Earnings Growth and Movements in Self-Reported Health

Timothy J. Halliday*

University of Hawai'i at Mānoa and IZA

December 11, 2015

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 causality

*Address: 2424 Maile Way; Saunders Hall 533; Honolulu, HI 96822. Tele: (808) 956-8615. Fax: (808) 956-4347 E-mail: halliday@hawaii.edu. This is a substantially revised version of an earlier paper titled, "Income Volatility and Health." I would like to thank Randy Akee, Jenny Aker, Meta Brown, Larry Katz, Pat Kline, Sumner La Croix, Jeremy Magruder, Ted Miguel, Enrico Moretti, Chris Paxson, Chris Ruhm and Lowell Taylor for useful feedback. Finally, I would like to thank the editor of this journal and two anonymous referees for invaluable comments. The usual disclaimer applies.

running from income to health for married women and men. In addition, spousal income appears to be protective for married women.

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. Traditionally, 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)). However recently, there does appear to be some reversal of this trend with recent work by Strully (2009) and Sullivan and von Wachter (2009) providing evidence of adverse consequences of job loss on health. On the other hand, public health experts and epidemiologists traditionally have tended to advocate the reverse causal pathway from SES to health. The most notable evidence of these causal pathways comes from the famous Whitehall studies of British civil servants *e.g.,* Marmot, Rose, Shipley, and Hamilton (1978) and Marmot, Smith, Stansfield, Patel, Head, White, Brunner, and Feeney (1991). While some have criticized some of the methodological foundations of these studies, recent work by Anderson and Marmot (2011) employs an instrumental variables strategy to address many of these critiques and still finds a significant and positive relationship running from SES to health. 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.

There are two primary approaches to unravel this correlation in the literature. The first uses quasi-experimental methods. For example, the studies by Strully (2009) and Sullivan and von Wachter (2009), which were mentioned above, as well as Browning, Dano, and Heinesen (2006) and Halliday (2014) look at the relationship between job loss (or unemployment in the case of Halliday (2014)) and health while adjusting for a variety of control variables which is essentially a "selection-on-observables" strategy. Others such as Frijters, Haisken-DeNew, and Shields (2005) and Meer, Miller, and Rosen (2003) have used exogenous variation generated by German reunification and inheritances, respectively, to produce credible evidence of causal effects of SES on health. The former produce some evidence of a causal effect of SES on health, whereas the latter do not.[1] The other approach to disentangle causal pathways 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 measures of SES while controlling for demographic characteristics and

[1] Note that there is not a contradiction between these two findings since both studies investigate the SES-health nexus along different margins.

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, on the whole, do not show 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, Gunasekara, Carter, and Blakely (2011) found that only 13 compared changes in health outcomes with changes in income. 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.

One recent paper that has addressed these concerns is Michaud and van Soest (2008) who employ dynamic panel techniques to investigate causality between wealth and health in the Health and Retirement Study. They find no evidence of causal effects of wealth on health in a population of older American couples. In this paper, we employ data from the Panel Study of Income Dynamics (PSID) to investigate income to health causality while taking these same issues into account.

There are three main points of departure between our paper and Michaud and van Soest (2008). First, we consider a younger population. In fact, in Michaud and van Soest (2008), they state, "The fact that we find no causal links from wealth to health for the age groups considered does not mean that such a causal link never operated earlier in life." The average age in our panel is 39.18, whereas Michaud and van Soest (2008) consider couples who are in their fifties. Second, we focus on income, whereas the other authors focus on wealth. We believe that earnings risk is probably more important than wealth shocks in the younger population that we consider. Moreover, work by Abowd and Card (1989) and Meghir and Pistaferri (2004) suggests that log-earnings changes have sizable permanent components. Third, based on the performance of specification tests, our core results utilize differenced GMM, whereas Michaud and van Soest (2008) use system GMM which requires stronger assumptions.

Specifically, we focus on the relationship between earnings *growth* which has been the subject of a large literature in labor economics 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 more meaningful time series variation. Finally, by the nature of our research design, we will identify the effects of relatively short-run fluctuations in SES on health status, whereas much of the epidemiology literature such as the Whitehall Studies has focused on longer-term exposures to SES.

