Sickness in the Family: Health Shocks and Spousal Labor Supply

Kerwin Kofi Charles
University of Michigan
e-mail:kcharles@umich.edu
phone:734.764.8075
August, 1999

JEL: I12, J12, J14, J22, J26

Abstract

Using evidence from the Health and Retirement Study (H.R.S.), this paper assesses how husbands' and wives' labor supply is affected when their partner is in poor health. I apply methods for dealing with the endogeneity and mismeasurement of dichotomous reported health status, which can only be used with data such as are present in the H.R.S. Contrary to results from the limited work which has been previously done on this subject, I find evidence of significant reactions to spousal bad health. Men reduce labor supply by substantial amounts in response to their partners' poor health whereas wives of ill husbands significantly increase theirs. The paper presents theoretical and statistical explanations for the different spousal responses, and explains also why these results differ from estimates in previous work.

1. Introduction

If married partners live and act as single entities, as most promise to do in their vows, then no event ever truly affects a spouse individually. Certainly, an event as serious as a bout of poor health or disabling illness should affect the welfare and behavior both of the person directly affected and of that person's partner. But while it is well established that poor health generally lowers a person's labor supply; adversely affects the activities they are able to conduct at home; and may affect the degree to which they are able to function effectively without assistance, there is very little evidence on the effects of such a health shock on his spouse. Yet, knowledge of these spousal effects is crucial for constructing sound health care and disability policy for the losses from imperfect health are incurred by the entire family unit to which a disabled person belongs. This paper assesses the impact of spousal health on labor supply using evidence from a new longitudinal survey.

Economists studying disability and poor health[1] have tended to focus on the large labor supply or earnings losses the disabled typically personally suffer. Few have measured other losses directly experienced by the victim, and fewer have looked at how these different losses affect the actions of the person's spouse. Authors looking at spousal response have gotten answers which are quite sensitive to the data and methods used. Cross section studies such as that by Parsons (1977) and Berger (1983) find that the wives of husbands in poor health work slightly more hours than do wives whose husbands are healthy. Berger and Fleisher (1984) look at the labor supply of wives in the National Longitudinal Study whose husbands' health changes and find evidence of very small increases in labor supply among wives whose husband's health worsens, and no effect on participation for the same women. Haurin (1989) also finds no statistically significant effect on labor supply for women whose husbands' health worsens. These very

small or non-existent estimated effects have led Weaver (1994) to conclude in a recent review piece that a spouse's labor supply is little affected by their partner's health.

Several questions are called to mind by these results. One is: Are there sound theoretical reasons why the onset of disability or poor health might genuinely produce the minor labor supply effects the empirical estimates suggest exist? A second set of questions derives from the fact that virtually all of the previous work examines only the effect of husbands' health on wives labor supply. How do husbands change their work behavior when their wives are ill? Are their labor supply reactions similarly modest? Finally, there is the question whether the few existing estimates of a spousal labor supply response are credible, given the various statistical biases which frustrate efforts to determine the true effect of spousal health on labor supply.

Simple Theory and the Spousal Effect of Health

A useful theoretical framework for studying the effects of disability on spouses is the standard family life cycle labor supply model. Variants of this model have been used by countless authors; its essential features are therefore well known and need not be formally reproduced.[2] In the most simple version of the model, a family is assumed to maximize a utility function which is strictly increasing in its two arguments - combined spousal income, and a home produced good. A natural modification which incorporates health assumes that this home produced good is “home care” - the time which spouses devote to providing care and attention to each other. The spouses are endowed with fixed amounts of leisure time which they can either sell in the labor market or use in home production. Husbands and wives may command different wages in the market, and the usefulness of and need for husband's and wife's leisure time in home production may differ as well.

In this simple model, the family's marginal valuation for a spouse's leisure time will be a decreasing function of the number of hours the person desires to work; an increasing function of the family's income, since leisure is a normal good; and an increasing function of the productivity or usefulness of that person's leisure time in home production. To theoretically predict the effect of health on spousal labor supply, it is sufficient to focus on how health affects the market wage or the family's marginal valuation of the person's leisure time. Anything which raises the market wage ought to increase hours of work and the probability of working positive hours; variables which raise the marginal valuation of leisure time should make work less likely.

It is easy to see that bad health or disability has a theoretically ambiguous effect on labor supply of the unaffected spouse. On one hand, since disability almost certainly decreases the labor supply of the affected spouse, it lowers the family's income. This income effect ought to make the other spouse work more. This "added-worker effect" (AWE) has long been recognized in the literature (see Mincer (1962) for example), and has previously been studied in the context of job loss or unemployment (eg. Speltzer (1992); Maloney (1987; 1991); Cullen and Gruber (1991) and Stephens (1998)). But, unlike events like job loss, disabling illness also unambiguously lowers a person's productivity in the home while simultaneously raising his need for attention. This effect raises the family's valuation for the non-disabled spouse's leisure time. Overall, it is not possible to know theoretically which of these opposite effects dominates.

Even if disability profoundly affects a family, it is theoretically possible that the two effects might off-set, on average, so that the observed spousal labor supply response to disability is very small. This offsetting tendency is likely to be strongest when the spouse suffering the disability actually contributes to the family's income - that is, when the disabled spouse previously worked positive hours. Since husbands are much more likely to be labor force participants than are wives, we would expect that small or

statistically insignificant effects would be found in studies which examine how wives respond to husband's disability. The focus on wives makes previous work particularly susceptible to this offsetting effect.

Two other important theoretical considerations likely to affect estimates of spousal health have to do with life-cycle and dynamic aspects of health shocks. Consider first how health evolves over the lifecycle. One thing about which every person can be certain is that, sooner or later, their health will worsen. Obviously, every spouse has this expectation about their partner's health. Aiding the spouse's ability to anticipate the *date* of the inevitable health decline is the fact that he or she is uniquely well informed about a partner's lifestyle and health habits. So, rational spouses may not be observed to change their labor supply when their partner's health worsens because they might not be "shocked" by changes in their partner's conditions at all; they may not make changes because they fully anticipate the trajectory the path that their partner's health takes.

