	(0.222)	(0.001)	(0.271)	(0.001)
Participation	-0.142*	-0.022	-0.301*	-0.248*	36.8*
	(0.000)	(0.659)	(0.000)	(0.000)	(0.000)
H12	-0.094*	-0.036	-0.078	-0.128	17.4
	(0.011)	(0.668)	(0.185)	(0.213)	(0.096)
Low BMI	0.102*	0.043	0.010	0.034	26.1*
	(0.000)	(0.418)	(0.827)	(0.595)	(0.006)
High BMI	0.007	0.036	0.062	0.067	3.7
	(0.777)	(0.396)	(0.082)	(0.190)	(0.978)
<u>The Estimated Effect of Health on Labour Supply:</u>					
Change in hours:	-351.1	261.7	-532.0	-285.0	
H12 (OLS)	-19.9	-7.1	-16.4	-24.8	
H12, BMI (OLS)	-31.4	-17.4	-22.4	-31.0	
H12 (IV)	-94.4	-108.9	-77.7	-383.4	
H12, BMI (IV)	-96.0	-113.8	-86.9	-381.6	
Change in participation:	-0.144	-0.035	-0.301	-0.261	
H12 (OLS)	-0.004	0.000	-0.003	-0.004	
H12, BMI (OLS)	-0.004	-0.001	-0.007	-0.007	
H12 (IV)	-0.011	0.000	-0.008	-0.011	
H12, BMI (IV)	-0.009	0.000	-0.012	-0.014	

Notes:

14. The estimated age profile is based on the difference in age coefficients from a regression of the dependent variable on covariates: age-group dummies (40-24, 45-29,…, 80+); year dummies for 1993, 1997; years of education; and province dummies. The Age 50 to 60 "profile" is the difference between the age 60-64 and age 50-54 coefficients, while the Age 60 to 70 "profile" is the difference between the age 70-74 and age 60-64 coefficients. P-values are reported for the hypothesis that the difference is zero.
15. The regressions are estimated over the pooled sample (1991, 1993, and 1997) for men 40 years and older. Sample size is 2359.
16. "RE" refers to the random effects specification, while "FE" refers to the (individual) fixed effects specification. The Hausman Test column reports the results of the Hausman test of random versus fixed effects (p-values in parentheses).
17. The dependent variables for the labour supply measures are: Hours (hours worked in the previous year, including zero); Participation (whether the individual worked at all the previous year); H12 (self-reported health in the top two categories); Low BMI (BMI less than the 20^{th} percentile); and High BMI (BMI greater than the 80^{th} percentile).
18. The estimated effect of health on labour supply is the predicted impact of the change in health (H12, or H12 and BMI) multiplied by the estimated "structural" effect of health on labour supply (from Table 3A). OLS refers to predictions based on OLS coefficients, while IV refers to the 2SLS-based coefficients.

Appendix Table A4
Health and Labour Supply: Women over 40
(p-values in parentheses, * indicates statistical significance at the 5% level)

	Age 50 to 60		Age 60 to 70		Hausman Test
Reduced Form Age Profiles:	**FE**	**RE**	**FE**	**RE**	**RE vs. FE**
Hours	-422.6*	-119.2	-394.3*	317.4	11.6
	(0.000)	(0.621)	(0.035)	(0.341)	(0.398)
Participation	-0.116*	-0.007	-0.244*	-0.102	21.7*
	(0.000)	(0.908)	(0.000)	(0.192)	(0.027)
H12	-0.089*	0.051	-0.096	-0.140	16.9
	(0.031)	(0.600)	(0.192)	(0.298)	(0.111)
Low BMI	0.024	-0.030	0.019	-0.020	12.5
	(0.451)	(0.626)	(0.734)	(0.816)	(0.330)
High BMI	0.028	0.129*	-0.068	0.023	11.0
	(0.350)	(0.021)	(0.196)	(0.767)	(0.445)
The Estimated Effect of Health on Labour Supply:					
Change in hours:	-422.6	-119.2	-394.3	317.4	
H12 (OLS)	-10.8	3.1	-11.6	-8.6	
H12, BMI (OLS)	-9.3	8.1	-4.3	-11.4	
H12 (IV)	-151.7	85.9	-163.3	-236.0	
H12, BMI (IV)	-145.2	89.8	-156.6	-242.3	
Change in participation:	-0.116	-0.007	-0.244	-0.102	
H12 (OLS)	-0.001	0.001	-0.002	-0.002	
H12, BMI (OLS)	-0.001	0.006	0.000	-0.001	
H12 (IV)	-0.009	-0.002	-0.010	0.006	
H12, BMI (IV)	-0.008	0.007	-0.009	0.006	

