University of Hawai'i
Department of Economics
Working Paper Series

Saunders Hall 542, 2424 Maile Way, Honolulu,
HI 96822 · Phone: (808) 956-8496
economics.hawaii.edu

Working Paper No. 11-17

Earnings Growth and Movements in Self-Reported Health

by
Timothy J. Halliday

December 6, 2011

Earnings Growth and Movements in Self-Reported Health

Timothy J. Halliday*

University of Hawai'i at Mānoa and IZA

December 6, 2011

Abstract

We employ data from the Panel Study of Income Dynamics to investigate income to health causality. To account for unobserved heterogeneity, we focus on the relationship between earnings *growth* and *changes* in self-reported health status. Causal claims are predicated upon appropriate moment restrictions and specification tests of their validity. We find evidence of Granger-type causality running from income to health for married men but not for women

*Address: 2424 Maile Way; Saunders Hall 533; Honolulu, HI 96822. Tele: (808) 956-8615. E-mail: halliday@hawaii.edu. This is a substantially revised version of an earlier paper titled, "Income Volatility and Health." The usual disclaimer applies.

or single men. These effects are more pronounced for younger men and the poor. The former may be the consequence of permanent earnings shocks.

Key words: Gradient, Health, Dynamic Panel Data Models

JEL: I0, I12, J1

1 Introduction

The relationship between economic circumstances and health or the gradient has been the subject of academic inquiry for quite some time. While these investigations have documented a strong positive correlation between socioeconomic status (SES) and health in a variety of contexts, they have failed to produce a consensus among scholars concerning the underlying causal pathways. Typically, economists have tended to champion the causal pathway from health to income (*e.g.* Smith (1999), Smith (2004), Adams, Hurd, McFadden, Merrill, and Ribeiro (2003)). On the other hand, public health experts and epidemiologists have tended to advocate the reverse causal pathway from SES to health (e.g. Marmot, Smith, Stansfield, Patel, Head, White, Brunner, and Feeney (1991)). Still others such as Fuchs (1982) have suggested that this correlation may have less to do with causality *per se* than it does with a selection mechanism in which certain personality traits lead to similar economic and health outcomes.

To unravel this correlation, much of this literature has focused on exercises in the spirit of Granger causality tests. One common approach can be found in Smith (1999), Smith (2004), and Adams, Hurd, McFadden, Merrill, and Ribeiro (2003) in which health outcomes are regressed on a battery of controls for SES while controlling for demographic characteristics and lagged or baseline health outcomes. These studies typically look at relationships either between levels of health and SES or between changes in health and levels of SES. In predominately older populations, these papers show no evidence of causality running from income to health.

While these studies have provided some very important and interesting insights into the gradient, methodologically, there are several areas where some improvements can be made. First, because these studies do not relate changes in SES with changes in health status, they do not adequately adjust for unobserved time-invariant characteristics or unobserved heterogeneity that may be associated with both health and income. In fact, exercises of this type are somewhat rare. In a comprehensive survey of 3393 articles screened, Gunasekara, Carter, and Blakely (2011) found that only 13 compared changes in health outcomes with changes in SES. Second, as discussed by Arellano and Honore (2001), claims that the parameters of dynamic models are causal typically are predicated on moments that restrict the dynamics of the model in meaningful ways. Importantly, these restrictions have testable implications. This

has also, to a large extent, been ignored.

In this paper, we employ data from the Panel Study of Income Dynamics (PSID) to investigate income to health causality while taking these issues into account. Specifically, we focus on the relationship between earnings *growth* which has been the subject of a large literature in labor economics (*e.g.* Abowd and Card (1989); Meghir and Pistaferri (2004)) and *changes* in Self-Reported Health Status (SRHS) which has received less attention. We employ appropriate moment restrictions and specification tests of their validity. Earnings and SRHS are the subjects of our analysis because unlike most other measures of health or SES (e.g. education, wealth, chronic health conditions, mortality), they exhibit meaningful time series variation.

We find evidence of a causal relationship running from earnings to self-reported health status for married men but not single men or women. These effects are present at the earlier stages of the life-cycle but not the later parts of it. This may reflect that differenced income can be viewed as a proxy for a permanent earnings innovation. Given this, we should expect larger effects of earnings on health for younger people. We also find that the effects of income tend to be strongest in the bottom quartile of the earnings distributions and that their effects decline monotonically with income.

The balance of this paper is laid out as follows. In the next section, we discuss our data. After that, we discuss our estimation methods. This is followed by a

discussion of our results. We then compare our results to results in several papers in the literature that are related to ours. Finally, we conclude.

2 Data

We use a sample from PSID waves 1984 to 1993 of people ages 25 to 60 (inclusive). We chose this age range as these are the ages that people are most likely to be in the labor force. Our measure of health is SRHS which is a five-point categorical variable used to assess a survey respondents health status (1 = excellent; 2 = very good; 3 = good; 4 = fair; 5 = poor). Our income measure is labor income which includes all money earned from (the labor part of) farm and business income, wages, bonuses, overtime, commissions, professional practice, and income from boarders. This is the same measure that was used in Meghir and Pistaferri (2004). We did not use data prior to 1984 because the SRHS question was not available prior to this year. As in Meghir and Pistaferri (2004), we did not use data beyond 1993 because our particular measure of labor income was not available after 1993 and we wanted to maintain consistency in all of our data across survey years. Also, following Meghir and Pistaferri (2004), we did not drop the Survey of Economic Opportunity (SEO) which is an over-sample of economically disadvantaged individuals. Our reasons for

doing so were twofold. First, our estimations are all in first differences which purges the model of fixed effects which, thus, ameliorates the initial conditions problem. Second, our estimations place large demands on the data and, so dropping the SEO would have greatly reduced our sample sizes which is something that we could not afford. Descriptive statistics and variable definitions can be found in Table 1.[1] Step-by-step details on the sample construction can be found in the appendix.

Our main justification for emphasizing labor income over other measures of SES is that it exhibits more variation over time than many other correlates of economic status such as education and wealth. This temporal variation is crucial in any study that seriously aims to control for unobserved heterogeneity. For married people, we acknowledge that there are issues concerning whether labor income is the most appropriate measure if the spouse is the main breadwinner. To address this, we run models that include own and spousal income for married people.

We map the five point categorical SRHS variable into a binary variable. Values of SRHS of four or five get mapped into zero and values of one, two or three get mapped into a one. This partition is standard in the literature, although note that, after our mapping, *higher* values denote better health (*i.e.* the ones are healthier

[1] Note that because we dropped the SEO, the percentage of blacks in the data is higher than in the US population.

than the zeros). The reason why we do not work with the five-point variable is that doing so would have required using a nonlinear ordered model with fixed effects and predetermined regressors. In our view, estimation and identification of this class of models is not fully developed in the literature. For more on this issue, we refer the reader to next section.