Because researchers, at best, observe a family at only a few points in its life cycle, someone observed to be ill at any point in time will have become ill some time he is observed. But suppose that adjustments to labor supply in response to spouses' health are not constant over the duration of the illness. This is certainly true in the context of parental response to sick children. A parent might stay home from work on the first day that a child is suffering from the flu but, once she is sure that the child is settled in and will be fine, might work *more* than she otherwise would in the second or third day of the child's illness. Similarly, if most of the labor supply adjustment to a partner's health occurs soon after the bad health event happens, there might well be no significant labor supply difference between people whose spouses are ill and those whose partners are not if observations are made sufficiently long after any adjustment has occurred.[3]

Not only is every health shock not a shock to a spouse, but health shocks are rarely permanent. Health varies a great deal throughout life, and people who become ill

most often recover. Thus, even if she is genuinely surprised that her husband has become ill, a rational wife may decide not to alter her labor supply behavior in response to his illness. Because of the impermanence of bouts of poor health, more likely than not, she is probably correct in presuming that her spouse will recover. It might thus be very unwise to alter labor supply in response to every health change on a partner's part, particularly if there are costs associated with changing labor supply.

Finally, the age at which a spouse suffers a health shock is likely a key determinant of how their partner responds to it.[4] For one thing, the older a person is, the more likely that the shock will be permanent or serious. This seriousness attends both to the loss in earnings that poor health may engender and to the degree of incapacitation the health condition may cause. Studies which examine people in the prime of their lives, as do many of the papers discussed previously, are likely to find small effects simply because the people suffering the disabilities are not old enough for them to be serious. This is unlikely to be a problem in the present study because, as will be discussed later, the sample under study is drawn from population of older adults.

Empirical Estimates of Effect of Health

Sound theoretical reasons may be posited in support of the small estimates of the effect of spousal health found in the literature. But, before we conclude that existing estimates are therefore credible, the possible confounding effects of various types of statistical bias must be discussed.[5] Specifically, estimates of spousal response to health are probably affected by the presence of measurement error and endogeneity bias in most empirical specifications. The measurement error problem arises, obviously, because health is almost never perfectly measured. Empirical tests usually rely on people's self-report about the particular aspect of health being studied. People may either genuinely make errors when they report their health, or may report poor health status either for

themselves or a spouse as a way of justifying non-work. This measurement error causes an attenuation bias when the effect of health on labor supply is estimated.

The endogeneity problem arises because health status probably depends in some fashion on latent factors which are themselves related to labor supply. Suppose for example that some people are "lazy". If husbands are being studied, and attention is on their hours of work, then we would expect that lazy men would work fewer hours than observationally identical men. But, it is also known from matching models of marriage that men who are lazy do not fare well on the marriage market. In particular, they may be more likely to marry women who are also lazy. If a woman's laziness affects her willingness to exercise, then the wives of lazy men will be *truly* unhealthy and a wife's health status will be endogeneous in her husband's labor supply equation. Even in the absence of any pre-marital sorting, wife's health is still probably endogenous since married couples live similar lifestyles and since lifestyle almost surely affects both health and labor supply behavior.

The confounding influence of these two biases means that spousal effects are, in general, not consistently estimated in most models. Nor is it easy to predict the direction in which they are biased. Below, I presents methods for dealing with the two types of bias.

2. **Empirical Framework**

The empirical analysis seeks to determine the effect of spouse's health on own labor. Suppose that

$$O_{it} = b_x x_{it} + d(\text{spouse's poor health}) + l\,T_{it} + a_i + m_{it} \qquad (1)$$

where O_t is a measure of an individual's labor supply behavior at time t; x_{it} is a vector of time varying determinants of the labor supply outcome; T is the length of time that the spouse has been disabled, a_i is a latent factor which is either fixed or slowly

changing over time; and m_{it} is a white noise error term. Our main interest is in the parameter d which measures the instantaneous and permanent effect of bad spousal health on own work behavior. It should be noted at once that, since there is rarely any information on the precise date of onset of the health condition, the very best that a researcher can do in a cross-section setting, even in the presence of no statistical biases whatever, is to estimate $d + l\bar{T}$, where $\bar{T}$ is the average duration of poor health among those in poor health. This is the instantaneous effect plus the mean of what we might call the "adjustment" effect for those person with a sick spouse.

An endogeneity problem discussed above arises because, in general

$$Cov(\text{spouse's poor health}, a_i) \neq 0. \tag{2}$$

As a result, the estimate of the effect of spouse's health in (1) will be biased in the direction of the covariance in (2). The other statistical problem is that "spouse's poor health" will, in general, be mis-measured. Health is typically measured in dichotomous terms; someone is "healthy" or "sick", and a patient's condition is listed as "critical", "fair" or "stable". In this study, we observe the dichotomous variable h_{it}, for which a value of 0 is supposed to indicate that the person is "non-disabled" and a value of 1 is supposed to indicate that the person is disabled.[6] Assuming that the dichotomous states of health to which these terms refer actually exist, because of measurement error. But because health is measured with error, the relationship between observed and true spouse's health is

$$h_{it} = h_{it}^{*} + v_{it} \tag{3}$$

where $h_{it}^{*} = 1$ refers to the state in which a spouse is truly disabled, and $h_{it}^{*} = 0$ to the case where they truly are not, and v_{it} is measurement error. Were we to use the variable h_{it} as the measure of spouse's poor health in (1), there would be the standard problem that the estimated effect on labor supply would suffer from attenuation bias. But the fact that the badly measured health variable is dichotomous makes the

traditional method of dealing with it – using instrumental variables (I.V.)– more complicated than the usual case.