Notes:

19. The estimated age profile is based on the difference in age coefficients from a regression of the dependent variable on covariates: age-group dummies (40-44, 45-49,…, 80+); year dummies for 1993, 1997; years of education; and province dummies. The Age 50 to 60 "profile" is the difference between the age 60-64 and age 50-54 coefficients, while the Age 60 to 70 "profile" is the difference between the age 70-74 and age 60-64 coefficients. P-values are reported for the hypothesis that the difference is zero.
20. The regressions are estimated over the pooled sample (1991, 1993, and 1997) for women 40 years and older. Sample size is 2159.
21. "RE" refers to the random effects specification, while "FE" refers to the (individual) fixed effects specification. The Hausman Test column reports the results of the Hausman test of random versus fixed effects (p-values in parentheses).
22. The dependent variables for the labour supply measures are: Hours (hours worked in the previous year, including zero); Participation (whether the individual worked at all the previous year); H12 (self-reported health in the top two categories); Low BMI (BMI less than the 20th percentile); and High BMI (BMI greater than the 80th percentile).
23. The estimated effect of health on labour supply is the predicted impact of the change in health (H12, or H12 and BMI) multiplied by the estimated "structural" effect of health on labour supply (from Table 5A). OLS refers to predictions based on OLS coefficients, while IV refers to the 2SLS-based coefficients.

Appendix Table A5
Structural Models: The Effect of Health on Labour Supply
Men 40 and Older
(absolute *t*-values in parentheses, * indicates statistical significance at the 5% level)

	Random Effects		Fixed Effects	
	OLS	**IV**	**OLS**	**IV**
Hours:				
H12	211.1*	999.4*	194.3*	2999.1*
	(3.0)	(2.7)	(2.1)	(2.8)
Wife's H12	-154.5*	-285.7	-76.7	-1419.3
	(2.3)	(0.7)	(0.9)	(1.6)
Low BMI	-106.7	-5.8	-285.3*	-234.0
	(1.6)	(0.1)	(2.5)	(1.0)
High BMI	-80.3	-148.5	54.0	147.4
	(1.0)	(1.3)	(0.4)	(0.4)
Participation:				
H12	0.049*	0.109	0.050*	0.158
	(2.9)	(1.2)	(2.3)	(0.8)
Wife's H12	-0.051*	-0.139	-0.034	-0.306
	(3.1)	(1.5)	(1.6)	(1.8)
Low BMI	-0.017	0.004	-0.069*	-0.034
	(1.0)	(0.2)	(2.5)	(0.7)
High BMI	-0.071*	-0.069*	-0.031	-0.032
	(3.5)	(2.6)	(0.9)	(0.5)

Notes:

9. The estimated coefficients are the coefficients from a regression of labour supply (hours or participation) on health variables (H12 = good health; wife's H12; and own Low and High BMI indicators), plus covariates (age group dummies (the same as table 2A), years of education, year and province dummies. Absolute *t*-values for the coefficients are shown in parentheses.
10. "OLS" is estimated on the full pooled sample of men 40 and older (1991,1993,1997), and is estimated by Random and Fixed Effects. The sample size is 2359.
11. "IV" is 2SLS estimation for the two available years (1991 and 1993) with an index of physical "infirmities" for the individual and spouse as instruments for individual and spouse H12. The sample size is 1395. The IV specification is also estimated by Random and Fixed Effects.
12. The *t*-value for the difference between the OLS and IV own H12 coefficients (i.e., the Hausman Tests) are (1) For RE Hours, 2.17; (2) For FE Hours, 2.63; (3) For RE Participation, 0.67; (4) For FE Participation, 0.55.