An exercise of this nature goes a long way towards alleviating the concern of Fuchs (1982) that different discount factors (which should be relatively time invariant) lead to similar investments in both human and health capital. It also helps to mitigate some of the lingering issues with some of the recent studies that rely on cross-sectional variation in job loss for identification (*e.g.,* Strully (2009) and Sullivan and von Wachter (2009)). In exchange for eliminating time-invariant individual-specific omitted variables, we must place some restrictions on the causal ordering between health and earnings which, to some extent, are testable. Importantly, these restrictions do allow health to impact earnings, but it does so with a lag. While we do not claim that this is an unimpeachable assumption, we do contend that our approach does a very thorough job of eliminating many confounding variables that may be problematic in other studies and also that many other panel studies of the effects of earnings on health assume that income is strictly exogenous which rules out *any* causality running from health to income (*e.g.,* Jones and Wildman (2008) and Contoyannis, Jones, and Rice (2004a)).

The remainder 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. 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 respondent's 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).[2] Following Meghir

[2] 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 for two reasons. First, PSID switched from paper and pencil collection to computer assisted collection. Second, PSID stopped releasing the final releases of the data after 1993. Both of these factors appear to have impacted the income measurements in the PSID. As such, many studies of earnings growth such as Meghir and Pistaferri (2004) and Storesletten, Telmer, and Yaron (2004) stop at 1993.

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.[3]

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.

For most of the results, 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 common in the literature (*e.g.*, Hoffman, Koyama, Albertsen, Barry, Daskivich, Goodman, Hamilton, Stan-

[3]Note that because we did not drop the SEO, the percentage of blacks in the data is higher than in the US population.

ford, Stroup, Potosky, et al. (2015)), although note that, after our mapping, *higher* values denote better health (*i.e.*, the ones are healthier than the zeros). Because of this transformation, we will, in effect, be estimating a linear probability model. One of the advantages of this is that the probability of being in good health is measured in cardinal units (*i.e.*, percentage points), whereas the five point SRHS variable is a cardinal measure. Finally, while SRHS are not perfect, they are one of the best measures of health status that exhibits temporal variation.[4]

3 Estimation Equation

We let i denote individuals, t denote time, and g denote a demographic group. Denoting log labor income by y_{it}, the binary SRHS variable by h_{it}, and age by a_{it}, we consider the following model:

$$h_{it} = \alpha_i + \gamma_g h_{i(t-1)} + \beta_g y_{it} + \phi_g a_{it} + \upsilon_{it} \quad (1)$$

for $i = 1, ..., N$, $t = 1, ...T$ and $g = 1, ..., G$.[5] To account for parameter heterogeneity across subgroups as in Meghir and Pistaferri (2004), we subscript the parameters g to

[4] Banks and Smith (2012) discuss many of the pros and cons of SRHS measures.
[5] We took the log of labor income plus one since we kept people with zero earnings in the sample.

allow for the possibility that the effect of income on health varies across demographic groups. This equation accounts for unobserved heterogeneity in the constant term, dynamics which operate via the lagged dependent variable, causality from income to health, and aging. To purge the model of unobserved heterogeneity, we will work with the model in first-differences:

$$\Delta h_{it} = \gamma_g \Delta h_{i(t-1)} + \beta_g \Delta y_{it} + \phi_g + \Delta \upsilon_{it}. \tag{2}$$

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.[6]

3.1 Moment Restrictions

[6]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 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.

Identification of β_g 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 are strictly exogenous. Specifically, if we adopt the notation that $z_i^t \equiv (z_{i1}, ..., z_{it})'$, then strict exogeneity requires that

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

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

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

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. This assumption implies that a health innovation today is uncorrelated with contemporaneous income, but can affect income tomorrow and beyond.[7]

[7] It imposes a particular causal ordering on health and income in which income at time t is allowed

There are pros and cons associated with the assumption in equation (4). Its primary virtue is that it imposes no relationship between the fixed effect and earnings. Consequently, it allows for substantial strides to be made towards addressing the critiques of Fuchs (1982). Also, it relies solely on time series variation in health and earnings for identification. As such, it eliminates many of the cross-sectional confounding variables that may be issues in recent work by Strully (2009) and Sullivan and von Wachter (2009). Another of its virtues is that it does allow for causality from health to earnings, albeit with a lag. Alas, this virtue may also be its vice since many readers may think that even this is too strong.[8]

However, it is important to bear in mind that this assumption is actually substantially weaker than what has been employed elsewhere. For example, Smith (1999), Smith (2004), and Adams, Hurd, McFadden, Merrill, and Ribeiro (2003) do not account for unobserved heterogeneity. In this sense, we are innovating upon these studies. Work by Jones and Wildman (2008), Contoyannis, Jones, and Rice (2004b), and Contoyannis, Jones, and Rice (2004a) does allow for unobserved heterogeneity,

---

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.