The reason for this complication is that the error term v_{it} is not classical. In particular, it is correlated with true health status. If a person is truly "well" , so that $h_{it}^{*} = 0$, the only type of mis-measurement which could occur would be $v_{it} = 1$. Similarly, if the person were truly "disabled", so that $h_{it}^{*} = 1$, the only type of error of measurement that could occur is $v_{it} = -1$. The fact that the error is non-classical means that the use of IV techniques to instrument for the observed dichotomous health measure produces estimates which are upwards biased. This topic has been the subject of interesting recent work. For example, Card (1995) has examined the case where the mis-measured variable is union status; and Kane, Rouse and Staiger (1998) have dealt with the mis-measurement of completed years of schooling. The major difference between these papers and the present study is that the variables which other researchers have sought to study, unlike health, actually truly exist only in dichotomous categories. People are either in a union truly or they are not. Someone either graduated college or they did not. But, poor health *truly exists* in nature as a latent *continuous* variable, for which we observe binary indicators.

Suppose that in any period, and for every individual, there is an underlying, latent variable s_{it}^{*} which measures an individual's "sickness" at a point in time, given by

$$s_{it}^{*} = gZ_{it} + G(a_i) + u_{it} \tag{4}$$

where the mean-zero and serially un-correlated error u_{it} represents shocks to poor health at a point in time, and the vector Z_{it} represent a set of known determinants of s_{it}^{*}, such as the difficulties that the person has with various activities in normal life. In (4), suppose that $G(a_i)$ is a function which assumes a fixed value for every level of a_i, so that there is as aspect of sickness which is time invariant. Since $G(a_i)$ is obviously correlated with the latent determinant of labor supply, a_i, then spouse's sickness, or

health, is endonegous in an own labor supply equation. We do not observe underlying sickness, s_{it}^*, but rather indicators which are supposed to summarize whether underlying sickness is greater than or less than some threshold. Suppose that people are truly "disabled" when $s_{it}^* > 0$. Because of measurement error, the poor health indicator h_{it} takes a value of 0 not when true health is good (s^* is very small), but rather when a latent *health reporting* measure $s_{it}^{**} < 0$ is very small, where

$$s_{it}^{**} = s_{it}^* + e_{it}. \tag{5}$$

To see how we might get around the measurement error problem, notice that we can use the functional limitations Z_{it} and the dichotomous information h_{it} to estimate a probit model on the latent equation

$$s_{it}^{**} = gZ_{it} + G(a_i) + u_{it} + e_{it}, \tag{6}$$

where

$$h_{it} = \begin{cases} 0 & \text{if and only if } s_{it}^{**} < 0 \\ 1 & \text{otherwise.} \end{cases} \tag{7}$$

The error term in (6) is $h_{it} = G(a_i) + u_{it} + e_{it}$. This error is Normally distributed, with mean 0. But notice that the components $G(a_i)$ and u_{it} are omitted variables in the probit equation in (6). The probit estimate of g is therefore:[7]

$$\hat{g} = \frac{g + s}{\sqrt{Var(h_{it})}} \tag{8}$$

where s is the correlation between $G(a_i)$ and the variables Z_{it}.

Because $Var(h_{it}) > Var(e_{it})$, the probit estimate is always biased – i.e., not equal to $\frac{g}{Var(e_{it})}$[8] - regardless of whether Z_{it} is correlated with the latent portion of spouse's health that is correlated with own labor supply , $G(a_i)$. However, if we want to use information on Z_{it} to instrument for spouse's health in (1), we are not necessarily interested in estimating a consistent estimate of g. In particular, notice that the index

$$\hat{s}_{it}^* = \hat{g}Z_{it} \tag{9}$$

is strongly correlated with true sickness s_{it}^*, and is free of measurement error. Thus, two of the conditions are easily met for the use of Z_{it} as instrument for underlying spousal health. If in addition $s = 0$ so that there is no correlation between Z_{it} and factors related to labor supply, except insofar as Z_{it} affects health, then all of the conditions are met for use of Z_{it} as instruments, and using an IV technique which uses the index (9) in place of the observed dichotomous measures for "poor spouse's health" in (1) yields consistent estimates of the labor supply parameter d. If instead, Z_{it} is correlated with labor supply, separately from how it affects spousal health, then its use as an instrument in (1) is invalidated.

In all likelihood, Z_{it} will affect labor supply both through health and other mechanisms. As one example, if a_i in (1) refers to laziness, we would expect this characteristics to be related to the self-assessed difficulty of doing tasks like walking several blocks or climbing stairs. As another example, one could easily envision co-workers or employers forming the assessment that someone who avoided certain basic activities, or did them with difficulty, *might* be lazy. Wages, promotions and consequently labor supply may therefore be affected.

A solution to this problem is at hand if there is longitudinal data on all of the key variables, as is true with the data used in this paper. If we assume that

$$Z_{it} = g(a_i) + e_{it}, \tag{10}$$

so that Z_{it} is related to labor supply independent of its effect on health in a latent time-invariant manner. Given (8), (9), and (10), it follows that over any two years, m and n, the difference in the index $\hat{s}_{it}^*$ is given by

$$\hat{s}_{in}^* - \hat{s}_{im}^* = \frac{g+s}{\sqrt{Var(h)}}(g(a_i) + e_{in}) - \frac{g+s}{\sqrt{Var(h)}}(g(a_i) + e_{im}) = \frac{g}{\sqrt{Var(h_{it})}}(e_{in} - e_{im}) \tag{11}$$

since, by assumption, the relationship between Z_{it} and latent work determinants is time invariant; as is $Var(h_{it})$. Expression (11) is completely unrelated to the latent work determinants a_i. It follows that if there are multiple waves of data, any difference or

fixed effect model run on the version of (1) in which health is measured with the index s_{it}^{*} given in (9), yields a consistent estimate of the labor supply parameter, d. This is true so long as the relationship between Z_{it} and a_i is time invariant, or slowly changing over time.