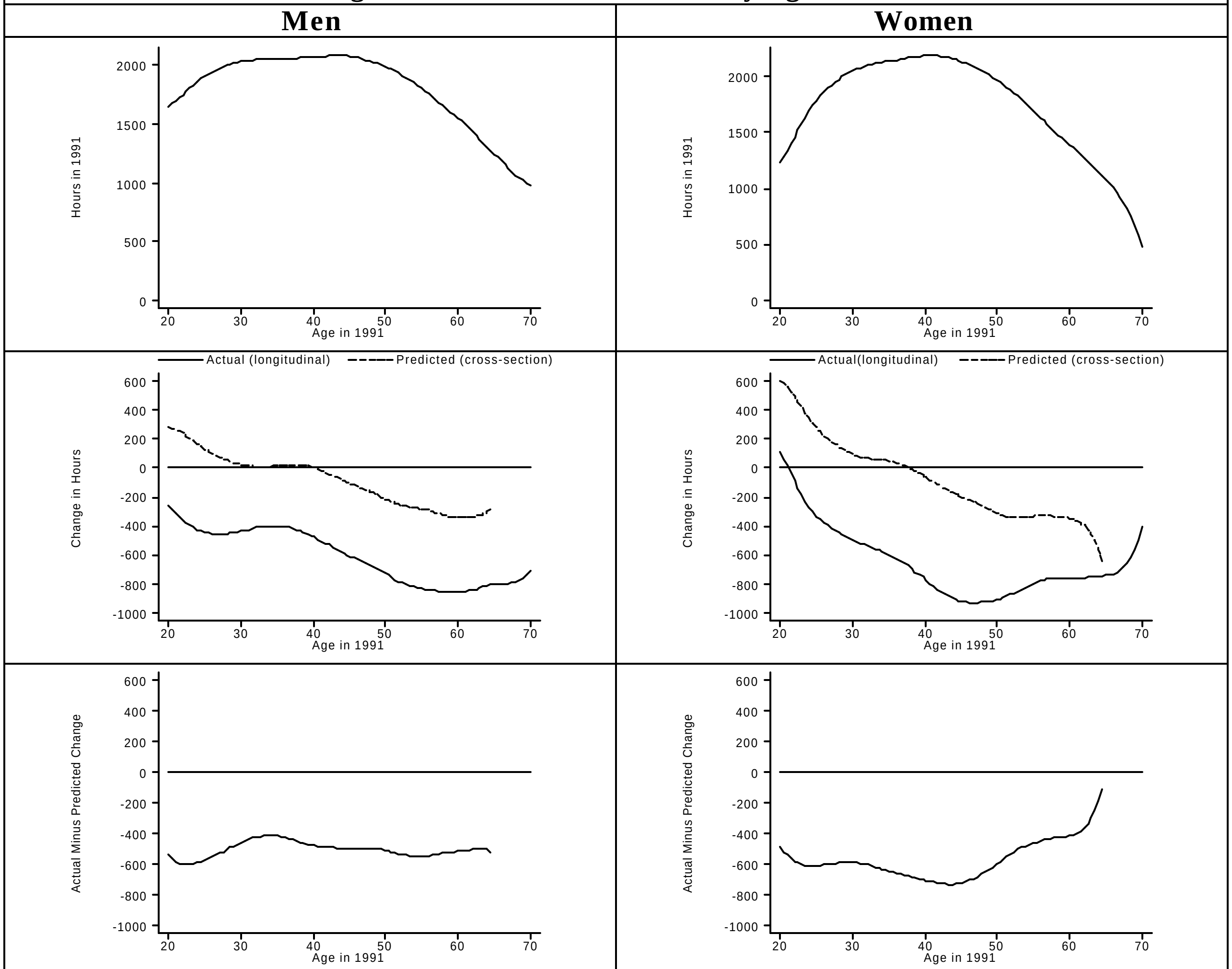
Appendix Table A6
Structural Models: The Effect of Health on Labour Supply
Women 40 and Older
(absolute *t*-values in parentheses, * indicates statistical significance at the 5% level)