Finally, while the SRHS variable can be criticized for being subjective, it does have several merits. First, unlike mortality or indicators of many chronic ailments, SRHS does vary substantially over time. Such variation is crucial in a study of the relationship between earnings growth and changes in health status. Moreover, Smith (2003) has shown that this variation in SRHS is correlated with changes in more objective measures of health. Second, it has been shown that many other more objective measures of health are not without flaws. For example, Baker, Stabile, and Deri (2004) compared self-reported measures of specific objective conditions, which are prone to errors, to their counterparts from medical records, which are not. They found that there was a large degree of measurement error in the objective measures which was also correlated with economic outcomes.

3 Estimation Equation

Denoting log labor income by y_{it}, the binary SRHS variable by h_{it}, and age by a_{it}, we consider the following model:

$$h_{it} = \alpha_i^H + \gamma^H h_{i(t-1)} + \beta^H y_{it} + \phi^H a_{it} + \upsilon_{it}^H \tag{1}$$

for $i = 1, ..., N$ and $t = 1, ...T$. This equation accounts for unobserved heterogeneity in the constant term, dynamics which operate via the lagged dependent variables, causality from income to health, and aging. To purge the model of unobserved heterogeneity, we will work with the model in first-differences:

$$\Delta h_{it} = \gamma^H \Delta h_{i(t-1)} + \beta^H \Delta y_{it} + \phi^H + \Delta \upsilon_{it}^H. \tag{2}$$

This will address any bias associated with time-invariant characteristics that are correlated with both health and income. For married people, we will also work with a modified version of equations (1) and (2) that includes spousal income which we denote by y_{it}^{sp}. Identification of the model in equation (1) will require restrictions on the timing of how income and health are allowed to affect each other.[2]

[2]The model that we work with is, in many ways, consistent with equation (1) in Smith (2004), but differs somewhat from Adda, Banks, and von Gaudecker (2009). The latter employ the permanent-transitory model that has become the standard model of earnings progression in the labor literature (e.g. Abowd and Card (1989) and Meghir and Pistaferri (2004)). It is important to note that while there is a preponderance of evidence suggesting that the permanent-transitory model is appropriate for earnings, there is much less of a consensus on how to model the dynamics

An alternative modeling strategy would be to employ "stress models" of health which are essentially the permanent-transitory model of earnings applied to health as in Halliday (2011), Adda, Banks, and von Gaudecker (2009) and Deaton and Paxson (1998). This avenue would require modeling the effects of permanent and transitory income shocks on health and *vise versa*. Identification would require an exclusion restriction as in Adda, Banks, and von Gaudecker (2009). However, it is very important to note that a major difference between their paper and this paper is that they estimate the model at the cohort level, whereas we estimate everything at the individual level. As such, their exclusion restriction is that permanent shocks to health at the cohort level do not affect income also at the cohort level. Importantly, they remain completely agnostic about causal pathways at the individual level. Consequently, any attempt to employ the permanent-transitory model of health here would require an assumption that is stronger than in Adda, Banks, and von Gaudecker (2009). This approach would require producing an exclusion restriction that is defensible at the same time that it provides moments that are informative of the model's parameters (*i.e.* avoids weak instruments). Our primary justification of estimating equation (1) is that identifying its parameters

of health. The model that we consider here is essentially a linear version of the models considered in Contoyannis, Jones, and Rice (2004a) and Contoyannis, Jones, and Rice (2004b), except that we will allow for feedback from income to health whereas the others do not.

requires moments that we believe to be both defensible and informative. We were not so sure that this would have been the case with the "stress model."

3.1 Moment Restrictions

Identification of β^H is achieved by restricting the causal ordering between health and income through assumptions on the model's residuals. The strongest assumption that we can make is that earnings is strictly exogenous. Specifically, if we adopt the notation that $z_i^t \equiv (z_{i1}, ..., z_{it})'$, then strict exogeneity requires that

$$E\left[\upsilon_{it}^H | h_i^{t-1}, y_i^T\right] = 0. \tag{3}$$

Assumption (3) says that innovations to health are uncorrelated with income at all leads and lags. This precludes any possibility that health today will affect earnings tomorrow or beyond. While this assumptions is strong, it does provide a useful benchmark.

A weaker assumptions is that

$$E\left[\upsilon_{it}^H | h_i^{t-1}, y_i^t\right] = 0. \tag{4}$$

Assumption (4) implies that the residuals at time t in the health equation are uncorrelated with income through time t. This assumption says that income is predetermined. It imposes a particular causal ordering on health and income in which income at time t is allowed to cause health at time t which is, in turn, allowed to cause income at time $t+1$. Importantly, it precludes any contemporaneous causality from health to income. This is a testable implication.

3.2 GMM Estimation

We will employ a GMM procedure to estimate the model's parameters. If we invoke the strict exogeneity assumptions, then no instruments are needed for income. However, instruments are still needed for the lagged dependent variables and so we will use h_i^{t-2} as instruments for $\Delta h_{i(t-1)}$. If we invoke the predeterminedness assumption, then we must also instrument for income and, so we will use y_i^{t-1} as instruments for Δy_{it}. We will follow the standard practice of reporting the one-step estimates as Arellano and Bond (1991) show that the two-step procedure has poor finite sample properties.

A final issue that we should address is that of using too many instruments. When using estimators of this type, there is a tendency for instrument proliferation as the

number of instruments will increase at a rate that is quadratic in T. As discussed by Roodman (2009), the fundamental issue here is that when there are too many instruments relative to the sample size, the R^2 on the first stage will approach unity and so the second stage estimator which be almost equivalent to OLS. To address this critique, when using the predetermined assumption, we cap the maximum number of lags that can be used as instruments at three.

3.3 Specification Tests

Arellano and Bond (1991) discuss several specification tests for dynamic panel data models such as those in equation (1). One test centers on the fact that the predetermined assumption restricts the serial correlation in the residuals. In particular, assumption (4) implies that

$$E\left[\Delta v_{it}^{H} \Delta v_{it-j}^{H}\right] = 0 \text{ for } j > 1. \tag{5}$$

Arellano and Bond (1991) develop a test that has a standard normal distribution when the null in equation (5) is true. We call this test m_2 following their notation. In addition, if the residuals are highly persistent so that they have close to a unit root, then a test based on m_2 will have no power. To address this, they note that

unit root residuals imply that

$$E\left[\Delta v_{it}^{H} \Delta v_{it-1}^{H}\right] = 0. \tag{6}$$

They also develop a test of the null implied by equation (6). This statistic is called m_1 (again following their notation). If m_1 is statistically different from zero and m_2 is not then this is a necessary condition for the model to be properly specified. The second is Hansen's overidentification test which is based on the Sargan statistic. This test statistic, which we call J, will have a chi-squared distribution when all of the overidentifying restrictions are valid.[3]

3.4 Assessing the Linearity Assumption

One important issue with equation (1) that we must address is that we adopted a linear model as opposed to a non-linear latent variable model. The only estimator of a non-linear latent variable model that allows for predetermined variables and unobserved heterogeneity that we know of is Arellano and Carrasco (2003). A major reason why we do not believe that this estimator would be appropriate for our

[3]It is important to note that evidence from Andersen and Sorensen (1996) and Bowsher (2002) suggests that the test has low power when the number of instruments is too large relative to the sample size. To mitigate this issue, we have truncated the number of instruments as discussed above.

application is that it requires observation of the entire history of the variables to be correctly specified. If it is not (which is the case in the PSID), then the model will be egregiously mis-specified as a mixture of normals will be assumed to be simply a normal distribution.