In the models estimated below, the vector Z_{it} is a set of measures detailing the difficulty that respondents have with performing routine tasks such as climbing several flights or stairs or walking several blocks. In the literature, these measures are known as "functional limitations" or "A.D.L".s. The full set used in the paper is presented in the Appendix. The functional limitation measures are generally assumed to be very measured both because the questions are very narrow, thereby decreasing the chance for honest error; and because the limitations questions are not defined in ways obviously related to work as is true for a health measure such as disability status, thereby decreasing the chance for strategic mis-reporting. In order to create the indexes given by (9), I estimate probit models for the disability measures and ordered-probit models for the three-part overall, self-rated health measure.

In the next section, I describe that data used in the project and present some summary information. I next present, in turn, the results of the various models described above. I first estimate pooled labor supply equations where I use the spousal dichotomous self reports about overall health status as the health measure. I next run pooled sample I.V. estimates in which I use information on functional limitations to instrument for underlying (continuous) health. This estimate should be purged of both measurement error and endogeneity bias, unless the functional limitations have independent effects on labor supply. Finally, to correct for this possibility I exploit the longitudinal structure of the data to estimate a set of fixed effect models in which health in any period is the

estimated health index derived from using the limitations as instruments. I will refer to this estimator as the within-I.V. estimate.

3. Data

The data used in this study are from the first two waves of the Health and Retirement Survey (H.R.S.). The H.R.S. is a nationally representative, longitudinal survey of households in which there is at least one person born between 1931 and 1941 (age eligibles). In each year of the survey, information is obtained of all age-eligible respondents and their wives or husbands. The ages of the people in the H.R.S. is ideal for a study of the effects of spousal health since the people in question are mature enough for health shocks to be significant, yet young enough that they are still making decisions about whether to be part of the labor force. I analyze a sample of married, opposite-sex couples in which both the husband and wife responded to the survey in Wave 1 (1992) and Wave 2 (1994); where both were between 45 and 65 years of age[9] as of Wave 1 of the survey; and where information on age, years of education, race and health status are non-missing for both spouses.

I use two different measures of health status in this study. Both have been used by authors studying other issues related to health, and at least one is available in most of the large data sets used by labor economists. The first measure is “disability”, a binary measures which is obtained from a respondent's response to the question “Do you have health limitation which limits the type or amount of work that you can do?”. The second measure asks respondents to rate their overall health relative to other people. There are five possible responses: “excellent”, “very good”, “good”, fair” and “poor”. I collapse the “excellent” and “very good” categories together and do the same thing for the “fair” and “poor” answers. The data set also contains, in each wave, detailed information about specific functional limitations (the previously mentioned F.L's) as well

as information about specific health conditions such as whether the person has ever been diagnosed with chronic heart disease or diabetes.

Table 1 summarizes the sample. The analysis sample consists of 3,594 married couples. Eighty per cent of both husbands and wives are white and analysis (not shown) indicates that there is virtually no inter-racial marriage. As of Wave 1 of the Survey, the couples had shared many years of married life; on average, they had been together for thirty years. The husbands in the sample were, on average, almost 57 years old in Wave 1 and were about 2 years older than their wives. Both the husbands and the wives in the data had completed about 12 years of schooling, though the variance in education was larger for husbands than for wives. The spouses differ greatly with respect to their labor market activity. In each year of the survey, husbands worked many more hours and weeks in the previous year and were much more likely to work for pay than were wives. For example, whereas only sixty per cent of wives worked for pay in the Wave 1, the comparable figure was eighty per cent for husbands. Because of this large difference in labor supply behavior, husbands' earnings in 1991 were almost three times as much as that of wives. The table also shows that for both husbands and wives, labor market activity decreased substantially between the Waves of the Survey. There was, for example, a 10 percentage point reduction in employment rates for husbands between 1992 and 1994 and comparable reductions in all of the other labor force participation measures, including the probability of being retired. For both men and women, the probability of being retired grew very much over the two years but was much larger for men than for women in each. The fact that both retirement and employment rates are so low for women suggests that people view retirement as the thing which comes at the end of a life spent working. Women who have not worked for all or most of their lives seem to view their continued non-work at mature ages not as retirement but simply as non-work.

Tables 2A and 2B examine spousal health over the two Waves of the Survey. I define a set of "health history" dummies for each of the two health measures which completely summarize an individual's health in the two waves of the study according to that measure. For the disability measure, there are four such dummies given by H_{00}^{D}, H_{01}^{D}, H_{10}^{D}, and H_{11}^{D}, where

$$H_{00}^{D} = \begin{cases} 1 & \text{if } h_{i1}^{D} = 0 \text{ and } h_{i2}^{D} = 0 \\ 0 & \text{otherwise} \end{cases}$$

$$H_{01}^{D} = \begin{cases} 1 & \text{if } h_{i1}^{D} = 0 \text{ and } h_{i2}^{D} = 1 \\ 0 & \text{otherwise} \end{cases}$$

$$H_{10}^{D} = \begin{cases} 1 & \text{if } h_{i1}^{D} = 1 \text{ and } h_{i2}^{D} = 0 \\ 0 & \text{otherwise} \end{cases}$$

$$H_{11}^{D} = \begin{cases} 1 & \text{if } h_{i1}^{D} = 1 \text{ and } h_{i2}^{D} = 1 \\ 0 & \text{otherwise} \end{cases}$$

For the other general health measure there are nine health history dummies defined in similar fashion and denoted by H_{00}^{R} to H_{11}^{R}.