[8] It is possible to weaken the moment condition to account for contemporaneous causality running from health to earnings. However, doing so substantially weakens the moments. Also, if the specification tests of the moment in equation (4) perform well then this suggests that using these weaker conditions might do more harm than good.

but they adopt a random effects approach. Importantly, they assume that earnings is strictly exogenous which is a stronger assumption than we employ. Overall, we do not believe that this moment condition is beyond reproach, but we do believe that it has numerous merits that have not been fully exploited in this literature.

3.2 GMM Estimation

We will employ a GMM procedure on the differenced model 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}. It has been shown that standard errors from the two-step procedure are biased, so we employ the correction developed by Windmeijer (2005).

Note that the results that we first present in this paper are based on what is called difference-GMM which contrasts with system-GMM which is what was utilized by Michaud and van Soest (2008). In this paper, we utilize both procedures. We discuss the key differences between these two procedures at greater length in Section 4.4.

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 will 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.[9]

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

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

[9] It made little difference if we capped the number of instruments at two or three. For example, the estimate in Table 4 for married men in Column 4 is 0.0536 with a standard error of 0.0278 whereas using only two lags for instruments, we obtain an estimate of 0.0517 with a standard error of 0.0283.

Arellano and Bond (1991) develop a test statistic 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

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

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.[10]

[10] 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.

4 Empirical Results

Throughout most of this section, we will estimate our models using four demographic subsamples: single men (SM), single women (SW), married men (MM), and married women (MW).[11] The reason that we do this is twofold. First and as we have already discussed, it is entirely plausible to suspect that the effects of income on health will vary across demographic groups. Second, when considering married people, spousal in addition to own income may matter and, so it makes sense to estimate the model separately for married and single people.

4.1 Weak Instruments

Before we discuss our GMM results, we will investigate whether or not weak instruments is an issue using a common procedure in which we regress the endogenous variables on the excluded instruments.[12] To do this, we consider the following

[11] 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. Finally, people may switch marital status while in the data. This is not problematic to the extent that our fixed effect approach eliminates the initial conditions problem, but may be problematic if marital status changes are related to υ_{it}. Note that we obtain similar results when we stratify only by gender in Table 7 and, so we do not view this as being problematic.

[12] It is also standard to include the included exogenous variables as well, but our only exogenous variable is age which is essentially unity when first differenced. So, the constants should account for this.

equations via Ordinary Least Squares:

$$\Delta y_{it} = \phi_0 + \phi_1 y_{i(t-1)} + \phi_2 y_{i(t-2)} + \phi_3 y_{i(t-3)} + u_{it}^Y \tag{7}$$

and

$$\Delta h_{it} = \eta_0 + \eta_1 h_{i(t-1)} + \eta_2 h_{i(t-2)} + \eta_3 h_{i(t-3)} + u_{it}^H. \tag{8}$$

Note that we have suppressed the g subscript for the ease of notation. 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 only include three lags in equations (7) and (8). As discussed in the weak instruments literature, the conventional distribution theory for the F-statistic is no longer applicable. Instead, we will use the rule-of-thumb 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, 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 2. In the top panel, the estimation of equation (7) reveals that weak instruments may be an issue for single and married women as the F-statistics are 3.74 and 4.12, respectively. 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 (8) 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.

4.2 Strictly Exogenous Income

We now discuss the GMM results obtained using the strict exogeneity assumption. These are reported in Table 3. First, the tests of serial correlation in the differenced residuals (*i.e.,* m_1 and m_2) perform well. 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 GMM results. We fail to reject null for all of the overidentification tests at the 5% level, although we do reject the null that our moments are valid at the 10% level in column 6 for married women. Note that the strict exogeneity moment condition uses no overidentifying restrictions for income. In other words, since income is assumed strictly exogenous, β_g is identified from the moment

$$E[\Delta \upsilon_{it} \Delta y_{it}] = 0. \tag{9}$$

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

We now turn to the estimates of the impact of earnings on health. First, we do not see any effects for women, although own and spousal income does appear to be moderately protective for married women. Second, income is protective for men. The estimates range from 0.018 for married men to 0.034 for single men. This indicates that a 1% increase in income results in between a 0.018 and 0.034 percentage point increase in the probability of being in good health.[13] Finally, there is some indication that spousal income is negatively associated with good health for married men.[14]

---

[13] 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, $\beta/100$ can be interpreted as the effect of a 1% change in earnings on the probability of being in good health. To obtain the effect in percentage points, we multiply one more time by 100.