One defense of the linearity assumption is that often binary choice models have conditional expectations that are linear over a large portion of the support of the index. We offer some suggestive evidence for this in Figures 1 through 4 where we plot fitted values from linear and probit regressions of the binary SRHS variable onto an age trend. While the figures do show some curvature in the fitted values from the probits, they are close to linear as evidenced by their proximity to the fitted OLS values.

4 Empirical Results

Throughout this section, we will estimate our models using four demographic subsamples: single men (SM), single women (SW), married men (MM), and married women (MW).[4]

[4]Note that these groups are not mutually exclusive as some people were single for parts of their duration in the PSID but married for others. As such, the sample sizes reported at the bottom of the tables in this section sum to a number that is greater than 6447 which is size of the sample that we report in the Appendix.

4.1 Autocorrelations in Income and Health

We begin by reporting estimates of autocorrelations of earnings growth and changes in SRHS which is an exercise that is common in the earnings dynamics literature (*e.g.* Abowd and Card (1989) and Meghir and Pistaferri (2004)). Results for singles are reported in Table 2 and for couples in Table 3. In order to make comparisons between income and health, we report correlations and not covariances.

The tables reveal a number of interesting patterns in the data. We see strong negative serial correlation in both income and health changes at the first order. While negative serial correlation in earnings growth has been well-established (*e.g.* Abowd and Card (1989) and Meghir and Pistaferri (2004)), negative serial correlation in changes in SRHS is relatively less established. If we were to have adopted the permanent-transitory model of earnings to health (*i.e.* the "stress model"), then these results are strongly indicative of an important role for transitory components in the health process. It is also interesting to note that the first-order autocorrelations are of a larger magnitude for health than for income. Next, the autocorrelations are not significant at the 5% level or greater past the first order for all demographic groups except for married men. This is suggestive that the residuals in equation (2) will be serially uncorrelated for orders greater than one. Finally, we note that the

autocorrelations in both health and income taper to zero. This suggests that there are no heterogeneous trends in neither health nor income. The result for income is consistent with Abowd and Card (1989) and Meghir and Pistaferri (2004). Also, the result for health is consistent with Halliday (2008).

4.2 OLS

Before we discuss the GMM estimation results, we present results for OLS estimation of the first-differenced health model in equation (2) in Table 4 as a benchmark. First, we see that the estimates of lagged health are all negative and highly significant. This is a reflection of the negative serial correlation shown in Tables 2 and 3 since OLS uses

$$E\left[\Delta\upsilon_{it}^{H}\Delta h_{i(t-1)}\right]=0 \tag{7}$$

as a moment condition. The GMM procedures do not use this as a moment and so we will see that the estimates of the lagged health coefficients for these procedures are generally positive as one would expect. Reflective of this large and negative estimate, we see that the R^2's are all large (i.e. around 20%) across demographic groups.

Moving to the estimates of the income coefficients, we see a positive and significant

estimate for men but insignificant estimates for women. For single men, the estimate of 0.038 suggests that a 1% increase in labor income is associated with an increased probability of being in the top three SRHS categories of 0.00038 or 0.038 percentage points.[5] For married men, the coefficient estimate is almost half as small at 0.02.

For couples, it is reasonable to suspect that own and spousal income might matter, particularly, for women. Interestingly, for married men, we see that a 1% increase in spousal income is associated a small but statistically significant reduction in the propensity to be in the top three SRHS categories of 0.001 percentage points. One interpretation of this result is that an increase in a wife's income might be a marker of financial distress within the household. For example, in our sample, we see that a 1% increase in the wife's income is associated with a 0.0023% reduction in the husband's income and this estimate is statistically significant ($t = -2.61$).[6] While it is true that this point-estimate is very small, it is commensurate with the small estimate of the effects of spousal income on health that we see for married men. The coefficient on spousal income for married women is positive but not significant.

[5] This level-log specification is somewhat uncommon. The proper interpretation of β in the model

$$y = \alpha + \beta \log x + u$$

is

$$\Delta y = \beta \frac{\Delta x}{x} = (\beta/100)\,\%\Delta x.$$

So, β can be interpreted as the effect of a 1% change in earnings on the probability of being in good health in percentage points.

[6] This result is not reported but is available from the author upon request.

4.3 Arellano-Bond Estimates

4.3.1 Strictly Exogenous Income

We now discuss the Arellano-Bond results obtained using the strict exogeneity assumption. These are reported in Table 5. First, the tests of serial correlation in the differenced residuals (*i.e.* m_1 and m_2) pass. We see that the differenced residuals are strongly negatively serially correlated at one lag but exhibit no serial correlation at higher lags. In general, the tests of serial correlation in the residuals perform quite well for all of our Arellano-Bond results. All of the overidentification tests pass at the 10% level, although we do reject the null at the 5% level in column 6 for married women. Note that the strict exogeneity moment condition uses no overidentifying restrictions for income. In other words, since is income is assumed strictly exogenous, β^H is identified from the moment

$$E[\Delta v_{it}^H \Delta y_{it}] = 0. \tag{8}$$

As such, the Sargan statistic here is not an indication of the validity of the strict exogeneity assumption.

We now turn to the parameter estimates. First, because we are no longer employing the moment condition in equation (7), we no longer see the highly significant

negative coefficients on lagged health. The estimates are positive and significant for single and married men and insignificant for everyone else. Second, we see very similar estimates of the income coefficients in Tables 4 and 5. Indeed, the estimates for single and married men are virtually the same in the two tables. The reason underlying this is that both estimators rely on the moment in equation (8). The only mildly noteworthy difference is that the estimate of the spousal income coefficient in column 6 for married women is positive and of borderline significance. This indicates that an increase in spousal income is modestly associated with better health for women.

4.3.2 Predetermined Income

We now turn to the results that use the predetermined assumption. These are reported in Table 6. We no longer see any effects of income for single men. While all three specification tests pass, it is important to bear in mind that the sample size is small ($N = 916$). So, this null results may be driven by the predetermined assumption resulting in less efficient estimates than the strict exogeneity assumption. Next, we see a very large *negative* estimate of the income coefficient for single women. However, a few issues must be borne in mind. First, its magnitude is implausibly large at -0.178. Second, we will see that the instruments in this estimation are weak

which is probably what is driving this odd estimate.