Table 2A and 2B are cross tabulations of husbands' and wives' health as given by the two health measures. The columns in the tables indicate the fraction of all husbands in the sample with particular health histories, and the rows entries indicate the same thing for wives. The shaded entries represent health histories where health worsened over two waves for one or both of the spouses. Reading the last row of numbers in Table 2A, almost three-quarters of the husbands in the sample were non-disabled in both waves and fourteen per cent were disabled in both Waves. There was a change in disability status for 12% of husbands: 8% experienced a worsening of their disability status worsen, while 4% of all husbands saw their disability status improve over the two years the survey covers. The evidence suggests that disability status was not permanent; for husbands who had *any* bouts of disability in their history, almost one-fifth say their disability status improve over time. Reading down the last column of numbers, the same general pattern is true for the wives. Three-quarters of the wives had no disability

at all over the two years they were observed and 12 per cent were disabled in both years. Nine percent of wives experienced health reductions and 4 per cent saw their health improve.

The inner table entries show the joint distribution of disability histories within a family. In particular, the entries allow us to ask whether husbands' and wives' disability histories are distributed independently or are correlated. If spousal disability histories were distributed independently, then the fraction of families in which husbands and wives jointly had particular histories would be equal to the product of the unconditional probability that a husband had his particular history multiplied by the probability that a wife had her particular history. Thus, a simple gauge of the independence of spousal disability history is to see whether any of the interior entries in the table is equal to the product of the associated entries in the far right column and bottom row. To see how to apply the reasoning, consider disability history where there is no disability in either survey year. Since 75% of wives were not disabled in either period, and 74% of husbands were not disabled in either period, we would expect that in a fraction $0.74 \times 0.75 = 0.56$ of all families, neither the husband nor the wife would be disabled in either year if spousal disability histories were distributed independently. The uppermost right entry in the table shows that the actual fraction was 58 per cent. Similar calculations for other pairs of disability histories indicate that the spouses' disability seems to distributed independently, though there is very mild evidence of a positive correlation.

Table 2B repeats the analysis above for self-rated health. There was a slight worsening in the overall incidence of good and bad for both men and women over the two years covered by the survey. Thus, 53% and 51% of husbands were in the best possible health in the first and second year of the survey, respectively. The comparable numbers were 57% and 53% for wives in the two years. That the overall incidence was essentially unchanged over time does not imply that health did vary much for *individual*

husbands and wives over time. For example, 19% of all husbands and 18% of wives experienced a health decline over the two waves. For both sexes, those experiencing declines were disproportionately in the best possible health to begin with. Also, declines in self-rated health seemed to be gradual, so that the fraction of people who went from the best possible health, to the worst was very small for both husbands (2%) and wives (2%).

For neither type of spouse was recovery rare. Fourteen per cent of husbands and 12% of wives saw their health get better over the survey years. But dramatic improvements were quite uncommon. Improvements were concentrated among those whose health was only somewhat bad, and only 1% of both husbands and wives who gave themselves the worst possible health rating in the first wave of the data report that their health is excellent in the second. In summary, both the health and disability measures indicate substantial evidence of poor health in both waves among both husbands and wives. The tables also indicate that health exhibits non-trivial variability over time for both husbands and wives in this age cohort, though the finer categorization permitted under the self-rated health indicates that these dynamics are only rarely dramatic. Health gets both worse and better over time but almost never becomes *much* better or worse. In the next section, I present the results for the empirical models.

4. Results

In the empirical analysis, I focus on two different measures of labor supply behavior.[10] One is a binary indicator which captures whether the individual is employed for pay as of the survey date. The other is the total number of hours that the person has worked in the previous year. This annual hours of work variable is the product of the number of hours usually worked in a week and the number of weeks usually worked in a year. For people who do not work for pay the annual hours variable equals zero.

Tables 3 and 4 present O.L.S.[11] estimates of the effect of health as measured by the dichotomous disability and overall self-rated health on labor supply of husbands and wives. Both Tables 3 and 4 show that many of the control variables have the same effect on labor supply behavior for husbands and wives. Being older unambiguously lowers labor supply, while being better educated has the opposite effect. The table also shows that the older one's spouse, the lower one's labor supply. One interesting control variable is race, where the estimated effects are quite different for husbands and wives. White men are more likely to be employed for pay and work more hours in a year, but are statistically no more likely than non-white husbands to be retired. White wives, by contrast, are less likely to be employed for pay and work fewer hours per year than their non-white counterparts, though there is no difference between the groups in the probability of being retired. On the whole, the regressions presented in the two tables do reasonably well at explaining the variation in the outcome variables, but the R^2 are larger for the various husband equations.

Table 3 confirms the well-established result that own disability adversely affects labor supply for both husbands and wives. For example, being disabled appear to lower annual hours of work by more than 950 hours a year for husbands and 655 hours a year for wives. The effect of own disability on the probability of being employed is similarly large for both men and women. For example, being disabled lowers the probability of being employed by 35 and 45 percentage points, for men and women respectively.

The results also show that the disability status of one's spouse likewise has significant effects on own labor supply behavior, but the sign of the effect appears to be different for husbands in wives in the pooled regressions. It appears that wives whose husbands are disabled work more. Their annual hours of work are larger by about 60 hours a year and the probability that they are employed is just a little over 1 percentage point larger. Husbands, on the other hand, appear to work less when their wives are

disabled. The probability that they are employed is smaller by 1 percentage point, and annual hours of work are lower by 35 hours a year in these pooled regressions.

The three-way self-rated health measure used in the regressions presented in Table 4 produces very similar results to the disability measure. The excluded health category is that for which the individual's health is rated as “excellent” or “very good”. Own poor health is associated with very large and statistically significant adverse effects on labor supply, with the adverse effects being larger, the worse own health. As with the disability measure, the results regarding spousal health are different for husbands and wives. The worse a husband's health, the more his wife works. For example, wives with husbands whose health is only somewhat bad work a little over 10 more hours per year, and those whose health is rated in the worst possible category work 40 more hours. Both of these effects are strongly statistically significant. For husbands, we see that as with the binary disability measure, the estimated effect is of spouse’s health is of different sign from that for the wives. Husbands appear to work less when their wives are in less than “excellent” health. Indeed, husbands whose wives are in the worst possible health work 35 hours less a year than their observationally identical counterparts.