[14] The estimates of the coefficient on lagged health are not significant for single men or married women. This indicates a lack of state dependence for these demographics. However, it should be borne in mind that when we use system GMM which employs stronger moment conditions, we find very strong evidence that health status is highly persistent. Our suspicion is that a combination of weak instruments and the relative inefficiency of difference GMM are responsible for the lack of evidence of positive state dependence for these subgroups.

4.3 Predetermined Income

We now turn to the results that use the predetermined assumption. These are reported in Table 4. We no longer see any effects of income for single men. However, it is important to bear in mind that both the sample size is small ($N = 916$) and the predeterminedness assumption utilizes weaker instruments than the strict exogeneity assumption. So, this null result may be driven by an inefficient estimator coupled with a small sample size. 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 -0.165. Second, as indicated in Table 2, the instruments for single women are very weak and so this result is most likely the consequence of this.

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_g is now 0.077 and significant at the 1% level. In contrast, we saw that with the strict exogeneity assumption, the point-estimate was more than half as small at 0.022. Moving to column 4 where we include spousal income, we see that own income is still significant with a similar point-estimate of 0.064. 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 3 where we invoke the strict exogeneity assumption, we saw that the estimate was smaller at -0.002 but much more significant. We see that the model performs 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; however, the estimates in column 6 are at the edge of significance. The models 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 4 are larger than their counterparts in Table 3 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 $t-1$.

4.4 System GMM

We have shown in Table 2 that weak instruments may be problematic, particularly for women. Blundell and Bond (1998) showed that one solution to the problem of weak instruments in models of the type that we are considering in this paper is to augment the model with an additional set of moments in which lagged differences act as instruments for the model in levels (which is the opposite of what we have done). Under additional conditions, these are valid moment conditions.

Note that by doing this, we are, in effect, trading less plausible moment conditions for stronger instruments. Hence, in the absence of weak instruments, the estimates in Table 4 should be viewed as superior to the system estimates. In addition, because the system estimates utilize a different set of moment conditions, theoretically, it is possible for a model that satisfies the specification tests when using only the moments in equation (4) to be rejected once we add the additional moment conditions in the system approach.

We provide a set of results using this "system" approach in Table 5. We found that three lags of health were needed for the specification tests to perform adequately which, as we just discussed, is theoretically possible. This is consistent with Michaud and van Soest (2008) who also found that more than one lag was necessary for the models to perform well.

The results in this table by-and-large buttress the previous results. We find no effects for singles, even with the more powerful instruments. We still find effects of earnings on the health of married men. Its point-estimate is 0.030 which is smaller than the "difference-GMM" results with predetermined income in Table 4 but larger than the difference-GMM with strictly exogenous income. We still see the negative effect of spousal earnings on male health, but this estimate is no longer signifcant at conventional levels. For married women, the stronger instruments appear to make a difference as we now see positive and significant effects of own income on women's health. Moreover, in the final column, we see that spousal income is positive and moderately significant.

We conclude by estimating the model using system GMM but with the full five point SRHS variable as the dependent variable as a sensitivity check.[15] First, we see that, on the whole, all of the models perform well. We only reject the overidentification test in column 4 for married men at the 10% level. Next, we still see that positive earnings innovations are associated with better health for married men. Finally, we see that positive income shocks are associated with better health for married women, but these estimates are only marginally significant. Interestingly, in column 6, we still see that spousal income is strongly protective of married women's health. Finally, the negative effects of spousal income on male health is not present in this table.

[15] Bear in mind that now a negative coefficient on income indicates that positive income shocks are associated with better health. In addition, we employed the binary SRHS variable as instruments for the 5-point SRHS variable. The results were similar either way, but the specification tests performed slightly better when we employed the former method.