We now turn to the results for married men in columns 3 and 4 and married women in columns 5 and 6. For this group, the sample sizes are on the order of 3000 for men and 2000 for women and, so efficiency should be less of an issue. In column 3, for married men, we see that the estimate of β^H is now 0.054 and significant at the 10% level. In contrast, we saw that with the strict exogeneity assumption, the point-estimate was more than half as small at 0.022 but much more significant. Once again, this probably reflects that the estimates in Table 5 were more efficiently estimated than in Table 6. Moving to column 4 where we include spousal income, we see that own income is no longer significant, although its point-estimate is still larger at 0.039 than in column 4 of Table 5 where it is 0.020. In the same column, we see that the estimate on spousal income is -0.003 and significant at the 10% level. In contrast, in Table 5 where we invoke the strict exogeneity assumption, we saw that the estimate is smaller at -0.002 but much more significant. We see that the specifications tests perform well in both columns. In columns 5 and 6, we do not see any effects of own or spousal income on health for married women. The specification tests in these columns also perform well.

It is important to point out that there is a general pattern in which the significant estimates in Table 6 are larger than their counterparts in Table 5 where we assumed

strict exogeneity. This is most likely a result of there being less attenuation bias from measurement error using the predetermined assumption in equation (4) than the strict exogeneity assumption in equation (3). The reason for this is that the former assumption uses Δy_{it} as its own instrument, whereas the latter uses $y_{i(t-1)}$ as an instrument for Δy_{it}. In the first case, measurement error bias will be present in periods t and $t-1$, but in the second it will only be present in $t-1$.

We conclude this subsection by investigating the possibilities of non-linearities in the health income relationship. To do this, we consider a modified model that allows for a spline in earnings

$$h_{it} = \alpha_i^H + \gamma^H h_{i(t-1)} + \beta^H y_{it} + \sum_{q=25,50,75} \psi^q y_{it} * 1\left(y_{it} \in (y_q, y_{q+25}]\right) + \phi^H a_{it} + \upsilon_{it}^H \quad (9)$$

where y_q is the qth percentile for earnings. Note that $[y_0, y_{25}]$ is the omitted bracket. This specification allows the effect of income to change as it increases. The effect of income for the bottom quartile is β^H, for the second lowest quartile is $\beta^H + \gamma^{25}$, and so on.

Estimation results are reported in Table 9 for married men.[7] First, we see that the estimate of β^H increases from 0.054 to 0.11 once we account for the non-linearities

[7] We did not find noteworthy results for the other three demographic groups and so we do not report them.

in earnings. The reason why the estimate increases by so much is that the variable y_{it} is strongly positively correlated with the variables $y_{it} * 1\left(y_{it} \in (y_q, y_{q+25}]\right)$ which are negatively correlated with health once we partial out income. So, including the terms inside the summation in equation (9) increases the estimate of β^H. Second, we see that the effects of earnings diminish with higher incomes as the estimates of ψ^q are all negative and highly significant. This suggests that the largest effects of income on health can be found in the bottom quartile.

4.3.3 Weak Instruments

To investigate whether or not weak instruments is an issue, we consider the following equations:

$$\Delta y_{it} = \phi_0 + \phi_1 y_{i(t-1)} + \phi_2 y_{i(t-2)} + \phi_3 y_{i(t-3)} + u_{it}^Y \tag{10}$$

and

$$\Delta h_{it} = \eta_0 + \eta_1 h_{i(t-1)} + \eta_2 h_{i(t-2)} + \eta_3 h_{i(t-3)} + u_{it}^H. \tag{11}$$

These two equations, while not a formal and rigorous test for weak instruments, will shed light on the power of the information embedded in the moment condition in equation (4). Since we only used a maximum of three lags in the estimations, we also only include three lags in equations (10) and (11). As discussed in the weak

instruments literature, the conventional distribution theory for the F-statistic is no longer applicable. Instead, we will use the conventional wisdom of seeing if the F-statistic of the nulls that $H_0 : \phi_1 = \phi_2 = \phi_3 = 0$ and $H_0 : \eta_1 = \eta_2 = \eta_3 = 0$ is above ten. As a justification for this rule-of-thumb, we note that the 5% critical values for the case of three instruments in Table 1 of Stock, Wright, and Yogo (2002) are typically around ten.

The estimation results are reported in Table 8. In the top panel, the estimation of equation (10) reveals that weak instruments are probably an issue for single and married women as the F-statistics are 3.74 and 4.12, respectively. This helps to make sense of the odd estimate of β^H that we reported in column 2 of Table 6 for single women. In addition, while the estimates for married women in columns 5 and 6 of the same table were more reasonable, the low F-statistic for married women does indicate that some degree of caution should be taken with these estimates as well. Turning to single and married men, weak instruments do not appear to be a problem here as the F-statistics are 26.48 and 171.22, respectively. Finally, in the bottom panel of the table, we report estimates of equation (11) and we see that the F-statistics are all well above 100 indicating that weak instruments is not an issue when instrumenting for lagged health.

4.4 Interpreting the Estimates

The estimates of β^H in column 3 of Tables 5 and 6 suggest that a 1% increase in labor income results in an increase in the probability of being in the top three SRHS categories of between 0.022 and 0.054 percentage points. To help us better understand why these effects are so large, it is useful to consider to what extent the parameter β^H is determined by the effects of permanent and transitory income shocks. To fix ideas, we appeal to the permanent-transitory model of earnings from Abowd and Card (1989) and Meghir and Pistaferri (2004):

$$\Delta y_{it} = e_{it} + \Delta \varepsilon_{it} \tag{12}$$

where e_{it} and ε_{it} are permanent and transitory shocks to log earnings. Essentially, what equation (12) says is that Δy_{it} in equation (2) can be interpreted as a proxy for a permanent income shock. As such, one can then interpret β^H as, to some extent, identifying the effects of permanent earnings shocks on health status.

If this interpretation is correct, then one should expect to see larger effects of income on health for younger people since a permanent shock will persist for a shorter time for older people. To investigate this, we estimate the model in equation (1) for people age 45 and under and for people over 45. In Table 9, we present

estimates of β^H for the four demographic groups broken down by age. Notably, in column 3, we estimate the parameter for married men and we see that it is large and significant for younger men, 0.086 with a t-statistic of 2.85, but small and statistically indistinguishable from zero for older men.

4.5 Testing the Causal Ordering

We conclude with a final check on the causal ordering implied by the assumption in equation (4). As discussed above, the predetermined assumption implies that income at time t causes health at time t which, in turn, may affect income at $t+1$. However, it precludes any contemporaneous causality running from health to income. Consequently, another way to check the validity of the predetermined assumption is to test for a contemporaneous causal relationship running from health to income.

To do this, we consider a model similar to equation (1) except with the roles of health and income reversed. Specifically, we consider a model of the form

$$y_{it} = \alpha_i^Y + \gamma^Y y_{i(t-1)} + \beta^Y h_{it} + \phi^Y a_{it} + \upsilon_{it}^Y \tag{13}$$

in which

$$E\left[\upsilon_{it}^Y | y_i^{t-1}, h_i^t\right] = 0. \tag{14}$$

If we find a positive and significant estimate of β^Y in conjunction with passing specification tests, then this would provide strong evidence against the predetermined assumption that we invoked throughout this paper.