On the whole, we are tempted to draw the tentative conclusion based on these results that the effect of spousal health has different effects for husbands and wives; that these effects may be causal; and, that though, statistically significant, the spousal effect is not that large for either spouse. The causal interpretation would certainly be premature, for the regressions in Tables (3) and (4) do not attempt to control for endogeneity bias in any way. Similarly, the economically small estimated effects may be due to attenuation bias from measurement error, which is also not addressed in these regressions. Also, before we conclude that husbands and wives respond differently to their partners' health problems, recall that the estimated effects in Tables 3 and 4

represent the sum of what I have previously called the instantaneous effect of a spouse being ill, d_k, plus the average of any adjustment effects $l\bar{T}$. It is therefore very possible that any apparent differences in the responses of men and women might be due to how men and women adjust over time, or to unobserved differences in the data in the duration of health problems suffered by husbands and wives. Even with these caveats, the results do suggest that that poor health of a spouse affects husbands and wives differently. After all, the same pattern shows up for both labor supply measures, and for both disability and overall health.

Tables (5) and (6) present results of the initial effort to deal with two types of statistical bias. These tables present the pooled IV estimates in which the dichotomous health measures from the regressions in Tables (3) and (4) are replaced by the estimated continuous health construct given by (9). If the functional limitations are truly good instruments – specifically, if they are un-correlated with labor supply apart from their effect on health - then these estimates should be purged of both endogenity and measurement error. (Later, I relax this assumption.) The estimates for the two sets of first stage non-linear models are presented in Appendix Table A1, which shows that for both husbands and wives, the functional limitations are very strong predictors of health.

The results in the two tables are, in broad outlines, similar to those in the simple pooled models. Importantly, most of the estimates are also statistically significant. To interpret the coefficients in the various models, it is useful to know that the mean of the disability index is 0.22 for men and 0.19 for women, with standard deviations within a gender of 0.16 and 0.16, respectively. For the self-rated health measure, the mean of the index for men is 0.03 for both men and women, with standard deviations of 0.45 for men and 0.43 for women. Table 5 shows that the own disability has a strong negative effect on labor supply for both husbands and wives. For example, an increase in own underlying continuous disability of one standard deviation causes annual hours of work

to fall by 164 (0.16×1025). For wives, a one standard deviation increase in underlying disability lowers the probability of being employed by 7 percentage points (0.16×0.45). All of the effects for own disability are strongly statistically significant.

The table reproduces the results from the pooled models that spousal disability has different effects for husbands and wives. A one standard deviation increase in spouse's disability raises wives annual hours of work by 30.7 hours (0.16×192). Since a man with the worst predicted in the sample is almost 5 standard deviations above a husband of average disability, this means that the wife of the husband with the worst predicted disability works 151.5 more hours per year than does the average wife. This effect is strongly statistically significant, and is also larger than the pooled estimates in Table (3) would imply. For husbands, a one standard deviation worsening of their wives health is associated with a *reduction* in annual hours of work of 32.5 hours. This implies that a husband whose wife's predicted disability takes the largest value in the sample works fully 162 fewer hours a year than does the husband of a wife whose disability is at the sample average. Again, this effect is very statistically significant. Notice also that it is much larger than the simple pooled estimates in Table (3) would suggest. Husbands and wives both appear to make large adjustments to their work effort in response to a spouse's poor health, though the direction of their adjustments is different. That all of the estimates are larger (in absolute value) than the pooled results using the dichotomous health measures suggests that measurement error, with its attendant attenuation bias, is a problem for both husbands and wives disabilities, and appear to affect both to approximately the same degree.

The estimates for the binary "work for pay" measure show the same pattern as the annual hours of work variable. Women are more likely to be employed, all else equal, the more disabled their husbands are; men are less likely to be employed the more disabled their wives are. All of these estimates are highly significant, unlike the simple

pooled results in Table (3) – none of which is statically different from 0. Again we see that husbands' responses to spousal disability are of about the same absolute value as that of women.

Table 6 presents the I.V. estimates for the second health measure. As with the pooled estimates with the dichotomous measures, these effects are much smaller than those for disability. Apart from this, the results are similar for disability, with effects of different sign for husbands and wives, and statistically significant estimates.

How credible are these IV estimates? Recall from the discussion in Section 2 that a major concern is that the instruments Z_{it} might have effects on labor supply quite apart from their effects on health. If this is so, then the IV estimates are biased. To explore this issue, Table 7 presents the two-step, IV-fixed effect estimates, in which I exploit the panel structure of the data to run fixed effect versions of the models run in Tables 5 and 6. That is, I run fixed effects on the model given by (1), where "spouse's poor health" and "own poor health" are replaced by the index (9). In these models, I include indicators for change of region of residence, and variables measuring completed years of schooling and race. These last two variables are time-invariant, and so might be presumed to drop out of the fixed effect. I include them here to account for the fact that the impact of *observable,* time invariant characteristics might change over time. For key time-varying variables such as age, the difference model causes these effects to be absorbed into the constant term (the change in age over any two years is the same for every person in the sample), so the overall explanatory power of the first difference model is predictably much smaller than is true for the pooled models as well as the IV models presented above. It should also be pointed out that the fixed effect models capture more cleanly any "instantaneous" effect of own or spouse's disability as measured by d since the focus on within-changes over a small interval renders it unlikely that the health estimates are confounded by any 'adjustment" effects.