	Random Effects		Fixed Effects	
	OLS	**IV**	**OLS**	**IV**
Hours:				
H12	120.5	1698.5*	61.2	1682.1
	(1.7)	(3.5)	(0.7)	(1.7)
Husband's H12	-123.5	-1134.5*	21.5	-2625.1
	(1.7)	(3.0)	(0.2)	(1.9)
Low BMI	141.4*	287.3*	253.7*	482.8
	(2.1)	(2.6)	(2.2)	(1.4)
High BMI	-67.6	-17.6	97.4	142.7
	(1.0)	(0.2)	(0.8)	(0.5)
Participation:				
H12	0.016	0.105	0.016	-0.042
	(1.0)	(1.0)	(0.8)	(0.2)
Husband's H12	-0.022	-0.036	-0.021	-0.287
	(1.3)	(0.5)	(0.9)	(1.2)
Low BMI	0.021	0.047*	0.003	0.091
	(1.2)	(2.0)	(0.1)	(1.5)
High BMI	-0.016	-0.002	0.041	0.096
	(0.9)	(0.1)	(1.4)	(1.7)

Notes:

13. The estimated coefficients are the coefficients from a regression of labour supply (hours or participation) on health variables (H12 = good health; husband's H12; and own Low and High BMI indicators), plus covariates (age group dummies (the same as table 4A), years of education, year and province dummies. Absolute *t*-values for the coefficients are shown in parentheses.
14. "OLS" is estimated on the full pooled sample of women 40 and older (1991,1993,1997), and is estimated by Random and Fixed Effects. The sample size is 2159.
15. "IV" is 2SLS estimation for the two available years (1991 and 1993) with an index of physical "infirmities" for the individual and spouse as instruments for individual and spouse H12. The sample size is 1190. The IV specification is also estimated by Random and Fixed Effects.
16. The *t*-value for the difference between the OLS and IV own H12 coefficients (i.e., the Hausman Tests) are (1) For RE Hours, 3.29; (2) For FE Hours, 1.64; (3) For RE Participation, 0.86; (4) For FE Participation, 0.28. For the effect of husband's H12 (on hours), the RE Hausman test is 2.72, and for FE it is 1.92.