Estimation results are reported in Table 10. First, we see that, with the exception of married women in columns 5 and 6, that the overidentifcation tests do not perform well. However, the rejection of the overidentification tests may be a consequence of equation (13) being an inadequate model for income. Indeed, in results that we do not report, we do see that the Sargan statistics for estimation of equation (13) with health excluded also have low p-values. Second, we see that, with the exception of single women, that the estimates of β^Y are all small and insignificant. As we already discussed, the large negative estimate for single women is most likely due to weak instruments. So, given this model of earnings (which may not be the best model of earnings progression according to the Sargan statistics), there is no evidence of a contemporaneous causal relationship running from health to income. This lends additional (albeit somewhat weak given the Sargan statistics) evidence in favor of the assumption that income is predetermined in equation (1).

5 Relationship to the Existing Literature

We now compare our findings to two important papers in this literature. These papers were chosen because of their emphasis on investigating the income to health causal pathway. We will focus on the following papers: Smith (2004), who estimates similar parameters and finds conflicting results, and Adda, Banks, and von Gaudecker (2009), who estimate different parameters but use similar methods to ours.

5.1 Smith (2004)

Smith (2004) regresses an indicator for the onset of a chronic condition on a battery of proxies for SES including income, wealth, and education while controlling for past health status, a set of health risk factors, and a conventional set of demographic controls. To fix ideas, consider a model such as

$$\Delta h_{it} = \kappa_0 + \kappa_1 h_{i(t-1)} + \kappa_2 y_{i(t-1)} + X_i'\theta + v_{it} \tag{15}$$

where $v_{it} = \tau_i + \epsilon_{it}$ and the demographics are contained in X_i. If we restrict $\kappa_1 = 0$ but include baseline health variables in X_i then we would obtain something similar to the estimation models in Smith (2004). Identification stems from assumptions

such as

$$E\left[\tau_i|h_i^T, y_i^T\right] = E\left[\epsilon_{it}|h_i^{t-1}, y_i^{t-1}\right] = 0.$$

The idea behind this exercise is that once we partial out these other factors, the estimates of the SES coefficients (κ_2 in our case) can be interpreted as causal in some Granger sense. The restrictions imposed on ϵ_{it} are similar to our predeterminedness assumption (in fact, they are weaker). However, the restrictions on τ_i actually contradicts the specification in equation (15) provided that $\kappa_1 \neq -1$. In other words, consistent estimation of the parameters of this model precludes a fixed effect. He finds in both the Health and Retirement Survey and the PSID that the only component of SES that matters is education and that income never matters.

Our specification differs from equation (15) since identification in this paper is driven by how *differences* in income lead to *differences* in SRHS. Claims of causality are grounded in appropriate moment restrictions and specification tests of the validity of these restrictions. In contrast, Smith (2004) investigates how *levels* of income are associated with *changes* in health once one controls for a battery of possible confounders. No discussion of model specification is provided.

To provide the reader with some idea of how specification matters, we present Table 11 where we estimate the model in equation (15). In the odd numbered

columns, we estimate the model without lagged health (*i.e.* $\kappa_1 = 0$). We see that income never matters. In the even numbered columns, we include lagged health and we now see that income matters a lot as the t-statistics are always above five. The reason for this discrepancy is that lagged health is highly and negatively correlated with changes in health but is also positively correlated with income. So, omitting lagged health in the model results in a severe negative bias. One could have simply reported the results in columns (1) and (3) without regard to the model in columns (2) and (4) and claimed that the effects of income on health are null.[8]

5.2 Adda, Banks, and von Gaudecker (2009)

Adda, Banks, and von Gaudecker (2009) also employ dynamic panel data techniques to conduct tests of "Granger-type" causality. However, an important difference between their paper and ours is that their focus is at the cohort level, whereas ours

[8] There are other potential reasons for the divergence between our results and his results. First, poor efficiency due to small sample sizes may be an issue. For example, the PSID results were run on single cross-sections corresponding to baselines at 1984, 1989, and 1994. Without pooling data across cross-sections, there may be an inability to detect some effects. Second, the dependent variables are different in his and our studies. There is certainly some debate in Health Economics about which measures of health are best and we certainly do not claim that SRHS is the best. However, as discussed earlier, there is evidence suggesting that there are non-trivial measurement errors in self-reported objective measures of chronic conditions (see Baker, Stabile, and Deri (2004)). Moreover, there may not be sufficient variation in these measures, particularly at earlier ages where we believe the bulk of the effects to lie, to be able to detect meaningful income effects.

is at the individual level. Borrowing their notation, they consider a model like

$$\begin{aligned} Y_{ct} &= Y_{c(t-1)} + \zeta_{ct}^{Y} + \xi^{C,P}\zeta_{ct}^{H} + (1-L)\left[\varepsilon_{ct}^{Y} + \xi^{C,T}\varepsilon_{ct}^{H}\right] \\ H_{ct} &= H_{c(t-1)} + \zeta_{ct}^{H} + \gamma^{C,P}\zeta_{ct}^{Y} + (1-L)\left[\varepsilon_{ct}^{H} + \gamma^{C,T}\varepsilon_{ct}^{Y}\right]. \end{aligned} \tag{16}$$

where Y_{ct} and H_{ct} denote the income and health of cohort c at time t, ζ_{ct}^{Y} and ζ_{ct}^{H} are permanent shocks to income and health, and ε_{ct}^{Y} and ε_{ct}^{H} are transitory shocks to income and health. The authors focus on the parameter $\gamma^{C,P}$ which is the impact of permanent shocks to income on health. Identification stems from the assumption that $\xi^{C,P} = 0$. They find significant estimates $\gamma^{C,P}$. Similar to Ruhm (2000), this suggests that positive cohort-specific earnings shocks, such as economic booms, lead to increased mortality. However, they find no effects using measures of morbidities suggesting that the mortality results are driven by things such as vehicular and on-the-job accidents.[9]

It is crucial to point out again that this model is specified at the cohort level but remains completely agnostic about causal pathways at the individual level. In other words, they require no restrictions on the individual level parameters in their

[9]This is consistent with the analysis of Miller, Page, Huff-Stevens, and Filipski (2009). It also helps to reconcile the results in Ruhm (2000) from those in Sullivan and von Wachter (2009a), Sullivan and von Wachter (2009b), and Strully (2009) who show that job loss is associated with worse health outcomes since it suggests that the procyclicality of mortality rates might be driven by a mechanism other than job loss.

equations (2) and (3). In this sense, there is no contradiction between their results and ours.