Table 7 shows that the within-I.V. estimates for own disability are *much* smaller in absolute value than the IV estimates.[12] For example, own disability is predicted to lower annual hours of work by only 56 hours a year and 176 hours a year, for men and women, respectively. When compared with the predicted reductions of 1025 and 755 annual hours for men and women for a one unit change in underlying disability predicted by the IV estimates, this estimates suggest a very substantial change in the predicted effect of disability. The same is true for the health measures. Much larger labor supply reductions are implied by the I.V. estimates than by the within-I.V. results. For example a unit increase in own poor health is predicted to lower employment probabilities by 22 percentage points for men according to the IV estimates, and by only 6 percentage points with the fixed effect estimates.

Oddly, given the results for own health, Table 7 shows that the spousal health estimates, for both men and women; for both health measures; and for both the binary "work for pay" and the annual hours of work variable are very similar in simple IV and estimates and the within-I.V. procedure. Most of the estimated within-I.V. effects are only slightly smaller, but are no less statistically significant than the I.V. estimates. For example, the results from the within-I.V. results predict that a unit increase in wife's underlying disability causes husbands to lower annual hours of work by 198.9 hours rather than the 203.9 hours predicted in the IV model. Similarly, the two-step procedure predicts that wives raise hours of work by 106 hours in a year rather than by the 192 hours the IV model predicts. This marks a reduction in the estimated effect, but one nowhere as large as those for own health across the two model. The similarity of the two sets of estimates of the spousal effect is reassuring. It appears that husbands and wives both change labor supply significantly in response to changes in their spouses' health, and do so in different directions.

Why, though, are the differences between the I.V. estimates and the within-I.V. estimates for own health so large relative to the same difference for spousal health. One possible explanation has to do with the adjustment effect to poor health. Suppose that most of the reaction to own bad health occurs gradually. Then, there would be a small instantaneous effect to own disability, and the estimated effect from the within-I.V. estimator – which consistently estimate the instantaneous effect – would be very small relative to the estimate from the pooled I.V. models which estimates the sum of the adjustment and instantaneous effect. If, on the other hand, nearly all of the response to spousal health were instantaneous, with very little adjustment, then the I.V. and within-I.V. procedures should give essentially the same results. This chain of reasoning holds together logically, but it is almost surely wrong. Why would the instantaneous effect be smaller for own disability than spouse's? In the former case, the person feels the effects immediately, and does not have to form an estimate about how seriously the person is affected. He or she is *immediately* and *directly* affected, and should be expected to modify labor supply accordingly.

A much more promising answer may have to do with the relative importance of endogeneity and measurement error bias for own and spousal health measures in an own labor supply equation. Even with marital sorting, shared lifestyles and similar considerations it is almost certainly the case that own health and own functional limitations are more correlated with the latent determinants of own labor supply than are spouse's health and limitations. If this is true, then the principal bias affecting the own health variables in pooled models would be endogeneity bias. The simple I.V. models which correct for measurement error yield estimates which therefore are biased up because of the correlation between the instruments and own labor supply. As this endogeneity is removed in the within-I.V. estimator, the size of the estimates fall accordingly. On the other hand, the chief source of bias on spouse's health variable in

an own labor supply model is probably measurement error bias. As this is already taken care of in the IV estimates, and because correlation between spouse's specific limitations and own labor supply are likely small, further correction for endogeneity bias in the within-I.V. estimator brings no appreciable change in the estimates.

But, why do husbands and wives respond differently to a spouse's illness? Section 1 outlines some possible theoretical explanations why men and women might behave differently when confronted with a sick spouse. But, a more basic explanation may have to do with the different constraints which men and women face in the labor market – specifically whether they are equally able to adjust their hours of work. Some evidence on this issue may be gleaned from the degree to which working husbands and wives in Wave 1 of the data are free to adjust their hours of work. Table 8 examines hours constraints and the desire to change hours of work by sex. The results indicate that, among working wives and husbands there are no differences by sex either in hours constraints or in the desire to change hours, conditional on being able to do so.

The most reasonable explanation for the different responses in the family has to do with how men and women (at least in the cohort studied here) tend to use their time. The very large average labor supply differences between husbands and wives implies that there is a significant degree of specialization within families; husbands appear to be the principal bread-winners, wives the principal home and care-takers. When poor health strikes, the affected party becomes unable to ably execute their specialized task and the other spouse takes up the slack. Thus, women begin to work, or work more hours when their husbands are ill and husbands cut back on work to help around the home.

5. Conclusion

In this paper, I assess how husbands' and wives' labor supply is affected when their partner is in poor health, using evidence from the Health and Retirement Survey. I

employ three different estimation techniques to control for the different types of biases which might make dis-entangling the effect of spousal health on own labor supply difficult. I also present some evidence on behavior of health and disability over time and within families. I find that women and men make significant adjustments to labor supply in response to their partner's poor health differently, but these responses are of different sign. Women work more according to two different labor supply measures while men either reduce labor supply. Also, the estimated effects for both husbands and wives are much larger when health is measured by a disability measure rather than an overall self-rated measure.

I estimate, in turn, a pooled model in which health is measured by dichotomous self-reports; then a pooled, I.V, model in which health is measured by a continuous construct which uses functional limitations as instruments; and finally a two step estimator which combines IV and fixed effect techniques, The results indicate that there is substantial evidence of both endogeneity bias and measurement error bias in the pooled regressions, though these biases appear to be matter differently for own and spouse's health in labor supply equations. I provide some evidence that there is little difference between men and women in their abilities to change hours of work (at least for those already working) and argue that a fair reading of the results is that husbands and wives generally specialize in market and home production, respectively, and argue further that the main effect of illness is to cause married partners to take up more of the activity typically done by the ill person.

The results also raise some interesting questions which will be the subject of future research. One task will be examine how outcomes other than changes in spousal labor supply are affected by health changes. In particular, an important research question is whether illness makes marriages more fragile. Does illness raise the likelihood

of marital dissolution? A second area would be to examine the impact of health changes on labor supply over longer intervals than the two years studied here.