4.5 Stratifying Exclusively by Gender

One potential issue with the results thus far is that we stratified by gender and marital status. We chose to do this because of the possible importance of spousal income for married people. However, the potential drawback of this is that marital status is endogenous and may be impacted by income and health. Consequently, we re-estimate our models separately for men (M) and women (W). We employ both difference and system GMM and both the binary and 5-point SRHS variables. All estimations treat income as predetermined.

The results are reported in Table 7. The main finding in this table is that we still find strong evidence of income to health causality when only stratifying by gender. The only insignificant estimate occurs when we employ difference GMM for women. In addition, the magnitudes are broadly in-line with previous estimates. For example, in column 1, we obtain an estimate of 0.073 for all men using difference GMM, whereas we obtained an estimate of 0.077 for married men in Table 4. Similarly, we obtain an estimate of 0.031 in column 3 using system GMM for men, whereas we obtained an estimate of 0.030 for married men in Table 5.

4.6 Using Different Income Transformations

Finally, we explore the robustness of our results to employing different transformations of income. We employ the system GMM estimator as we did for the results in Tables 5 and 6.[16] We consider four transformations of income: income (divided by 1000) and dummies for income being above the 10th, 50th and 90th percentiles. As in the previous section, we estimate the model separately for men and women.

The results are reported in Table 8. First, we note that there is evidence of income to health causality for all income transformations and for both genders. Second, when we use the untransformed income variable, we see that \$10,000 extra in income is associate with a 0.46 and 2.61 percentage point increase in the probability of being in good health for men and women, respectively. Third, looking at the income percentile results for men, we see that movements out of the bottom 10th percentile far-and-away have the largest impacts. Fourth, looking at the income percentile results for women, we see the opposite, namely, that the largest impacts occur for movements into the top 10th percentile.

One issue with the results using the income quantiles is that their magnitudes can be implausible. Given this, we admonish the reader not to take these results too seriously. However, they do indicate two issues. The first is that there is evidence of

[16] We employed lags of the binary health variable the logged income as instruments.

income to health causality with alternative transformations of income. The second is that they provide some suggestive (but not conclusive) evidence that the bulk of the effects for men are concentrated in the bottom of the income distribution, whereas the bulk of the effects for women are concentrated in the top of the distribution.

5 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 conducted tests for income to health causality and found evidence of a positive relationship between earnings growth and improvements in health for married men and women. We also found some evidence that spousal income was protective for married women.

Moving forward, we propose two avenues for research. First, while our work does suggest that there are causal effects of income on health, the mechanism is not clear. Two possibilities are a stress channel as has been suggested by the seminal Whitehall studies *e.g.,* Marmot, Rose, Shipley, and Hamilton (1978) and Marmot, Smith, Stansfield, Patel, Head, White, Brunner, and Feeney (1991) and loss of employer-based health insurance which is an issue in our sample of Americans.[17]

[17] Unfortunately, the PSID does not ask about the respondent's health insurance status regularly until 1999.

Second, our work suggests that the combination of a permanent and a transitory earnings shock has a causal effect on health status, but it is somewhat silent on how much each of these two components might matter. In future work, researchers should attempt to disentangle these two shocks.

6 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. 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. 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.

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.

Anderson, M., and M. Marmot (2011): "The Effects of Promotions on Heart Disease: Evidence from Whitehall," forthcoming Economic Journal.

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 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.

Banks, J., and J. P. Smith (2012): "International comparisons in health economics: evidence from aging studies," *Annual review of economics*, 4, 57.

Blundell, R., and S. Bond (1998): "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics*, 87(1), 115–143.

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

Browning, M., A. Dano, and E. Heinesen (2006): "Job Displacement and Stress-Related Health Outcomes," *Health Economics*, 15(10), 1061–1075.

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.

Frijters, P., J. Haisken-DeNew, and M. Shields (2005): "The Causal Effect of Income on Health: Evidence from German Ruinification," *Journal of Health Economics*, 24(5), 997–1017.

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. J. (2014): "Unemployment and Mortality: Evidence from the PSID," *Social Science & Medicine*, 113, 15–22.

Hoffman, R. M., T. Koyama, P. C. Albertsen, M. J. Barry, T. J. Daskivich, M. Goodman, A. S. Hamilton, J. L. Stanford, A. M. Stroup, A. L. Potosky, et al. (2015): "Self-reported health status predicts other-cause mortality in men with localized prostate cancer: results from the prostate cancer outcomes study," *Journal of general internal medicine*, pp. 1–11.