6 Conclusions

In this paper, we estimated the effects of earnings growth on movements in self-reported health status using the Panel Study of Income Dynamics. We documented that changes in log income and self-rated health exhibit similar temporal patterns. In particular, both exhibit strong negative autocorrelation at one lag suggesting that transitory components may be important in both earnings and health. We then conducted tests for income to health causality (similar but not identical to Granger causality) and found evidence of a positive relationship between earnings growth and improvements in health for married men but not single men or women. These effects were only present for men younger than age 45 and strongest in the bottom quartile of the earnings distribution.

7 Appendix: Sample Selection

Initially, we started with 20,338 heads of household and their spouses who were in the PSID between 1984 and 1993. Next, we dropped people with incomplete information

on SRHS which dropped the sample size to 20,222. As in Meghir and Pistaferri (2004), we then dropped people whose first-difference log income was smaller than -1 or greater than 5.[10] This dropped the sample size to 18,073. Next, we kept people who were between ages 25 and 60 (inclusive) which left us with 14,670 individuals. We then dropped people whose ages declined by more than a year or increased by more than 2 years across years which brought the sample size to 12,899. Finally, we dropped people who were not in the panel continuously which further dropped the sample size to 10,502. Year-by-year sample sizes can be found in Table 12. Finally, we kept people who were in the panel for at three years which brought the sample size to 6447 individuals.

References

ABOWD, J., AND D. CARD (1989): "On the Covariance Structure of Earnings and Hours Changes," *Econometrica*, 57(2), 411–45.

[10] The magnitudes of the effects for men in this paper are larger than those in ?. Restricting the analysis to people whose differenced log-earnings was between -1 and 5 is what is driving these differences. The earlier version did not use this restriction. Using it, eliminates individuals whose incomes fluctuate too wildly from year-to-year and, thus, mitigates attentuation bias due to measurement errors.

Adams, P., M. Hurd, D. McFadden, A. Merrill, and T. Ribeiro (2003): "Healthy, Wealthy, and Wise? Tests for Direct Causal Paths between Health and Socioeconomic Status," *Journal of Econometrics*, 112(1), 3–56.

Adda, J., J. Banks, and H. von Gaudecker (2009): "The Impact of Income Shocks on Health: Evidence from Cohort Data," *Journal of the European Economics Association*, 7(6), 1361–99.

Andersen, T., and B. Sorensen (1996): "GMM Estimation of a Stochastic Volatility Model: A Monte Carlo Study," *Journal of Business and Economic Statistics*, 14(3), 328–352.

Arellano, M., and S. Bond (1991): "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies*, 58(2), 277–297.

Arellano, M., and R. Carrasco (2003): "Binary Choice Panel Data Models with Predetermined Variables," *Journal of Econometrics*, 115(1), 125–157.

Arellano, M., and B. Honore (2001): "Panel Data Models: Some Recent Developments," in *Handbook of Econometrics, vol. 5*, ed. by J. J. Heckman, and E. Leamer. North-Holland, Amsterdam.

Baker, M., M. Stabile, and C. Deri (2004): "What Do Self-Reported, Objective Measures of Health Measure," *Journal of Human Resources*, 39(4), 1067–93.

Bowsher, C. (2002): "On Testing Overidentifying Restrictions in Dynamic Panel Data Models," *Economic Letters*, 77(2), 211–220.

Contoyannis, P., A. Jones, and N. Rice (2004a): "The Dynamics of Health in the British Panel Survey," *Journal of Applied Econometrics*, 19(4), 473–503.

——— (2004b): "Simulation-Based Inference in Dynamic Panel Probit Models: An Application to Health," *Empirical Economics*, 29(1), 49–77.

Deaton, A., and C. Paxson (1998): "Aging and Inequality in Income and Health," *American Economic Review: Papers and Proceedings*, 88(2), 248–253.

Fuchs, V. (1982): "Time Preference and Health: an Exploratory Study," NBER Working Paper.

Gunasekara, F., K. Carter, and T. Blakely (2011): "Change in Income and Change in Self-Rated Health: Systematic Review of Studies using Repeated Measures to Control for Confounding Bias," *Social Science and Medicine*, 72(2), 193–201.

HALLIDAY, T. (2008): "Heterogeneity, State Dependence, and Health," *Econometrics Journal*, 11(3), 499–516.

——— (2011): "Health Inequality over the Life-Cycle," *The B.E. Journal of Economic Analysis and Policy*, 11(3), Article 5.

MARMOT, M., G. SMITH, S. STANSFIELD, C. PATEL, J. HEAD, I. WHITE, E. BRUNNER, AND A. FEENEY (1991): "Health Inequalities among British Civil Servants: the Whitehall II Studies," *Lancet*, 337(8754), 1387–1393.

MEGHIR, C., AND L. PISTAFERRI (2004): "Income Variance Dynamics and Heterogeneity," *Econometrica*, 72(1), 1–32.

MILLER, D., M. PAGE, A. HUFF-STEVENS, AND M. FILIPSKI (2009): "Why are recessions good for your health?," *American Economic Review*, 99(2), 122–127.

ROODMAN, D. (2009): "A Note on the Theme of Too Many Instruments," *Oxford Bulletin of Economics and Statistics*, 71(1), 135–58.

RUHM, C. (2000): "Are Recssions Good for your Health?," *Quarterly Journal of Economics*, 115(2), 617–50.

Smith, J. (1999): "Healthy Bodies and Thick Wallets: The Dual Relationship between Health and Economic Status," *Journal of Economic Perspectives*, 13(2), 145–166.

——— (2003): "Health and SES across the Life Course," Unpublished Mimeo.

——— (2004): "Unraveling the SES: Health Connection," *Population and Development Review*, 30(Supplement: Aging, Health, and Public Policy), 108–132.

Stock, J., J. Wright, and M. Yogo (2002): "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments," *Journal Business and Economic Statistics*, 20(4), 518–529.

Strully, K. (2009): "Job Loss and Health in the US Labor Market," *Demography*, 46(2), 221–246.

Sullivan, D., and T. von Wachter (2009a): "Average Earnings and Long-Term Mortality: Evidence from Administrative Data," *American Economic Review*, 99(2), 133–138.

——— (2009b): "Job Displacement and Mortality: An Analysis Using Administrative Data," *Quarterly Journal of Economics*, 124(3), 1265–1306.