Figure 1: Annual Hours Worked by Age in 1991



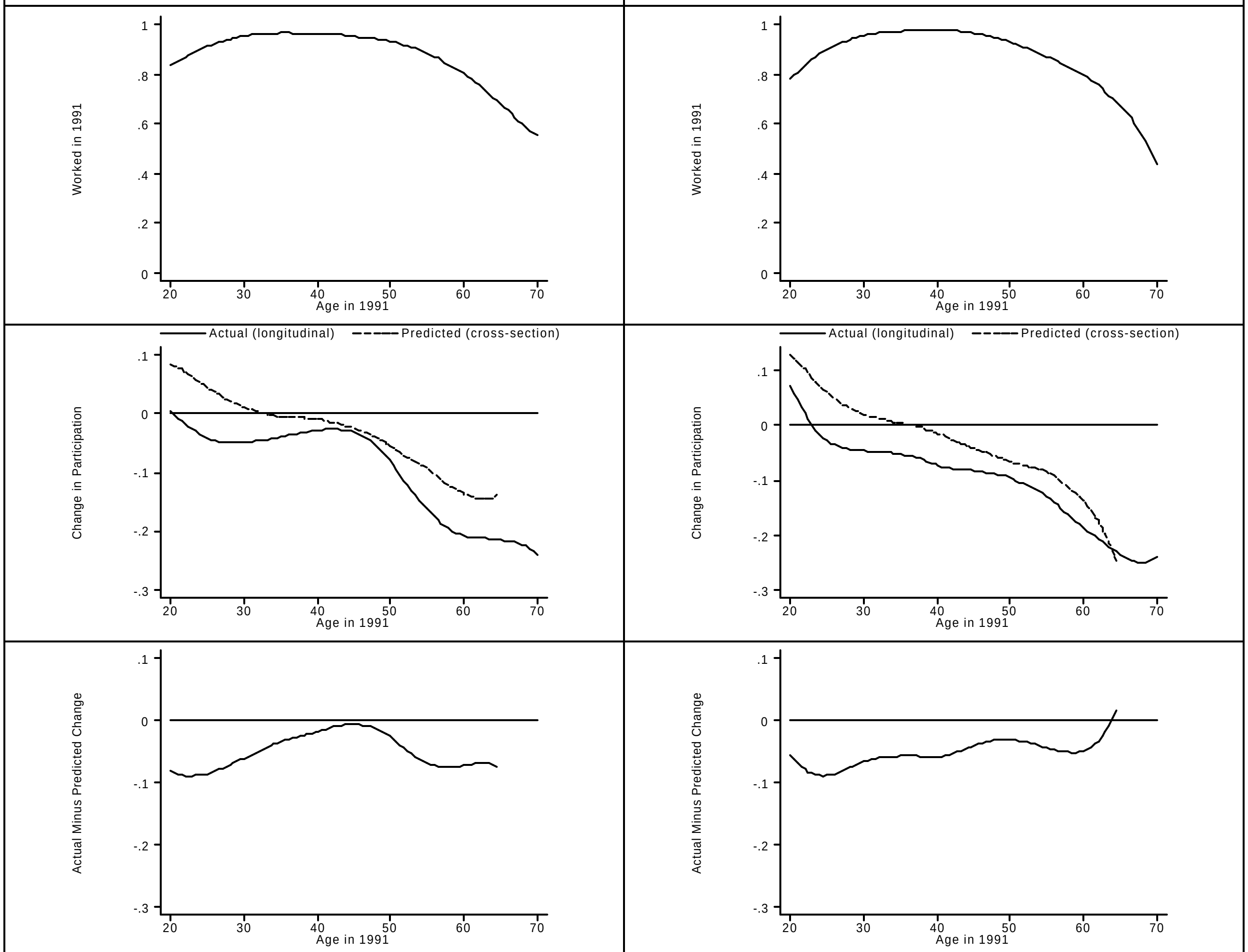
Source: CHNS, 1991,1993,1997

Notes:

1. Each figure represents a non-parametric simple regression of the dependent variable ("y-axis") against the independent variable ("x-axis");
2. The top panels show the 1991 cross-section relationship between hours (including zeroes) and age, while the bottom panels follow the panel households between 1991 and 1997, plotting the changes in hours worked (between 1997 and 1991) and age in 1991.