Jones, A., and J. Wildman (2008): "Health, Income and Relative Deprivation: Evidence from the BHPS," *Journal of Health Economics*, 27(2), 308–324.

Marmot, M., G. Rose, M. Shipley, and P. Hamilton (1978): "Employment Grade and Coronary Heart Disease in British Civil Servants," *Journal of Epidemiology and Community Health*, 32(4), 244–249.

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.

Meer, J., D. Miller, and H. Rosen (2003): "Exploring the Health-Wealth Nexus," *Journal of Health Economics*, 22(5), 713–730.

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

Michaud, P.-C., and A. van Soest (2008): "Health and Wealth of Elderly Couples: Causality Tests using Dynamic Panel Models," *Journal of Health Economics*, 27(5), 1312–1325.

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

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

——— (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.

Storesletten, K., C. Telmer, and A. Yaron (2004): "Consumption and Risk Sharing over the Life Cycle," *Journal of Monetary Economics*, 51(3), 609–633.

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

Sullivan, D., and T. von Wachter (2009): "Job Displacement and Mortality: An Analysis Using Administrative Data," *Quarterly Journal of Economics*, 124(3), 1265–1306.

Windmeijer, F. (2005): "A finite sample correction for the variance of linear efficient two-step GMM estimators," *Journal of econometrics*, 126(1), 25–51.

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)
Age	Respondent's age	39.18 (9.26)
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: 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 3: Difference GMM Estimates: Strictly Exogenous Income

	SM	SW	MM	MM	MW	MW
	(1)	(2)	(3)	(4)	(5)	(6)
$h_{i(t-1)}$	-0.007	0.128	0.063	0.061	0.009	-0.004
	(-0.10)	(2.94)	(2.14)	(2.03)	(0.27)	(-0.10)
y_{it}	0.034	-0.027	0.022	0.018	0.012	0.009
	(1.59)	(-1.17)	(2.85)	(2.37)	(1.51)	(1.11)
y_{it}^{sp}				-0.002		0.003
				(-2.32)		(1.35)
age_{it}	-0.004	-0.006	-0.003	-0.003	-0.004	-0.004
	(-1.06)	(-2.28)	(-3.23)	(-3.21)	(-3.27)	(-2.85)
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 4: Difference GMM Estimates: Predetermined Income

	SM	SW	MM	MM	MW	MW
	(1)	(2)	(3)	(4)	(5)	(6)
$h_{i(t-1)}$	-0.013	0.114	0.073	0.076	-0.004	0.001
	(-0.20)	(2.94)	(2.69)	(2.78)	(-0.13)	(0.03)
y_{it}	-0.005	-0.165	0.077	0.064	0.035	0.073
	(-0.09)	(-1.82)	(2.85)	(2.38)	(0.65)	(1.49)
y_{it}^{sp}				-0.003		0.006
				(-1.77)		(1.08)
age_{it}	-0.006	-0.001	-0.004	-0.003	-0.004	-0.005
	(-1.72)	(-0.22)	(-4.08)	(-4.00)	(-2.21)	(-2.49)
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 5: System GMM Estimates

	SM	SW	MM	MM	MW	MW
	(1)	(2)	(3)	(4)	(5)	(6)
$h_{i(t-1)}$	0.275	0.220	0.261	0.265	0.202	0.222
	(5.39)	(5.31)	(9.55)	(9.50)	(5.43)	(6.13)
$h_{i(t-2)}$	0.128	0.085	0.110	0.106	0.099	0.110
	(2.68)	(2.12)	(4.88)	(4.57)	(2.82)	(3.11)
$h_{i(t-3)}$	0.018	0.066	0.049	0.054	0.077	0.081
	(0.37)	(1.73)	(2.13)	(2.26)	(2.82)	(3.01)
y_{it}	0.022	0.009	0.030	0.028	0.023	0.020
	(0.94)	(0.67)	(3.60)	(3.54)	(3.70)	(3.01)
y_{it}^{sp}				−0.001		0.004
				(−1.20)		(1.59)
age_{it}	−0.004	−0.006	−0.003	−0.002	−0.002	−0.002
	(−2.33)	(−5.65)	(−6.47)	(−6.46)	(−2.38)	(−2.34)
m_1	−5.95	−9.00	−12.68	−12.24	−10.40	−10.02
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
m_2	−0.15	−0.42	−0.87	−0.73	1.19	1.15
	[0.880]	[0.450]	[0.386]	[0.468]	[0.232]	[0.251]
J	67.72	98.20	57.53	97.50	71.06	96.32
	[0.487]	[0.010]	[0.813]	[0.327]	[0.376]	[0.359]
# of IV	74	74	74	99	74	99
N	764	941	2771	2771	1855	1855