Table 1: Summary Statistics

Variable	Notes	Mean (Std Dev)
Health	Binary indicator for SRHS <=3	0.89 (0.31)
Income	Individual labor income (2007 $)	42008.01 (39533.51)
White	Indicator for being white	0.69 (0.46)
Black	Indicator for being black	0.30 (0.46)
Married	Indicator for being married	0.78 (0.42)
Sex	Indicator for being male	0.55 (0.50)
College	Indicator for having a college degree	0.29 (0.45)
High School	Indicator for having >= 12 yrs of school	0.50 (0.50)

Table 2: Autocorrelations - Singles

	$Cor\left(\Delta y_t, \Delta y_{t-s}\right)$	$Cor\left(\Delta h_t, \Delta h_{t-s}\right)$	$Cor\left(\Delta y_t, \Delta y_{t-s}\right)$	$Cor\left(\Delta h_t, \Delta h_{t-s}\right)$
	Men		Women	
$s=1$	-0.2066^{***}	-0.4829^{***}	-0.1193^{***}	-0.4420^{***}
	[0.000]	[0.000]	[0.000]	[0.000]
$s=2$	0.0148	0.0461	-0.0262	-0.0305
	[0.627]	[0.243]	[0.328]	[0.172]
$s=3$	-0.0621^{*}	-0.0382	0.0207	0.0212
	[0.052]	[0.427]	[0.429]	[0.423]

p-values in brackets. Standard errors were computed with the bootstrap.
*** sig at 99%, ** sig at 95%, * sig at 90%

Table 3: Autocorrelations - Couples

	$Cor\left(\Delta y_t, \Delta y_{t-s}\right)$	$Cor\left(\Delta h_t, \Delta h_{t-s}\right)$	$Cor\left(\Delta y_t, \Delta y_{t-s}\right)$	$Cor\left(\Delta h_t, \Delta h_{t-s}\right)$
	Men		Women	
$s=1$	-0.2296^{***}	-0.4442^{***}	-0.0442	-0.4768^{***}
	[0.000]	[0.000]	[0.238]	[0.000]
$s=2$	-0.0373^{***}	-0.0270	-0.0232	0.0029
	[0.002]	[0.120]	[0.159]	[0.896]
$s=3$	-0.0123	-0.0062	0.0084	-0.0020
	[0.399]	[0.752]	[0.629]	[0.937]

p-values in brackets. Standard errors were computed with the bootstrap.
*** sig at 99%, ** sig at 95%, * sig at 90%

Table 4: OLS Estimates: Income to Health

	SM	SW	MM	MM	MW	MW
	(1)	(2)	(3)	(4)	(5)	(6)
$h_{i(t-1)}$	−0.499	−0.448	−0.441	−0.440	−0.479	−0.481
	(−18.58)	(−23.58)	(−35.95)	(−35.86)	(−31.61)	(−31.73)
y_{it}	0.038	−0.007	0.020	0.018	0.011	0.009
	(2.06)	(−0.36)	(3.05)	(2.71)	(1.47)	(1.18)
y_{it}^{sp}				−0.001		0.002
				(−2.23)		(1.26)
age_{it}	−0.010	−0.005	−0.003	−0.004	−0.004	−0.003
	(−2.52)	(−1.45)	(−2.94)	(−3.40)	(−2.28)	(−1.96)
R^2	0.2349	0.1955	0.1978	0.1973	0.2275	0.2294
N	916	1103	3103	3058	2156	2114

t-statistics in parentheses. All standard errors clustered by individual.

Table 5: AB Estimates: Income to Health, Strictly Exogenous Income

	SM	SW	MM	MM	MW	MW
	(1)	(2)	(3)	(4)	(5)	(6)
$h_{i(t-1)}$	−0.039	0.092	0.061	0.053	0.004	0.004
	(−0.92)	(3.13)	(3.53)	(3.03)	(0.18)	(0.20)
y_{it}	0.037	−0.018	0.022	0.020	0.012	0.010
	(2.03)	(−0.92)	(2.92)	(2.50)	(1.32)	(1.07)
y_{it}^{sp}				−0.002		0.003
				(−2.41)		(1.72)
age_{it}	−0.007	−0.005	−0.003	−0.003	−0.004	−0.004
	(−2.06)	(−2.10)	(−4.45)	(−4.66)	(−4.03)	(−3.43)
m_1	−7.18	−11.12	−14.70	−14.21	−11.77	−11.61
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
m_2	0.998	0.35	−0.11	−0.22	0.348	0.23
	[0.319]	[0.724]	[0.911]	[0.826]	[0.728]	[0.817]
J	35.82	41.99	37.91	37.54	35.20	47.96
	[0.429]	[0.194]	[0.338]	[0.354]	[0.459]	[0.071]
# of IV	39	39	39	40	39	40
N	916	1103	3103	3058	2156	2114

t-statistics in parentheses. p-values in brackets.

Table 6: AB Estimates: Income to Health, Predetermined

	SM	SW	MM	MM	MW	MW
	(1)	(2)	(3)	(4)	(5)	(6)
$h_{i(t-1)}$	0.010	0.103	0.077	0.073	−0.004	0.004
	(0.29)	(4.21)	(6.23)	(5.78)	(−0.21)	(0.24)
y_{it}	−0.033	−0.178	0.054	0.039	0.014	0.014
	(−0.49)	(−2.32)	(1.93)	(1.39)	(0.28)	(0.31)
y_{it}^{sp}				−0.003		0.001
				(−1.87)		(0.25)
age_{it}	−0.004	−0.000	−0.004	−0.003	−0.005	−0.004
	(−1.22)	(0.05)	(−4.21)	(−4.00)	(−2.45)	(−1.98)
m_1	−7.07	−11.24	−15.51	−15.16	−12.36	−12.57
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
m_2	0.845	0.27	0.10	0.11	0.12	0.29
	[0.398]	[0.790]	[0.924]	[0.912]	[0.903]	[0.771]
J	54.96	66.43	54.70	67.25	64.57	91.30
	[0.552]	[0.184]	[0.562]	[0.824]	[0.229]	[0.163]
# of IV	61	61	61	84	61	84
N	916	1103	3103	3058	2156	2114

t-statistics in parentheses. p-values in brackets.

Table 7: Income Spline for Married Men

	MM
y_{it}	0.11
	(3.22)
$y_{it} * 1(y_{it} \in (y_{25}, y_{50}])$	-0.006
	(-3.92)
$y_{it} * 1(y_{it} \in (y_{50}, y_{75}])$	-0.010
	(-4.00)
$y_{it} * 1(y_{it} \in (y_{75}, y_{100}])$	-0.012
	(-3.81)
m_1	-15.24
	[0.000]
m_2	-0.15
	[0.879]
J	114.12
	[0.705]
N	3103

This table reports estimates of equation (9).
We only report the estimates of coefficients on the income variables. t-statistics are in parentheses.
p-values are in brackets.

Table 8: First Stage Regressions

	SM	SW	MM	MW
	Dep Var = Δy_{it}			
$y_{i(t-1)}$	−0.27	−0.17	−0.29	−0.07
	(−7.50)	(−3.01)	(−20.56)	(−1.45)
$y_{i(t-2)}$	0.21	0.13	0.19	0.05
	(5.44)	(2.49)	(11.79)	(0.97)
$y_{i(t-3)}$	0.05	0.04	0.09	0.02
	(1.52)	(1.24)	(7.51)	(1.34)
F	26.48	3.74	171.22	4.12
	Dep Var = Δh_{it}			
$h_{i(t-1)}$	−0.68	−0.66	−0.62	−0.66
	(−18.68)	(−26.57)	(−32.85)	(−30.35)
$h_{i(t-2)}$	0.28	0.21	0.21	0.25
	(7.51)	(8.37)	(11.83)	(11.50)
$h_{i(t-3)}$	0.17	0.21	0.19	0.22
	(4.79)	(8.62)	(11.27)	(10.56)
F	121.18	252.13	366.65	311.94

t-statistics in parentheses.