Figure 2: Proportion Working During Previous Year by Age in 1991

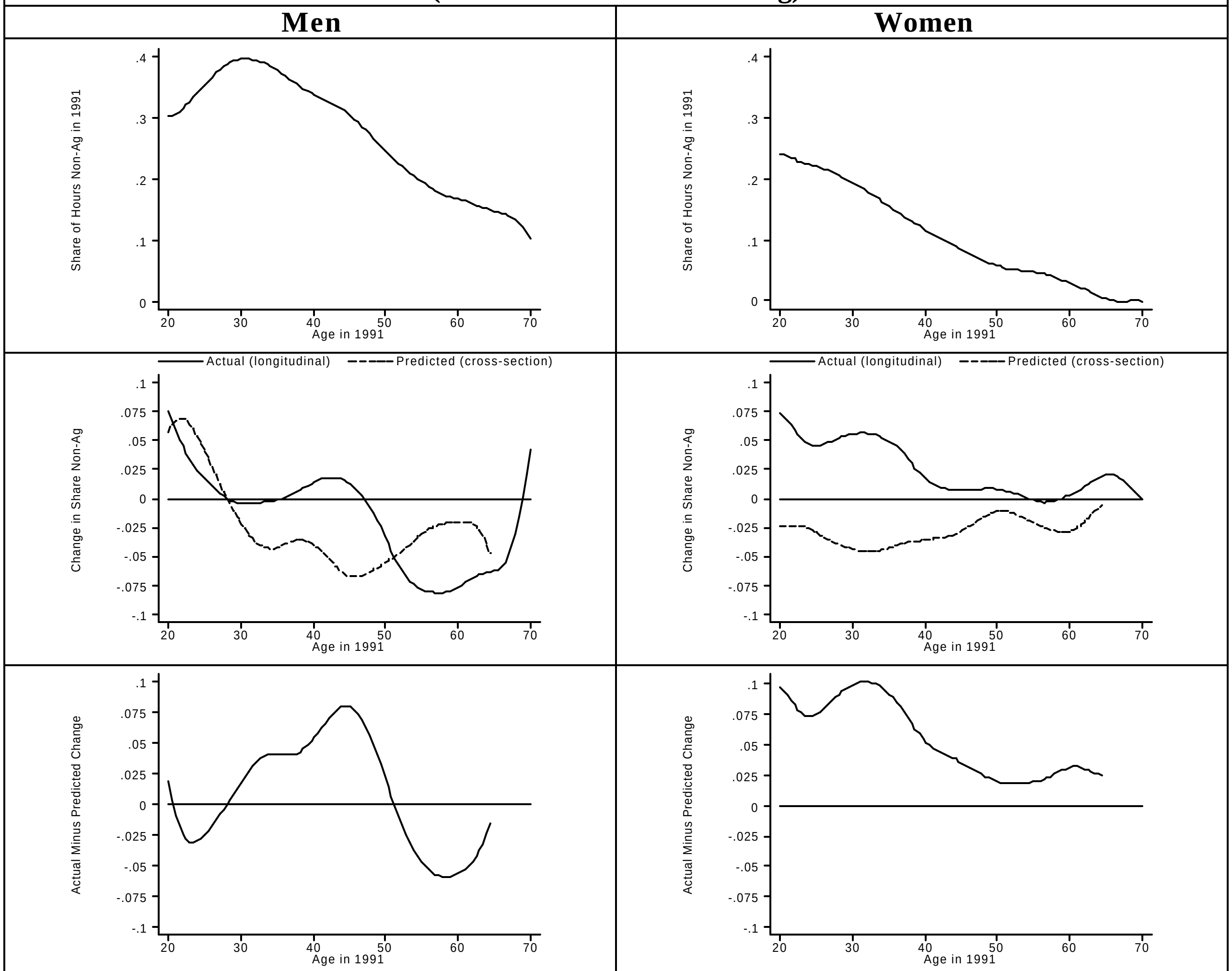
Men	**Women**



Source: CHNS, 1991,1993,1997
Notes:

1. Each figure represents a non-parametric simple regression of the dependent variable ("y-axis") against the independent variable ("x-axis");
2. The top panels show the 1991 cross-section relationship between participation (positive hours worked) and age, while the bottom panels follow the panel households between 1991 and 1997, plotting the changes in participation (between 1997 and 1991) and age in 1991.

Figure 3: Proportion of Work Hours in Non-Agricultural Employment (CONDITIONAL on working)