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

Table 6: System GMM Estimates Using 5-Point SRHS Variable

	SM	SW	MM	MM	MW	MW
	(1)	(2)	(3)	(4)	(5)	(6)
$h_{i(t-1)}$	0.360	0.274	0.365	0.362	0.268	0.274
	(6.69)	(5.73)	(11.44)	(11.29)	(6.24)	(6.71)
$h_{i(t-2)}$	0.130	0.096	0.170	0.167	0.092	0.092
	(2.21)	(2.21)	(6.24)	(5.95)	(2.23)	(2.32)
$h_{i(t-3)}$	0.052	0.112	0.052	0.054	0.040	0.059
	(0.81)	(2.54)	(1.82)	(1.90)	(1.01)	(1.52)
y_{it}	-0.039	0.045	-0.089	-0.065	-0.026	-0.035
	(-0.55)	(1.36)	(-2.88)	(-2.22)	(-1.39)	(-1.84)
y_{it}^{sp}				0.002		-0.016
				(0.38)		(-2.01)
age_{it}	0.012	0.018	0.007	0.008	0.014	0.011
	(2.53)	(5.28)	(4.59)	(4.89)	(4.33)	(3.83)
m_1	-7.34	-9.63	-15.17	-14.94	-12.61	-12.57
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
m_2	0.33	1.14	-0.34	-0.18	1.38	1.58
	[0.740]	[0.254]	[0.730]	[0.861]	[0.168]	[0.115]
J	70.60	68.68	80.86	112.78	83.23	107.08
	[0.391]	[0.454]	[0.136]	[0.070]	[0.101]	[0.135]
# of IV	74	74	74	99	74	99
N	764	941	2771	2771	1855	1855

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

Table 7: Difference and System GMM Results: Stratifying Exclusively By Gender

	M	W	M	W	M	W
	(1)	(2)	(3)	(4)	(5)	(6)
$h_{i(t-1)}$	0.083	0.062	0.268	0.200	0.363	0.248
	(3.78)	(2.74)	(11.42)	(7.37)	(14.00)	(7.55)
$h_{i(t-2)}$			0.113	0.100	0.162	0.078
			(5.48)	(3.93)	(6.54)	(2.56)
$h_{i(t-3)}$			0.052	0.064	0.051	0.056
			(2.51)	(2.87)	(1.98)	(1.89)
y_{it}	0.073	-0.022	0.031	0.030	-0.078	-0.027
	(2.67)	(-0.46)	(3.78)	(4.72)	(-2.87)	(1.89)
age_{it}	-0.004	-0.003	-0.003	-0.003	0.008	0.016
	(-4.59)	(-2.00)	(-6.96)	(-4.43)	(5.44)	(6.55)
m_1	-18.014	-17.421	-14.75	-14.92	-18.03	-16.50
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
m_2	1.136	0.523	-0.50	0.31	-0.50	1.85
	[0.256]	[0.601]	[0.618]	[0.760]	[0.618]	[0.064]
J	75.269	71.641	65.39	81.11	79.25	91.29
	[0.053]	[0.092]	[0.567]	[0.132]	[0.165]	[0.031]
Health Variable	Bin	Bin	Bin	Bin	5-Point	5-Point
Estimation Method	Dif	Dif	Sys	Sys	Sys	Sys
# of IV	61	61	74	74	74	74
N	3486	2961	3110	2574	3110	2574

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

Income is predetermined in all specifications.

Table 8: System GMM Results: Different Income Transformations

	M	W
	(1)	(2)
Income/1000	0.00046	0.00261
	(3.10)	(4.67)
Income > p10	0.455	0.281
	(2.25)	(4.51)
Income > p50	0.064	0.153
	(3.47)	(4.53)
Income > p90	0.078	0.526
	(3.25)	(3.60)

t-statistics in parentheses. Income is predetermined in all specifications. The dependent variable is the binary health variable. Each cell corresponds to an estimate from a separate regression. All regressions employed the system GMM estimator.