Table 9: Estimations Broken Down by Age

	SM	SM	MM	MW
	(1)	(2)	(3)	(4)
			≤ 45	
β^H	0.017	-0.164	0.086	-0.004
	(0.24)	(-1.89)	(2.85)	(-0.09)
N	771	791	2495	1689
			> 45	
β^H	-0.071	-0.178	-0.010	-0.026
	(-0.73)	(-1.85)	(-0.18)	(-0.31)
N	187	415	963	691

This table reports estimations of the same model from Table 6 broken down by age. We only report the estimates of β^H. t-statistics are reported in parentheses.

Table 10: AB Estimates: Health to Income, Predetermined

	SM	SW	MM	MM	MW	MW
	(1)	(2)	(3)	(4)	(5)	(6)
$y_{i(t-1)}$	0.010	0.506	0.372	0.368	0.585	0.568
	(1.10)	(6.94)	(8.79)	(9.03)	(11.29)	(11.47)
h_{it}	−0.003	−0.046	0.004	0.002	0.000	0.003
	(−0.09)	(−2.06)	(0.21)	(0.13)	(0.02)	(0.14)
h_{it}^{sp}				−0.002		−0.010
				(−0.16)		(−0.60)
age_{it}	0.003	0.006	0.008	0.008	0.004	0.006
	(0.57)	(1.89)	(5.55)	(5.66)	(1.63)	(2.07)
m_1	−4.27	−7.45	−12.40	−12.77	−8.22	−8.44
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
m_2	0.953	−0.06	2.52	2.48	−0.123	−0.161
	[0.340]	[0.949]	[0.012]	[0.013]	[0.902]	[0.872]
J	78.81	70.17	109.17	121.29	60.34	86.56
	[0.029]	[0.113]	[0.000]	[0.002]	[0.356]	[0.263]
# of IV	61	61	61	84	61	84
N	916	1103	3103	3103	2156	2156

t-statistics in parentheses. p-values in brackets.

Table 11: Alternative Model Results

	(1)	(2)	(3)	(4)
	Dep Var = Δh_{it}			
	SM		SW	
$h_{i(t-1)}$	-	-0.585	-	-0.485
		(-16.77)		(-22.11)
$y_{i(t-1)}$	-0.001	0.037	0.000	0.017
	(-0.14)	(8.71)	(0.04)	(8.63)
R^2	0.0010	0.2801	0.0001	0.2412
N	980	980	1123	1123
	MM		MW	
$h_{i(t-1)}$	-	-0.501	-	-0.492
		(-27.22)		(-23.33)
$y_{i(t-1)}$	0.000	0.031	-0.001	0.008
	(0.03)	(14.18)	(-1.30)	(7.72)
R^2	0.0001	0.2389	0.0001	0.2365
N	3157	3157	2177	2177

This table contains estimates of the alternative model in equation (15). Controls for education and race were included but not reported. t-statistics are reported in parentheses.

Table 12: Sample Sizes by Year

Year	Sample Size
1984	4752
1985	4752
1986	4920
1987	4913
1988	4898
1989	4907
1990	4921
1991	4937
1992	4588
1993	4083

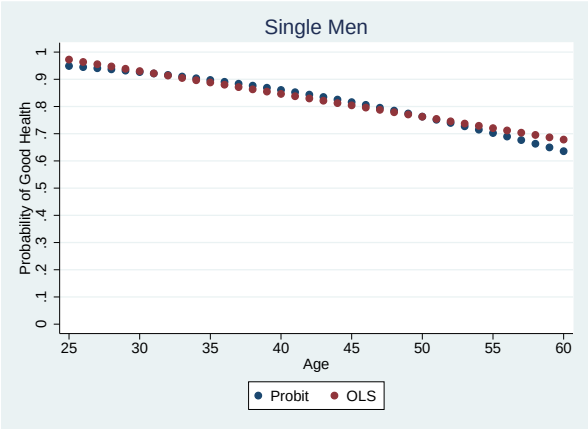

Figure 1

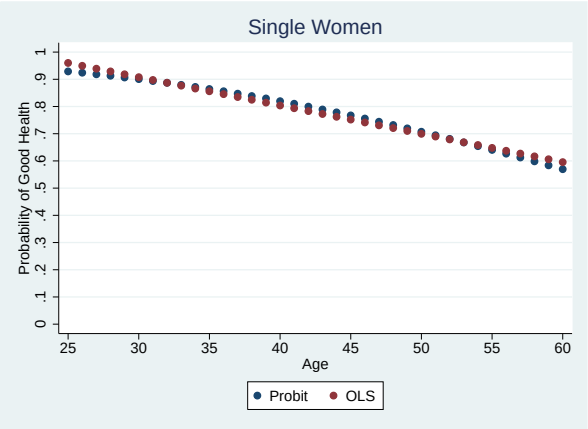

Figure 2

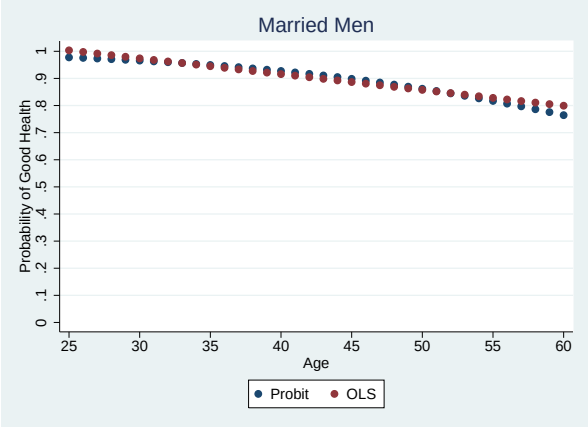

Figure 3

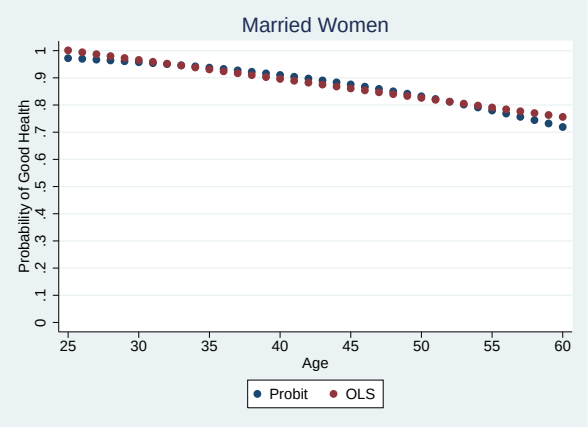

Figure 4