Bibliography

Becker, G. *A Treatise on the Family*. Boston: Harvard University Press, 1991.

Becker, G. Landes, E. and Michael, R. "An Economic Analysis of Marital Instability." *Journal of Political Economy*, 1977, 85. pp. 1141-1187.

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

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

Berkowitz, Monroe and William G. Johnson, "Health and Labor Force Participation." *Journal of Human Resources*, 1974, 9(1). pp. 117-128.

Bound, John. " Self-Reported versus Objective Measures of Health in Retirement Models." *Journal of Human Resources*, 1991, 26. pp. 106-138.

Card, David. "The Effects of Unions on the Structure of Wages: A Longitudinal Analysis." *Econometrica*, July 1996, 64(4). pp. 957-979.

Charles, Kerwin Kofi. "The Longitudinal Structure of Earnings Losses from Disability." unpublished manuscript, University of Michigan.

Davis, J. M., "Impact of Health on Earnings and Labor Market Activity." *Monthly Labor Review*, October 1972, 95(10). pp. 46-48.

Freeman, Richard, "Longitudinal Analyses of the Effects of Trade Unions." *The Journal of Labor Economics*, January 1984, 2(1). pp. 1-26.

Griliches, Zvi, "Errors in Variables and Other Unobservables." *Econometrica*, November 1974, 42 (6). pp. 971-998.

Gruber, Jonathan and Julie Berry Cullen, "Spousal Labor Supply as Insurance: Does Unemployment Insurance Crowd out the Added-Worker Effect?." 1996 National Bureau of Economic Research Working Paper #5608,.

Haurin, Donald R. "Women's Labor Market Reactions to Family Disruptions." *The Review of Economics and Statistics*, February 1989, 71(1). pp. 54-61.

Kane, Thomas, Rouse, C. and Douglas Staiger, "Estimating the Returns to Schooling When Schooling is Misreported", unpublished manuscript.

Killingsworth, Mark R. *Labor Supply*, Cambridge: Cambridge University Press, 1983.

Lundberg, Shelly. " Labor Supply of Husbands and Wives: A Simultaneous Equations Approach." *Review of Economics and Statistics*, May 1988., 70(2). pp. 224- 35.

Luft, Harold S., "The Impact of Poor Health on Earnings," *Review of Economics and Statistics*, Feb 1975. 57(1). pp. 43-57.

Maloney, Tim, "Employment Constraints and the Labor Supply of Married Women," *Journal of Human Resources*, Winter 1987, 22(1). pp. 51-61.

Maloney, Tim, "Unobserved Variables and the Elusive Added Worker Effect." *Economica*, May 1991, 58(230). pp. 173-87.

Mincer, Jacob "Labor Force Participation of Married Women," In *Aspects of Labor Economics*, ed. H.G. Lewis. Princeton, N.J.: Princeton University Press, (1962).

Parsons, Donald, "Health, Family Structure, and Labor Supply," *American Economic Review* , September 1977, 67(4), pp .703-712.

Spletzer, James R. "Reexamining the Added Worked Effect," Bureau of Labor Statistics Working Paper no. 237, 1992.

Stephens, Melvin Jr., "Worker Displacement and the Added Worker Effect," unpublished manuscript, University of Michigan, 1998.

Weaver, David A., "The Work and Retirement Decisions of Older Women: A Literature Review." *Social Security Bulletin*, 1994, Spring, 57(1). pp. 3-24.

Yatchew, Adonis and Zvi Griliches, "Specification Errors in Probit Models." *The Review of Economics and Statistics*, February 1985, 67(1). pp. 134-139.

End Notes

[1] Henceforth, I use the expression "disabled" and "poor health" interchangeably. In the empirical work, I distinguish between the two concepts by using two different health measures - one for disability and another for self-rated overall health.

[2] Theoretical foundations for the model are provided by Becker. Empirical versions of the model have been used by many authors. See Killingsworth (1983).

[3] Retrospective information on the onset of bad health is not available in most data sources used by empirical economists.

[4] Charles (1997) finds that there are significant age-of-onset differences in the effect of disability on earnings, and on the rate of recovery after disability has occurred.

[5] See Bound (1991) for a careful discussion about the different biases present in empirical models of the effects of health.

[6] Actually two indicators of overall health are used in the analysis. The second measure is a rating of overall health, with 3 possible categories. For simplicity, the arguments in this section focus only on the disability measures for which there are only two categories, but both measures are used in the empirical work.

[7] See Yatchew and Griliches (1985) for a description of biases in the probit model.

[8] It is conventional to assume that, in probit model, the variance of the underlying, true error is 1.

[9] I did not restrict the sample to families where *both* spouses are age eligible because the sampling frame of the H.R.S. is meant to be representative of households in *at least* one spouse is age eligible. The age 45-65 age restriction on spouses' ages is imposed so that families with overly large age difference between the spouses are excluded from the analysis. That these families are atypical with respect to the how far apart in age the partners are suggests that their labor supply behavior might also not be particularly representative.

[10] I also estimated linear probability models for "whether retired". As this variable is quite difficult to clearly identify for the sample under study, I do not these results in the body of the paper. The results are summarized in the Appendix in Table A2.

[11] For the binary "work for pay" variable, probit models were estimated in addition to the linear probability models presented here. The probit results are qualitatively quite similar to those presented here and are available upon request.

[12] As a sensitivity test, I also estimated the two step model in the other order. That is, in addition to running the fixed effect models (which may be thought of as a type of difference estimate) on the instrumented health, I also ran the fixed effect estimates on the dichotomous health measures, *then* replaced the change in the dichotomous health measures with their associated indexes created by using the coefficients from a set of ordered probit models. The basic patterns described in the text were virtually identical.