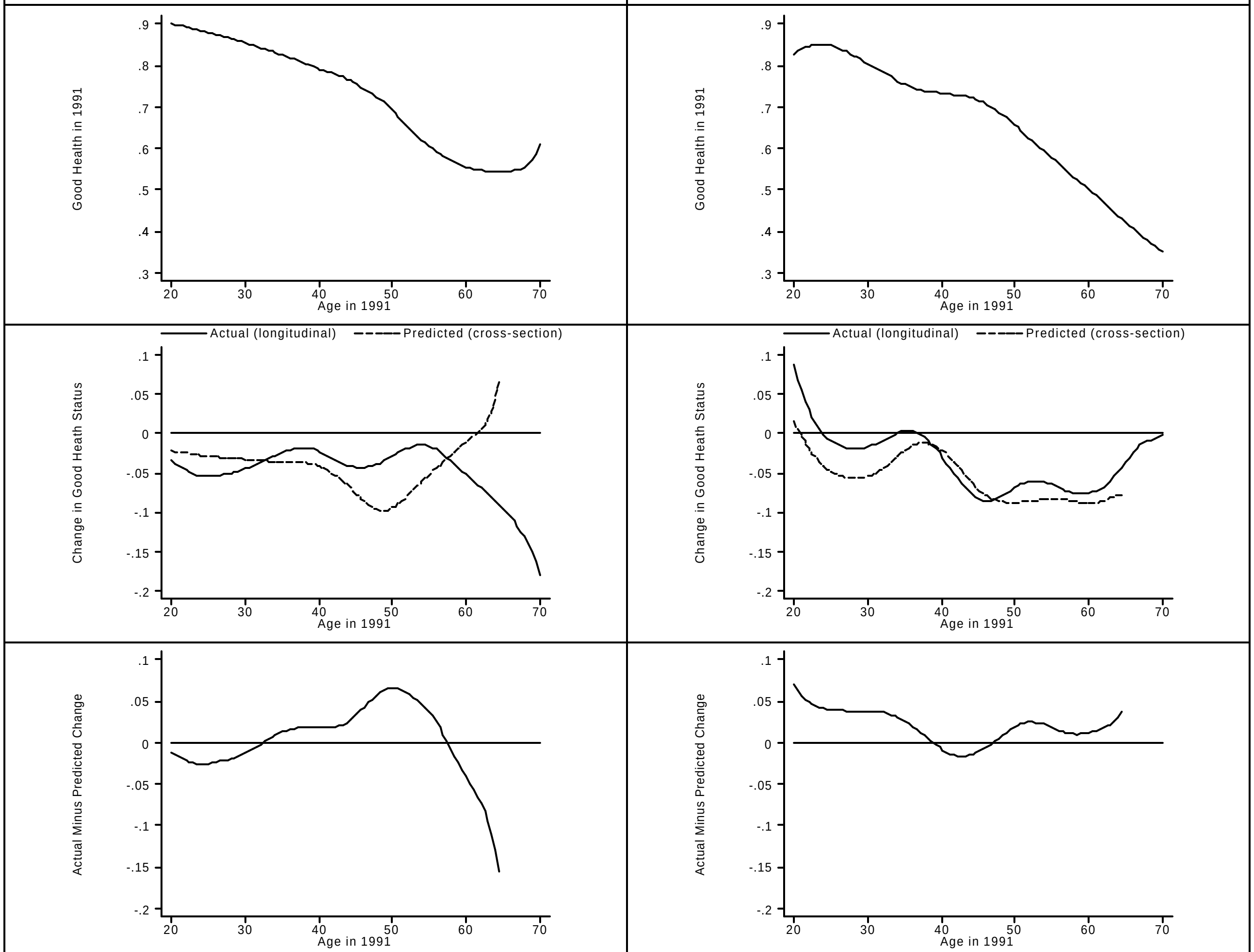
Source: CHNS, 1991,1993,1997

Notes:

1. Each figure represents a non-parametric simple regression of the dependent variable ("y-axis") against the independent variable ("x-axis");
2. The top panels show the 1991 cross-section relationship between the share of work hours in non-agricultural employment (off-farm, or family run businesses, conditional on positive hours worked) and age, while the bottom panels follow the panel households between 1991 and 1997, plotting the changes in proportion of time spent in non-agricultural pursuits (between 1997 and 1991) and age in 1991.

Figure 4: Self-Reported Health Status by Age in 1991

Men	**Women**



Source: CHNS, 1991,1993,1997

Notes:

1. Each figure represents a non-parametric simple regression of the dependent variable (“y-axis”) against the independent variable (“x-axis”);
2. The top panels show the 1991 cross-section relationship between being in “good health” (Self-reported health status in top two categories, i.e., “H12”) and age, while the bottom panels follow the panel households between 1991 and 1997, plotting the changes in proportion in good health (between 1997 and 1991) and age in 1991.